Министерство общего и профессионального
образования Российской Федерации

Тверской государственный университет

Ю.С.ХОХЛОВ

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебное пособие

Часть I

Тверь 1997

УДК 519.2

Хохлов Ю.С. Теория вероятностей и математическая статистика: Ч. I. Учебное пособие/ ТвГУ. — Тверь, 1997. — 74 с.

Настоящее пособие является первой частью курса лекций по теории вероятностей и математической статистики, в которой излагаются основные понятия теории вероятностей: вероятностное пространство, свойства вероятностей, условная вероятность, последовательность независимых испытаний, случайные величины и векторы и их распределения.

Рекомендуется студентам математических специальностей, а также экономистам.

Библиогр. 9.

Рецензенты

Кафедра математической статистики Московского госуниверситета им. М.В. Ломоносова;

доктор физико-математических наук В.В. Сенатов

Часть I

Теория вероятностей

1.1 Предмет теории вероятностей

Теория вероятностей является математической наукой. Как и другие разделы математики, она имеет дело не напрямую с объектами и явлениями реального мира, а с их математическими моделями. Процесс построения математической модели, ее исследования и применения - довольно сложный многоступенчатый процесс. Вначале, наблюдая за реальным объектом, проводя эксперименты и накапливая факты, строят модели конкретных наук, таких, как физика, химия, биология, экономика и т.д. Далее, сравнивая различные модели, выделяя их существенные черты, отвлекаются от конкретных объектов и исследдуют только структуру моделей. Это и приводит нас к математической модели. Изучая свойства математической модели, приходят к открытию новых эффектов и предсказанию поведения этого объекта в новых условиях, где он ранее не наблюдался. Таким образом, мы вновь возвращаемся к реальному объекту.

До возникновения теории вероятностей предметом математики были только модели таких явлений, в которых исход тех или иных экспериментов однозначно определялся заданием некоторого комплекса начальных условий. Классическим примером являются механические модели. Эксперименты такого типа будем называть **детерминированными**. Но нетрудно привести примеры и таких экспериментов, когда при возможной точности фиксации начальных условий исход эксперимента однозначно не определен, т.е. некоторое событие $\mathbf{A}$ при заданном комплексе условий $\mathbf{K}$ иногда происходит, а иногда нет. Такие эксперименты мы будем называть **экспериментами с неопределенным исходом** (термин "случайный" мы сохраним для более точных формулировок). Классическим примером является подбрасывание монеты. Какого типа закономерности можно изучать для таких неопределенных ситуаций? Сразу же оговоримся, что теория вероятностей имеет дело не с любыми экспериментами с неопределенным исходом, а только с так называемыми **массовыми явлениями**, когда экперимент проводится большое число раз в одинаковых условиях или рассматривается большая совокупность однородных объектов. Типичный вопрос теории вероятностей (с практической точки зрения) - насколько часто происходит данное событие **A** при заданном комплексе условий **K** в длинной серии испытаний. Чтобы перейти к более аккуратным формулировкам, уточним некоторые понятия. Пусть мы N раз провели некоторый эксперимент, в котором событие A произошло в $N(A)$ испытаниях. **Относительной частотой** появления события A в N испытаниях называется число

$$h_N(A) \;=\; \frac{N(A)}{N}\;.$$

Чтобы понять, модели каких экспериментов изучаются в теории вероятностей, рассмотрим простейший пример - подбрасывание монеты. Выполним 6 серий по 100 подбрасываний и в каждой серии подсчитаем число $N(A)$ появлений события A, когда монета падала гербом вверх.

$N = 100.$

Серия	N(A)	$h_N(A)$
1	56	0,56
2	48	0,48
3	52	0,52
4	50	0,50
5	47	0,47
6	46	0,46

Сгруппируем эти результаты в три серии по 200 подбрасываний.

$N = 200$

Серия	N(A)	$h_N(A)$
1	104	0,520
2	104	0,510
3	93	0,465

Можно увеличивать число испытаний в одной серии и следить за поведением относительной частоты $h_N(A)$. В результате мы получаем последовательно: 0,56; 0,52; 0,52; 0,515; 0,506; 0,498. На этом примере мы видим, что выполнены следующие свойства:

1) относительная частота $h_N(A)$ события A в длинной серии испытаний "тяготеет" к некоторому постоянному неслучайному числу;

2) в разных сериях испытаний, но проводимых в одинаковых условиях, относительные частоты приблизительно равны;

3) если мы из данной серии испытаний выберем некоторую подсерию, не используя информацию о результатах эксперимента, то новая относительная частота тяготеет к тому же числу.

Будем говорить, что для некоторого эксперимента выполнено свойство **статистической устойчивости частот**, если выполнены свойства 1-3. **Случайным экспериментом** будем называть такой, в котором выполнено свойство устойчивости частот.

Теория вероятностей - это раздел матеметики, где изучаются модели массовых случайных явлений, для которых выполняется свойство устойчивости частот.

Число, к которому тяготеет относительная частота $h_n(A)$, будем называть **вероятностью** этого события. Вероятность события A измеряет меру возможности его появления в случайном эксперименте.

В теории вероятностей изучаются свойства вероятностей различных событий. *Основная задача теории вероятностей как математической науки состоит в том, чтобы, зная вероятности одних событий, вычислить вероятности других, так или иначе связанных с первыми.*

1.2 Дискретное вероятностное пространство

1.2.1 Определение вероятностного пространства

При построении математической модели мы должны найти компромисс между двумя обстоятельствами. С одной стороны, она должна быть достаточно подробной, чтобы учесть все существенные черты изучаемого явления. С другой стороны, необходимо отбросить все несущественные детали, затемняющие суть дела. Излишняя подробность затрудняет изучение свойств модели, а чрезмерное упрощение может привести к неправильным выводам относительно поведения реальной системы.

Мы начинаем изучение курса теории вероятностей с исследования свойств моделей таких случайных экспериментов, которые имеют конечное или счетное число исходов. **Элементарным исходом** мы будем называть такое событие, которое однозначно (с опреленной точки зрения) говорит о том, чем закончился эксперимент. Это сразу же накладывает на множество элементарных исходов следующее важное ограничение: в каждом испытании происходит один и только один элементарный исход.

Чтобы понять, как должна выглядеть наша модель, рассмотрим пример. Однородный игральный кубик в одинаковых условиях подбрасывают много раз и отмечают число очков, выпавших на верхней грани. Ясно, что в этом эксперименте есть 6 элементарных исходов, которые мы обозначим $\omega_1, \omega_2, \ldots, \omega_6$ (ω_k означает, что выпало k очков). Пусть $h_N(\omega_k)$ - относительная частота появления исхода ω_k. Тогда эти частоты обладают следующими свойствами:

1) $\quad h_N(\omega_k) \geq 0, \ \forall k,$

2) $\quad \sum\limits_{k=1}^{6} h_N(\omega_k) = 1.$

Как отмечалось выше, частоты тяготеют к некоторым числам, которые мы будем называть вероятностями этих исходов. Ясно, что они должны наследовать свойства частот. Эти предварительные рассмотрения приводят нас к следующему определению.

Определение 1 . **Дискретным вероятностным пространством** *называется пара* (Ω, P), *где* Ω *- конечное или счетное множество,* P *- вещественная функция, заданная на* Ω, *такая, что*

1) $P(\omega) \geq 0, \ \forall \omega \in \Omega,$

2) $\sum\limits_{\omega \in \Omega} P(\omega) = 1.$

Множество Ω называется **пространством элементарных исходов**, его элементы ω - **элементарными исходами**, а число $P(\omega)$ - **вероятностью появления** элементарного исхода ω.

Пример 1. Симметричную монету подбрасывают один раз. Здесь два элементарных исхода: выпал герб - Г, выпала цифра - Ц. Таким образом, $\Omega = \{\text{Г}, \text{Ц}\}$. В силу симметрии естественно положить $P(\text{Г}) = 1/2, \ P(\text{Ц}) = 1/2$.

Пример 2. Однородный симметричный игральный кубик подбрасывают один раз. В этом случае

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \ P(\omega) = 1/6.$$

Другие примеры будут приведены на практических занятиях. Важную роль играет следующий частный случай дискретного вероятностного пространства.

Определение 2 . *Говорят, что мы имеем задачу на классическое определение вероятности, если* $\Omega = \{\omega_1, \ldots, \omega_n\}$*-конечное множество и для всех* ω_i, $P(\omega_i) = 1/n$, *т.е. все исходы* **равновозможны**.

Обычно предположение о равновозможности исходов делается из соображений симметрии задачи. Но так ли это на самом деле (т.е. верна ли модель), можно установить только из сравнения с экспериментальными данными.

1.2.2 События и операции над ними.

До сих пор мы рассматривали только элементарные исходы, т.е. в некотором смысле простейшие события. Но кроме них нас могут интересовать и другие, более сложные события. В примере 2 мы можем рассмотреть событие A, состоящее в том, что выпало четное число очков. В теории вероятностей о каждом событии мы хотим знать только одно: произошло оно или нет в данном испытании. Каждое испытание (т.е. однократное проведение эксперимента) заканчивается появлением одного из элементарных исходов, которые однозначно описывают то, чем закончился эксперимент. В частности, по элементарному исходу ω можно определить, произошло событие A или нет. Поэтому все элементарные исходы делятся на две группы: те ω, которые приводят к появлению события A (назовем их **благоприятными** этому событию), и все остальные. С точки зрения их появления в рассматриваемом эксперименте событие A и множество благоприятных для него исходов являются для нас эквивалентными. Таким образом, мы приходим к следующему определению.

Определение 3 . **Случайным событием** *назовем произвольное подмножество A пространства элементарных исходов Ω.*

Будем говорить, что событие произошло, если появился элементарный исход, ему принадлежащий, т.е. благоприятный.

Пример 3. Подбрасывают игральный кубик, A - выпала четная цифра. Тогда $A = \{2, 4, 6\}$.

В силу того, что каждое случайное событие отождествляется с некоторым подмножеством A пространства элементарных исходов Ω, различные операции над множествами позволяют определить некоторые операции над событиями. С точки зрения теории вероятностей каждое событие характеризуется только тем, когда оно происходит, а когда нет. Поэтому определения операций над событиями даются именно в этих терминах. С другой стороны, они соответствуют определенным операциям над множествами. Отсюда появляется определенная двойственность терминологии.

Определение 4 . *1) Событие называется* **достоверным**, *если оно происходит всегда, и* **невозможным**, *если оно никогда не происходит.*

Этим событиям соответствуют все пространство Ω и пустое множество $\emptyset$.

2) **Объединением** *двух событий A и B называется такое событие C, которое происходит тогда и только тогда, когда происходит хотя бы одно из этих двух событий.*

На языке теории множеств это соответствует операции объединения множеств и обозначается как $C = A \cup B$.

3) **Пересечением** *двух событий A и B называется такое событие C, которое происходит тогда и только тогда, когда происходят оба эти события одновременно.*

Это соответствует операции пересечения множеств и обозначается $C = A \cap B$ или $C = AB$.

4) События A и B называются **несовместными (непересекающимися)***, если они не могут происходить одновременно.*

Это соответствует непересекающимся множествам и обозначается $AB = \emptyset$.

5) **Суммой** *событий A и B называется их объединение в случае, когда они несовместны.*

Это не новая операция, а частный случай определения 2 и обозначается $A + B$.

6) Событие $\overline{A}$ называется **противоположным** *к событию A, если оно происходит тогда и только тогда, когда не происходит событие A.*

На языке теории множеств это соответствует переходу к дополнению множества A.

7) **Разностью** *двух событий A и B называется такое событие C, которое происходит тогда и только тогда, когда происходит A и не происходит B.*

Это соответствует операции разности множеств и обозначается $C = A \setminus B$.

8) Говорят, что событие A **влечет** *событие B, если при появлении события A, обязательно происходит и событие B.*

Это означает, что множество A есть часть (подмножество) множества B и обозначается $A \subset B$.

Чтобы наглядно представлять себе операции над событиями, полезно рисовать их в виде некоторых фигур на плоскости, например кругов. Картинки такого рода называются **диаграммами Венна**.

Конкретные примеры событий и операций над ними будут рассмотрены далее , а также на практических занятиях. Некоторые свойства операций над событиями собраны в следующем предложении.

Предложение 1
1) $A \cup A = A$, 2) $A \cap A = A$, 3) $A \cup \overline{A} = \Omega$,
4) $A \cap \overline{A} = \emptyset$, 5) $A \cup \Omega = \Omega$, 6) $A \cap \Omega$,
7) $A \cup \emptyset = A$, 8) $A \cap \emptyset = \emptyset$, 9) Если $A \subset B, B \subset C$, то $A \subset C$,
10) Если $A \subset B$, то $B = A + (B \setminus A)$, 11) $A \setminus B = A \cap \overline{B}$,
12) $A \cup B = A + (B \setminus A) = A + A\overline{B}$, 13) $A \cup (B \cup C) = (A \cup B) \cup C$,
14) $A \cap (B \cap C) = (A \cap B) \cap C$,
15) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
16) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
17) $\bigcup\limits_{\alpha \in I} \overline{A_\alpha} = \overline{\bigcap\limits_{\alpha \in I} A_\alpha}$,
18) $\bigcap\limits_{\alpha \in I} \overline{A_\alpha} = \overline{\bigcup\limits_{\alpha \in I} A_\alpha}$,
19) $A_1 \cup A_2 \cup \ldots \cup A_n = A_1 + \overline{A_1}A_2 + \overline{A_1 A_2}A_3 + \ldots + \overline{A_1} \ldots \overline{A_n}A_n$.

Задача 1. Доказать предложение 1.

Во многих задачах нас интересует не все множество событий, связанных с данным экспериментом, а только некоторые из них. Но всегда нам хотелось бы, чтобы определенные выше операции над событиями не выводили нас за пределы рассматриваемого множества событий. В связи с этим полезно следующее понятие.

Определение 5 . *Некоторый класс $\mathcal{A}$ событий называется* **алгеброй событий**, *если*
1) $\Omega \in \mathcal{A}$,
2) если $A \in \mathcal{A}$, *то* $\overline{A} \in \mathcal{A}$,
3) если $A, B \in \mathcal{A}$, *то* $A \cup B \in \mathcal{A}$.

Задача 2. Доказать, что все определенные выше операции не выводят нас за пределы алгебры $\mathcal{A}$.

С практической точки зрения выбор некоторой алгебры событий соответствует определенному взгляду на случайный эксперимент. Алгебра событий - это только те события, которые нас интересуют с этой точки зрения (например, те, которые доступны наблюдению).

1.2.3 Вероятности событий и их свойства.

До сих пор мы рассматривали только вероятности элементарных исходов. Теперь мы определим вероятности событий и исследуем некоторые их свойства.

Определение 6 . **Вероятностью** *события A называется число*

$$P(A) := \sum_{\omega \in A} P(\omega). \tag{2.1}$$

Пример 4. Симметричный игральный кубик подбрасывают один раз. Найти вероятность события A, состоящего в том, что выпала четная цифра.

В этом случае $A = \{2, 4, 6\}$, $P(A) = P(2) + P(4) + P(6) = 3/6 = 1/2$.

Пример 5. Пусть мы имеем задачу на классическое определение вероятности. Если $\Omega = \{\omega_1, \dots, \omega_n\}$, где n - общее число элементарных исходов, а m - число благоприятных исходов для события A, то

$$P(A) = \underbrace{\frac{1}{n} + \ldots + \frac{1}{n}}_{m \text{ раз}} = \frac{m}{n}.$$

Именно этот результат обычно приводят в качестве определения в элементарных учебниках по теории вероятностей.

Соберем некоторые простейшие свойства вероятностей в виде следующего предложения.

Предложение 2 . *Пусть выделена некоторая алгебра $\mathcal{A}$ событий, для которых определены вероятности по формуле (1). Тогда справедливы следующие свойства:*
1. $P(A) \geq 0$.
2. $P(\Omega) = 1$.
3. Если $A_1,\ A_2,\ \ldots,\ A_n$-попарно несовместны, то

$$P(\sum_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k).$$

Это свойство называется **теоремой сложения** *или свойством* **конечной аддитивности** *вероятности.*
4. $P(\overline{A}) = 1 - P(A)$.
5. $P(A \cup B) = P(A) + P(B) - P(AB)$.
6. Если $A \subset B,$ то $P(B \setminus A) = P(B) - P(A)$.
7. Если $A \subset B,$ то $P(A) \leq P(B)$.
8. $P(\bigcup_{k=1}^{n}) \leq \sum_{k=1}^{n} P(A_k)$ - *свойство полуаддитивности вероятности.*
9. $P(A) \leq 1$.
10. $P(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k) - \sum_{i<j} P(A_iA_j) + \sum_{i<j<k} P(A_iA_jA_k) - \ldots (-1)^n P(A_1 \ldots A_n)$
- формула включения и исключения.

Доказательство. Основными являются свойства 1-3. Только здесь мы будем использовать в явном виде то, что мы работаем в рамках дискретного вероятностного пространства. Все остальные свойства будут выведены из этих трех.

Свойства 1 и 2 вытекают непосредственно из определений дискретного вероятностного пространства и вероятности события.

Докажем свойство 3. Пусть, вначале, $n = 2$. Тогда мы имеем два события A_1 и A_2. Так как они несовместны, то исходы $\omega \in A_1 + A_2$ распадаются на два непересекающихся класса: те, что принадлежат A_1, и те, что принадлежат A_2. В силу свойств рядов с неотрицательными членами имеем

$$P(A_1 + A_2) = \sum_{\omega \in A_1 + A_2} P(\omega) = \sum_{\omega \in A_1} P(\omega) + \sum_{\omega \in A_2} P(\omega) = P(A_1) + P(A_2).$$

Для произвольного n доказательство проводится по индукции. Предлагается это сделать самостоятельно.

Свойство 4 очевидно следует из определения вероятности события и того, что сумма вероятностей всех элементарных исходов равна 1.

События A и $\overline{A}$ несовместны, и $A+\overline{A}=\Omega$. Используя свойства 2 и 3, имеем:

$$1 = P(\Omega) = P(A+\overline{A}) = P(A) + P(\overline{A}) \ .$$

Представим события A, B и $A \cup B$ в следующем виде:

$$A = AB+(A\setminus B),\ B = AB+(B\setminus A),\ A\cup B = AB+(A\setminus B)+(B\setminus A).$$

Далее используем свойство 3:

$$P(A) = P(AB) + P(A \setminus B),\ P(B) = P(AB) + P(B \setminus A),$$

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

Этим доказано свойство 5.

Свойства 6,7,8,9 можно доказать аналогично (сделать самостоятельно!).

Докажем свойство 10. Для упрощения доказательства введем одно новое понятие, которое будет полезным и в других вопросах.

Индикатором события A называется функция $I(A,\omega)$, заданная на пространстве элементарных исходов Ω по правилу:

$$I(A,\omega) = \begin{cases} 1 & , \quad \omega \in A, \\ 0 & , \quad \omega \notin A. \end{cases}$$

Легко доказать следующие свойства индикаторов событий:

1) $\quad I(\Omega,\omega) \equiv 1,\ I(\emptyset,\omega) \equiv 0,$

2) $\quad I(\overline{A},\omega) = 1 - I(A,\omega),$

3) $\quad I(\bigcap\limits_{k=1}^{n} A_k,\omega) = \prod\limits_{k=1}^{n} A_k.$

Используя понятие индикатора события, можно в следующем виде записать определение вероятности события:

$$P(A) = \sum_{\omega \in \Omega} P(\omega) I(A, \omega). \tag{2.2}$$

Применим этот результат к доказательству свойства 10. В силу предложения 2.1 мы имеем $\overline{\bigcup\limits_{k=1}^{n} A_k} = \bigcap\limits_{k=1}^{n} \overline{A_k}$. Тогда для индикаторов получаем

$$\begin{aligned}
1 - I(\bigcup_{k=1}^{n} A_k, \omega) &= I(\overline{\bigcup_{k=1}^{n} A_k}, \omega) = I(\bigcap_{k=1}^{n} \overline{A_k}, \omega) \\
&= \prod_{k=1}^{n} I(\overline{A_k}, \omega) = \prod_{k=1}^{n} (1 - I(\overline{A_k}, \omega)) = \\
&= 1 - \sum_{k=1}^{n} I(A_k, \omega)) + \sum_{i<j} I(A_i A_j, \omega)) - \\
&\quad - \sum_{i<j<k} I(A_i A_j A_k, \omega)) + \ldots
\end{aligned}$$

Дальше воспользуемся формулой (2), свойством аддитивности суммы ряда и перейдем к противоположному событию. □

Из вышеизложенного следует, что основой нашей модели является некоторое множество (алгебра) событий и вероятности этих событий, обладающие свойствами 1-3 из предложения 2.2. Поэтому мы можем дать новое определение вероятностного пространства.

Определение 7 . **Вероятностным пространством** *(в слабом смысле) называется тройка* $(\Omega, \mathcal{A}, P)$*, где* Ω *- произвольное множество,* $\mathcal{A}$ *- некоторая алгебра его подмножеств, а* P *- вещественная функция, заданная на* $\mathcal{A}$ *и обладающая свойствами:*
1) $\forall A \in \mathcal{A}, P(A) \geq 0,$
2) $P(\Omega) = 1,$
3) если $A_1, \ldots, A_n$ *-попарно несовместны , то*

$$P(\sum_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k).$$

Множество Ω *называется* **пространством элементарных исходов**, *элементы алгебры* $\mathcal{A}$ *называется* **событиями**, *число* $P(A)$ - **вероятностью** *события* A.

Преимущество такого определения в том, что оно применимо и к некоторым ситуациям, в которых мы имеем несчетное множество элементарных исходов. Одним из первым был рассмотрен следующий частный случай.

Определение 8 . *Говорят, что мы имеем задачу на* **геометрическое определение вероятности**, *если Ω есть ограниченное борелевское подмножество в R^d, $\mathcal{A}$ -алгебра всех его борелевских подмножеств, а вероятность событий задается по правилу*

$$P(A) = \frac{L(A)}{L(\Omega)},$$

где $L(A)$-мера Лебега(длина, площадь, объем, ...) множества A.

Пример 6. Из отрезка [0,1] случайным образом выбирают точку. Найти вероятность того, что она лежит на расстоянии не более 1/4 от точки 0.

В результате такого эксперимента мы получаем некоторую точку $\omega \in [0,1]$. Поэтому естественно положить $\Omega = [0,1]$. Если воспользоваться геометрическим определением вероятности,то получаем $A = [0, 1/4]$ и

$$P(A) = \frac{L(A)}{L(\Omega)} = \frac{1/4}{1} = 1/4.$$

1.3 Элементы комбинаторики. Модели случайного выбора.

Вопросы, которые рассматриваются в этом параграфе, не имеют особого значения с точки зрения общей теории, так как здесь изучается некоторая частная модель, связанная с классическим определением вероятности. Но модели случайного выбора широко используются в экономике, социологии, демографии и других науках. В силу этого они очень важны с практической точки зрения.

Как мы видели выше, при классическом определении вероятности нам нужно решить две задачи: вычислить число всех элементарных исходов n и число m благоприятных исходов для события

A. Обычно это сводится к подсчету того, сколькими способами может быть выполнено некоторое действие, т. е. сколько существует вариантов. Задачами такого типа занимается специальный раздел математики, называемый **комбинаторикой**. Мы начнем изучение основ комбинаторики с одной элементарной, но очень полезной леммы.

Лемма 1 (**умножения**). *Пусть некоторое действие осуществляется в два этапа. На первом этапе мы имеем n_1 вариантов его осуществления, а на втором, независимо от того, что произошло на первом этапе, n_2 - вариантов. Тогда общее число вариантов осуществления такого действия равно $n_1 n_2$.*

Доказательство этой леммы сводится к установлению взаимно однозначного соответствия между двухэтапными вариантами и клетками таблицы размера $n_1 \times n_2$.

Нас в основном будут интересовать задачи комбинаторики, которые можно сформулировать в терминах выбора некоторого количества предметов из заданной совокупности $U = \{a_1, \ldots, a_k\}$. При этом необходимо учитывать два обстоятельства: как производился выбор и как фиксировался результат выбора. В первом случае мы имеем два варианта: **выбор с повторением**, когда на каждом этапе выбранный предмет возвращается назад и выбор производится из одной и той же совокупности; **выбор без повторения**, когда выбранный на некотором шаге предмет уже не используется в дальнейшем. При фиксации результата мы также имеем два варианта: **с учетом порядка** (в этом случае мы знаем, какие выбраны предметы и в каком порядке) и **без учета порядка** (здесь мы знаем только, какие были выбраны предметы, но не имеем информации о порядке их появления). Имея в виду вероятностные и статистические применения рассматриваемых задач, мы будем называть результат выбора элементарным исходом или **выборкой**. Число элементов в выборке называется ее **объемом**. Рассмотрим несколько стандартных примеров.

Определение 1 . Размещением с повторением *из k элементов по l называется произвольная упорядоченная выборка с возвращением, объема l из совокупности U.*

В этом случае $\omega = (\omega_1, \ldots, \omega_l)$, где $\omega_i \in U$. Используя лемму умножения, легко подсчитать ,что число различных элементарных исходов будет равно $n = k^l$.

Определение 2 . Размещением *(без повторения) из k элементов по l называется произвольная упорядоченная выборка без возвращения объема l из совокупности U.*

Здесь $\omega = (\omega_1, \ldots, \omega_l),\ \omega_i \in U,\ \omega_i \neq \omega_j,\ i \neq j$. Вновь по лемме умножения получаем ,что число различных размещений равно

$$A_k^l = k(k-1)\ldots[k-(l-1)].$$

Если $l = k$, то такое размещение называется **перестановкой**. Число различных перестановок k элементов равно

$$P_k = k(k-1)\ldots 1 = k!$$

Определение 3 . Сочетанием *(без повторений) из k элементов по l называется произвольная неупорядоченная выборка без возвращения объема l из совокупности U.*

В этом случае мы отмечаем только, какие именно элементы вошли в эту выборку, и не учитываем порядок их появления. Поэтому для каждого элемента достаточно указать, входит он в выборку или нет.Таким образом, $\omega = (\omega_1, \ldots, \omega_k)$, где $\omega_i = 1$, если a_i входит в выборку ω, и равно 0 в противном случае. Из каждого сочетания, переставляя его элементы, мы получаем $l!$ различных размещений. В силу этого число сочетаний в $l!$ раз меньше числа размещений, т. е.

$$C_k^l = \frac{A_k^l}{l!} = \frac{k(k-1)\ldots[k-(l-1)]}{l!} = \frac{k!}{l!(k-l)!}.$$

Определение 4 . Сочетанием с повторением *из k элементов по l называется произвольная неупорядоченная выборка с возвращением объема l из совокупности U.*

Так как теперь один элемент может входить в выборку несколько раз, то элементарный исход имеет вид $\omega = (\omega_1, \dots, \omega_k)$, где $\omega_i = q$ означает, что элемент a_i входит в выборку q раз. Для подсчета числа таких выборок применим следующий вспомогательный прием. Выпишем сначала все элементы a_1, затем a_2 и так далее. Между элементами разных типов поставим перегородки. Таким образом, мы имеем l элементов и $k-1$ перегородку. Если каких-то элементов нет, то перегородки стоят рядом. Для задания сочетания с повторением достаточно расставить $k-1$ перегородку на $k-1+l$ возможных мест без учета порядка. Это можно сделать C_{k-1+l}^{k-1} способами.

Рассмотрим два важных с практической точки зрения примера.

Пример 1. Пусть мы имеем некоторую совокупность, содержащую N предметов, из которых N_1 - одного типа и N_2 - другого типа. С возвращением выбираем k предметов из этой совокупности. Найти вероятность события того,что среди них будет l предметов первого типа.

Так как выбор осуществляется с возвращением, то мы имеем всего N^k различных элементарных исходов в этой задаче. Благоприятные исходы легче пересчитать, используя следующее замечание. Сначала мы должны выбрать, на каких шагах мы будем отбирать предметы первого типа. Это можно сделать C_k^l способами. Затем на каждом из этих шагов мы имеем N_1^l вариантов, а на остальных- N_2^{k-l} вариантов. По лемме умножения число благоприятных исходов будет равно $m = C_k^l N_1^l N_2^{k-l}$. Если обозначить через $p = N_1/N$ долю предметов первого типа во всей совокупности, то получим

$$P(A) = C_k^l p^l (1-p)^{k-l}.$$

Пример 2. Пусть мы имеем ту же задачу, что и в примере 1, но выбор осуществляется без возвращения.

В этом примере удобнее фиксировать результат без учета порядка следования элементов. Тогда общее число элементарных исходов $n = C_N^k$. Для получения благоприятного исхода нужно в два этапа набрать l элементов первого типа (из N_1) и $k-l$ элементов второго типа (из N_2), т. е. по лемме умножения $m = C_{N_1}^l C_{N_2}^{k-l}$. Тогда по классическому определению вероятности получаем

$$P(A) = \frac{C_{N_1}^l C_{N_2}^{k-l}}{C_N^l}.$$

Задача 1. Пусть в примере $N, N_1, N_2 \to \infty$ так ,что $N_1/N \to p \in (0,1)$, а k и l - фиксированы. Тогда

$$P(A) \to C_k^l p^l (1-p)^{k-l}.$$

Таким образом, в этом случае выбор с возвращением и выбор без возвращения эквивалентны. Легко понять, почему так получилось. Дело в том, что для больших совокупностей удаление небольшого числа элементов не меняет практически пропорций, т. е. мы имеем на каждом шаге выбор из той же совокупности.

Пример 3. В некотором городе живет 100 000 человек, среди которых 60 000 женщин. Для проведения социологического обследования производят выборку объема 500. Найти вероятность того, что среди них будет 300 женщин.

Так как отбор 500 человек не изменит существенно пропорций, то можно считать, что мы имеем выбор с возвращением. Здесь $N = 100000$, $N_1 = 60000, p = N_1/N = 0.6, k = 500, l = 300$. Тогда

$$P(A) = C_{500}^{300} (0.6)^{300} (0.4)^{200}.$$

Примеры 1 и 2 можно обобщить на случай выбора из совокупностей, где есть предметы r типов ($r \geq 2$).

Примеры других задач будут приведены на практических занятиях.

1.4 Условная вероятность. Независимость событий.

Как отмечалось в начале нашего курса, мы подразумеваем, что эксперимент проводится при некотором фиксированном комплексе условий **K**. Если эти условия изменились, то изменяется и вероятность событий, относящихся к этому эксперименту. Такое изменение всегда можно понимать как появление некоторого события (кроме исходного комплекса условий **K**). Чтобы понять, как определить в этом случае новую (условную) вероятность, рассмотрим соответствующие частоты. Пусть эксперимент проведен N раз, событие B произошло $N(B)$ раз, а события A и B вместе $N(AB)$ раз. Тогда "условная" частота события A среди тех экспериментов, где произошло событие B, равна

$$h_N(A|B) = \frac{N(AB)}{N(B)} = \frac{N(AB)/N}{N(B)/N} = \frac{h_N(AB)}{h_N(B)}.$$

Имея в виду, что вероятность наследует свойства частот, можно дать следующее

Определение 1 . **Условной вероятностью** *события A при условии, что произошло событие B ($P(B) \neq 0$), называется число*

$$P(A|B) := \frac{P(AB)}{P(B)}.$$

Иногда применяют и другое обозначение $P_B(A)$.

Пример 1 . *Симметричную монету подбрасывают два раза. Известно, что выпал один герб (событие B). Найти вероятность события A, состоящего в том, что герб выпал при первом бросании.*

Легко вычислить, что $P(B) = 1/2$, а $P(AB) = 1/4$. Отсюда следует, что $P(A|B) = (1/4)/(1/2) = 1/2$.

Нетрудно проверить, что при фиксированном B условная вероятность обладает следующими свойствами:

1. $P(A|B) \geq 0$,

2. $P(\Omega|B) = 1$,

3. Если $A_1, \dots, A_n$ - попарно несовместны, то

$$P(\sum_{k=1}^{n} A_k|B) = \sum_{k=1}^{n} P(A_k|B).$$

Таким образом, условная вероятность обладает всеми основными свойствами вероятности.

Очень важную роль играет следующая

Теорема умножения. *Пусть A и B - два события и $P(B) > 0$. Тогда*

$$P(AB) = P(B)P(A|B).$$

Ее доказательство следует из определения условной вероятности. Польза этой теоремы состоит в том, что иногда мы можем вычислить условную вероятность непосредственно и затем использовать это для вычисления $P(AB)$.

Пример 2 . *В урне 5 шаров - 3 белых и 2 черных. Без возвращения выбираем два шара. Найти вероятность того, что оба шара белые.*

Пусть событие A_1 состоит в том, что первый шар белый, а событие A_2 - второй шар белый. Легко вычислить, что $P(A_1) = 3/5$. После того, как мы вынули один шар и знаем, что он белый, мы имеем 4 шара и среди них 2 белых. Тогда $P(A_2|A_1) = 2/4$. По теореме умножения

$$P(A_1A_2) = P(A_1)P(A_2|A_1) = 3/5 \cdot 2/4 = 3/10.$$

Теорему умножения легко распространить на любое конечное число событий.

Следствие 1 . *Пусть $A_1, \dots, A_n$ - случайные события, тогда*

$$P(A_1 \dots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2) \dots P(A_n|A_1 \dots A_{n-1}).$$

Если появление события B не меняет вероятности события A, т. е. $P(A|B) = P(A)$, то такие события естественно назвать независимыми. В этом случае по теореме умножения мы получаем $P(AB) = P(A)P(B)$. Последнее соотношение симметрично относительно A и B и имеет смысл при $P(B) = 0$. Поэтому мы возьмем его в качестве определения.

Определение 2 . *События A и B называются* **независимыми**, *если*

$$P(AB) = P(A)P(B).$$

Пример 3 . *Подбрасывают две симметричных монеты. Событие A состоит в том, что на первой монете выпал герб, а событие B - на второй монете выпал герб.*

Интуитивно ясно, что такие события должны быть независимыми. Действительно, $P(A) = 1/2$, $P(B) = 1/2$, $P(AB) = 1/4 = 1/2 \cdot 1/2 = P(A)P(B)$. Таким образом A и B - независимы в смысле определения . Менее очевидно, что независимы события A и C, где C означает, что выпал только один герб (доказать !).

Сложнее определяется независимость более двух событий.

Определение 3 . *События $A_1, \ldots, A_n$ называем* **независимыми в совокупности**, *если для любого $2 \leq k \leq n$ и любых событий $A_{i_1}, \ldots, A_{i_k}$ из рассматриваемых справедливо*

$$P(A_{i_1} \ldots A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k}).$$

Покажем на примерах, что попарной независимости и выполнения последнего равенства для перечня всех событий недостаточно для независимости в совокупности.

Пример 4 . *Правильный тетраэдр окрашен тремя цветами: одна грань - в синий цвет, вторая - в красный, третья - в зеленый, а на четвертой присутствуют все три цвета. Этот тетраэдр подбрасывают и отмечают, какой гранью он выпал.*

Пусть A_1 означает появление синего цвета, A_2 - красного, A_3- зеленого. Тогда $P(A_1) = 2/4 = 1/2,\ P(A_2) = 1/2,\ P(A_3) = 1/2,\ P(A_1A_2) = P(A_2A_3) = P(A_1A_3) = P(A_1A_2A_3) = 1/4$. Отсюда мы получаем, что $P(A_1A_2) = 1/4 = 1/2 \cdot 1/2 = P(A_1)P(A_2)$. Аналогично для других пар. Таким образом, мы имеем попарную независимость. Но $P(A_1A_2A_3) = 1/4 \neq 1/2 \cdot 1/2 \cdot 1/2 = P(A_1)P(A_2)P(A_3)$.

Задача 1 . *Придумать пример эксперимента и трех событий A_1, A_2, A_3, для которых $P(A_1A_2A_3) = P(A_1)P(A_2)P(A_3)$, но которые не являются попарно независимыми.*

Можно дать следующее более общее

Определение 4 . *Пусть $\mathcal{M}, \ldots, \mathcal{M}_n$- некоторые классы событий. Они называются* **независимыми**, *если любые события $A_1 \in \mathcal{M}_1, \ldots, A_n \in \mathcal{M}_n$ - независимы в совокупности.*

Типичная ситуация описана в следующем примере.

Пример 5 . *Симметричный игральный кубик подбрасывают два раза. $\mathcal{M}_1$ обозначает набор событий, связанных с результатом первого бросания. $\mathcal{M}_2$ определяется аналогично для результата второго бросания. Тогда $\mathcal{M}_1$ и $\mathcal{M}_2$-независимы.*

Во многих задачах является полезным следующий результат.

Предложение 1 . *Если события A и B независимы, то независимы и любые два из следующих: A и $\overline{B}$, $\overline{A}$ и B, $\overline{A}$ и $\overline{B}$.*

Доказательство. Докажем независимость A и $\overline{B}$.

$$\begin{aligned} P(A\overline{B}) &= P(A) - P(AB) = P(A) - P(A)P(B) \\ &= P(A)(1 - P(B)) = P(A)P(\overline{B}). \end{aligned}$$

Независимость остальных пар событий предлагается доказать самостоятельно.

Во многих ситуациях мы встречаемся с такими экспериментами, которые можно разложить на два (или более) этапов. На первом этапе мы имеем несколько вариантов, а спрашивается что-либо о том, что произошло в конце - на втором этапе. В этом случае чрезвычайно полезен приводимый ниже результат. Начнем со следующего определения.

Определение 5 . *События* $H_1, \ldots, H_n$ *образуют* **полную группу событий** *(***разбиение пространства** Ω*), если*

1. $H_i H_j = \emptyset, \; i \neq j,$

2. $H_1 + \ldots + H_n = \Omega.$

Теорема 1 . *Пусть события* $H_1, \ldots, H_n$ *образуют полную группу событий,* $P(H_k) > 0$ *для всех* k *и* A *- произвольное событие. Тогда*

$$P(A) = \sum_{k=1}^{n} P(H_k) P(A|H_k) \; - \; \textbf{формула полной вероятности}.$$

Доказательство. Так как события $H_1, \ldots, H_n$ образуют полную группу, то мы имеем

$$(AH_i) \cap (AH_j) = AH_i H_j = \emptyset$$

и

$$A = A \cap \Omega = A(H_1 + \ldots + H_n) = AH_1 + \ldots + AH_n.$$

Отсюда получаем

$$P(A) = \sum_{k=1}^{n} P(AH_k) = \sum_{k=1}^{n} P(H_k) P(A|H_k),$$

где мы использовали теорему умножения.

Пример 6 . *На некоторой фабрике 30% продукции производится машиной* A*, 25% продукции - машиной* B*, а остальная продукция - машиной* C*. У машины* A *в брак идет 1% производимой ей продукции, у машины - 1,2% , а у машины* C *- 2%. Из всей произведенной продукции случайно выбрано одно изделие. Какова вероятность того, что оно бракованное?*

Пусть H_1 обозначает событие, состоящее в том, что выбранная деталь изготовлена на машине A, H_2 - на машине B, H_3 - на машине C. Обозначим через D событие, состоящее в том, что выбранная деталь бракованная. События H_1, H_2, H_3 образуют полную группу событий. По условию задачи

$$P(H_1) = 0.3,\ P(H_2) = 0.25,\ P(H_3) = 0.45,$$

$$P(D|H_1) = 0.01,\ P(D|H_2) = 0.012,\ P(D|H_3) = 0.02.$$

По формуле полной вероятности получаем

$$\begin{aligned} P(D) &= P(H_1)P(D|H_1) + P(H_2)P(D|H_2) + P(H_3)P(D|H_3) = \\ &= 0.3 \cdot 0.01 + 0.25 \cdot 0.012 + 0.45 \cdot 0.02 = 0.016. \end{aligned}$$

В статистических приложениях обычно приходится решать обратную задачу. Мы имеем несколько вариантов $H_1, \ldots, H_n$ условий проведения эксперимента. Для каждого из этих вариантов на основе прошлой информации нам известна вероятность $P(H_k)$ его реализации в данном испытании. В результате проведения эксперимента мы получили некоторую информацию (произошло событие A). Теперь мы хотим оценить вероятность того, что реализовался вариант H_k. Это можно сделать с помощью следующей теоремы.

Теорема 2 . *Если $H_1, \ldots, H_n$ образуют полную группу событий, A - произвольное событие и $P(H_k) > 0,\ k = \overline{1, n},\ P(A) > 0$, то*

$$P(H_k|A) = \frac{P(H_k)P(A|H_k)}{P(A)} = \frac{P(H_k)P(A|H_k)}{\sum\limits_{j=1}^{n} P(H_j)P(A|H_j)}$$

- формула **Байеса**.

Доказательство. По определению условной вероятности и теореме умножения

$$P(H_k|A) = \frac{P(AH_k)}{P(A)} = \frac{P(H_k)P(A|H_k)}{P(A)}.$$

Для вычисления $P(A)$ используем формулу полной вероятности.

Вероятности $P(H_k)$ называются **априорными вероятностями гипотез** H_k, а $P(H_k|A)$-**апостериорными вероятностями**.

Пример 7 . *В урне находятся три шара, которые могут быть либо черными, либо белыми, но конкретный состав урны не известен. С возвращением выбираем 5 шаров. Среди них оказалось 3 белых (событие A). Какой состав шаров наиболее вероятен?*

Мы имеем три варианта состава урны:
H_0 - нет белых шаров,
H_1 - один белый шар,
H_2 - два белых шара,
H_3 - три белых шара.

Мы предполагаем, что априорные вероятности $P(H_k) = 1/4,\ k = 0, 1, 2, 3$. Выбор производится с возвращением. Используя результат примера 3.1, получаем
$P(A|H_0) = 0,$
$P(A|H_1) = C_5^3(1/3)^3(2/3)^2 = 40/243,$
$P(A|H_2) = C_5^3(2/3)^3(1/3)^2 = 80/243,$
$P(A|H_3) = 0.$
По формуле полной вероятности

$$P(A) = 1/4 \cdot 0 + 1/4 \cdot 40/243 + 1/4 \cdot 80/243 + 1/4 \cdot 0 = 30/243.$$

Окончательно по формуле Байеса получаем

$$P(H_k|A) = \frac{P(H_k)P(A|H_k)}{P(A)},$$

$$P(H_0|A) = 0,\ P(H_1|A) = 1/3,\ P(H_2|A) = 2/3,\ P(H_3|A) = 0.$$

Таким образом, наиболее вероятно, что в урне было два белых шара.

1.5 Последовательности испытаний

1.5.1 Определение последовательности испытаний

До сих пор мы в основном рассматривали случайные экперименты, состоящие как бы из одного этапа. Хотя встречались примеры и "многоступенчатых" экспериментов. В этом разделе мы дадим описание математической модели эксперимента, который состоит из нескольких шагов или этапов.

Пусть $X_1, \ldots, X_n$ - некоторые конечные множества. Элементы множества X_k - это возможные результаты эксперимента на k-м шаге. Тогда в качестве элементарного исхода такого эксперимента, состоящего из n шагов, естественно взять $\omega = (\omega_1, \ldots, \omega_n)$, где $\omega_k \in X_k,\ k = \overline{1, n}$. В этом случае $\Omega = X_1 \times \ldots \times X_n$. Чтобы полностью задать вероятностью модель, необходимо задать вероятности

$$P(\omega) = P(\omega_1, \ldots, \omega_n)$$

элементарных исходов.

Определение 1 . *Вероятностное пространство* (Ω, P)*, где* $\Omega = X_1 \times \ldots \times X_n$*, называется* **последовательностью испытаний** *с множеством значений* X_k *на* k*-м шаге* $(k = \overline{1, n})$.

В этом определении вероятность задана сразу для всей последовательности $\omega = (\omega_1, \ldots, \omega_n)$. Но в реальных задачах результаты эксперимента появляются последовательно шаг за шагом. Поэтому и вероятностную структуру естественно задавать с помощью вероятностей появления ω_k, если мы знаем, что на предыдущих шагах уже появились $\omega_1, \ldots, \omega_{k-1}$. Используя теорему умножения, мы можем записать

$$\begin{aligned} P(\omega_1, \ldots, \omega_n) &= P(\omega_1) \cdot P(\omega_2|\omega_1) \cdot \ldots \cdot P(\omega_n|\omega_1, \ldots, \omega_{n-1}) \\ &= P^{(1)}(\omega_1) \cdot P^{(2)}(\omega_2|\omega_1) \cdot \ldots \cdot P^{(n)}(\omega_n|\omega_1, \ldots, \omega_{n-1}) \ . \end{aligned}$$

$P^{(1)}$ называется **начальным распределением**, а $P^{(2)}, \ldots, P^{(n)}$ - **вероятностями переходов** на соответствующих шагах для последовательности из n испытаний.

Пример (модель Эренфестов). Пусть мы имеем две урны. В начальный момент в первой урне два белых шара, во второй - два черных. Сначала мы выбираем случайно один шар из первой урны, отмечаем его цвет и кладем его во вторую урну. Затем выбираем случайно один шар из второй урны, отмечаем его цвет и кладем в первую урну и т.д. Пусть проводим n испытаний. На каждом шаге $X_k = \{б, ч\}$. На первом шаге с вероятностью 1 выбирают белый шар. После $(k-1)$ шагов мы знаем состав очередной урны и можем найти вероятность появления белого или черного шара, используя классическое определение вероятностей.

Если в этом примере отмечать не цвет появившегося шара, а вновь полученный состав урн (что эквивалентно), то вероятности перехода $P^{(k)}$ будут зависеть от результата эксперимента только на последнем шаге, а не на всех предыдущих шагах. Это свойство имеет место для многих реальных задач. Поэтому целесообразно выделить этот случай отдельно.

Определение 2 . *Последовательность испытаний называется* **цепью Маркова**, *если для всех* $k = \overline{2, n}$ *и всех* $\omega_1, \ldots, \omega_k$

$$P^{(k)}(\omega_k|\omega_1, \ldots, \omega_{k-1}) = P^{(k)}(\omega_k|\omega_{k-1}) \ . \tag{1.1}$$

Еще более простую ситуацию мы получаем, если $P^{(k)}$ вообще не зависят от условий $\omega_1, \ldots, \omega_{k-1}$.

Определение 3 . *Говорят, что мы имеем* **последовательность** n **независимых испытаний**, *если* $\forall \omega = (\omega_1, \ldots, \omega_n)$

$$P(\omega) = P^{(1)}(\omega_1) \cdot \ldots \cdot P^{(n)}(\omega_n) \ . \tag{1.2}$$

Нетрудно проверить, что в этом случае для любого $k = \overline{1, n}$ имеет место

1) $P^{(k)}(\omega_k) \geq 0, \ \omega_k \in X_k$,
2) $\sum\limits_{\omega_k \in X_k} P^{(k)}(\omega_k) = 1$.

Определение 4 . *Вероятностное пространство* (Ω, P) *называется* **последовательностью** n **независимых и одинаковых испытаний**, *если это последовательность независимых испытаний,* $X_1 = \ldots = X_n = X$ *и* $P^{(1)} = \ldots = P^{(n)}$.

Можно показать, что события, связанные с разными испытаниями, являются независимыми.

Простейшим является случай, когда $X = \{0,1\}$, т.е. множество исходов на каждом шаге состоит из двух элементов. Модель такого экперимента называется **схемой Бернулли** и изучается подробно в следующем разделе.

1.5.2 Схема Бернулли. Биномиальное распределение

Начнем с неформального определения. **Схемой Бернулли или последовательностью** n **независимых одинаковых испытаний с двумя исходами** называется случайный эксперимент, в котором:

1) проводится n независимых испытаний;

2) каждое испытание кончается одним из двух исходов (один исход называется "успех" и обозначается 1, а второй - "неуспех" и обозначается 0);

3) вероятность появления "успеха" одна и та же в каждом испытании и равна p.

Числа n и p называются **параметрами** схемы Бернулли. Формальное описание модели такого эксперимента дано в следующем определении.

Определение 5 . **Схемой Бернулли** *с параметрами* n *и* p *называется дискретное вероятностное пространство* (Ω, P), *где* Ω *состоит из элементарных исходов вида* $\omega = (\omega_1, \ldots, \omega_n)$, $\omega_k = 0, 1$, $k = \overline{1, n}$, *а вероятности элементарных исходов* ω *задаются по правилу*

$$P(\omega) = p^m (1-p)^{n-m} \ , \tag{1.3}$$

где m *- число единиц в исходе* ω.

Пример 1 . *Симметричную монету подбрасывают 5 раз и на каждом шаге отмечают какой стороной выпала монета.*

Выпадение герба будем считать "успехом". Это схема Бернулли с параметрами $n = 5$ и $p = 1/2$.

Обычно в рамках схемы Бернулли мы хотим вычислить вероятность не отдельного элементарного исхода, а некоторого более сложного события. Например, в предыдущем примере нас может интересовать вероятность того, что выпало ровно 3 герба. Такой вопрос является наиболее типичным для схемы Бернулли. Пусть A_m есть событие, состоящее в том, что в n независимых испытаниях мы получили ровно m успехов. По определению вероятности события

$$P(A_m) = \sum_{\omega \in A_m} P(\omega) = C_n^m p^m (1-p)^{n-m} \ ,$$

где мы воспользовались тем, что вероятности элементарных исходов в A_m все одинаковые и вычисляются по формуле (3), а число различных исходов равно размещению m единиц по n мест без учета порядка.

Если нас интересует число успехов, а не когда именно они появились, то мы можем построить менее подробную модель.

Определение 6 . **Биномиальной моделью** *с параметрами* n *и* p *называется вероятностное пространство* (Ω, P), *где* $\Omega = \{\omega_0, \ldots, \omega_n\}$ *и*

$$P(\omega_m) = b(n,p,m) = C_n^m p^m (1-p)^{n-m} \ , \tag{1.4}$$

$m = \overline{0,n}, \ 0 < p < 1.$

Вероятности $b(n,p,m)$, вычисляемые по формуле (4), образуют так называемое **биномиальное распределение**. Нетрудно проверить, что они обладают обычными свойствами вероятностей:

1) $b(n,p,m) \geq 0$,

2) $\sum\limits_{m=0}^{n} b(n,p,m) = 1$,

и, кроме того,

3) $b(n,p,m) = b(n,1-p,n-m)$.

Часто в прикладных задачах нас интересует, какое число успехов m_0 имеет наибольшую вероятность. Для этого сравним вероятность двух соседних значений.

$$\frac{b(n,p,m+1)}{b(n,p,m)} = \frac{C_n^{m+1}p^{m+1}(1-p)^{n-m-1}}{C_n^m p^m (1-p)^{n-m}} = \frac{n-m}{m+1}\frac{p}{1-p} \ .$$

Последнее выражение больше 1 при $m < (n+1)p+1$ и меньше 1 при $m > (n+1)p-1$. Это приводит нас к следующим свойствам биномиальных вероятностей:

4) если $m < (n+1)p-1$, то при переходе от m к $m+1$ вероятность возрастает,

5) если $m > (n+1)p-1$, то при переходе от m к $m+1$ вероятность убывает,

6) если число $(n+1)p$ - целое, то имеем два наиболее вероятных значения для числа успехов: $m_0 = (n+1)p-1$ и m_0+1,

если число $(n+1)p$ - дробное, то имеем одно наиболее вероятное значение - $m_0 = [(n+1)p]$.

Полезно нарисовать графики изменения вероятностей $b(n,p,m)$ с изменением m для разных n и p.

Пример 2 . *Симметричную игральную кость подбрасывают 6 раз. Найти вероятность того, что выпадут ровно два герба и наиболее вероятное число появлений шестерки.*

В этой задаче $n = 6, \ p = 1/6$. Тогда

$$b(6,1/2,2) = C_6^2 0.5^2 0.5^4 = 0.235 \ .$$

Число $(n+1)p = 7/6$ является дробным, поэтому наиболее вероятное число появлений шестерки есть $m_0 = [(n+1)p] = 1$.

1.5.3 Предельные теоремы в схеме Бернулли

Выше мы получили формулу, по которой можно рассчитать вероятность того, что в серии из n испытаний Бернулли с вероятностью успеха p будет получено ровно m успехов. А именно

$$b(n,p,m) = C_n^m p^m (1-p)^{n-m} \ .$$

В реальных задачах число испытаний бывает достаточно большим, и производить расчеты по этой формуле становится затруднительным. В этих случаях обычно стараются найти более простые выражения, которые асимптотически эквивалентны точным формулам, когда те или иные параметры меняются определенным образом. Для нашей модели существуют две аппроксимации, которые находят широкие приложения в практических задачах и, как мы увидим позднее, имеют и самостоятельное значение.

Теорема Пуассона. *Пусть мы имеем схему Бернулли с параметрами n и p. Пусть $n \to \infty$, $p \to 0$ так, что $np \to \lambda$, $0 < \lambda < \infty$. Тогда для любого фиксированного m*

$$b(n,p,m) \to \pi_m(\lambda) = \frac{\lambda^m}{m!} e^{-\lambda} \ .$$

Доказательство. Зафиксируем некоторое целое неотрицательное m. Тогда

$$b(n,p,m) = C_n^m p^m (1-p)^{n-m} = \frac{n!}{m!(n-m)!} \cdot p^m (1-p)^{n-m} =$$

$$\frac{1}{m!} \cdot n(n-1)\cdots[n-(m-1)] \cdot p^m (1-p)^{-m} (1-p)^n =$$

$$= \frac{1}{m!} \cdot n(n-1)\cdots[n-(m-1)] \cdot \left[\frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right]^m \times$$

$$\times \left[1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right]^{-m} \cdot \left[1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right]^n$$

$$\to \frac{1}{m!} \cdot \lambda^m \cdot e^{-\lambda} \ .$$

Более аккуратный анализ позволяет доказать, что

$$\sum_{m=0}^{\infty} |b(n,p,m) - \pi_m(\lambda)| \le np^2 \ ,$$

если $np = \lambda$ и $b(n,p,m) = 0$ для $m \ge n$.

Пример 3 . *На некоторой телефонной станции 10 000 номеров. В день через станцию поступает в среднем 30 000 вызовов. Найти вероятность того, что по некоторому конкретному номеру будет ровно два звонка.*

Предположим, что вызов по любому номеру является равновероятным и при каждом вызове номер выбирается независимо от других вызовов. Тогда мы имеем схему Бернулли с параметрами $p = 10^{-4}$, $n = 3 \cdot 10^4$. Используя теорему Пуассона, получаем ($\lambda = 3$)

$$b(n,p,2) = C_{30000}^2 p^2 (1-p)^{29998} \approx \frac{3^2}{2!} e^{-3} = 0.1804 \ .$$

Для распределения Пуассона составлены таблицы.

Другой асимптотический результат получается, если $n \to \infty$, а m выбрано так, что $m \to \infty$, но величина $x_{n,m} = \frac{m-np}{\sqrt{np(1-p)}}$ лежит в некотором фиксированном интервале (a,b), где $-\infty < a < b < \infty$.

Локальная теорема Муавра-Лапласа. *Пусть мы имеем схему Бернулли с параметрами n и p. Тогда при $n \to \infty$ и фиксированном p*

$$b(n,p,m) \sim \frac{1}{\sqrt{np(1-p)}} \varphi(x_{n,m})$$

равномерно по всем m, для которых $-\infty < a < x_{n,m} < b < \infty$. Здесь

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

Для функции $\varphi(x)$ составлены таблицы. Отметим, что $\varphi(x) = \varphi(-x)$.

Пример 4 . *Симметричную монету подбрасывают 100 раз. Найти вероятность того, что герб выпадет ровно 50 раз.*

В этой задаче $n = 100$, $p = 0.5$, $m = 50$. Тогда

$$x_{n,m} = \frac{50 - 100 \cdot 0.5}{\sqrt{100 \cdot 0.5 \cdot 0.5}} = 0 \ .$$

Используя локальную теорему Муавра-Лапласа, получаем

$$b(100, 0.5, 50) \approx \varphi(0) \frac{1}{\sqrt{100 \cdot 0.5 \cdot 0.5}} = \frac{0.3989}{5} \approx 0.08 \ .$$

Как видно из последнего примера, при больших n и фиксированном p вероятности $b(n, p, m)$ для всех значений m очень малы. Поэтому обычно интересуются не тем, какое конкретное число успехов будет в нашем эксперименте, а в каких пределах оно окажется. Например, мы может вычислить вероятность того, что число появлений герба при 100 подбрасываниях будет лежать в пределах от 40 до 60. В таких задачах оказывается полезной следующая

Интегральная теорема Муавра-Лапласа. *Пусть мы имеем схему Бернулли с параметрами n и p. Если $n \to \infty$, а p - фиксировано, то равномерно по всем $m_1 < m_2$*

$$P(m_1 \le S_n < m_2) \sim \Phi(x_{n,m_2}) - \Phi(x_{n,m_1}) \ ,$$

где S_n - число успехов в n испытаниях, а

$$\Phi(x) = \int\limits_{-\infty}^{x} \varphi(y) dy \ .$$

Более того, для любых $m_1 < m_2$ имеет место оценка

$$|P(m_1 \le S_n < m_2) - \Phi(x_{n,m_2}) + \Phi(x_{n,m_1})| \le \frac{p^2 + (1-p)^2}{\sqrt{np(1-p)}} \ .$$

Для функции $\Phi(x)$, называемой **функцией распределения стандартного нормального закона**, составлены таблицы. Она обладает свойствами:

1) $\Phi(-\infty) = 0,\ \Phi(\infty) = 1,\ \Phi(0) = 0.5$,
2) $\Phi(-x) = 1 - \Phi(x)$.
В силу свойства 2 таблицы обычно составляют только для положительных или только для отрицательных x. Так как $\Phi(3.9) = 0.999$, то для $x \geq 4$ можно считать, что $\Phi(x) \approx 1$. Очень часто вместо $\Phi(x)$ используют функцию

$$\Phi_0(x) = \int\limits_0^x \varphi(y)dy \ ,$$

которая называется **функцией Лапласа** или **интегралом вероятностей**. Она обладает следующими свойствами:
1) $\Phi_0(0) = 0,\ \Phi_0(\infty) = 0.5$,
2) $\Phi_0(-x) = -\Phi_0(x)$,
3) $\Phi(x) = \Phi_0(x) + 0.5$.
Для нее также составлены таблицы (для $x > 0$). В силу свойства 3 ее можно использовать в интегральной теореме Муавра-Лапласа вместо функции $\Phi(x)$.

Пример 5 . *Симметричную монету подбрасывают 100 раз. Найти вероятность того, что число появившихся гербов будет лежать в пределах от 40 до 60.*

В этой задаче $n = 100$, $p = 0.5$, $m_1 = 40$, $m_2 = 60$. В силу интегральной предельной теоремы Муавра-Лапласа

$$P(m_1 \leq S_n < m_2) \approx \Phi_0(x_{n,m_2}) - \Phi_0(x_{n,m_1}) \ .$$

$$x_{n,m_2} = \frac{60 - 100 \cdot 0.5}{\sqrt{100 \cdot 0.5 \cdot 0.5}} = \frac{10}{5} = 2 \ ,$$

$$x_{n,m_1} = \frac{40 - 100 \cdot 0.5}{\sqrt{100 \cdot 0.5 \cdot 0.5}} = -\frac{10}{5} = -2 \ ,$$

$$P(40 \leq S_n < 60) \approx \Phi_0(2) - \Phi_0(-2) = 2\Phi_0(2) = 2 \cdot 0.4772 = 0.9544 \ .$$

Применяя эту аппроксимацию, мы допускаем ошибку, которая не превышает величины

$$\frac{0.5^2 + 0.5^2}{\sqrt{100 \cdot 0.5 \cdot 0.5}} = \frac{0.5}{5} = 0.1 \ .$$

Более точный анализ показывает, что эта ошибка гораздо меньше.

Интегральная предельная теорема Муавра-Лапласа является частным случаем более общего результата, называемого **центральной предельной теоремой**, доказательство которого будет приведено позднее.

1.5.4 Полиномиальное распределение

Последняя модель, а именно биномиальное распределение, имеет очевидное обобщение на случай, когда число исходов в каждом испытании одинаково, но, возможно, отлично от двух.

Пусть мы имеем последовательность независимых испытаний, каждое из которых кончается одним из r исходов, вероятности которых равны соответственно $p_1, \ldots, p_r$, и они одни и те же во всех испытаниях. Пусть, далее, m_k есть число появлений k-го исхода в этих n испытаниях. Тогда вектор $m = (m_1, \ldots, m_r)$ дает нам описание того, чем закончился такой эксперимент. Используя те же рассуждения, что и в биномиальной модели, нетрудно доказать, что вероятность такого исхода вычисляется по формуле

$$P(m) = P(m_1, \ldots, m_r) = \frac{n!}{m_1! \ldots m_r!} p_1^{m_1} \ldots p_r^{m_r} \ . \qquad (1.5)$$

Суммируя все вышеизложенное, приходим к определению.

Определение 7 . *Вероятностное пространство* (Ω, P), *в котором элементарные исходы* ω *имеют вид* $\omega = (m_1, \ldots, m_r)$, *где* $(m_1, \ldots, m_r)$ *- целые неотрицательные числа такие, что* $m_1 + \ldots + m_r = n$, *а их вероятности вычисляются по формуле*

$$P(\omega) = P(m_1, \ldots, m_r) = \frac{n!}{m_1! \ldots m_r!} p_1^{m_1} \ldots p_r^{m_r} \ ,$$

называется **полиномиальным распределением**.

Пример 6 . *Симметричный игральный кубик подбрасывают 10 раз. Найти вероятность события* A, *состоящего в том, что выпадут 2 шестерки и одна пятерка.*

В этом эксперименте проводится 10 испытаний, в которых естественно фиксировать три различных исхода: выпали шестерка, пятерка и другая цифра, вероятности которых равны 1/6, 1/6 и 4/6 соответственно. Тогда

$$P(A) = \frac{10!}{2!1!7!}\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{4}{6}\right)^7 = \frac{5}{8}\left(\frac{2}{3}\right)^{10} .$$

При больших n расчеты по формуле (5) становятся затруднительными. В этом случае применяются асимптотические формулы, аналогичные тем, что мы рассматривали для биномиального распределения.

Задача 1 . *Пусть мы провели n испытаний с тремя исходами, вероятности которых равны p_1, p_2 и p_3 соответственно. Предположим, что $n \to \infty$, $p_1 \to 0$, $p_2 \to 0$ таким образом, что $np_1 \to \lambda_1$, $np_2 \to \lambda_2$. Тогда при фиксированных m_1 и m_2*

$$P(m_1, m_2, n - m_1 - m_2) \to \frac{\lambda^{m_1}}{m_1!}e^{-\lambda_1}\frac{\lambda^{m_2}}{m_2!}e^{-\lambda_2} .$$

1.6 Определение вероятностного пространства в общем случае

1.6.1 Постановка задачи

До сих пор мы работали только с простейшей моделью случайного эксперимента, а именно с дискретным вероятностным пространством. Но многие реальные задачи невозможно описать в рамках этой модели, так как в них число элементарных исходов несчетно. Одним из таких примеров была задача на геометрическую вероятность, в которой рассматривается модель случайного выбора точки из некоторой области. Необходимо дать общее определение вероятностного пространства, которое охватывало бы и такие ситуации. Мы начнем с того, что напомним определение вероятностного пространства в слабом смысле.

Определение 1 . **Вероятностным пространством (в слабом смысле)** *называется тройка $(\Omega, \mathcal{A}, P)$, где Ω - произвольное мно-*

жество, $\mathcal{A}$ - алгебра подмножеств пространства Ω, P - вещественная функция, заданная на $\mathcal{A}$ и обладающая свойствами:

1) $P(A) \geq 0, \ \forall A \in \mathcal{A}$,

2) $P(\Omega) = 1$,

3) если $A_1, \ldots, A_n \in \mathcal{A}$ *- попарно несовместны, то*

$$P(\sum_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k) \ .$$

Мы рассматриваем Ω как пространство элементарных исходов некоторого эксперимента, $\mathcal{A}$ - это алгебра интересующих нас событий, $P(A)$ - вероятность события. Чтобы понять, почему такого определения недостаточно для наших целей, рассмотрим простой пример.

Пример 1 . *Из единичного квадрата случайным образом выбирают точку. Событие K состоит в том, что выбранная точка принадлежит кругу с центром $(1/2, 1/2)$ радиуса 1/4.*

Это типичная задача на геометрическое определение вероятностей. Здесь $\omega = (x_1, x_2), \ x_1, x_2 \in [0,1], \ \Omega = [0,1] \times [0,1] = [0,1]^2$. Интересующее нас событие K естественно отождествить с указанным кругом. Тогда по геометрическому определению вероятности

$$P(K) = \frac{S(K)}{S(\Omega)} \ ,$$

где $S(K)$ - площадь круга. Но круг - это довольно сложная фигура, и непросто определить, что такое его площадь. Из школьного курса геометрии мы знаем, как вычислить площадь прямоугольника. Затем мы исчерпываем круг такими прямоугольниками и определяем его площадь как сумму площадей этих прямоугольников. Но таких прямоугольников будет бесконечное число.

Приведенный выше анализ показывает, что для таких событий, как попадание случайной точки в круг K, необходимо применять бесконечное число операций над более простыми событиями (попадание в прямоугольник), а для определения его вероятности нужно более сильное свойство, чем конечная аддитивность, так как мы складываем площади бесконечного числа прямоугольников.

1.6.2 Определение σ-алгебры

Определение 2 . *Система $\mathcal{A}$ подмножеств пространства Ω называется σ-**алгеброй**, если выполнены следующие свойства:*

1) $\Omega \in A$,

2) если $A \in \mathcal{A}$, *то* $\bar{A} \in \mathcal{A}$,

3) если $A_1, A_2, \ldots \in \mathcal{A}$, *то* $\bigcup\limits_{k=1}^{\infty} A_k \in \mathcal{A}$.

Замечание. Можно показать, что применение любых ранее определенных операций над событиями, выполненными не более чем в счетном числе, не выводит нас за пределы σ-алгебры.

Примеры. 1. $\mathcal{A} = \{\emptyset, \Omega\}$ - тривиальная σ-алгебра, соответствует случаю, когда мы ничего не знаем о рассматриваемом эксперименте.

2. $\mathcal{A}$ - σ-алгебра всех подмножеств пространства Ω соотвествует полной информации об эксперименте.

3. Любая конечная алгебра $\mathcal{A}$ (т. е. содержащая конечное число событий) является σ-алгеброй (задача!).

Обычно в реальной задаче мы начинаем с некоторого класса событий $\mathcal{M}$, который, возможно, не является σ-алгеброй. Например, в рассмотренном выше примере это был класс прямоугольников. Хотелось бы дополнить его так, чтобы получить σ-алгебру.

Определение 3 . *Пусть $\mathcal{M}$ - некоторая система подмножеств пространства Ω. Класс $\mathcal{A} = \sigma(\mathcal{M})$ подмножеств пространства Ω называется* σ-**алгеброй, порожденной системой** $\mathcal{M}$, *если выполнены следующие свойства:*

1) $\mathcal{M} \subset \mathcal{A}$,

2) $\mathcal{A}$ - *σ-алгебра* ,

3) если $\mathcal{A}_1$ - некоторая σ-алгебра, содержащая $\mathcal{M}$, то $\mathcal{A} \subset \mathcal{A}_1$.

Замечания. 1. Свойство 3 означает, что $\sigma(\mathcal{M})$ в определенном смысле самая маленькая σ-алгебра, содержащая систему $\mathcal{M}$.

2. Порожденная σ-алгебра всегда существует. Для доказательства нужно рассмотреть семейство всех σ-алгебр, содержащих $\mathcal{M}$

(такие существуют), образовать их пересечение и доказать,что это σ-алгебра (провести подробное доказательство самостоятельно!).

Пример. *Пусть* $\Omega = R^1$, $\mathcal{M}$ *- класс всех интервалов.* $\mathcal{A} = \sigma(\mathcal{M})$ *называется* **борелевской** σ*-алгеброй, а элементы* A *из* $\mathcal{A}$ *называются* **борелевскими** *множествами.*

Если $\Omega = R^n$, *а* $\mathcal{M}$ *- класс всех открытых шаров в* R^n, *то* $\mathcal{A} = \sigma(\mathcal{M})$ - **борелевская** σ*-алгебра подмножеств в* R^n*. Элементы этой* σ*-алгебры называются* **борелевскими подмножествами** *в* R^n.

Если у нас есть вероятностное пространство в слабом смысле, то мы можем расширить алгебру событий до порожденной σ-алгебры. Теперь необходимо продолжить вероятность P на эту более широкую систему. Это можно сделать только в том случае, когда P обладает определенными свойствами непрерывности.

1.6.3 Свойства непрерывности вероятности

Лемма 1 . *Если* $(\Omega, \mathcal{A}, P)$ *- вероятностное пространство в слабом смысле, то следующие свойства вероятности* P *эквивалентны:*

1) если $A_1, A_2, \ldots \in \mathcal{A}$ *- попарно несовместны и* $A = \bigcup\limits_{k=1}^{\infty} A_k \in \mathcal{A}$, *то*

$$P(\sum_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$$

- **счетная аддитивность вероятности**;

2) если $A_1, A_2, \ldots \in \mathcal{A}$, $A_1 \subset A_2 \subset \ldots$ *и* $A = \bigcup\limits_{k=1}^{\infty} A_k \in \mathcal{A}$, *то*

$$P(A_k) \nearrow P(A)$$

- **непрерывность снизу**;

3) если $A_1, A_2, \ldots \in \mathcal{A}$, $A_1 \supset A_2 \supset \ldots$ *и* $A = \bigcap\limits_{k=1}^{\infty} A_k \in \mathcal{A}$, *то*

$$P(A_k) \searrow P(A)$$

- **непрерывность сверху**,

4) если $A_1, A_2, \ldots \in \mathcal{A}$, $A_1 \supset A_2 \supset \ldots$ *и* $A = \bigcap\limits_{k=1}^{\infty} A_k = \emptyset$, *то*

$$P(A_k) \searrow 0$$

- **непрерывность сверху на** $\emptyset$.

Доказательство. 1. $\boxed{1} \Rightarrow \boxed{2}$.

Пусть $A_1, A_2, \ldots \in \mathcal{A}$, $A_1 \subset A_2 \subset \ldots$ и $A = \bigcup\limits_{k=1}^{\infty} A_k \in \mathcal{A}$. Определим $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $\ldots$, $B_k = A_k \setminus A_{k-1}$, $\ldots$. События $B_1, B_2, \ldots$ - попарно несовместны, $A_k = B_1 + \ldots + B_k$ и $A = \sum\limits_{k=1}^{\infty} B_k$. В силу свойства 1

$$P(A_k) = \sum_{k=1}^{n} P(B_k) \nearrow \sum_{k=1}^{\infty} P(B_k) = P(A) \ .$$

2. $\boxed{2} \Rightarrow \boxed{3}$.

Пусть $A_1, A_2, \ldots \in \mathcal{A}$, $A_1 \supset A_2 \supset \ldots$ и $A = \bigcap\limits_{k=1}^{\infty} A_k \in \mathcal{A}$. Определим $B_k = \overline{A_k} \in \mathcal{A}$. Новая последовательность событий обладает свойствами: $B_1 \subset B_2 \subset \ldots$ и $B = \bigcup\limits_{k=1}^{\infty} B_k = \bigcup\limits_{k=1}^{\infty} \overline{A_k} = \overline{\bigcap\limits_{k=1}^{\infty} A_k} = \overline{A}$. Исходя из свойства 2 и используя свойства вероятностей событий, получаем

$$1 - P(A_k) = P(\overline{A_k}) = P(B_k) \nearrow P(B) = P(\overline{A}) = 1 - P(A) \ .$$

Отсюда следует, что $P(A_k) \searrow P(A)$.

3. $\boxed{3} \Rightarrow \boxed{4}$. Эта импликация тривиальна, так как свойство 4 есть частный случай свойства 3.

4. $\boxed{4} \Rightarrow \boxed{1}$.

Пусть $A_1, A_2, \ldots \in \mathcal{A}$ - попарно несовместны и $A = \sum\limits_{k=1}^{\infty} A_k \in \mathcal{A}$. Обозначим $B_n = A_1 + \ldots A_n$, $C_n = \sum\limits_{k=n+1}^{\infty} A_k$. Последовательность событий C_n обладает свойствами: $C_1 \supset C_2 \supset \ldots$ и $\bigcap\limits_{n=1}^{\infty} C_n = \emptyset$. Кроме того, $A = B_n + C_n$. Используя конечную аддитивность вероятности и свойство 4, получаем $P(C_n) \searrow 0$ и, так как $B_n \cap C_n = \emptyset$,

$$P(A) = P(B_n) + P(C_n) = \sum_{k=1}^{n} P(A_k) + P(C_n) = \sum_{k=1}^{\infty} P(A_k) + 0 \ .$$

$\square$

Как мы отмечали выше, нам необходимо продолжить вероятность, заданную на некоторой алгебре событий $\mathcal{A}_0$, на σ-алгебру $\mathcal{A} = \sigma(\mathcal{A}_0)$, порожденную этой алгеброй. Немецкий математик Каратеодори доказал, что если вероятность P обладает свойством счетной аддитивности, то ее можно продолжить с алгебры $\mathcal{A}_0$ на σ-алгебру $\mathcal{A} = \sigma(\mathcal{A}_0)$, причем единственным образом.

Теорема о продолжении вероятности (Каратеодори). Пусть $(\Omega, \mathcal{A}_0, P)$ - вероятностное пространство в слабом смысле и вероятность P обладает свойством счетной аддитивности. Тогда существует единственная счетно-аддитивная вероятность Q на $\mathcal{A} = \sigma(\mathcal{A}_0)$:

$$P(A) = Q(A) \ \ \forall A \in \mathcal{A}_0 \ .$$

Приведенные выше рассуждения и примеры приводят нас к следующему определению.

Определение 4 . Вероятностным пространством *называется тройка $(\Omega, \mathcal{A}, P)$, где Ω - произвольное множество, $\mathcal{A}$ - некоторая σ-алгебра его подмножеств, P - вещественная функция на $\mathcal{A}$:*

1) $P(A) \geq 0, \ \forall A \in \mathcal{A}$,

2) $P(\Omega)$,

3) если $A_1, A_2, \ldots$ - попарно несовместны, то

$$P(\sum_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k) \ .$$

Такое определение впервые было предложено А.Н. Колмогоровым в книге "Основные понятия теории вероятностей", опубликованной в 1933 г. на немецком языке (русский перевод - 1936 г.)

Замечание. Отметим, что все свойства операций над событиями и свойства вероятностей, доказанные ранее, остаются справедливыми.

Примеры. 1. **Дискретное вероятностное пространство**.

Пусть мы имеем пару (Ω, P), где Ω - конечное или счетное множество, P - вещественная функция на Ω:

1) $P(\omega) \geq 0, \ \forall \omega \in \Omega$,

2) $\sum\limits_{\omega \in \Omega} P(\omega) = 1$.

Пусть $\mathcal{A}$ – σ-алгебра всех подмножеств пространства Ω, а P задана на $\mathcal{A}$ по правилу

$$P(A) := \sum_{\omega \in A} P(\omega) \ .$$

Задача. *Доказать, что заданная таким образом вероятность P на $\mathcal{A}$ обладает свойством счетной аддитивности.*

Таким образом, $(\Omega, \mathcal{A}, P)$ удовлетворяет определению 4.

2. Пусть $\Omega \subset R^n$ - ограниченное борелевское множество, $f(x)$ - неотрицательная вещественная функция, заданная на Ω:

$$\int_{\Omega} f(x) dx = 1 \ .$$

Пусть, далее, $\mathcal{A}$–σ-алгебра борелевских подмножеств пространства Ω, а вероятность P задается на $\mathcal{A}$ по правилу:

$$P(A) = \int_{A} f(x) dx \ .$$

Тройка $(\Omega, \mathcal{A}, P)$ определяет вероятностное пространство.

Если $f(x) \equiv C = const$, то приходим вновь к геометрическому определению вероятности.

1.7 Случайная величина и ее распределение

1.7.1 Определение случайной величины и ее распределения

Во многих практических задачах мы имеем дело с такими экспериментами, в которых мы изучаем некоторые числовые характеристики. Приведем несколько примеров из тех, что встречались нам ранее.

Примеры. 1. Симметричную монету подбрасываем три раза и отмечаем число выпавших гербов.

2. Симметричную кость подбрасываем два раза и отмечаем сумму выпавших очков.

3. Пусть мы имеем схему Бернулли с n испытаниями и подсчитываем число успехов.

Во всех этих примерах мы видим, что в результате эксперимента мы получаем некоторое число, которое однозначно определяется элементарным исходом. Это приводит нас к определению случайной величины как функции на пространстве элементарных исходов. Элементарные соображения, связанные с решением практических задач, показывают, что эта функция не может быть произвольной, а должна удовлетворять определенным ограничениям. Действительно, как отмечалось выше, вероятностное пространство есть тройка $(\Omega, \mathcal{A}, P)$. В качестве событий рассматриваются только те подмножества пространства Ω, которые принадлежат σ -алгебре $\mathcal{A}$. Только им мы можем приписать некоторую вероятность. С практической точки зрения хотелось бы, чтобы все множества вида $\{\omega : \ a < \xi(\omega) < b\}$, где ξ - случайная величина, были событиями и им можно было приписать вероятность. Это приводит нас к следующему определению.

Определение 1 . *Пусть $(\Omega, \mathcal{A}, P)$ - вероятностное пространство, а $(R^1, \mathcal{B})$ - вещественная прямая с выделенной на ней борелевской σ-алгеброй подмножеств.* **Случайной величиной** *называется функция $\xi : \ \Omega \to R^1$, которая обладает следующим свойством: $\forall B \in \mathcal{B}$*

$$\xi^{-1}(B) := \{\omega \in \Omega : \ \xi(\omega) \in B\} \in \mathcal{A} \ . \tag{7.1}$$

Такая функция ξ называется **измеримой**. Таким образом, случайными величинами мы будем называть вещественные измеримые функции на пространстве Ω. Случайные величины мы будем обозначать в дальнейшем греческими буквами $\xi, \ \eta, \ \zeta$ и т. д.

Замечание. С практической точки зрения достаточно было бы потребовать выполнения свойства 1 для интервалов, т. е. когда $B = (a, b)$. Но нетрудно доказать, что тогда оно справедливо и для любых борелевских множеств (задача!).

Определение 2 . **Распределением случайной величины** ξ *называется функция* P_ξ, *заданная на борелевской* σ-*алгебре* $\mathcal{B}$ *по правилу:* $\forall B \in \mathcal{B}$

$$P_\xi(B) := P\{\xi \in B\} \ . \tag{7.2}$$

Распределение вероятностей случайной величины ξ показывает, какова вероятность попадания случайной величины в то или иное множество. Необходимо отметить, что наша модель случайной величины как функции на пространстве элементарных исходов - это некоторая абстракция. В реальном эксперименте мы производим измерение и получаем конкретное число. По большому числу независимых измерений мы можем вычислить частоты, а значит, и вероятности попадания в различные множества и больше ничего. Таким образом, объективной характеристикой случайной величины является ее распределение, так как только его мы можем восстановить на основе результатов эксперимента.

Но распределение случайной величины - это довольно сложный объект, так как надо задать вероятность $P_\xi(B)$ для всех борелевских множеств $B \in \mathcal{B}$, которых достаточно много. Для более компактного описания распределения вводится понятие функции распределения.

Определение 3 . **Функция распределения** $F_\xi(x)$ *случайной величины* ξ *определяется по правилу:* $\forall x \in R^1$

$$F_\xi(x) = P(\xi < x). \tag{7.3}$$

Используя свойства вероятностей событий, нетрудно доказать следующее

Предложение 1 . *Если* $F_\xi(x)$- *функция распределения случайной величины* ξ, *то*

1) $\forall x \in R^1,\ 0 \le F_\xi(x) \le 1$*,*
2) если $x_1 \le x_2$*, то* $F_\xi(x_1) \le F_\xi(x_2)$*,*
3) $F_\xi(x)$ *- непрерывна слева,*
4) $\lim\limits_{x\to-\infty} F_\xi(x) = 0,\ \lim\limits_{x\to\infty} F_\xi(x) = 1.$
5) $P(a \le \xi < b) = F_\xi(b) - F_\xi(a).$

Доказательство. Свойство 1 следует из свойств вероятностей событий. Определим событие $A(x) := (\xi < x)$. Если $x_1 \le x_2$, то $A(x_1) \subset A(x_2)$ и

$$F_\xi(x_1) = P(A(x_1)) \le P(A(x_2)) = F_\xi(x_2).$$

Пусть последовательность $\{x_n\}$ монотонно возрастает и $\lim_n x_n = x$. Тогда последовательность событий $\{A(x_n)\}$ также монотонно возрастает и $\bigcup\limits_{n=1}^{\infty} A(x_n) = A(x)$. Используя свойства непрерывности вероятности, получаем

$$F_\xi(x_n) = P(A(x_n)) \to P(A(x)) = F_\xi(x).$$

Аналогично доказывается свойство 4.

Нетрудно заметить, что

$$(a \le \xi < b) = (\xi < b) - (\xi < a).$$

Тогда

$$P(a \le \xi < b) = P(\xi < b) - P(\xi < a) = F_\xi(b) - F_\xi(a).$$

Замечание. Зная функцию распределения $F_\xi(x)$ случайной величины ξ, мы можем восстановить и все распределение. Наметим схему доказательства.

1. Для интервалов вида $[a, b)$ вероятность находится из свойства 5 функции распределения.

2. Если борелевское множество B есть сумма конечного или счетного числа непересекающихся интервалов, то вероятность попадания в такое множество равна сумме вероятностей попадания в составляющие его интервалы.

3. Произвольное борелевское множество можно аппроксимировать (в определенном смысле) множествами из пункта 2 так, что вероятность попадания в это борелевское множество является пределом вероятностей попадания в аппроксимирующие множества (пример - площадь круга).

1.7.2 Классификация распределений

.

В реальных задачах нам редко приходится работать с распределениями общего типа. Чаще всего мы имеем дело с так называемыми дискретными и абсолютно непрерывными распределениями и их смесями. Ниже мы приводим соответствующие определения и примеры.

Определение 4 . *Случайная величина ξ имеет* **дискретное распределение**, *если существует такое конечное или счетное множество $X = \{x_1, x_2, \ldots\}$, что $P(\xi \in X)$. Числа $x_1, x_2, \ldots$ называются* **значениями** *случайной величины ξ, а $p_k = P(\xi = x_k)$ - вероятностями этих значений.*

Предложение 2 . *Пусть случайная величина ξ имеет дискретное распределение с множеством значений $X = \{x_1, x_2, \ldots\}$ и вероятностями этих значений $\{p_k\}$. Тогда*

1. $p_k \geq 0$.

2. $\sum\limits_k p_k = 1$.

3. $\forall B \in \mathcal{B},\ P_\xi(B) = \sum\limits_{x_k \in B} p_k$.

В частности,

4. $\forall x \in R^1,\ F_\xi(x) = \sum\limits_{x_k < x} p_k$.

Обратно

5. $\forall x_n \in X,\ p_n = P(\xi = x_n) = F_\xi(x_n + 0) - F_\xi(x_n)$.

6. $P(a < \xi < b) = \sum\limits_{a < x_k < b} p_k$.

Задача 1 *Доказать предложение 2.*

Из свойства 3 мы видим, что, зная множество значений X и вероятности $\{p_k\}$ всех возможных значений, можно восстановить и все распределение P_ξ. Поэтому пару $(X, \{p_k\})$ называют распределением дискретной случайной величины (что, строго говоря, не совсем верно) и записывают ее в виде **таблицы распределения**

x	x_1	x_2	...	x_n	...
p	p_1	p_2	...	p_n	...

Пример 1 . *Симметричную монету подбрасывают три раза, случайная величина ξ есть число выпавших гербов. Это дискретная величина с таблицей распределения*

x	*0*	*1*	*2*	*3*
p	*1/8*	*3/8*	*3/8*	*1/8*

Определение 5 . *Распределение случайной величины ξ называется* **абсолютно непрерывным**, *если существует такая вещественная функция $\rho_\xi(x)$, что $\forall B \in \mathcal{B}$*

$$P_\xi(B) = P(\xi \in B) = \int_B \rho_\xi(x)dx. \tag{7.4}$$

Функция $\rho_\xi(x)$ называется **плотностью распределения** *случайной величины ξ.*

В дальнейшем абсолютно непрерывные распределения будут называться просто непрерывными.

Предложение 3 . *Пусть случайная величина ξ имеет абсолютно непрерывное распределение с плотностью распределения $\rho_\xi(x)$. Тогда справедливы следующие свойства.*

1. $\forall x \in R^1,\ \rho_\xi(x) \geq 0.$

2. $\int\limits_{-\infty}^{\infty} \rho_\xi(x)dx = 1.$

3. $\forall B \in \mathcal{B},\ P_\xi(B) = P(\xi \in B) = \int\limits_B \rho_\xi(x)dx.$

В частности,

4. $\forall x \in R^1,\ F_\xi(x) = \int\limits_{-\infty}^{x} \rho_\xi(y)dy.$

5. $P(a \le \xi < b) = F_\xi(b) - F_\xi(a) = \int\limits_a^b \rho_\xi(x)dx$.

Обратно

6. $\forall x \in R^1$, *где* $\rho_\xi(x)$ *непрерывна* $\rho_\xi(x) = \frac{d}{dx}F_\xi(x)$.

7. $\forall x \in R^1$ $P(\xi = x) = 0$.

Задача 2 . *Доказать предложение 3.*

Свойства 1 и 2 являются характеристическими свойствами плотности распределения. Любая функция $\rho(x)$, обладающая свойствами 1 и 2, является плотностью распределения некоторой случайной величины ξ. Из свойства 3 мы видим, что, зная плотность распределения, мы можем восстановить и все распределение.

Пример 2 . *Из отрезка* $[0,1]$ *случайным образом выбирают точку,* ξ*-координата выбранной точки.*

Используя геометрическое определение вероятности, получаем

$$F_\xi = \begin{cases} 0 & , \quad x \le 0, \\ x & , \quad 0 < x \le 1, \\ 1 & , \quad x > 1, \end{cases}$$

и

$$\rho_\xi(x) = \begin{cases} 1 & , \quad x \in [0,1], \\ 0 & , \quad x \notin [0,1]. \end{cases}$$

Это равномерное распределение на отрезке $[0,1]$.

Кроме дискретных и абсолютно непрерывных, существуют еще так называемые **сингулярные** распределения. В нашем курсе мы не будем рассматривать распределения такого типа.

Определение 6 . *Пусть* $X = \{x_1, x_2, \cdots\}$ *- конечное или счетное множество,* $\{p_k\}$ *- некоторый набор положительных чисел, а* $\rho_\xi(x)$ *- некоторая неотрицательная функция. Распределение случайной величины называется* **смешанным***, если*

$$P_\xi(B) = P(\xi \in B) := \sum_{x_k \in B} p_k + \int\limits_B \rho_\xi(x)dx.$$

$(X, \{p_k\})$ *называется* **дискретной компонентой**, *а плотность* $\rho_\xi(x)$ - **непрерывной компонентой** *распределения с.в.* ξ. *Числа*

$$\alpha_d := \sum_k p_k,\ \ \alpha_c := \int_{-\infty}^{\infty} \rho_\xi(x)dx$$

называются **весами** *соответствующих компонент. Ясно, что* $\alpha_d, \alpha_c \geq 0,\ \ \alpha_d + \alpha_c = 1$. *Если распределение содержит только одну компоненту, то оно называется* **чистым**.

Пример 3 . *Из отрезка* $[0,1]$ *случайным образом выбираем точку* ω.

$$\xi = \begin{cases} 0 & , \quad \omega < 1/4, \\ \omega - 1/4 & , \quad 1/4 \leq \omega < 3/4, \\ 1/2 & , \quad \omega \geq 3/4. \end{cases}$$

Такого типа преобразования часто применяются в теории страхования. Случайная величина ξ имеет две дискретные точки $x_1 = 0$ и $x_2 = 1/2$ с вероятностями $p_1 = P(\xi = 0) = 1/4,\ p_2 = P(\xi = 1/2) = 1/4$ и равномерное распределение на отрезке [1/4,3/4], т.е. $\rho_\xi(x) = 1$ при $x \in [1/4, 3/4]$. Веса компонент равны $\alpha_d = 1/2$ и $\alpha_c = 1/2$.

1.7.3 Примеры стандартных распределений

В этом разделе мы приведем несколько примеров дискретных и непрерывных распределений, которые применяются как при исследовании теоретических вопросов, так и при решении практических задач.

1. Случайная величина ξ имеет **вырожденное в точке** x **распределение**, если $P(\xi = x) = 1$.

Фактически мы имеем не случайную величину, а константу.

2. Случайная величина ξ называется **индикатором события** A, если

$$\xi = (\omega) = I_A(\omega) = \begin{cases} 1 & , \quad \omega \in A, \\ 0 & , \quad \omega \notin A. \end{cases}$$

Это дискретная случайная величина с множеством значений $X = \{0, 1\}$ и вероятностями значений $p = P(\xi = 1)$ и $q = 1 - p = P(\xi = 0)$. Ее распределение называется **распределением Бернулли** с параметром p. Обозначение: $\xi \in Bi(1, p)$.

Индикатор события является в некотором смысле "элементарным кирпичиком" при построении произвольных дискретных случайных величин.

Задача 3 . *Пусть ξ есть дискретная случайная величина с множеством значений $X = \{x_1, x_2, \cdots\}$ и вероятностями значений $p_k = P(\xi = x_k)$. Обозначим $A_k = (\xi = x_k)$. Тогда*

1) $A_i A_j = \emptyset, \ i \neq j$,

2) $A_1 + A_2 + \cdots = \Omega$,

3) $\xi(\omega) = \sum\limits_k x_k I_{A_k}(\omega)$.

3. Дискретная случайная величина ξ имеет **биноминальное распределение** с параметрами n и p, если $X = \{0, 1, \cdots, n\}$ и

$$P(\xi = m) = C_n^m p^m (1 - p)^{n-m}.$$

В этом случае будем писать $\xi \in Bi(n, p)$. Здесь n - целое число, $n \geq 1, \ 0 < p < 1$. Эта случайная величина есть число успехов в n независимых испытаниях схемы Бернулли. Такие величины часто появляются в теории страхования, социологии, экономике, физике и других науках.

4. Дискретная случайная величина ξ имеет **геометрическое распределение** с параметром p, если $X = \{0, 1, 2, \ldots\}$ и

$$P(\xi = m) = p(1 - p)^m, \ m = 0, 1, \ldots .$$

Здесь $0 < p < 1$. Эта случайная величина равна числу испытаний в схеме Бернулли, предшествующих появлению первого успеха.

5. Дискретная случайная величина ξ имеет **отрицательное биноминальное распределение** с параметрами r и p, $r \geq 1$-целое, $0 < p < 1$, если $X = \{0, 1, 2, \ldots\}$ и

$$P(\xi = m) = C_m^{r-1} p^r (1 - p)^{m-r+1}.$$

При $r = 1$ получаем геометрическое распределение. Эти случайные величины равны числу испытаний в схеме Бернулли, предшествующих появлению r-го успеха. Это распределение часто используется в теории страхования при описании числа исков, поступивших в страховую компанию за определенный промежуток времени.

6. Дискретная случайная величина ξ имеет **распределение Пуассона** с параметром λ, если $X = \{0, 1, 2, \ldots\}$ и

$$P(\xi = m) = \frac{\lambda^m}{m!} e^{-\lambda}, \ m = 0, 1, \ldots .$$

Появляется как предельный случай для биноминального распределения, если $p \to 0, n \to \infty, np = \lambda$. Часто используется в теории страхования, теории массового обслуживания, теории надежности и других прикладных разделах теории вероятностей. Описывает, как правило, число исков, заявок, отказов, поступивших за определенный промежуток времени.

7. Случайная величина ξ имеет **равномерное** на отрезке $[a, b]$ распределение, если у нее существует плотность

$$\rho_\xi(x) = \begin{cases} \frac{1}{b-a} & , \quad x \in (a, b) \ , \\ 0 & , \quad x \notin (a, b) \ . \end{cases}$$

Эта модель часто используется для описания распределения случайного момента времени, если известно, что он меняется в ограниченном интервале.

8. Случайная величина ξ имеет **показательное распределение** с параметром λ, если она обладает плотностью

$$\rho_\xi(x) = \begin{cases} \lambda e^{-\lambda x} & , \quad x > 0, \\ 0 & , \quad x \leq 0. \end{cases}$$

Это распределение обладает целым рядом замечательных свойств и часто используется при описании времени между поступлениями двух последовательных заявок, исков, отказов и т.п.

9. Случайная величина ξ имеет **гамма-распределение** с параметрами (α, β), если оно обладает плотностью

$$\rho_\xi(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & , \quad x > 0, \\ 0 & , \quad x \le 0, \end{cases}$$

где

$$\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1} e^{-x} dx$$

есть гамма-функция Эйлера. Напомним, что гамма-функция обладает следующими свойствами:

1) $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$,

2) $\Gamma(1/2) = \sqrt{\pi},\ \Gamma(1) = 1$,

3) $\Gamma(n+1) = n!$.

Это распределение находит применение в теории массового обслуживания, теории надежности, теории страхования и риска и других прикладных разделах теории вероятностей. При $\alpha = 1,\ \beta = \lambda$ получаем показательное распределение с параметром λ.

10. Случайная величина ξ имеет **нормальное (гауссовское)** распределение с параметрами (a, σ^2), $a \in R^1, \sigma^2 > 0$, если оно обладает плотностью

$$\rho_\xi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}},\ x \in R^1.$$

Мы будем использовать обозначение $\xi \in N(a, \sigma^2)$.

Если $a = 0,\ \sigma^2 = 1$, то мы имеем **стандартное нормальное распределение**. В этом случае

$$\rho_\xi(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},\ x \in R^1.$$

Функция

$$F_\xi(x) = \Phi(x) = \int\limits_{-\infty}^x \varphi(y) dy$$

есть **функция распределения** стандартного нормального закона. Часто используется другая функция

$$\Phi_0(x) = \int\limits_0^x \varphi(y)dy,$$

которая называется **функцией Лапласа** или **интегралом вероятностей**. Справедливы следующие соотношения:

1) $\Phi(x) = \Phi_0(x) + 1/2$,

2) $\Phi_0(+\infty) = 1/2$,

3) $\Phi_0(-x) = -\Phi_0(x)$.

Для функций $\varphi(x)$, $\Phi(x)$ и $\Phi_0(x)$ составлены подробные таблицы (смотри, например, Большев Л.Н., Смирнов Н.И. "Таблицы математической статистики"). Ниже будет показано: если $\xi \in N(a, \sigma^2)$, то случайная величина

$$\xi_0 = \frac{\xi - a}{\sigma}$$

имеет стандартное нормальное распределение. Таким образом, зная $\Phi(x)$ или $\Phi_0(x)$, мы можем вычислять вероятности для случайных величин с произвольным нормальным распределением. Например,

$$P(|\xi - a| \leq 3\sigma) = \Phi(3) - \Phi(-3) = 2\Phi_0(3) \approx 0.9973,$$

т. е. практически достоверно, что случайная величина ξ отклоняется от точки a на расстояние не более 3σ. Этот результат часто применяется на практике и носит название "правило трех σ".

1.7.4 Функциональные преобразования случайной величины.

Одной из наиболее важных прикладных задач, связанных со случайными величинами, является задача нахождения распределения функции от случайной величины. При этом в качестве функционального преобразования нужно брать такое, чтобы мы получили вновь случайную величину.

Определение 7 . *Функция* $g : R^1 \to R^1$ *называется* **борелевской**, *если* $\forall B \in \mathcal{B}$ *множество*

$$g^{-1}(B) := \{x \in R^1 : g(x) \in B\}$$

также является борелевским.

Если $\xi : \Omega \to R^1$ есть случайная величина, а $g : R^1 \to R^1$ - борелевская функция, то $\eta = f(\xi)$ также является случайной величиной. Действительно,

$$\Omega \xrightarrow{\xi} R^1 \xrightarrow{g} R^1 .$$

Тогда $\eta^{-1}(B) = \xi^{-1}(g^{-1}(B))$ является случайным событием для любого $B \in \mathcal{B}$.

Найдем распределение с.в. η. По определению

$$P_\eta(B) = P(\eta \in B) = P(g(\xi) \in B) = P(\xi \in g^{-1}(B)) = P_\xi(g^{-1}(B))$$

т. е.

$$P_\eta(B) = P_\xi(g^{-1}(B)) .$$

Покажем, как это можно записать в терминах функций распределения. В общем случае это сделать не удается, но в некоторых специальных случаях можно получить удовлетворительные результаты.

Предложение 4 . *Пусть $y = g(x)$ - строго монотонная функция. Тогда для $\eta = g(\xi)$ мы имеем*

$$F_\eta(y) = F_\xi(g^{-1}(y)) ,$$

если $g(x)$ строго возрастает, и

$$F_\eta(y) = 1 - F_\xi(g^{-1}(y) + 0) ,$$

если $g(x)$ строго убывает.

Доказательство. Докажем результат для случая строго возрастающей функции $g(x)$. Второй случай доказывается аналогично. По определению функции распределения

$$F_\eta(y) = P(\eta < y) = P(g(\xi) < y) = P(\xi < g^{-1}(y)) = F_\xi(g^{-1}(y)) .$$

Примеры. 1. $y = g(x) = \sigma x + a,\ \sigma > 0,\ a \in R^1$. Тогда для $\eta = \sigma\xi + a$ получаем

$$F_\eta(y) = P(\sigma\xi + a < y) = P\left(\xi < \frac{y-a}{\sigma}\right) = F_\xi\left(\frac{y-a}{\sigma}\right) .$$

В частности, если ξ имеет плотность $\rho_\xi(x)$, то существует

$$\rho_\eta(y) = \frac{d}{dy}F_\eta(y) = \frac{d}{dy}F_\xi\left(\frac{y-a}{\sigma}\right) = \frac{1}{\sigma}\rho_\xi\left(\frac{y-a}{\sigma}\right) .$$

2. Если $\sigma \in N(0,1)$, то $\eta = \sigma\xi + a \in N(a, \sigma^2)$.

Рассмотрим теперь подробнее случаи дискретного и абсолютно непрерывного распределений.

Если ξ имеет дискретное распределение с множеством значений $x = \{x_1, x_2, \ldots\}$ и вероятностями появления значений $p_n = P(\xi = x_n)$, то случайная величина η также будет иметь дискретное распределение с множеством значений $Y = \{y_1, y_2, \ldots\}$, где $y_k = g(x_m)$ для некоторого x_m, и вероятностями значений

$$q_k = P(\eta = y_k) = \sum_{x_n : g(x_n) = y_k} p_n .$$

Пример. Случайная величина ξ имеет дискретное распределение следующего вида:

x	-2	-1	0	1	2
p	0.3	0.1	0.1	0.3	0.2

Тогда случайная величина $\eta = \xi^2$ является дискретной и принимает значения 0,1 и 4 с вероятностями

$$P(\eta = 0) = P(\xi = 0) = 0.1 ,$$

$$P(\eta = 1) = P(\xi = -1) + P(\xi = 1) = 0.4 ,$$
$$P(\eta = 4) = P(\xi = -2) + P(\xi = 2) = 0.5 .$$

Если ξ имеет непрерывное распределение с плотностью $\rho_\xi(x)$, то распределение случайной величины $\eta = g(\xi)$ может и не являться непрерывным. Более того, оно может быть даже дискретным.

Пример. ξ имеет равномерное на $[0,1]$ распределение,

$$g(x) = \begin{cases} 0 \quad , \quad x \in [0, 0.5] \ , \\ 1 \quad , \quad x \in [0.5, 1] \ . \end{cases}$$

Случайная величина $\eta = g(\xi)$ принимает два значения 0 и 1 с равными вероятностями 1/2.

Таким образом, на функцию $g(x)$ необходимо наложить некоторые дополнительные ограничения. Один частный, но практически важный случай рассматривается в следующей теореме.

Теорема 1 . *Пусть с.в. ξ имеет непрерывное распределение с плотностью $\rho_\xi(x)$, $y = g(x)$ - строго монотонная дифференцируемая функция. Тогда с.в. $\eta = g(\xi)$ также имеет непрерывное распределение и ее плотность имеет вид*

$$\rho_\eta(y) = \rho_\xi(g^{-1}(y)) \left| \frac{d}{dx} g(x)_{x=g^{-1}(y)} \right|^{-1} \ .$$

Доказательство. Проведем доказательство для случая, когда $y = g(x)$ строго возрастает. Пусть $y \in R^1$ такая точка, что $0 < P(\eta < y) < 1$. Тогда

$$F_\eta(y) = P(\eta < y) = P(g(\xi) < y) = P(\xi < g^{-1}(y)) = F_\xi(g^{-1}(y)) \ .$$

Используя правила дифференцирования сложной и обратной функций, получаем

$$\rho_\eta(y) = \frac{d}{dy} F_\eta(y) = \rho_\xi(g^{-1}(y)) \left(\frac{d}{dx} g(g^{-1}(y)) \right)^{-1} \ .$$

Доказательство для строго убывающей функции аналогично.

Аналогичную теорему можно сформулировать и для случая немонотонных функций $y = g(x)$. Но такого рода результат редко используется в реальных задачах. Обычно легче провести заново все расчеты в каждом конкретном случае.

Пример. *С.в. $\xi \in N(0,1)$, найти распределение с.в.$\eta = \xi^2$.*

Новая случайная величина $\eta = \xi^2$ принимает только положительные значения. Пусть $y > 0$.

$$F_\eta = P(\eta < y) = P(\xi^2 < y) = P(|\xi| < \sqrt{y})$$

$$= P(-\sqrt{y} < \xi < \sqrt{y}) F_\xi(\sqrt{y}) - F_\xi(-\sqrt{y}).$$

$$\rho_\eta(y) = \frac{d}{dy} F_\eta(y) = \rho_\xi(\sqrt{y}) \frac{1}{2\sqrt{y}} + \rho(-\sqrt{y}) \frac{1}{2\sqrt{y}} =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} e^{-\frac{1}{2}y} .$$

1.8 Случайный вектор

1.8.1 Распределение случайного вектора

Во многих реальных задачах мы имеем не одну, а несколько случайных величин в одном и том же эксперименте. Иногда их удобно рассматривать как единый объект. Это приводит нас к следующему определению.

Определение 1 . *n-мерным случайным вектором называется набор $\xi = (\xi_1, \ldots, \xi_n)$ случайных величин, заданных на одном и том же вероятностном пространстве $(\Omega, \mathcal{A}, P)$.*

Фактически случайный вектор ξ есть отображение $\xi : \Omega \to R^n$. Нетрудно показать (задача !), что это отображение является **борелевским**, т.е. для любого борелевского подмножества $B \subset R^n$ (σ-алгебру всех борелевских подмножеств в R^n мы будем обозначать $\mathcal{B}_n$) мы имеем $\xi^{-1}(B) \in \mathcal{A}$. Как и для случайных величин, можно дать следующее

Определение 2 . *Распределением случайного вектора ξ называется функция P_ξ , заданная на σ-алгебре $\mathcal{B}_n$ по правилу*

$$P_\xi(B) = P(\xi \in B).$$

Распределение является объективной характеристикой случайного вектора, которую можно однозначно восстановить из эксперимента. Но распределение, будучи удобной характеристикой в теоретических исследованиях, является довольно сложным для реальных задач. Как и в одномерном случае, используют понятие функции распределения.

Определение 3 . **Функцией распределения** *случайного вектора $\xi = (\xi_1, \dots, \xi_n)$ называется функция $F_\xi(x)$, $x \in R^n$, такая, что $\forall x = (x_1, \dots, x_n) \in R^n$,*

$$F_\xi(x_1, \dots, x_n) = P(\xi_1 < x_1, \dots, \xi_n < x_n).$$

Основные свойства функции распределения случайного вектора собраны в следующем предложении.

Предложение 1 . *Функция распределения $F_\xi(x) = F_\xi(x_1, \dots, x_n)$ случайного вектора $\xi = (\xi_1, \dots, \xi_n)$ обладает следующими свойствами:*

1. $\forall x \in R^n$, $0 \le F_\xi(x) \le 1$.

2. $F_\xi(x)$ *не убывает по каждому аргументу* x_i, $i = \overline{1,n}$.

3. $F_\xi(x)$ *- непрерывна слева по каждому аргументу* x_i, $i = \overline{1,n}$.

4. $F_\xi(x) \to 0$, *если* **некоторое** $x_i \to -\infty$, $i = \overline{1,n}$.
$F_\xi(x) \to 1$, *если* **все** $x_i \to \infty$, $i = \overline{1,n}$.

5. $\forall x = (x_1, \dots, x_n) \in R^n$

$$P(x_1 \le \xi_1 < x_1 + h_1, \dots, x_n \le \xi_n < x_n + h_n) =$$
$$\Delta h_1 \dots \Delta h_n F_\xi(x_1, \dots, x_n) \ ,$$

где

$$\Delta h_i F_\xi(x_1, \dots, x_n) =$$
$$F_\xi(x_1, \dots x_i + h_i, \dots, x_n) - F_\xi(x_1, \dots x_i, \dots, x_n) \ .$$

6. $F_\xi(x_1, \dots x_{i-1}, \infty, x_{i+1}, \dots, x_n)$ *есть функция распределения случайного вектора* $(\xi_1, \dots \xi_{i-1}, \xi_{i+1}, \dots, \xi_n)$.

Задача 1 . *Доказать предложение 1.*

Замечание. В силу свойства 5 по функции распределения $F_\xi(x)$ можно найти вероятности попадания в множества $B = [a_1, b_1) \times \ldots \times [a_n, b_n)$. Далее, так же как и в одномерном случае, можно восстановить распределение ξ для любых борелевских множеств B, аппроксимируя их параллелограммами.

1.8.2 Классификация распределений

Как и в одномерном случае, мы выделим два важных частных случая распределений, которые наиболее часто используются на практике. Конечно, бывают и более общие примеры, но мы не будем их подробно рассматривать в нашем курсе.

Определение 4 . *Случайный вектор $\xi = (\xi_1, \ldots, \xi_n)$ имеет* **дискретное распределение**, *если существует конечное или счетное множество $X \subset R^n$, такое, что $P(\xi \in X) = 1$.*

Если $x = (x_1, \ldots, x_n) \in X$- одно из возможных значений случайного вектора ξ, то $p(x) = P(\xi = x)$ называется **вероятностью появления значения** *x.*

Обычно используют следующую стандартную форму описания распределения дискретного случайного вектора. Ясно, что каждая координата ξ_k случайного вектора ξ имеет дискретное распределение. Пусть $X^{(k)} = \{x_1^{(k)}, x_2^{(k)}, \ldots, x_{i_k}^{(k)}, \ldots\}$ есть множество значений случайной величины ξ_k. Образуем множество

$$X = X^{(1)} \times \ldots X^{(n)} = \{x = (x_1, \ldots, x_n) : x_k \in X^{(k)},\ k = \overline{1, n}\}$$

в R^n.

Задача. Доказать, что $P(\xi \in X) = 1$, т.е. X можно взять в качестве множества значений случайного вектора ξ.

Для произвольного вектора $x = (x_{i_1}^{(1)}, \ldots, x_{i_n}^{(n)})$, где $x_{i_k}^{(k)} \in X^{(k)}$, обозначим через

$$P_{i_1 \ldots i_n} = P(\xi = x) = P(\xi_1 = x_{i_1}^{(1)}, \ldots, \xi_n = x_{i_n}^{(n)}) \tag{8.1}$$

вероятность появления значения x случайного вектора ξ. При таком выборе множества X некоторые его элементы будут появляться с вероятностью 0.

Пример. Случайный вектор $\xi = (\xi_1, \xi_2)$ имеет два значения $(1, 1)$ и $(2, 2)$, которые появляются с вероятностями $1/2$. Точка $x = (1, 2)$ входит в построенное выше множество X, но
$P(\xi = x) = 0$.

Пару $(X, \{P_{i_1 \ldots i_n}\})$ будем называть **распределением дискретного случайного вектора** ξ, хотя, строго говоря, это не совсем точно. Для $n = 2$ распределение дискретного случайного вектора обычно задают в виде следующей таблицы, называемой
таблицей распределения:

$\xi_2 \setminus \xi_1$	x_1	x_2	$\ldots$	x_i	$\ldots$
y_1	P_{11}	P_{21}	$\ldots$	P_{i1}	$\ldots$
y_2	P_{12}	P_{22}	$\ldots$	P_{i2}	$\ldots$
$\vdots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
y_j	P_{1j}	P_{2j}	$\ldots$	P_{ij}	$\ldots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Здесь $\{x_1, x_2, \ldots, x_i, \ldots\}$ - множество значений для ξ_1, $\{y_1, y_2, \ldots, y_j, \ldots\}$ - множество значений для ξ_2, а $P_{ij} = P(\xi_1 = x_i, \xi_2 = y_j)$.

Предложение 2 . *Распределение* $(X, \{P_{i_1,\ldots,i_n}\})$ *дискретного случайного вектора* $\xi = (\xi_1, \ldots, \xi_n)$ *обладает следующими свойствами:*

1) $\forall x = (x^{(1)}_{i_1}, \ldots, x^{(n)}_{i_n}) \in X \quad P_{i_1,\ldots,i_n} \geq 0$,

2) $\sum\limits_{i_1,\ldots,i_n} P_{i_1,\ldots,i_n} = 1$,

3) $\forall B \in \mathcal{B}_n$

$$P_\xi(B) = P(\xi \in B) = \sum_{(x^{(1)}_{i_1},\ldots,x^{(n)}_{i_n}) \in B} P_{i_1,\ldots,i_n} \ ,$$

4) $P(\xi_1 = x^{(1)}_{i_1}, \ldots \xi_{k-1} = x^{(k-1)}_{i_{k-1}}, \xi_{k+1} = x^{(k+1)}_{i_{k+1}}, \ldots, \xi_n = x^{(n)}_{i_n}) =$

$$= \sum_{i_k} P_{i_1,\ldots,i_{k-1},i_k,i_{k+1},\ldots,i_n} \ .$$

Все эти свойства легко следуют из приведенных выше определений и свойств вероятностей. Поэтому доказательство этого предложения предлагается в виде задачи.

Пример. Пусть мы приводим n независимых испытаний, каждое из которых может закончиться одним из r исходов $(r \geq 2)$ и вероятности появления этих исходов одни и те же в каждом испытании и равны $p_1, \ldots, p_r$. Пусть ξ_k есть число появлений k-го исхода в этих n испытаниях. Тогда $\xi = (\xi_1, \ldots, \xi_r)$ есть дискретный случайный вектор. Его значениями являются векторы $m = (m_1, \ldots, m_r)$, такие, что m_k - целые неотрицательные числа и $m_1 + \ldots m_r = n$. Как было показано выше, при изучении последовательностей независимых испытаний

$$P(\xi = m) = \frac{n!}{m_1! \ldots m_r!} p_1^{m_1} \ldots p_r^{m_r} \ . \tag{8.2}$$

Такое распределение называется **полиномиальным распределением** с параметрами $(n; p_1, \ldots, p_r)$.

Определение 5 . *Случайный вектор* $\xi = (\xi_1, \ldots, \xi_n)$ *имеет* **абсолютно непрерывное распределение**, *если существует вещественная функция* $\rho_\xi(x)$, $x \in R^n$, *такая, что* $\forall B \in \mathcal{B}_n$

$$P_\xi(B) = P(\xi \in B) = \int\limits_B \rho_\xi(x) dx \ .$$

Функция $\rho_\xi(x)$ *называется* **плотностью распределения** *случайного вектора* ξ.

Нетрудно доказать следующее утверждение, доказательство которого предлагается в качестве задачи.

Предложение 3 . *Случайный вектор* $\xi = (\xi_1, \ldots, \xi_n)$ *имеет абсолютно непрерывное распределение с плотностью* $\rho_\xi(x)$, $x \in R^n$. *Тогда справедливы следующие свойства:*

1) $\forall \in R^n \ \rho_\xi(x) \geq 0$;

2) $\int\limits_{R^n} \rho_\xi(x) dx = 1$;

3) $\forall B \in \mathcal{B}_n \; P(\xi \in B) = \int\limits_B \rho_\xi(x)dx$ *;*

4) $\forall x = (x_1, \ldots, x_n) \in R^n$

$$F_\xi(x_1, \ldots, x_n) = \int\limits_{-\infty}^{x_1} \cdots \int\limits_{-\infty}^{x_n} \rho_\xi(y_1, \ldots, y_n)dy_n \ldots dy_1 \; ;$$

5) если $(x_1, \ldots, x_n)$ *- точка непрерывности плотности* $\rho_\xi(x)$ *, то*

$$\rho_\xi(x_1, \ldots, x_n) = \frac{\partial^n F_\xi(x_1, \ldots, x_n)}{\partial x_1 \ldots \partial x_n} \; ;$$

6) плотность случайного вектора $\tilde{\xi} = (\xi_1, \ldots, \xi_{k-1}, \xi_{k+1}, \ldots, \xi_n)$ *можно вычислить по формуле*

$$\rho_{\tilde{\xi}}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) =$$

$$\int\limits_{-\infty}^{\infty} \rho_\xi(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n)dx_k \; .$$

Замечание. Если мы имеем некоторый случайный вектор $\xi = (\xi_1, \ldots, \xi_n)$, то, выбирая некоторые из его координат, например первые m, мы получаем новый случайный вектор $\tilde{\xi} = (\xi_1, \ldots, \xi_m)$, который называют **подвектором** вектора ξ. Выше было показано, как найти распределение подвектора, когда убирают одну из координат. Применяя эту процедуру несколько раз, мы сможем найти распределение произвольного подвектора. Распределение отдельно взятой координаты ξ_i вектора ξ называется **одномерным** или **маргинальным** распределением.

Как и в одномерном случае, можно ввести понятие смеси распределений, но мы не будем его рассматривать подробно так как здесь не возникает ничего нового.

Примеры. 1. Случайный вектор $\xi = \xi_1, \ldots, \xi_n)$ имеет равномерное распределение в области D, если он обладает плотностью распределения следующего вида:

$$\rho_\xi(x) = \begin{cases} \frac{1}{L(D)} & , \quad x \in D \; , \\ 0 & , \quad x \notin D \; , \end{cases}$$

где $L(D)$ - мера Лебега области D. Фактически мы имеем дело с геометрическим определением вероятности.

2. Случайный вектор $\xi = (\xi_1, \xi_2)$ имеет **двумерное нормальное распределение**, если он обладает плотностью распределения следующего вида:

$$\rho_\xi(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}}\times$$

$$\exp\{-\frac{1}{2(1-\rho^2)}[\frac{(x_1-a_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-a_1)}{\sigma_1}\frac{(x_2-a_2)}{\sigma_2} + \frac{(x_1-a_1)^2}{\sigma_1^2}]\} \ .$$

Числа $a_1, a_2 \in R^1$, $\sigma_1^2, \sigma_2^2 > 0$, $\rho \in (-1, 1)$ называются **параметрами** двумерного нормального распределения. Их вероятностный смысл будет выяснен позднее.

1.8.3 Независимые случайные величины

При изучении свойств вероятностей случайных событий мы видели, что понятие независимости событий играет важную роль при вычислении вероятностей сложных событий. Аналогично понятие независимости является центральным понятием в теории случайных величин, их функциональных преобразований и других вопросах.

Определение 6 . *Случайные величины $\xi_1, \ldots, \xi_n$ называются* **независимыми***, если для любых борелевских $B_1, \ldots, B_n \in \mathcal{B}_1$*

$$P(\xi_1 \in B_1, \ldots, \xi_n \in B_n) = P(\xi_1 \in B_1) \cdot \ldots \cdot P(\xi_n \in B_n) \ . \tag{8.3}$$

Дадим эквивалентные формулировки понятия независимости случайных величин в терминах функций распределения, а также для случаев дискретных и непрерывных распределений.

Предложение 4 . *Пусть мы имеем случайный вектор $\xi = (\xi_1, \ldots, \xi_n)$. Его компоненты $\xi_1, \ldots, \xi_n$ независимы тогда и только тогда, когда*

$$F_\xi(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n) \ . \tag{8.4}$$

В случае дискретных распределений условие независимости эквивалентно условию

$$P(\xi_1 = x_{i_1}^{(1)}, \ldots, \xi_n = x_{i_n}^{(n)}) = P(\xi_1 = x_{i_1}^{(1)}) \cdot \ldots \cdot P(\xi_n = x_{i_n}^{(n)}) \ , \quad (8.5)$$

а в случае непрерывных - условию

$$\rho_\xi(x_1, \ldots, x_n) = \rho_{\xi_1}(x_1) \cdot \ldots \cdot \rho_{\xi_n}(x_n) \ . \quad (8.6)$$

Доказательство. Рассмотрим множества

$$B_i = \{x \in R^1 : \ x < x_i\}, \ i = \overline{1, n} \ .$$

Для них из (3) следует (4). Обратно, из (4) легко получить (3) для параллелепипедов, а затем аппроксимировать произвольные B_i с помощью сумм отрезков. Свойство (6) получается из (4) дифференцированием. Свойство (5) следует непосредственно из определения независимости.

Пример 1. Пусть мы имеем схему Бернулли с параметрами n и p. Пусть $\varepsilon_i = 1$, если в i-м испытании был "успех", и равно 0 в противном случае. Тогда случайные величины $\varepsilon_1, \ldots, \varepsilon_n$-независимы. Кстати, число успехов S_n в этих n испытаниях представимо в виде $S_n = \varepsilon_1 + \ldots + \varepsilon_n$.

Пример 2. Пусть $\xi = (\xi_1, \xi_2)$ имеет двумерное нормальное распределение. ξ_1 ξ_2 будут независимы тогда и только тогда, когда $\rho = 0$ (задача!).

Нетрудно доказать следующий полезный результат (задача!).

Предложение 5 . *Пусть случайные величины $\xi_1, \ldots, \xi_n, \xi_{n+1}, \ldots, \xi_{n+m}$ - независимы, а $y = f(x_1, \ldots, x_n)$ и $y = g(x_1, \ldots, x_m)$ - борелевские функции. Тогда случайные величины $\eta_1 = f(\xi_1, \ldots, \xi_n)$ и $\eta_2 = g(\xi_{n+1}, \ldots, \xi_{n+m})$ также являются независимыми.*

Пример 3. Пусть мы имеем схему Бернулли с $n_1 + n_2$ испытаниями. Тогда число успехов S_{n_1} в первых n_1 испытаниях и число успехов S_{n_2} в последующих n_2 испытаниях - независимые случайные величины.

1.8.4 Функциональные преобразования случайных векторов

Как и в одномерном случае, важной с практической точки зрения является задача о вычислении распределения функционального преобразования случайного вектора.

Определение 7 . *Отображение $g : R^n \to R^m$ называется* **борелевским**, *если $\forall B \in \mathcal{B}_m$ мы имеем $g^{-1}(B) \in \mathcal{B}_n$.*

Если g - борелевское отображение, $\xi = (\xi_1, \dots, \xi_n)$ - случайный вектор, то $\eta = (\eta_1, \dots, \eta_m) = g(\xi)$ вновь является случайным вектором. Действительно, если $B \in \mathcal{B}_m$, то $g^{-1}(B) \in \mathcal{B}_n$, а $\eta^{-1}(B) = \xi^{-1} \circ g^{-1}(B) \in \mathcal{A}$. Отсюда нетрудно получить выражение для распределения вектора η, если мы знаем распределение вектора ξ : $\forall B \in \mathcal{B}_m$

$$P_\eta(B) = P(\eta \in B) = P(g(\xi) \in B) = \\ P(\xi \in g^{-1}(B)) = P_\xi(g^{-1}(B)) \ . \tag{8.7}$$

Рассмотрим теперь отдельно случаи дискретного и непрерывного распределений.

Если ξ имеет дискретное распределение с множеством значений $X = \{x_1, x_2, \dots\}$ и вероятностями $p_n = P(\xi = x_n)$ появления этим значений, то ясно, что случайный вектор $\eta = g(\xi)$ также имеет дискретное распределение с множеством значений $Y = \{y_1, y_2, \dots, y_m, \dots\}$, где каждое $y_m = g(x_n)$ для некоторого x_n, а вероятности $q_m = P(\eta = y_m)$ появления значения можно вычислить по формуле

$$q_m = \sum_{n:g(x_n)=y_m} p_n \ . \tag{8.8}$$

Пример. Пусть $\xi = (\xi_1, \xi_2)$ - двумерный случайный вектор с дискретным распределением, $p(x, y) = P(\xi_1 = x, \xi_2 = y)$ - вероятность появления некоторого значения (x, y). Рассмотрим функцию $z = g(x, y) = x + y$. Тогда $\eta = \xi_1 + \xi_2$ есть дискретная случайная величина и вероятность того, что $\eta = z$, где z - одно из возможных значений суммы $\xi_1 + \xi_2$, можно рассчитать по формуле

$$P(\eta = z) = \sum_{(x,y):x+y=z} p(x,y) = \sum_x p(x, z-x) \ . \tag{8.9}$$

Если ξ_1 и ξ_2 независимы, то

$$p(x,y) = P(\xi_1 = x, \xi_2 = y) = P(\xi_1 = x) \cdot P(\xi_2 = y) = P_1(x)P_2(y) \ .$$

В этом случае мы получаем

$$P(\xi_1 + \xi_2 = z) = \sum_x P_1(x)P_2(z-x) \ . \tag{8.10}$$

Это **формула свертки** для дискретных распределений.

Пусть теперь $\xi = (\xi_1, \ldots, \xi_n)$ - случайный вектор, который имеет плотность распределения $\rho_\xi(x)$, $x \in R^n$. Как и в одномерном случае, распределение случайного вектора $\eta = g(\xi)$ может не иметь плотности и даже быть дискретным. Необходимы некоторые дополнительные ограничения на функцию $y = g(x)$. Рассмотрим один частный, но практически важный случай.

Предложение 6 . *Пусть $\xi = (\xi_1, \ldots, \xi_n)$ - случайный вектор, который имеет плотность распределения $\rho_\xi(x)$, $g : R^n \to R^n$ - взаимно однозначное и непрерывно дифференцируемое отображение. Тогда распределение случайного вектора $\eta = g(\xi)$ является абсолютно непрерывным и его плотность $\rho_\eta(y)$ можно вычислить по формуле*

$$\rho_\eta(y) = \rho_\xi(g^{-1}(y))|J(g^{-1}(y))|^{-1} \ , \tag{8.11}$$

где $J(x)$ - якобиан отображения $y = g(x)$.

Доказательство этого предложения дословно повторяет доказательство в одномерном случае, но теперь мы должны сделать замену переменных в R^n.

Пример. Пусть $y = Ax$, где A - невырожденная квадратная матрица размера n, т.е. мы имеем линейное отображение в R^n. В этом случае $J(x) = detA$ и

$$\rho_\eta(y) = \rho_\xi(A^{-1}y) \cdot |detA|^{-1} \ .$$

В тех случаях, когда $n \neq m$, последнее предложение не применимо. Но часто можно дополнить отображение $g : R^n \to R^m$ еще одним отображением $f : R^n \to R^{n-m}$ так, чтобы отображение $(g, f) : R^n \to R^n$ уже обладало нужными свойствами.

Пример. Пусть случайный вектор $\xi = (\xi_1, \xi_2)$ имеет плотность распределения $\rho_\xi(x_1, x_2)$. Найдем плотность распределения случайной величины $\eta_1 = \xi_1 + \xi_2$. Здесь $n = 2, m = 1$. Рассмотрим еще одну случайную величину $\eta_2 = \xi_2$. Тогда в целом мы имеем следующее линейное отображение $y_1 = x_1 + x_2,\ y_2 = x_2$. Матрица A этого отображения имеет вид

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} ,$$

$det A = 1$, а обратная матрица A^{-1} равна

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} .$$

В предыдущем примере мы получили, что

$$\rho_\eta(y_1, y_2) = \rho_\xi(A^{-1}y) \cdot |det A|^{-1} .$$

Чтобы найти плотность распределения для η_1, достаточно проинтегрировать по координате y_2, т.е.

$$\rho_{\eta_1}(y_1) = \int_{-\infty}^{\infty} \rho_\xi(y_1 - y_2, y_2) dy_2 . \tag{8.12}$$

Если ξ_1 и ξ_2 - независимы, то $\rho_\xi(x_1, x_2) = \rho_{\xi_1}(x_1)\rho_{\xi_2}(x_2)$. Заменяя y_1 на y, получаем

$$\rho_{\xi_1+\xi_2}(y) = \int_{-\infty}^{\infty} \rho_{\xi_1}(y - x)\rho_{\xi_2}(x) dx . \tag{8.13}$$

Это **формула свертки** для непрерывных распределений.

В более сложных ситуациях, когда не удается свести задачу к предложению 6, необходимо провести прямые расчеты, вычисляя распределение P_η (например, функцию распределения), а затем находя плотность. Технически это сводится к нахождению множества $g^{-1}(B)$ и вычислению интеграла от $\rho_\xi(x)$ по этому множеству. Чтобы продемонстрировать, как работает этот метод, рассмотрим тот же самый пример: $\eta = \xi_1 + \xi_2$. Вычислим для η функцию распределения:

$$F_\eta(y) = P(\eta < y) = P(\xi_1 + \xi_2 < y) \ .$$

Фактически нам нужно найти вероятность попадания случайного вектора $\xi = (\xi_1, \xi_2)$ в множество

$$B(y) = \{(x_1, x_2) : \ x_1 + x_2 < y\} \subset R^2 \ .$$

Тогда мы имеем

$$F_\eta(y) = P((\xi_1, \xi_2) \in B(y)) = \int\int_{B(y)} \rho_\xi(x_1, x_2) dx_1 dx_2 \ .$$

Далее, расписывая двойной интеграл в виде повторного, получаем

$$F_\eta(y) \ = \ \int_{-\infty}^{\infty} \int_{-\infty}^{y-x_2} \rho_\xi(x_1, x_2) dx_1 dx_2 = \left| \begin{array}{l} x = x_1 + x_2 \\ z = x_2 \end{array} \right| =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{y} \rho_\xi(x - z, z) dx dz == \int_{-\infty}^{y} [\int_{-\infty}^{\infty} \rho_\xi(x - z, z) dz] dx \ .$$

Дифференцируя по y, окончательно получаем

$$\rho_\eta(y) = \frac{d}{dy} F_\eta(y) = \int_{-\infty}^{\infty} \rho_\xi(y - z, z) dz \ .$$

Таким образом, мы имеем тот же результат, что и ранее.

1.9 Задачи для самостоятельного решения

В дополнение к тем задачам, что были приведены в основном тексте, мы предлагаем несколько задач стандартного типа, которые помогут студентам подготовиться к экзаменам.

1. Для социологического обследования из группы в 100 человек, среди которых 40 мужчин и 60 женщин, случайно отбирают 10 человек. Какова вероятность того, что среди них будет четверо мужчин?

2. Из отрезка $[0,1]$ случайно и независимо выбирают две точки. Какова вероятность того, что расстояние между ними будет больше $1/2$?

3. Два человека играют в следующую игру: независимо друг от друга они выбирают одну из монет в 10 или 20 рублей и отмечают полученную сумму. Если эта сумма равна 20, то выигрывает первый, если она равна 40, то выигрывает второй, в противном случае фиксируется ничья. Какова вероятность того, что первый игрок не проиграет, если они выбирают монеты с одинаковыми вероятностями?

4. Есть две монеты: одна обыкновенная, а другая - с двумя гербами. Случайно выбрали одну из монет и подбросили два раза. Выпало два герба. Какова вероятность того, что это обыкновенная монета?

5. Две трети секретарей большого стенографического бюро имеют водительские права. Для участия в некоторой поездке случайно отобраны 4 секретаря. Какова вероятность того, что по крайней мере двое из них имеют водительские права?

6. Вероятность выигрыша в некоторой лотерее равна 0.3. Вы купили 100 билетов. Не менее какого числа выигрышей мы можем гарантировать с вероятностью не менее 0.9?

7. По некоторому каналу связи передаются сообщения. Вероят-

ность ошибки при передаче одного символа равна 0.01. При каком количестве символов вероятность того, что мы имеем в сообщении не менее двух ошибок, не менее 0.95?

8. Симметричную монету подбрасывают 4 раза. ξ - число выпавших гербов. Найти распределение вероятностей с.в. ξ. Вычислить ее математическое ожидание, дисперсию и $P(\xi \geq 2)$.

9. Для случайной величины ξ из задачи 8 определим $\eta = (\xi - 2)^2$. Найти:

1) распределение η,
2) $M\eta,\ D(\eta)$.

10. Есть три монеты. Вероятность появления герба равна 0.3 для первой монеты, 0.5 - для второй и 0.6 - для третьей. Пусть ξ равна числу выпавших гербов. Найти распределение ξ и вычислить $M\xi$, $D(\xi)$.

11. Из отрезка [0,1] случайно и независимо выбирают две точки. Найти плотность распределения разности координат этих точек. Найти математическое ожидание и дисперсию этой случайной величины.

12. Для случайной величины ξ из задачи 11 определим $\eta = \xi^2$.

1) распределение η,
2) $M\eta,\ D(\eta)$.

13. Распределение дискретного случайного вектора $\xi = (\xi_1, \xi_2)$ задано таблицей

$\xi_2 \setminus \xi_1$	-1	0	1
0	0.05	0.1	0.15
1	0.2	0.05	0.05
2	0.1	0.2	0.1

Найти:

1) распределение с.в. ξ_1 и ξ_2,
2) проверить независимость ξ_1 и ξ_2,
3) вероятность того, что $\xi_1 \xi_2 = 0$,

4) коэффициент корреляции для ξ_1 и ξ_2,

5) совместное распределение случайного вектора $\eta = (\eta_1, \eta_2)$, где $\eta_1 = \xi_1,\ \eta_2 = \xi_1\xi_2$.

14. ξ_1 и ξ_2 независимы и имеют показательное распределение с параметром λ. Найти плотность распределения случайной величины $\eta = \xi_1 - \xi_2$.

15. Случайный вектор $\xi = (\xi_1, \xi_2)$ имеет распределение с плотностью

$$\rho_\xi(x_1, x_2) = \begin{cases} C(x_1 + 2x_2) & , \quad x_1, x_2 \in [0, 1], \\ \qquad 0 & , \quad \text{в пр. сл..} \end{cases}$$

Найти: 1) константу C,

2) $\rho_{\xi_1}(x_1)$,

3) проверить независимость ξ_1 и ξ_2,

4) $P(\xi_1 > \xi_2)$.

5) $M(\xi_1), M(\xi_2), cov(\xi_1, \xi_2), \rho(\xi_1, \xi_2)$.

16. ξ_1 и ξ_2 независимы и одинаково распределены,

$$\rho_{\xi_1}(x) = \rho_{\xi_2}(x) = \begin{cases} 2x & , \quad x \in [0, 1] \ , \\ \ 0 & , \quad \text{впр.сл..} \end{cases}$$

$\eta = \xi_1 + 2\xi_2$. Найти $\rho_\eta(y)$.

Список литературы

а) Основной

1. Севастьянов Б.А. Курс теории вероятностей и математической статистики. - М.: Наука, 1982.

2. Чистяков В.П. Курс теории вероятностей. - М.: Наука, 1982.

3. Ширяев А.Н. Вероятность. - М.: Наука, 1980.

4. Агапов Г.И. Задачник по теории вероятностей: Учебное пособие для студентов втузов. - М.: Высш. шк., 1986.

5. Зубков А.М., Севастьянов Б.А., Чистяков В.П. Сборник задач по теории вероятностей. - М.: Наука, 1989.

b) Дополнительный

6. Колмогоров А.Н. Основные понятия теории вероятностей. - М.: Наука, 1974.

7. Феллер В. Введение в теорию вероятностей. Т. 1,2. - М.: Мир, 1984.

8. Хургин Я.И. Да, нет или может быть. - М.: Наука, 1977.

9. Хургин Я.И. Как объять необъятное. - М.: Знание, 1979.