Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования

**Санкт-Петербургский государственный технологический
университет растительных полимеров**

МАТЕМАТИКА.

Руководство к решению задач Теории вероятностей и математической статистики

Учебно-методическое пособие

Санкт-Петербург
2007

0108

Федеральное агентство по образованию

Государственное образовательное учреждение
высшего профессионального образования

«Санкт-Петербургский государственный технологический
университет растительных полимеров»

Математика

Руководство к решению задач теории вероятностей и математической статистики

Учебно-методическое пособие

Санкт-Петербург
2007

УДК 51(07.07)

Математика. Руководство к решению задач теории вероятностей и математической статистики: Учебно-методическое пособие/ сост П.П. Смышляев, Е. А. Титова, Н.К. Брыксенкова, О.Е. Куляхтина, Е.Г.Иванова, Р.В.Пелюхов; ГОУВПО СПбГТУРП. СПб., 2007.- 46 с.

Излагаются основные положения теории вероятностей и математической статистики. Ключевым понятиям курса уделено особое внимание. Рассмотрены типовые задачи. Рекомендуется для студентов и исследователей технического (технологического) профиля заочного отделения.

Рецензент: зав. кафедрой прикладной и вычислительной математики ГОУВПО СПбГТУРП, д-р техн. наук, проф. О.К.Федоров

Подготовлено и рекомендовано к печати кафедрой высшей математики ГОУВПО Санкт-Петербургского государственного технологического университета растительных полимеров (протокол № 7 от 16 марта 2007 г.).

Утверждено к изданию методической комиссией факультета промышленной энергетики ГОУВПО СПбГТУРП (протокол № 8 от 12 мая 2007 г.).

Рекомендовано к изданию Редакционно-издательским советом университета в качестве учебно-методического пособия.

Редактор и корректор Т. А.Смирнова
Техн.редактор Л.Я.Титова

Подп. к печати 31.05.2007. Формат 60x84/16. Бумага тип № 1.
Печать офсетная. Объём 3,0 печ.л., 3,0 уч.-изд. л.
Тираж 100 экз. Изд. № 83 .Цена «С» . Заказ 1628.

Ризограф ГОУВПО Санкт-Петербургского государственного технологического университета растительных полимеров, 198095, Санкт-Петербург, ул. Ивана Черных, 4.

Введение

В практической деятельности мы встречаемся с различными явлениями, из которых можно выделить две большие группы явлений – детерминированные и случайные. К первой группе относятся события, исход которых мы можем заранее предсказать, а ко второй – события, исход которых заранее предсказать невозможно (например, номер выигрышного лотерейного билета).

Теория вероятностей занимается математическим описанием случайных явлений. Было замечено, что одни случайные события происходят более часто, чем другие. Количественная характеристика частоты наступления случайного события, выраженная в долях единицы, называется вероятностью события. Чем ближе вероятность к единице, тем чаще происходит наблюдаемое случайное событие, и наоборот, если вероятность близка к нулю, то наблюдаемое событие происходит сравнительно редко.

Основной математической моделью теории вероятностей является теоретико-вероятностный эксперимент (опыт, испытание), в рамках которого происходят определённые случайные события. Математический анализ этих событий и составляет предмет теории вероятностей.

1. Основные понятия теории вероятностей

1.1. Основные комбинаторные понятия

Наблюдаемые нами события можно подразделить на следующих три вида : достоверные, невозможные и случайные.

События называется случайным, если оно может либо произойти, либо не произойти. Например, при бросании игральной кости может выпасть шестерка, либо какое-то другое количество очков (следовательно, шестерка не выпадет).

Событие называется достоверным, если оно всегда происходит в эксперименте. Достоверное событие принято обозначать буквой I. Например, если у вас на полке стоят только учебники по математике, то какую книгу вы не возьмете, она окажется по этой дисциплине. Соответственно, взять учебник по математике – достоверное событие, поскольку других вариантов у вас нет.

Событие называется невозможным, если оно никогда не происходит в эксперименте. Невозможное событие обозначается символом $\varnothing$. Например, вытащить красный шар из коробки с зелеными является невозможным событием.

Суммой событий A и B называется такое событие, которое состоит в том, что происходит хотя бы одно из событий A и B (при этом не исключается возможность наступления обоих событий). Сумму событий A и B записывают $A+B$, и геометрически она показана на рис. 1.1 штриховкой.

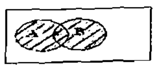


Рис. 1.1

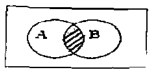


Рис. 1.2

Произведением событий A и B называется такое событие, которое состоит в том, что происходят оба события. Обозначают произведение событий A и B следующим образом : $A\cdot B$, на рис. 1.2 оно выделено штриховкой.

Событие называется противоположным событию A, если оно заключается в том, что событие A не наступило. Противоположное событие обозначается как $\overline{A}$ (рис. 1.3).

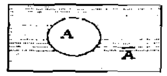


Рис. 1.3

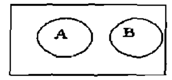


Рис. 1.4

События A и B называются несовместными если их произведение – невозможное событие (рис. 1.4).

Для операций сложения и умножения удобно использовать следующие соотношения:

$$
\begin{array}{lll}
1.\ A+\overline{A}=I; & 5.\ A\cdot\overline{A}=\varnothing; & 9.\ \overline{A+B}=\overline{A}\cdot\overline{B};\\
2.\ A+A=A; & 6.\ A\cdot A=A; & 10.\ \overline{A\cdot B}=\overline{A}+\overline{B};\\
3.\ A+\varnothing=A; & 7.\ A\cdot\varnothing=\varnothing; & 11.\ \overline{\overline{A}}=A.\\
4.\ A+I=I; & 8.\ A\cdot I=A. &
\end{array}
\tag{1.1}
$$

Рассмотрим несколько примеров.

Задача 1.1. Опыт состоит в бросании двух монет. Определить противоположные для следующих событий: 1- выпадение двух орлов;2- выпадение орла и решки; 3- выпадение хотя бы одного орла.

Решение.

1. Событие – выпадение двух орлов – не наступит, если при бросании появится хотя бы одна решка.

2. Событие – выпадения орла и решки – не наступит, если выпадут либо два орла, либо две решки.

3. Событие – выпадение хотя бы одного орла – не произойдёт, если выпадут две решки.

Задача 1.2. Два стрелка производят залп по мишени. Обозначим через A_i событие, заключающееся в том, что i-й стрелок поразил мишень. Выразить с помощью операций над событиями A_1 и A_2 следующие события: 1) мишень поражена;2) в мишень попала только одна пуля;3) мишень не поражена;4) в мишень попало две пули; 5) промахнулся второй стрелок.

Решение.

1. Если мишень поражена, то в неё попал хотя бы один стрелок, поэтому искомое событие равно $A_1 + A_2$. Т. е. происходит либо событие A_1, либо событие A_2, либо оба одновременно.

2. События A_1 и A_2 соответствуют попаданиям стрелков по мишени, а события $\overline{A_1}$ и $\overline{A_2}$ – промахам. Если в мишень попала только одна пуля, то такая ситуация возможна, когда первый стрелок попал, а второй – промахнулся (событие $A_1\overline{A_2}$) или когда первый стрелок промахнулся, а второй – попал (событие $\overline{A_1}A_2$). Поэтому искомое событие есть сумма этих событий $A_1\overline{A_2}+\overline{A_1}A_2$.

3. Если мишень не поражена, то промахнулись оба стрелка, следовательно, искомое событие является произведением $\overline{A_1}\cdot\overline{A_2}$.

4. Если в мишень попало две пули, то оба стрелка поразили мишень, поэтому искомое событие равно $A_1 \cdot A_2$.

5. Событие – промахнулся второй стрелок – противоположно событию A_2, следовательно, это событие равно $\overline{A_2}$.

Для решения многих задач теории вероятностей требуется подсчитать число благоприятных исходов, т.е. тех событий, вероятность которых необходимо определить. Этот подсчёт можно упростить, используя простейшие комбинаторные формулы.

Рассмотрим множество, состоящее из n элементов $a_1, a_2, ..., a_n$. Выбираем k элементов ($k \le n$). Группа из этих k элементов называется соединениями из n по k или выборками k элементов из n элементов.

Различают два вида соединений: сочетания и размещения.

Сочетания – это такие соединения, которые отличаются друг от друга только составом элементов.

Размещения – это такие соединения, которые отличаются друг от друга как составом элементов, так и порядком их следования.

Число различных сочетаний из n по k обозначается C_n^k и равно:

$$C_n^k = \frac{n!}{k!(n-k)!},$$

а число различных размещений из n по k обозначается A_n^k и равно :

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

Задача 1.3. Каким количеством различных способов можно посадить двух пассажиров в пятивагонную электричку так, чтобы они не сидели в одном вагоне.

Решение. Посадка двух пассажиров в разные вагоны является, по существу, выборкой двух вагонов из пяти. Эта выборка является упорядоченной, поскольку посадка первого пассажира в i-й вагон, а второго – в j-й вагон отличается от посадки первого пассажира в j-й вагон, а второго в i-й. Поэтому данная выборка есть размещение из 5 по 2. Следовательно, число различных посадок равно:

$$A_5^2 = \frac{5!}{2!} = 3 \cdot 4 \cdot 5 = 60 .$$

Задача 1.4. Каким количеством различных способов можно выбрать из урны, содержащей пять белых и три чёрных шара, два шара одного цвета.

Решение. Событие – выбор двух шаров одного цвета – является суммой двух событий: выбор двух шаров белого цвета и выбор двух шаров чёрного цвета. Обозначим число благоприятных для этих событий исходов через m_1 и m_2 соответственно. Тогда общее число различных выборок одноцветных шаров $m = m_1 + m_2$. Найдём значения m_1 и m_2. Выборка шаров одного цвета является неупорядоченной, поскольку нас интересует только состав выборки. Поэтому числа m_1 и m_2 находятся по формуле числа сочетаний. Так как белые шары выбираются только из белых, а чёрные только из чёрных, то имеем

$$m_1 = C_5^2 = \frac{5!}{2! \cdot 3!} = 10$$

и

$$m_2 = C_3^2 = \frac{3!}{1! \cdot 2!} = 3 .$$

В итоге получим

$$m = 10 + 3 = 13 .$$

1.2. Вероятность событий

Анализ случайных событий показывает, что одни события происходят более часто, чем другие. Количественная мера, характеризующая частоту появления события и выраженная в долях единицы, называется вероятностью случайного события. Чем ближе к единице вероятность события, тем чаще оно наступает в испытаниях, и, наоборот, чем ближе к нулю вероятность случайного события, тем реже оно наступает в теоретико-вероятностном эксперименте. Если вероятность события равна единице, то это означает, что событие наступает практически всегда в эксперименте, а если она равна нулю, то событие практически никогда не наступает в нём. Вероятность события A обозначают $P(A)$.

Вероятность обладает следующими свойствами:

- вероятность события $P(A)$ не меньше нуля и не больше единицы

$$0 \le P(A) \le 1 ;$$

- вероятность достоверного события $P(I)$ равна единице

$$P(I) = 1 ;$$

- вероятность суммы двух несовместных событий A и B равна сумме вероятностей этих событий.

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

Эти свойства вероятности являются аксиомами. Приведём важнейшие следствия этих аксиом:

- вероятность невозможного события равна нулю

$$P(\varnothing) = 0 ;$$

- вероятность противоположного события равна разности между единицей и вероятностью исходного события

$$P(\overline{A}) = 1 - P(A) ;$$

- если событие A влечёт событие B, то вероятность события A не превосходит вероятности события B :

$$P(A) \le P(B) .$$

Для определения вероятности события A рассмотрим подробнее события, которые происходят в эксперименте. Их можно разбить на два класса: простые (элементарные) события и сложные (составные) события.

Элементарным событием или исходом называется такое событие, которое для любого события A рассматриваемого эксперимента находится в одном из двух отношений: либо оно влечёт событие A, либо оно влечёт противоположное событие $\overline{A}$. Из определения элементарного события следует, что оно не может быть суммой других событий нашего эксперимента, а сложное событие A есть сумма всех элементарных событий, которые являются частными случаями A. Совокупность элементарных событий называется пространством элементарных событий. Оно обозначается буквой Ω, а элементарные события, входящие в него, обозначаются $\omega_1, \omega_2, \ldots, \omega_n, \ldots$ Если число элементарных событий в эксперименте конечно, то такой эксперимент называется конечным.

Для конечных экспериментов вероятность события A определяется формулой

$$P(A)= \sum_{\omega_k \in A} P(\omega_k), \qquad (1.2)$$

где суммирование ведётся по всем элементарным событиям, являющимся частными случаями события A (т.е. влекущими A).

Пример. Из урны, в которой пять белых шаров и три чёрных, вынимается шар. Для этого эксперимента элементарным событием является выборка какого-то шара. Всего в этом эксперименте восемь элементарных событий. Сложными событиями здесь будут: выборка белого шара и выборка чёрного шара, поскольку они являются суммой, соответственно, пяти и трёх элементарных событий.

Если элементарные события равновозможны, то формула (1.2) может быть записана следующим образом:

$$P(A) = \frac{m}{n}, \qquad (1.3)$$

где m – число благоприятных исходов (элементарных событий, влекущих событие A), n – общее число исходов (элементарных событий).

Формула (1.3) выражает так называемое классическое определение вероятности, согласно которому вероятность события A есть отношение благоприятного числа исходов к общему числу исходов эксперимента.

Задача 1.5. В ящике находятся 3 зеленых, 2 желтых и 5 белых шаров. Какова вероятность того, что случайным образом вытянутый шар окажется зеленым?

Решение. Общее количество исходов соответствует общему количеству шаров. Следовательно, $n = 10$. Благоприятным исходом является вытянутый зеленый шар. Таком образом, получаем, что $m = 3$. По определению вероятности события:

$$P(A) = \frac{m}{n} = \frac{3}{10} = 0.3.$$

Задача 1.6. В партии из десяти деталей три детали бракованные. Из партии выбираются три детали. Какова вероятность того, что две из них будут стандартными, а одна бракованной?

Решение. Обозначим через A событие – две детали стандартные, а одна бракованная. Общее количество исходов n равно числу сочетаний C_{10}^3, поскольку для нашей выборки важен только состав элементов. Число благоприятных исходов m равно произведению числа способов выбора двух стандартных деталей из семи C_7^2 на число способов выбора одной дефектной из трёх C_3^1. Окончательно получаем:

$$P(A) = \frac{m}{n} = \frac{C_7^2 C_3^1}{C_{10}^3} = \frac{7!\cdot 3!}{2!\cdot 5!\cdot 1!\cdot 2!} : \frac{10!}{3!\cdot 7!} = \frac{21}{40}.$$

Задача 1.7. Из колоды карт в 36 листов выбираются три карты. Какова вероятность того, что две из них будут одной масти, а одна - другой.

Решение. Обозначим искомую вероятность через P. В колоде карт четыре масти – пики, трефы, бубны и черви, а в каждой масти девять карт. Общее число различных выборок трёх карт из колоды вычисляется по формуле числа сочетаний (выборка неупорядоченная)

$$n = C_{36}^{3} = \frac{36!}{3!\cdot 33!} = 7140.$$

Для подсчёта благоприятных исходов m заметим, что процесс выбора карт можно представить в виде двух этапов. Сначала выбираются две масти (первый этап), а затем выбираются карты из этих мастей (второй этап). Число различных выборок двух мастей из четырёх вычисляется по формуле числа размещений, так как число карт первой масти не равно числу карт второй масти. Действительно, наборы карт (♠;♠;♦) и (♦;♦;♠) будут различны, хотя выбраны одинаковые масти.

Когда масти выбраны, то число различных выборок карт этих мастей равно числу выборок карт первой масти, умноженному на число выборок карт второй масти. Поскольку здесь не важен порядок выбора, то эти числа находятся по формуле числа сочетаний. Таким образом, число благоприятных исходов m равно:

$$m = A_4^2 C_9^2 C_9^1 = \frac{4!}{2!}\cdot\frac{9!}{2!\cdot 7!}\cdot\frac{9!}{1!\cdot 8!} = 3888.$$

Следовательно, вероятность выбора двух карт одной масти и одной карты другой масти равна:

$$P = \frac{m}{n} = \frac{3888}{7140} \approx 0{,}545.$$

2. Основные формулы теории вероятностей

2.1. Формулы сложения и умножения вероятностей

Для любых двух событий A и B вероятность их суммы равна:

$$P(A+B) = P(A) + P(B) - P(AB). \qquad (2.1)$$

Эта формула называется формулой сложения вероятностей, она обобщает аксиому сложения вероятностей, сформулированную в п. 1.1 для несовместных событий. Следует отметить, что если события A и B независимы, то формула (2.1) примет вид (2. 2):

$$P(A+B) = P(A) + P(B). \qquad (2.2)$$

Во многих случаях приходится определять вероятность события, когда известно, что другое событие произошло. Вероятность события A, вычисленная при условии, что событие B наступило, называется условной вероятностью события A относительно события B и обозначается $P(A|B)$. Условная вероятность определяется по формуле

$$P(A \mid B) = \frac{P(AB)}{P(B)}. \qquad (2.3)$$

Приведем еще одно определение независимости событий. События A и B называются независимыми, если вероятность наступления одного из них не зависит от наступления другого, то есть

$$P(A \mid B) = P(A)$$

и

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

Из формулы (2.3) следует формула вычисления вероятности произведения событий A и B:

$$P(AB) = P(A \mid B)\,P(B), \tag{2.4}$$

которая называется формулой умножения вероятностей. Для независимых событий A и B она принимает более простой вид:

$$P(AB) = P(A)P(B). \tag{2.5}$$

Рассмотрим задачи, при решении которых используются формулы сложения и умножения вероятностей.

Задача 2.1. Два стрелка поражают мишень с вероятностью 0,6 и 0,8 соответственно. Они произвели залп по мишени. Требуется определить вероятность того, что мишень будет поражена.

Решение.

Способ 1. Обозначим через A_1 и A_2 – попадание в мишень первым и вторым стрелками соответственно, а через B – поражение мишени. Мишень будет поражена, если в неё попадёт хотя бы один стрелок, поэтому событие B равносильно сумме $A_1 + A_2$. Тогда, учитывая формулы сложения и умножения вероятностей, а также то, что события A_1 и A_2 независимы (меткость – это внутреннее свойство стрелка), получим

$$P(B) = P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1)P(A_2) =$$
$$= 0{,}6 + 0{,}8 - 0{,}6 \cdot 0{,}8 = 0{,}92.$$

Способ 2. Событие – мишень поражена – противоположно событию – оба стрелка промахнулись. Поэтому

$$P(B) = 1 - P(\overline{B}) = 1 - P(\overline{A_1\,A_2}) = 1 - P(\overline{A_1})P(\overline{A_2}) = 1 - 0{,}4 \cdot 0{,}2 = 0{,}92.$$

Задача 2.2. Из колоды карт в 36 листов последовательно выбираются две карты. Первая выбранная карта оказалась красной масти. Определить вероятность того, что и вторая карта будет красной масти.

Решение. Пусть A и B – появление карты красной масти при первой и второй выборках соответственно. В колоде 18 карт красной масти и 18

карт чёрной масти. Так как при первой выборке произошло событие A, то состав колоды изменился: в колоде стало 17 карт красной масти и 18 карт чёрной. Вычисляя $P(B \mid A)$ как отношение благоприятного числа исходов к общему, получим

$$P(B \mid A) = \frac{17}{35}.$$

Задача 2.3. В урне 3 белых и 5 чёрных шаров; из урны последовательно вынули два шара, которые оказались разноцветными. Какова вероятность, что первым был вынут белый шар?

Решение. Введём обозначения для событий:

A_i – появление белого шара при i-й выборке;

B_i – появление чёрного шара при i-й выборке;

C – вынуты разноцветные шары.

В задаче требуется найти $P(A_1 \mid C)$, поэтому, исходя из определения условной вероятности, определим $P(C)$ и $P(AC)$.

Событие – выборка разноцветных шаров – есть сумма двух несовместных событий: первый вынутый шар белый, а второй – чёрный; первый вынутый шар чёрный, а второй – белый. Поэтому

$$P(C) = P(A_1B_2) + P(B_1A_2).$$

Из этого равенства, учитывая формулу умножения вероятностей (2.4), получим

$$P(C) = P(B_2 \mid A_1)P(A_1) + P(A_2 \mid B_1)P(B_1).$$

По условию в урне 3 белых и 5 чёрных шаров, а их общее число равно 8. Вероятности событий A_1 и B_1 определяем как отношение благоприятного числа исходов к общему:

$$P(A_1) = \frac{3}{8}$$

и

$$P(B_1) = \frac{5}{8}.$$

Если произошло событие A_1 или B_1, то состав шаров в урне изменился. Условные вероятности событий B_2 и A_2 находим аналогично:

$$P(B_2 \mid A_1) = \frac{5}{7}$$

и

$$P(A_2 \mid B_1) = \frac{3}{7}.$$

Подставляя полученные значения, вычислим вероятность события C:

$$P(C) = \frac{5}{7} \cdot \frac{3}{8} + \frac{3}{7} \cdot \frac{5}{8} = \frac{15}{28}.$$

Для того, чтобы определить вероятность $P(A_1C)$, заметим, что события A_1 и C происходят совместно только в том случае, когда происходит событие B_2. Это означает, что

$$A_1C = A_1B_2$$

и, следовательно,

$$P(A_1C) = P(A_1B_2) = P(B_2 \mid A_1)\, P(A_1).$$

Вероятности $P(B_2 \mid A_1)$ и $P(A_1)$ были вычислены ранее, поэтому

$$P(A_1C) = \frac{5}{7} \cdot \frac{3}{8} = \frac{15}{56}$$

и, окончательно,

$$P(A_1 \mid C) = \frac{15}{56} : \frac{15}{28} = \frac{1}{2}.$$

В данном примере вероятность $P(C)$ можно определить проще: как отношение благоприятного числа исходов к общему:

$$P(C) = \frac{C_3^1 \cdot C_5^1}{C_8^2} = \frac{3 \cdot 5}{28} = \frac{15}{28}.$$

Задача 2.4. Для сигнализации об аварии установлены три независимо работающих устройства. Вероятность того, что при аварии сработает первое устройство равна 0,85, второе - 0,92, третье – 0,75. Найти вероятность того, что при аварии сработает : а) только одно устройство; б) только два устройства; в) все три устройства.

Решение: Пусть А – сработало первое устройство, В – второе, С – третье. По формулам умножения (2.5) и сложения (2.1) вероятностей имеем:

а)

$$P(A\overline{B}\overline{C} + \overline{A}B\overline{C} + \overline{A}\overline{B}C) = P(A\overline{B}\overline{C}) + P(\overline{A}B\overline{C}) + P(\overline{A}\overline{B}C) =$$
$$0{,}85 \cdot (1-0{,}92)(1-0{,}75) + (1-0{,}85) \cdot 0{,}92 \cdot (1-0{,}75) + (1-0{,}85) \cdot$$
$$(1-0{,}92) \cdot 0{,}75 = 0{,}1835 \quad ;$$

б)

$$P(AB\overline{C} + A\overline{B}C + \overline{A}BC) = 0{,}85 \cdot 0{,}92 \cdot 0{,}25 + 0{,}85 \cdot 0{,}08 \cdot 0{,}75 +$$
$$0{,}15 \cdot 0{,}92 \cdot 0{,}75 = 0{,}247 \quad ;$$

в) $P(ABC) = 0{,}85 \cdot 0{,}92 \cdot 0{,}75 = 0{,}586$.

2.2. Формула полной вероятности. Формулы Байеса

События A_1, $A_2, \ldots A_n$ образуют полную группу событий, если выполнены два условия: они попарно несовместны и сумма всех событий является достоверным событием. Если события A_1, $A_2, \ldots, A_n$ образуют полную группу, то вероятность любого события B можно определить по формуле

$$P(B) = P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \ldots + P(B \mid A_n)P(A_n), \quad (2.6)$$

которая называется формулой полной вероятности.

Задача 2.5. Рабочие изготавливают однотипные детали. Вероятность брака у первого рабочего – 0,01, а у второго – 0,05. Из партии в пятьсот деталей, двести из которых сделаны первым рабочим, а триста – вторым, выбирается одна деталь. Какова вероятность того, что эта деталь бракованная?

Решение. Рассматриваемая ситуация является двухэтапной: первый этап – изготовление деталей рабочими, а второй – выбор детали из партии. Пусть событие A_i – изготовление детали i-м рабочим, а событие B – выбор бракованной детали. Заметим, что любая деталь изготавливается только одним рабочим, поэтому события A_1 и A_2 несовместны. Кроме того, каждая деталь из партии изготавливается только этими рабочими и, следовательно, сумма событий A_1 и A_2 есть достоверное событие. Это означает, что события A_1 и A_2 образуют полную группу событий и для определения вероятности события B можно использовать формулу полной вероятности

$$P(B) = P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2).$$

Определим значения вероятностей, входящих в правую часть этого равенства. Так как общее число деталей пятьсот, а двести и триста деталей изготовлены первым и вторым рабочим соответственно, то вероятности $P(A_1)$ и $P(A_2)$ равны:

$$P(A_1) = \frac{200}{500} = \frac{2}{5}$$

и

$$P(A_2) = \frac{300}{500} = \frac{3}{5}.$$

По условию, если деталь изготовлена первым рабочим, то вероятность брака равна 0,01, а если вторым, то 0,05. Поэтому

$$P(B \mid A_1) = 0{,}01$$

и

$$P(B \mid A_2) = 0{,}05.$$

Подставляя полученные значения в формулу полной вероятности, получим

$$P(B) = 0{,}01 \cdot \frac{2}{5} + 0{,}05 \cdot \frac{3}{5} = \frac{17}{500} = 0{,}034.$$

Задача 2.6. У рыбака два излюбленных места лова, которые он посещает с вероятностями 0,2 и 0,6 соответственно. На первом месте лова он вылавливает 2 кг рыбы с вероятностью 0,7, а на втором с вероятностью 0,4. С рыбалки он возвращается с уловом в 2 кг. Какова вероятность того, что рыбак ловил рыбу на первом месте?

Решение. Пусть A_i обозначает событие – рыбак ловил рыбу на i-м месте, а B – улов рыбы 2 кг. Из условия задачи имеем

$$P(A_1) = 0{,}2;\ P(A_2) = 0{,}6;\ P(B \mid A_1) = 0{,}7;\ P(B \mid A_2) = 0{,}4.$$

Вероятность $P(B)$ вычисляем по формуле полной вероятности

$$P(B) = 0{,}7 \cdot 0{,}2 + 0{,}4 \cdot 0{,}6 = 0{,}38,$$

а искомую вероятность $P(A_1 \mid B)$ по формулам Байеса

$$P(A_1 \mid B) = \frac{0{,}7 \cdot 0{,}2}{0{,}38} = \frac{0{,}14}{0{,}38} = \frac{7}{19} \approx 0{,}37.$$

2.3. Последовательность независимых испытаний. Формулы Бернулли, Муавра-Лапласа, Пуассона

Большой круг задач практического характера может быть сведён к теоретико-вероятностной модели так называемой последовательности независимых испытаний. Суть этой модели состоит в следующем. Рассматривается n одинаковых и независимых друг от друга опытов, в которых наблюдается событие A. Вероятность наступления события A одна и та же для каждого опыта. В этих условиях ставятся две основные задачи: определение вероятности наступления события A ровно k раз в n опытах; определение вероятности того, что событие A произойдёт не менее k_1 раз и не более k_2 раз в n опытах. Соответствующие вероятности обозначаются через $P_n(k)$ и $P_n(k_1;\, k_2)$.

Решение первой задачи может быть получено по формуле Бернулли

$$P_n(k) = C_n^k p^k q^{n-k}, \qquad (2.7)$$

где p – вероятность наступления события A в одном опыте, а $q = 1 - p$ является вероятностью события $\overline{A}$.

Однако применение формулы Бернулли для больших n затруднительно ввиду сложности вычислений. Поэтому для решения первой задачи при больших n используют приближённые формулы подсчёта $P_n(k)$ – локальную формулу Муавра-Лапласа и формулу Пуассона.

Согласно локальной формуле Муавра-Лапласа, величина $P_n(k)$ приближенно равна

$$\mathrm{P_n(k)} \approx \frac{1}{\sqrt{\mathrm{npq}}}\varphi(\mathrm{x}), \qquad (2.8)$$

где $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x = \dfrac{k - np}{\sqrt{npq}}$.

При этом следует отметить, что чем больше количество испытаний, тем точнее полученный результат. Таблица значений функции $\varphi(x)$ на промежутке [0,4] приведена в Приложении 1, для определения значений $\varphi(x)$ на промежутке [-4,0] следует использовать четность функции $\varphi(x)$.

Существует еще одна постановка этой задачи. Нас будет интересовать вероятность того, что в n опытах интересующее событие произойдет не менее k_1 раз и не более k_2. При решении данной задачи также используются приближенные методы. Вероятность $P_n(k_1;k_2)$ согласно интегральной формуле Муавра-Лапласа, равна:

$$P_n(k_1;k_2) = \Phi(x_2) - \Phi(x_1), \qquad (2.9)$$

где $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}} dt$ – функция Лапласа,

$$x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}.$$

Таблица значений функции Лапласа на промежутке [0,5] приведена в Приложении 2. Для значений $x \geq 5$ $\Phi(x)$ полагают равной 0,5; для определения функции Лапласа при отрицательных x используют нечетность функции $\Phi(x)$.

Задача 2.7. Бросается пять раз игральная кость. Какова вероятность, что в этой серии бросков выпадет: первое – ровно две тройки; второе – не менее четырех троек; третье – хотя бы одна тройка?

Решение. Событие A – выпадение тройки, вероятность этого события $P(A) = p = \frac{1}{6}$. Вероятность выпадения двух троек определяем по формуле (2.7), для $n=5$ и $k=2$ находим

$$P_5(2)=C_5^2\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3=\frac{1250}{7776}\approx 0{,}1608.$$

Вероятность выпадения не менее четырех троек определяем по формуле (2.8), имеем

$$P_5(4;5)=P_5(4)+P_5(5)\approx 0{,}0032+0{,}0001=0{,}0033.$$

Вероятность выпадения хотя бы одной тройки также можно вычислить по формуле (2.8), но можно определить ее проще, переходя к противоположному событию – не выпадению ни одной тройки. Тогда

$$P_5(0)=C_5^0\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^5=\frac{3125}{7776}\approx 0{,}4019,$$

откуда искомая вероятность

$$P_5(1;5)=1-P_5(0)\approx 0{,}5981.$$

Задача 2.8. Стрелок поражает мишень с вероятностью 0,8. Он произвел серию из ста выстрелов. Требуется определить вероятности следующих событий: первое – попадание в мишень равно 75 раз; второе – попадание в мишень не менее 70 раз и не более 80 раз; третье – попадание в мишень не более 70 раз.

Решение. Поскольку число испытаний n достаточно велико, а произведение np существенно больше единицы, то для ответа на первый вопрос используем локальную формулу Муавра-Лапласа (2.9). Имеем n=100, k=75, p=0,8 и q=0,2, тогда

$$x=\frac{75-80}{\sqrt{100\cdot 0{,}8\cdot 0{,}2}}=-\frac{5}{4}=-1{,}25\ .$$

Учитывая, что функция $\varphi(x)$ четная, т.е. $\varphi(x)=\varphi(-x)$, по таблице (Приложение 1) находим значение $\varphi(1{,}25)=0{,}1826$, а затем искомую вероятность

$$P_{100}(75)\approx\frac{1}{4}\cdot 0{,}1826\approx 0{,}0457\ .$$

Для решения второго вопроса используем интегральную формулу Муавра- Лапласа (2.11). Имеем n=100, k_1=70, k_2=80, p=0,8 и q=0,2. Вычисляем значения x_1 и x_2

$$x_1=\frac{70-80}{\sqrt{100\cdot 0{,}8\cdot 0{,}2}}=-\frac{10}{4}=-2{,}5,\quad x_2=0.$$

По таблице (Приложение 2) находим значения $\Phi(0)=0$ и $\Phi(2{,}5)=0{,}4938$.Так как функция Лапласа нечетная, то $\Phi(-2{,}5)=-\ 0{,}4938$. Искомая вероятность равна:

$$P_{100}(70;80)\approx 0-(-0{,}4938)=0{,}4938.$$

Аналогично решается и третий пункт задачи. Имеем

$$x_1=-20,\ x_2=-2{,}5;$$

$$\Phi(-20)=-0{,}5,\ \Phi(-2{,}5)=-0{,}4938.$$

Искомая вероятность равна

$$P_{100}(0;70)\approx\Phi(-2{,}5)-\Phi(-20)\approx -0{,}4938+0{,}5=0{,}0062\ .$$

3. Случайные величины

3.1. Закон распределения случайной величины

Случайной величиной называется такая величина, которая в результате теоретико-вероятностного эксперимента может принимать то или иное значение, причем неизвестно заранее, какое именно.

Среди случайных величин выделяются два основных типа – дискретные и непрерывные случайные величины. Случайная величина называется дискретной, если она принимает изолированные значения, и непрерывной, если ее значения сплошь заполняют некоторый промежуток <a,b>. Например, номера выигрышных лотерейных билетов являются дискретной случайной величиной, а время ожидания автобуса на остановке – непрерывной.

С теоретико-вероятностной точки зрения, случайная величина полностью определена, если известна вероятность, с которой она принимает значения из любого заданного интервала. Соотношения, определяющие эти вероятности, называются законом распределения случайной величины. Закон распределения случайной величины может иметь различные формы: ряд распределения, многоугольник распределения, функция распределения, плотность вероятности и т.п.

Рядом распределения дискретной случайной величины X называется таблица, в которой приведены все возможные значения случайной величины x_1, x_2, x_3, … и соответствующие им вероятности $p_i = P\{X=x_i\}$.

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

Отметим, что $\sum\limits_i p_i = 1$. Вероятность принятия случайной величиной значений из данного промежутка $<a,b>$ определяется по формуле

$$P\{X \in <a,b>\} = \sum_{x_i \in <a,b>} p_i \,. \qquad (3.1)$$

Графическое изображение ряда распределения называется многоугольником распределения (рис. 3.1).

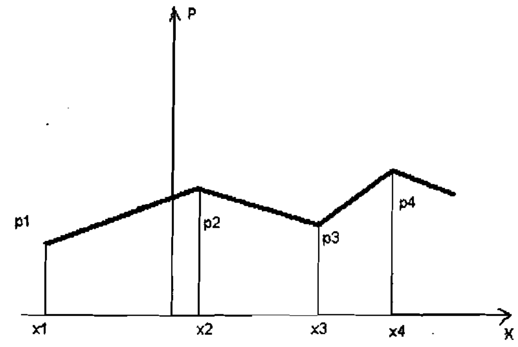


Рис. 3.1

Следует отметить, что приведенные формы закона распределения применимы лишь к дискретным случайным величинам. Однако имеется универсальная форма закона распределения, применимая для описания случайных величин обоих типов – функция распределения.

3.2. Функция распределения и плотность распределения случайной величины

Функцией распределения случайной величины называется функция $F(x)$, определяющая вероятность того, что X примет значение меньше x

$$F(x)=P\{X < x\}. \qquad (3.2)$$

Из определения $F(x)$ следует, что она принимает значения в промежутке от нуля до единицы. Перечислим основные свойства функции распределения: $F(x)$ является неубывающей функцией; вероятность принятия случайной величиной X значений из промежутка $[a,b)$ определяется формулой

$$P\{a \le X < b\} = F(b) - F(a). \qquad (3.3)$$

Задача 3.1. По ряду распределения случайной величины X

X	0	2	3	6
p	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$

найти функцию распределения, построить ее график и определить вероятности принятия случайной величиной значений в промежутках (2,5), (-1,0) и (3,8).

Решение. Пусть $x \le 0$, тогда событие $X \in (-\infty, x)$ является невозможным, так как случайная величина X не принимает отрицательных значений и поэтому $F(x)$=0. Если $0 < x \le 2$, то в промежутках $(-\infty, x)$ имеется только одно значение случайной величины, а именно нулевое. Поэтому, учитывая формулу (3.1), имеем

$$F(x) = P\{X < x\} = p_1 = \frac{1}{8}.$$

Аналогично определяются значения функции распределения на промежутках: (2, 3], (3, 6], (6, $+\infty$). В итоге получим :

$$F(x) = \begin{cases} 0, & x \le 0; \\ \frac{1}{8}, & 0 < x \le 2; \\ \frac{3}{8}, & 2 < x \le 3; \\ \frac{3}{8}, & 3 < x \le 6; \\ 1, & 6 < x. \end{cases}$$

График функции распределения приведен на рис. 3.2.

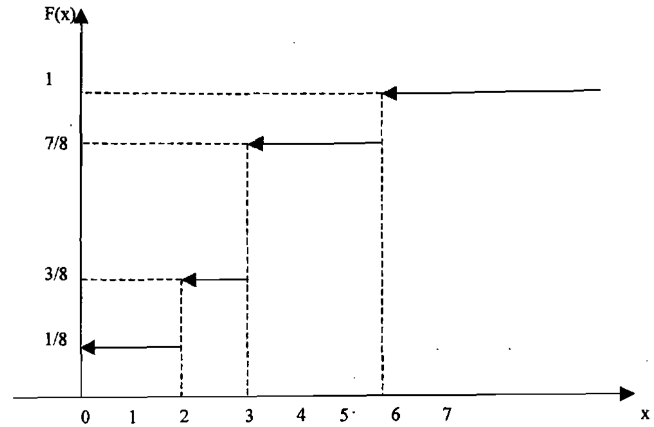


Рис. 3.2

Плотностью распределения случайной величины X называется неотрицательная функция *f(x)*, удовлетворяющая равенству

$$F(x) = \int_{-\infty}^{x} f(t)dt, \tag{3.4}$$

где $F(x)$ – функция распределения.

Если $F(x)$ дифференцируемая функция, то $F'(x)=f(x)$, поэтому плотность называют также дифференциальной функцией распределения. Основные свойства плотности:

1. $f(x) \ge 0$;
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$; (3.5)
3. $P\{a < X < b\} = \int_{a}^{b} f(x)dx$.

Заметим, что в соотношении 3 любой знак строгого неравенства можно заменить на знак нестрогого неравенства.

3.3. Числовые характеристики случайных величин

Математическим ожиданием или средним значением случайной величины X называется число, определенное равенствами:

$$M[X] = \sum_i x_i p_i ,$$

для дискретных случайных величин;

$$M[X] = \int_{-\infty}^{+\infty} x f(x) dx , \qquad (3.6)$$

для непрерывных случайных величин.

Математическое ожидание – это такое значение, вокруг которого группируются значения случайной величины.

Перечислим основные свойства математического ожидания:

- математическое ожидание постоянной есть сама постоянная

$$M[C] = C;$$

- постоянный множитель можно выносить за знак математического ожидания

$$M[CX] = CM[X];$$

- математическое ожидание суммы двух независимых случайных величин равно сумме математических ожиданий этих величин

$$M[X + Y] = M[X] + M[Y];$$

- математическое ожидание произведения двух независимых случайных величин равно произведению математических ожиданий этих величин

$$M[XY] = M[X]M[Y].$$

Разность между случайной величиной и математическим ожиданием X - $M[X]$ называется отклонением случайной величины. Математическое

ожидание квадрата отклонения случайной величины называется дисперсией. Дисперсия характеризует степень концентрации значений случайной величины около среднего значения. Чем меньше дисперсия, тем больше значений случайной величины появляется вблизи математического ожидания и наоборот. Квадратный корень из дисперсии называется среднеквадратичным отклонением $\sigma[X] = \sqrt{D}[x]$. Дисперсия обозначается либо $D[X]$, либо $\sigma^2[X]$.

Основные свойства дисперсии:

- дисперсия постоянной равна нулю $D[X] = 0$;
- постоянный множитель можно выносить за знак дисперсии, возводя его в квадрат $D[CX] = C^2D[X]$;
- дисперсия неотрицательна $D[X] \geq 0$;
- дисперсия суммы независимых случайных величин равна сумме дисперсий этих величин $D\ [X+Y] = D[X] + D[Y]$.

Приведем расчетные формулы для вычисления дисперсии:

$$D[X] = \sum_i (x_i - m)^2 p_i ,$$

для дискретных случайных величин;

$$D[X] = \int_{-\infty}^{+\infty} (x - m)^2 f(x) dx , \qquad (3.7)$$

для непрерывных случайных величин, где m – математическое ожидание случайной величины X.

На практике удобнее пользоваться следующей формулой для вычисления дисперсии:

$$D[X] = M[X^2] - (M[X])^2. \qquad (3.8)$$

Задача 3.2. Для дискретной случайной величины X, заданной рядом распределения

X	0	3	4	8
P	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

определить математическое ожидание, дисперсию и среднеквадратичное отклонение.

Решение. По формулам (3.6) имеем

$$M[X]=0\cdot\frac{1}{2}+3\cdot\frac{1}{2}+4\cdot\frac{1}{8}+8\cdot\frac{1}{8}=3.$$

Для определения дисперсии воспользуемся формулой (3.8), тогда

$$D[X]=0\cdot\frac{1}{2}+9\cdot\frac{1}{2}+16\cdot\frac{1}{8}+64\cdot\frac{1}{8}-(3)^2=5{,}5$$

и

$$\sigma[X]=\sqrt{5{,}5}\approx 2{,}345.$$

Задача 3.3. Определить математическое ожидание, дисперсию и среднеквадратичное отклонение случайной величины X, плотность распределения которой

$$f(x)=\begin{cases}0, & x<0;\\ 2e^{-2x}, & x\ge 0.\end{cases}$$

Решение. По формулам (3.6) имеем

$$M[X]=\int_{-\infty}^{+\infty}xf(x)dx=\int_{0}^{+\infty}x2e^{-2x}dx=-(xe^{-2x}+\frac{1}{2}e^{-2x})\Big|_0^{\infty}=\frac{1}{2}.$$

Для определения дисперсии используем формулу (3.8):

$$D[X]=\int_{-\infty}^{+\infty}x^2f(x)dx-(M[X])^2=\int_0^{\infty}x^2 2e^{-2x}dx-\frac{1}{4}=$$

$$=-(x^2e^{-2x}+xe^{-2x}+\frac{1}{2}e^{-2x})\Big|_0^{\infty}-\frac{1}{4}=\frac{1}{2}-\frac{1}{4}=\frac{1}{4};$$

$$\sigma[X]=\sqrt{\frac{1}{4}}=\frac{1}{2}.$$

Задача 3.4. Дискретная случайная величина X может принимать только два значения x_1 и x_2, причём $x_1<x_2$. Известны – вероятность p_1 значения x_1, математическое ожидание $M(X)$ и дисперсия $D(X)$. Найти закон распределения этой случайной величины, если $p_1=0{,}95, M(X)=3{,}2, D(X)=0{,}1$.

Решение. $M(X)=x_1\cdot 0{,}95+x_2\cdot 0{,}05=3{,}2$;

$$D(X)=(x_1-3{,}2)^2\cdot 0{,}95+(x_2-3{,}2)^2\cdot 0{,}05=0{,}1.$$

Имеем два уравнения для нахождения неизвестных x_1 и x_2. Решая эту систему уравнений, определяем $x_1=3{,}128, x_2=3{,}564$. Таким образом, закон распределения случайной величины X имеет вид:

X	3,128	3,564
P	0,95	0,05

Задача 3.5. Случайная величина X задана функцией распределения $F(x)$. Найти плотность распределения $f(x)$, математическое ожидание $M(X)$ и дисперсию $D(X)$ случайной величины X.

$$F(x)=\begin{cases}0, & x\le 3,\\ \frac{x}{3}-1, & 3<x\le 6,\\ 1, & x>6.\end{cases}$$

Решение. Согласно формуле (3.4) имеем: $f(x)=F'(x)$, то есть

$$f(x)=\begin{cases}0, & x\le 3,\\ \dfrac{1}{3}, & 3<x\le 6,\\ 0, & x>6.\end{cases}$$

Из формул (3.6) и (3.7)

$$M(X)=\int_3^6 x\cdot f(x)dx=\int_3^6 \frac{x}{3}dx=\frac{1}{3}\cdot\frac{x^2}{2}\Big|_3^6=\frac{1}{6}(36-9)=\frac{9}{2};$$

$$D(X)=\int_{-\infty}^{\infty}\left(x-\frac{9}{2}\right)^2 f(x)dx=\int_3^6\left(x-\frac{9}{2}\right)^2\frac{1}{3}dx=\frac{3}{4}.$$

3.4. Нормальное распределение

Случайная величина X распределена по нормальному закону с параметрами a и σ, если ее плотность $f(x)$ равна

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}. \tag{3.9}$$

Вероятностный смысл параметров распределения следующий: a-это математическое ожидание, σ - среднеквадратичное отклонение случайной величины.

Замена переменной

$$Y=\frac{X-a}{\sigma} \tag{3.10}$$

позволяет любую случайную величину X, распределенную по нормальному закону с параметрами a и σ, привести к случайной величине Y, распределенной по так называемому стандартному нормальному закону с параметрами 0 и 1. Для стандартного нормального закона распределения плотность равна

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \tag{3.11}$$

а функция распределения

$$F(x)=\frac{1}{2}+\Phi(x), \tag{3.12}$$

где $\Phi(x)$ – функция Лапласа:

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{z^2}{2}}dz. \tag{3.13}$$

Задача 3.6. На станке изготавливают деталь. Длина детали X является случайной величиной с параметрами a=20 см и σ =0,2 см. Найти вероятность того, что длина детали будет заключена между 19,7 см и 20,3 см, т.е. отклонение в ту или иную сторону не превысит 0,3 см.

Решение. Отклонение $|X - a|<\varepsilon$ случайной величины X равносильно отклонению $|Y|<\frac{\varepsilon}{\sigma}$ случайной величины Y, распределенной по стандартному нормальному закону (см. (3.10)). Для нашего случая $-\frac{3}{2}<Y<\frac{3}{2}$. Вероятность попадания случайной величины Y в интервал $(-\frac{3}{2},\frac{3}{2})$ равна

:

$$P\{-\frac{3}{2}<Y<\frac{3}{2}\}=F(\frac{3}{2})-F(-\frac{3}{2})=\Phi(\frac{3}{2})-\Phi(-\frac{3}{2})=2\Phi(\frac{3}{2}).$$

Определив по таблице Приложения П2 значение $\Phi(1,5)$, получим

$$P\{|x-20|<0,3\}=P\{|Y|<1,5\}=2\cdot 0,4332=0,8664.$$

Задача 3.7. Случайная величина X распределена по нормальному закону с параметрами 3 и 5. Определить вероятность ее попадания в интервал (-1,2).

Решение. Неравенство $-1<X<2$ для случайной величины X равносильно неравенству $\frac{-1-a}{\sigma}<Y<\frac{2-a}{\sigma}$ для стандартно распределенной случайной величины Y (3.10). Поэтому, учитывая, что $a=3, \sigma=5$, имеем

$$P\{-1<X<2\}=P\{-0{,}8<Y<-0{,}2\}=\Phi(-0{,}2)-\Phi(-0{,}8)=$$
$$=-0{,}5793-(-0{,}7881)=0{,}2088\ .$$

4. Элементы математической статистики

4.1. Основные понятия

К основным задачам математической статистики относятся:

1. Нахождение по опытным (статистическим) данным закона распределения наблюдаемой случайной величины – статистического закона распределения.
2. Нахождение по статистическим данным параметров распределения наблюдаемой случайной величины – точечное и интервальное оценивание параметров распределения.
3. Статистическая проверка гипотезы о теоретическом законе распределения наблюдаемой случайной величины.

Первая и вторая задачи решаются посредством выборки большого и малого объёма. Для статистической проверки гипотезы о законе распределения (задача 3) используются специальные критерии (Пирсона, Колмогорова и т.п.).

4.2. Статистические характеристики опытных данных

Простой статистический ряд. Это – первичная форма записи статистического материала. Например, случайная величина X при 15 наблюдениях приняла следующие значения:

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X_i	-20	-60	-10	30	60	70	-10	-30	120	-100	-80	20	40	-60	-10

Простой статистический ряд может быть обработан различными способами.

Статистическая функция (закон) распределения. Так называется частота события $X<x$ в данном статистическом материале

$$F^*(x)=P^*(X<x)\ .$$

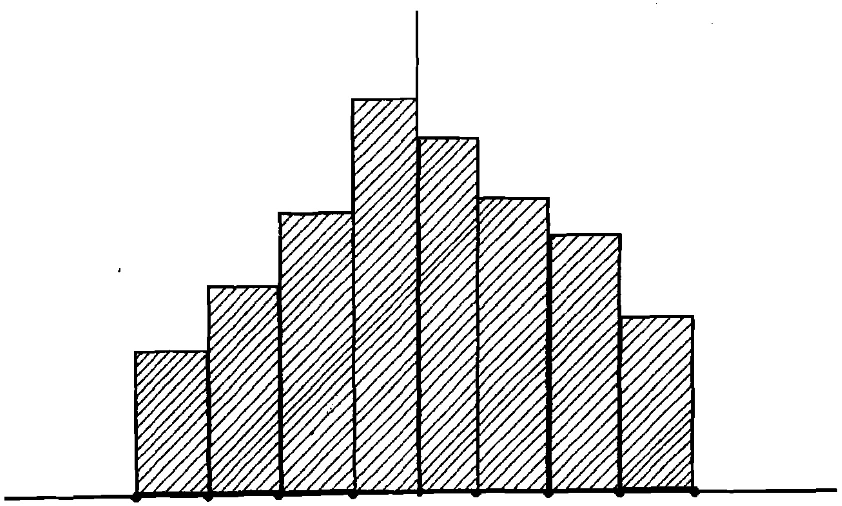
Так, для рассмотренного простого статистического ряда при $X=x=-100$ будет $F^*(-100)=0$; когда $-100<x<-80$ будет $F(x)=1/15$; в точке $x=-80$ $F^*(x)$ возрастёт ещё на 1/15; в точке $x=-60$ $F^*(x)$ возрастает на 2/15 и так далее до $F^*(x)=1$ при $x>120$. Таким образом, $F^*(x)$ терпит разрывы в точках $x=-100$, $x=-80$, $x=-60$ и т.д. График её – ступенчатая линия, расположенная между $F^*(x)=0$, $x<-100$, и $F^*(x)=1$ при $x>120$.

Всё множество значений, наблюдамых в результате опыта, называется ***генеральной совокупностью***, а число наблюдений – её ***объёмом*** (n=15 –объём в рассмотренном примере).

При большом числе наблюдений удобнее пользоваться некоторыми усреднёнными характеристиками. Разделим весь диапазон наблюдаемых значений на интервалы ***(разряды)*** и подсчитаем количество значений m_i, приходящихся на каждый ι-й разряд; это число разделим на общее число наблюдений **n** и найдём частоту, соответствующую данному разряду $P_i^* = m_i / n$ (сумма частот равна 1). Составим таблицу: в первой строке приводим разряды, во второй – количество значений в данном разряде, в третьей – соответствующие частоты; эта таблица называется ***статистическим рядом.*** Пусть $n = 500$, I_ι = разряд:

ι	1	2	3	4	5	6	7	8
I_ι	-4; -3	-3; -2	-2; -1	-1; 0	0; 1	1; 2	2; 3	3; 4
P_ι^*	0,012	0,05	0,144	0,266	0,24	0,176	0,092	0,02

График статистического ряда – *гистограмма*; в рассмотренном примере она имеет вид:



Гистограмма приближённо соответствует эмпирической плотности распределения непрерывной случайной величины X.

4.3. Числовые характеристики статистического распределения

Эмпирическим аналогом математического ожидания является среднее арифметическое наблюдённых значений случайной величины

$$M^*(X) = \frac{1}{n}\sum_{i=1}^{n} x_i = m_x^* \qquad (4.1)$$

- статистическое среднее случайной величины. При неограниченном увеличении числа опытов статистическое среднее приближается к математическому ожиданию. $M^*(X)$ является случайной величиной, которая тем не менее даёт некоторое представление о неизвестном значении $M(X)$.

Аналогично определяется ***статистическая дисперсия*** случайной величины

$$D^*(x) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - m_x^*\right)^2. \qquad (4.2)$$

При не заметно большом количестве опытов применяют следующий приём – воспользуемся статистическим рядом, полученным из той же генеральной совокупности. Тогда будем иметь

$$m_x^* = M^*(X) = \sum_{i=1}^{k} \tilde{x}_i p_i^* = \bar{x}\,, \qquad (4.3)$$

где $\bar{x}$ выборочная средняя , $\tilde{x}_i$ - "представитель" i-ого разряда (например, среднее арифметическое значение x для данного разряда – середина интервала (x_i, x_{i+1})); p_i^* - частота; k - количество разрядов. Аналогично

$$D^*(x)=\sum_{i=1}^{k}\left(\widetilde{x}_i-m_x^*\right)^2 P_i^*, \quad \sqrt{D^*(x)}=\sigma \,, \tag{4.4}$$

где $D^*(x)$ - выборочная дисперсия; σ - ***среднее квадратическое отклонение выборки.***

4.4. Оценки числовых характеристик

Из рассмотрения статистического материала можно сделать предположение о законе распределения случайной величины. Тогда возникает задача отыскания параметров, которыми определяется это распределение; например, для нормального закона необходимо найти величины a и σ. Такое приближённое значение параметра является случайной величиной и называется его ***оценкой.*** Пусть $\widetilde{a}$ - статистическая оценка неизвестного параметра a, выбранного (предполагаемого) закона распределения. Желательно, чтобы при увеличении числа испытаний n она приближалась, сходилась по вероятности к величине a. Оценка, обладающая таким свойством, называется ***состоятельной.***

Кроме того, необходимо, чтобы, заменяя величину a приближённо на $\widetilde{a}$, мы не делали систематической ошибки, т.е. чтобы выполнялось условие $M(\widetilde{a})=a$. Такая оценка называется ***несмещённой.***

Наконец, желательно, чтобы выбранная несмещённая оценка имела минимальное рассеяние около математического ожидания, т.е. обладала по сравнению с другими оценками минимальной дисперсией $D(\widetilde{a})=\min$. Такая оценка называется ***эффективной.***

Точечной называется оценка, которая определяется одним числом. При небольшом числе наблюдений точечная оценка неизвестного параметра может значительно отличаться от оцениваемой величины, т.е. приводить к грубым ошибкам. Поэтому в большинстве задач требуется не только найти

для параметра a подходящее численное значение, но и оценить его точность, надёжность, т.е. нужно установить, к каким ошибкам приведёт замена параметра a его оценкой $\widetilde{a}$, и с какой степенью уверенности можно утверждать, что эти ошибки не выйдут за данные пределы.

4.5. Доверительный интервал. Доверительная вероятность (надёжность)

Пусть для параметра a получена из опыта несмещённая оценка $\widetilde{a}$. Оценим возможную ошибку. Назначим некоторую, близкую к 1, вероятность (например, $\beta=0{,}9$; 0,95 или 0,99). Так что рассматриваемое случайное событие с вероятностью β можно считать практически достоверным. Найдём такое число $\varepsilon>0$, что будет

$$P\left(\left|\widetilde{a}-a\right|<\varepsilon\right)=\beta \,, \tag{4.5}$$

т.е. вероятность того, что оценка $\widetilde{a}$ отличается от истинного (теоретического) значения параметра a меньше, чем на ε, равна β. Тогда возможные ошибки практически не выходят из интервала $(-\varepsilon;\varepsilon)$. Большие ошибки могут появиться с малой вероятностью $\alpha=1-\beta$. Из (4.5) следует

$$P\left(\widetilde{a}-\varepsilon<a<\widetilde{a}+\varepsilon\right)=\beta \,, \tag{4.6}$$

что означает, что с вероятностью β параметр a попадает в интервал

$$I_\beta=\left(\widetilde{a}-\varepsilon,\widetilde{a}+\varepsilon\right). \tag{4.7}$$

Здесь параметр a неслучаен, а интервал I_β - случайный – случайно его положение на оси ОХ, определяемое центром $\widetilde{a}$; случайна и его длина 2ε. Величину β необходимо толковать как вероятность того, что случайный интервал I_β покрывает неизвестный параметр a. Величина β называется ***доверительной вероятностью*** или надёжностью, а интервал I_β - ***доверительным интервалом.*** Границы доверительного интервала $a_1 = \tilde{a} - \varepsilon$, $a_2 = \tilde{a} + \varepsilon$ называются ***доверительными границами.***

Библиографический список

Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. М.: Изд-во «Высшая школа», 1975.

Гмурман В.Е. Теория вероятностей и математической статистике. М.: Изд-во «Высшая школа», 1975.

ПРИЛОЖЕНИЯ

Приложение 1

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

X	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290

Окончание табл. П1

X	0	1	2	3	4	5	6	7	8	9
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0111	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Приложение 2

Таблица значений функции $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{z^2}{2}}dz$

X	$\Phi(x)$	X	$\Phi(x)$	X	$\Phi(x)$	X	$\Phi(x)$
0,00	0,0000	0,22	0,0871	0,44	0,1700	0,66	0,2454
0,01	0,0040	0,23	0,0910	0,45	0,1736	0,67	0,2486
0,02	0,0080	0,24	0,0948	0,46	0,1772	0,68	0,2517
0,03	0,0120	0,25	0,0987	0,47	0,1808	0,69	0,2549
0,04	0,0160	0,26	0,1026	0,48	0,1844	0,70	0,2580
0,05	0,0199	0,27	0,1064	0,49	0,1879	0,71	0,2611
0,06	0,0239	0,28	0,1103	0,50	0,1915	0,72	0,2642
0,07	0,0279	0,29	0,114	0,51	0,1950	0,73	0,2673
0,08	0,0319	0,30	0,1179	0,52	0,1985	0,74	0,2703
0,09	0,0359	0,31	0,1217	0,53	0,2019	0,75	0,2734
0,10	0,0398	0,32	0,1255	0,54	0,2054	0,76	0,2764
0,11	0,0438	0,33	0,1293	0,55	0,2088	0,77	0,2794
0,12	0,0478	0,34	0,1331	0,56	0,2123	0,78	0,2823
0,13	0,0517	0,35	0,1368	0,57	0,2157	0,79	0,2852
0,14	0,0557	0,36	0,1406	0,58	0,2190	0,80	0,2881
0,15	0,0596	0,37	0,1443	0,59	0,2224	0,81	0,2910
0,16	0,0636	0,38	0,1480	0,60	0,2257	0,82	0,2939
0,17	0,0675	0,39	0,1517	0,61	0,2291	0,83	0,2967
0,18	0,0714	0,40	0,1554	0,62	0,2324	0,84	0,2995
0,19	0,0753	0,41	0,1591	0,63	0,2357	0,85	0,3023
0,20	0,0793	0,42	0,1628	0,64	0,2389	0,86	0,3051
0,21	0,0832	0,43	0,1664	0,65	0,2422	0,87	0,3078

Продолжение табл. П2

X	$\Phi(x)$	X	$\Phi(x)$	X	$\Phi(x)$	X	$\Phi(x)$
0,88	0,3106	1,13	0,3708	1,38	0,4162	1,63	0,4484
0,89	0,3133	1,14	0,3729	1,39	0,4177	1,64	0,4495
0,90	0,3159	1,15	0,3749	1,40	0,4192	1,65	0,4505
0,91	0,3186	1,16	0,3770	1,41	0,4207	1,66	0,4515
0,92	0,3212	1,17	0,3790	1,42	0,4222	1,67	0,4525
0,93	0,3238	1,18	0,3810	1,43	0,4236	1,68	0,4535
0,94	0,3264	1,19	0,3830	1,44	0,4251	1,69	0,4545
0,95	0,3289	1,20	0,3849	1,45	0,4265	1,70	0,4554
0,96	0,3315	1,21	0,3869	1,46	0,4279	1,71	0,4564
0,97	0,3340	1,22	0,3883	1,47	0,4292	1,72	0,4573
0,98	0,3365	1,23	0,3907	1,48	0,4306	1,73	0,4582
0,99	0,3389	1,24	0,3925	1,49	0,4319	1,74	0,4591
1,00	0,3413	1,25	0,3944	1,50	0,4332	1,75	0,4599
1,01	0,3438	1,26	0,3962	1,51	0,4345	1,76	0,4608
1,02	0,3461	1,27	0,3980	1,52	0,4357	1,77	0,4616
1,03	0,3485	1,28	0,3997	1,53	0,4370	1,78	0,4625
1,04	0,3508	1,29	0,4015	1,54	0,4382	1,79	0,4633
1,05	0,3531	1,30	0,4032	1,55	0,4394	1,80	0,4641
1,06	0,3554	1,31	0,4049	1,56	0,4406	1,81	0,4649
1,07	0,3577	1,32	0,4066	1,57	0,4418	1,82	0,4655
1,08	0,3599	1,33	0,4082	1,58	0,4429	1,83	0,4664
1,09	0,3621	1,34	0,4099	1,59	0,4441	1,84	0,4671
1,10	0,3643	1,35	0,4115	1,60	0,4452	1,85	0,4678
1,11	0,3665	1,36	0,4131	1,61	0,4463	1,86	0,4686
1,12	0,3686	1,37	0,4147	1,62	0,4474	1,87	0,4693

Окончание табл. П2

X	$\Phi(x)$	X	$\Phi(x)$	X	$\Phi(x)$	X	$\Phi(x)$
1,88	0,4699	2,12	0,4830	2,48	0,4934	2,84	0,4977
1,89	0,4706	2,14	0,4838	2,50	0,4938	2,86	0,4979
1,90	0,4713	2,16	0,4846	2,52	0,4941	2,88	0,4980
1,91	0,4719	2,18	0,4854	2,54	0,4945	2,90	0,4981
1,92	0,4726	2,20	0,4861	2,56	0,4948	2,92	0,4982
1,93	0,4732	2,22	0,4868	2,58	0,4951	2,94	0,4984
1,94	0,4738	2,24	0,4875	2,60	0,4953	2,96	0,4985
1,95	0,4744	2,26	0,4881	2,62	0,4956	2,98	0,4986
1,96	0,4750	2,28	0,4887	2,64	0,4959	3,00	0,49865
1,97	0,4756	2,30	0,4893	2,66	0,4961	3,20	0,49931
1,98	0,4761	2,32	0,4898	2,68	0,4963	3,40	0,49966
1,99	0,4767	2,34	0,4904	2,70	0,4965	3,60	0,499841
2,00	0,4772	2,36	0,4909	2,72	0,4967	3,80	0,499928
2,02	0,4783	2,38	0,4913	2,74	0,4969	4,00	0,499968
2,04	0,4793	2,40	0,4918	2,76	0,4971	4,50	0,499997
2,06	0,4803	2,42	0,4922	2,78	0,4973	5,00	0,499997
2,08	0,4812	2,44	0,4927	2,80	0,4974		
2,10	0,4821	2,46	0,4931	2,82	0,4976		

Оглавление