МИНОБРНАУКИ РОССИИ

Санкт-Петербургский государственный электротехнический
университет “ЛЭТИ” им. В. И. Ульянова (Ленина)

ТЕОРИЯ ВЕРОЯТНОСТЕЙ
В ПРИМЕРАХ И ЗАДАЧАХ

Санкт-Петербург
2017

МИНОБРНАУКИ РОССИИ

Санкт-Петербургский государственный электротехнический университет “ЛЭТИ”

ТЕОРИЯ ВЕРОЯТНОСТЕЙ В ПРИМЕРАХ И ЗАДАЧАХ

Учебное пособие

Санкт-Петербург
Издательство СПбГЭТУ “ЛЭТИ”
2017

УДК 519.21(075)
ББК В171я7
Т33

Авторы: Гоголева Н. Г., Жукова Е. Е., Колбина С. А., Непомнящая Т. В., Фролова Е. В., Шевченко Е. А.

Т33 Теория вероятностей в примерах и задачах: учеб. пособие. СПб.: Изд-во СПбГЭТУ „ЛЭТИ“, 2017. 80 с.

ISBN ?

Рецензенты: кафедра высшей математики ВШТ СПбГУПТД (Б. Ф. Иванов); д-р физ.-мат. наук проф. Л. М. Баскин (СПбГУТ им. проф. М.А.Бонч-Бруевича).

Предназначено для студентов технических факультетов, обучающихся по всем направлениям и специальностям дневной, вечерней и заочной форм обучения.

Утверждено
редакционно-издательским советом университета
в качестве учебного пособия

ISBN ?

1. ДИСКРЕТНОЕ ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО

1.1. Построение модели случайного эксперимента. Вероятностное пространство

Объект изучения теории вероятностей – случайное явление, т. е. любое явление, результат которого невозможно полностью предсказать из-за каких-либо факторов, не поддающихся учету. Теория вероятностей изучает закономерности случайных явлений. Основным понятием теории вероятностей является понятие вероятностного пространства. Универсальная модель вероятностного пространства, основанная на теоретико-множественном аксиоматическом подходе разработана академиком А. Н. Колмогоровым в 20-х г.г. XX в. В настоящее время эта модель общепринята во всем мире под названием „аксиоматика Колмогорова". В данной главе мы ограничимся изучением частного случая этой модели — дискретного вероятностного пространства.

Опыт, в котором наблюдается случайное явление, называется случайным экспериментом. Результат случайного эксперимента называется его исходом. Два исхода различны если они не могут произойти одновременно в одном испытании. Все возможные различные исходы эксперимента называются „элементарными исходами" или „элементарными событиями". Множество Ω, элементами которого являются все возможные исходы эксперимента, называется „множеством всех элементарных исходов". Если число элементов множества Ω конечно либо счетно, модель случайного эксперимента называется **дискретной**. В этом случае распределение вероятностей задается следующим образом.

Определение 1.1.1. *Вероятностным распределением на Ω называется функция $P : \Omega \to [0;1]$, сопоставляющая каждому элементарному событию $\omega \in \Omega$ число $P(\omega) \in [0,1]$ называемое вероятностью элементарного события ω. Причем выполняется равентсво (аксиома нормировки)*

$$\sum_{\omega \in \Omega} P(\omega) = 1.$$

Определение 1.1.2. *Пара (Ω, P) называется дискретным вероятностным пространством.*

В дискретном вероятностном пространстве любое подмножество $A \subset \Omega$ называется событием. Если $\omega_i \in A$, то говорят, что элементарный исход ω_i благоприятен для события A. Вероятность случайного события A равна сумме вероятностей всех исходов благоприятных для события A:

$$P(A) = \sum_{i:\omega_i \in A} p_i.$$

Заметим, что задание вероятностного пространства является совершенно произвольным и зависит исключительно от желаемого эксперимента (реального или мысленного). Подтверждение актуальности заданного вероятностного пространства можно получить только в ходе реального эксперимента, что является вопросом, относящимся к математической статистике. С точки зрения теории вероятностей вероятностное пространство можно задать сколь угодно фантастическим образом.

Приведем примеры вероятностных пространств, событий в этих пространствах и их вероятностей.

Пример 1.1.1. Рассмотрим эксперимент по подкидыванию монетки. Зададим $\Omega = \{$упадет гербом, упадет решкой, встанет на ребро, закатится под диван, вылетит за окно$\}$. Первым двум событиям припишем вероятности равные $2/5$, последним трем – $1/15$. Проверьте корректность задания вероятностного пространства. Одним из возможных событий является событие A – „экспериментатор утратил возможность наблюдения за монеткой“. Найдите вероятность этого события.

Решение. Сумма $2 \cdot 2/5 + 3 \cdot 1/15 = 1$, следовательно, вероятностное пространство задано корректно.

Событие A состоит из последних двух элементарных событий, и его вероятность $P(A) = 1/15 + 1/15 = 2/15$.

Даже если ограничиться всего двумя исходами подбрасывания монеты, можно задать различные вероятностные пространства.

Пример 1.1.2. Эксперимент состоит в подбрасывании монеты. Тогда $\Omega = \{\text{упадет ребром}, \text{упадет решкой}\} = \{\omega_1, \omega_2\}$. Зададим на Ω вероятностное распределение. Ясно, что это можно сделать многими способами:

1. $P(\omega_1) = P(\omega_2) = 1/2$;
2. $P(\omega_1) = 1/3$, $P(\omega_2) = 2/3$;

3. $P(\omega_1) = p$, $P(\omega_2) = 1 - p$, $0 \le p \le 1$.

Во всех случаях аксиомы вероятностного распределения выполнены. Следовательно, различные варианты задания вероятностного распределения на Ω порождают различные модели одного и того же эксперимента. Какую из этих моделей считать адекватной эксперименту?

Если монета „правильная“, то, как подсказывает жизненный опыт, обе ее стороны выпадают одинаково часто. Следовательно, частота выпадения герба или цифры будет группироваться около числа 1/2. Английский статистик К. Пирсон, например, подбрасывал „правильную“ монету 24 000 раз, и герб выпал 12 012 раз. Логические умозаключения позволяют сделать вывод о равновозможности выпадения как герба, так и цифры. Значит, естественно считать вероятность выпадения герба „правильной“ монеты равной 1/2. Поэтому первая модель — это модель для „правильной“ монеты, и она будет более точно, чем другие модели, описывать эксперимент с „правильной“ монетой. Тогда вторая модель будет описывать эксперимент с „неправильной“ монетой, у которой цифра выпадает в 2 раза чаще, чем герб.

С помощью третьей модели можно описать эксперимент с любой монетой. Для этого достаточно в качестве числа p выбрать частоту выпадения герба. Вопрос, как это сделать, относится к математической статистике, к которой обратимся в конце данного пособия. А сейчас приведем еще несколько примеров задания вероятностного пространства.

Пример 1.1.3. Шулер сделал кубик, на гранях которого написаны числа от 1 до 6. Вероятность выпадения шестерки равна 1/3, пятерки – 1/4, четверки – 1/6, тройки и двойки – по 1/10, наконец, единица выпадает с вероятностью 1/20. Таким образом, Ω состоит из шести элементарных событий ω_i, заключающихся в том, что на кубике выпало число i. Соответствующие элементарные вероятности $p_1 = 1/20$, $p_2 = p_3 = 1/10$, $p_4 = 1/6$, $p_5 = 1/4$, $p_6 = 1/3$. Проверьте корректность задания вероятностного пространства. Найдите вероятность события A, состоящего в том, что на кубике выпало четное число, и события B, состоящего в том, что на кубике выпало число не меньше пятерки.

Решение. Заметим, что сумма $\sum\limits_{i=1}^{6} p_i = 1$, поэтому задание вероятностного пространства корректно.

Событие A имеет вероятность $P(A) = p_2+p_4+p_6 = 1/10+1/6+1/3 = 3/5$. Событие B имеет вероятность $P(B) = p_5 + p_6 = 1/4 + 1/3 = 7/12$.

Еще раз подчеркнем, что элементарные вероятности задаются условием задачи и могут быть заданы совершенно произвольно.

Пример 1.1.4. Шулер сделал кубик, на гранях которого написаны числа от 1 до 6. На этом кубике вероятность выпадения шестерки равна 1/4, пятерки и четверки – по 1/6, тройки – 2/9, двойки – 1/9, наконец, единица выпадает с вероятностью 1/12. Таким образом, Ω состоит из шести элементарных событий ω_i, заключающихся в том, что на кубике выпало число i. Соответствующие элементарные вероятности $p_1 = 1/12$, $p_2 = 1/9$, $p_3 = 2/9$, $p_4 = p_5 = 1/6$, $p_6 = 1/4$. Проверьте корректность задания вероятностного пространства. Найдите вероятность события A, состоящего в том, что на кубике выпало четное число, и события B, состоящего в том, что на кубике выпало число не меньше пятерки.

Решение. Заметим, что сумма $\sum\limits_{i=1}^{n} p_i = 1$, поэтому задание вероятностного пространства столь же корректно как и в предыдущем примере.

Событие A имеет вероятность $P(A) = p_2+p_4+p_6 = 1/9+1/6+1/4 = 19/36$. Событие B имеет вероятность $P(B) = p_5 + p_6 = 1/6 + 1/4 = 5/12$.

Задачи для самостоятельного решения

1. В пражском отеле на ресепшен дежурит один из семи сотрудников. Помимо чешского и английского языков, которые знают все сотрудники, трое из них знают немецкий язык, двое – русский, один – французский и один – итальянский. В отель приехал русский турист, немного говорящий по-французски и не знающий других иностранных языков. Постройте вероятностное пространство и определите вероятность того, что сотрудник отеля и турист смогут понять друг друга.

2. 550 абитуриентов случайным образом распределяют по трем аудиториям, две из которых вмещают 200 человек, а третья – 150. Постройте вероятностное пространство и найдите вероятность, что Василий окажется в аудитории вместимостью 150 человек.

Ответы к задачам для самостоятельного решения

1. 3/7. 2. 3/11.

1.2. Теорема сложения

В теории вероятностей используется следующая терминология:

Ω – достоверное событие, из определения следует $P(\Omega) = 1$;

$\emptyset$ – невозможное событие, из определения следует $P(\emptyset) = 0$;

$A \cup B$ – сумма событий, состоит из всех элементарных исходов, благоприятных для A или для B;

$A \cap B$ – произведение событий, состоит из всех элементарных исходов, благоприятных и для A и для B;

$A \setminus B$ – разность событий, состоит из всех элементарных исходов, благоприятных для A, но неблагоприятных для B;

$\overline{A}$ – событие противоположное A, состоит из всех элементарных исходов, неблагоприятных для A.

Одна из основных теорем, относящихся к основам теории вероятностей, носит название теоремы сложения. Для классической формулировки потребуется следующее определение:

Определение 1.2.1. *Если $A \cap B = \emptyset$, то события A и B называются несовместными.*

Теорема 1.2.1. (Теорема сложения). *Если события A и B несовместны, то*

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

Следствие 1. *Если события A и B несовместны, то*

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

Поскольку события A и $\overline{A}$ несовместны и $A \cup \overline{A} = \Omega$, из теоремы сложения получим

Следствие 2. $P(\overline{A}) = 1 - P(A).$

Если события не являются несовместными, то используют обобщение теоремы сложения.

Теорема 1.2.2. *Для произвольных событий A и B*

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

Пример 1.2.1. В вероятностном пространстве, моделирующем эксперимент по подбрасыванию кубика из примера 1.1.3., найдите вероятность того, что выпавшее число четно или не меньше пяти.

Решение. Введем события: A – выпавшее число четное, B – выпавшее число не меньше пяти. По Теореме 1.2.1 имеем $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 3/5 + 7/12 - 1/3 = 17/20$. Здесь мы использовали вероятности событий A и B, сосчитанные в 1.1. Пересечение событий A и B состоит из одного элементарного события, заключающегося в том, что выпало 6 и имеющего вероятность p_6.

Пример 1.2.2. Решите аналогичную задачу в вероятностном пространстве, моделирующем эксперимент по подбрасыванию кубика из примера 1.1.4.

Решение. Аналогично предыдущему примеру имеем $P(A \cup B) = = P(A) + P(B) - P(A \cap B) = 19/36 + 5/12 - 1/4 = 25/36$.

Пример 1.2.3. Рассмотрим рулетку с 50 красными и 50 черными полями, пронумерованными числами от 1 до 50. Все красные поля одинаковы и вдвое больше, чем каждое из черных полей. Найдите вероятности следующих событий:

A – выпало четное число;

C – выпало четное число не делящееся на три;

D – выпало четное красное число или нечетное черное;

E – выпало красное число делящееся на 3, но не делящееся на 6;

F – выпало красное число, делящееся на 3 или на 5;

G – выпало число, делящееся на 3 или на 5.

Решение. Прежде всего следует перевести словесную формулировку задачи на математический язык, четко определив вероятностное пространство. Описанной рулетке соответствует множество

$$\Omega = \{r1, r2, \ldots, r50, b1, b2, \ldots, b50\}$$

с элементарными вероятностями
$p_1 = p_2 = \ldots = p_{50} = 2/150$, $p_{51} = p_{51} = \ldots = p_{100} = 1/150$.

Через R, $R2$, $R3$, $R5$, $R6$, $R15$ обозначим события, заключающиеся в том, что выпало произвольное красное число; выпало красное число, делящееся на 2, на 3, на 5, на 6 и на 15 соответственно. Найдем их вероятности, просуммировав соответствующее количество элементарных вероятностей: $P(R) = \frac{100}{150} = \frac{2}{3}$, $P(R2) = \frac{50}{150} = \frac{1}{3}$, $P(R3) = \frac{32}{150} = \frac{16}{75}$, $P(R5) = \frac{20}{150} = \frac{2}{15}$, $P(R6) = \frac{16}{150} = \frac{8}{75}$, $P(R15) = \frac{6}{150} = \frac{1}{25}$.

Аналогичные множества для черных полей имеют вероятности:
$P(B) = \frac{50}{150} = \frac{1}{3}$, $P(B2) = \frac{25}{150} = \frac{1}{6}$, $P(B3) = \frac{16}{150} = \frac{8}{75}$,
$P(B5) = \frac{10}{150} = \frac{1}{15}$, $P(B6) = \frac{8}{150} = \frac{4}{75}$, $P(B15) = \frac{3}{150} = \frac{1}{50}$.

Теперь $A = R2 \cup B2$ и по теореме сложения
$P(A) = P(R2) + P(B2) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$.

$P(\text{выпало число, делящееся на } 6) = P(R6 \cup B6) = \frac{8}{75} + \frac{4}{75} = \frac{12}{75}$.

Далее находим
$P(C) = P(A) - P(R6 \cup B6) = \frac{1}{2} - \frac{12}{75} = \frac{51}{150} = \frac{17}{50}$.

$P(D) = P(R2 \cup (B \setminus B2)) = P(R2) + (P(B) - P(B2)) = \frac{1}{3} + (\frac{1}{3} - \frac{1}{6}) = \frac{1}{2}$.

$P(E) = P(R3 \cap \overline{R6}) = P(R3 \setminus R6) = P(R3) - P(R6) = \frac{16}{75} - \frac{8}{75} = \frac{8}{75}$.

$P(F) = P(R3 \cup R5) = P(R3) + P(R5) - P(R15) = \frac{16}{75} + \frac{2}{15} - \frac{1}{25} = \frac{46}{150}$.

$P(G) = P(E \cup (B3 \cup B5)) = P(E) + P(B3 \cup B5) =$
$= \frac{46}{150} + P(B3) + P(B5) - P(B15) = \frac{46}{150} + \frac{8}{75} + \frac{1}{15} - \frac{1}{50} = \frac{23}{50}$.

Задачи для самостоятельного решения

1. Как изменятся вероятности в примере 1.2.3., если добавить еще зеленое поле “зеро“, размер которого в пять раз больше, чем черное поле.

2. Балет „Лебединое озеро“ в Мариинском театре дают каждый вторник, а в Александринском – один раз в 5 дней. Иностранный турист приехал в Петербург на три дня. Какова вероятность, что он сможет посмотреть балет в одном из этих театров (предполагается, что билеты можно купить без проблем).

3. Куб с окрашенными гранями распилен на 27 кубиков одинакового размера. Полученные кубики тщательно перемешаны. Наудачу извлекается один кубик. Постройте вероятностное пространство и найдите вероятность того, что извлеченный кубик имеет ровно две окрашенные грани.

4. Из урны, содержащей 29 пронумерованных шаров, наудачу выбирается один шар. Какова вероятность того, что его номер делится на 2 или на 3?

5. Начинающий шулер пометил карты в колоде, содержащей 52 карты, так, что он с равными вероятностями вытаскивает туза, короля, даму или валета, и эта вероятность для каждой из этих карт в 10 раз больше, чем для каждой из остальных карт. Его противник вытаскивает одну из карт, оставшихся в колоде после вытаскивания шулером своей карты, с равной вероятностью. Постройте вероятностную модель и найдите вероятность, что карта шулера окажется старше карты противника.

Ответы к задачам для самостоятельного решения

1. $P(A) = \dfrac{15}{31}$; $P(C) = \dfrac{51}{155}$; $P(D) = \dfrac{15}{31}$; $P(E) = \dfrac{46}{155}$: $P(F) = \dfrac{74}{155}$; $P(G) = \dfrac{69}{155}$.

2. $\dfrac{27}{35}$

3. Пронумеруем кубики в некотором порядке и составим вероятностное пространство из элементарных событий, соответствующих извлечению каждого из кубиков. Логично приписать им элементарные вероятности $\dfrac{1}{27}$. Искомая вероятность $\dfrac{12}{27}$.

4. $\dfrac{19}{29}$.

5. Вероятностное пространство состоит из пар карт, вытащенных шулером и его противником. Элеметнарные вероятности равны $\dfrac{10}{9996}$, если первая карта туз, король, дама или валет и $\dfrac{1}{9996}$ в остальных случаях. Искомая вероятность $\dfrac{608}{833}$.

1.3. Классическая вероятностная модель

Рассмотрим простейший случайный эксперимент, в котором пространство всех элементарных исходов — конечное множество. Если условия случайного эксперимента таковы, что естественным является предположение о равновозможности исходов, то мы приходим к так называемой „классической“ вероятностной модели:

$$\forall \omega \in \Omega, \quad p(\omega) = \frac{1}{|\Omega|}, \qquad\qquad \forall A \subset \Omega, \quad p(A) = \frac{|A|}{|\Omega|}.$$

Здесь через $|A|$ обозначено число элементов множества A.

Полученная формула означает, что в классической вероятностной модели вероятность случайного события A равна отношению количества исходов, благоприяных для события A, к количеству всех возможных элементарных исходов эксперимента.

Гипотеза о равновозможности исходов не всегда имеет место. Применять классическую вероятностную модель допустимо только в том случае, когда в формулировке задачи содержатся сведения о том, что гипотеза о равновозможности исходов соответствует данному случайному эксперименту. Например, „...бросается правильная игральная кость...“, „... из 10 неотличимых на ощупь шаров случайным образом выбираются два...“.

При решении задач в рамках классической вероятностной модели важно умение подсчитывать число элементов конечного множества, что является одной из основных задач комбинаторики.

Приведем основные правила и формулы комбинаторики.

Утверждение 1.3.1. (Правило умножения). *Пусть имеется k конечных множеств A_i, $i = 1, ...k$. Число элементов множества A_i обозначим через n_i. Общее число N всевозможных упорядоченных наборов $(a_1, a_2, ..., a_k)$, $\quad a_i \in A_i$, $\quad$ равно: $N = n_1 n_2 ... n_k$.*

Утверждение 1.3.2. (Правило сложения). *Пусть имеется два конечных множества A_i, $i = 1, 2$. Число элементов множества A_i обозначим через n_i. Предположим, что множества A_1 и A_2 не имеют общих элементов. Выбор одного элемента или из A_1 или из A_2 можно осуществить $n_1 + n_2$ способами.*

Размещения без повторений

Пусть имеется множество, состоящее из n элементов. Будем составлять из элементов этого множества упорядоченные k-элементные наборы. Понятно, что первый элемент можно выбрать n способами, второй выбирается из множества, в котором на один элемент меньше, и т. д.

Количество всевозможных упорядоченных k-элементных наборов из n элементного множества называется **число размещений** и обозначается A_n^k. По правилу умножения

$$A_n^k = n(n-1)(n-2)...(n-k+1) = \frac{n!}{(n-k)!}.$$

$A_n^n = n!$ называется числом **перестановок**.

Размещения с повторениями

Пусть имеется множество, состоящее из n элементов. Будем составлять из элементов этого множества упорядоченные k-элементные наборы, каждый раз возвращая отобранный элемент. Тогда на каждое место в упорядоченном наборе выбирается один из n элементов. По правилу умножения количество размещений с повторениями равно n^k.

Сочетания

Количество всевозможных неупорядоченных k-элементных наборов из n-элементного множества называется **число сочетаний** и обозначается C_n^k.

$$C_n^k = A_n^k \frac{1}{k!} = \frac{n!}{k!(n-k)!}.$$

Из этой формулы следуют очевидные соотношения:

$$C_n^k = C_n^{n-k}, \qquad C_n^1 = C_n^{n-1} = n.$$

Пример 1.3.1. В ящике лежат 15 одинаковых шаров (неотличимых на ощупь). Четыре шара выкрашены в красный цвет, шесть в зеленый и пять в синий. Случайным образом из ящика достают 4 шара. Найдите вероятности случайных событий:

A – достали 2 зеленых и 2 синих;

B – среди этих четырех шаров есть хотя бы один красный;

D – среди этих четырех шаров есть шары всех цветов.

Решение. Поскольку по условию шары одинаковые, можно принять гипотезу, что все исходы данного эксперимента равновероятны, и воспользоваться формулой для вычисления вероятности в классической вероятностной модели. Наборы шаров не упорядочены, поэтому количество всех возможных исходов эксперимента посчитаем по формуле для числа соче-

таний:
$$|\Omega| = C_{15}^4 = \frac{15!}{4!11!}.$$

Количество исходов, благоприятных для события A, найдем по правилу умножения:
$$|A| = C_6^2 C_5^2 = \frac{6!}{2!4!}\frac{5!}{2!3!}.$$
$$p(A) = \frac{|A|}{|\Omega|} - \frac{5!6!4!11!}{2!3!2!4!15!} = \frac{5 \cdot 6!}{12 \cdot 13 \cdot 14 \cdot 15} = \frac{10}{91}.$$

Для вычисления вероятности случайного события B удобно перейти к противоположному событию $\overline{B}$ — среди этих четырех шаров красных нет.
$$|\overline{B}| = C_{11}^4 = \frac{11!}{4!7!}, \qquad p(B) = 1 - p(\overline{B}) = 1 - \frac{11!11!4!}{7!4!15!} = 1 - \frac{22}{91} = \frac{69}{91}.$$

Случайное событие D является суммой трех попарно несовместных событий:
$$D = D_1 + D_2 + D_3,$$

D_1 – достали один зеленый, один синий и два красных шара;

D_2 – достали один зеленый, один красный и два синих шара;

D_3 – достали один красный, один синий и два зеленых шара.

По правилу умножения
$$|D_1| = C_6^1 C_5^1 C_4^2, \quad |D_2| = C_6^1 C_4^1 C_5^2, \quad |D_3| = C_4^1 C_5^1 C_6^2.$$

Согласно теореме сложения вероятность суммы несовместных событий равна сумме их вероятностей:
$$p(D) = p(D_1) + p(D_2) + p(D_3) = \frac{6 \cdot 5 \cdot C_4^2 + 6 \cdot 4 \cdot C_5^2 + 5 \cdot 4 \cdot C_6^2}{C_{15}^4} = \frac{48}{91}.$$

Пример 1.3.2. Найти вероятность того, что в семизначном телефонном номере все цифры различны (телефонный номер может начинаться с нуля).

Решение. За элементарный исход эксперимента естественно принять упорядоченный набор из семи цифр:

$$\omega = (a_1, a_2, a_3, a_4, a_5, a_6, a_7), \qquad a_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}.$$

По формуле для вычисления количества размещений с повторениями, имеем $|\Omega| = 10^7$.

Количество телефонных номеров с различными цифрами равно числу размещений (без повторений) из 10 по 7. Считая, что любой из телефонных номеров появляется с одинаковой вероятностью, воспользуемся формулой для вычисления вероятности случайного события в классической вероятностной модели. Найдем вероятность случайного события А – в семизначном номере телефона все цифры различны:

$$p(A) = \frac{|A|}{|\Omega|} = \frac{A_{10}^7}{10^7} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{10^7}.$$

Пример 1.3.3. Шесть шаров случайным образом раскладывают по трем ящикам. Найти вероятности случайных событий:

А – во всех ящиках разное число шаров;

В – во всех ящиках одинаковое число шаров.

Решение. За элементарный исход эксперимента примем упорядоченный набор из шести цифр

$$\omega = (a_1, a_2, a_3, a_4, a_5, a_6), \qquad a_i \in \{1, 2, 3\},$$

здесь a_i – номер ящика, в который попал i-ый шар. По формуле для вычисления количества размещений с повторениями, имеем $|\Omega| = 3^6$.

Если во всех ящиках разное число шаров, то в одном из ящиков три шара, в двух других – два и один. Разобьем шары на три группы: из одного шара, двух шаров, трех шаров. Учтем, что эти группы можно по-разному разместить в трех ящиках. Таким образом найдем количество элементарных исходов, благоприятных для события А:

$$|A| = C_6^3 \cdot C_3^2 \cdot 3! = \frac{6!}{3!3!} \cdot 3 \cdot 3! = \frac{6!}{2}.$$

Поскольку шары разбрасываются случайным образом, естественно принять гипотезу, что любой шар попадает в каждый из ящиков с одинаковой вероятностью, поэтому воспользуемся формулой для вычисления вероятности случайного события в классической вероятностной модели:

$$p(A) = \frac{|A|}{|\Omega|} = \frac{6!}{2 \cdot 3^6} = \frac{40}{81}.$$

Посчитаем количество элементарных исходов, благоприятных для события B. В этом случае в каждом ящике оказывается два шара, следовательно, в упорядоченном наборе две единицы, две двойки и две тройки. Два места для единиц можно выбрать C_6^2 способами, из оставшихся четырех два места для двоек можно выбрать C_4^2 способами. Таким образом:

$$|B| = C_6^2 \cdot C_4^2, \qquad p(B) = \frac{|B|}{|\Omega|} = \frac{C_6^2 \cdot C_4^2}{3^6} = \frac{10}{81}.$$

Пример 1.3.4. Восемь человек случайным образом выстраиваются в ряд один за другим. Какова вероятность того, что два конкретных человека не окажутся рядом?

Решение. Количество размещений восьми человек в ряд равно числу перестановок: $|\Omega| = 8!$.

Введем случайное событие A – два выбранных человека рядом не стоят. Удобно перейти к противоположному событию $\overline{A}$ – два выбранных человека стоят рядом. Выбранная пара может располагаться на семи местах (первое и второе, второе и третье, ... , седьмое и восьмое). Вариантов размещения оставшихся шести человек на шесть местах – 6!. Поскольку наша модель учитывает порядок, надо также учесть, что выбранные два человека могут меняться местами, поэтому

$$|\overline{A}| = 6! \cdot 7 \cdot 2, \qquad p(A) = 1 - \frac{2 \cdot 7!}{8!} = \frac{3}{4}.$$

Пример 1.3.5. Группа из 15 студентов пишет контрольную работу из трех вариантов (каждый вариант – 5 человек). Случайным образом из группы выбирают пять студентов. Найти вероятность случайного события A – среди них есть писавшие все три варианта.

Решение. Случайный эксперимент состоит в выборе 5 студентов из 15. Поскольку порядок в данном эксперименте не важен, общее число исходов эксперимента найдем по формуле для числа сочетаний

$$|\Omega| = C_{15}^5 = \frac{15!}{10! \cdot 5!}.$$

Представим A в виде суммы попарно несовместных событий:

$$A = A_{311} + A_{131} + A_{113} + A_{122} + A_{212} + A_{221},$$

где событие A_{ijk} – первый вариант писали i студентов, второй – j студентов, третий – k. Поскольку в задаче число вариантов одинаковое,

$$p(A_{311}) = p(A_{131}) = p(A_{113}) = \frac{C_5^3 C_5^1 C_5^1}{C_{15}^5} = \frac{10 \cdot 5 \cdot 5}{7 \cdot 13 \cdot 3 \cdot 11},$$

$$p(A_{122}) = p(A_{212}) = p(A_{221}) = \frac{C_5^2 C_5^2 C_5^1}{C_{15}^5} = \frac{10 \cdot 10 \cdot 5}{7 \cdot 13 \cdot 3 \cdot 11},$$

по теореме сложения

$$p(A) = 3p(A_{113}) + 3p(A_{122}) = \frac{750}{1001} \approx 0,74925.$$

Задачи для самостоятельного решения

1. Студент знает 20 вопросов из 30. Какова вероятность того, что на экзамене он верно ответит на два случайным образом выбранных вопроса?

2. Найти вероятность того, что в семизначном номере телефона 4 цифры совпадают, а остальные различны (считаем что телефонный номер может начинаться с нуля).

3. На станции семь человек случайным образом выбирают один из шести вагонов поезда. Найти вероятность того, что хотя бы в один вагон никто не сядет.

4. В коробке лежат три конфеты в красных фантиках, пять — в желтых и семь — в зеленых. Случайным образом достают четыре конфеты. Какова вероятность того что достали конфеты в фантиках всех цветов?

5. Найти вероятность того, что дни рождения 10 человек приходятся на разные месяцы года.

6. Два грузовых и семь легковых автомобилей случайным образом занимают места на стоянке (места располагаются в один ряд). Найти вероятность того, что между грузовыми автомобилями будут стоять два легковых.

7. Бросаются две правильные игральные кости. Какова вероятность того, что сумма выпавших очков больше семи?

Ответы к задачам для самостоятельного решения

1. $\dfrac{38}{87}$
2. $\dfrac{441}{25000}$
3. $\dfrac{613}{648}$
4. $\dfrac{6}{13}$
5. $\dfrac{A_{12}^{10}}{12^{10}}$
6. $\dfrac{1}{6}$
7. $\dfrac{15}{36}$

1.4. Условная вероятность. Теорема умножения

Пусть Ω – пространство элементарных событий с вероятностным распределением P. Пусть A и B – некоторые события на Ω, причем $P(A) > 0$.

Определение 1.4.1. *Число $P(B|A) = \dfrac{P(A \cap B)}{P(A)}$ называется условной вероятностью события B при условии, что событие A наступило (коротко говорят „вероятностью B при условии A“).*

Пример 1.4.1. Студенческая группа состоит из 12 студентов. При этом в ней 7 спортсменов, 5 отличников и из спортсменов 3 являются отличниками. На соревнования выбрали одного спортсмена. Какова вероятность, что он отличник?

Решение. Определим события. $A =${выбран спортсмен}, $B =${выбран отличник}, $A \cap B =${выбран спортсмен-отличник}. Надо найти условную вероятность $P(B|A)$.

$P(A) = \dfrac{7}{12}$, $P(A \cap B) = \dfrac{3}{12}$, тогда по формуле из определения условной вероятности получаем: $P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{3}{7}$.

Пример 1.4.2. Игральная кость подбрасывается один раз. Известно, что выпало более трёх очков. Какова вероятность того, что выпало чётное число очков?

Решение. Пусть событие $A =$\{выпало более трех очков\}, а событие $B =$\{выпало четное число очков\}. По условию задачи событие A наступило. Требуется найти условную вероятность $P(B|A)$.

При одном подбрасывании выпало более трех очков – это значит, что выпало 4, 5 или 6 очков. Следовательно, $P(A) = \dfrac{3}{6} = \dfrac{1}{2}$. Событие $A \cap B =$\{выпало 4 или 6 очков\}, его вероятность $P(A \cap B) = \dfrac{2}{6} = \dfrac{1}{3}$.

Наконец, искомая вероятность $P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{2}{3}$.

Пример 1.4.3. Четыре человека А, Б, В, Г становятся в очередь в случайном порядке. Найдите условную вероятность того, что А первый, при условии, что Б – последний.

Решение. Обозначим события $A =$\{А – первый в очереди\}, $B =$\{Б – последний в очереди\}. Надо найти условную вероятность $P(A|B)$. Для этого рассмотрим событие $A \cap B =$\{А – первый, а Б – последний в очереди\}. Количество элементарных событий, благоприятных для него, $|A \cap B| = 2$, так как А и Б должны занять определенные места в очереди: первое и последнее соответственно, а для В и Г остаются два места, которые они могут занять двумя способами. Всего элементарных событий $|\Omega| = 4!$, т. е. количество способов занять четыре места в очереди четырьмя людьми. Следовательно, $P(A \cap B) = \dfrac{2}{4!} = \dfrac{1}{12}$.

Вычислим вероятность события B. Количество элементарных событий, благоприятных для него, $|B| = 3!$, поскольку Б должен занять последнее место в очедеди, оставшиеся три человека могут занять оставшиеся три места произвольно. Значит $P(B) = \dfrac{3!}{4!} = \dfrac{1}{4}$. По формуле из определения условной вероятности получаем $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{1}{3}$.

Перепишем формулу из определения условной вероятности в виде

$$P(A \cap B) = P(A)P(B|A).$$

Эту формулу традиционно называют **теоремой умножения**, хотя доказывать здесь ничего не нужно, лишь бы $P(A)$ была больше нуля. Теорема умножения позволяет записать вероятность произведения событий через произведение безусловной и условной вероятностей этих событий.

Понятие условной вероятности позволяет описывать целую группу случайных экспериментов со сложной структурой.

Пример 1.4.4. В каждой из двух урн имеется по 2 белых и 3 черных шара. Из первой урны наудачу извлекают один шар и перекладывают во вторую урну. Затем из второй урны наудачу достают один шар и перекладывают в первую. Найти вероятность того, что состав шаров в урнах не изменится.

Решение. В данной задаче имеем дело со сложным экспериментом, состоящим из двух частей. Исходы первой части эксперимента: $\delta_1 =$\{из первой урны вынули белый шар\}, $r_1 =$\{из первой урны вынули черный шар\}. Вероятности этих исходов $P(\delta_1) = \dfrac{2}{5}$, $P(r_1) = \dfrac{3}{5}$.

Исходы второй части эксперимента обозначим $\delta_2 =$\{из второй урны вынули белый шар\} и $r_2 =$\{из второй урны вынули черный шар\}. Вероятности этих исходов зависят от исходов первой части эксперимента, т. е. это условные вероятности $P(\delta_2|\delta_1) = \dfrac{3}{6} = \dfrac{1}{2}$, $P(\delta_2|r_1) = \dfrac{2}{6} = \dfrac{1}{3}$, $P(r_2|\delta_1) = \dfrac{3}{6} = \dfrac{1}{2}$, $P(r_2|r_1) = \dfrac{4}{6} = \dfrac{2}{3}$.

Интересующее нас событие обозначим $C =$\{состав шаров в урнах не изменится\}. Оно может произойти, если из первой и из второй урны извлекают шары одного цвета. Вероятность того, что из первой и второй урны достанут белые шары, на основании теоремы умножения равна

$$P(\delta_1 \cap \delta_2) = P(\delta_1)P(\delta_2|\delta_1) = \frac{2}{5} \cdot \frac{1}{2} = \frac{1}{5}.$$

Вероятность того, что из обеих урн достанут черные шары

$$P(r_1 \cap r_2) = P(r_1)P(r_2|r_1) = \frac{3}{5} \cdot \frac{2}{3} = \frac{2}{5}.$$

По теореме сложения

$$P(C) = \frac{1}{5} + \frac{2}{5} = \frac{3}{5}.$$

Пример 1.4.5. В трамвайном парке имеются 15 трамваев маршрута №1 и 10 трамваев маршрута №2. Трамваи выходят на линию в произвольном порядке. Какова вероятность того, что вторым по счету на линию выйдет трамвай маршрута №1?

Решение. Пусть событие A_1 ={трамвай маршрута №1 вышел на линию первым}, B_1 ={трамвай маршрута №2 вышел на линию первым}, A_2 ={трамвай маршрута №1 вышел на линию вторым}, B_2 ={трамвай маршрута №2 вышел на линию вторым}.

Рассмотрим все события, которые могут произойти в условиях нашей задачи:

$A_1 \cap A_2$ ={трамвай маршрута №1 выйдет на линию первым и вторым по счету};

$A_1 \cap B_2$ ={трамвай маршрута №1 выйдет на линию первым, а №2 вторым по счету};

$B_1 \cap A_2$ ={трамвай маршрута №2 выйдет на линию первым, а №1 вторым по счету};

$B_1 \cap B_2$ ={трамвай маршрута №2 выйдет на линию первым и вторым по счету}.

Из них нас будут интересовать только первое и третье, когда вторым выйдет трамвай маршрута №1. По теореме умножения

$P(A_1 \cap A_2) = P(A_1)P(A_2|A_1) = \frac{15}{25} \cdot \frac{14}{24}$,

$P(B_1 \cap A_2) = P(B_1)P(A_2|B_1) = \frac{10}{25} \cdot \frac{15}{24}$.

Используя теорему сложения, получаем вероятность интересующего нас события A ={вторым по счету выйдет трамвай маршрута №1}.

$P(A) = \frac{15}{25} \cdot \frac{14}{24} + \frac{10}{25} \cdot \frac{15}{24} = \frac{3}{5}$.

Вернемся к определению условной вероятности и рассмотрим случай, когда вероятность события B не зависит от того, произошло событие A или нет, т. е. $P(B|A) = P(B)$. Из определения получаем:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = P(B)$$

или

$$P(A \cap B) = P(A)P(B).$$

Следовательно, если событие B не зависит от события A, то и A не зависит от B, поскольку последнее равенство симметрично.

Определение 1.4.2. *Cобытия A и B называются независимыми,* *если выполняется равенство* $P(A \cap B) = P(A)P(B)$.

Смысл этого формального определения заключается в том, что если произошло одно из двух независимых событий, то это никак не влияет на вероятность другого события. Но в таком случае, если первое событие не произошло, то это также не должно влиять на вероятность второго.

Утверждение 1.4.1. *Если события A и B независимы, то события $\overline{A}$ и B также независимы.*

Если события A и B независимы, то события $\overline{A}$ и $\overline{B}$ также независимы.

Если для событий A и B не выполнено равенство из определения независимости событий, то они называются зависимыми.

Определение независимости событий можно обобщить.

Определение 1.4.3. *События $A_1, A_2, \ldots, A_n$ называются независимыми в совокупности, если для любых k из них ($k \leq n$) выполняется условие*

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = \prod_{j=1}^{k} P(A_{i_j}).$$

Если соотношение выполняется только при $k = 2$, то события называются попарно независимыми.

Пример 1.4.6. Из 100 студентов, находящихся в аудитории, 50 человек знают английский язык, 40 – французский и 35 – немецкий. Английский и французский языки знают 20 студентов, английский и немецкий – 8, французский и немецкий – 10. Все три языка знают 5 человек. Один из студентов вышел из аудитории. Рассмотрим следующие события: $A =${вышедший знает английский язык}, $B =${вышедший знает французский язык}, $C =${вышедший знает немецкий язык}.

Указать все пары независимых событий. Установить, являются ли события A, B и C независимыми в совокупности.

Решение. Вычислим вероятности указанных событий:
$P(A) = \dfrac{50}{100} = \dfrac{1}{2}$, $P(B) = \dfrac{40}{100} = \dfrac{2}{5}$, $P(C) = \dfrac{35}{100} = \dfrac{7}{20}$.
Исследуем события на попарную независимость.

Вероятность того, что вышедший студент знает и английский, и французский языки, равна $P(A \cap B) = \dfrac{20}{100} = \dfrac{1}{5}$, что в свою очередь равно произведению вероятностей $P(A)P(B) = \dfrac{1}{2} \cdot \dfrac{2}{5} = \dfrac{1}{5}$. Следовательно, по определению события A и B независимы.

Проверим, зависимы или независимы события A и C: $P(A \cap C) = \dfrac{8}{100} = \dfrac{2}{25}$, однако, произведение вероятностей $P(A)P(C) = \dfrac{1}{2} \cdot \dfrac{7}{20} = \dfrac{7}{40}$. То есть равенство из определения независимости событий не выполняется, а значит, события A и C зависимы.

Для событий B и C: $P(B \cap C) = \dfrac{10}{100} = \dfrac{1}{10}$, $P(B)P(C) = \dfrac{2}{5} \cdot \dfrac{7}{20} = \dfrac{7}{50}$. Очевидно, равенство опять не выполняется, поэтому события B и C зависимы.

События A, B и C не являются независимыми в совокупности, поскольку они не являются даже попарно независимыми. Это только что было показано.

Покажем, что попарной независимости недостаточно для независимости в совокупности.

Пример 1.4.7. (Бернштейна). Пусть одна из граней правильного тетраэдра окрашена в красный цвет, вторая в зеденый, а третья в желтый. На четвертой грани есть и красный и зеленый и желтый цвета. Эксперимент состоит в подбрасывании тетраэдра. Покажите что события:

А – тетраэдр упал на грань, содержащую красный цвет;

В – тетраэдр упал на грань, содержащую зеленый цвет;

С – тетраэдр упал на грань, содержащую желтый цвет,

попарно независимы, но не являются независимыми в совокупности.

Решение. $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, где элементарное событие ω_i состоит в выпадении i-ой грани. Поскольку тетраэдр правильный, $P(\omega_i) = 1/4$, $i = 1, 2, 3, 4$. $A = \{\omega_1, \omega_4\}$, $B = \{\omega_2, \omega_4\}$, $C = \{\omega_3, \omega_4\}$, поэтому

$$A \cap B = B \cap C = A \cap C = A \cap B \cap C = \{\omega_4\}.$$

Так как $P(A) = P(B) = P(C) = 1/2$, а $P(\omega_4) = 1/4 = 1/2 \cdot 1/2$, события A, B, C являются попарно независимыми. Однако они не являются независимыми в совокупности:

$$P(A \cap B \cap C) = P(\omega_4) = 1/4, \qquad P(A)P(B)P(C) = 1/2 \cdot 1/2 \cdot 1/2 = 1/8.$$

Пример 1.4.8. Пусть в мешке лежат 10 шаров, из которых 6 белых и 4 чёрных. Какова вероятность события: два раза подряд вытащен чёрный шар?

Решение. Задача поставлена некорректно: требуется уточнить, что делать с вытащенным шаром: возвращать в мешок, так что следующий шар снова вытаскивается из мешка со всеми десятью шарами, или откладывать в сторону, и тогда к следующему вытаскиванию остаются девять шаров.

В первом случае события A ={первый вытащенный шар чёрный} и B ={второй вытащенный шар чёрный} независимы: вероятность получить второй раз чёрный шар не зависит от того, что было в первый раз. Вероятности этих событий равны $P(A) = P(B) = \dfrac{4}{10} = \dfrac{2}{5}$.

Поэтому вероятность события C ={дважды вынули черный шар} равна произведению вероятностей

$$P(C) = P(A \cap B) = P(A)P(B) = \frac{2}{5} \cdot \frac{2}{5} = \frac{4}{25}.$$

Если же первый вытащенный шар обратно не возвращается, то в мешке перед вторым вытаскиванием остаются девять шаров. В этом случае вероятность события B ={второй вытащенный шар чёрный} зависит от того, произошло событие A или нет. События A и B являются зависимыми, и по теореме умножения получаем

$$P(C) = P(A \cap B) = P(A)P(B|A) = \frac{4}{10} \cdot \frac{3}{9} = \frac{2}{15}.$$

Пример 1.4.9. Подбрасывают наудачу три игральные кости. Установите, зависимы или независимы события A ={хотя бы на одной из костей выпадет “шестерка”} и B ={„пятерка” не выпадет}.

Решение. Найдем вероятности указанных событий. Вероятность события A удобнее искать через вероятность противоположного события.

$\overline{A}$ ={„шестерка" не выпадет}. То есть на всех трех костях может выпасть любое число от 1 до 5. Количество всех элементарных событий $|\Omega| = 6^3$, а количество элементарных событий, благоприятных для $\overline{A}$, $|\overline{A}| = 5^3$. Следовательно,

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{5^3}{6^3} = \frac{91}{216}.$$

Количество элементарных событий, благоприятных для B, такое же, как и для $\overline{A}$, поэтому

$$P(B) = P(\overline{A}) = \frac{5^3}{6^3} = \frac{125}{216}.$$

Теперь рассмотрим событие $A \cap B$ ={хотя бы на одной из костей выпадет „шестерка", причем „пятерка" не выпадет}. Это событие является суммой трех несовместных событий: C_1 ={выпала одна „шестерка" и нет „пятерок"}, C_2 ={выпали две „шестерки" и нет „пятерок"}, C_3 ={выпали три „шестерки"}.

Определим количество элементарных событий, благоприятных для C_1. „Шестерка" может выпасть на любой кости из трех, а на двух оставшихся костях – любое из чисел 1,2,3,4. Следовательно, $|C_1| = 3 \cdot 4 \cdot 4$. Для события C_2: две „шестерки" – на двух костях, и на одной кости может выпасть 1,2,3,4, причем, эта кость может быть любой из трех. Поэтому $|C_2| = 3 \cdot 4$. Очевидно, что $|C_3| = 1$.

По теореме сложения получаем

$$P(A \cap B) = P(C_1) + P(C_2) + P(C_3) = \frac{48}{6^3} + \frac{12}{6^3} + \frac{1}{6^3} = \frac{61}{216}.$$

Осталось проверить равенство в определении независимости событий:

$$P(A)P(B) = \frac{91}{216} \cdot \frac{125}{216} = \frac{11375}{46656} \cong 0,2438,$$

$$P(A \cap B) = \frac{61}{216} \cong 0,2824.$$

Равенство не выполняется, следовательно, события A и B зависимы.

Задачи для самостоятельного решения

1. В урне 3 белых и 3 черных шара. Из урны дважды вынимают по од-

ному шару, не возвращая их обратно. Найти вероятность появления белого шара при втором испытании (событие B), если при первом испытании был извлечен черный шар (событие A).

2. Четыре человека А, Б, В, Г становятся в очередь в случайном порядке. Найдите:

а) условную вероятность того, что А первый, если А не последний;

б) условную вероятность того, что А первый, если Б не последний;

в) условную вероятность того, что А первый, если Б стоит в очереди позже А;

г) условную вероятность того, что А стоит в очереди раньше Б, если известно, что А раньше В.

3. Какова вероятность того, что 2 карты, вынутые из колоды в 36 карт, окажутся одной масти?

4. В урне 4 белых и 5 черных шаров. Какова вероятность вытащить белый шар четыре раза подряд, если а) шары берутся с возвращением; б) шары берутся без возвращения. В каком случае вероятность больше?

5. Бросают два кубика: красный и синий. Будут ли события $A =${на красном кубике выпала чётная цифра} и $B =${на синем кубике выпала „пятёрка" или „шестёрка"} независимы?

6. Подбрасывают наудачу три игральные кости. Установить, зависимы или независимы события $A =${„шестерка“ выпадет не менее двух раз} и $B =${на одной кости выпадет „пятерка“}.

Ответы к задачам для самостоятельного решения

1. $\dfrac{3}{5}$.

2. а) $\dfrac{1}{3}$; б) $\dfrac{2}{9}$; в) $\dfrac{1}{2}$; г) $\dfrac{2}{3}$.

3. $\dfrac{8}{35}$.

4. а) $\dfrac{256}{6561} \cong 0,039$; б) $\dfrac{1}{126} \cong 0,008$.

5. да.

6. зависимы.

1.5. Формула полной вероятности

Следствием двух основных теорем теории вероятностей – теоремы сложения и теоремы умножения – являются формула полной вероятности и формула Байеса.

Определение 1.5.1. *События $H_1, H_2, \ldots, H_n$, являющиеся попарно несовместными, т. е. $H_i \cap H_j = \emptyset$ при $i \neq j, i, j = 1, 2...., n$ и дающие в объединении все пространство, т. е. $H_1 \cup H_2 \cup \ldots \cup H_n = \Omega$, называются полной системой событий (или гипотезами).*

Теорема 1.5.1. (Формула полной вероятности). *H_1, H_2,..., H_n – полная система событий, и A произвольное событие. Тогда справедлива следующая формула, которую называют* **формулой полной вероятности***:*

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + ... + P(H_n)P(A|H_n) =$$
$$= \sum_{i=1}^{n} P(H_i)P(A|H_i).$$

Пример 1.5.1. В ящике содержится 12 деталей, изготовленных на заводе № 1, 20 деталей, изготовленных на заводе № 2, и 18 деталей, изготовленных на заводе № 3. Вероятность того, что деталь, изготовленная на заводе № 1, отличного качества, равна 0,9; для деталей, изготовленных на заводах № 2 и № 3, эти вероятности соответственно равны 0,6 и 0,8. Найти вероятность того, что наугад извлеченная деталь отличного качества.

Решение. Введем обозначения:
событие H_1 — извлеченная деталь изготовлена на заводе № 1;
событие H_2 — извлеченная деталь изготовлена на заводе № 2;
событие H_3 — извлеченная деталь изготовлена на заводе № 3;
событие A — извлеченная деталь отличного качества.

Всего в ящике $12 + 20 + 18$ деталей. Из них 12 деталей, изготовленных на первом заводе. Поэтому вероятность выбрать деталь, изготовленную на первом заводе $P(H_1) = 12/50 = 0,24$. Аналогично вероятности выбрать детали, изготовленные на втором и третьем заводах, составляют соответственно $P(H_2) = 20/50 = 0,4$ и $P(B_3) = 18/50 = 0,36$. По условию вероятность того, что деталь, изготовленная на заводе № 1 отличного качества, равна 0,9, т. е. $P(A|H_1) = 0,9$, вероятность того, что деталь, изготовленная на заводе № 2 отличного качества, равна $0,6$, т. е.$P(A|H_2) = 0,6$, вероятность того, что деталь, изготовленная на заводе № 3 отличного качества, равна $0,8$, т. е. $P(A|H_3) = 0,8$.

Используя формулу полной вероятности получаем:

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + P(H_3)P(A|H_3) =$$
$$= 0,24 \cdot 0,9 + 0,4 \cdot 0,6 + 0,36 \cdot 0,8 = 0,744.$$

Пример 1.5.2. Из первой урны, содержащей 8 белых и 2 черных шара, перекладывают два шара во вторую урну, содержащую 6 белых и 4 черных шара, а затем из второй урны вынимают один шар. Какова вероятность того, что этот шар черный?

Решение. Введем обозначения:
событие H_1 – из первой урны вынули два белых шара;
событие H_2 – из урны вынули два черных шара;
событие H_3 – из урны вынули один белый и один черный шар;
событие A – случайно выбранный из второй урны шар оказался черным.

По теореме умножения $P(H_1) = \dfrac{8}{10} \cdot \dfrac{7}{9} = \dfrac{56}{90}$; $P(H_2) = \dfrac{2}{10} \cdot \dfrac{1}{9} = \dfrac{2}{90}$; $P(H_3) = \dfrac{8}{10} \cdot \dfrac{2}{9} + \dfrac{2}{10} \cdot \dfrac{8}{9} = \dfrac{32}{90}$;

Если произошло событие H_1, то во второй урне стало 8 белых и 4 черных шара. В этом случае вероятность вытащить из второй урны черный шар составляет $P(A|H_1) = \dfrac{4}{12}$. Если произошло событие H_2, то во второй урне стало 6 белых и 6 черных шаров. В этом случае вероятность вытащить из второй урны черный шар составляет $P(A|H_2) = \dfrac{6}{12}$. Если произошло событие H_3, то во второй урне стало 7 белых и 5 черных шаров. В этом случае вероятность вытащить из второй урны черный шар составляет $P(A|H_3) = \dfrac{5}{12}$.

По формуле полной вероятности:

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + P(H_3)P(A|H_3) =$$
$$= \frac{56}{90} \cdot \frac{4}{12} + \frac{2}{90} \cdot \frac{6}{12} + \frac{32}{90} \cdot \frac{5}{12} = \frac{11}{30}$$

Задачи для самостоятельного решения

1. В магазине продаются электролампы производства трех заводов,

причем доля первого завода 30, второго 50, третьего 20 %. Брак в их продукции сотавляет 5, 3 и 2 % соответственно. Найти вероятность того, что случайно выбранная лампа оказалась бракованной.

2. Из первой урны, содержащей 6 белых и 4 черных шара, перекладывают 2 шара во вторую урну, содержащую 8 белых и 2 черных шара, а затем из второй урны вынимают 1 шар. Какова вероятность того, что этот шар белый?

3. В первой урне находится 6 белых и 2 черных шара, во второй 6 белых и 4 синих, в третьей 5 белых и 5 красных. Из первой и второй урны наудачу извлекают по одному шару и кладут в третью. После этого из третьей урны вынимают один шар. Найти вероятность того, что этот шар будет белым.

Ответы к задачам для самостоятельного решения

1. 0,034 2. $\frac{23}{30}$ 3. $\frac{127}{240}$

1.6. Формула Байеса

Вероятности событий $H_1, H_2, ..., H_n$, составляющих полную систему гипотез, называют априорными. Если же известно, что произошло некоторое событие A, то могут быть вычислены апостериорные вероятности $P(H_i|A)$. На бытовом языке это означает, что появление информации о каком-то событии может изменять вероятности других событий (что и заложено в понятии условной вероятности). Апостериорные вероятности могут быть вычислены при помощи формулы, получившей название по фамилии математика Томаса Байеса.

Теорема 1.6.1. (Теорема Байеса). *Пусть H_1, $H_2, \ldots$, H_n – полная система событий, A – произвольное событие. Тогда справедлива следующая формула:*

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + ... + P(H_n)P(A|H_n)}.$$

Пример 1.6.1. Противотанковая батарея состоит из 10 орудий. Для первой группы из 6 орудий вероятности того, что при одном выстреле произойдет недолет, попадание или перелет, равны соответственно $0,1$; $0,7$;

$0,2$. Для каждого из остальных четырех орудий вероятности тех же самых событий равны соответственно $0,2$; $0,6$; $0,2$. Наудачу выбранное орудие произвело три выстрела по цели, в результате чего было зафиксировано одно попадание, один недолет и один перелет. Какова вероятность того, что стрелявшее орудие принадлежит первой группе?

Решение. Полная система событий состоит из двух событий: H_1 – выбрано орудие первой группы, и H_2 – выбрано орудие второй группы. Очевидно, $P(H_1) = 0,6$ и $P(H_2) = 0,4$. Событие A – при трех выстрелах зафиксированы одно попадание, один недолет и один перелет. $P(A|H_1) = 6 \cdot 0,1 \cdot 0,7 \cdot 0,2 = 0,084$, так как всего существует шесть последовательностей из трех различных исходов (см. количество перестановок). Аналогично $P(A|H_2) = 6 \cdot 0,2 \cdot 0,6 \cdot 0,2 = 0,024$. Теперь применим формулу Байеса для получения искомой вероятности:
$P(H_1|A) = \dfrac{0,6 \cdot 0,084}{0,6 \cdot 0,084 + 0,4 \cdot 0,024} = 7/15 \approx 0,4666.$

Пример 1.6.2. Прибор, установленный на борту самолета, может работать в двух режимах: в условиях крейсерского полета и в условиях перегрузки при взлете и посадке. Крейсерский режим полета занимает 80 % всего вермени полета, а полет в условиях перегрузки — 20 %. Вероятность отказа прибора за время крейсерского полета равна 0,1, в условиях перегрузки – 0,4. Найдите надежность (вероятность безупречной работы) прибора во время полета. Известно, что во время полета прибор отказал. Найти вероятность того, что полет проходил а) в крейсерском режиме и б)в условиях перегрузки.

Решение. Событие А – во время полета произойдет отказ прибора. Выдвинем гипотезы: H_1 – прибор работает в условиях нормального крейсерского полета; H_2 – прибор работает в условиях перегрузки.

Тогда $P(H_1) = 0,8; P(H_2) = 0,2;\ P(A|H_1) = 0,1;\ P(A|H_2) = 0,4.$

Используя формулу полной вероятности, найдем вероятность отказа прибора:

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) = 0,8 \cdot 0,1 + 0,2 \cdot 0,4 = 0,16.$$

Тогда вероятность того, что прибор работает безупречно $P(\bar{A}) = 0,84$.

Используя формулу Байеса получим:

$$P(H_1|A) = \frac{P(H_1)P(A|H_1)}{P(A)} = \frac{0,8 \cdot 0,1}{0,16} = 0,5.$$

$$P(H_2|A) = \frac{P(H_1)P(A|H_2)}{P(A)} = \frac{0,2 \cdot 0,4}{0,16} = 0,5.$$

Задачи для самостоятельного решения

1. На заводе, изготавливающем болты, первая машина производит 25, вторая 35, третья 40 % всех изделий. В их продукции брак составляет соответственно 5, 4 и 2 %. Найдите:

a) вероятность того, что случайно выбранный болт оказался дефектным;

b) вероятность того, что случайно выбранный болт, оказавшийся дефектным, был произведен первой, второй и третьей машиной.

2. В торговую фирму поступили телевизоры от трех поставщиков в отношении 1:4:5. Практика показала, что телевизоры, поступившие от 1-го, 2-го и 3-го поставщиков, не потребуют ремонта в течение гарантийного срока соответственно в 98, 88 и 92 % случаев. Найдите:

a) вероятность того, что поступивший в торговую фирму телевизор не потребует ремонта в течение гарантийного срока;

b) от какого поставщика вероятнее всего поступил проданный телевизор, потребовавший ремонта в течение гарантийного срока.

3. Известно, что в среднем 95 % выпускаемой продукции удовлетворяют страндарту. Упрощенная схема контроля признает пригодной продукцию с вероятностью 0.98, если она станартна, и с вероятностью 0.06, если она нестанартна. Определите вероятность того, что взятое наудачу изделие пройдет упрощенный контроль. Какова вероятноять, что изделие стандартное, если оно а) прошло упрощенный контроль; б) дважды прошло упрощенный контроль.

4. В первой урне 3 белых и 5 черных шаров, во второй – 4 белых и 4 черных. Случайным образом из первой урны достают два шара и перекладывают во вторую. После этого из второй урны вынимают шар. Он

оказался белым. Найдите вероятность того, что в первой урне только один из шаров белый.

5. Однотипные приборы выпускаются тремя заводами в количественном отношении $5 : 3 : 7$. Вероятности брака для этих заводов равны соответственно 0,1; 0,3 и 0,2. Прибор, приобретенный научно-исследовательским институтом, оказался бракованным. Какова вероятность того, что данный прибор произведен первым заводом?

Ответы к задачам для самостоятельного решения

1. 0.0345; $\dfrac{25}{69}$; $\dfrac{28}{69}$; $\dfrac{16}{69}$.
2. 0.91; от третьего поставщика.
3. 0.934; 0.997; 0.9998.
4. 18/133.
5. $5/28 \approx 0,17857$.

1.7. Испытания Бернулли

Рассмотрим случайный эксперимент с двумя возможными исходами: ω_1 (назовем его „успех”) и ω_2 (назовем его „неудача”). Пусть вероятность „успеха” $P(\omega_1) = p$, соответственно, $P(\omega_2) = 1 - p = q$.

Эксперимент повторяется n раз, причем результаты разных экспериментов не зависят друг от друга, т. е. события вида $\{\omega_j$ — исход i-го эксперимента$\}$ $(i = 1, \dots, n, \quad j = 1, 2)$ являются независимыми в совокупности. Такая модель называется последовательностью независимых испытаний Бернулли.

Нас интересует, сколько раз наступит „успех” в этих n испытаниях. Вероятность того, что „успех” наступит ровно k раз (безразлично, в какой последовательности), равна

$$P_n(k) = C_n^k p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}.$$

Данное соотношение называется **формулой Бернулли**.

Пример 1.7.1. Симметричная монета подбрасывается 5 раз. Найти вероятности следующих событий: $A =$\{в каждом подбрасывании выпадет „орел”\}, $B =$\{„орел” выпадет три раза\}, $C =$\{три раза выпадет „решка”\}, $D =$\{„орел” не выпадет ни разу\}.

Решение. В качестве „успеха" будем рассматривать выпадение „орла" при одном подбрасывании монеты. Поскольку монета симметричная, вероятность „успеха" равна вероятности „неудачи" $p = q = 1/2$. Независимых испытаний всего $n = 5$. Значит

$$P(A) = P_5(5) = p^5 = \left(\frac{1}{2}\right)^5 = \frac{1}{32},$$

$$P(B) = P_5(3) = C_5^3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = \frac{5!}{3!2!}\left(\frac{1}{2}\right)^5 = \frac{10}{32}.$$

Так как вероятности „успеха" и „неудачи" равны, то $P(C) = P(B) = \dfrac{10}{32}$. $P(D) = P_5(0) = (1-p)^5 = \left(\dfrac{1}{2}\right)^5 = \dfrac{1}{32}$.

Пример 1.7.2. Вероятность того, что телевизор имеет скрытые дефекты, равна 0,2. На склад поступило 20 телевизоров. Какое событие вероятнее: в этой партии имеется два телевизора со скрытыми дефектами или три?

Решение. Пусть, „успех" означает что телевизор имеет скрытые дефекты, вероятность „успеха" $p = 0,2$. Количество независимых испытаний Бернулли $n = 20$. Обозначим интересующие нас события: $A =${в партии два телевизора со скрытыми дефектами}, $B =${в партии три телевизора со скрытыми дефектами}. Осталось сравнить их вероятности.

$$P(A) = P_{20}(2) = C_{20}^2\left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^{18} = \frac{20!}{2!18!}\left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^{18} \approx 0,137.$$

$$P(B) = P_{20}(3) = C_{20}^3\left(\frac{1}{5}\right)^3\left(\frac{4}{5}\right)^{17} = \frac{20!}{3!17!}\left(\frac{1}{5}\right)^3\left(\frac{4}{5}\right)^{17} \approx 0,205.$$

Очевидно, $P(B) > P(A)$, т. е. вероятность получить три телевизора со скрытыми дефектами больше вероятности получить только два таких телевизора.

Пример 1.7.3. Вероятность выигрыша по одному лотерейному билету равна 0,3. Вы купили 8 билетов. Найти вероятность того, что а) хотя бы один билет выигрышный; б) менее трех билетов выигрышные.

Решение. Логично определить „успех" как выигрышный билет. Очевидно, $p = 0,3$, $n = 8$ в формуле Бернулли.

Рассмотрим событие $A =${хотя бы один из купленных билетов выигрышный}. Подсчитывая вероятность этого события, удобнее перейти к противоположному событию $\overline{A}$ ={нет ни одного билета с выигрышем}.

$$P(A) = P_8(k \geq 1) = 1 - P_8(k < 1) = 1 - P_8(0) = 1 - (1-p)^8 = 1 - (0,7)^8 \approx 0,942.$$

Событие B ={менее трех билетов выигрышные} является сложным, его вероятность равна сумме вероятностей трех событий;

$$P(B) = P_8(k < 3) = P_8(0) + P_8(1) + P_8(2) = (0,7)^8 + C_8^1(0,3)(0,7)^7 + C_8^2(0,3)^2(0,7$$

Пример 1.7.4. Стрелок попадает в мишень с вероятностью 0,8. Найти вероятность того, что стрелок поразил мишень с пятой попытки.

Решение. Будем считать „успехом“ попадание в мишень. По условию вероятность „успеха“ $p = 0,8$. Следовательно, вероятность „неудачи“ $q = 0.2$. Случайное событие A – первый „успех“ наступил в пятом испытании. Вследствие независимости испытаний

$$P(A) = q^4 p = 0,2^4 \cdot 0,8 = 0,00128.$$

Задачи для самостоятельного решения

1. Играют равносильные противники. Что вероятнее: выиграть не менее трех партий из четырех или не менее шести из восьми? (Ничьи не учитываются).

2. Из 100 аккумуляторов за год хранения 5 выходит из строя. Наудачу выбирают 9 аккумуляторов. Определите вероятность того, что среди них 7 исправных.

3. Какова вероятность, что игрок, который слабее своего оппонента в два раза, выиграет две партии из трех?

4. Вероятность того, что родившийся ребенок — мальчик, равна 0,51. Какова вероятность того, что в семье из шести детей: одна или две девочки?

5. Производится N независимых выстрелов зажигательными снарядами по резервуару с горючим. Каждый снаряд попадает в резервуар с ве-

роятностью p. Если в резервуар попал один снаряд, то горючее воспламеняется с вероятностью 0,7, если два снаряда — с полной достоверностью. Найти вероятность того, что при N выстрелах горючее воспламенилось.

Ответы к задачам для самостоятельного решения

1. $P_4(3)+P_4(4)=\dfrac{5}{16}$; $P_8(6)+P_8(7)+P_8(8)=\dfrac{37}{256}$. Вероятнее выиграть не менее трех партий из четырех;

2. 0,063;

3. $\dfrac{2}{9}$;

4. 0,345;

5. $1-q^N-Npq^{N-1}0,3$.

2. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

2.1. Дискретная случайная величина

Определение 2.1.1. *Дискретной случайной величиной называется вещественная функция, заданная на множестве элементарных событий*

$$\xi:\Omega\to\mathrm{R},$$

множество значений которой конечно или счетно.

Пусть вероятность того, что случайная величина ξ принимает значение x_i равна p_i , т.е. $P(\xi=x_i)=p_i$.

Определение 2.1.2. *Набор значений случайной величины вместе с вероятностями принятия значений*

$$\begin{array}{cccccc} \xi: & x_1 & x_2 & \ldots & x_n & \ldots \\ p_i: & p_1 & p_2 & \ldots & p_n & \ldots, \end{array}$$

где $p_i\in[0,1]$ *и* $\sum\limits_i p_i=1$ *называется рядом распределения дискретной случайной величины.*

Определение 2.1.3. *Функцией распределения случайной величины* ξ *называется функция*

$$F_\xi(x)=P(\xi<x),\quad x\in\mathrm{R},$$

определенная на всей вещественной оси.

Для дискретной случайной величины функция распределения — непрерывная слева кусочно-постоянная функция с разрывами в точках x_i. Причем, в каждой точке x_i значение функции увеличивается на величину p_i.

Пусть значения x_i пронумерованы в порядке возрастания и нумерация начинается с 1. Тогда ряду распределения (x_i, p_i) соответствует функция распределения $F_\xi(x) = \sum\limits_{i: x_i < x} p_i$ или

$$F_\xi(x) = \begin{cases} 0, & \text{если } x \le x_1; \\ p_1, & \text{если } x_1 < x \le x_2; \\ p_1 + p_2, & \text{если } x_2 < x \le x_3; \\ \ldots, & ....; \\ p_1 + p_2 + ... + p_k, & \text{если } x_k < x \le x_{k+1}; \\ \ldots, & ....; \\ 1, & \text{если } x > sup_i(x_i). \end{cases}$$

Пример 2.1.1. Ряд распределения дискретной случайной величины

x_i	-1	2	4	5
p_i	0,1	0,6	0,1	0,2

. Найдите функцию распределения.

Решение. Считая суммы вероятностей в соответствии с общей формулой, получаем $F_\zeta(x) = \begin{cases} 0, & x \le -1, \\ 0.1, & -1 < x \le 2, \\ 0.7, & 2 < x \le 4, \\ 0.8, & 4 < x \le 5, \\ 1, & x > 5. \end{cases}$

Ряд распределения или функция распределения дают полную информацию о дискретной случайной величине. Если же мы хотим представить информацию о ней в сжатом виде, то описываем её числовыми характеристиками. Наиболее важное место среди числовых характеристик занимает математическое ожидание или среднее значение случайной величины.

Определение 2.1.4. *Математическим ожиданием дискретной случайной величины ξ называется число, равное сумме произведений всех ее* *значений на соответствующие им вероятности:*

$$M\xi = \sum_i x_i p_i.$$

Сумма конечная если случайная величина принимает конечное число значений или сумма ряда (бесконечная), если случайная величина принимает счетное число значений. Математическое ожидание существует, если этот ряд сходится абсолютно. В противном случае говорят, что случайная величина ξ не имеет математического ожидания. Математическое ожидание обладает свойством линейности и имеет ту же размерность, что и случайная величина.

Пример 2.1.2. Стрелок поражает мишень с вероятностью 0,7. Он стреляет до первого попадания, но у него только четыре патрона. Найти математическое ожидание числа выстрелов.

Решение. Пусть случайная величина ξ — число выстрелов. Так как у стрелка 4 патрона, ξ принимает значения $1, 2, 3, 4$. $\xi = 1$ означает что стрелок поразил мишень первым выстрелом, $P(\xi = 1) = 0,7$; $\xi = 2$ означает что стрелок поразил мишень вторым выстрелом (первый раз промазал). $P(\xi = 2) = 0,3 \cdot 0,7$. Аналогично $P(\xi = 3) = (0,3)^2 \cdot 0,7$. То, что потребовался четвертый выстрел, означает промах в первых трех и $P(\xi = 4) = (0,3)^3$.

$$M\xi = 1 \cdot 0,7 + 2 \cdot 0,21 + 3 \cdot 0,063 + 4 \cdot 0,027 = 1,417.$$

Другая важная характеристика — степень рассеивания, разброса значений, принимаемых случайной величиной, вокруг математического ожидания. Мерой этого разброса является дисперсия.

Определение 2.1.5. *Дисперсией $D\xi$ дискретной случайной величины ξ называется математическое ожидание квадрата ее отклонения от математического ожидания:*

$$D\xi = M((\xi - M\xi)^2) = \sum_i (x_i - M\xi)^2 p_i.$$

Воспользовавшись свойствами математического ожидания, можно доказать, что эта формула эквивалентна следующей:

$$D\xi = M((\xi)^2) - (M\xi)^2,$$

где $M((\xi)^2) = \sum_i x_i^2 p_i$ – математическое ожидание ξ^2.

Дисперсия имеет размерность квадрата размерности случайной величины ξ, поэтому в качестве характеристики разброса значений удобнее использовать другую числовую характеристику — среднее квадратическое отклонение(сокращенно СКО).

Определение 2.1.6. *Средним квадратическим отклонением случайной величины (СКО) ξ называется квадратный корень из дисперсии:*

$$\sigma_\xi = \sqrt{D\xi}.$$

Математическое ожидание и дисперсия являются частными случаями более общих понятий — начальных и центральных моментов случайной величины порядка k.

Определение 2.1.7. *Начальным моментом порядка k случайной величины ξ называется математическое ожидание k-ой степени этой величины, обозначается α_k: $\alpha_k = M(\xi^k)$.*

Для дискретной случайной величины начальный момент равен сумме $\alpha_k = \sum\limits_i x_i^k p_i$. В частности, $\alpha_1 = M\xi$, (математическое ожидание равно начальному моменту 1-го порядка).

Определение 2.1.8. *Центральным моментом порядка k случайной величины ξ называется математическое ожидание $(\xi - M\xi)^k$, обозначается μ_k: $\mu_k = M((\xi - M\xi)^k)$.*

Для дискретной случайной величины: $\mu_k = \sum_i (x_i - M\xi)^k p_i$. В частности, $\mu_1 = 0$, а $\mu_2 = D\xi$, (центральный момент второго порядка – дисперсия).

Пример 2.1.3. Дискретная случайная величина ξ задана своим рядом распределения

x_i	-1	0	2	3
p_i	0,1	0,2	?	0,3

Найти $P(\xi = 2)$, функцию распределения, математическое ожидание, дисперсию.

Решение. Из условия $\sum_i p_i = 1$ найдем $P(\xi = 2) = 1 - (0,1 + 0,2 + 0,3) = 1 - 0,6 = 0,4$. Функция распределения равна

$$F_\xi(x) = \begin{cases} 0, & \text{если } x \le -1; \\ 0{,}1, & \text{если } -1 < x \le 0; \\ 0{,}1{+}0{,}2{=}0{,}3, & \text{если } 0 < x \le 2; \\ 0{,}1{+}0{,}2{+}0{,}4{=}0{,}7, & \text{если } 2 < x \le 3; \\ 1, & \text{если } x > 4. \end{cases}$$

Математическое ожидание вычисляем по формуле

$$M\xi = \sum_i x_i p_i = -1 \cdot 0,1 + 0 \cdot 0,2 + 2 \cdot 0,4 + 3 \cdot 0,3 = 1,6,$$

дисперсию по формуле

$$D\xi = M((\xi)^2) - M^2\xi = (-1)^2 \cdot 0,1 + (0)^2 \cdot 0,2 + (2)^2 \cdot 0,4 + (3)^2 \cdot 0,3 - 1,6^2 = 1,84.$$

Приведем примеры распределений часто встречающихся дискретных случайных величин.

Пример 2.1.4. ξ — индикатор события A. Он равен 1, если событие A произошло, и равен 0, если событие A не произошло.

$$P(A) = p, \qquad P(\overline{A}) = 1 - p = q.$$

Его ряд распределения:

x_i	0	1
p_i	q	p

Функция распределения:

$$F_\xi(x) = \begin{cases} 0, & \text{при } x \le 0; \\ q, & \text{при } 0 < x \le 1; \\ 1, & \text{при } x > 1. \end{cases}$$

Математическое ожидание и дисперсия:

$$M\xi = 0 \cdot q + 1 \cdot p = p, \qquad D\xi = M(\xi^2) - (M\xi)^2 = p - p^2 = pq.$$

Пример 2.1.5. Биномиальное распределение $B(p, n)$. Случайная величина ξ, равная числу успехов в n испытаниях Бернулли, распределена по биномиальному закону $B(p, n)$, где p — вероятность успеха в одном испытании. Эта случайная величина может принять одно из значений $k = 0, 1, ..., n$. Вероятность принятия каждого из значений находится по формуле Бернулли

$$P(\xi = k) = P_n(k) = C_n^k p^k q^{n-k}.$$

Математическое ожидание найдем применив свойство сложения математических ожиданий к случайной величине из примера 2.1.4: $M\xi = np$. Дисперсия $D\xi = npq$.

Пример 2.1.6. Геометрическое распределение. Геометрическое распределение с параметром $p \in (0, 1)$ имеет случайная величина, которая принимает натуральные значения, а их вероятности задаются формулой

$$P(\xi = k) = P(k) = pq^{k-1}, \qquad q = 1 - p, \qquad k \in \mathbb{N}.$$

Математическое ожидание $M\xi = 1/p$, дисперсия $D\xi = q/p^2$.

По геометрическому закону распределена случайная величина ξ — номер испытания Бернулли, в котором наступил первый успех, если p вероятность успеха в одном испытании.

Пример 2.1.7. Распределение Пуассона. Случайная величина, распределенная по закону Пуассона с параметром $\lambda > 0$, может принимать значения $0, 1, 2, 3, ...$ Ряд распределения задается формулой

$$P(\xi = k) = P(k) = \frac{\lambda^k}{k!}e^{-\lambda}, \qquad k = 0, 1, 2, ...$$

Математическое ожидание $M\xi = \lambda$, дисперсия $D\xi = \lambda$.

Распределение Пуассона является предельным случаем биномиального распределения при $n \to \infty$ и $p \to 0$ так, что $np \to \lambda$ (закон редких событий). Применяется для описания числа радиоактивных распадов за время T, в теории информации и теории массового обслуживания при описании пуассоновского потока событий.

Задачи для самостоятельного решения

1. В урне 19 белых и 6 черных шаров. Наугад из урны вынули 3 шара. Случайная величина ξ — число черных шаров, среди них. Постройте ряд

распределения ξ.

2. Дискретная случайная величина ξ задана своим рядом распределения

x_i	-2	0	2	4
p_i	0,1	0,2	?	0,1

. Найдите функцию распределения, математическое ожидание, дисперсию, СКО, $P(\xi \geq 0)$.

3. Игральную кость бросают до тех пор, пока не выпадет 6 очков. Случайная величина ξ — число произведенных бросков. По какому закону распределена ξ? Найдите математическое ожидание, дисперсию, СКО, $P(\xi > 6)$.

4. Известно, что в среднем 5 % студентов носят очки. На курсе 100 студентов. Случайная величина ξ – число студентов курса, пользующихся очками. Предполагая, что ξ распределена по закону Пуассона, найдите $M\xi$, $D\xi$, $P(\xi < M\xi)$.

5. Дискретная случайная величина, принимающая 5 возможных значений, распределена по биномиальному закону. Найдите дисперсию этой случайной величины, если $M\xi = 1$.

Ответы к задачам для самостоятельного решения

1.

x_i	0	1	2	3
p_i	0,42	0,45	0,12	0,01

.

2.

$$F_\xi(x) = \begin{cases} 0, & \text{если } x \leq -2; \\ 0{,}1, & \text{если } -2 < x \leq 0; \\ 0{,}3, & \text{если } 0 < x \leq 2; \\ 0{,}9, & \text{если } 2 < x \leq 4; \\ 1, & \text{если } x > 4. \end{cases}$$

1,4; 2,44; 1,56; 0,9.

3. По геометрическому. 6; 30; 5,48; 0,333.

4. 5; 5; 0,440.

5. 0,75.

2.2. Абсолютно непрерывная случайная величина

Если множество элементарных событий Ω несчетно, то согласно аксиоматической теории Колмогорова событиями являются не все подмножества Ω а лишь некоторые из подмножеств. При построении вероятностного пространства указывается также сигма-алгебра событий, на которой и задается вероятностное распределение [1]. Рассмотрим частный случай $\Omega = \mathbb{R}$. В этом случае в качестве сигма-алгебры событий обычно выбирают борелевскую сигма-алгебру на прямой. Она состоит из всевозможных промежутков вида

$$(a,b), \quad [a,b), \quad (a,b], \quad [a,b], \qquad -\infty \leq a \leq b \leq +\infty$$

и объединений конечного либо счетного числа таких промежутков.

Если результату случайного эксперимента сопоставляется число, то как и в дискретном случае возникает понятие случайной величины.

Определение 2.2.1. *Случайная величина ξ называется абсолютно-непрерывной, если существует функция $f_\xi(x)$ — плотность распределения случайной величины, такая, что для любого борелевского множества $A \subset \mathbb{R}$*

$$P(\xi \in A) = \int\limits_A f_\xi(x)dx.$$

Свойства плотности распределения вероятности:

1. $f_\xi(x) \geq 0$ для любого $x \in \mathbb{R}$.

2. Условие нормировки $\int\limits_{-\infty}^{+\infty} f_\xi(x)dx = 1$.

В этом параграфе рассмотрим только такие случайные величины, для которых плотность распределения кусочно-непрерывна. Для абсолютно непрерывной случайной величины вероятность того, что случайная величина примет конкретное значение $x \in \mathbb{R}$ равна 0: $P(\xi = x) = 0$. Имеет смысл говорить только о вероятности попадания в интервал.

Определение 2.2.2. *Функцией распределения абсолютно непрерывной случайной величины ξ с кусочно-непрерывной плотностью называ-*

ется функция

$$F_\xi(x) = P(\xi < x) = \int\limits_{-\infty}^{x} f_\xi(t)dt, \qquad x \in \mathrm{R},$$

определенная на всей вещественной оси.

Зная функцию распределения можно найти вероятность попадания случайной величины в любой интервал вида $[a, b)$:

$$P(\xi \in [a, b)) = F_\xi(b) - F_\xi(a).$$

Если построить график плотности вероятности, площадь под графиком будет равна 1. Вероятность того, что случайная величина попадет в интервал $[a, b]$ равна площади криволинейной трапеции под графиком плотности в пределах промежутка $[a, b]$.

Пример 2.2.1. Плотность распределения абсолютно непрерывной случайной величины равна

$$f_\xi(x) = \begin{cases} 0, \quad x < -2, \\ 1/2, \quad -2 < x < -1, \\ -4x - 2, \quad -1 < x < -1/2, \\ 0, \quad x > -1/2. \end{cases}$$

Найти функцию распределения.

Решение. Воспользуемся определением функции распределения и свойством аддитивности интеграла. Если $x \in (-2, -1)$, то

$$F_\xi(x) = \int\limits_{-\infty}^{x} f_\xi(t)dt = \int\limits_{-2}^{x} \frac{1}{2}dt = 1 + \frac{x}{2}.$$

Если $x \in (-1, -1/2)$, то

$$F_\xi(x) = \int\limits_{-\infty}^{x} f_\xi(t)dt = \int\limits_{-2}^{-1} \frac{1}{2}dt + \int\limits_{-1}^{x} (-4t - 2)dt = \frac{1}{2} - 2x^2 - 2x.$$

Следовательно,

$$F_\xi(x) = \begin{cases} 0, \quad x \le -2, \\ 1/2x + 1, \quad -2 < x \le -1, \\ -2x^2 - 2x + 1/2, \quad -1 < x \le -1/2, \\ 1, \quad x > -1/2. \end{cases}$$

Пример 2.2.2. Функция распределения абсолютно-непрерывной случайной величины равна

$$F_\xi(x) = \begin{cases} 0, \quad x \le 0, \\ \dfrac{x^2}{2}, \quad 0 < x \le 1, \\ -\dfrac{1}{3}x^2 + \dfrac{5}{3}x - \dfrac{13}{12}, \quad 1 < x \le 5/2, \\ 1, \quad x > 5/2. \end{cases}$$

Найдите плотность распределения.

Решение. Воспользовавшись теоремой Барроу о дифференцировании интеграла с переменным верхним пределом, находим

$$f_\xi(x) = \begin{cases} 0, \quad x < 0, \\ x, \quad 0 < x < 1, \\ -\dfrac{2}{3}x + \dfrac{5}{3}, \quad 1 < x < 5/2, \\ 0, \quad x > 5/2. \end{cases}$$

В плотности вероятности распределения или в функции распределения содержится полная информация об абсолютно непрерывной случайной величине. Если же достаточно кратких сведений о случайной величине, то описываем её числовыми характеристиками.

Определение 2.2.3. *Математическим ожиданием абсолютно-непре-*

рывной случайной величины ξ называется интеграл:

$$M\xi = \int\limits_{-\infty}^{+\infty} x f_{\xi}(x) dx.$$

Математическое ожидание существует, если этот интеграл сходится абсолютно. В противном случае говорят, что случайная величина ξ не имеет математического ожидания.

Пример 2.2.3. Плотность распределения абсолютно-непрерывной СВ равна
$f_{\xi}(x) = \begin{cases} 0, & x < 4, \\ C\dfrac{1}{x^{3/2}}, & x > 4. \end{cases}$ Найдите постоянную C и математическое ожидание.

Решение. По условию нормировки для плотности распределения абсолютно непрерывной случайной величины, имеем

$$C \int\limits_{4}^{+\infty} \frac{1}{x^{3/2}} dx = 1,$$

откуда $C = 1$. Так как интеграл

$$M\xi = \int\limits_{4}^{+\infty} x \frac{1}{x^{3/2}} dx = \int\limits_{4}^{+\infty} \frac{1}{x^{1/2}} dx$$

расходится, случайная величина ξ не имеет математического ожидания.

Определение 2.2.4. *Дисперсией $D\xi$ абсолютно-непрерывной случайной величины ξ называется математическое ожидание квадрата ее отклонения от математического ожидания:*

$$D\xi = M((\xi - M\xi)^2) = \int\limits_{-\infty}^{+\infty} (x - M\xi)^2 f_{\xi}(x) dx.$$

По свойствам математического ожидания ([1]), эта формула эквивалентна следующей:

$$D\xi = M(\xi^2) - (M\xi)^2,$$

где

$$M((\xi)^2) = \int\limits_{-\infty}^{+\infty} x^2 f_\xi(x)dx.$$

Дисперсия существует, если этот интеграл сходится абсолютно. В противном случае говорят, что случайная величина ξ не имеет дисперсии.

Определение 2.2.5. *Для абсолютно непрерывной случайной величины начальный момент порядка k равен*

$$\alpha_k = M(\xi^k) = \int\limits_{-\infty}^{+\infty} x^k f_\xi(x)dx,$$

центральный момент порядка k равен

$$\mu_k = \int\limits_{-\infty}^{+\infty} (x - M\xi)^k f_\xi(x)dx.$$

Важным примером распределения абсолютно непрерывной случайной величины является **нормальное распределение** $N(a, \sigma)$, т.е. распределение с плотностью

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x-a)^2}{2\sigma^2}}.$$

Математическое ожидание $M\xi = a$; дисперсия $D\xi = \sigma^2$.

Нормальное распределение с параметрами $a = 0$, $\sigma = 1$ называется **стандартным нормальным распределением**.

Функция распределения стандартного нормального распределения — **функция Лапласа**

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\dfrac{t^2}{2}} dt.$$

Нормальное распределение занимает в теории вероятностей и статистике центральное место, так как сумма большого числа случайных величин при определенных условиях имеет распределение, близкое к нормальному. Применяется для описания ошибок измерений, параметров изделий при массовом производстве. Это предельный случай биномиального распределения при $n \to \infty$. Поэтому составлены подробные таблицы значений функции Лапласа.

Пример 2.2.4. Плотность распределения абсолютно непрерывной СВ равна

$$f_\xi(x) = \begin{cases} a, & x \in (-3/5, 0), \\ 2/5x + 1/5, & x \in (0, a), \\ 0, & x \notin (-3/5, a). \end{cases}$$

Найти параметр a, математическое ожидание и дисперсию случайной величины $\eta = 3\xi - 5$.

Решение. Чтобы найти параметр a, воспользуемся свойством нормировки плотности $\int\limits_{-\infty}^{+\infty} f_\xi(x)dx = 1$. Следовательно,

$$\int\limits_{-3/5}^{0} adx + \int\limits_{0}^{a}(2/5x + 1/5)dx = 1$$

и параметр a – корень квадратного уравнения $a^2+4a-5 = 0$. Это уравнение имеет два корня $a = 1$ и $a = -5$. Поскольку плотность вероятности не может быть отрицательной, параметр $a = 1$.

Найдем математическое ожидание случайной величины ξ:

$$M\xi = \int\limits_{-\infty}^{+\infty} xf_\xi(x)dx = \int\limits_{-3/5}^{0} xdx + \int\limits_{0}^{1}(2/5x + 1/5)xdx$$

$$= \frac{x^2}{2}\Big|_{-3/5}^{0} + \frac{2}{15}x^3\Big|_0^1 + \frac{1}{10}x^2\Big|_0^1 = -\frac{9}{50} + \frac{2}{15} + \frac{1}{10} = \frac{4}{75}.$$

По свойству линейности математического ожидания

$$M\eta = M(3\xi - 5) = 3M\xi - 5 = \frac{4}{25} - 5 = -4,84.$$

Найдем математическое ожидание случайной величины ξ^2:

$$M(\xi^2) = \int\limits_{-\infty}^{+\infty} x^2 f_\xi(x)dx = \int\limits_{-3/5}^{0} x^2 dx + \int\limits_{0}^{1}(2/5x + 1/5)x^2 dx$$

$$= \frac{x^3}{3}\Big|_{-3/5}^{0} + \frac{2}{20}x^4\Big|_0^1 + \frac{1}{15}x^3\Big|_0^1 = \frac{9}{125} + \frac{1}{10} + \frac{1}{15} = \frac{179}{750}.$$

Найдем дисперсию ξ по формуле $D\xi = M(\xi^2) - (M\xi)^2$:

$$D\xi = \frac{179}{750} - \left(\frac{4}{75}\right)^2 = \frac{2653}{11250} \approx 0.2358.$$

Воспользовавшись свойствами дисперсии, находим

$$D\eta = D(3\xi - 5) = 3^2 D\xi = 2,1224.$$

Задачи для самостоятельного решения

1. Плотность распределения абсолютно-непрерывной случайной величины ξ равна $f_\xi(x) = \begin{cases} 0, & x < 0, \\ 2x, & 0 \le x \le a, \\ 0, & x > a. \end{cases}$
Найдите постоянную a и $P(\xi \in (-2, 1/2))$.

2. В кадр попали часы. Случайная величина ξ — положение минутной стрелки. По какому закону распределена ξ? Найдите $M\xi$, $D\xi$, σ_ξ, $P(10 \le \xi \le 25)$.

3. Непрерывная случайная величина задана функцией распределения вероятностей $F_\xi(x) = \begin{cases} 0, & x \le -2, \\ Ax + B, & -2 < x < 3, \\ 1, & x \ge 3. \end{cases}$
Найдите A, B, плотность вероятности, $M\xi$, $D\xi$, σ_ξ.

4. Плотность распределения случайной величины ξ равна

$$f_\xi(x) = \frac{1}{\sqrt{6\pi}} e^{-\dfrac{(x-2)^2}{6}}.$$

Случайная величина $\eta = -3\xi + 7$. Найдите $M\eta$, $D\eta$.

Ответы к задачам для самостоятельного решения

1. $a = 1$, $P(\xi \in (-2, 1/2)) = \dfrac{1}{4}$.

2. Равномерное распределение на $[0; 60]$. 30; 300; 17,32; 0,25.

3. $A = 0,2$; $B = 0,4$; $M\xi = 0,5$; $D\xi = 2,08$; $\sigma_\xi = 1,44$. Равномерное распределение на $[-2; 3]$.

4. 1; 27.

2.3. Дискретный случайный вектор

Ограничимся рассмотрением двумерного случайного вектора. В дискретном случае распределение двумерного случайного вектора — совместное распределение двух дискретных случайных величин, являющихся координатами (компонентами) случайного вектора задается с помощью таблицы:

$\xi \setminus \eta$	b_1	b_2	...	b_m	...
a_1	p_{11}	p_{12}	...	p_{1m}	...
a_2	p_{21}	p_{22}	...	p_{2m}	...
.	.	.	...	.	...
a_k	p_{k1}	p_{k2}	...	p_{km}	...
.	.	.	...	.	...

Каждая из строк этой таблицы соответствует одному значению первой компоненты вектора $\xi = a_i$, а каждый из столбцов — одному значению второй компоненты $\eta = b_j$. В таблице указываются вероятности

$$p_{ij} = P(\xi = a_i; \quad \eta = b_j).$$

Зная таблицу распределения дискретного случайного вектора, можно найти распределение каждой из координат, суммируя вероятности появления всех пар, в которых данная координата не меняется.

Определение 2.3.1. *Случайные величины ξ и η (координаты случайного вектора) называются независимыми, если*

$$P\left(\xi = a_i; \quad \eta = b_j\right) = P\left(\xi = a_i\right) \cdot P\left(\eta = b_j\right)$$

при всех возможных значениях индексов.

Очевидно, что для случайного вектора с независимыми координатами можно по распределениям координат восстановить таблицу распределения случайного вектора. Если координаты случайного вектора не являются независимыми, это невозможно.

Математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий:

$$M(\xi\eta) = M\xi \cdot M\eta.$$

Дисперсия суммы независимых случайных величин равна сумме их дисперсий: $D(\xi + \eta) = D\xi + D\eta$.

Пример 2.3.1. Дана таблица распределения дискретного случайного вектора

$\xi \setminus \eta$	2	3	4	7
0	1/16	3/16	1/8	1/16
1	1/16	1/16	1/16	1/8
2	1/16	1/16	1/16	1/16

Являются ли его координаты независимыми?

Решение. Найдем ряд распределения каждой из координат случайного вектора

ξ:	0	1	2
$p(\xi = (\cdot))$	7/16	5/16	1/4

η :	2	3	4	7
$p(\eta = (\cdot))$	3/16	5/16	1/4	1/4

Возьмем в таблице значение

$$p\,(\xi = 0; \eta = 2) = \frac{1}{16}.$$

Найдем вероятности принятия соответствующих значений координат

$$p\,(\xi = 0) = \frac{7}{16}, \qquad p\,(\eta = 2) = \frac{3}{16}.$$

Поскольку $\frac{7}{16} \cdot \frac{3}{16}$ не совпадает с $\frac{1}{16}$, то координаты случайного вектора зависимы.

Числовыми характеристиками связи между случайными величинами также являются корреляционный момент и коэффициент корреляции.

Определение 2.3.2. *Корреляционным моментом или ковариацией двух случайных величин* ξ *и* η *называется математическое ожидание произведения отклонений этих случайных величин от их математических ожиданий:*

$$cov(\xi, \eta) = M\left((\xi - M\xi)(\eta - M\eta)\right).$$

Воспользовавшись свойствами математического ожидания, получим эквивалентную формулу для вычисления ковариации:

$$cov(\xi, \eta) = M\left(\xi\eta\right) - M\xi \cdot M\eta.$$

Размерность равна произведению размерностей случайных величин, поэтому в качестве числовой характеристики зависимости удобнее иметь дело с безразмерной величиной, а именно коэффициентом корреляции.

Определение 2.3.3. *Коэффициентом корреляции двух случайных величин* ξ *и* η *называется отношение их ковариации (корреляционного момента) к произведению их средних квадратических отклонений:*

$$\rho_{\xi,\eta} = \frac{cov(\xi, \eta)}{\sqrt{D\xi}\sqrt{D\eta}} = \frac{cov(\xi, \eta)}{\sigma_\xi \sigma_\eta}.$$

Свойства коэффициента корреляции:

1. Коэффициент корреляции по абсолютной величине не превосходит 1, т. е. $|\rho_{\xi,\eta}| \le 1$ или $-1 \le \rho_{\xi,\eta} \le 1$.

2. Если ξ и η независимы, то $\rho_{\xi,\eta} = 0$.

3. Если ξ и η связаны линейной зависимостью, т. е. $\eta = a\xi + b$, то $|\rho_{\xi,\eta}| = 1$, причем при $a > 0$ имеем $\rho_{\xi,\eta} = 1$ и при $a < 0$ имеем $\rho_{\xi,\eta} = -1$.

Заметим, что коэффициент корреляции характеризует меру именно линейной зависимости. Случайные величины могут быть сильно, но не линейно связаны при коэффициенте корреляциии, равном нулю.

Определение 2.3.4. *Случайные величины называются некоррелированными, если их ковариационный момент, а соответственно и коэффициент корреляции равен нулю.*

Если случайные величины независимы, то они и некоррелированные. Обратное неверно.

Пример 2.3.2. Дана таблица распределения дискретного случайного вектора

$\xi \setminus \eta$	-1	0	1
0	1/4	3/16	1/4
1	1/8	1/16	1/8

.

Найдите законы распределения ξ и η, корреляционный момент и коэффициент корреляции. Являются ли ξ и η независимыми случайными величинами?

Решение. Найдем закон распределения ξ. Для этого просуммируем вероятности в строках

$$P(\xi = 0) = 1/4 + 3/16 + 1/4 = 11/16.$$

$$P(\xi = 1) = 1/8 + 1/16 + 1/8 = 5/16.$$

Ряд распределения ξ имеет вид:

ξ_i	0	1
p_i	11/16	5/16

Вычислим математическое ожидание ξ:

$$M\xi = 0 \cdot 11/16 + 1 \cdot 5/16 = 5/16.$$

Аналогично найдем закон распределения η. Для этого просуммируем вероятности в столбцах

$$P(\eta = -1) = 1/4 + 1/8 = 3/8, \qquad P(\eta = 0) = 3/16 + 1/16 = 1/4,$$

$$P(\eta = 1) = 1/4 + 1/8 = 3/8.$$

Ряд распределения η имеет вид

η_i	-1	0	1
p_i	3/8	1/4	3/8

.

Вычислим математическое ожидание η:

$$M\eta = -1 \cdot 3/8 + 0 \cdot 1/4 + 1 \cdot 3/8 = 0.$$

Найдем корреляционный момент:

$$cov(\xi, \eta) = M\xi\eta - M\xi M\eta = 0 \cdot (-1) \cdot 1/4 + 0 \cdot 0 \cdot 3/16 + 0 \cdot 1 \cdot 1/4$$
$$+1 \cdot (-1) \cdot 1/8 + 1 \cdot 0 \cdot 1/16 + 1 \cdot 1 \cdot 1/8 - 5/16 \cdot 0 = 1/8 - 1/8 - 0 = 0.$$

Коэффициент корреляции тоже равен нулю и, следовательно, компоненты случайного вектора ξ и η некоррелированы. Проверим, являются ли

ξ и η независимыми:

$$P(\xi = 0; \eta = -1) = 1/4,$$

$$P(\xi = 0) \cdot P(\eta = -1) = 11/16 \cdot 3/8 = 33/128 \neq 1/4,$$

значит ξ и η зависимые, но некоррелированые.

Задачи для самостоятельного решения

1. Дана таблица распределения дискретного случайного вектора

$\xi \setminus \eta$	-2	-1	1
2	$0,3$	$0,1$	$0,05$
4	$0,15$	$0,15$	$0,25$

Являются ли его координаты независимыми? Найдите коэффициент корреляции.

2. Дана таблица распределения дискретного случайного вектора.

$\xi \setminus \eta$	-1	1	2
-3	$1/16$	$1/64$	$3/64$
-2	$1/8$	$1/32$	$3/32$
1	$5/16$	$5/64$	$15/64$

Являются ли его координаты независимыми? Найдите коэффициент корреляции.

Ответы к задачам для самостоятельного решения

1. Зависимы, $0,4215$.

2. Независимы, 0.

2.4. Случайный вектор с абсолютно-непрерывным распределением

Рассмотрим только случай двумерного абсолютно непрерывного случайного вектора с компонентами (ξ, η).

Определение 2.4.1. *Случайный вектор (ξ, η) называется абсолютно непрерывным, если существует такая функция $f_{\xi,\eta}(x, y)$, называемая совместной плотностью распределения вероятности (ξ, η) или плотностью распределения случайного вектора, что для любого борелевского множества [1] $A \subset \mathbb{R}^2$ (в частности для любой области)*

$$P((\xi, \eta) \in A) = \iint\limits_A f_{\xi,\eta}(x, y)dxdy.$$

Плотность распределения случайного вектора обладает следующими свойствами:

1. $f_{\xi,\eta}(x, y) \geq 0$
2. Условие нормировки:

$$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f_{\xi,\eta}(x, y)dxdy = 1$$

Зная плотность распределения случайного вектора, можно определить плотность распределения вероятности каждой из компонент:

$$f_\xi(x) = \int\limits_{-\infty}^{+\infty} f_{\xi,\eta}(x,y)dy, \qquad f_\eta(y) = \int\limits_{-\infty}^{+\infty} f_{\xi,\eta}(x,y)dx$$

и функции распределения вероятности компонент вектора ξ и η:

$$F_\xi(x) = \int\limits_{-\infty}^{x} \left(\int\limits_{-\infty}^{+\infty} f_{\xi,\eta}(t,y)dy \right) dt, \qquad F_\eta(y) = \int\limits_{-\infty}^{y} \left(\int\limits_{-\infty}^{+\infty} f_{\xi,\eta}(x,s)dx \right) ds.$$

Определение 2.4.2. *Случайные величины ξ и η называются независимыми, если для любых x, y события $\xi < x$ и $\eta < y$ являются независимыми, т. е.*

$$P(\xi < x, \eta < y) = P(\xi < x) \cdot P(\eta < y).$$

Чтобы случайные величины ξ и η были независимы, необходимо и достаточно, чтобы в точках непрерывности функция плотности распределения вероятности вектора (ξ, η) была равна произведению функций плотности распределения вероятности ξ и η.

$$f_{\xi,\eta}(x,y) = f_\xi(x) \cdot f_\eta(y).$$

Пример 2.4.1. Дана плотность распределения вероятности двумерной случайной величины

$$f_{\xi,\eta}(x,y) = \begin{cases} C(x^2+y^2), & 0 \le x \le 1, 0 \le y \le 1; \\ 0, & (x,y) \notin [0,1] \times [0,1]. \end{cases}$$

Найти C, $f_\xi(x)$, $f_\eta(y)$, $M\xi$, $M\eta$, σ_ξ, σ_η. Выяснить, зависимы ли ξ и η.

Решение. Коэффициент C находим из условия

$$\int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} f_{\xi,\eta}(x,y)dxdy = 1.$$

$$C\int_0^1 dx\int_0^1 (x^2+y^2)dy = C\int_0^1 dx\ (x^2y+\frac{y^3}{3})\Big|_0^1 =$$

$$= C\int_0^1 (x^2+\frac{1}{3})dx = C\ (\frac{x^3}{3}+\frac{x}{3})\Big|_0^1 = \frac{2}{3}C = 1.$$

Получаем $C = \dfrac{3}{2}$.

Для вычисления числовых характеристик компонент случайного вектора найдем их плотности вероятности. При $x \in [0;1]$

$$f_\xi(x) = \int_{-\infty}^{+\infty} f_{\xi,\eta}(x,y)dy = \frac{3}{2}\int_0^1 (x^2+y^2)dy = \frac{3}{2}(x^2y+\frac{y^3}{3})\Big|_0^1 = \frac{3x^2}{2}+\frac{1}{2}$$

и $f_\xi(x) = 0$ при $x \notin [0;1]$. Из симметричности формулы и области следует, что $f_\eta(y) = \dfrac{3y^2}{2}+\dfrac{1}{2}$ при $y \in [0;1]$ и $f_\eta(y) = 0$ при $y \notin [0;1]$ Зная плотности вероятности ξ и η можно найти математические ожидания:

$$M\xi = M\eta = \int_{-\infty}^{+\infty} x f_\xi(x)dx = \frac{3}{2}\int_0^1 x(x^2+\frac{1}{3})dx = \frac{3}{2}(\frac{x^4}{4}+\frac{x^2}{6})\Big|_0^1 = \frac{3}{2}(\frac{1}{4}+\frac{1}{6}) = \frac{5}{8}.$$

Вычислим $M(\xi^2) = M(\eta^2)$:

$$M(\xi^2) = \int_{-\infty}^{+\infty} x^2 f_\xi(x)dx = \frac{3}{2}\int_0^1 x^2(x^2+\frac{1}{3})dx = \frac{3}{2}(\frac{x^5}{5}+\frac{x^3}{9})\Big|_0^1 = \frac{3}{2}(\frac{1}{5}+\frac{1}{9}) = \frac{7}{15}.$$

Найдем дисперсию по формуле $D\xi = M(\xi^2) - (M\xi)^2$:

$$D\eta = D\xi = \frac{7}{15} - (\frac{5}{8})^2 = \frac{73}{960} \approx 0,076.$$

Средне квадратические отклонения:

$$\sigma_\xi = \sigma_\eta = \sqrt{D\xi} \approx \sqrt{0,076} \approx 0,276.$$

Случайные величины ξ и η не являются независимыми, так как плотность распределения вероятности вектора (ξ,η) не равна произведению плотностей распределения вероятности ξ и η. Например, при $0 \le x \le 1$, $0 \le y \le 1$

$$\frac{3}{2}(x^2+y^2) \neq (\frac{3x^2}{2}+\frac{1}{2})(\frac{3y^2}{2}+\frac{1}{2}).$$

Определение 2.4.3. *Для непрерывных случайных величин ковариация равна:*

$$cov(\xi,\eta) = M((\xi-M\xi)(\eta-M\eta)) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}(x-M\xi)(y-M\eta)f_{\xi,\eta}(x,y)dxdy.$$

Можно также использовать формулу

$$cov(\xi,\eta) = M(\xi\eta) - M\xi\cdot M\eta = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xyf_{\xi,\eta}(x,y)dxdy - M\xi\cdot M\eta.$$

Пример 2.4.2. Для случайных величин ξ и η из предыдущего примера найти ковариацию.

Решение. Найдем $M(\xi\eta)$:

$$M(\xi\eta) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xyf_{\xi,\eta}(x,y)dxdy = \frac{3}{2}\int\limits_{0}^{1}\int\limits_{0}^{1} xy(x^2+y^2)dxdy =$$

$$= \frac{3}{2}\int\limits_{0}^{1} dx\ (x^3\frac{y^2}{2}+x\frac{y^4}{4})\Big|_0^1 = \frac{3}{2}\int\limits_{0}^{1}(x^3\frac{1}{2}+x\frac{1}{4})dx = \frac{3}{2}\ (\frac{x^4}{8}+\frac{x^2}{8})\Big|_0^1 = \frac{3}{2}(\frac{1}{8}+\frac{1}{8}) = \frac{3}{8}$$

и ковариацию

$$cov(\xi,\eta) = \frac{3}{8}-\frac{25}{64} = -\frac{1}{64}.$$

Пример 2.4.3. Дана плотность распределения вероятности двумерной случайной величины

$$f_{\xi,\eta}(x,y) = \begin{cases} C, & (x,y)\in D; \\ 0, & (x,y)\notin D \end{cases}$$

где D—четверть круга $x^2 + y^2 \leq 1$, в которой $x \geq 0$, $y \geq 0$.

Найти $f_\xi(x)$, $f_\eta(y)$, $M\xi$, $M\eta$, σ_ξ, σ_η.

Выяснить, зависимы ли ξ и η.

Решение. Из условия нормировки $C = \dfrac{1}{S_D} = \dfrac{4}{\pi}$. При $x \in [0;1]$

$$f_\xi(x) = \int\limits_{-\infty}^{+\infty} f_{\xi,\eta}(x,y)dy = \int\limits_{0}^{\sqrt{1-x^2}} \frac{4}{\pi}dy = \frac{4}{\pi}\sqrt{1-x^2},$$

при $x \notin [0;1]$ $f_\xi(x) = 0$.

Аналогично, при $y \in [0;1]$, $f_\eta(y) = \dfrac{4}{\pi}\sqrt{1-y^2}$, при $y \notin [0;1]$, $f_\eta(y) = 0$.

$$M\xi = \int\limits_{-\infty}^{+\infty} x f_\xi(x)dx = \int\limits_{0}^{1} \frac{4}{\pi}x\sqrt{1-x^2}dx$$

$$= -\frac{2}{\pi}\int\limits_{0}^{1} \sqrt{1-x^2}d(1-x^2) = -\frac{4}{3\pi}\left(1-x^2\right)^{\frac{3}{2}}\Big|_0^1 = \frac{4}{3\pi}.$$

Аналогично $M\eta = \dfrac{4}{3\pi}$.

$$M(\xi^2) = M(\eta^2) = \int\limits_{-\infty}^{+\infty} x^2 f_\xi(x)dx = \frac{4}{\pi}\int\limits_{0}^{1} x^2\sqrt{1-x^2}dx =$$

$$= \frac{4}{\pi}\int\limits_{0}^{\pi/2} sin^2 t cos^2 t dt = \frac{1}{\pi}\int\limits_{0}^{\pi/2} sin^2(2t)dt = \frac{1}{2\pi}\int\limits_{0}^{\pi/2} (1-\cos(4t))dt =$$

$$= \frac{1}{2\pi}(t - \frac{1}{4}\sin(4t))\Big|_0^{\frac{\pi}{2}} = \frac{1}{2\pi}\cdot\frac{\pi}{2} = \frac{1}{4};$$

$$D\xi = D\eta = M(\xi^2) - (M\xi)^2 = \frac{1}{4} - (\frac{4}{3\pi})^2 \approx 0,070;$$

$$\sigma_\xi = \sigma_\eta = \sqrt{D\xi} \approx 0,264.$$

Найдем ковариацию:

$$M(\xi\eta) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xyf_{\xi,\eta}(x,y)dxdy = \frac{4}{\pi}\iint_D xydxdy =$$

$$= \frac{4}{\pi}\int_0^1 xdx\int_0^{\sqrt{1-x^2}} ydy = \frac{4}{\pi}\int_0^1 x\Big(\frac{y^2}{2}\Big|_0^{\sqrt{1-x^2}}\Big)dx = \frac{2}{\pi}\int_0^1 x(1-x^2)dx =$$

$$= \frac{2}{\pi}\left(\frac{x^2}{2}-\frac{x^4}{4}\right)\Big|_0^1 = \frac{2}{\pi}\Big(\frac{1}{2}-\frac{1}{4}\Big) = \frac{2}{\pi}\frac{1}{4} = \frac{1}{2\pi} \approx 0,160;$$

$$cov(\xi,\eta) = \frac{1}{2\pi} - \left(\frac{4}{3\pi}\right)^2 \approx -0,021;$$

Задачи для самостоятельного решения

1. Дана плотность распределения вероятности двумерной случайной величины

$$f_{\xi,\eta}(x,y) = \begin{cases} Cx^2y^2, & (x,y)\in[0;1]\times[0;1], \\ 0, & (x,y)\notin[0;1]\times[0;1], \end{cases}$$

Найти C, $M\xi$, $M\eta$, σ_ξ, σ_η, $\rho_{\xi,\eta}$. Выяснить, зависимы ли ξ и η.

2. Плотность распределения двумерной случайной величины

$$f_{\xi,\eta}(x,y) = \begin{cases} C(x+y), & (x,y)\in[0;1]\times[0;1], \\ 0, & (x,y)\notin[0;1]\times[0;1], \end{cases}$$

Найти C, $M\xi$, $M\eta$, σ_ξ, σ_η, $\rho_{\xi,\eta}$. Выяснить, зависимы ли ξ и η.

3. Случайный вектор распределен равномерно в D — полукруге: $x^2 + y^2 \le 1$, $y \ge 0$. Найти $f_{\xi,\eta}(x,y)$, $f_\xi(x)$, $f_\eta(y)$, $M\xi$, $M\eta$, σ_ξ, σ_η, $\rho_{\xi,\eta}$. Выяснить, зависимы ли ξ и η.

4. Плотность распределения двумерной случайной величины

$$f_{\xi,\eta}(x,y) = \frac{1}{12\pi}exp\Big(-\frac{1}{2}\Big(\frac{(x-3)^2}{4}+\frac{(y-5)^2}{9}\Big)\Big)$$

Найти $M\xi$, $M\eta$, σ_ξ, σ_η, $\rho_{\xi,\eta}$. Выяснить, зависимы ли ξ и η.

5. Плотность распределения двумерной случайной величины

$$f_{\xi,\eta}(x,y) = \frac{1}{\pi\sqrt{0,75}} exp\Big(-\frac{2}{3}\Big(x^2 - 2x(y-2) + 4(y-2)^2\Big)\Big)$$

Найти $M\xi$, $M\eta$, σ_ξ, σ_η, $\rho_{\xi,\eta}$. Выяснить, зависимы ли ξ и η.

Ответы к задачам для самостоятельного решения

1. $C = 9$; $M\xi = M\eta = \frac{3}{4}$; $\sigma_\xi = \sigma_\eta = \frac{\sqrt{3}}{4\sqrt{5}} \approx 0,194$; $\rho_{\xi,\eta} = 0$. Независимы.

2. $C = 1$; $M\xi = M\eta = \frac{7}{12} \approx 0,583$; $\sigma_\xi = \sigma_\eta = \frac{\sqrt{11}}{12} \approx 0,276$; $\rho_{\xi,\eta} = -\frac{1}{11} \approx -0,091$. Зависимы.

3. $M\xi = 0$; $M\eta = \frac{4}{3\pi} \approx 0,424$, $\sigma_\xi = 0,5$; $\sigma_\eta \approx 0,264$; $\rho_{\xi,\eta} = 0$. Зависимы.

4. $M\xi = 3$; $M\eta = 5$; $\sigma_\xi = 2$; $\sigma_\eta = 3$; $\rho_{\xi,\eta} = 0$. Независимы.

5. $M\xi = 0$; $M\eta = 2$; $\sigma_\xi = 1$; $\sigma_\eta = 0,5$; $\rho_{\xi,\eta} = 0,5$. Зависимы.

2.5. Неравенство Чебышева

Теорема 2.5.1. *Пусть случайная величина ξ имеет математическое ожидание $M\xi$ и дисперсию $D\xi$. Для любого $\varepsilon > 0$ справедливо* **неравенство Чебышева***:*

$$P(|\xi - M\xi| \geq \varepsilon) \leq \frac{D\xi}{\varepsilon^2}.$$

Это неравенство определяет верхнюю границу вероятности того, что отклонение значения случайной величины ξ от ее математического ожидания не меньше некоторого заданного числа ε. Примечательно, что такая оценка дается для случайной величины, распределение которой неизвестно, известны лишь ее математическое ожидание и дисперсия.
На практике часто используется следствие из этого неравенства:

Следствие 1.

$$P(|\xi - M\xi| < \varepsilon) \geq 1 - \frac{D\xi}{\varepsilon^2}.$$

Пример 2.5.1. В 350 испытаниях Бернулли вероятность успеха в каждом испытании равна 0,7. С помощью неравенства Чебышева оцените вероятность того, что разница между числом успехов в этих испытаниях и средним числом успехов не превысит 10.

Решение. Известно, что число успехов распределено по закону Бернулли, поэтому математическое ожидание и дисперсия числа успехов равны

$$M\xi = np = 350 \times 0,7 = 245,$$

$$D\xi = npq = 350 \times 0,7 \times 0,3 = 73,5.$$

Применяя неравенство Чебышева, получим

$$P(|\xi - 245| < 10) \geq 1 - \frac{73,5}{10^2} = 1 - \frac{73,5}{100} = 0,265.$$

Пример 2.5.2. Для освещения придомовой территории многоквартирного дома 20 ламп накаливания были заменены на такое же количество светодиодных светильников с датчиками движения. Вероятность того, что в течение 10 лет светильник не выйдет из строя, равна 0,8. Пользуясь неравенством Чебышева, оцените вероятность того, что за 10 лет абсолютная величина разности между числом светильников, не вышедших из строя, и математическим ожиданием числа светильников, проработавших безотказно, окажется а) меньше 3; б) не меньше 3.

Решение. Обозначим через ξ дискретную случайную величину — число светильников, не вышедших из строя за 10 лет. Тогда

$$M\xi = np = 20 \cdot 0,8 = 16,$$

$$D\xi = npq = 20 \cdot 0,8 \cdot 0,2 = 3,2.$$

Воспользуемся неравенством Чебышева:

$$P(|\xi - 16| < 3) \geq 1 - \frac{3,2}{3^2} = 1 - \frac{3,2}{9} \approx 0,64.$$

$$P(|\xi - 16| \geq 3) \leq \frac{3,2}{3^2} \approx 0,36.$$

Пример 2.5.3. Известны числовые характеристики случайной величины ξ: $M\xi = 0,4$; $D\xi = 0,01$. Оцените снизу вероятности следующих событий: $A = \{0,2 < \xi < 0,6\}$, $B = \{\xi < 0,9\}$.

Решение.

$$P(A) = P(0,2 < \xi < 0,6) = P(-0,2 < \xi - 0,4 < 0,2) =$$

$$= P(|\xi - 0,4| < 0,2) \geq 1 - \frac{0,01}{0,2^2} = 0,75.$$

$$P(B) = P(\xi < 0,9) = P(-\infty < \xi < 0,9) \geq P(-0,1 < \xi < 0,9) =$$

$$= P(0,4 - 0,5 < \xi < 0,4 + 0,5) = P(|\xi - 0,4| < 0,5) \geq 1 - \frac{0,01}{0,5^2} = 0,96.$$

Следовательно

$$P(A) \geq 0,75, \qquad P(B) \geq 0,96.$$

Задачи для самостоятельного решения

1. Для подсветки экспонатов музея используются 400 лампочек, работающих независимо друг от друга. Вероятность того, что какая-нибудь из них перестанет работать при резком перепаде напряжения в сети, равна 0,5. С помощью неравенства Чебышева оцените вероятность того, что число испортившихся от перепада напряжения лампочек будет заключено между 100 и 300.

2. По статистическим данным вероятность того, что человек доживет до 50 лет равна 0,87. С помощью неравенства Чебышева оценить вероятность того, что из 1 000 новорождённых доля (относительная частота) доживших до 50 лет будет отличаться от вероятности дожития не более, чем на 0,04 (по модулю).

3. Известно, что ξ — нормально распределенная случайная величина, у которой математическое ожидание равно 0, а среднеквадратическое отклонение равно 1. Используя неравенство Чебышева, оцените снизу вероятность события $A = \{-3 < \xi < 3\}$.

Ответы к задачам для самостоятельного решения

1. $0,99$; 2. $0,93$; 3. $\dfrac{8}{9}$.

2.6. Предельные теоремы

В теории вероятностей и ее приложениях часто рассматриваются последовательности событий, повторяющихся в неизменных условиях. Поэтому для приложений большую роль играют предельные теоремы.

Теорема 2.6.1. (Закон больших чисел). *Пусть задана последовательность независимых случайных величин* $\xi_1, \xi_2, \dots, \xi_n, \dots$ *с одинаковыми математическими ожиданиями* $M\xi_i = a$ *и равномерно ограниченными дисперсиями* $D\xi_i \leq C$*. Тогда последовательность средних арифметических первых* n *величин сходится по вероятности к математическому ожиданию* a*, т.е. для любого* $\varepsilon > 0$ *справедливо равенство*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - a\right| \geq \varepsilon\right) = 0.$$

Суть закона больших чисел заключается в том, что при возрастании числа слагаемых случайных величин с одинаковыми математическими ожиданиями среднее арифметическое этих слагаемых мало отличается от математического ожидания a.

Теорема 2.6.2. (Центральная предельная теорема). *Задана бесконечная последовательность независимых, одинаково распределенных случайных величин* $\xi_1, \xi_2, \dots, \xi_n, \dots$ *с математическими ожиданиями* $M\xi_i = a$ *и конечными дисперсиями* $D\xi_i = \sigma^2, i = 1, 2, \dots, n, \dots$*. Для любого вещественного* x *справедливо*

$$\lim_{n\to\infty} P\left(\frac{\xi_1 + \xi_2 + \ldots + \xi_n - na}{\sigma\sqrt{n}} < x\right) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{\dfrac{-y^2}{2}}\,dy = \Phi(x).$$

Суть центральной предельной теоремы заключается в том, что сумма независимых, одинаково распределенных случайных величин при надлежащем „центрировании“ и „нормировании“ с увеличением числа слагаемых $n \to \infty$ ведет себя почти как стандартно распределенная случайная величина (имеющая нормальное распределение с математическим ожиданием $a = 0$ и средним квадратическим отклонением $\sigma = 1$).

Из центральной предельной теоремы следует формула для подсчета вероятности того, что в n испытаниях Бернулли при достаточно большом значении n, число успехов будет заключено между m_1 и m_2:

$$P(m_1 \leq m \leq m_2) = P\left(\frac{m_1 - np}{\sqrt{npq}} \leq \frac{m - np}{\sqrt{npq}} \leq \frac{m_2 - np}{\sqrt{npq}}\right) =$$

$$= \varPhi\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \varPhi\left(\frac{m_1 - np}{\sqrt{npq}}\right),$$

где $\varPhi(x)$ – функция Лапласа (функция распределения стандартного нормального закона). Этот результат носит название **интегральная теорема Муавра-Лапласа**.

При большом количестве испытаний Бернулли подсчет вероятности попадания числа успехов в заданный интервал с использованием формулы Бернулли становится очень трудоемким. Наличие приближенной формулы значительно облегчает вычисления. Интегральной теоремой Муавра-Лапласа можно пользоваться, если величина $np(1-p)$ составляет хотя бы несколько десятков [1]. Если же величина np порядка нескольких единиц, используют другую предельную теорему для биномиального закона, основанную на распределении Пуассона.

Теорема 2.6.3. (Теорема Пуассона). *Пусть задана $\{\xi_n\}$ – последовательность случайных величин с биномиальными распределениями, причем параметр p меняется так, что величина $\lambda = np$ остается постоянной. Тогда эта последовательность сходится по распределению к случайной величине, имеющей распределение Пуассона с параметром λ. Это означает, что при любом неотрицательном k справедливо соотношение*

$$\lim_{n\to\infty} P\left(\xi_n = k\right) = \frac{\lambda^k}{k!}e^{-\lambda}, \quad \textit{где} \;\; \lambda = np.$$

Пример 2.6.1. В 50 % случаев в мастерскую по ремонту мобильных телефонов обращаются с программными проблемами. В течение месяца в ремонт сдано 150 телефонов. Какова вероятность того, что число обращаний по причине программных проблем отличается от 75 не более чем на 5?

Решение. Из условия задачи следует, что число мобильных телефонов с программными проблемами должно находиться в пределах от 70 до 80. По теореме Муавра–Лапласа

$$P(70 \le m \le 80) = \Phi\left(\frac{80 - 150 \cdot \frac{1}{2}}{\sqrt{150 \cdot \frac{1}{2} \cdot \frac{1}{2}}}\right) - \Phi\left(\frac{70 - 150 \cdot \frac{1}{2}}{\sqrt{150 \cdot \frac{1}{2} \cdot \frac{1}{2}}}\right) =$$

$$= \Phi\left(\frac{5}{\sqrt{37,5}}\right) - \Phi\left(-\frac{5}{\sqrt{37,5}}\right) \approx 2 \cdot \Phi(0,82) \approx 0,59.$$

Пример 2.6.2. Используя условия предыдущей задачи, укажите, в каких границах с вероятностью 0,99 находится число обращаний по причине программных проблем.

Решение. Из условия следует, что

$$P\left(\left|\frac{m - np}{\sqrt{npq}}\right| \le t\right) \approx 0,99.$$

Получаем

$$2\Phi(t) = 0,99.$$

По таблице значений функции Лапласа находим $t = 3$, значит число обращаний по причине программных проблем лежит в пределах $np \pm 3\sqrt{npq}$. Подставляя значения n, p, q, получаем 75 ± 18.

Пример 2.6.3. По заказу фирмы, занимающейся установкой пластиковых окон, завод изготовил 1000 москитных сеток. Вероятность брака (размеры изделия не соответствуют размерам, указанным в договоре) равна 0,5 %. Какова вероятность того, что в изготовленной партии москитных сеток ровно 3 не соответствуют размерам заказчика; не менее трех не соответствуют размерам заказчика?

Решение. По условию задачи $n = 1000$; $p = 0,005$; $q = 0,995$. Воспользуемся формулой Пуассона. Получаем

$$\lambda = 1000 \cdot 0,005 = 5,$$

$$P(m = 3) = \frac{5^3}{3!}e^{-5} \approx 0,14.$$

Ответим на второй вопрос задачи:

$$P(m \geq 3) = 1 - P(m < 3) = 1 - P(m = 0) - P(m = 1) - P(m = 2) =$$

$$= 1 - \frac{5^0}{0!}e^{-5} - \frac{5^1}{1!}e^{-5} - \frac{5^2}{2!}e^{-5} \approx 0,88.$$

Задачи для самостоятельного решения

1. Задана последовательность независимых случайных величин $\xi_1, \xi_2, ..., \xi_n$, имеющих следующий закон распределения:

ξ_n	-$\sqrt{n}$	0	$\sqrt{n}$
P	1/n	1-2/n	1/n

Применим ли к этой последовательности закон больших чисел?

2. Известно, что масса некоторой микросхемы является случайной величиной ξ, имеющей равномерное распределение на отрезке от 1 до 2 г. В каких пределах с вероятностью 0,99 будет находиться суммарный вес 10000 микросхем?

3. В продукции цеха детали отличного качества составляют 50 %. Детали укладываются в коробки по 200 штук в каждой. Какова вероятность того, что число деталей отличного качества в коробке отличается от 100 не более, чем на 5?

Ответы к задачам для самостоятельного решения

1. Да 2. $15kg \pm 75g$. 3. 0,5.

3. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

3.1. Числовые характеристики выборки

Математическая статистика занимается систематизацией и обработкой экспериментальных данных с целью получения научных и практических выводов. Множество всех возможных объектов, объединенных по некоторому качественному или количественному признаку, называется **генеральной совокупностью**. Обычно исследуется более узкий, конечный набор объектов, называемый выборочной совокупностью. Предположим, например, что нужно получить сведения о средней зарплате в Санкт-Петербурге. Понятно, что произвести опрос всего трудоспособного населения города не представляется возможным, поэтому опрашивается только часть населения. Полученные сведения представляются в виде набора чисел. Этот набор называется выборкой. Задача математической статистики – на основе выборочных данных получить сведения о генеральной совокупности. Для этого разработаны методы обработки и систематизации данных, методы оценки параметров распределения и подтверждения гипотез о законе распределения.

Приемы и методы, используемые в математической статистике, опираются на классические результаты теории вероятностей. Генеральной совокупности соответствует вероятностное пространство, в котором рассматривается наблюдаемая случайная величина ξ. Если в n повторных наблюдениях ξ приняла численные значения $x_1, x_2, \ldots, x_n$, то числовой вектор

$$X = [x_1, x_2, ..., x_n]^T$$

называется выборкой из распределения наблюдаемой случайной величины или просто **выборкой** (часто выборку записывают в виде матрицы-строки). Математической моделью выборки является случайный вектор $[\xi_1, \xi_2, \ldots, \xi_n]^T$, где $\xi_i = \xi$.

Количество чисел в выборке называется **объем выборки**.

Работать с выборкой будет удобнее, если упорядочить числа по возрастанию. Выборка, в которой $x_i \leq x_{i+1}$ называется **вариационным рядом**. Если выборка X — вариационный ряд, $x_n - x_1$ называется **размахом выборки**. Значение находящееся в середине вариационного ряда называется **выборочной медианой**. Если объем выборки нечетное число $n = 2k + 1$, то $x_{med} = x_{k+1}$, если объем выборки — четное число $n = 2k$, то $x_{med} = \dfrac{x_k + x_{k+1}}{2}$.

Если в выборке одни и те же числа встречаются по нескольку раз, то разумно перечислить все различные значения, указав, сколько раз каждое из них встречается:

x	x_1	x_2	x_3	...	x_k
n	n_1	n_2	n_3	...	n_k

Так записанная выборка называется **статистическим рядом**, $n_1 + n_2 + ... + n_k = n$

Варианта с наибольшей частотой называется **выборочной модой** и обозначается x_{mod}. Может быть несколько вариант с одинаковой частотой, в этом случае выборочная мода не определяется.

Основными выборочными числовыми характеристиками являются

1. **Выборочное среднее** $\bar{x} = \dfrac{1}{n}\sum\limits_{i=1}^{n} x_i$,

2. **Выборочная дисперсия** $\hat{\sigma}^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} (x_i - \bar{x})^2$.

Любая функция выборки (выборочных наблюдений) называется **статистикой**. Статистика, значение которой используется в качестве приближенного значения параметра, называется **точечной оценкой параметра**.

Определение 3.1.1. *Статистика* $\hat{\theta}_n$ *называется несмещенной точечной оценкой параметра* θ, *если ее математическое ожидание совпадает с оцениваемым параметром при любом объеме выборки* n*:*

$$M(\hat{\theta}_n) = \theta.$$

Определение 3.1.2. *Статистика* $\hat{\theta}_n$ *называется асимптотически несмещенной точечной оценкой параметра* θ, *если*

$$\lim_{n\to+\infty} M(\hat{\theta}_n) = \theta.$$

Определение 3.1.3. *Статистика* $\hat{\theta}_n$ *называется состоятельной точечной оценкой параметра* θ, *если при неограниченном увеличении объема выборки она сходится по вероятности к оцениваемому параметру:*

$$\lim_{n\to+\infty} P\left(|\hat{\theta}_n - \theta| < \varepsilon\right) = 1, \qquad \forall\varepsilon > 0.$$

Покажем, что выборочное среднее является несмещенной состоятельной оценкой математического ожидания наблюдаемой случайной величины. Пусть $M\xi = a$, $D\xi = \sigma^2$. Тогда

$$M\bar{x} = \frac{1}{n}\sum_{i=1}^{n} M\xi_i = \frac{1}{n}\sum_{i=1}^{n} a = a,$$

следовательно оценка несмещенная.

$$D\bar{x} = \frac{1}{n^2}\sum_{i=1}^{n} D\xi_i = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n}.$$

По неравенству Чебышева при $\forall\varepsilon > 0$

$$P\left(|\bar{x} - a| \geq \varepsilon\right) \leq \frac{D\bar{x}}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0, \qquad n \to +\infty,$$

значит

$$\lim_{n\to+\infty} P\left(|\bar{x} - a| < \varepsilon\right) = 1$$

и $\bar{x}$ — состоятельная оценка параметра a.

В [1] показано, что выборочная дисперсия является состоятельной оценкой дисперсии наблюдаемой случайной величины, но является смещенной оценкой. Поэтому вводится так называемая **исправленная выборочная дисперсия**

$$\hat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2,$$

которая дает несмещенную оценку дисперсии.

Пример 3.1.1. Дана выборка $X = (-2, 5, -6, -3, 0, 0, -4, 2)$. Найти выборочное среднее, моду, медиану и исправленную выборочную дисперсию.

Решение. Выборочное среднее равно

$$\frac{-2+5-6-3-4+2}{8} = -1.$$

Так как число 0 встречается в данной выборке два раза, а все остальные числа по одному, $x_{mod} = 0$.

Чтобы найти выборочную медиану, упорядочим выборку по возрастанию (запишем соответствующий вариационный ряд):

$$X = (-6, -4, -3, -2, 0, 0, 2, 5).$$

Выборочная медиана равна $x_{med} = \dfrac{-2+0}{2} = -1.$

Исправленная выборочная дисперсия равна

$$\hat{S}^2 = \frac{1}{7}\sum_{i=1}^{8}(x_i+1)^2 = \frac{25+9+4+1+1+1+9+36}{7} = \frac{86}{7} = 12\frac{2}{7}.$$

Рассмотрим методы получения точечных оценок параметров: метод выборочных моментов и метод максимального правдоподобия.

3.2. Метод выборочных моментов

Метод моментов был предложен английским математиком Карлом Пирсоном (1857–1936) в 1894 г.

Суть метода заключается в приравнивании некоторого числа выборочных моментов к соответствующим теоретическим, которые являются функциями неизвестных параметров $\theta_1, \theta_2, ..., \theta_k$. Количество уравнений должно совпадать с количеством оцениваемых параметров.

Выборочные начальные моменты k-го порядка вычисляются по формуле

$$\hat{\alpha}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k,$$

а центральные выборочные моменты — по формуле

$$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^k.$$

Замечание. Уравнения для неизвестных параметров $\theta_1, \theta_2, ..., \theta_k$ можно составлять, используя как начальные, так и центральные моменты. Поэтому в методе моментов есть неопределенность. Оценки часто получаются смещенными.

Итак, оценки параметров $\theta_1, \theta_2, ..., \theta_k$ являются решением системы уравнений

$$\alpha_i(\theta_1, \theta_2, ..., \theta_k) = \hat{\alpha}_i$$

или

$$\mu_i(\theta_1, \theta_2, ..., \theta_k) = \hat{\mu}_i$$

для некоторых $i = i_1, i_2, ..., i_k$.

Система уравнений для моментов может не иметь решений в элементарных функциях или вообще быть несовместной.

Если распределение определяется одним параметром, то для построения оценок один теоретический момент приравнивается к одному эмпирическому моменту того же порядка (обычно первого).

Пример 3.2.1. Случайная величина ξ (число попаданий стрелком в „десятку“ при n выстрелах) имеет биномиальное распределение с неизвестным параметром p. Ниже приведено эмпирическое распределение числа попаданий в 10 сериях по 5 выстрелов в каждой (в первой строке указано число x_i попаданий в серии; во второй строке указана частота n_i – количество серий, в которых наблюдалось x_i попаданий):

x_i	3	4	5
n_i	2	5	3

Найти методом моментов точечную оценку параметра p биномиального распределения.

Решение. Математическое ожидание биномиального распределения равно $M\xi = np$. Приравниваем математическое ожидание к выборочному среднему, так как оно является начальным моментом первого порядка. Получаем $np = \bar{x}$. Значит $\hat{p} = \dfrac{\bar{x}}{n}$. В нашем случае $\bar{x} = (3\times2+4\times5+5\times3)/10 = 4,1$; $\hat{p} = \dfrac{4,1}{5} = 0,82$.

Пример 3.2.2. Компания производит соединительные пластины, причем контролируется отклонение их толщины от номинального размера — ξ. Случайная величина ξ подчинена нормальному закону распределения с неизвестными параметрами a и σ. Ниже приведена таблица наблюдаемых отклонений от номинала, подвергнутых группировке, для $n = 100$ изделий (в первой строке указаны середины интервалов отклонений x_i (мм); во второй строке приведены соответствующие частоты):

x_i	0,3	0,5	0,7	0,9	1,1	1,5
n_i	35	15	16	24	6	4

Найдите методом моментов точечные оценки неизвестных параметров a и σ нормального распределения.

Решение. Необходимо составить два уравнения, так как требуется найти два неизвестных параметра. Первое уравнение получим, приравнивая начальный теоретический момент первого порядка к начальному эмпирическому моменту первого порядка:

$$a^* = \hat{\alpha}_1 = \bar{x}.$$

Второе уравнение составим, приравнивая центральный теоретический момент второго порядка к центральному эмпирическому моменту второго порядка

$$(\sigma^*)^2 = \hat{\mu}_2 = \hat{\sigma}^2.$$

Находим по выборке $\bar{x}$ и $\hat{\sigma}^2$:

$$\bar{x} = \frac{1}{100}\sum_{i=1}^{6} x_i n_i = (0,3{\cdot}35+0,5{\cdot}15+0,7{\cdot}16+0,9{\cdot}24+1,1{\cdot}6+1,5{\cdot}4)/100 = 0,634,$$

$$\hat{\sigma}^2 = \frac{1}{100}\sum_{i=1}^{6} x_i^2 n_i - \bar{x}^2 \approx 0,102.$$

Поскольку $\hat{\sigma}^2$ — смещенная оценка дисперсии, найдем исправленную выборочную дисперсию

$$S^2 = \frac{100}{99}\hat{\sigma}^2 \approx 0,103.$$

Таким образом, получаем: $a^* = 0,634$ (мм), $\sigma^* \approx 0,321$ (мм).

Пример 3.2.3. Найдите методом моментов оценку параметра p (вероятности „успеха“) для геометрического распределения.

Решение. Напомним, что для геометрического распределения $P(\xi = x_i) = (1-p)^{x_i - 1}p$, где ξ — случайная величина (номер первого успешного испытания в схеме Бернулли), x_i — число испытаний, произведенных до первого появления события. При этом $M\xi = \dfrac{1}{p}$.

Приравниваем начальный теоретический момент первого порядка к начальному эмпирическому моменту первого порядка. Получим $M\xi = \bar{x}$. Значит $\dfrac{1}{p} = \bar{x}$. Следовательно, $\hat{p} = \dfrac{1}{\bar{x}}$.

Пример 3.2.4. Найдите методом моментов оценку параметра θ распределения Рэлея. Плотность распределения Рэлея равна

$$f_\xi(x) = \begin{cases} 0, & x < 0, \\ \dfrac{x}{\theta} e^{-\dfrac{x^2}{2\theta}}, & x \geq 0. \end{cases}$$

Решение. Требуется оценить один параметр, поэтому достаточно приравнять начальный теоретический момент первого порядка к начальному эмпирическому моменту первого порядка. Найдем

$$\alpha_1 = M\xi = \int\limits_{-\infty}^{+\infty} x f_\xi(x) dx = \frac{1}{\theta} \int\limits_0^{+\infty} x^2 e^{-\dfrac{x^2}{2\theta}} dx.$$

Сделаем замену переменной $x = \sqrt{2\theta}t$ и проинтегрируем по частям

$$\alpha_1 = 2\sqrt{2\theta} \int\limits_0^{+\infty} t^2 e^{-t^2} dt = \sqrt{2\theta} \left(-te^{-t^2}\Big|_0^{+\infty} + \int\limits_0^{+\infty} e^{-t^2} dt \right) = \sqrt{\frac{\pi\theta}{2}}.$$

Приравняем $\alpha_1 = \hat{\alpha}_1 = \bar{x}$: $\sqrt{\dfrac{\pi\theta}{2}} = \bar{x} \quad \Rightarrow \quad \hat{\theta} = \dfrac{2(\bar{x})^2}{\pi}$.

Если приравнивать вторые начальные моменты, то получается другая оценка параметра: $\hat{\theta} = \dfrac{\hat{\alpha}_2}{2}$. Приравнивание вторых центральных моментов приводит к оценке $\hat{\theta} = \dfrac{2S^2}{4 - \pi}$.

Теоретически можно показать, что на самом деле $\theta = \sigma^2$.

Задачи для самостоятельного решения

1. Найдите методом моментов оценку параметра λ распределения Пуассона.

2. Найдите методом моментов оценку параметра θ для геометрического распределения с вероятностью „успеха“ $p = 1/(1 + \theta), \theta \geq 0$.

3. Случайная величина ξ (задержка доставки товаров курьерской службой) имеет показательное распределение. В таблице приведены сгруппированные данные по продолжительности задержки (в часах) для двухсот

заказов, выполненных не вовремя.

x_i	2,5	6	7	22,5	23	25,5
n_i	135	45	15	2	2	1

Найдите методом моментов точечную оценку неизвестного параметра λ показательного распределения.

4. Прибор состоит из двух блоков — основного и резервного. Если основной блок выходит из строя, включается резервный. Времена службы блоков показательно распределены со средними θ_1 и θ_2. Выборочные испытания для 10 приборов показали средний срок службы 35 часов и среднее квадратическое отклонение 25 часов. Оценить средние времена службы основного и резервного блоков методом моментов в предположении, что $\theta_1 > \theta_2$.

Ответы к задачам для самостоятельного решения

1. $\hat{\lambda} = \bar{x}$
2. $\hat{\theta} = \bar{x} - 1$
3. $\hat{\lambda} \approx 0,24$.
4. 20 и 15 часов

3.3. Метод максимального правдоподобия

Пусть $[x_1, x_2, ..., x_n]$ — выборка, полученная в результате проведения n независимых наблюдений над случайной величиной ξ. Предположим, что известен вид закона распределения случайной величины ξ, но неизвестен параметр θ, которым определяется этот закон. Требуется по выборке оценить параметр θ.

Метод максимального правдоподобия (ММП), предложенный Р.Фишером, является одним из основных методов получения точечных оценок параметров. Основу метода составляет функция правдоподобия, выражающая плотность вероятности (либо вероятность) совместного появления результатов выборки $[x_1, x_2, ..., x_n]$.

Если ξ – абсолютно непрерывная случайная величина, то функцией правдоподобия, построенной по выборке $X = [x_1, x_2, \ldots, x_n]$, называется функция

$$L_n(X, \theta) = \prod_{i=1}^{n} f_\xi(x_i, \theta),$$

где $f_\xi(x, \theta)$ – плотность распределения наблюдаемой случайной величины ξ.

Если ξ – дискретная случайная величина, то функция правдоподобия имеет вид

$$L_n(X, \theta) = \prod_{i=1}^{n} p(x_i, \theta),$$

где $p(x_i, \theta) = P(\xi = x_i, \theta)$.

Из определения следует, что чем больше значение функции правдоподобия $L_n(X, \theta)$, тем более вероятно (правдодподобнее) появление в результате наблюдений чисел $x_1, x_2, ..., x_n$. За точечную оценку параметра θ согласно ММП берут такое его значение $\hat{\theta}$, при котором функция правдоподобия достигает максимума. Эту оценку называют оценкой максимального правдоподобия.

Поскольку максимизируемая функция представляет собой произведение большого числа сомножителей, обычно бывает удобнее искать максимум не самой функции, а ее натурального логарифма. На результат такая замена не повлияет, поскольку сама функция положительна, а натуральный логарифм является монотонно возрастающей функцией, определенной при положительных значениях аргумента. Функцию $lnL_n(X, \theta)$ называют логарифмической функцией правдоподобия.

Точку максимума функции $lnL_n(X, \theta)$ ищут следующим образом:

1. Находят производную $\dfrac{dlnL_n(X, \theta)}{d\theta}$.
2. Приравнивают производную нулю и находят критические точки – корни полученного уравнения, которое называют уравнением правдоподобия (необходимое условие экстремума).
3. Проверяют, является ли критическая точка точкой максимума с помощью достаточного условия экстремума. Например, проверяют что

$$\frac{d^2lnL_n(X, \theta)}{d^2\theta} < 0.$$

Если оценке подлежат несколько параметров $\theta_1, \theta_2, ..., \theta_k$ распределения, то точечные оценки параметров определяются из системы k уравнений правдоподобия.

ММП имеет ряд достоинств: оценки максимального правдоподобия, вообще говоря, состоятельны (но они могут быть смещенными), распределены асимптотически нормально (при больших значениях n приближенно нормальны) и имеют наименьшую дисперсию по сравнению с другими асимптотически нормальными оценками. Если для оцениваемого параметра θ существует эффективная [1] оценка $\hat{\theta}$, то уравнение правдоподобия имеет единственное решение $\hat{\theta}$. Этот метод наиболее полно использует данные выборки об оцениваемом параметре, поэтому он особенно полезен в случае малых выборок. Недостаток метода состоит в том, что он часто требует сложных вычислений.

Пример 3.3.1. Найдите с помощью ММП по выборке $[x_1, x_2, ..., x_n]$ точечную оценку неизвестного параметра λ показательного распределения, плотность которого $f_\xi(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0. \end{cases}$

Решение. Составим функцию правдоподобия:

$$L_n(X, \theta) = f_\xi(x_1, \theta) \cdot f_\xi(x_2, \theta) \cdot ... \cdot f_\xi(x_n, \theta).$$

Учитывая, что $\theta = \lambda$ и, следовательно, $f_\xi(x, \theta) = f_\xi(x, \lambda) = \lambda e^{-\lambda x}$.

$$L_n(X, \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x} = \lambda^n e^{-\lambda \sum\limits_{i=1}^{n} x_i}.$$

Найдем логарифмическую функцию правдоподобия:
$lnL_n(X, \lambda) = nln\lambda - \lambda \sum\limits_{i=1}^{n} x_i$.

Найдем первую производную по параметру λ: $\dfrac{dlnL_n(X, \lambda)}{d\lambda} = \dfrac{n}{\lambda} - \sum\limits_{i=1}^{n} x_i$. Напишем уравнение правдоподобия, приравняв первую производную нулю: $\dfrac{n}{\lambda} - \sum\limits_{i=1}^{n} x_i = 0$.

Найдем критическую точку, для чего решим полученное уравнение относительно $\lambda : \lambda = \dfrac{n}{\sum\limits_{i=1}^{n} x_i} = \dfrac{1}{\bar{x}}$.

Найдем вторую производную логарифмической функции правдоподобия по λ : $\dfrac{d^2 \ln L_n(X,\lambda)}{d\lambda^2} = -\dfrac{n}{\lambda^2}$. Вторая производная отрицательна, следовательно критическая точка является точкой максимума. Значит в качестве оценки максимального правдоподобия параметра λ показательного распределения надо принять величину, обратную выборочному среднему $\hat{\lambda} = \dfrac{1}{\bar{x}}$.

Пример 3.3.2. Проведена серия из n опытов. Каждый опыт состоит в проведении m испытаний Бернулли. В результате получена выборка $X = [x_1, x_2, ..., x_n]$ значений случайной величины ξ — числа успехов в m испытаниях Бернулли. Найдите с помощью ММП точечную оценку неизвестного параметра p биномиального распределения (вероятность успеха в одном испытании).

Решение. Составим функцию правдоподобия:

$$L_n(X,\theta) = p(x_1,\theta) \cdot p(x_2,\theta) \cdot ... \cdot p(x_n,\theta),$$

Случайная величина ξ распределена по биномиальному закону с параметрами p, m

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

Учитывая, что неизвестен параметр $\theta = p$, получим:

$$L_n(X,p) = C_m^{x_1} C_m^{x_2} ... C_m^{x_n} p^{x_1+x_2+...x_n} (1-p)^{nm-(x_1+x_2+...+x_n)}.$$

Найдем логарифмическую функцию правдоподобия;

$$lnL_n(X,p) = ln(C_m^{x_1} C_m^{x_2} ... C_m^{x_n}) + \sum_{i=1}^{n} x_i ln(p) - \left(nm - \sum_{i=1}^{n} x_i\right) ln(1-p).$$

Найдем первую производную по параметру p:

$$\frac{dlnL_n(X,p)}{dp} = \frac{\sum\limits_{i=1}^{n} x_i}{p} - \left(nm - \sum_{i=1}^{n} x_i\right) \frac{1}{1-p}.$$

Приравняв первую производную нулю и решив полученное уравнение, получим критическую точку

$$p = \frac{\sum\limits_{i=1}^{n} x_i}{nm} = \frac{\bar{x}}{m}.$$

Найдем вторую производную логарифмической функции правдоподобия по р:

$$\frac{d^2 lnL(x,p)}{dp^2} = -\frac{\sum\limits_{i=1}^{n} x_i}{p^2} - \frac{nm - \sum\limits_{i=1}^{n} x_i}{(1-p)^2}.$$

Легко убедиться в том, что при $p = \frac{\bar{x}}{m}$ вторая производная отрицательна. Следовательно, эта точка есть точка максимума и ее надо принять в качестве оценки максимального правдоподобия неизвестного параметра p биномиального распределения: $\hat{p} = \frac{\bar{x}}{m}$.

Пример 3.3.3. Случайная величина ξ (число появлений события в 10 независимых испытаниях) подчинена биномиальному закону распределения с неизвестным параметром p. Проведено 110 опытов по 10 испытаний Бернулли в каждом опыте. В результате получена выборка значений случайной величины ξ – числа успехов в 10 испытаниях. Соответствующий статистический ряд:

x_i	0	1	2	3	4	5	6	7
n_i	2	3	10	22	26	20	12	15

.

В первой строке указано x_i — число появлений события в одном опыте из 10 испытаний, во второй строке приведена частота n_i — число опытов, в которых наблюдалось x_i появлений события A. Найдите с помощью ММП точечную оценку неизвестного параметра p биномиального распределения.

Решение. В предыдущем примере было показано, что в качестве оценки неизвестного параметра p биномиального распределения надо взять $\hat{p} = \frac{\bar{x}}{m}$. Подставляя данные таблицы, получаем $\hat{p} = \frac{\sum\limits_{i=1}^{8} n_i x_i}{110 \cdot 10} = \frac{47}{110} \approx 0,4273$.

Пример 3.3.4. Найдите с помощью ММП оценки параметров a и σ нормального распределения с плотностью

$$f_\xi(x, a, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x-a)^2}{2\sigma^2}},$$

если в результате n испытаний случайная величина ξ приняла значения $x_1, x_2, \ldots, x_n$.

Решение. Оцениваемые параметры распределения $\theta_1 = a$ и $\theta_2 = \sigma$. Составим функцию правдоподобия:

$$L_n(X, a, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x_1-a)^2}{2\sigma^2}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x_2-a)^2}{2\sigma^2}} \ldots \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x_n-a)^2}{2\sigma^2}}$$
$$= \frac{1}{\sigma^n(\sqrt{2\pi})^n} e^{-\sum\limits_{i=1}^{n}(x_i-a)^2/(2\sigma^2)}.$$

Найдем логарифмическую функцию правдоподобия:

$$lnL_n(X, a, \sigma) = -n \cdot ln(\sigma) + ln\frac{1}{(2\pi)^n} - \frac{\sum\limits_{i=1}^{n}(x_i-a)^2}{2\sigma^2}.$$

Найдем частные производные по a и по σ:

$$\frac{\partial lnL_n(X, a, \sigma)}{\partial a} = \frac{\sum\limits_{i=1}^{n} x_i - na}{\sigma^2}, \qquad \frac{\partial lnL_n(X, a, \sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum\limits_{i=1}^{n}(x_i-a)^2}{\sigma^3}.$$

Приравняв частные производные нулю и решив полученную систему двух уравнений относительно a и σ^2, получим:

$$a = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}, \qquad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \hat{\sigma}_n^2 = \hat{\mu}_2.$$

Итак, искомые оценки максимального правдоподобия:

$$\hat{a} = \bar{x} \qquad \hat{\sigma} = \sqrt{\hat{\mu}_2}.$$

Заметим, что первая оценка несмещенная, а вторая смещенная.

Задачи для самостоятельного решения

1. Случайная величина ξ распределена по закону Пуассона с неизвестным параметром λ:

$$P(\xi = x_i) = \lambda^{x_i} \frac{e^{-\lambda}}{x_i!}.$$

С помощью ММП Найдите по выборке $[x_1, x_2, ..., x_n]$ точечную оценку неизвестного параметра λ распределения Пуассона.

2. Случайная величина ξ (число поврежденных стеклянных изделий в одном контейнере) распределена по закону Пуассона с неизвестным параметром λ. Ниже приведено эмпирическое распределение числа поврежденных изделий в 500 контейнерах (в первой строке указано количество x_i поврежденных изделий в одном контейнере, во второй строке приведена частота n_i —число контейнеров, содержащих x_i поврежденных изделий):

x_i	0	1	2	3	4	5	6	7
n_i	199	169	87	31	9	3	1	1

Найдите с помощью ММП точечную оценку неизвестного параметра λ распределения Пуассона.

3. Найдите с помощью ММП по выборке $[x_1, x_2, ..., x_n]$ точечную оценку параметра p геометрического распределения:

$$P(\xi = x_i) = (1-p)^{x^i - 1} p,$$

где x_i – число испытаний Бернулли, проведенных до первого успеха; p – вероятность успеха в одном испытании.

Ответы к задачам для самостоятельного решения

1. $\bar{\lambda} = \frac{1}{n} \sum\limits_{i=1}^{n} x_i = \bar{x}$, где $n = \sum\limits_{i=1}^{n} n_i$.   2. $\hat{\lambda} = 1$.   3. $\hat{p} = \frac{1}{\bar{x}}$.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Даугавет А. И., Постников Е. В., Червинская Н. М. Введение в теорию вероятностей: учеб. пособие. СПб.: Изд-во СПбГЭТУ „ЛЭТИ“, 2012.

2. Даугавет А. И., Постников Е. В., Солынин А. А. Математическая статистика: учеб. пособие. СПб.: Изд-во СПбГЭТУ „ЛЭТИ“, 2012.

5. Фадеева Л. Н. Математика для экономистов. Теория вероятностей и математическая статистика. курс лекций, М., Эксмо, 2006.

ОГЛАВЛЕНИЕ

1. ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО 3
1.1. Построение модели случайного эксперимента. Вероятностное пространство. 3
1.2. Теорема сложения. 7
1.3. Классическая вероятностная модель. 10
1.4. Условная вероятность. Теорема умножения. 17
1.5. Формула полной вероятности. 25
1.6. Формула Байеса. 28
1.7. Испытания Бернулли. 31
2. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ 34
2.1. Дискретная случайная величина. 34
2.2. Абсолютно-непрерывная случайная величина. 41
2.3. Дискретный случайный вектор. 48
2.4. Случайный вектор с абсолютно-непрерывным распределением. 52
2.5. Неравенство Чебышева. 58
2.6. Предельные теоремы. 61
3. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА 64
3.1. Числовые характеристики выборки. 64
3.2. Метод выборочных моментов. 68
3.3. Метод максимального правдоподобия. 72
Список литературы 79

Гоголева Надежда Генриховна,
Жукова Екатерина Евгеньевна,
Колбина Светлана Анатольевна ,
Непомнящая Татьяна Владимировна,
Фролова Елена Вениаминовна,
Шевченко Елена Аркадьевна.

**Теория вероятностей
в примерах и задачах.**
Учебное пособие

Редактор О. Р. Крумина

Подписано в печать Формат 60×84 1/16.
Бумага офсетная. Печать цифровая. Печ. л. 5,0.
Гарнитура „Times New Roman“ Тираж экз. Заказ

Издательство СПбГЭТУ “ЛЭТИ”
197376, С.-Петербург, ул. Проф. Попова, 5