течка
нных
для
истов
тиков

Э. М. Борель,
Р. Дельтейль,
Роже Юрон

ВЕРОЯТНОСТИ, ОШИБКИ

БИБЛИОТЕЧКА
ИНОСТРАННЫХ КНИГ
ДЛЯ ЭКОНОМИСТОВ
И СТАТИСТИКОВ

PROBABILITÉS, ERREURS

É. BOREL, R. DELTHEIL, ROGER HURON

9^{me} édition entièrement réfondue

Paris

ЭМ. БОРЕЛЬ, Р. ДЕЛЬТЕЙЛЬ, Р. ЮРОН

ВЕРОЯТНОСТИ, ОШИБКИ

Перевод с французского
Альб. Л. ВАЙНШТЕЙНА и Н. С. ЧЕТВЕРИКОВА

Издательство «Статистика»
Москва 1972

БИБЛИОТЕЧКА ИНОСТРАННЫХ КНИГ
ДЛЯ ЭКОНОМИСТОВ И СТАТИСТИКОВ

Издательство «Статистика» выпускает на русском языке серию книг иностранных авторов по статистике, рассчитанных на круг читателей, нуждающихся в пополнении своих математических и статистических знаний. Задача этих книг — ознакомить статистиков и экономистов на не очень сложном материале с современными методами, которые за рубежом применяются в экономическом анализе и в различных хозяйственных расчетах.

Среди намеченных к выпуску книг как книги по общим вопросам статистики, так и книги, посвященные статистическому анализу в отдельных областях экономики. Издательство старается подбирать работы, не перегруженные сложными теоретическими изысканиями, но подводящие к применению результатов таких изысканий на практике.

Уже вышли из печати книги:

1. М. Броуди. **О статистическом рассуждении.**
2. А. Бернстейн. **Справочник статистических решений.**
3. У. Дж. Рейхман. **Применение статистики.**
4. Х. Крыньский. **Математика для экономистов.**
5. С. Даймонд. **Мир вероятностей.**
6. А. Хьютсон. **Дисперсионный анализ.**
7. С. Лизер. **Эконометрические методы и задачи.**

Подготавливается к изданию книга

«Статистические методы исследования корреляций в экономике».

$\frac{1-8-3 \text{ и } 2-2-3}{61-71}$

ПРЕДИСЛОВИЕ К РУССКОМУ ИЗДАНИЮ

Книга «Вероятности, ошибки» — очередной выпуск серии «Библиотечка иностранных книг для экономистов и статистиков». Она принадлежит перу французских ученых Э. Бореля, Р. Дельтейля и Р. Юрона. Среди них наиболее известен Э. Борель[1]. Он был крупнейшим математиком-исследователем, организатором научной работы, популяризатором и вместе с тем инициатором и основателем высших центров исследовательской работы математиков и физиков своей страны; с этим совмещалась большая общественная и политическая деятельность.

В 1895 г. были опубликованы его знаменитые «Тезисы» в которых был высказан ряд новых идей, из которых для развития теории вероятностей была особенно важна идея «меры» бесконечных множеств. В 1905 г. в Бюллетене Общества математиков Франции появилась его первая работа, посвященная теории вероятностей: «Заметки по некоторым вопросам теории вероятностей».

С 1925 г. начала выходить организованная им серия работ: «Traité de Calcul des Probabilités et de ses Applications», которая была закончена в 1939 г. В 1934 г. Борель был избран Президентом Французской академии наук.

Общее число написанных Борелем работ — книг, статей, заметок — превышает 300, в том числе 35 книг.

Книга «Вероятности, ошибки» вышла в свет в 1923 г. и до 1954 г. выдержала 8 изданий. В последнем издании к двум авторам (Э. Борель и Р. Дельтейль) присоединился третий — Р. Юрон и книга была в корне переработана с учетом новейшего развития исчисления вероятностей. Однако план и направленность на прикладные цели в общем сохранились. Книга сохранила и характер «популяризации на

[1] В переводах на русский язык были изданы: «Случай» (1923), «Пространство и время» (1924) и «Вероятность и достоверность», 1969.

высоком уровне». Она предполагает у читателя знание интегрального исчисления, некоторых понятий из теории множеств и т. п.; благодаря ясному и простому изложению чтение этой книги серьезных затруднений вызвать не может.

Через всю книгу ясно проступает ее общий план, основанный на разделении случайных переменных на три категории: принимающих конечное число значений или же бесконечное, либо счетно-бесконечное, либо непрерывное множество. Особенно подробно и с большим мастерством изложена глава о геометрических вероятностях, в которой Э. Борель вводит свой метод решения задач путем построения «элементарных вероятностей» или вернее, согласно оговорке самого автора, «элементарных плотностей статочностей (шансов)». Этим путем Э. Борель побивает скептицизм Ж. Бертрана, показывающего в своем классическом трактате «Исчисление вероятностей» возможность разных решений в зависимости от «подхода» и не замечающего, что разные «подходы» соответствуют различным задачам (см. третью главу).

Э. Борель, полемизируя с Ж. Бертраном, требовал достаточно полной и точной формулировки задач, указывая, например, способ экспериментальной проверки решений, соответствующий поставленной задаче. Такое требование, характерное для стиля мышления Э. Бореля, действительно обеспечивало бы непротиворечивость и полноту условий задачи. Обобщая свой метод «элементарной вероятности», Э. Борель придает особо важное значение идее А. Пуанкаре о возможности подчинения геометрических вероятностей произвольным функциям, характеризующим распределение статочностей в поле возможных значений.

В последнем издании книги несколько смягчена направленность в сторону физики, включено много чисто теоретического материала, но обилие задач осталось прежним. Все изложение построено так, чтобы дать читателю в ясной и строгой форме то, что необходимо для пользования исчислением вероятностей в его применениях к позитивным наукам. Этим настоящая книга отличается от курсов университетских математических факультетов.

В приложении решено сохранить вывод интеграла синуса в степени m предыдущего французского издания; этот вывод предшествует выводу формулы Стирлинга и для облегчения чтения книги целесообразно избавить читателя от обращения к другим источникам, как это предлагают авторы 9-го издания.

Необходимо обратить внимание читателя на символику и терминологию французских авторов, которая иной раз расходится с принятой в советской литературе: например, был сохранен при переводе термин «типическое отклонение», которым французы называют «среднее квадратическое отклонение», но второй термин тоже встречается во французском оригинале.

Книга с пользой послужит тем, кто от общих теоретических университетских курсов захочет перейти к специальной литературе, связанной с применением исчисления вероятностей в статистике и экономике.

ИЗДАТЕЛЬСТВО

ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ

Исчисление вероятностей — одна из наиболее увлекательных и вместе с тем доходчивых отраслей математики. Только из соображений традиции, чтобы не сказать рутины, основы этой дисциплины не вошли в программу средней школы, где они с большим успехом могли бы заменить многие предметы, которые там задерживаются только потому, что никто не дает себе труда их из программы исключить.

Истоки исчисления вероятностей, как и многих других наук, были очень скромными, и основатели этой дисциплины, по-видимому, не подозревали все значение, которое новую науку ожидало в будущем.

Поводом для установления основ исчисления вероятностей послужили для Паскаля и Ферма (Pascal, Fermat) задачи, поставленные игроками в азартные игры, в частности в играх в кости; для основателей исчисления вероятностей эти понятия служили как бы забавлявшим их отдыхом от других более отвлеченных трудов. В XVIII в. дискуссии о парадоксальности Петербургской задачи, об опытах Бюффона (Buffon) с иглой все еще носят характер в такой же мере игры в смекалку, как и научных проблем, и только Лаплас (Laplace), по-видимому, был первым, кто осознал значение исчисления вероятностей для положительных наук. Это значение быстро усиливалось в течение XIX столетия по мере расширения области применения исчисления вероятностей. Ныне вероятности господствуют в современной физике и их роль возрастает; вместе с успехом атомной теории она стала значительной и в биологии, и в антропологии. Нужно ли еще вспоминать о страховании, о применении теории вероятностей в правилах стрельбы?

Мы сделали попытку собрать в небольшой книжке важнейшие основы исчисления вероятностей, придав им возможно более простую форму, но все же изложить их достаточно полно, чтобы наши читатели не испытывали затруднений при изучении областей применения теории вероятностей.

Январь 1923 г.

ПРЕДИСЛОВИЕ К ДЕВЯТОМУ ИЗДАНИЮ

Со времени первого издания этой книжки исчисление вероятностей и его применения претерпели новое и значительное развитие, а вместе с тем и к самому изложению и к символике было приложено немало усилий к их нормализации.

То, что эта книга выдержала с успехом восемь изданий с небольшими лишь исправлениями некоторых деталей, укрепило авторов и издателей в намерении полностью пересмотреть ее, не изменяя общего характера элементарного руководства, с тем чтобы модернизировать терминологию и символику. Настоящее девятое издание, проредактированное с участием нового соавтора Роже Юрона (Roger Huron) профессора Тулузского университета, привело к однообразному обозначению случайных переменных и терминов, относящихся к законам распределения и к отклонениям, а также к стандартному отклонению, к биномиальному закону вероятностей и к изучению функций этого последнего, как и нормального закона, для которого все же и старые показатели упомянуты при изложении.

Июнь 1954 г.

1. ОПРЕДЕЛЕНИЯ И ОСНОВНЫЕ ПОЛОЖЕНИЯ

I. ПОНЯТИЕ ВЕРОЯТНОСТИ. СЛОЖЕНИЕ ВЕРОЯТНОСТЕЙ. ПЕРЕМНОЖЕНИЕ ВЕРОЯТНОСТЕЙ

1. ПОНЯТИЕ СЛУЧАЙНОГО СОБЫТИЯ. ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТИ

Если мы подбросим монету и заметим, какой стороной она легла наверх, то мы произведем испытание, исход которого мы не в состоянии предвидеть. Мы считаем его результатом случая и мы говорим, что два возможных события: выпадение «орла» и выпадение «решки», представляются событиями случайными.

Теория вероятностей часто обращается к испытаниям, подобным игре в «орлянку» по случайному характеру возможных результатов: такие опыты называются испытаниями. Приведем несколько простых примеров:

бросание игральной кости,

вытаскивание карты из тщательно перетасованной колоды,

извлечение наудачу шара из урны, содержащей шары, различающиеся только по цвету, и т. п.

Повторить испытание значит произвести его снова в тех же самых условиях. Ряд повторенных испытаний называется *серией испытаний*; существенно важно, прежде чем производить какие-либо расчеты, точно определить рассматриваемую серию; этим путем мы избавимся от кажущихся парадоксов в исчислении вероятностей.

Если мы обратим внимание на возможные результаты в каждом виде опытов, указанные выше в качестве примеров, то мы не найдем никакого основания, чтобы один вид результатов осуществлялся бы преимущественно перед другим, и мы уславливаемся говорить в таком случае, что они одинаково вероятны. Так, например, если мы подбросим игральную кость, имеющую форму куба и сделанную из однородного материала, то одинаково вероятно будет для нее выпасть любой из ее сторон. Если положить в урну шары одинакового размера и сделанные из одного и того же материала, отличающиеся между собой одной лишь окраской, а затем сунуть в урну руку и извлечь шар, то представляется *одинаково возможным*, что извлеченным окажется любой из находящихся в урне шаров.

Условие, позволяющее рассматривать все случаи как равновозможные, называется *принципом симметрии*. Однако на опыте эта симметрия никогда в точности не осуществляется. Какими бы до тонкости одинаковыми ни были грани кубической кости, они должны быть распознаваемы, чтобы сделать игру возможной; это выполняется нанесенными на них очками, число которых на разных гранях, очевидно, не должно быть одинаковым.

Поэтому, рассматривая при бросании кости выпадение всех ее граней как равновозможное, мы делаем допущение, которое на деле никогда в точности не соблюдается.

То же самое происходит и в любых других проблемах теории вероятностей. Получаемые результаты исчисления могут быть применены с тем бо́льшим успехом, чем точнее будет соблюдено условие равновозможности шансов (статочностей), предполагаемое в принципе при расчетах.

Приняв раз навсегда указанные оговорки относительно принципа симметрии, мы будем рассматривать серию испытаний, исходы которых равновозможны, как образующие определенное число элементарных случаев (статочностей), и будем среди них различать те, которые удовлетворяют некоторым выставленным нами требованиям, т. е. осуществляют некоторое событие E.

По определению *вероятность события E есть отношение числа n тех случаев, в которых событие E осуществляется, и которые именуются ему благоприятными, к общему числу N всех элементарных равновозможных случаев (статочностей)*. Мы пишем

$$Pr(E) = p = \frac{n}{N}.$$

Когда же событие E не осуществляется, то говорят об осуществлении противоположного события; имеется, очевидно, $N - n$ случаев, благоприятствующих противоположному событию, и его вероятность равна

$$q = \frac{N-n}{N} = 1 - p.$$

Таким образом, вероятность какого-нибудь события и вероятность *события, ему противоположного*, вместе составляют единицу.

Вероятность представляет собой правильную дробь, числитель которой меньше или равен знаменателю. В последнем случае, событие реализуется при всех статочностях — оно *достоверно*: его *вероятность равна единице*. Когда числитель очень близок по величине к знаменателю, то вероятность *весьма близка к единице*. Часто в обыденной речи говорят, что событие чрезвычайно вероятно или что вероятность ему совершиться очень велика, но, выражаясь таким образом, слову вероятность придают несколько иной смысл, подразумевая отношение числа благоприятных шансов к числу неблагоприятных.

2. ИССЛЕДОВАНИЕ ИГРЫ В ОРЛЯНКУ

Если игра ограничивается одной партией, то в силу принципа симметрии и данного выше определения вероятность получить «орла» равна $\frac{1}{2}$.

И она всегда оказывается такой, *каковы бы ни были результаты предшествовавших партий*; и хотя бы двадцать предшествующих подбрасываний давали бы орла, вероятность получить «решку» при двадцать первом была бы равна $\frac{1}{2}$. Результаты последовательных партий совершенно не зависят один от другого, и монета, по выражению Жозефа Бертрана (Joseph Bertrand), «не имеет ни совести, ни памяти».

Если рассматривать результаты двух последовательных партий, то здесь имеются четыре возможных случая: либо первая партия дает орла, а вторая — орла или решку, либо первая партия дает решку а вторая может также дать или орла, или решку. Все четыре случая равновероятны; их можно обозначить символами:

$$OO,\ OP,\ PO,\ PP.$$

Эти обозначения можно получить, производя умножение:

$$(O+P)(O+P)=OO+OP+PO+PP.$$

Третья партия может дать орла или решку, каковы бы ни были результаты первых двух партий. Поэтому серия из трех партий может дать восемь различных результатов, одинаково вероятных. Их можно обозначить такими же символами:

$$OOO\ \ OPO\ \ POO\ \ PPO$$
$$OOP\ \ OPP\ \ POP\ \ PPP,$$

которые получаются из произведения

$$(O+P)(O+P)(O+P).$$

Те же рассуждения, повторенные столько раз, сколько требуется, показывают, что если рассматривать результаты n партий, то они представляют собой 2^n возможных равновероятных случая, представленных членами произведения

$$(O+P)(O+P)\ldots(O+P)$$

из n множителей, которые в развернутом виде дадут

$$OO\ldots OO+OO\ldots OP+\ldots$$

Если в результатах n партий интересоваться только общим числом орлов или решек, то число возможных случаев сведется к $n+0$: n орлов и 0 решек, $n-1$ орлов и 1 решка и т. д. до 0 орлов и n решек. Но вероятности этих случаев уже не будут одинаковыми.

Возьмем как пример случай трех партий. Из вышеприведенной таблицы различных возможностей видно, что событие OOO встречается среди остальных только один раз, его вероятность равна $\frac{1}{8}$. Но если интересоваться только итогами выходов O и P, то события OPO, POO, OOP уже не различаются между собой, так как в них во всех решка появлялась лишь в одной из трех партий и вероятность такого события равна $\frac{3}{8}$.

Если в общем виде рассматривать серию из n партий, то числа элементарных случаев, благоприятствующих событию, когда орел появлялся 0 раз, 1 раз, ..., k раз, ..., n раз, представлены *коэффициентами биномиального ряда*

$$(O+P)^n = O^n + C_n^1 O^{n-1} P + \ldots + C_n^k O^{n-k} P^k + \ldots + P^n.$$

Вероятность выкинуть орла k раз при n подбрасываниях равна, следовательно,

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

3. ДРУГИЕ ПРИМЕРЫ РАСЧЕТОВ

Проблемы исчисления вероятностей, поскольку уточнен основной вопрос о равновозможных статочностях, сводятся к двойному подсчету всех возможных случаев и случаев, благоприятствующих событию.

Большое число вопросов может быть разрешено этим путем без применения тех общих положений, с которыми мы ознакомимся ниже. Вот несколько весьма простых примеров.

Задача 1. *Урна содержит a белых шаров и b черных. Мы извлекаем два шара сразу. Какова вероятность того, что они окажутся различных цветов?*

Имеем число

$$C_{a+b}^2 = \frac{(a+b)(a+b-1)}{2}$$

различных парных комбинаций из общего числа $a+b$ шаров, которые образуются, если вторые шары вынимать, не вернувши первых шаров обратно в урну. Среди всех этих комбинаций имеется ab таких, которые содержат один шар белый и один черный.

Искомая вероятность равна, следовательно,

$$p = \frac{2ab}{(a+b)(a+b-1)}.$$

Задача 2. *Из колоды в 52 карты берем наудачу 13 карт. Какова вероятность того, что среди этих вынутых 13 карт окажутся все 4 туза?*

Мы рассматриваем как равновозможные все комбинации из 52 карт по 13. Общее число таких комбинаций равно:

$$C_{52}^{13}=\frac{52!}{13!\,39!}.$$

Таково число всех возможных случаев. Чтобы определить число благоприятных статочностей, заметим, что сдачу из 13 карт, содержащую 4 туза, можно получить, прибавляя к четырем тузам любую комбинацию из 9 карт, взятых из колоды в 48 карт, не содержащей тузов. При этом мы ничего не пропустим и не повторим. Число благоприятных статочностей будет, следовательно,

$$C_{48}^{9}=\frac{48!}{9!\,39!}.$$

Таким образом, для вероятности мы получим

$$p=\frac{48!\,13!}{52!\,9!}=\frac{10.11.12.13}{49.50.51.52}=\frac{11}{4\,165}$$

или, приблизительно 0,0026, что означает, говоря попросту, около 26 шансов из 10 000.

Задача 3. *Найти вероятность выкинуть 12 очков хотя бы один раз из n партий при бросании двух игральных костей.*

Бросаем две кости; при этом возможны 36 элементарных равновозможных случаев, в которых каждое число на первой кости сочетается с любым числом на второй кости; только одно сочетание из них соответствует появлению шестерки на каждой из костей, являющейся способом получения итога в 12 очков.

Если играть n партий, то надо рассмотреть все 36 возможных сочетаний.

А чтобы получить вероятность выбросить 12 очков по меньшей мере один раз, мы рассмотрим противоположное событие, т. е. n событий, когда будет каждый раз появляться иное (не 12) число очков.

Это можно осуществлять 35 способами и вероятность такого события равна, следовательно,

$$q=\left(\frac{35}{36}\right)^n,$$

а поэтому искомая вероятность составляет

$$p=1-q=1-\left(\frac{35}{36}\right)^n.$$

Можно показать, что вероятность будет больше $\frac{1}{2}$, начиная с $n=25$ партий.

События (в любом числе) называют *исключающими друг друга* или взаимно несовместимыми, если никакие из них, взятые из числа данных, не могут осуществляться совместно. Если, например, при игре в «бридж» извлекается червонная карта, то появление карты пик — явление несовместимое с первым, извлечение же червонной масти и извлечение туза не представляются несовместимыми.

Суммой некоторого множества событий $E_1, E_2, \ldots, E_k$ называется событие S, состоящее из осуществления какого-либо из событий $E_1, E_2, \ldots, E_k$; это условно записывается в форме:

$$S = E_1 + E_2 + \ldots + E_k. \tag{1}$$

Правило сложения вероятностей относится по существу к тому случаю, когда события $E_1, E_2, \ldots, E_k$ взаимно несовместимы. Его можно формулировать следующим образом.

Правило сложения вероятностей. Если событие S может осуществляться различными способами $E_1, E_2, \ldots, E_k$ взаимно несовместимыми, то вероятность такого события S равна сумме частных вероятностей, присущих различным способам осуществления события S.

Это доказывается непосредственно. Предположим, что среди всех N элементарных статочностей a_1 благоприятствует событию E_1, a_2 — событию E_2, ..., a_k — событию E_k. Мы имеем равенства

$$Pr(E_1) = \frac{a_1}{N}, \quad Pr(E_2) = \frac{a_2}{N}, \ldots, Pr(E_k) = \frac{a_k}{N}.$$

Так как $E_1, E_2, \ldots, E_k$ взаимно несовместимы, то случаи, благоприятствующие этим событиям, образуют *непересекающиеся множества* и мы имеем $a_1 + a_2 + \ldots + a_k$ случаев, благоприятствующих событию S, вероятность которого имеет, следовательно, величину

$$Pr(S) = \frac{a_1 + a_2 + \ldots + a_k}{N} = Pr(E_1) + Pr(E_2) + \ldots + Pr(E_k). \tag{2}$$

Пример. Игра в «passe—sept» («больше семи очков»). Рассмотрим ряд испытаний, состоящий из подбрасывания двух игральных костей с шестью гранями каждая и найдем вероятность события S, когда число очков, выпавших на обеих костях, превышает 7.

Приводимая здесь таблица, составленная по образцу таблицы Пифагора, показывает итоги очков, выпавших во всех 36 равновозможных случаях. Кости обозначены буквами A и B. Среди 36 случаев 5 благоприятствуют появлению 8 очков; 4, 3, 2 и 1 появлению соответственно 9, 10, 11 и 12 очков. События, здесь указанные, очевидно несовместимы одно с другим; вероятность получить сумму очков, превышающую 7, равна

$$\frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{15}{36}, \text{ или } \frac{5}{12}.$$

Число очков на кости A

Число очков на кости B	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Замечание. Условие взаимного исключения различных способов осуществления события S является, очевидно, основоположным. Если оно не соблюдено, то $a_1 + a_2 + \ldots + a_k$ случаев, благоприятствующих S, перестают быть «непересекающимися»; раздельных благоприятствующих случаев оказывается меньше, чем $a_1 + a_2 + \ldots + a_k$, и мы имеем неравенство:

$$Pr(S) < Pr(E_1) + Pr(E_2) + \ldots + Pr(E_k). \tag{3}$$

Если возможности $E_1, E_2, \ldots, E_k$ взаимно несовместимы и полностью охватывают все элементарные статочности, относящиеся к ряду рассматриваемых испытаний, то совокупность называют системой составляющих (systéme de constituants). Сумма S таких возможностей представляет собой явление достоверное и мы имеем соотношения:

$$a_1 + a_2 + \ldots + a_k = N$$

$$Pr(S) = Pr(E_1) + Pr(E_2) + \ldots + Pr(E_k) = 1.$$

Какая-либо система событий, взаимно несовместимых, сумма которых существует с достоверностью, называется также *исчерпывающей системой события.*

5. ПРАВИЛО ПЕРЕМНОЖЕНИЯ ВЕРОЯТНОСТЕЙ

Произведением двух случайных событий A и B, обозначаемым символом AB (или $A \cdot B$, или $A \times B$), называется событие, состоящее в том, что *осуществляются совместно и* A *и* B.

Из этого определения, а также из определения суммы двух событий получаются соотношения

$$(4) \qquad \begin{aligned} AB &= BA, \\ A(B+C) &= AB + AC. \end{aligned}$$

Чтобы удостовериться в их справедливости достаточно заменить знак сложения союзом «или», а знак умножения — союзом «и».

Рассмотрим событие AB. В отношении к A и B может иметь место

1) что эти два события осуществляются как то, так и другое
2) что осуществляется A, но не осуществляется B,
3) что осуществляется B, но не осуществляется A,
4) наконец, что не осуществляется ни A, ни B.

Пусть n_1, n_2, n_3, n_4 — элементарные статочности, соответствующие этим различным возможностям, тогда мы имеем

$$n_1+n_2+n_3+n_4=N,$$

где N означает общее число всех случаев (статочностей).

Вероятность сложного события AB равна

$$p=\frac{n_1}{n_1+n_2+n_3+n_4},$$

а вероятность простого события A равна

$$\omega=\frac{n_1+n_2}{n_1+n_2+n_3+n_4}.$$

Если мы знаем, что событие A произошло, то вероятность осуществления B равна

$$\omega'=\frac{n_1}{n_1+n_2}.$$

Мы видим, что

$$p=\omega\cdot\omega'. \tag{5}$$

В этом заключается правило перемножения вероятностей.

Правило перемножения вероятностей. Если событие предполагает осуществление двух событий A и B, то его вероятность равна произведению вероятности A на вероятность B, когда известно, что событие A уже произошло.

Соотношение (5) можно записать также и в таком виде:

$$Pr(AB)=Pr(A)\cdot Pr(B/A), \tag{6}$$

где символ $Pr\ (B/A)$ означает вероятность B, когда A уже осуществилось; такую вероятность B называют *условной вероятностью B по отношению к A*. Вероятность же $Pr\ (B)$, т. е. вероятность B, независимо от того, произошло ли событие A или нет, называется *вероятностью à priori B*.

Предыдущие рассуждения, если в них A и B поменять местами, показывают, что

$$Pr(AB)=Pr(BA)=Pr(B)\cdot Pr(A/B). \tag{7}$$

Правило перемножения вероятностей, естественно, обобщается на случай многих событий, входящих в состав сложного события. Следует заметить, что в такой обобщенной формулировке каждая вероятность, входящая в произведение вероятностей, предполагается относящейся к соответственному событию в *предположении, что все ранее введенные в рассмотрение события уже осуществились.*

Так, например, мы имеем такое соотношение:

$$Pr(ABC) = Pr(A) \cdot Pr(B/A) \cdot Pr(C/AB). \tag{8}$$

6. СОБЫТИЯ НЕЗАВИСИМЫЕ

Интуитивно мы утверждаем, что B независимо от A, если осуществление A не влияет на вероятность B, т. е. если имеет место

$$Pr(B/A) = Pr(B).$$

Согласно с принятой выше символикой это соотношение можно заменить такой формой

$$\frac{n_1}{n_1+n_2} = \frac{n_1+n_3}{n_1+n_2+n_3+n_4},$$

которая приводит к виду:

$$\frac{n_1}{n_2} = \frac{n_3}{n_4};$$

и отсюда непосредственно вытекает, что независимость между A и B — *взаимна.*

В случае двух независимых событий правило перемножения вероятностей переходит в соотношение:

$$Pr(AB) = Pr(A) \cdot Pr(B). \tag{9}$$

Замечание. Пользуясь обозначениями § 5, мы можем написать:

$$Pr(A+B) = \frac{n_1+n_2+n_3}{n_1+n_2+n_3+n_4},$$

$$Pr(A) = \frac{n_1+n_2}{n_1+n_2+n_3+n_4},$$

$$Pr(B) = \frac{n_1+n_3}{n_1+n_2+n_3+n_4}.$$

Отсюда вытекает, что

$$Pr(A+B) \leqslant Pr(A) + Pr(B) - Pr(AB). \tag{10}$$

Если A и B взаимно исключают одно другое, то $n_1 = 0$, а поэтому и $Pr(AB) = 0$ и мы возвращаемся к правилу сложения вероятностей.

7. ПРИМЕНЕНИЕ К ПОСЛЕДОВАТЕЛЬНОМУ ИЗВЛЕЧЕНИЮ ШАРОВ ИЗ ОДНОЙ И ТОЙ ЖЕ УРНЫ

Задача 4. *Урна содержит N шаров, в том числе a белых. Последовательно извлекают из урны три шара. Какова вероятность, что все они окажутся белыми?*

Применим правило сложения вероятностей. Вероятность того, что первый шар окажется белым, равна $\frac{a}{N}$; после того, как это первое событие произошло, в урне осталось $N-1$ шар, из которых белых осталось $a-1$; вероятность вынуть белый шар при втором извлечении равна $\frac{a-1}{N-1}$; после того, как это предполагаемое извлечение совершилось, урна стала содержать $N-2$ шара, в том числе белых $a-2$. Вероятность извлечения белого шара в третий раз равна, следовательно, $\frac{a-2}{N-2}$; наконец сложная вероятность равна

$$\frac{a}{N}\cdot\frac{a-1}{N-1}\cdot\frac{a-2}{N-2}=\frac{a(a-1)(a-2)}{N(N-1)(N-2)}.$$

Например, колода из 32 карт содержит 4 короля. Для вероятности того, что, вынимая из колоды наудачу три карты, мы получим трех королей, мы найдем, приняв в нашей формуле $N = 32$ и $a = 3$,

$$\frac{4.3.2}{30.31.32}=\frac{1}{10.31.4}=\frac{1}{1240},$$

т. е. один шанс из 1240.

Равным образом колода из 52 карт содержит 13 треф; вероятность получить при сдаче три трефы равна

$$\frac{13.12.11}{52.51.50}=\frac{11}{850}=0{,}0129,$$

мы имеем приближенно тринадцать шансов из тысячи.

С колодой в 32 карты аналогичная вероятность окажется равной

$$\frac{8.7.6}{32.31.30}=0{,}0112,$$

т. е. несколько меньше предыдущей.

Задача 1. (см. стр. 12). Пусть $a+b=N$ показывает число шаров в урне. Требуется вычислить вероятность того, что при двух извлечениях шары окажутся разного цвета. Непосредственный подсчет дал нам

$$P=\frac{2ab}{(a+b)(a+b-1)}=\frac{2a(N-a)}{N(N-1)}.$$

Решение этой задачи можно рассматривать как комбинацию правил сложения и перемножения вероятностей: ожидаемое событие может наступить или путем извлечения белого шара при первом испытании и затем черного — при втором, или же при первом извлечении — черного шара и при втором — белого. Следовательно, мы можем применить правило сложения и сумма вероятностей окажется равной

$$\frac{a}{N}\ \frac{N-a}{N-1}+\frac{N-a}{N}\ \ \frac{a}{N-1}=\frac{2a(N-a)}{N(N-1)}.$$

II. ЗАКОН ВЕРОЯТНОСТЕЙ ДЛЯ ОДНОГО ИЛИ ДЛЯ НЕСКОЛЬКИХ СЛУЧАЙНЫХ ПЕРЕМЕННЫХ С КОНЕЧНЫМ ЧИСЛОМ ВОЗМОЖНЫХ ЗНАЧЕНИЙ

8. СЛУЧАЙНАЯ ПЕРЕМЕННАЯ ДИСКРЕТНАЯ, КОНЕЧНОГО ПОРЯДКА

Всякая переменная, значение которой определяется в результате испытания, принадлежащего к некоторой серии, называется *случайной переменной.*

Для обозначения случайной переменной пользуются, как правило, заглавными буквами, тогда как принимаемые ими значения указываются соответственными строчными буквами, снабженными, если нужно, индексами.

Примеры. 1. X представляет собой сумму очков, полученных при бросании двух игральных костей с шестью гранями каждая. Эта случайная переменная может принимать значения 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 с соответстующими вероятностями (согласно таблице в § 4)

$$\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36}.$$

2. X обозначает случайное событие E с вероятностью p и становится равным 1, когда E осуществляется, в противном же случае приравнивается 0; X называется *указателем* события E.

Закон вероятностей. Пусть нам дана в общем виде случайная переменная, способная принимать значения $x_1, x_2, \ldots, x_k$ в зависимости от осуществления взаимно независимых событий $E_1, E_2, \ldots, E_k$, в конечном числе образующих исчерпывающую систему. X называется случайной переменной порядка k. Вероятности $p_1, p_2, \ldots, p_k$ событий $E_1, E_2, \ldots, E_k$ удовлетворяют равенству

$$p_1+p_2+\ldots+p_k=1.$$

Система связанных одна с другой величин

$$(x_1, p_1), (x_2, p_2), \ldots, (x_k, p_k)$$

определяет закон вероятности случайной переменной X.

В первом приведенном выше примере X является *случайной переменной порядка 11*, закон вероятности которой дан следующим рядом:

Значения x_i:	2	3	4	5	6	7	8	9	10	11	12
вероятности p_i:	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$.

9. ФОРМЫ ПРЕДСТАВЛЕНИЯ ЗАКОНА ВЕРОЯТНОСТИ СЛУЧАЙНОЙ ПЕРЕМЕННОЙ

Пусть X случайная переменная; вероятность того, что X примет значение меньшее, чем x, является функцией от x, которая называется функцией распределения X, а также *функцией накопленных вероятностей*, присущих X.

Обозначив эту функцию через $F(x)$, мы можем написать

$$Pr(X < x) = F(x).$$

Обращаясь к нашему первому примеру и применяя к нему правило накопления вероятностей, мы увидим, что $F(x)$ равна 0 для $x < 2$; $\frac{1}{36}$ для $2 \leqslant x < 3$; $\frac{1}{36} + \frac{2}{36}$ для $3 \leqslant x < 4$ ... вплоть до $\frac{36}{36} = 1$ для $x \geqslant 12$.

График функции распределения случайной переменной дает наглядное представление о законе вероятностей этой переменной; если X порядка k и если возможные его значения даны в возрастающем порядке

$$-\infty < x_1 < x_2 < \ldots < x_{k-1} < x_k < +\infty,$$

то график состоит из k ступеней (рис. 1), $F(x)$, принимающих значение 0 для $x < x_1$; p_1 для $x_1 \leqslant x < x_2$; $p_1 + p_2$ для $x_2 \leqslant x < x_3$ ... до 1 для $x \geqslant x_k$. Функция иногда называется *ступенчатой*.

Закон вероятностей можно представить еще и в иной форме, сочетая с каждой точкой абсциссы $x_1, x_2, \ldots, x_k$ некоторые массы, равные соответственным вероятностям $p_1, p_2, \ldots, p_k$. Эту форму мы будем называть *механическим представлением* закона вероятностей и в дальнейшем мы увидим присущие ей выгоды. В этом случае измеряют сумму масс, абсциссы которых меньше x.

10. МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ

Рассмотрим случайную переменную X порядка k, которая может принимать значения $x_1, x_2, \ldots, x_k$ с соответствующими вероятностями $p_1, p_2, \ldots, p_k$ в сумме равными единице. По определению, *математическим ожиданием X называется величина*

$$p_1 x_1 + p_2 x_2 + \ldots + p_k x_k, \tag{11}$$

обозначаемая символом $E(X)$, который читается «E от X заглавного». Эту величину называют также и *вероятным и средним*[1] *значением* X, употребляя также символ $\overline{X}$.

Такое название — математическое ожидание — очевидно подсказано изучением азартных игр, послужившим началом теории вероятностей. Переменную X можно всегда рассматривать как выигрыш игрока, а возможные отрицательные значения — как его проигрыш; в таком случае математическое ожидание игрока представляет собой алгебраическую сумму произведений возможных выигрышей на вероятности осуществления.

[1] Хотя по своей форме математическое ожидание представляет собой среднюю (теоретическую ибо опирается на вероятности), но ее следует всегда отличать от собственно «средней» (опирающейся на частости) вследствие принципиального различия в их логическом содержании. — *Прим. перев.*

Например, Петр получает 100 франков, когда при игре в орлянку выпадает орел; вероятность этого равна 0,5; его математическое ожидание составит, следовательно, 50 франков. Павел же получает 1000 франков, когда в двух последовательных партиях в обеих выпадут орлы, поэтому его математическое ожидание равно

$$1\,000 \times 0{,}25 = 250 \text{ франков.}$$

Мы видим, что выражение *математическое ожидание* должно рассматриваться как единый нераздельный термин с точно определенным смыслом, которому не следует пытаться придавать обычный смысл двух слов: *ожидание* и *математический*. В обычном смысле слова Павел может надеяться заполучить 1000 франков, если же эта надежда не оправдается, то он ничего не получит, ни в каком случае он не получит 250 франков в силу своего *математического ожидания*.

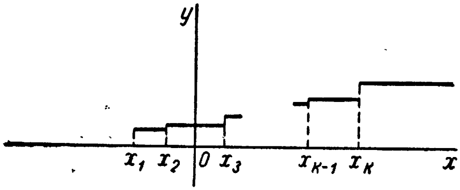

Рис. 1

Игру называют справедливой, если математическое ожидание игрока равно его ставке; поэтому математическое ожидание, равное некоторой определенной сумме, может быть обменено на эту сумму в *том случае, если находится игрок, согласный принять на себя справедливую игру*.

11. ПРИМЕНЕНИЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ

Пользование математическим ожиданием может в известных случаях значительно упростить расчеты по сравнению с вычислением, непосредственно опирающимся на вероятности. Вот пример тому.

Задача 5. *Урна содержит n шаров, перенумерованных числами от 1 до n; их вынимают в последовательном порядке. Какова вероятность того, что по крайней мере в одном случае номер на шаре совпадет с порядковым номером извлечения?*

Эта задача относится к группе игр «на встречу»; мы ограничимся исчислением вероятности того, что произойдет по меньшей мере одна встреча.

Пусть p_i будет вероятностью встречи при i-м испытании. Существует всего n возможных порядков выхода для n шаров, и если отложить в сторону шар с номером i, идущим на i-м извлечении в каждом благоприятном случае, то остальные $n-1$ шаров могут образовать $(n-1)!$ возможных перестановок, следовательно,

$$p_i = \frac{1}{n},$$

каково бы ни было i.

Аналогичные рассуждения показывают, что вероятность p_{ij} дл[я] двух встреч с порядковыми номерами i и j равна

$$p_{ij}=\frac{1}{n(n-1)},$$

каковы бы ни были i и j. А вероятность встреч в заданном количеств[е] будет

$$p_{ij\ldots l}=\frac{1}{n(n-1)(n-2)\ldots(n-h+1)},$$

каковы бы ни были порядковые номера. В частности, вероятность n встреч равна $\frac{1}{n!}$ — результат очевидный непосредственно.

Имеем

$$\Sigma p_i=np_i=1,$$

$$\Sigma p_{ij}=\frac{n(n-1)}{1\cdot 2}p_{ij}=\frac{1}{1\cdot 2},$$

$$\cdots\cdots\cdots\cdots$$

$$\Sigma p_{ij\ldots l}=\frac{1}{h!}.$$

Установив это, представим себе игрока, который выигрывает франк при каждой встрече и вносит франк при каждой группе из двух встреч; выигрывает франк при трех встречах и т. д. Его математическое ожидание было бы

$$E=1-\frac{1}{1.2}+\frac{1}{1.2.3}-\ldots\pm\frac{1}{n!}.$$

Я утверждаю, что это и есть вероятность того, чтобы была по меньшей мере одна встреча. В самом деле, эта вероятность и есть как раз математическое ожидание игрока, который не получает ничего, если не произойдет ни одной встречи и получает франк, если случится хотя бы одна встреча. Легко усмотреть, что эти условия совпадают с условиями задачи. На самом деле, если не произойдет ни одной встречи, то игрок не получает ничего, если же встреч будет h, то эти h образуют группы по две встречи числом $\frac{h(h-1)}{1.2}$, или по три встречи числом $\frac{h(h-1)(h-2)}{1.2.3}$ и т. д.; так что наш игрок, согласно условиям задачи, получит количество франков, равное

$$h-\frac{h(h-1)}{1\cdot 2}+\frac{h(h-1)(h-2)}{1\cdot 2\cdot 3}-\ldots\pm 1,$$

или $1-(1-1)^h=1$, каково бы ни было h.

Предложение доказано. Интересно отметить, что E образовано членами знакопеременного ряда, сумма которого равна $1-\frac{1}{e}$; принимая за сумму это ее приближенное значение $1-\frac{1}{e}$, мы совершаем ошибку, меньшую, чем $\frac{1}{(n+1)!}$.

12. МОМЕНТЫ СЛУЧАЙНОЙ ПЕРЕМЕННОЙ. КОЛЕБЛЕМОСТЬ. ТИПИЧЕСКОЕ ОТКЛОНЕНИЕ

Пусть X случайная переменная с дискретными значениями конечного порядка. Для большего удобства мы будем в дальнейшем обозначать этот порядок буквой n вместо k. x_i — одно из возможных значений $x_1, x_2, \ldots, x_n$ переменной X и p_i — соответственная вероятность этого значения.

Если каждому значению X сопоставить некоторое значение, указываемое данной функцией $f(X)$, то $f(X)$ будет некоторой новой случайной переменной, для которой при $i = 1, 2, \ldots, n$ будут справедливы соотношения:

$$Pr\,[f(X) = f(x_i)] = Pr\,(X = x_i);$$

поэтому и математическое ожидание $f(X)$ будет

$$E\,[f(X)] = \sum_{i=1}^{i=n} p_i f(x_i). \tag{12}$$

Моментом k-го порядка случайной переменной X называют величину $E(X^k)$, где k — целое положительное число или равно 0. Для упрощения пишут:

$$E\,(X^k) = m_k. \tag{13}$$

Момент m_0 равен по своей величине единице; момент m_1 есть *математическое ожидание* $E(X)$.

Отклонением называют новую случайную переменную Z

$$Z = X - m_1, \tag{14}$$

она обладает тем свойством, что $E(Z) = E(X) - m_1 = 0$.

Среднее абсолютное отклонение, по определению, есть величина

$$E\,(|Z|); \tag{15}$$

центрированный момент второго порядка переменной X есть математическое ожидание значений Z^2, т. е. величина

$$\mu_2 = E\,(Z^2) = E\,[(X - m_1)^2], \tag{16}$$

которая иначе называется колеблемостью (рассеянием) или *дисперсией* X.

Квадратный корень из дисперсии

$$\sigma = \sqrt{\mu_2} \tag{17}$$

называется *средним квадратическим отклонением*, а в настоящее время чаще именуется *типическим отклонением* (écart-type). Эта величина является наиболее важной среди «типических величин», характеризующих закон вероятностей; она имеет основополагающее значение в современном развитии теории вероятностей и ее применений.

Дисперсия может быть очень просто выражена через моменты первых двух порядков. На самом деле, согласно (16) мы имеем

$$\mu_2 = \Sigma p_i (m_1 - x_i)^2 = m_1^2 \Sigma p_i - 2m_1 \Sigma p_i x_i + \Sigma p_i x_i^2,$$

откуда в силу того, что $\Sigma p_i = 1$, $\Sigma p_i x_i = m_1$, получаем

$$\mu_2 = m_2 - m_1^2, \tag{18}$$

которое можно написать еще в таком виде:

$$E(X^2) = [E(X)]^2 + \mu_2.$$

Заметим, что дисперсия μ_2, будучи суммой квадратов

$$\Sigma p_i (m_i - x_i)^2,$$

с положительными коэффициентами не может равняться нулю за исключением случая, когда все x_i равны m_1 и, следовательно, равны между собой. Таким образом, мы получаем условие:

$$E(X^2) \geqslant [E(X)]^2,$$

в котором равенство возможно лишь в случае $x_1 = x_2 = \ldots = x_n$.

13. НЕРАВЕНСТВО БЬЭНЭМЕ (BIENAYME) — ЧЕБЫШЕВА

Пусть X случайная переменная, математическое ожидание которой равно $m_1 = m$, а типическое отклонение равно σ. Мы будем исследовать вероятность

$$Pr[|X - m| > t\sigma],$$

где t означает произвольное положительное число и где неравенство выражает требование, чтобы отклонение $X - m$ превышало или равнялось абсолютной величине $t\sigma$.

Нанесем на ось $0X$ точку ω с абсциссой m и точки H' и H с абсциссами $m - t\sigma$ и $m + t\sigma$ (рис. 2).

Если M соответствует абсциссе X, то задача приводится к отысканию вероятности того, что точка M *лежит вне сегмента* $H'H$.

По определению, мы имеем

$$\sigma^2 = \sum_{i=1}^{i=n} p_i (x_i - m)^2. \tag{19}$$

В этой сумме мы всегда вправе предполагать, что x_i расположены в порядке их возрастания и писать $\sigma^2 = A + B + C$.

A представляет совокупность членов, зависящих от M_i, расположенных влево от H' (включая и точку H'). B — совокупность членов, зависящих от M_i, расположенных вправо от H (включая и точку H).

Так как все члены суммы (19) либо положительны, либо равны нулю, то мы можем написать неравенство

$$\sigma^2 \geqslant A + C. \tag{20}$$

Но для всех точек M_i, расположенных вне сегмента $H'H$, квадраты $(x_i - m)^2$ превосходят $t^2\sigma^2$, поэтому неравенство (20) á fortiori приводит к неравенству

$$\sigma^2 \geqslant \sigma^2 t^2 [\Sigma' p_i],$$

в котором суммирование $\sum'$ распространяется лишь на M_i, лежащие вне сегмента $H'H$. Сумма $\Sigma' p_i$ выражает как раз ту вероятность, кото-

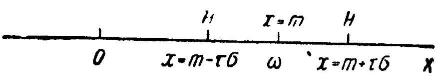

Рис. 2

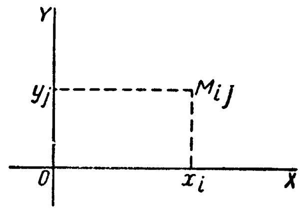

Рис. 3

рую мы ищем, и она оказывается меньшей или равной $\frac{1}{t^2}$, что мы и запишем в виде

$$Pr(|X - m| \geqslant t\sigma] \leqslant \frac{1}{t^2}. \tag{21}$$

В этом и состоит основное неравенство Бьэнэме—Чебышева, которое показывает, что с увеличением t вероятность отклонения, большего или равного $t\sigma$, убывает по меньшей мере как $\frac{1}{t^2}$; для $t = 10$ эта вероятность меньше или равна $\frac{1}{100}$.

14. СИСТЕМА, СОСТОЯЩАЯ ИЗ ДВУХ СЛУЧАЙНЫХ ПЕРЕМЕННЫХ КОНЕЧНОГО ПОРЯДКА

Рассмотрим две случайные переменные X и Y, для которых ряды возможных значений определяются общим для них рядом испытаний и которые мы на этом основании будем называть совместными случайными переменными.

Каждой паре возможных значений x_i, y_j этих переменных мы можем отнести точку M_{ij} на плоскости с осями прямоугольных координат $0X$ и $0Y$ и сообщить этой точке массу p_{ij}, являющуюся мерой вероятности события $(X = x_i,\ Y = y_j)$ (рис. 3).

Таким образом, мы обобщим механическое представление, определенное нами в § 9 для закона вероятностей одной случайной переменной.

При этом индекс i изменяется от 1 до m, а индекс j — от 1 до n; пре полагается, что p_{ij} подчинено условию

$$(22) \qquad \sum_{i=1}^{i=m} \sum_{j=1}^{i=n} p_{ij} = 1,$$

т. е. общая сумма всех масс в механическом представлении равна ед нице.

Зададимся функцией $f(X, Y)$ и каждой паре x_i, y_j возможных зн чений X и Y соотнесем число $f(x_i, y_j)$. Математическое ожидание сл чайной переменной, которое мы таким образом определим, буд

$$(23) \qquad E[f(X, Y)] = \sum_{i}^{m} \sum_{j}^{n} p_{ij} f(x_i, y_j).$$

Один частный случай имеет первостепенную важность, когд $f(X, Y)$ является *суммой* $X + Y$; для него имеем соотношения

$$(24) \qquad E(X+Y) = \sum_i \sum_j p_{ij}(x_i + y_j) = \sum_i (x_i \sum p_{ij}) + \sum_j \Big(y_j \sum_i p_{ij}\Big).$$

Обозначим через p_{x_i} сумму $\sum_j p_{ij}$ тех масс, которые расположен в точке M_{ij}, находящейся на прямой $X = x_i$; эту массу можно пред ставить себе сосредоточенной на оси $0X$ в точке с абсциссой x_i. Сово купность этих масс p_{x_i} определяет для X, рассматриваемого отдель но, распределение, которое вполне устанавливает закон вероятностей так как удовлетворяет условию

$$\sum_i p_{x_i} = \sum_i \Big(\sum_j p_{ij}\Big) = 1;$$

Такой закон вероятностей называется *законом краевых вероятносте* переменной X. Таким же образом, пользуясь $p_{y_j} = \sum_i p_{ij}$, можн определить и закон краевых вероятностей для Y.

Этим законам краевых вероятностей соответствуют математически ожидания:

$$E(X) = \sum_i x_i p_{x_i} = \sum_i \Big(x_i \sum_j p_{ij}\Big),$$

$$E(Y) = \sum_j y_j p_{y_j} = \sum_j \Big(y_j \sum_i p_{ij}\Big),$$

которые и доказывают, согласно (24), основные соотношения

$$(25) \qquad E(X+Y) = E(X) + E(Y),$$

выражающие собой то, что *математическое ожидание суммы двух слу чайных переменных равно сумме математических ожиданий каждо из них, взятой в отдельности.*

Если мы зафиксируем X, приняв его равным x_i, то соответствующими точками M_{ij} будут те, которые расположены на прямой $x = x_i$. Распределение масс p_{ij}, присущих этим точкам, не может служить для определения закона вероятностей соответственных значений Y, ибо сумма этих масс $p_{x_i} = \sum_j p_{ij}$, вообще говоря, меньше единицы.

Однако согласно правилу перемножения вероятностей вероятность p_{ij} равна произведению вероятности p_{x_i} на вероятность для $Y = y_j$, при условии, что X принял значение x_i; эта вероятность, следовательно, напишется так:

$$Pr\,[Y = y_j/X = x_i] = \frac{p_{ij}}{p_{x_i}}. \tag{26}$$

Для каждого значения i мы, таким образом, получим закон вероятностей для Y при условии, что $X = x_i$; этот закон называется *законом условной вероятности* Y.

Если X и Y представляют собой две *независимые случайные переменные*, то вероятность для $Y = y_j$, когда $X = x_i$, равна попросту вероятности $Y = y_j$; и это условие должно быть соблюдено для всех пар возможных значений i и j, откуда вытекает соотношение

$$p_{ij} = p_{x_i} p_{y_j}, \tag{27}$$

которое означает для всех пар возможных значений i и j, равенство вероятностей

$$Pr\,[X = x_i/Y = y_j] = p_{y_j}$$

и, следовательно, образует необходимую и достаточную систему условий независимости[1].

Исследуем, предполагая эти условия осуществленными, математическое ожидание произведения XY; по определению мы имеем

$$E\,(XY) = \sum_i \sum_j p_{ij} x_i y_j = \sum_i x_i \left(\sum_j p_{ij} y_j\right);$$

[1] Авторы недостаточно отчетливо разделяют два понятия: независимость и некоррелированность.

Независимость имеет место, если условный закон вероятностей одной переменной остается неизменным, когда другая переменная пробегает все возможные для нее значения.

Некоррелированность имеет место, если математическое ожидание одной случайной переменной не изменяется при изменениях другой переменной. При этом, однако, другие параметры закона вероятностей первой случайной переменной (например, ее типическое отклонение и т. п.) могут реагировать на изменения значений другой переменной.

Теорема о перемножении математических ожиданий справедлива при втором, более слабом, условии, а именно при некоррелированности. — *Прим. перев.*

последняя сумма, стоящая в скобках, приводится, согласно (27), к

$$\sum_j p_{ij} y_j = \sum_j p_{x_i} p_{y_j} y_j;$$

отсюда

(28) $$E(XY) = \left(\sum_i x_i p_{x_i}\right)\left(\sum_j y_j p_{y_j}\right) = E(X) \cdot E(Y).$$

Таким образом, *математическое ожидание произведения двух независимых случайных переменных равно произведению математических ожиданий каждого из этих переменных, взятого отдельно.*

16. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ

Пусть X и Y две совместно осуществляющиеся случайные переменные, математические ожидания которых $\sum\limits_i x_i p_{x_i}$ и $\sum\limits_j y_j p_{y_j}$ мы обозначим через $\overline{X}$ и $\overline{Y}$. Мы знаем, что величины

$$E[(X-\overline{X})^2 = \sigma_x^2, \quad E[(Y-\overline{Y})^2] = \sigma_y^2$$

представляют собой дисперсии X и Y.

Математическому ожиданию

(29) $$E[(X-\overline{X})(Y-\overline{Y})] = \sum_i \sum_j p_{ij}(x_i - \overline{X})(y_j - \overline{Y})$$

присвоено наименование ковариации случайных переменных X и Y; отношение же

(30) $$r = \frac{E[(X-\overline{X})(Y-\overline{Y})]}{\sigma_X \sigma_Y}$$

является их коэффициентом корреляции.

Если случайные переменные независимы, то же самое имеет место и для $X - \overline{X}$ и $Y - \overline{Y}$; в этом случае математическое ожидание их произведения, согласно (28), равно произведению $E(X-\overline{X})E(Y-\overline{Y})$. Оно, следовательно, равно нулю, так как оба множителя в этом произведении равны нулю по самому определению $\overline{X}$ и $\overline{Y}$.

Таким образом, *если две случайные переменные независимы, то их коэффициенты корреляции равны нулю.* Однако обратное положение не справедливо.

17. ТЕОРЕМА О СЛОЖЕНИИ ДИСПЕРСИЙ

Понятие совместных случайных переменных, определенное в § 14, без затруднений может быть обобщено на случай любого числа случайных переменных.

Рассмотрим n случайных переменных совместимых или несовместимых, предполагая их *независимыми*. Обозначим их через X_1, X_2,

..., X_n и пусть $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$ будут их дисперсии. Обратимся к рассмотрению новой случайной переменной

(31) $$Z = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n,$$

являющейся линейным сочетанием переменных $X_1, X_2, \ldots, X_n$ с постоянными коэффициентами; эти коэффициенты не являются «случайными величинами».

Математическое ожидание Z выражается, даже в случае отсутствия независимости между случайными переменными, через

$$\overline{Z} = a_1 \overline{X}_1 + a_2 \overline{X}_2 + \ldots + a_n \overline{X}_n.$$

Что же касается дисперсии Z, то она дается выражением

$$\sigma_Z^2 = E[(Z-\overline{Z})^2] = E\{[a_1(X_1-\overline{X}_1) + a_2(X_2-\overline{X}_2) + \\ + \ldots + a_n(X_n-\overline{X}_n)]^2\},$$

которое развертывается в ряд, состоящий из n квадратов и $\frac{n(n-1)}{2}$ произведений попарно всех переменных. Так как попарные произведения в этом разложении в ряд *все равны нулю, поскольку* X_1, X_2, ..., X_n *между собой независимы*, а члены-квадраты представляют собой произведения $a_1^2, a_2^2, \ldots, a_n^2$ на дисперсии $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, то в результате разложения получаем более краткую форму:

(32) $$\sigma_Z^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \ldots + a_n^2\sigma_n^2,$$

которая и выражает собой *теорему о сложении дисперсий* для линейного сочетания независимых случайных переменных с участием постоянных коэффициентов.

III. ВЕРОЯТНОСТИ СЧЕТНО-БЕСКОНЕЧНЫХ МНОЖЕСТВ. ВЕРОЯТНОСТИ НЕПРЕРЫВНЫХ МНОЖЕСТВ (КОНТИНУУМОВ)

18. СЛУЧАЙНЫЕ ПЕРЕМЕННЫЕ, МНОЖЕСТВО ЗНАЧЕНИЙ КОТОРЫХ СЧЕТНО-БЕСКОНЕЧНО

Первые две части настоящей главы опираются на обычное определение вероятности, допускающее существование равновозможных статочностей. При более строгой теоретической точке зрения это определение следует заменить другим, согласно которому каждой случайной переменной приписывается некоторый коэффициент p, заключенный в пределах $0 \leqslant p \leqslant 1$, и принять в качестве двух основных положений (аксиом) правило сложения вероятностей и правило их перемножения.

Становясь на эту новую точку зрения, можно понятие случайной переменной конечного порядка обобщить, принимая существование случайной переменной X, которая может принимать счетно-бесконечное число значений

$$x_0, x_1, x_2, \ldots, x_n, \ldots$$

с соответственными вероятностями

$$p_0, p_1, p_2, \ldots, p_n, \ldots$$

конечно, при соблюдении условия, что ряд

(33) $$p_0+p_1+p_2+\ldots+p_n+\ldots$$

сходится и в сумме дает единицу.

Пример. Рассмотрим закон, называемый *законом Пуассона* (Poisson), в котором

(34) $$x_n=n, \quad p_n=\frac{\lambda^n}{n!}e^{-\lambda},$$

где λ — данное положительное число (мы принимаем условно, что $0!=1$); соответствующий ряд дает, конечно, в сумме единицу.

Математическое ожидание $E(X)$ и в этом обобщенном случае представляется суммой ряда, общий член которого имеет форму $p_n x_n$. В примере (34) он имеет вид:

$$m_1=e^{-\lambda}\left[0+\lambda+\frac{\lambda^2}{1!}+\frac{\lambda^3}{2!}+\ldots+\frac{\lambda^n}{(n-1)!}+\ldots\right],$$

т. е. $m_1=\lambda$, так как сумма, заключенная в скобки, есть не что иное, как развернутое в ряд произведение λe^{λ}.

Математическое ожидание какой-либо данной функции $f(X)$ определяется как сумма ряда, общий член которого имеет вид $p_n f(x_n)$. Вычислим, например, момент второго порядка для случая (34); он выражается как

$$m_2=e^{-\lambda}\left[0+1^2\frac{\lambda}{1!}+2^2\frac{\lambda^2}{2!}+\ldots+n^2\frac{\lambda^n}{n!}+\ldots\right],$$

и нетрудно усмотреть, что каждый коэффициент ряда, заключенного в прямые скобки, можно представить в виде

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

и поэтому

(35) $$m_2=\lambda+\lambda^2.$$

Отсюда следует, что в случае (34) дисперсия получает значение

(36) $$\mu_2=m_2-m_1^2=\lambda.$$

19. СЛУЧАЙНАЯ ПЕРЕМЕННАЯ, ПРИНИМАЮЩАЯ НЕПРЕРЫВНЫЙ РЯД ЗНАЧЕНИЙ

Обобщение понятия случайной переменной можно продолжить и рассматривать X как случайную переменную, которая может принимать *все значения в некотором интервале* (a, b).

Закон вероятностей для X будет определен, если мы знаем вероятность $F(x)$ осуществления неравенства

$$a \leqslant X < x,$$

и для этой функции мы сохраним прежнее наименование *функции распределения* случайной переменной X. Механическое представление этого закона вероятностей получается в виде некоторого тела, *масса которого равна единице* и распределена на сегменте AB, соединяющем точки a и b на оси $0X$, причем вся масса, расположенная влево от точки M с абсциссой x, равна по величине $F(x)$.

Мы ограничимся случаем, в котором $F(x)$ *имеет на сегменте* (a, b) *непрерывную производную*; если обозначить в этом случае производную через $\rho(x)$, то мы получим соотношение

$$Pr\,(X < x) = F(x) = \int_a^x \rho(x)\,dx \tag{37}$$

и

$$Pr\,[x_1 < X < x_2] = F(x_2) - F(x_1) = \int_{x_2}^{x_1} \rho(x)\,dx. \tag{38}$$

Функция $\rho(x)$ играет здесь роль такую же, как плотность в механике; соответственный дифференциал

$$\rho(x)\,dx \tag{39}$$

следует, естественно, назвать *плотностью вероятности*, являющейся основной частью соответствующего бесконечно малого dx, т. е. вероятности того, что X примет значения, лежащие между x и $x + dx$. Этой вероятности дано наименование *элементарной вероятности*.

Мы видим, что если существует непрерывная плотность вероятности, то событие $x = \alpha$, где α означает любую величину на сегменте (a, b), *имеет вероятность, равную нулю*. Такое событие представляется невозможным в случае вероятностей переменной с непрерывными значениями, а также переменной со значениями в счетно-бесконечном числе. Однако в случае (39) такое строгое суждение не имеет места, поскольку мы предположили, что X, по определению, может принимать любые значения в сегменте (a, b), в таком случае говорят о событии *почти невозможном*.

Замечание. Можно перейти к еще большему обобщению и рассматривать случайную переменную, значения которой распределены на сегменте с такой плотностью, которая в некоторых определенных точках принимает бесконечно большие значения. Но мы не станем исследовать эту категорию переменных в этом элементарном руководстве и ограничимся случаем, в котором закон вероятности определен элементарной вероятностью $\rho(x)dx$. Такой закон вероятностей называется абсолютно непрерывным.

20. МОМЕНТЫ СЛУЧАЙНОЙ ПЕРЕМЕННОЙ АБСОЛЮТНО НЕПРЕРЫВНЫЙ ЗАКОН ВЕРОЯТНОСТЕЙ

Распространяя на рассматриваемый случай те определения, которые были даны в § 10 и 12, мы назовем *математическим ожиданием* $f(x)$ величину

$$E[f(X)] = \int_a^b f(x)\rho(x)\,dx. \tag{40}$$

Моментом порядка k от X будет, в частности для всякого целого положительного числа k или для нулевого значения, служить интеграл

$$m_k = E(X^k) = \int_a^b (f(x))^k \rho(x)\,dx; \tag{41}$$

момент же нулевого порядка, очевидно, будет удовлетворять условию:

$$m_0 = \int_a^b \rho(x)\,dx = 1.$$

Важное замечание. Все, что будет сказано в дальнейшем распространяется и на тот случай, когда интервал (a, b) заменяется интервалом $(-\infty, +\infty)$ при непременном соблюдении условия, что интеграл

$$\int_{-\infty}^{+\infty} \rho(x)\,dx$$

имеет смысл и равен единице. Можно даже *считать случай интервала* $(-\infty, +\infty)$ *за общий случай*, если положить, что X может принимать только значения, расположенные внутри сегмента (a, b), а что $\rho(x)$ для значений $x < a$ и $x > b$ предполагается равным нулю.

Формулы (40) и (41) в том случае, который сейчас рассматривается, принимают вид:

$$E[f(X)] = \int_{-\infty}^{+\infty} f(x)\rho(x)\,dx,$$

$$m_k = E[X^k] = \int_{-\infty}^{+\infty} x^k \rho(x)\,dx,$$

если предполагать, как это и естественно, что интегралы в правых частях равенств *имеют смысл* (*т. е. существуют*).

Если мы затем положим

$$m_1 = E(X) = \int_{-\infty}^{+\infty} x\rho(x)\,dx,$$

то *дисперсия* или центральный момент второго порядка будет равен:

$$\mu_2 = \int_{-\infty}^{+\infty} (x - m_1)^2 \rho(x)\,dx,$$

а *типическое отклонение* равно:

$$\sigma = \sqrt{\mu_2}.$$

21. ПРИМЕРЫ
ЗАКОН ГАУССА

Закон Гаусса (Fr. Gauss). Наиболее часто встречающийся в практике закон вероятностей абсолютно непрерывный и определенный для значений x, лежащих между $-\infty$ и $+\infty$, для элементарной вероятности

$$\rho(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx, \tag{42}$$

это *закон Гаусса*, имеющий основополагающее значение во всей теории вероятностей. При помощи функции от $\alpha : \lambda = \frac{x-m}{\sigma}$, которая называется приведенной (нормированной) получается *закон Гаусса в приведенном виде*, в основании которого лежит элементарная вероятность

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{\lambda^2}{2}} d\lambda.$$

Закон Гаусса, конечно, удовлетворяет условию $m_0 = 1$, ибо, пользуясь *приведенной формой* с помощью общеизвестного расчета[1], мы находим:

$$\int_{-\infty}^{+\infty} e^{-\frac{\lambda^2}{2}} d\lambda = \sqrt{2\pi}.$$

Случайная переменная x имеет *среднюю* величину m и *типическое отклонение* σ. Средняя для λ равна нулю, так как в *приведенном законе* плотность вероятности представляет собой четную функцию от λ. Что же касается момента $E(x^2)$, то для его вычисления мы имеем

$$m_2 = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx,$$

[1] См. Leconte et Deltheil, Éléments de calcul différentiel et intégral (Collection Armand Colin № 72—73, t. II, p. 51).

откуда, переходя на функцию от λ и учитывая предыдущее, получае

$$m_2 - m^2 = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \lambda^2 e^{-\frac{\lambda^2}{2}} d\lambda. \tag{43}$$

Этот интеграл легко вычисляется интегрированием по частям мы без труда найдем результат:

$$m_2 - m^2 = \sigma^2.$$

22. ЗАКОН ВЕРОЯТНОСТЕЙ, АБСОЛЮТНО НЕПРЕРЫВНЫХ ДЛЯ ДВУХ СЛУЧАЙНЫХ ПЕРЕМЕННЫХ

Пусть X и Y две «совмещенные» случайные переменные, которы мы будем рассматривать как координаты *случайно расположенной точки* на плоскости $X0Y$. Предположим, что закон вероятностей системы (X, Y) представлен механически пр помощи массы в сумме, равной единице, размещенной непрерывно на некоторой области (D) на плоскости (рис. 4). При этом мы ограничимся случаем, когда (Δ) представляет собой некоторый участок, расположенный внутри (D); вероятность того, что точка (X, Y) попадет в (Δ), измеряется двойным интегралом *плотности непрерывной вероятности:*

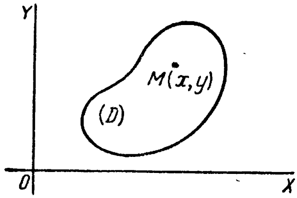

Рис. 4

$$\iint_{\Delta} \rho(x, y)\, dx\, dy.$$

Выражение $\rho(x, y)dxdy$ измеряет в бесконечно малых второго порядка вероятность совместного осуществления неравенств:

$$x < X < x + dx \quad \text{и} \quad y < Y < y + dy,$$

являясь, таким образом, *элементарной вероятностью.*

Можно всегда исходить из предположения, что областью (D) служит вся плоскость $X0Y$; в таком случае достаточно допустить, что $\rho(x, y) = 0$ всюду вне области (D).

Краевые плотности. Если (D), как мы только что допустили, занимает всю плоскость $X0Y$, то вероятность для X того, что эта переменная имеет значение, меньшее, чем x, выражается интегралами:

$$Pr[X < x] = \int_{-\infty}^{x} dx \int_{-\infty}^{+\infty} \rho(x, y)\, dy. \tag{44}$$

Интеграл по y представляет собой непрерывную функцию от x и равен

$$(45) \qquad f(x) = \int_{-\infty}^{+\infty} \rho(x, y)\, dy.$$

Этот интеграл называется краевой плотностью переменной X ввиду того, что простая элементарная вероятность для X и равна как раз $f(x)dx$. Таким же путем определится и *краевая плотность* для Y, а именно:

$$(46) \qquad g(y) = \int_{-\infty}^{+\infty} \rho(x, y)\, dx.$$

23. УСЛОВНЫЕ ВЕРОЯТНОСТИ В СЛУЧАЕ ЗАКОНА АБСОЛЮТНОЙ НЕПРЕРЫВНОСТИ ДВУХ СЛУЧАЙНЫХ ПЕРЕМЕННЫХ

Обозначим через A событие $X = x$ и через B — событие $Y < y$. Событие B/A состоит в том, что $Y < y$, когда известно, что $X = x$; *поставим себе целью вычислить вероятность события* B/A.

Согласно правилу перемножения вероятностей мы имеем:

$$(47) \qquad Pr(B/A) = \frac{Pr(AB)}{Pr(A)},$$

здесь правая часть имеет форму неопределенности $\frac{0}{0}$.

Чтобы раскрыть эту неопределенность мы рассмотрим вместо A событие $x \leqslant X < x + \Delta x$, которое мы назовем A'. Вероятность произведения $A'B$ имеет величину:

$$Pr(A'B) = \int_{x}^{x+\Delta x} dx \int_{-\infty}^{y} \rho(x, y)\, dy;$$

и мы положим

$$\int_{-\infty}^{y} \rho(x, y)\, dy = \Phi(x, y).$$

Теорема о среднем значении интеграла позволяет написать

$$Pr(A'B) = \Delta x\, \Phi(x + \theta\Delta x, y),$$

где θ означает число, заключенное между 0 и 1. С другой стороны, вероятность A' имеет значение

$$\int_{x}^{x+dx} f(x)\, dx = \Delta x f(x + \omega\Delta x),$$

2*

где $f(x)$ означает краевую плотность (45), а ω — еще одно число, заключенное между 0 и 1. Отсюда мы выводим соотношение:

$$Pr(B/A') = \frac{Pr(A'B)}{Pr(A')} = \frac{\Phi(x+\theta\Delta x, y)}{f(x+\omega\Delta x)}.$$

А теперь мы будем считать само собой разумеющимся тот факт, что вероятность $Pr(B/A)$ служит пределом для $Pr(B/A')$, когда dx стремится к нулю; тогда мы, наконец, получим

$$Pr(B/A) = \frac{1}{f(x)} \int_{-\infty}^{y} \rho(x, y)\, dy, \tag{48}$$

и мы видим, что плотность условной вероятности Y в зависимости от значений X равна:

$$\delta(Y/X) = \frac{\rho(x, y)}{f(x)}. \tag{49}$$

Тем же путем определяется и плотность условной вероятности X в зависимости от значений Y:

$$\delta(X/Y) = \frac{\rho(x, y)}{g(y)}. \tag{50}$$

Необходимое и достаточное условие взаимной независимости X и Y.

Мы говорим, что Y *не зависит от* X, если при любых значениях y вероятность $Pr(B/A)$ не зависит от значений x, т. е. при любых значениях x оказывается одинаковой с $Pr(B/A)$. Для этого необходимо и достаточно, чтобы второй член в равенстве (48) был тождественно одинаков с $Pr(B)$; это можно выразить таким образом:

$$\int_{-\infty}^{y} dy \int_{-\infty}^{+\infty} \rho(x, y)\, dx = \int_{-\infty}^{y} g(y)\, dy,$$

откуда и вытекает необходимое и достаточное условие независимости

$$\rho(x, y) = f(x)\, g(y), \tag{51}$$

влекущее за собой в силу симметрии независимость X от Y.

Проще говоря, *необходимо и достаточно, чтобы плотность* $\rho(x, y)$ *была бы произведением функции от* x, *взятого самого по себе, на функцию от* y, *взятого также самого по себе.* Двойной интеграл этого произведения, взятый по всей плоскости, величина которого составляет единицу, равен в этом случае произведению интегралов от $f(x)$ и от $g(y)$, взятых от $-\infty$ до $+\infty$. Мы не умалим общности нашего утверждения, считая, что каждый из этих двух интегралов равен единице, что равносильно утверждению, что $f(x)$ и $g(y)$ — краевые плотности x и y соответственно.

Полагая, что область (D) всегда может быть заменена всей плоскостью $X0Y$, мы имеем

$$\overline{X} = E(X) = \int_{-\infty}^{+\infty} xf(x)\,dx = \iint_D x\rho(x, y)\,dxdy,$$

$$\overline{Y} = E(Y) = \int_{-\infty}^{+\infty} yg(y)\,dy = \iint_D y\rho(x, y)\,dx\,dy,$$

так что точка g с координатами $x = \overline{X}$; $y = \overline{Y}$ является не чем иным, как центром тяжести той массы, которая представляет распределение переменных в области (D); эту точку часто называют *средней точкой.*

Кроме того, мы уже имеем установленные определения в развернутом виде

$$\sigma_x^2 = E[(X-\overline{X})^2] = \iint_D (x-\overline{X})^2 \rho(x, y)\,dx\,dy$$

и

$$\sigma_y^2 = E[(Y-\overline{Y})^2] = \iint_D (y-\overline{Y})^2 \rho(x, y)\,dx\,dy.$$

Ковариация X и Y также выражается через

$$E[(X-\overline{X})(Y-\overline{Y})] = \iint_{(D)} (x-\overline{X})(y-\overline{Y})\,\rho(x, y)\,dx\,dy \tag{52}$$

и из этих выражений можно составить *коэффициент корреляции,* пользуясь формулой (30).

Если случайные переменные X и Y *независимы,* то ковариация согласно (51) представится нам произведением

$$\int_{-\infty}^{+\infty} (x-\overline{X})\,f(x)\,dx \cdot \int_{-\infty}^{+\infty} (y-\overline{Y})\,g(y)\,dy$$

математических ожиданий $(x - \overline{X})$ на $(y - \overline{Y})$, которые оба равны нулю; а следовательно, и коэффициент корреляции также должен равняться нулю.

В рассматриваемом общем случае коэффициент корреляции удовлетворяет двойному неравенству

$$-1 \leqslant r \leqslant +1. \tag{53}$$

В самом деле, положим $X_1 = X - \overline{X}$ и $Y_1 = Y - \overline{Y}$ и рассмотрим случайную переменную $Z = Y_1 - kX_1$, где k имеет произвольное действительное значение. Математическое ожидание $E(Z^2)$ положительно или равно нулю, поскольку оно выражается интегралом, в котором все значения подынтегрального выражения положительны или равны нулю. Развертывая квадрат, взятый в скобки, получи неравенства

$$E(Y_1^2)-2kE(X_1Y_1)+k^2E(X_1^2)\geqslant 0,$$

которые можно написать в таком виде:

$$k^2\sigma_x^2-2rk\sigma_x\sigma_y+\sigma_y^2\geqslant 0,$$

где r означает коэффициент корреляции $\frac{E(X_1Y_1)}{\sigma_x\sigma_y}$. Левая часть н равенства представляет собой трехчлен второго порядка относител но k и не может сделаться отрицательным, а из этого сразу же вы кает условие: $r^2\leqslant 1$.

25. ПОЛНОЕ (ПОДРОБНОЕ) ИССЛЕДОВАНИЕ ОДНОЙ ПЕРЕМЕННОЙ

Рассмотрим закон вероятностей, определенный на всей плоскос XOY выражением:

$$\rho(x,y)=H\cdot e^{-\left(\frac{x^2}{a^2}-2\frac{xy}{k}+\frac{y^2}{b^2}\right)}, \tag{54}$$

где $a>0$, $b>0$, $|k|>ab$, это последнее условие заключает в се то, что кривая

$$\Omega=\frac{x^2}{a^2}-2\frac{xy}{k}+\frac{y^2}{b^2}=1$$

является эллипсом; H — постоянная, выбранная так, чтобы двойно интеграл J функции $\rho(x,y)$, взятый по всей плоскости, равнялся бы

Определение H. Ω можно представить в форме

$$\left(\frac{1}{a^2}-\frac{b^2}{k^2}\right)x^2+\left(\frac{y}{b}-\frac{bx}{k}\right)^2,$$

в таком случае двойной интеграл напишется в виде:

$$J=H\int_{-\infty}^{+\infty}e^{-\left(\frac{1}{b^2}-\frac{b^2}{k^2}\right)x^2}dx\int_{-\infty}^{+\infty}e^{-\left(\frac{y}{b}-\frac{bx}{k}\right)^2}dy.$$

В интеграле по y произведем замену переменной $\frac{y}{b}-\frac{bx}{k}=\frac{u}{\sqrt{2}}$ тогда он примет вид

$$\frac{b}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-\frac{u^2}{2}}du=b\sqrt{\pi}\ (\text{ср. § 21}),$$

откуда следует:

$$J=Hb\sqrt{\pi}\int_{-\infty}^{+\infty}e^{-\left(\frac{1}{a^2}-\frac{b^2}{k^2}\right)x^2}dx,$$

полагая

$$x\sqrt{\frac{1}{a^2}-\frac{b^2}{k^2}}=\frac{v}{\sqrt{2}},$$

мы получим значение J:

$$J=\frac{\pi Hb}{\sqrt{\frac{1}{a^2}-\frac{b^2}{k^2}}}.$$

Условие $J=1$ нам дает, следовательно (при $k>0$):

(55) $$H=\frac{1}{\pi abk}\sqrt{k^2-a^2b^2}.$$

Коэффициент корреляции. Читателю предоставляется в качестве упражнения доказать с помощью аналогичных выкладок, что

(56) $$\sigma_X^2=\frac{1}{2}\frac{a^2k^2}{k^2-a^2b^2};$$

$$\sigma_Y^2=\frac{1}{2}\frac{b^2k^2}{k^2-a^2b^2},$$

$$r\sigma_X\sigma_Y=\frac{1}{2}\frac{a^2b^2k}{k^2-a^2b^2},$$

откуда

$$r=\frac{ab}{k}.$$

Краевые вероятности. Для краевых плотностей вероятностей получаются выражения:

(57) $$f(x)=\frac{1}{\sigma_X\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma_X^2}};$$

$$g(y)=\frac{1}{\sigma_Y\sqrt{2\pi}}e^{-\frac{y^2}{2\sigma_Y^2}},$$

из которых непосредственно следует, что $\overline{X}=0$ и $\overline{Y}=0$, а средняя точка g помещается в начале координат.

Условные вероятности. Согласно формуле (49) плотность условной вероятности Y в зависимости от X является функцией одних лишь величин σ_x, σ_y и r:

(58) $$\delta(Y/X)=\frac{1}{\sigma_Y\sqrt{2\pi(1-r^2)}}e^{-\frac{1}{2(1-r^2)}\left(\frac{y}{\sigma_Y}-r\frac{x}{\sigma_X}\right)^2};$$

меняя x и y местами, мы можем вывести плотность условной вероят ности $\delta(X/Y)$.

Мы видим, что краевые законы вероятностей, так же как и закон условных вероятностей, являются *законами Гаусса* (ср. § 21), которы легко написать в приведенной форме, произведя соответствующу замену переменных.

УПРАЖНЕНИЯ К ГЛАВЕ ПЕРВОЙ

1. Играя в орлянку, один игрок подбрасывает три монеты, а дру гой две. Выигрывает тот, у кого выпадет большее число орлов. Како ва вероятность выигрыша у каждого из игроков, если они ограничи вают игру одной партией?

Тот же вопрос ставится при условии, что в случае одинаково числа орлов игра повторяется.

Ответ на второй вопрос: $\frac{8}{11}$ и $\frac{3}{11}$.

2. В лотерее имеются 100 билетов, в том числе 3 выигрышных Петр взял 40 билетов. Какова вероятность, что ему достанется вы грыш и притом только один?

Ответ: $\frac{236}{539}$.

3. *Задача о пульке.* Три игрока — A, B и C, играют в орлянку н следующих условиях: A и B играют первую партию, C заменяет пр игравшего и игра продолжается; игрок, исключенный из какой-либ партии, сражается против выигравшего эту партию и это продолжает ся до тех пор, пока один из игроков не выиграет две партии подряд.

Какова вероятность выигрыша для каждого из трех игроков пере тем, как они приступили к игре?

Ответ: $\frac{5}{14}$, $\frac{5}{14}$ и $\frac{2}{7}$.

4. Тот же вопрос, поставленный после второй партии в предполо жении, что в первой партии выиграл игрок B.

Ответ: $\frac{1}{7}$, $\frac{4}{7}$ и $\frac{2}{7}$.

5. Рассмотрим неограниченное число урн U_1, U_2, ..., таких, чт урна U_n содержит a_n белых и b_n черных шаров. Извлекают один ша из урны U_1, если он окажется черным, то производят второе извлече ние из урны U_2, и так продолжают дальше до тех пор, пока не появит ся белый шар. Какова вероятность извлечь белый шар, если обще число извлечений не должно превышать k?

Ответ:

$$P_k = \frac{a_1}{a_1+b_1} + \frac{b_1}{a_1+b_1}\cdot\frac{a_2}{a_2+b_2} + \frac{b_1}{a_1+b_1}\cdot\frac{b_2}{a_2+b_2}\cdot\frac{a_3}{a_3+b_3} + \ldots + \frac{b_1 b_2 \ldots a_k}{(a_1+b_1)\ldots(a_k+b_k)},$$

или же

$$P_k = 1 - \frac{b_1 b_2 \ldots b_k}{(a_1+b_1)\ldots(a_k+b_k)},$$

как это сразу получится, если рассмотреть событие противоположное.

6. Пусть X случайная переменная, следующая закону Пуассона (§ 18) и имеющая математическое ожидание, равное λ.

1) Положим для r, равного 1, 2, ..., что

$$Pr(X < r) = F(r)$$

Показать, что

$$F(r) = \frac{1}{(r-1)!} \int_{\lambda}^{+\infty} e^{-x} x^{r-1} dx.$$

2) Показать, обозначив через r какой-либо параметр, что математическое ожидание $f(x)e^x$ равно

$$\varphi(t) = E(e^{tx}) = e^{\lambda}(e^t - 1).$$

3) Показать, что для всех целых положительных значений для функции $\varphi(t)$ и ее производных справедливы соотношения:

$$\varphi^{n+1}(t) = \lambda e^t [\varphi + C_n^1 \varphi' + C_n^2 \varphi'' + \ldots + C_n^n \varphi^{(n)}].$$

7. Случайная непрерывная переменная может принимать любое положительное значение, причем закон вероятностей определяется элементарной вероятностью

$$\frac{\alpha p}{\Gamma(p)} e^{-\alpha x} x^{p-1} dx,$$

где α означает какое-либо данное положительное число, p — число равное и большее 1 и $\Gamma(p)$ — функция гамма[1].

$$\Gamma(p) = \int_0^{+\infty} t^{p-1} e^{-t} dt;$$

так определенное распределение называется *распределением гамма.*

Показать, что для $k = 0, 1, 2, \ldots$ момент k-го порядка дается выражением:

$$m_k = \frac{1}{\alpha^k} \frac{\Gamma(p+k)}{\Gamma(p)},$$

и что центральный момент второго порядка равен $\mu_2 = \frac{p}{\alpha^2}$.

[1] См. Leconte et Deltheil, o. c. t 11, стр. 22.

8. Рассмотреть случайную переменную X, данную элементарной вероятностью

$$\frac{1}{\pi} \frac{dx}{1+(x-\mu)^2},$$

для значений $-\infty < x < +\infty$ (распределение Коши).

Показать, что при таком законе распределения моменты m_1 и m_2 не существуют и что

$$\lim_{l \to \infty} \int_{-l}^{+l} \frac{x\,dx}{1+(x-\mu)^2} = \mu.$$

9. Рассмотрим закон вероятностей, определенный на всей плоскости XOY согласно элементарной вероятности

$$\frac{\sqrt{k^2-a^2 b^2}}{\pi abk} e^{-\left(\frac{x^2}{a^2} - 2\frac{xy}{k} + \frac{y^2}{b^2}\right)} dx\,dy,$$

полагая $k > \sqrt{ab}$ (ср. § 25).

1) Вычислить в функциях от дисперсий σ_x^2, σ_y^2 и коэффициента корреляции r типичное отклонение законов условных вероятностей для X в зависимости от Y и для Y в зависимости от X.

Ответ:

$$\sqrt{1-r^2}\,\sigma_X \text{ и } \sqrt{1-r^2}\,\sigma_Y.$$

2) На каждой прямой $X = x$ точка на ординате, соответствующая $E[Y/(X = x)]$, есть средняя точка распределения соответственных значений Y. Показать, что при изменении x геометрическое место этих точек есть прямая, проходящая через начало координат; эта прямая называется прямой регрессии Y по X.

2. ТЕОРИЯ ПОВТОРНЫХ ИСПЫТАНИЙ НОРМАЛЬНЫЙ ЗАКОН ОТКЛОНЕНИЙ

1. ПОВТОРНЫЕ ИСПЫТАНИЯ С УРНОЙ, СОДЕРЖАЩЕЙ ШАРЫ ДВУХ ЦВЕТОВ

26. БИНОМИАЛЬНЫЙ ЗАКОН ВЕРОЯТНОСТЕЙ

Представим себе, что производится серия испытаний, причем осуществляется либо та, либо другая альтернатива, вероятности которых остаются постоянными.

Мы можем, например, конкретизировать такую серию испытаний в виде извлечений шаров из урны, содержащей белые и черные шары в количествах, пропорциональных p и q. После каждого извлечения, записав цвет вынутого шара, мы его возвращаем обратно в урну для того, чтобы ее состав оставался неизменным. О таких испытаниях говорят, что *производится испытание по схеме возвращенного шара из урны, содержащей элементы двух сортов.*

Если считать благоприятствующей альтернативу, когда появляется белый шар, то ее вероятность равна p. Если произвести n последовательных извлечений, то вероятность того, что между всеми n испытаниями k испытаний будут иметь благоприятный исход, а $n-k$ — неблагоприятный, равна $p^k q^{n-k}$, если мы при этом будем обращать внимание на порядок чередования тех и других результатов. Если же на этот порядок внимания не обращать, то указанное произведение надо умножить на C_n^k — число возможных сочетаний из n по k. Таким образом, мы приходим к величине

$$P_k = \frac{n!}{k!\,(n-k)!}\, p^k q^{n-k}, \tag{59}$$

которая представляет собой член порядка k в ряде разложения би-
нома

$$(p+q)^n = q^n + nq^{n-1}p + \ldots + C_n^k q^{n-k} p^k + \ldots + p^n.$$

Пусть X случайная переменная, показывающая число благоприят-
ных исходов в серии из n испытаний; X может принимать значения 0,
1, 2, ..., k, ..., n с вероятностями P_1, P_2, ..., P_k, ..., P_n, указанными
формулой (59).

Закон вероятностей, получаемых таким путем, называется *бино-
миальным законом*; он имеет весьма важное значение.

Моменты. Вычисление моментов биномиального закона не
представляет трудности, если рассматривать функцию

$$\varphi(t) = (q+pt)^n = \sum_{k=0}^{n} P_k t^k, \tag{60}$$

которая называется *производящей функцией моментов.*

Момент порядка 0 равен $\varphi(1) = \Sigma P_k$ и его величина равна $(p +$
$+ q)^n = 1$, как это и должно быть. Если мы возьмем производную от
тождества (60) по параметру t, то получим

$$\varphi'(t) = np(q+pt)^{n-1} = \Sigma kP_k t^{k-1},$$

и отсюда, если приравнять t единице, найдем

$$np = \Sigma kP_k = E(X) = m_1.$$

Возьмем производную вторично и снова положим $t = 1$, тогда
получим

$$n(n-1)p^2 = \Sigma k(k-1)P_k = m_2 - m_1.$$

Следовательно, момент второго порядка дается выражением

$$m_2 = n(n-1)p^2 + np = npq + n^2p^2,$$

а для типического отклонения получается значение

$$\mu_2 = m_2 - m_1^2 = npq, \quad \sigma = \sqrt{npq}. \tag{61}$$

27. ЗАКОН ЧАСТОСТЕЙ. НАИБОЛЕЕ ВЕРОЯТНАЯ ЧАСТОСТЬ

Если благоприятная альтернатива в серии из n испытаний осущест-
вилась k раз, то ее *частость* равна

$$f = \frac{k}{n}.$$

Случайную переменную, представляющую различные значения час-
тости f, мы обозначим через $Y = \frac{X}{n}$; ее математическое ожидание равно

$$\overline{Y} = E(Y) = \frac{1}{n} E(X) = p,$$

а дисперсия равна

$$E[(Y-p)^2]=\frac{1}{n^2}E[X-np)^2]=\frac{pq}{n};$$

таким образом, частость подчиняется закону вероятностей с математическим ожиданием p и с типическим отклонением $\frac{pq}{n}$.

Поставим себе теперь задачей найти, какая частость является наиболее вероятной? Очевидно, это будет та частость, которой соответствует наибольшее значение P_k.

Рассмотрим отношение

$$\mu=\frac{P_k}{P_{k-1}}=\frac{n-k+1}{k}\cdot\frac{p}{q}.$$

Пока это отношение превышает единицу, члены ряда по своей величине возрастают; когда же отношение становится меньше единицы, они начинают убывать. Наибольший член будет иметь порядковый номер h такой, при котором

$$P_{h-1}<P_h>P_{h+1},$$

т. е.

$$(n-h+1)p>hq,$$

$$(n-h)p<(h+1)q,$$

или, пользуясь равенством $p+q=1$, напишем

$$np-q<h<np+p,$$

так что h представляет собой единственное целое число, заключенное между величинами $np-q$ и $np+p$, которые разнятся между собой на единицу. В одном лишь случае получается неопределенность, когда оба числа $np-q$ и $np+p$ оказываются целыми числами. Соответственные значения

$$h_1=np-q,$$

$$h_2=np+p=h_1+1$$

относятся к двум равновеликим членам в ряде разложения бинома $(p+q)^n$.

В окончательном виде это можно выразить так: когда P_k принимает наибольшее значение, то k и $n-k$ оказываются целыми числами, наиболее близкими к np и nq. Этот результат можно формулировать очень просто в таких словах: *при данном n наиболее вероятная частость является одной из тех, которые заключают в себе вероятность*; если, в частности, np целое число, то наиболее вероятная частость совпадает с вероятностью. Мы в дальнейшем вернемся к этому результату огромной важности.

28. ТЕОРЕМА ЯКОВА БЕРНУЛЛИ (JACQUES BERNOULLI), ИЛИ ЗАКОН БОЛЬШИХ ЧИСЕЛ

Мы показали выше, что случайная переменная Y, представляющая собой частость f, следует закону с математическим ожиданием, равным p, и с типическим отклонением, равным $\sqrt{\frac{pq}{n}}$.

На основании равенства Бьенэмэ—Чебышева (§ 13) мы питаем уверенность в том, что вероятность для $(Y-p)$ превзойти величину $t\sqrt{\frac{pq}{n}}$ меньше, чем $\frac{1}{t^2}$, и это можно написать в такой форме:

$$Pr\left[|Y-p|>t\sqrt{\frac{pq}{n}}\right]\leqslant\frac{1}{t^2}, \tag{62}$$

где t — произвольное положительное число.

Пусть ε будет положительным и сколь угодно малым числом. Если мы положим

$$t\sqrt{\frac{pq}{n}}=\varepsilon,$$

то неравенство (62) дает нам верхнюю границу вероятности того, что отклонение $(Y-p)$ по своей абсолютной величине превзойдет ε; на самом деле

$$Pr\left[|Y-p|>\varepsilon\right]\leqslant\frac{pq}{n\varepsilon^2}. \tag{63}$$

Отсюда следует, что для любого заданного значения ε можно взять n настолько большим, чтобы указанная выше вероятность была меньше данной величины α, которая также может быть взята сколь угодно малой; действительно, достаточно, чтобы n превышало величину $\frac{pq}{\alpha\varepsilon^2}$.

Отсюда вытекает следующая теорема Якова Бернулли (1654—1705), опубликованная в 1713 г. в его посмертно изданном сочинении Ars conjectandi.

Если дано сколь угодно малое число ε, то вероятность для отклонения частости f благоприятного события в некоторой серии испытаний от вероятности p этого события получит абсолютную величину, превышающую ε, стремится к нулю, когда число испытаний неограниченно возрастает.

Введение в эту формулировку числа испытаний, достаточно большого, объясняет, почему теорема получила наименование *закона больших чисел*, данного теореме Якова Бернулли[1].

Представленное выше доказательство ее опирается на неравенство Бьенэмэ—Чебышева и основано на следующих допущениях:

[1] Название «закона больших чисел» Пуассон присвоил своему выводу теоремы, аналогичной теореме Якова Бернулли. Существует обоснованное мнение (Бьенэмэ), что теорема Пуассона по математическому существу своему не внесла ничего нового в теорему Якова Бернулли. — *Прим. перев.*

1) средняя величина существует,

2) типическое отклонение при данной численности испытаний имеет определенное значение,

3) это значение стремится к нулю одновременно с $\frac{1}{n}$.

Эта теорема справедлива, следовательно, не только для биномиального распределения, но также и для всякого закона вероятностей, при котором соблюдаются вышеуказанные условия.

II. СЕРИИ, СОДЕРЖАЩИЕ БОЛЬШОЕ ЧИСЛО ИСПЫТАНИЙ

29. ФОРМУЛА СТИРЛИНГА (STIRLINQ)

Чтобы получать численные значения различных вероятностей, в особенности такие, какие до сих пор рассматривались в этой главе, нужно обращаться к *факториалам* вида

$$n! = 1 \cdot 2 \cdot 3 \ldots n,$$

которые быстро возрастают, когда n выходит за пределы немногих единиц.

Отсюда понятен интерес к быстрому вычислению $n!$, даже если такое вычисление будет лишь приближенным. Необходимо уточнить, что мы понимаем под словами «приближенное вычисление». Приближенное значение, которое мы укажем для факториала $n!$ в действительности намного отличается от его точной величины — в том смысле, что разность между приближенной и точной величиной весьма велика и будет тем большей, чем больше число n. Но *отношение* приближенной величины к точной будет очень мало отличаться от единицы и будет отличаться тем меньше, чем больше будет число n. Другими словами, если *абсолютная ошибка очень велика, то относительная ошибка весьма мала*. Но важна именно относительная ошибка, так как вероятность получается делением факториалов.

Формула, которой на практике пользуются для приближенного вычисления $n!$, принадлежит Стирлингу и имеет следующий вид:

$$n! = n^n e^{-n} \sqrt{2\pi n}\,(1 + \varepsilon_n). \tag{64}$$

В этой формуле n означает какое-либо целое число, e — основание натуральных логарифмов ($e = 2{,}71828\ldots$), π — отношение длины окружности к диаметру ($\pi = 3{,}14159\ldots$) и ε_n — переменная величина, зависящая от n, которая *стремится к нулю, когда n увеличивается беспредельно*. При некоторых применениях может оказаться полезным знать, что $n\varepsilon_n$ стремится к $\frac{1}{12}$, когда n беспредельно возрастает. Полагают $12n\varepsilon_n = 1 + \theta_n$.

Что касается доказательства формулы Стирлинга, то его читатель найдет в приложении в конце книги.

Легко себе представить в случае числового примера относительную ошибку, получающуюся в результате замены $n!$ на

$$\varphi(n) = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}.$$

Возьмем $n = 20$. Для него найдем
$20! = 2432902008176640000$ (19 цифр),
$\varphi(20) = 2422786385510400000$ (19 цифр).

Относительное расхождение между ними будет заключаться между 0,004 и 0,005. Само собой разумеется, что значение $\varphi(20)$ приближенное, а нули, которыми оно заканчивается, должны быть заменены значащими цифрами и за ними должен был бы идти бесконечный ряд десятичных знаков.

30. ПРИМЕНЕНИЕ К УРАВНИВАНИЮ ПОЛОЖЕНИЯ ИГРОКОВ ПРИ ИГРЕ В ОРЛЯНКУ

Чтобы дать сейчас же пример применения формулы Стирлинга, вычислим вероятность выкинуть орла n раз подряд в $2n$ партиях игры в орлянку. На основании изложенного в предыдущей главе эта вероятность равна

$$\frac{1}{2^{2n}} C_{2n}^{n} = \frac{(2n)!}{2^{2n}(n!)^2}.$$

Пользуясь формулой Стирлинга, мы получим для искомой величины

$$\frac{(2n)^{2n} e^{-2n} \sqrt{4\pi n}\left(1+\frac{1+\theta_{2n}}{24n}\right)}{2^{2n} n^{2n} e^{-2n} 2\pi n\left(1+\frac{1+\theta_n}{12n}\right)^2},$$

или, если произвести сокращения и заменить через $1+\varepsilon_n$ множитель, стремящийся к единице при беспредельном возрастании n (ε_n стремится при этом к нулю), получим

$$\frac{1+\varepsilon_n}{\sqrt{\pi n}}.$$

Чтобы представить себе степень приближения, возьмем, например, $n = 5$; вероятность в этом случае окажется равной $\frac{252}{1024}$; соответственное значение ε_n отрицательно и близко к 0,01.

Если мы примем во внимание, что $\sqrt{\pi}$ приблизительно $= 1{,}772$, то для $n = 100$ вероятность 100 раз получить орла при 200 подбрасываниях монеты равна:

$$\frac{1}{10 \times 1{,}772} = \frac{1}{17{,}72},$$

т. е. после округления, $\frac{1}{18}$. Вероятность получить 10 000 раз орла при 20 000 подбрасываниях монеты точно таким же путем вычисляется почти точно, равной $\frac{1}{177}$.

Отсюда видно, что вероятность для игрока в орлянку остаться без выигрыша и без проигрыша уменьшается, когда *заранее фиксированное* число партий увеличивается; вероятность того, что выигрыш и проигрыш уравновесятся при 200 партиях, равна $\frac{1}{12}$, вероятность такого же уравновешивания при 20 000 партиях равна лишь $\frac{1}{177}$, а при 2 000 000 подбрасываниях составляет всего лишь $\frac{1}{1772}$.

Можно отчетливо представить себе это уменьшение шансов равновесия выигрыша и проигрыша, вычисляя математическое ожидание для зрителя игры, которому оба игрока вносили бы определенную сумму всякий раз, когда возникало бы равновесие выигрыша и проигрыша. Мы сейчас произведем этот расчет, принимая приближенную формулу за точную; ошибка будет очень мала, так как ошибка ощутительна только при небольшом числе партий.

Искомое математическое ожидание есть сумма математических ожиданий в каждой из гипотез относительно наступления равновесий выигрыша и проигрыша, т. е. после 2 партий, после 4 партий, после 6 партий, ..., после $2n$ партий. Следовательно, его величина после $2n$ партий будет:

$$\frac{1}{\sqrt{\pi}}\left(1+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}+\ldots+\frac{1}{\sqrt{n}}\right).$$

Заменяя выражение в скобках интегралом

$$\int_0^n \frac{dx}{\sqrt{x}}=2\sqrt{n},$$

получим для приближенного значения математического ожидания величину

$$\frac{2}{\sqrt{\pi}}\sqrt{n},$$

т. е., приближенно, $1{,}128\sqrt{n}$. Таким образом, для 200 партий величина математического ожидания равна 11; для 800 партий она только удвоится, т. е. составит 22; а для 10 000 000 партий она будет около 1128.

31. АСИМПТОТИЧЕСКОЕ ЗНАЧЕНИЕ P_k ДЛЯ ОЧЕНЬ БОЛЬШИХ ЗНАЧЕНИЙ n

Применим формулу Стирлинга для вычисления приближенных значений P_k при очень больших n; такое приближенное значение часто называется асимптотическим. Введем вместо k переменную t, определяемую соотношением

$$k = np + t\sqrt{n},$$

из которого следует

$$n - k = nq - t\sqrt{n},$$

так как $p + q = 1$.

Заметим, что формуле Стирлинга можно придать логарифмическую форму:

$$\log n! = \left(n + \frac{1}{2}\right)\log n - n + \log\sqrt{2\pi} + \varepsilon_n',$$

где символ log означает натуральный логарифм, а ε_n' представляет количество вида

$$\varepsilon_n' = \frac{1 + \theta_n}{12n},$$

стремящееся к нулю при n возрастающем беспредельно.

В применении к выражению для P_k мы имеем:

$$\log P_k = \log n! - \log(np + t\sqrt{n})! - \log(nq - t\sqrt{n})! + \\ + k\log p + (n - k)\log q,$$

а, с другой стороны, мы можем написать

$$(n+1)\log n + k\log p + (n-k)\log q = \left(k + \frac{1}{2}\right)\log np + \\ + \left(n - k + \frac{1}{2}\right)\log nq - \frac{1}{2}\log pq,$$

что вытекает из элементарных свойств логарифмической функции[1].

Отсюда следует

$$\log P_k = -\frac{1}{2}\log npq - \log\sqrt{2\pi} - \left(np + t\sqrt{n} + \frac{1}{2}\right)\log\left(1 + \frac{t}{p\sqrt{n}}\right) - \\ - \left(nq - t\sqrt{n} + \frac{1}{2}\right)\log\left(1 - \frac{t}{q\sqrt{n}}\right) + \frac{1 + \theta_n}{12n} -$$

---

[1] Мы получим это тождество, если представим $n + 1$ в виде суммы $\left(k + \frac{1}{2}\right) + \left(n - k + \frac{1}{2}\right)$ и, кроме того, вычтем и прибавим к левой части $\log\sqrt{pq}$. — *Прим. перев.*

$$-\frac{1-\theta_k}{12k}-\frac{1+\theta_{n-k}}{12(n-k)}.$$

Полагая $\frac{t}{p\sqrt{n}}$ меньшим единицы, имеем

$$\left(np+t\sqrt{n}+\frac{1}{2}\right)\log\left(1+\frac{t}{p\sqrt{n}}\right)=\left(np+t\sqrt{n}+\frac{1}{2}\right)\left(\frac{t}{p\sqrt{n}}-\frac{t^2}{2p^2 n}+\right.$$
$$\left.+\frac{t^3}{3p^3n\sqrt{n}}-\ldots\right);\quad \left(nq-t\sqrt{n}+\frac{1}{2}\right)\log\left(1+\frac{t}{q\sqrt{n}}\right)=$$
$$=\left(nq-t\sqrt{n}+\frac{1}{2}\right)\left(\frac{t}{q\sqrt{n}}-\frac{t^2}{2q^2n}+\frac{t^3}{3q^3n\sqrt{n}}\ldots\right).$$

Таким путем без труда получим

$$\log P_k=-\frac{1}{2}\log 2\pi npq-\frac{t^2}{2pq}+\frac{t^3}{6\sqrt{n}}\left(\frac{1}{p^2}-\frac{1}{q^2}\right)+$$
$$+\frac{t}{2\sqrt{n}}\left(\frac{1}{p}+\frac{1}{q}\right)+\ldots,$$

причем члены опущенные значительно меньше членов сохраненных.

Предполагая $\frac{1}{pq\sqrt{n}}$ очень малым, мы считаем возможным пренебречь последним из написанных членов; предшествующий ему член $\frac{t^3}{6\sqrt{n}}\left(\frac{1}{p^2}-\frac{1}{q^2}\right)$ может его превысить, если t само превышает единицу. Легко, однако, видеть из соображений, в подробное рассмотрение которых мы здесь не входим, что этим членом практически можно пренебречь, учитывая, что $\frac{1}{pq\sqrt{n}}$ очень мало. Оба члена исчезают, когда $p=q$ (как при игре в орлянку).

Сохраняя лишь главные члены, мы получим

$$\log P_k=-\frac{1}{2}\log 2\pi npq-\frac{t^2}{2pq},$$

откуда и вытекает асимптотическая формула:

$$P_k=\frac{1}{\sqrt{2\pi npq}}e^{-\frac{t^2}{2pq}}. \tag{65}$$

Эта формула содержит в себе, как частный случай, формулу § 27, относящуюся к установлению равновесия выигрыша и проигрыша при игре в орлянку; в этом случае $p=q=\frac{1}{2}$, $t=0$ и нужно взять $2n$ вместо n. Тогда остается лишь

$$P_0=\frac{1}{\sqrt{\pi n}}$$

Выведенная выше асимптотическая формула пригодна лишь для очень малого по сравнению с $\sqrt{n}$; однако ее распространяют на все *возможные значения* t. Поступая так, не совершают заметной ошибки, *если* n *велико*; это обеспечивается тем лучше, чем больше p отличается как от очень малых значений, так и от очень близких к 1. Значения P (точнее их асимптотические значения) в действительности очень малы когда t принимает заметные размеры; и если даже относительная ошибка, которую мы совершаем, пользуясь асимптотическим значением может оказаться значительной, то ошибка абсолютная все же очень мала.

Следствия, вытекающие из формулы (65), очень важны; и стоит ли подчеркивать, что эта формула не содержит в себе ничего другого, как выводы из комбинаторного анализа и не предполагает иных гипотез.

Согласно основным определениям § 12 разность

$$h = k - np$$

называется отклонением. В частном случае, когда np целое число, h есть разность между зарегистрированным числом благоприятных событий и его наиболее вероятным числом np.

Величина

$$l = \frac{h}{n} = \frac{k}{n} - p = f - p$$

называется *относительным отклонением* это есть разность для данной серии испытаний между частостью и вероятностью наблюдаемого события.

Наконец, соотношение

$$\lambda = \frac{h}{\sigma} = \frac{h}{\sqrt{npq}} = \frac{k - np}{\sqrt{npq}} \tag{66}$$

называется *приведенным* (нормированным) *отклонением*; его значение — основоположно.

Внесение приведенного отклонения в формулу (65) приводит ее к виду:

$$P_k = \frac{1}{\sqrt{2\pi npq}}\, e^{-\frac{\lambda^2}{2}}. \tag{67}$$

Пусть k_1 и k_2 два значения k, различающиеся между собой на несколько единиц; соответственные значения для λ будут:

$$\lambda_1 = \frac{k_1 - np}{\sqrt{npq}} = \lambda(k_1),\ \lambda_2 = \frac{k_2 - np}{\sqrt{npq}} = \lambda(k_2),$$

а их разность

$$\lambda_2 - \lambda_1 = \frac{k_2 - k_1}{\sqrt{npq}}$$

очень мала, если n велико и если p не слишком мало, или не слишком близко к 1.

Установив все это, мы найдем вероятность того, что k примет одно из значений k_1, k_1+1, ..., k_2-1, число которых равно k_2-k_1; по правилу сложения вероятностей величина искомой вероятности равна:

$$P=\frac{1}{\sqrt{2\pi npq}}\left[e^{-\frac{1}{2}\lambda^2 k_1}+e^{-\frac{1}{2}\lambda^2(k_1+1)}+\ldots+e^{-\frac{1}{2}\lambda^2(k_2-1)}\right].$$

Здесь скобки содержат k_2-k_1 членов, очень близких друг к другу по своей величине, так как λ_2 и λ_1 между собой разнятся очень мало; поэтому мы можем написать, подставляя вместо λ' одну и ту же величину, взятую из промежуточных значений между λ_1 и λ_2:

$$(68)\qquad P=\frac{k_2-k_1}{\sqrt{2\pi npq}}\,e^{\frac{-\lambda'^2}{2}}=\frac{\lambda_2-\lambda_1}{\sqrt{2\pi}}\,e^{\frac{-\lambda'^2}{2}}.$$

Эта формула замечательна тем, что в нее входят одни лишь приведенные отклонения λ, в ней проявляется особая однородность проблем вероятностей.

Возьмем, например, игру в орлянку и положим $n=400$, $k_1=205$, $k_2=207$. Поставим задачу: определить вероятность того, что орел выпадет 205 или 206 раз при 400 подбрасываниях. В этом примере

$$\sigma=\sqrt{npq}=10;\quad \lambda_1=0{,}5;\quad \lambda_2=0{,}7.$$

Значения λ^2, встречающиеся в членах, стоящих в квадратных скобках, равны 0,25 и 0,36; мы можем приближенно взять $\lambda'^2=0{,}3$, откуда

$$P=\frac{0{,}2}{\sqrt{2\pi}}\,e^{-0{,}15}=0{,}068\ldots$$

Теперь возьмем $n=40\,000$, $k_1=20\,050$, $k_2=20\,070$. Значения λ_1, λ_2 и λ' остаются такими же, поэтому и *вероятность остается такой же.*

Этот пример хорошо выясняет, в чем состоит однородность условий. Было бы неправильно сказать, что вероятность получить 20 050 орлов при 40 000 подбрасываниях та же самая, как вероятность получить 205 орлов при 400 подбрасываниях; в действительности она примерно в 10 раз меньше. Вероятность выкинуть орла 205 раз при 400 подбрасываниях соответствует вероятности при 40 000 испытаниях выкинуть орла в количестве, заключенном между 20 050 и 20 059 (или, если угодно, между 20 045 и 20 054), т. е. в сумме десяти вероятностей, почти равных друг другу.

33. ПОЛЬЗОВАНИЕ НЕПРЕРЫВНОЙ ПЕРЕМЕННОЙ

Если n очень велико, то значения λ, изменяющиеся равными ступенями по $\frac{1}{\sqrt{2\pi npq}}$, можно считать за образующие непрерывный ряд;

это приводит к замене формулы (68), дающей величину P, следующей формулой:

$$P=\frac{1}{\sqrt{2\pi}}\int\limits_{\lambda_1}^{\lambda_2} e^{-\frac{\lambda^2}{2}}\,d\lambda. \tag{69}$$

Заметим к тому же, что в силу теоремы о среднем значении интеграла, значение данного интеграла можно получить из формулы (68), в которой λ' имеет значение, заключенное между λ_1 и λ_2. Преимущество формулы (68) состоит в том, что она не предполагает обязательно близости значений λ_1 и λ_2, ибо если предполагать $\lambda_1<\lambda<\lambda_2$ и если написать

$$P_{1,2}=\frac{1}{\sqrt{2\pi}}\int\limits_{\lambda_1}^{\lambda_2} e^{-\lambda^2/2}\,d\lambda,$$

$$P_{2,3}=\frac{1}{\sqrt{2\pi}}\int\limits_{\lambda_2}^{\lambda_3} e^{-\lambda^2/2}\,d\lambda,$$

то мы получим

$$P_{1,2}+P_{2,3}=\frac{1}{\sqrt{2\pi}}\int\limits_{\lambda_1}^{\lambda_3} e^{-\lambda^2/2}\,d\lambda.$$

Это можно обобщить на любое число последовательных интервалов изменений λ.

Вероятность для λ лежать между $-a$ и $+a$ равна:

$$P_a=\frac{1}{\sqrt{2\pi}}\int\limits_{-a}^{+a} e^{-\lambda^2/2}\,d\lambda. \tag{70}$$

Если n увеличивается беспредельно, то эта вероятность стремится к пределу:

$$P_\infty=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{-\frac{\lambda^2}{2}}\,d\lambda;$$

а мы знаем (§ 21), что значением этого интеграла является 1.

Этот результат требует некоторых замечаний. Если вместо пользования приближенными формулами мы применим точные, то ясно, что полный итог вероятностей для приведенных отклонений λ принять все значения, которые она только может принимать, должен равняться единице. В таком случае положим

$$\lambda=\frac{h}{\sqrt{npq}},$$

причем h может изменяться от $-np$ до $+nq$.

Следовательно, λ может варьировать

$$\text{от} - \sqrt{\frac{np}{q}} \text{ до } \sqrt{\frac{nq}{2p}},$$

а не от $-\infty$ до $+\infty$. Но как только n становится сколько-нибудь значительным, так $e^{-\lambda^2}$ делается настолько малым, что им можно пренебречь, когда λ принимает значения порядка $\sqrt{n}$; так, если

$$\frac{np}{2q} = 200,$$

то величиной $e^{-\frac{np}{2q}}$, меньшей, чем 10^{-45}, вполне, можно пренебречь. Поэтому мы можем, *не совершая никакой заметной ошибки*, распространить пределы интегрирования на весь интервал от $-\infty$ до $+\infty$. Общая вероятность, получаемая таким путем, в точности равна единице. Этот замечательный результат объясняется тем, что точная формула дает единицу, каковы бы ни были значения n, и что ошибка, присущая приближенной формуле, может быть меньше любой величины, если взять n достаточно большим. Эта приближенная формула не включает в себя n и поэтому может дать точное значение, равное единице.

34. КОЛОКОЛООБРАЗНАЯ КРИВАЯ ЛАПЛАСА — ГАУССА

Положим для $t > 0$

$$\psi(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{x^2}{2}} dx. \tag{71}$$

$\psi(t)$ после того, что было сказано выше, представляет собой такую величину, которая, чем больше число n, тем точнее определяет вероятность того, что приведенное отклонение $\lambda = \frac{h}{\sqrt{npq}}$ заключено между 0 и t.

Таблица I, приведенная в конце книги, дает величину $\psi(t)$ для значений t от 0 до 5.

Кривая, соответствующая уравнению

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

называется *нормальной приведенной кривой* или *колоколообразной кривой* Лапласа—Гаусса; ось Oy служит для нее осью симметрии, а ось Ox — ее асимптотой; $\psi(t)$ дает величину площади, ограниченной этой кривой, осью Ox и двумя прямыми: $x = 0$ и $x = t$.

На рис. 5 представлена эта кривая, в которой масштабная единица для ординат взята в 10 раз большей, чем для абсцисс; абсцисса $x = 1$ соответствует типическому отклонению закона приведенных (нормализованных) отклонений. Непосредственно видно, что точки K' и K с абсциссами -1 и $+1$ являются точками перегибов; их ординаты одинаковы и равны 0,241..., а максимальная ордината имеет высоту, равную 0,398...

Каковы бы ни были величины a и b ($a < b$), из § 33 следует, что вероятность отклонения находиться в пределах между a и b измеряет-

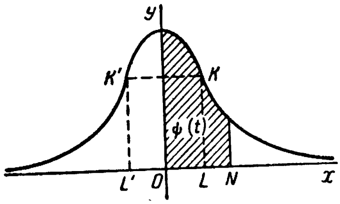

Рис. 5

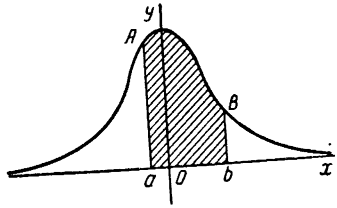

Рис. 6

ся площадью, ограниченной сверху колоколообразной кривой, снизу осью Ox, а с боков прямыми $x = a$ и $x = b$ (см. на рис. 6 заштрихованную площадь).

Вероятность приведенного отклонения, меньшего, чем a, или большего, чем b, равна:

$$P_r(\lambda < a \text{ или } \lambda > b) = 1 - Pr(a < \lambda < b);$$

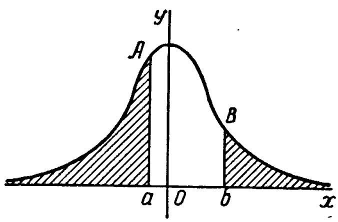

Рис. 7

она измеряется площадью, заштрихованной на рис. 7.

Значения вероятности $Pr\ [a < \lambda < b]$ можно сразу получить из значений $\psi(t)$, указанных в таблице.

Для положительных a и b вероятность $Pr[a < \lambda < b] = \psi(b) - \psi(a)$. Для отрицательных a и b $Pr(a < \lambda < b) = \psi(-a) - \psi(-b)$. И наконец, если a положительно, то

$$Pr[|\lambda| > a] = 1 - 2\psi(a),$$

например, вероятность того, что приведенное отклонение, взятое по своей абсолютной величине, превосходит 5, равно (см. табл. I):

$$1 - 2\psi(5) = 1 - 2 \times 0{,}499999713 < 6 \cdot 10^{-7}.$$

Такой величиной практически можно пренебречь, чем и объясняется то, что обычные таблицы $\psi(t)$ дают значения этой функции лишь для значений t между 0 и 5.

35. ПРИМЕРЫ ПОЛЬЗОВАНИЯ ТАБЛИЦЕЙ

Таблица $\psi(t)$ позволяет решать с хорошим приближением, поскольку n достаточно велико (при условии, что p не слишком близко по своей величине к 0 или к 1), различные практические задачи, связанные с исследованием определенной серии повторных испытаний. Приведем два примера.

З а д а ч а 6. *Дана серия из 400 партий игры в орлянку. Какова вероятность того, что орел выпадет меньше, чем 180 раз?*

Мы имеем здесь

$$n=400,\ p=q=\frac{1}{2},$$

откуда

$$np=200,\ \sigma=\sqrt{npq}=10,$$

приведенное отклонение, соответствующее 180 орлам и 220 решкам, равно:

$$\lambda=\frac{180-200}{10}=-2.$$

Поэтому мы имеем

$$Pr(\lambda<-2)=\frac{1}{2}\left(1-2\psi(2)\right)=0{,}0227\ldots,$$

что равносильно примерно 2 шансам из 100.

З а д а ч а 7. *Играют 800 партий в «экартэ». Какова вероятность, что король окажется на вскрышке более чем 110 раз?*

Здесь мы имеем

$$n=800,\ p=\frac{4}{32}=\frac{1}{8},\ q=\frac{7}{8},$$

откуда

$$np=100;\ \sigma=\sqrt{npq}=10\sqrt{\frac{7}{8}}.$$

Приведенное отклонение, соответствующее тому, чтобы король лег на вскрышке 110 раз, равно:

$$\lambda=\frac{110-100}{10\sqrt{\frac{7}{8}}}=\sqrt{\frac{8}{7}}=1{,}07\ \text{(приблизительно)}.$$

При этих условиях мы получим:

$$Pr(\lambda>1{,}07)=\frac{1}{2}\left[1-2\psi(1{,}07)\right]=0{,}14\ \text{(приблизительно)},$$

что соответствует 14 шансам из 100.

План решения остается таким же для всех задач такого рода; сперва выписывают приведенное отклонение, а затем ищут решение простым считыванием с таблицы $\psi(t)$.

Часто оказывается полезным хранить в памяти следующие цифры: вероятность того, что приведенное отклонение по своей абсолютной величине окажется большей, чем

1	не более		32%
2	»	»	5%
3	»	»	0,5%

Мы изложили в § 28 теорему Бернулли, согласно которой представляется маловероятным, чтобы при большом числе испытаний наблюденная часть какой-нибудь альтернативы, названной нами «благоприятной», заметно отличалась от ее вероятности.

Для доказательства мы воспользовались неравенством Бьенэмэ—Чебышева. Однако важно заметить, что в таком простом случае, как биномиальный закон, верхний предел, указываемый неравенством для вероятности отклонений, превышающих какую-либо заданную величину, может быть значительно улучшен с помощью функции $\psi(t)$.

Возьмем простой пример.

Задача 8. *Подбросим много раз одну и ту же игральную кость. Возможность выкинуть 6 равно $\frac{1}{6}$. Составим отношение между числом выпадений шестерки и общим числом подбрасываний. Сколько нужно сделать испытаний, чтобы вероятность для этих отношений отличаться от $\frac{1}{6}$ не более чем на 0,001, превышала бы $1 - 10^{-6}$?*

Мы имеем в этом случае

$$p = \frac{1}{6}, \quad q = \frac{5}{6}$$

и, пользуясь обозначениями § 28, получим

$$\varepsilon = 10^{-3}, \quad \alpha = 10^{-6}.$$

Неравенство Бьенэмэ—Чебышева гарантирует требуемый результат числом испытаний:

$$\frac{pq}{\alpha\varepsilon^2} = \frac{5}{36}10^{12}, \text{ т. е. около } 14 \cdot 10^{10} \text{ испытаний.}$$

Обратимся теперь к нормальному закону: серия в n испытаний, которая дает относительное отклонение в $\frac{1}{1000}$, должна соответствовать приведенному отклонению

$$\beta = \frac{k - np}{\sqrt{npq}} = \frac{n}{1000\sqrt{n\frac{1}{6}\cdot\frac{5}{6}}} = \frac{6\sqrt{n}}{1000\sqrt{5}}.$$

Поставленное в задаче условие можно выразить неравенством

$$2\psi(\beta) > 1 - 10^{-6},$$

или

$$\psi(\beta) > 0{,}4999995.$$

Таблица показывает, что такой результат может получаться, начиная с $\beta = 4{,}5$; следовательно, начиная с числа испытаний, равного

$$\frac{5}{36}(4{,}5)^2\,10^6 = 2812500 \text{ испытаний.}$$

37. ЭМПИРИЧЕСКИЙ ЗАКОН СЛУЧАЙНЫХ СОБЫТИЙ

Все применения теории вероятностей опираются в основе своей на следующее положение.

В многочисленной серии испытаний благоприятное событие осуществляется с частостью, близкой по величине к вероятности. Небесполезно несколько остановиться на значении для практики и на философской стороне этого утверждения.

Это утверждение представляет собой постулат, его можно наименовать *эмпирическим законом случайного.*

Подтверждение, идущее со стороны теоремы Бернулли, не является доказательством. Теорема Бернулли исходит из существования вероятности p благоприятного события и из правил сложения и перемножения вероятностей и приводит к выводу, относящемуся лишь к *вероятности* расхождения определенной величины между p и частостью $\frac{k}{n}$.

Частость *стремится к вероятности* при беспредельном увеличении числа испытаний n; но то понятие предела, которое входит в это утверждение совсем не то, каким пользуются в анализе. Можно сказать вместе с Кантелли (Cantelli), что $\frac{k}{n}$ стремится к p *в смысле теории вероятностей.* Для того чтобы от этого понятия предела перейти к понятию классического предела, нельзя миновать ссылки на эмпирический закон случайного.

Следует остерегаться делать из этого тот вывод, что теория вероятностей со всеми своими утверждениями, которыми исчисление вероятностей приводится к более простым формам выкладок, не может дать ничего достоверного. На самом же деле, она позволяет в некоторых случаях, и притом случаях первостепенной важности, *предвидеть будущее событие с такой достоверностью, которую не может превзойти никакое человеческое знание.*

Событие, вероятность осуществления которого отличается от единицы на величину, крайне малую, порядка, скажем, 10^{-100}, должно рассматриваться как достоверное *в человеческом смысле.* Один из авторов[1] этой книги когда-то привел в качестве примера дублирование томов Национальной библиотеки обезьянами, стукавшими по клавишам пишущих машинок. Подобное событие, хотя с абстрактной точки зрения не содержит в себе ничего невозможного, все же в столь высокой степени маловероятно, что все люди на всем свете согласятся считать его практически невозможным. Достоверность этой невозможности дает до очевидности наглядный образ достоверности тех результатов, к которым приводит закон больших чисел в некоторых случаях своего применения.

Рассмотрим один очень простой пример: пусть A и B два одинаковых сосуда, наполненных одним и тем же газом при одинаковых темпе-

[1] См. Emile Borel. Le hasard. Русский перевод: Эм. Борель. Случай. Госиздат 1923, стр. 114.

ратуре, и давлении. Представим себе, что между обоими сосудами устанавливается сообщение, позволяющее газам смешиваться. Кинетическая теория газов утверждает, что по прошествии достаточно долгого времени, когда установится равновесие, вероятность для отдельной данной молекулы газа оказаться в сосуде A будет равна вероятности очутиться в сосуде B. Все происходит так, как если бы $2n$ молекул были распределены по сосудам A и B в зависимости от результатов $2n$ партий игры в орлянку.

Примем для большей наглядности, что $n = 2 \times 10^{24}$, тогда типическое отклонение окажется равным

$$\sigma = \sqrt{2npq} = 10^{12}.$$

В таком случае вероятность получить отклонение, превышающее 100σ, равна 0, (2 173) 2688 (цифры, взятые в скобки, показывают, сколько нулей следует поставить после запятой): малость этой величины невообразима. Между тем такое отклонение вызвало бы разность на 10^{14} молекул между сосудами A и B, т. е. неоднородность порядка 10^{-10} в десять раз меньшую, чем в одну миллиардную, которую невозможно даже уловить экспериментально.

Отсюда видно, как закон случайного приводит к однородному состоянию при смешении двух газов.

38. ФУНКЦИЯ $\theta(x)$

Во всем предыдущем изложении, говоря о биномиальном законе, мы называли приведенным отклонением отношение действительного отклонения $k - np$ к *типическому отклонению* $\sigma = \sqrt{npq}$. В предыдущих изданиях этой книги приведенным отклонением называлось отношение действительного отклонения в собственном смысле слова к *единичному отклонению* или к *единице отклонения*

$$u = \sqrt{2npq}.$$

Вследствие этого нормальному закону отклонений была придана несколько иная форма по сравнению с той, какая изучалась в настоящем издании. Нам представляется необходимым посвятить несколько слов терминологии принятой в теории отклонений и обозначениям, основанным на этой единице; ибо эти обозначения применяются во многих трудах и статьях, которые нисколько не теряют в своем значении и остаются во всем остальном пригодными для пользования ими в исследовании ошибок наблюдений.

Если положить $\mu = \frac{k - np}{\sqrt{2npq}}$, сохраняя символ λ для отношения $\frac{k - np}{\sqrt{npq}}$, то мы будем иметь соотношение

$$\lambda = \mu\sqrt{2}, \tag{72}$$

из которого вытекает равенство

$$Pr\,[a < \lambda < b] = Pr\left[\frac{a}{\sqrt{2}} < \mu < \frac{b}{\sqrt{2}}\right],$$

а так как первая часть равенства имеет значение, равное (по § 69)

$$Pr\,[a < \lambda < b] = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}\,dx,$$

то отсюда мы выводим (произведя замену переменных $x = t\sqrt{2}$ и полагая $a = \alpha\sqrt{2}$ и $b = \beta\sqrt{2}$), что

$$Pr\,[\alpha < \mu < \beta] = \frac{1}{\sqrt{\pi}}\int_\alpha^\beta e^{-t^2}\,dt. \tag{73}$$

Через $\Theta(x)$ обозначают для $x > 0$ интеграл

$$\Theta(x) = \frac{1}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt = 2\psi\left(x\sqrt{2}\right). \tag{74}$$

Таблица II в конце книги дает значения этой функции, а также функции $1 - \Theta(x)$; пользование ею совершенно аналогично пользованию таблицей I.

39. ОТКЛОНЕНИЕ НАИБОЛЕЕ ВЕРОЯТНОЕ, ОТКЛОНЕНИЕ ВЕРОЯТНОЕ И ОТКЛОНЕНИЕ МЕДИАННОЕ

Элементарная вероятность нормального закона отклонений имеет следующий вид:

$$\frac{1}{\sqrt{2\pi}}e^{-\frac{\lambda^2}{2}}\,d\lambda = \frac{1}{\sqrt{\pi}}e^{-\mu^2}\,d\mu.$$

Эта вероятность имеет максимальную величину, когда

$$\lambda = \sqrt{2}\,\mu = 0 \ \left(\text{сохраняя соотношение } \lambda = \sqrt{2}\,\mu\right),$$

т. е. *наиболее вероятно то отклонение, которое равно нулю.*

Важно вполне понять смысл этого утверждения. Допустим, что мы набрали большую серию, членами которой являются группы из 200 партий игры в орлянку, и что для каждой группы установлено, сколько раз монета выпадала орлом. Когда орел выпадал 100 раз, отклонение равнялось нулю, когда орел выпадал 101 раз, отклонение равнялось $+1$, а когда число орлов было 99, то отклонение равнялось -1. При большом числе групп чаще всего будут встречаться отклонения, равные нулю; они будут в большем числе , чем отклонения $+1$ или также -1, но все же нулевые отклонения будут *менее многочисленны, чем отклонения $+1$ и -1, взятые вместе.*

Вместе с тем формулы сочетаний дают более точную, чем приближенные формулы, величину отношения между вероятностью отклонения, равного нулю, и отклонением, равным $+1$. Это отношение равно

$$C_{200}^{100} : C_{200}^{101} = \frac{101}{100} = 1{,}01.$$

Часто называют вероятным отклонением или средним отклонением *математическое ожидание абсолютной величины отклонений.*

Средняя величина приведенного отклонения μ равна, таким образом, средней величине приведенных отклонений, взятых по их абсолютным значениям:

(75) $$\frac{2}{\sqrt{\pi}} \int_0^{+\infty} te^{-t^2} dt = \frac{1}{\sqrt{\pi}} = 0{,}5642\ldots$$

Средняя величина отклонения в подлинном виде $k - np$ представляет собой произведение ее величины на единичное отклонение u, так, для серии из 20 000 партий игры в орлянку средняя величина отклонения равна $\frac{100}{\sqrt{\pi}} = 56,\ 42\ldots$

Принято называть *медианным отклонением* такое, которое служит гранью, делящей отдельные конкретные отклонения на две равночисленные группы. Этому отклонению соответствует, следовательно, такое значение μ, для которого

$$\Theta(\mu) = \frac{1}{2}.$$

Из таблицы видно, что эта величина μ равна 0,4769.

Медианное отклонение у артиллеристов называется «вероятным отклонением».

III. РАЗЛИЧНЫЕ ДОПОЛНЕНИЯ К ТЕОРИИ НОРМАЛЬНЫХ СЛУЧАЙНЫХ ПЕРЕМЕННЫХ И ПОВТОРНЫХ ИСПЫТАНИЙ

40. СЛОЖЕНИЕ НОРМАЛЬНЫХ СЛУЧАЙНЫХ ПЕРЕМЕННЫХ, КОГДА ОНИ ВЗАИМНО НЕЗАВИСИМЫ

Пусть нам даны две случайные независимые между собой переменные X_1 и X_2, каждая из которых подчиняется нормальному приведенному закону. Требуется определить закон вероятностей случайной переменной, определенной выражением

$$X = X_1 \cos\varphi + X_2 \sin\varphi,$$

в котором угол φ нам дан.

Примем X_1 и X_2 за прямоугольные координаты точки M (рис. 8). Согласно нашим предположениям и правилу перемножения вероятностей вероятность

$$Pr\,[x_1 \leqslant X_1 < x_1 + dx_1,\ x_2 \leqslant X_2 < x_2 + dx_2]$$

равна произведению

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(x_1^2\right)} dx_1 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2} dx_2,$$

откуда для плотности вероятности в точке M следует выражение

(76) $$\frac{1}{2\pi} e^{-\frac{1}{2}\left(x_1^2+x_2^2\right)}.$$

Пусть (D) — прямая, уравнение которой

$$X_1 \cos\varphi + X_2 \sin\varphi = a.$$

Она перпендикулярна к оси $0X$, получаемой путем поворота амплитуды φ около 0, в силу чего уравнением прямой относительно этой оси $0X$ и оси $0Y$ к ней перпендикулярной оказывается $X = a$.

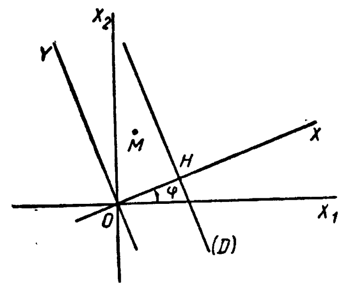

Рис. 8

Вероятность $Pr[x \leqslant a]$ при механическом представлении измеряется общей массой, распределенной на полуплоскости $X \leqslant a$ с плотностью (76). Как и для всех точек M этой полуплоскости $x_1^2 + x_2^2 = x^2 + y^2$, мы имеем соотношение

$$Pr[X \leqslant a] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx,$$

или, учитывая, что интеграл, взятый по dy имеет значение $\sqrt{2\pi}$,

(77) $$Pr[X \leqslant a] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx.$$

Из этого мы видим, что *случайная переменная X следует приведенному нормальному закону.*

Общий случай. Теперь пусть T_1 и T_2 — случайные переменные, подчиненные нормальным законам, со средними отклонениями m_1 и m_2 и с типическими отклонениями σ_1 и σ_2; займемся изучением случайной переменной, связанной с T_1 и T_2 соотношением

(78) $$U = a_1 T_1 + a_2 T_2,$$

в котором a_1 и a_2 означают два заданных числа.

Случайные переменные

$$X_1 = \frac{T_1 - m_1}{\sigma_1}, \quad X_2 = \frac{T_2 - m_2}{\sigma_2}$$

следуют приведенному нормальному закону и U может быть выражено как функция X_1 и X_2 соотношением

$$U = a_1 \sigma_1 X_1 + a_2 \sigma_2 X_2 + a_1 m_1 + a_2 m_2.$$

Введем случайную переменную

(79) $$X = \frac{U - a_1 m_1 - a_2 m_2}{\sqrt{a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2}} = \frac{a_1 \sigma_1 X_1 + a_2 \sigma_2 X_2}{\sqrt{a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2}};$$

тогда можно назначить такой угол φ, чтобы

$$\cos \varphi = \frac{a_1 \sigma_1}{\sqrt{a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2}}, \quad \sin \varphi = \frac{a_2 \sigma_2}{\sqrt{a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2}},$$

и мы получим, согласно вышеизложенному, что

$$X = X_1 \cos \varphi + X_2 \sin \varphi$$

следует *приведенному нормальному закону.*

Положим в тех же условиях

$$m = a_1 m_1 + a_2 m_2,$$

$$\sigma = \sqrt{a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2}, \quad \alpha = a\sigma + m;$$

затем в интеграле (77) произведем замену переменных

$$t = a\sigma + m,$$

отсюда получим соотношение:

(80) $$Pr[U \leqslant a] = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-\frac{1}{2}\left(\frac{t-m}{\sigma}\right)^2} dt.$$

откуда заключаем, что *U следует нормальному закону со средне[м]* $m = a_1 m_1 + a_2 m_2$ и с *типическим отклонением*

$$\sigma = \sqrt{a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2}.$$

Переход к случаю, когда

$$U = a_1 T_1 + a_2 T_2 + \ldots + a_n T_n,$$

осуществляется шаг за шагом.

41. ПРИМЕНЕНИЕ К УРНОВОЙ СХЕМЕ ПУАССОНА

Имеем две урны; одна содержит a_1 белых шаров и b_1 черных, другая a_2 белых и b_2 черных. Извлекаем n_1 шаров из первой урны [и] n_2 шаров из второй (каждый раз возвращая взятый шар обратно [в] урну). Мы имеем здесь $p_1 = \frac{a_1}{a_1 + b_1}$ и $p_2 = \frac{a_2}{a_2 + b_2}$, причем за бла[го]приятное событие будем считать извлечение белого шара.

Обозначим через X_1 и X_2 случайные переменные, представляющи[е] собой благоприятные события, осуществляющиеся при n_1 извлечения[х] из первой урны и n_2 извлечениях из второй. X_1 и X_2 представляют собой две независимые случайные переменные, следующие нормальному закону со средними $m_1 = p_1 n_1$ и $m_2 = p_2 n_2$ и с типическими отклонениями

$$\sigma_1 = \sqrt{n_1 p_1 q_1}\,,\quad \sigma_2 = \sqrt{n_2 p_2 q_2}.$$

Случайная переменная

$$X = X_1 + X_2 \tag{81}$$

показывает общее число благоприятных случаев; она согласно § 40 следует нормальному закону со средней

$$m = n_1 p_1 + n_2 p_2 \tag{82}$$

и с типическим отклонением

$$\sigma = \sqrt{n_1 p_1 q_1 + n_2 p_2 q_2}. \tag{83}$$

Новая случайная переменная имеет, следовательно, вид:

$$\lambda = \frac{X - m}{\sigma} \tag{84}$$

и следует, таким образом, приведенному нормальному закону.

42. ЗАМЕНА ДВУХ ГРУПП ИСПЫТАНИЙ ОДНОЙ ГРУППОЙ

Предположим, что p и q определяются соотношениями

$$\begin{aligned} np &= n_1 p_1 + n_2 p_2, \\ nq &= n_1 q_1 + n_2 q_2, \end{aligned} \tag{85}$$

в которых n_1, n_2, p_1, q_1, p_2, q_2 сохраняют значения, какие им были даны в предшествующем параграфе, так что

$$p + q = 1.$$

Если произвести n испытаний, в которых вероятность благоприятного события равна p, то математическое ожидание числа благоприятных событий равно np; ее значение — та же величина m, как в двух группах, рассмотренных ранее.

Напротив, типическое отклонение не будет тем же; его величина окажется равной

$$\Sigma = \sqrt{npq}$$

вместо

$$\sigma = \sqrt{n_1 p_1 q_1 + n_2 p_2 q_2}.$$

Согласно (85) мы имеем:

$$npq = \frac{(n_1 p_1 + n_2 p_2)(n_1 q_1 + n_2 q_2)}{n_1 + n_2}.$$

откуда следует, что

$$n\left(\sum^{2}-\sigma^{2}\right)=\left(n_{1} p_{1}+n_{2} p_{2}\right)\left(n_{1} q_{1}+n_{2} q_{2}\right)-$$
$$-\left(n_{1}+n_{2}\right)\left(n_{1} p_{1} q_{1}+n_{2} p_{2} q_{2}\right)$$

из чего после некоторых выкладок получим:

$$\sum^{2}-\sigma^{2}=\frac{n_{1} n_{2}}{n}\left(p_{1}-p_{2}\right)\left(q_{2}-q_{1}\right),$$

а так как

$$\left(q_{2}-q_{1}\right)=\left(1-p_{2}\right)-\left(1-p_{1}\right)=p_{1}-p_{2},$$

то получается, что

$$\sum^{2}-\sigma^{2}=\frac{n_{1} n_{2}}{n}\left(p_{1}-p_{2}\right)^{2}. \tag{86}$$

Отсюда видно, что если p_1 отлично от p_2 или, говоря другими словами, если две группы испытаний между собой существенно различаются, то Σ^2 всегда больше σ^2, т. е. что *типическое отклонение в объединенной группе испытаний всегда больше*[1].

Это правило очевидно справедливо в предельном случае, когда единая группа из $n = n_1 + n_2$ испытаний замещает серию из n_1 извлечений, производимых из урны, содержащей одни белые шары, и n_2 извлечений из другой урны, содержащей только черные шары. В этом случае имеем

$$p_{1}=1,\ q_{1}=0;\ p_{2}=0,\ q_{2}=1,$$

откуда согласно (85) $n_1 = np$ и $n_2 = nq$.

Очевидно, что суммарная группа извлечений, сделанных в таких условиях, не будет иметь отклонений: в ней $\sigma = 0$; тогда как

$$\sum^{2}=\frac{n_{1} n_{2}}{n_{1}+n_{2}}.$$

Положим теперь

$$n_{1}=n_{2},\ p_{1}=q_{2},\ p_{2}=q_{1},$$

отсюда согласно (85) получается

$$p=q=\frac{1}{2},\ n=2 n_{1},\ \sum^{2}=n p q=\frac{n_{1}}{2}.$$

В этом случае мы имеем

$$\frac{\sum^{2}-\sigma^{2}}{\sum^{2}}=\left(p_{1}-p_{2}\right)^{2};$$

[1] В рассмотренном случае числа извлечений из каждой урны — участницы в опыте — остаются неизменными от серии к серии, тогда как для объединенной серии содержание урн смешивается в общей урне. В этих условиях вывод справедлив. (Случай вероятности постоянного состава. — Вл. Борткевич). — *Прим. перев.*

предполагая же $p_1 = \frac{3}{4}$ и, следовательно, $p_2 = \frac{1}{4}$, получим

$$\Sigma = \frac{2\sigma}{\sqrt{3}}.$$

Следовательно, если, вместо того чтобы производить извлечения из одной (общей) урны, содержащей белые и черные шары поровну, мы произведем то же число извлечений, пользуясь поочередно двумя урнами, из которых первая содержит втрое больше белых шаров, чем черных, а вторая урна втрое больше черных, чем белых, то типическое отклонение в первом случае равно типическому отклонению во втором случае, умноженному на $\frac{2}{\sqrt{3}}$, т. е. приблизительно на 1,15.

Эти расчеты имеют большое значение в применении теории вероятностей к статистике. Их можно шаг за шагом распространить на случай многих урн.

Предположим, что производятся μ групп испытаний, причем числа испытаний в каждой из групп равны соответственно n_1, n_2 ..., n_μ; вероятности благоприятных событий — p_1, p_2, ..., p_μ; а вероятности событий противоположных — q_1, q_2, ..., q_μ.

Если рассматривать такую совокупность испытаний, то типическое отклонение получается согласно формуле:

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_\mu^2, \tag{87}$$

полагая же

$$\sigma_1^2 = n_1 p_1 q_1, \quad \sigma_2^2 = n_2 p_2 q_2, \ldots, \quad \sigma_\mu^2 = n_\mu p_\mu q_\mu, \tag{88}$$

т. е. принимая σ_1, σ_2, ..., σ_μ за типические отклонения соответственных групп, взятых каждая в отдельности.

Если мы обозначим через

$$n_1 p_1 + n_2 p_2 + \ldots + n_\mu p_\mu + \lambda\sigma \tag{89}$$

число благоприятных событий, то приведенное отклонение λ подчиняется приведенному нормальному закону, иначе сказать: вероятность того, что отклонение окажется заключенным в пределах x и $x + dx$ дается выражением

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Объединенная серия испытаний, в которой вероятности благоприятных и противоположных событий равны p и q, будет равнозначна предыдущей совокупности из μ групп, если

$$\begin{aligned} n &= n_1 + n_2 + \ldots + n_\mu \\ np &= n_1 p_1 + n_2 p_2 + \ldots + n_\mu p_\mu, \\ nq &= n_1 q_1 + n_2 q_2 + \ldots + n_\mu q_\mu. \end{aligned} \tag{90}$$

Однако типическое отклонение в объединенной серии будет большим, чем σ; расчет, аналогичный тому, с помощью которого была выведена формула (86), дает соотношение

$$\Sigma^2 - \sigma^2 = \frac{1}{n} \sum n_i n_k (p_i - p_k)^2, \tag{91}$$

где суммировка распространяется на всевозможные сочетания индексов i и k (безотносительно к их порядку).

Разность существенно положительна и может равняться нулю только в том случае, когда все вероятности p_k одинаковы.

43. ПОВТОРНЫЕ ИЗВЛЕЧЕНИЯ ИЗ УРНЫ, СОДЕРЖАЩЕЙ ЭЛЕМЕНТЫ ТРЕХ ВИДОВ. МУЛЬТИНОМИАЛЬНЫЙ ЗАКОН

Обозначим виды элементов через I, II, III, и для конкретности положим, что в урне имеются шары белые, синие и красные.

Пусть p_1, p_2, p_3 — соответственные пропорции шаров каждого цвета; мы имеем при этом соотношение

$$p_1 + p_2 + p_3 = 1. \tag{92}$$

Если мы проведем серию из n извлечений (при условии возвращения взятого шара обратно в урну), то пусть при этих испытаниях вынутся n_1 шаров I вида, n_2 шаров II вида и n_3 шаров III вида, причем

$$n_1 + n_2 + n_3 = n. \tag{93}$$

Обозначим через X_1, X_2, X_3 случайные переменные, представляющие собой числа возможных выходов шаров каждого вида в серии из n извлечений; эти случайные переменные независимы друг от друга, ибо всегда должно иметь место равенство

$$X_1 + X_2 + X_3 = n.$$

Пусть N — общее число шаров в урне, N_1, N_2 и N_3 — число шаров каждого вида. Если мы согласимся не различать между собой шары II и III видов, то урна будет содержать шары лишь двух видов: белые и цветные, для которых

$$p = \frac{N_1}{N} = p_1, \quad q = \frac{N_2 + N_3}{N} = p_2 + p_3;$$

и вероятность события $X_1 = n_1$ согласно (59) будет:

$$Pr\,[X_1 = n_1] = C_n^{n_1} p^{n_1} q^{n-n_1} = C_n^{n_1} p_1^{n_1} (p_2 + p_3)^{n_2+n_3}. \tag{94}$$

Теперь станем различать виды II и III, т. е. шары синие и красные. Вероятность того, что синих шаров появится n_2 (если *известно, что белых шаров было вынуто* n_1), одинакова с вероятностью в $n - n_1$ извлечениях из урны, содержащей N_2 синих и N_3 красных шаров, извлечь n_1 синих шаров. Она равна:

$$(95)\qquad Pr\,[X_2 = n_2/X_1 = n_1] = \\ = C_{n-n_1}^{n_2}\left(\frac{N_2}{N_2+N_3}\right)^{n_2}\left(\frac{N_3}{N_2+N_3}\right)^{n_3},$$

или ее можно записать иначе:

$$C_{n_2+n_3}^{n_2}\frac{p_2^{n_2}\,p_3^{n_3}}{(p_2+p_3)^{n_2+n_3}}.$$

Вероятность события $(X_1 = n_1,\ X_2 = n_2,\ X_3 = n_3)$, если применить правило перемножения вероятностей и произведение

$$C_n^{n_1}\,C_{n_2+n_3}^{n_2}$$

заменить выражением

$$\frac{n!}{n_1!\,n_2!\,n_3!},$$

будет равна

$$(96)\qquad T = \frac{n!}{n_1!\,n_2!\,n_3!}\,p_1^{n_1}\,p_2^{n_2}\,p_3^{n_3},$$

и мы видим, что различные значения T представляют собой члены разложения тринома

$$(p_1+p_2+p_3)^n.$$

Обобщение этого вывода на случай урны с шарами более трех цветов приводит к подобному же результату. Таким путем мы находим закон вероятностей для двух, трех... переменных, который называется *законом мультиномиальным*.

44. АСИМПТОТИЧЕСКИЕ ЗНАЧЕНИЯ T

Как в § 31, так и здесь, мы воспользуемся формулой Стирлинга; нужно для этого предположить, что n_1, n_2, и n_3 — числа очень большие.

После простых преобразований, ход которых читателю предоставляется выяснить самому, выражение (96) с помощью формулы Стирлинга приводится к виду:

$$T = A\left(\frac{np_1}{n_1}\right)^{n_1+\frac{1}{2}}\left(\frac{np_2}{n_2}\right)^{n_2+\frac{1}{2}}\left(\frac{np_3}{n_3}\right)^{n_3+\frac{1}{2}},$$

где

$$(97)\qquad A = \frac{1}{2\pi n\sqrt{p_1 p_2 p_3}}.$$

Переводя это на язык логарифмов, мы получаем

$$(98)\qquad \log T - \log A = \sum_{i=1}^{i=3}\left(n_i+\frac{1}{2}\right)\log\frac{np_i}{n_i}.$$

Примем, что ($i = 1, 2, 3$)

$$x_i = \frac{n_i - np_i}{\sqrt{n}}; \tag{99}$$

причем $x_1 + x_2 + x_3 = 0$. Тогда формула (98) преобразуется:

$$\log T - \log A = -\sum_i \left(np_i + x_i\sqrt{n} + \frac{1}{2}\right) \log\left(1 + \frac{x_i}{p_i\sqrt{n}}\right);$$

если же в правой части равенства каждый логарифм заменить его ра ложением в ряд подобно тому, как это было сделано в § 28, то м после упрощений, получим

$$\log T - \log A \sim -\frac{1}{2}\left(\frac{x_1^2}{p_1} + \frac{x_2^2}{p_2} + \frac{x_3^3}{p_3}\right),$$

откуда и выводится асимптотическое выражение

$$T = \frac{1}{2\pi n\sqrt{p_1 p_2 p_3}} e^{-\frac{1}{2}\left(\frac{x_1^2}{p_1} + \frac{x_2^2}{p_2} + \frac{x_3^2}{p_3}\right)}. \tag{100}$$

45. ПОЛЬЗОВАНИЕ НЕПРЕРЫВНЫМИ ПЕРЕМЕННЫМИ

Пусть α_1 и β_1 представляют собой два целых числа, отличающие одно от другого небольшим числом единиц; пусть α_2 и β_2 — друга пара целых чисел, взятых при том же условии (разности $\beta_1 -$ и $\beta_2 - \alpha_2$ могут быть различными).

По правилу сложения вероятностей вероятность того, что X_1 пр мет одно из значений $\alpha_1, \alpha_1 + 1, \ldots, \beta_1 - 1$ и что X_2 примет одно значений $\alpha_2, \alpha_2 + 1, \ldots, \beta_2 - 1$, равна сумме

$$\sum_{n_1=\alpha_1}^{\beta_1-1} \sum_{n_2=\alpha_2}^{\beta_2-1} Pr\,[X_1 = n_1,\ X_2 = n_2], \tag{101}$$

состоящей из $(\beta_1 - \alpha_1)(\beta_2 - \alpha_2)$ слагаемых.

Припишем каждому значению n_1 и каждому значению n_2 соответс венные значения x_1 и x_2, определяемые формулой (99), и положи

$$\Delta x_1 = x_1(\beta_1) - x_1(\alpha_1) = \frac{\beta_1 - \alpha_1}{\sqrt{n}};$$

$$\Delta x_2 = \frac{\beta_2 - \alpha_2}{\sqrt{n}};$$

Δx_1 и Δx_2 — весьма малы, так как n очень велико и разности $\beta_1 - \alpha_1$, $\beta_2 - \alpha_2$ незначительны; отсюда следует, поскольку n_1 и n_2 изм няются в соответствующих им границах, что x_1, x_2, а следовательн и x_3 колеблются весьма мало.

Поэтому мы получим приближенное значение для вероятности (101), заменяя в правой части каждое из слагаемых через $Pr[X_1 = n_1, X_2 = n_2]$ и получая таким путем

$$(\beta_1-\alpha_1)(\beta_2-\alpha_2) Pr\,[X_1=\alpha_1,\ X_2=\alpha_2],$$

откуда, подставляя $\sqrt{n}\Delta x_1$ вместо $(\beta_1-\alpha_1)$, а $\sqrt{n}\Delta x_2$ вместо $(\beta_2-\alpha_2)$ и пользуясь формулой (100), мы и находим выражение

(102) $$\frac{1}{2\pi\sqrt{p_1 p_2 p_3}} e^{-\frac{1}{2}\left(\frac{x_1^2}{p_1}+\frac{x_2^2}{p_2}+\frac{x_3^2}{p_3}\right)} \Delta x_1\,\Delta x_2,$$

для вероятности, что $\frac{X_1-np_1}{\sqrt{n}}$ будет лежать в пределах между x_1 и $x_1+\Delta x_1$, а $\frac{X_2-np_2}{\sqrt{n}}$ — в пределах между x_2 и $x_2+\Delta x_2$.

Если, таким образом, X_1 и X_2 показывают, сколько раз осуществляются события I и II в серии испытаний (извлечений с возвращением вынутого шара обратно в урну) с урной, содержащей троякого вида шары в пропорциях p_1, p_2, p_3, то система новых случайных переменных

$$\frac{X_1-np_1}{\sqrt{n}},\quad \frac{X_2-np_2}{\sqrt{n}}$$

может быть при большом числе n отождествлена с системой непрерывных переменных x_1, x_2 с элементарной вероятностью

(103) $$\frac{1}{2\pi\sqrt{p_1 p_2 p_3}} e^{-\frac{1}{2}\left(\frac{x_1^2}{p_1}+\frac{x_2^2}{p_2}+\frac{x_3^2}{p_3}\right)} dx_1\,dx_2,$$

где *переменные x_i связаны условием* $x_1+x_2+x_3=0$.

46. УКАЗАНИЯ ОТНОСИТЕЛЬНО ЗАКОНА χ^2

Поставим задачу определить в условиях § 45 закон вероятностей случайной переменной

(104) $$\chi^2=\frac{(X_1-np_1)^2}{np_1}+\frac{(X_2-np_2)^2}{np_2}+\frac{(X_3-np_3)^2}{np_3},$$

являющейся линейной функцией квадратов отклонений X_1, X_2, ..., X_n от их средних значений.

Так как X_1, X_2, X_3 имеют значения n_1, n_2 и $n_3=n-n_1-n_2$, соответствующие согласно (99) определенным значениям x_1, x_2 и $x_3 = -(x_1+x_2)$, то χ^2 принимает значение

$$\frac{x_1^2}{p_1}+\frac{x_2^2}{p_2}+\frac{x_3^2}{p_3}.$$

Так как вероятность системы (x_1, x_2) определяется элементарной вероятностью (103), то мы видим, что вероятность

$Pr(\chi^2 \leqslant a^2)$ измеряется двойным интегралом:

$$J = \frac{1}{2\pi\sqrt{p_1 p_2 p_3}} \iint e^{-\frac{1}{2}\left(\frac{x_1^2}{p_1}+\frac{x_2^2}{p_2}+\frac{(x_1+x_2)^2}{p_3}\right)} dx_1\, dx_2,$$

взятым на плоскости (x_1, x_2), отнесенной к системе прямоугольных координат, в области, ограниченной эллипсом (E), уравнение которого имеет вид:

$$\frac{x_1^2}{p_1}+\frac{x_2^2}{p_2}+\frac{(x_1+x_2)^2}{p_3}=a^2.$$

Чтобы вычислить J, мы предварительно определим бесконечно малую долю этого интеграла, заключенную в полоске между гомотетическими[1] по отношению к (E) эллипсами, имеющими центр в 0 и отношения подобия λ к $\lambda + d\lambda$.

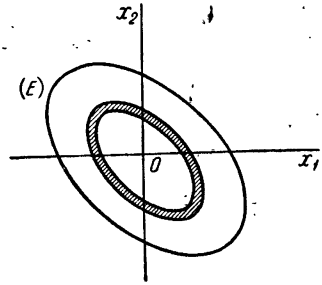

Рис. 9

Первый из этих эллипсов имеет уравнение

$$\frac{x_1^2}{p_1}+\frac{x_2^2}{p_2}+\frac{(x_1+x_2)^2}{p_3}=a^2\lambda^2,$$

а его площадь равна $\lambda^2 S$, если через S обозначить площадь, заключенную в эллипсе (E). Отсюда следует, что площадь упомянутой выше и заштрихованной на рис. 9 полосы, имеет величину $2\lambda S d\lambda$ и что доля dJ двойного интеграла, заключенная в этой полосе, дана выражением:

$$dJ = \frac{S}{\pi\sqrt{p_1 p_2 p_3}} e^{-a^2\frac{\lambda^2}{2}} \lambda\, d\lambda.$$

если же произвести замену переменной $a\lambda = u$, то

$$dJ = \frac{S}{\pi a^2\sqrt{p_1 p_2 p_3}} e^{-\frac{u^2}{2}} u\, du,$$

откуда получается, если принять во внимание, что λ изменяется от 0 до 1, а u от 0 до a, что

$$J = \frac{S}{\pi a^2\sqrt{p_1 p_2 p_3}} \int_0^a e^{-\frac{u^2}{2}} u\, du$$

Постоянный множитель $\frac{S}{a^2}$ определится весьма просто из того условия, что J стремится к 1, когда a возрастает неограниченно. Поэтому мы можем окончательно написать

[1] Гомотетическими, т. е. подобными и подобно расположенными. — *Прим. перев.*

$$(105)\qquad Pr(\chi_2 \leqslant a^2) = \frac{\int\limits_0^a e^{-\frac{u^2}{2}} u\,du}{\int\limits_0^{+\infty} e^{-\frac{u^2}{2}} u\,du}.$$

Так выражается закон, который в математической статистике называется *законом* χ^2 *при двух степенях свободы.*

47. ОБОБЩЕНИЕ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

Совсем нетрудно обобщить результаты, полученные в § 45 и 46 на неопределенно большое число случайных переменных. Мы ограничимся указаниями на результаты такого обобщения.

1) Для урны, содержащей шары k расцветок, вероятность получить в серии из n извлечений комбинацию из k окрасок так, чтобы первая окраска появилась бы в количестве n_1, вторая — в количестве n_2..., последняя — в количестве n_k, так что

$$n_1 + n_2 + \ldots + n_k = n,$$

такая вероятность выразится через члены мультинома:

$$(106)\qquad T = \frac{n!}{n_1!\, n_2! \ldots n_k!}\, p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k}.$$

2) Если $n_1, n_2, \ldots, n_k$ достаточно велики, то для T можно получить *асимптотическое значение*:

$$(107)\qquad \frac{1}{(2\pi n)^{\frac{k-1}{2}} \sqrt{p_1 p_2 \ldots p_k}}\, e^{-\frac{1}{2}\left(\frac{x_1^2}{p_1} + \frac{x_2^2}{p_2} + \ldots + \frac{x_k^2}{p_k}\right)};$$

в нем для $i = 1, 2, \ldots, k$; $x_i = \dfrac{n_i - np_i}{\sqrt{n}}$; $\sum x_i = 0$.

3) Случайная переменная

$$\frac{X_i - np_i}{\sqrt{n}} \quad (i = 1, 2, \ldots, k)$$

подчиняется закону, определенному с помощью элементарной вероятности:

$$(108)\qquad \frac{1}{(2\pi)^{\frac{k-1}{2}} \sqrt{p_1 p_2 \ldots p_k}}\, e^{-\frac{1}{2}\sum \frac{x_i^2}{p_i}}\, dx_1\, dx_2 \ldots dx_k.$$

при непременном условии, что

$$x_1 + x_2 + \ldots + x_k = 0.$$

4) Наконец, если положить

$$\chi^2 = \sum_i \frac{(X_i - np_i)^2}{np_i},$$

то χ^2 окажется случайной переменной, которая (при условии, что n достаточно велико) варьирует согласно закону вероятности:

$$(109) \qquad Pr\,[\chi^2 \leqslant a^2] = \frac{\int\limits_0^a e^{-\frac{u^2}{2}}\,du}{\int\limits_0^{+\infty} e^{-\frac{u^2}{2}}\,du},$$

который называется *законом χ^2 при $k-1$ степенях свободы.*

Вероятность $Pr(\chi^2 > a^2)$ дополнительная к предыдущей, выражается в случае $k-1=\nu$ степеней свободы следующей формулой, которую нетрудно получить из формулы (109) и в которой C_ν означает некоторую постоянную, величина которой определяется из условия, что $Pr(\chi^2 > 0) = 1$.

$$(110) \qquad Pr\,[\chi^2 > a^2] = C_\nu \int\limits_{a^2}^{+\infty} e^{-\frac{u^2}{2}} (u^2)^{\frac{\nu-2}{2}}\,du^2.$$

Читателю в виде упражнения предоставляется доказать, что для каждого целого положительного значения ν

$$m_1 = E(\chi^2) = \nu,$$
$$m_2 = E[(\chi^2)^2] = \nu(\nu+2),$$

так что

$$\mu_2 = m_2 - m_1^2 = 2\nu.$$

Существуют таблицы, дающие величину интеграла (110) для различных значений a^2 и ν. Таблица III, которую мы приводим в конце книги, дает эти величины лишь для значений ν от 1 до 30 и для тех a^2, для которых вероятность (110), что оно будет превзойдено, χ^2 равна 0,05 или 0,01. Такая сводка достаточна для наиболее обычных случаев применения метода χ^2.

48. ПРИМЕР

Следующий пример относится к закону Менделя (Gr. Mendel), лежащему в основе классической области применения теории вероятностей к естествоведению.

Если даны две пары аллеломорфных признаков (A, a) и (B, b), то по теоретическим соображениям можно ожидать, что при скрещивании фенотипы появятся в пропорциях

[1] См. G. Darmois. Statistique et applications. Ch. V.

$$p_1=\frac{9}{16}, \quad p_2=\frac{3}{16}, \quad p_3=\frac{3}{16} \text{ и } p_4=\frac{1}{16}.$$

В выборке из $n = 560$ особей оказалось

$$n_1=327, \ n_2=118, \ n_3=90 \text{ и } n_4=25$$

особей, принадлежащих к указанным выше типам.

Можно ли считать это экспериментально полученное распределение согласным с законом Менделя?

Приравняем исследованную популяцию к урне с шарами четырех цветов, из которой наудачу берется вышеприведенная выборка. Популяция предполагается достаточно многочисленной, так, чтобы выборку можно было бы считать не влияющей на относительные численности различных шаров в урне. Эту популяцию по отношению к выборке называют *генеральной совокупностью* (родительской популяцией).

Математические ожидания $E(X_i) = np_i$ здесь равны: $np_1 = 315$, $np_2 = 105$, $np_3 = 105$ и $np_4 = 35$.

Мы, следовательно, имеем

$$\chi^2=\frac{(327-315)^2}{315}+\frac{(118-105)^2}{105}+\frac{(90-105)^2}{105}+\frac{(25-35)^2}{35},$$

т. е. $\chi^2 = 7{,}07$ при $\nu = 4 - 1 = 3$ степенях свободы.

Из таблицы III мы находим для $\nu = 3$, что вероятность получить $\chi^2 > 7{,}81$ равна 5% и для $\chi^2 > 9{,}35$ равна 1%; отсюда мы заключаем, что *вероятность получить в силу чистого случая* $\chi^2 \geqslant 7{,}07$ *больше 5%*.

Согласно определению χ^2 для наиболее общего случая величина этого параметра при данных n и p_i тем значительнее, чем сильнее наблюденные значения n_i *отклоняются* от математических ожиданий np_i. Эмпирический закон случайности принуждает нас считать в том случае, когда эти расхождения слишком велики, что по крайней мере одна из наших гипотез неверна.

Но возникает, естественно, вопрос, какие же случайные значения χ^2 можно считать предельно допустимыми? *Практическое правило*, которого обычно придерживаются, состоит в следующем: если вероятность случайно превзойти наблюденную величину χ^2 *превосходит 5%*, то следует считать, что отклонения недостаточно велики, чтобы опорочить принятую систему предположений; про такие *отклонения* говорят, что *они несущественны* (*непоказательны*). Но если вероятность того, что величина χ^2, полученная из наблюдений, может быть превзойдена в силу случайности, *равна лишь 1%*, то следует заключить, что данные наблюдением расхождения слишком значительны, чтобы их можно было приписать влиянию одной лишь случайности выборки, и что по меньшей мере одно из сделанных предположений не справедливо.

Если вероятность, о которой речь шла выше, лежит между 5% и 1%, то следует воздержаться от заключений и продолжить наблюдения.

В примере, разобранном выше, вероятность получить $\chi^2 \geqslant 7{,}0$[?] при $\nu = 3$ превышает 5%, отсюда мы можем заключить, что наблю[денные] денные частости не опорочивают закон Менделя, но мы не говорим, что они его подтверждают.

49. ОЧЕНЬ МАЛЫЕ ВЕРОЯТНОСТИ И ЗАКОН ПУАССОНА

Приближения, которые мы выше допускали, переходя от биномиального закона к нормальному, становятся неприемлемыми, когда p или q оказываются очень малыми, если только число n при этом не чрезвычайно велико. Мы действительно в § 31 считали член $\frac{t}{pq\sqrt{n}}$ ничтожно малым, а это означает, что, полагая величину $\lambda = \frac{1}{\sqrt{pq}}$ равной нескольким единицам, членом $\frac{1}{\sqrt{npq}}$ можно пренебречь. Если же мы предположим, например, что $p = 10^{-6}$, то нужно брать величину n порядка 10^{12}, для того чтобы $\frac{1}{\sqrt{npq}}$ имело бы величину порядка $\frac{1}{1000}$.

Так как очень малые вероятности встречаются главным образом в молекулярной физике, то представляется чрезвычайно важным иметь в своем распоряжении такую приближенную форму, которая давала бы возможность рассчитывать P_k в случае очень малых, p без того, чтобы n было настолько большим, чтобы можно было применять теорию § 28. Напишем:

$$P_k = C_n^k p^k (1-p)^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!\,(1-p)^k} p^k (1-p)^n.$$

В такой форме

$$P_k = \frac{np\,(np-p)\ldots[np-(k-1)\,p]}{k!\,(1-p)^n}\left(1-\frac{np}{n}\right)^n$$

и, закрепив k, положим $np = \lambda$. Тогда получится, что

$$P_k = \frac{\lambda\,(\lambda-p)\ldots[\lambda-(k-1)\,p]}{k!\,(1-p)^n}\left(1-\frac{\lambda}{n}\right)^n.$$

Так как p очень мало, то λ, равная np, по сравнению с p очень велико; числитель вышенаписанной дроби мало отличается от λ^k, а множитель $(1-p)^k$ — от единицы. Множитель же $\left(1-\frac{\lambda}{n}\right)^n$ стремится к $e^{-\lambda}$ при беспредельном возрастании, и мы, таким образом, получаем приближенное выражение

$$(111) \qquad P_k = \frac{\lambda^k}{k!} e^{-\lambda},$$

в котором узнаем закон Пуассона, рассмотренный нами в § 18. Таблицы дают значения P_k как функции от λ и k.

УПРАЖНЕНИЯ К ГЛАВЕ ВТОРОЙ

1. В стране насчитывается 10 000 000 избирателей, из которых 5 500 000 принадлежат к партии A, и 4 500 000 принадлежат к партии B. Назначаются жребием 20 000 выборщиков. Какова вероятность того, что большинство выборщиков окажется сторонниками партии B?

Ответ: $\frac{1}{2}\left[1-\theta(10)\right]$.

2. Петр и Павел играют 1000 партий в орлянку со ставкой в 1 франк на каждую партию для каждого из игроков. Петр имеет m франков, а Павел имеет n франков. Расчет будет произведен лишь после того, как все 1000 партий будут сыграны. Какова вероятность того, что такой суммарный расчет окажется возможным?

Ответ: $$\frac{1}{\sqrt{\pi}}\int\limits_{\frac{m\sqrt{2}}{1000}}^{\frac{m\sqrt{2}}{1000}} e^{-t^2}\,dt.$$

3. *О разорении игроков.*

У Петра m франков, а у Павла — n франков. Они играют неограниченное число партий в такую игру, в которой вероятности выиграть для каждого из партнеров равны p и q. Если их ставки в игре составляют: для Петра a франков, а для Павла b франков, то какова вероятность того, что Петр будет разорен Павлом?

Эта задача одна из наиболее известных в теории вероятностей; она связана с работами де Муавра (de Moivre), Лагранжа (Lagrange), Ампера (Ampère) и совсем недавно с исследованиями по теории игр Руше (Rouché).

Ее можно разбирать общим *методом решения уравнений в конечных* разностях; значение этих уравнений в математической теории вероятностей очень велико, и для того чтобы интегрировать в конечных разностях Лаплас построил свою теорию производящих функций, которая в настоящее время представляет лишь ограниченный интерес, поскольку развились другие методы.

Если $f(x)$ вероятность того, что Петр разорится, обладая капиталом в x франков, то, применяя к следующей очередной партии правило сложения вероятностей, получим уравнение

$$(E) \qquad f(x)=pf(x+b)+qf(x-a),$$

которое представляет собой уравнение в конечных разностях с постоянными коэффициентами.

Мы принимаем без доказательств, что существует общее решение для (E) в форме

$$f(x) = A + B\alpha^x$$

при условии, что

$$(E') \qquad 1 = p\alpha + \frac{q}{\alpha}.$$

Так как

$$f(0) = 1, \; f(m+n) = 0,$$

то мы имеем

$$\left| \begin{array}{l} A + B = 1 \\ A + B\alpha^{m+n} = 0, \end{array} \right.$$

откуда

$$P = f(m) = \frac{\alpha^m(\alpha^n - 1)}{\alpha^{m+n} - 1}.$$

Уравнение (E') удовлетворяется при $\alpha = 1$, но это решение надо отбросить, приняв другое, отличное от 1. Например, при игре в рулетку, если Петр держит банк, $x = \frac{19}{37}$ и $y = \frac{18}{37}$.

Уравнение (E') напишется в виде:

$$1 = \frac{19}{37}\alpha + \frac{18}{37\alpha}.$$

и решением, отличным от единицы, служит $\alpha = \frac{18}{19}$. Предел P для n бесконечно большого (т. е. вероятность того, что банк никогда не будет сорван) равен:

$$\alpha^m = \left(\frac{18}{19}\right)^m$$

Эта вероятность очень мала, если m значительно.

Если игра ведется на справедливых условиях, то $p = q$ и (E') допускает двойное решение $\alpha = 1$ (т. е. оба решения совпадают). В этом случае P становится равным $\frac{n}{m+n}$ и стремится к 1 при n бесконечном.

Отсюда вытекает, что если Петр играет на справедливых условиях с любым партнером, пожелавшим вступить в игру (что равносильно тому, что его партнер обладает неограниченным капиталом), то разорение Петра наступит с несомненностью. Реальное существование банкомета показывает, что этот вывод перестает быть правильным, коль скоро *игра перестает вестись в условиях* справедливости и что какое-либо преимущество, как бы мало оно ни было, предотвращает неизбежность разорения (по Ж. Бертрану).

4. Если в формулах § 31 не пренебрегать членом $\frac{t(q-p)}{2pq\sqrt{n}}$, то можно положить

$$t+\frac{q-p}{2\sqrt{n}}=t',$$

если мы замечаем, что максимум для P_k приходится на $t'=0$, а не на $t=0$. Соответственное значение k будет:

$$k=np+t\sqrt{n},$$

$$k=np+\frac{p-q}{2},$$

т. е. что при $p \neq q$ ось симметрии кривой сдвинута на расстояние $\frac{q-p}{2}$. Вычислите при таких условиях, чему равна наибольшая величина P_k.

Ответ:

$$\frac{1}{\sqrt{2\pi npq}}e^{\frac{(p-q)^2}{8npq}}.$$

5. *Извлечение (без возвращения шара в урну) из урны, содержащей шары двух цветов.*

Пусть общее количество шаров равно N, белые и черные шары содержатся в пропорции $p=\frac{A}{N}$, $q=\frac{B}{N}$, n — количество последовательных извлечений без возвращения шара обратно в урну.

1) Показать, что вероятность вынуть k белых шаров равна:

$$P_k=\frac{n(n-1)\ldots(n-k+1)A(A-1)\ldots(A-k+1)B(B-1)\ldots(B-n+k+1)}{k!\,N(N-1)\ldots(N-n+1)}.$$

2) Пусть X — случайная переменная, которая может принимать значения $k=0, 1, 2, \ldots, n$ с вероятностями P_k, найденными в (1). Построить отношение

$$E(X)=np;\ E[X(X-1)]=\frac{n(n-1)A(A-1)}{N(N-1)};$$

вывести из них величину $E(X^2)$ и показать, что типическое отклонение имеет вид:

$$\sigma=\sqrt{n\frac{N-n}{N-1}pq}.$$

3) Показать, пользуясь формулой Стирлинга, что, положив $h=k-np$, мы получим для P_k асимптотическую величину

$$P_k=\frac{1}{\sqrt{2\pi n\frac{N-n}{N-1}pq}}e^{-\frac{h^2}{2n\frac{N-n}{N-1}pq}}.$$

6. *Опыт Велдона* (Weldon).

Велдон подбросил 26 306 раз 12 игральных костей, иначе сказать, 315 672 раза одну кость. Он отмечал в качестве благоприятного события каждый случай выпадения 5 или 6; опыт дал 106 602 благоприятных случая.

Показать, что нельзя считать, что кости были сделаны с полной точностью, ибо вероятность полученного результата была бы столь мала, что подобное предположение необходимо взять под сомнение для костей, которыми пользовался Велдон.

7. Принимая обозначения § 44, рассмотрим результаты $n_1 = 328$, $n_2 = 122$, $n_3 = 77$, $n_4 = 25$.

Вычислите χ^2 и сделайте выводы.

Ответ $P < 5\%$.

3. ГЕОМЕТРИЧЕСКИЕ ВЕРОЯТНОСТИ

50. ОБЩИЕ ЗАМЕЧАНИЯ

В первой главе нашей книги мы прежде всего дали определение вероятности случайного события, положив в основу этого определения предположение об элементарных равновероятных статочностях, данных в конечном числе; затем мы ввели общее понятие случайной переменной и привели несколько определений и положений относительно законов вероятностей, которым подчиняются отдельные случайные переменные или системы случайных переменных, рассматривая последовательно случаи вероятностей дискретных значений, счетно-бесконечных и непрерывных.

Эти общие понятия нашли свое применение в предыдущей (второй) главе, посвященной теории повторных испытаний. Говоря о некоторых проблемах дискретных вероятностей, рассмотренных в свете элементарных статочностей, данных в конечном числе, мы выяснили некоторые предельные законы, относящиеся к сериям из очень большого числа испытаний: закон Пуассона, являющийся законом вероятностей для счетно-бесконечного числа испытаний, и в особенности нормальный закон Лапласа—Гаусса, представляющийся законом для непрерывно меняющихся значений случайной переменной.

В настоящей (третьей) главе мы рассмотрим ворос: какие законы вероятностей непрерывных значений надо устанавливать при изучении проблем геометрических вероятностей, которые можно изложить следующим образом.

Пусть E и E' — два непрерывных множества геометрических элементов одного и того же рода, подчиненных условию, что все элементы E' входят в состав E. Если мы возьмем наудачу какой-либо элемент из множества E, то какова вероятность того, что он принадлежит E'?

Простейшие из этого рода задач относятся к положениям точек на прямой.

Берем наудачу точку M на отрезке прямой AB; какова вероятность того, что она попадет между точками C и D, на отрезке AB?

Представляется естественным принять условие, что вероятность для точки M попасть на какой-либо отрезок AB (рис. 10) пропорциональна длине этого отрезка. Если принять такое условие, то вероятность, что точка M окажется между C и D, равна $\frac{CD}{AB}$.

Выберем на прямой AB начало абсцисс и направление, принимаемое за положительное; если X — случайная абсцисса точки M и если a, b, c, d — абсциссы, соответствующие точкам A, B, C и D, то принятое условие позволяет написать:

$$Pr\,[c < X < d] = \frac{d-c}{b-a} = \int_c^d \frac{dx}{b-a},$$

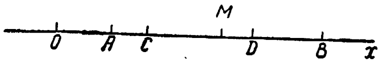

Рис. 10

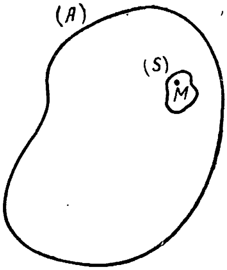

Рис. 11

т. е. положить, что случайная переменная X подчиняется равномерному закону, в котором плотность вероятности постоянна и равна $\frac{1}{b-a}$.

Но каким бы естественным ни казалось принятое выше условие, оно все же далеко не достаточно, чтобы избежать всех затруднений.

Действительно, пусть $Y = \varphi(X)$ — данная нам функция от X, представляющаяся непрерывной и монотонно возрастающей или монотонно убывающей, когда X изменяется от a до b. Связь между X и Y не имеет характера случайности: она достоверна, а это значит, что при непрерывности X, Y также непрерывен. Поэтому представляется вполне допустимым принять равенство

$$(112) \qquad Pr\,[c < X < d] = Pr\,[\varphi(c) < Y < \varphi(d)],$$

написанное в предположении, что Y функция возрастающая, когда X возрастает, и измеряет вероятность, указанную в правой части, отношением

$$P = \frac{\varphi(d) - \varphi(c)}{\varphi(b) - \varphi(a)}.$$

Однако, выбирая для Y подходящую форму $\varphi(X)$, можно получить для P значение *совершенно произвольное*; так, полагая, например, $a = 1$, $b = 6$, $c = 3$, $d = 4$, мы получим для P либо

$$P = \frac{d-c}{b-a} = \frac{1}{5},$$

либо же, взяв $Y = X^2$, получим

$$P = \frac{16-9}{25-1} = \frac{7}{24};$$

либо, взяв для $Y = \frac{1}{X}$, найдем

$$P = \frac{1/3-1/4}{1-1/6} = \frac{1}{10}.$$

Таким образом, мы находим три различных значения в зависимости от того, что для нашей переменной Y принимаем в качестве равномерного закона — закон, пропорциональный X, X^2 или $\frac{1}{X}$; или, говоря другими словами, принимаем X, X^2 или $\frac{1}{X}$ за *независимую переменную*.

В этом и состоит затруднение, которое мы старались выявить со всей возможной отчетливостью, но его практическая важность не слишком велика. Нужно просто-напросто (и этим не следует пренебрегать) *обращать внимание на те ошибки, которые могут повлечь за собой плохой выбор независимой переменной.* Однако в действительности этот выбор почти всегда с очевидностью вытекает из самой постановки задачи, когда дело идет о конкретной, а не об абстрактной проблеме, и можно почесть за простую шутку претензию изменить этот выбор переменной, являющийся лишь аналитической уловкой, не имеющей отношения к действительности. К тому же следует сказать, что выбор, предписываемый самой формулировкой задачи, имеет смысл только в том случае, когда эта формулировка настолько ясна и закончена, что допускает хотя бы попытку экспериментальной проверки результата.

52. ПОЛОЖЕНИЕ ТОЧКИ НА ПЛОСКОСТИ ИЛИ В ПРОСТРАНСТВЕ

Условность, которой мы сейчас займемся, касается того выбора независимой переменной в зависимости от постановки задачи, когда речь идет о положении точки уже не на прямой, а на плоскости или в пространстве; если точка M должна находиться на площади A, то вероятность для точки M оказаться внутри определенного участка (S) этой площади *пропорциональна величине площади* (S); она равна отношению $\frac{S}{A}$, если буквами S и A обозначить площади плоских фигур S и A (рис. 11).

Аналитически эта вероятность выражается отношением двух двойных интегралов:

$$\frac{\iint\limits_{S} dx\,dy}{\iint\limits_{A} dx\,dy}.$$

Если же точка M занимает место в пространстве внутри объема V, то для нее вероятность находиться в части U этого объема равна отношению $\frac{U}{V}$ и может быть вычислена как отношение двух тройных интегралов:

$$\frac{\iiint\limits_{U} dx\,dy\,dz}{\iiint\limits_{V} dx\,dy\,dz}.$$

Пользуясь языком эвклидовой геометрии n измерений, можно обобщить это определение на случай, когда число измерений превышает 3. Это обобщение часто оказывается весьма полезным в проблемах математической физики.

53. ОБЩЕЕ ПОНЯТИЕ ЭЛЕМЕНТАРНОЙ ВЕРОЯТНОСТИ

Замечания § 51, естественно, распространяются на многомерные проблемы, затронутые в § 52, и приводят, если подходить к теории геометрических вероятностей с чисто абстрактной точки зрения, к тому, чтобы ввести, как показал Пуанкаре (H. Poincaré), *произвольную положительную функцию* в определение вероятности.

Вернемся прежде всего к тому случаю, где речь идет о положении точки на прямой, которое определяется абсциссой x. Тогда говорят, что для точки, подчиненной некоторым условиям, определяющим множество E всех возможных случаев, вероятность занять положение, удовлетворяющее другим условиям, определяющим множество E' благоприятных случаев, равна отношению значений интеграла

(113) $$I = \int f(x)\,dx,$$

распространенного соответственно на множества E' и E. Функция $f(x)$ подчинена одному лишь требованию — *никогда не быть отрицательной*.

Это определение удовлетворяет правилу сложения вероятностей, если E' образовано из частичных множеств (подмножеств) E'_1 и E'_2, не имеющих общих элементов, то, разумеется, имеем

$$I' = I'_1 + I'_2,$$

где I', I'_1 и I'_2 — значения I, распространенные соответственно на множества E', E'_1 и E'_2; интегралы I', I'_1 и I'_2 — суть числители дробей с общим знаменателем, которые и равны соответственным вероятностям.

Ясно, что $f(x)$ можно заменить величиной ей пропорциональной, например, такой, какую мы указали в § 19. Дифференциальный элемент

$$d\omega = f(x)\,dx$$

называют элементарной вероятностью, хотя это название не вполне корректно, ибо следовало бы назвать *элементарной вероятностью* отношение $d\omega$ к интегралу I, распространенному на все множество возможных случаев. Но ради удобства мы сохраняем этот термин.

Если положение точки зависит от двух параметров x и y, то элементарная вероятность определяется сходным образом:

$$d\omega = f(x,\ y)\,dx\,dy \tag{114}$$

в зависимости от произвольной положительной функции двух переменных x и y. И это понятие легко обобщается. Элементарная вероятность в n-мерном пространстве с измерениями $x_1,\ x_2,\ \ldots,\ x_n$ имеет общее выражение:

$$d\omega = f(x_1,\ x_2,\ \ldots,\ x_n)\,dx_1\,dx_2\ldots dx_n. \tag{115}$$

54. ЗАМЕНА ПЕРЕМЕННЫХ

Остановимся на случае двух переменных и положим

$$d\omega = f(x,\ y)\,dx\,dy$$

в качестве элементарной вероятности. Если произвести замену переменных

$$x = \varphi(\alpha,\ \beta),$$
$$y = \psi(\alpha,\ \beta),$$

то элементарная вероятность примет новую форму

$$d\omega = F(\alpha,\ \beta)\,d\alpha\,d\beta,$$

в которой

$$F(\alpha,\ \beta) = f(\varphi,\ \psi)\cdot\left|\frac{\partial(\varphi,\ \psi)}{\partial(\alpha,\ \beta)}\right|.$$

Здесь $\frac{\partial(\varphi,\ \psi)}{\partial(\alpha,\ \beta)}$ означает дифференциальный детерминант

$$\begin{vmatrix} \frac{\partial\varphi}{\partial\alpha} & \frac{\partial\varphi}{\partial\beta} \\ \frac{\partial\psi}{\partial\alpha} & \frac{\partial\psi}{\partial\beta} \end{vmatrix},$$

который следует брать по его абсолютной величине.

Новые переменные α и β можно выбрать бесконечным числом способов, но так, чтобы

$$F(\alpha, \beta) = 1.$$

Элементарная вероятность точки будет:

(116) $$d\omega = d\alpha d\beta.$$

Говорят в этом случае, что α и β — суть нормальные переменные для функции $f(x, y)$. Очевидно, что замена $f(x, y)$ другой положительной функцией переменных приводит нас к подстановке вместо α и β других нормальных переменных λ, μ, выбранных целесообразно. Тогда новая элементарная вероятность принимает вид:

$$d\omega = d\lambda\, d\mu.$$

55. ОБЩЕЕ УСЛОВИЕ ПРИ ВЫБОРЕ ЭЛЕМЕНТАРНОЙ ВЕРОЯТНОСТИ

В очень большом числе случаев этот выбор диктуется следующим основным условием.

Результат исчисления должен оставаться неизменным при произвольном перемещении всей рассматриваемой фигуры.

Если, например, дело идет о точке на плоскости, то элементарная вероятность

$$f(x, y)\, dx\, dy,$$

выраженная в прямоугольных координатах, должна оставаться неизменной при преобразованиях:

$$x = x' \cos\alpha - y' \sin\alpha + a,$$

$$y = x' \sin\alpha + y' \cos\alpha + b,$$

а так как $dxdy = dx'dy'$,
то необходимо, чтобы

$$f(x, y) = f(x', y'),$$

что возможно только тогда, когда $f = \text{const}$.

Необходимо поэтому в качестве элементарной вероятности взять элемент площади

$$d\omega = dx\, dy.$$

Но если в рассматриваемом случае имеется в виду положение прямой на плоскости, определяемое уравнением прямой

$$ux + vy + 1 = 0,$$

то преобразование, соответствующее смещению прямой, даст новое уравнение

$$u(x' \cos\alpha - y' \sin\alpha + a) + v(x' \sin\alpha + y' \cos\alpha + b) + 1 = 0,$$

в котором

$$u' x' + v' y' + 1 = 0$$

при

$$u' = \frac{u\cos\alpha + v\sin\alpha}{au + bv + 1},$$

$$v' = \frac{-u\sin\alpha + v\cos\alpha}{au + bv + 1}.$$

Путем несложных выкладок получим

$$du'\,dv' = \frac{du\,dv}{(au + bv + 1)^3}$$

и так как

$$u'^2 + v'^2 = \frac{u^2 + v^2}{(au + bv + 1)^2},$$

то окончательно имеем:

$$\frac{du'\,dv'}{(u'^2 + v'^2)^{3/2}} = \frac{du\,dv}{(u^2 + v^2)^{3/2}}.$$

Таким образом, для прямых линий следует принимать в качестве элементарной вероятности

(117) $$d\omega = \frac{du\,dv}{(u'^2 + v^2)^{3/2}},$$

которая, если представить уравнение прямой в самом общем виде его каноническим уравнением

$$x\cos\theta + y\sin\theta - p = 0,$$

то элементарная вероятность получит вид:

(118) $$dp\,d\theta.$$

Очевидно, что p и Θ представляют собой систему нормальных координат для прямых на плоскости.

Аналогичные соображения позволяют выбрать типичную форму для элементарной вероятности и для случая плоскости или прямой в пространстве.

В заключение отметим случай *точки на сфере*. Результат исчисления можно легко предугадать простой интуицией. Элементарной вероятностью служит *элемент сферической поверхности*, который выражается как

$$d\omega = R\frac{dx\,dy}{z}$$

на сфере с уравнением

$$x^2 + y^2 + z^2 = R^2.$$

Этот элемент может быть представлен в форме

(119) $$d\omega = R^2 \sin\varphi\, d\theta\, d\varphi,$$

если положить

$$x = R\cos\theta\sin\varphi,\quad y = R\sin\theta\sin\varphi,\quad z = R\cos\varphi,$$

вводя географические координаты φ и θ.

56. ИССЛЕДОВАНИЯ НЕКОТОРЫХ ПРИМЕРОВ

Теперь займемся мы исследованием некоторых простых примеров, в которое входят элементарные эвклидовы вероятности, определенные в предыдущем параграфе.

З а д а ч а 9. *На отрезке AB длиной a берут наудачу две точки P и Q. Какова вероятность, что расстояние PQ будет меньше b?*

Положим $AP = x$, $AQ = y$. Вероятность того, что AP будет находиться между x и $x + dx$, равна $\frac{dx}{a}$, а вероятность того, что AQ окажется между y и $y + dy$, равна $\frac{dy}{a}$.

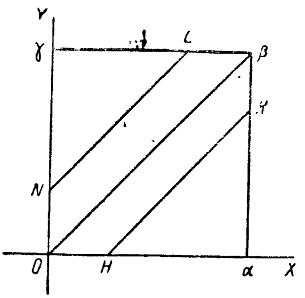

Рис. 12

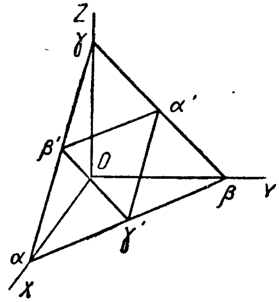

Рис. 13

Элементарной вероятностью для пары точек будет, следовательно, элемент площади

$$dx\,dy$$

на плоскости с двумя осями координат: Ox и Oy. Совокупность всех возможных случаев соответствует точкам квадрата $O\alpha\beta\gamma$ со стороной a. Совокупность же благоприятных случаев соответствует точкам, удовлетворяющим неравенству

$$|x-y|<b$$

и находящимся внутри квадрата. Этой совокупности может быть соотнесена площадь, ограниченная прямыми.

$$x-y=b,$$
$$x-y=-b,$$

взятыми лишь в пределах квадрата (рис. 12). Длина b меньше a, иначе все возможные случаи были бы вместе с тем и благоприятными. Следовательно, мы должны отрезать от квадрата $O\alpha\beta\gamma$ два равнобедренных прямоугольных треугольника $H\alpha k$ и $L\gamma N$; сумма этих двух площадей составляет $(a - b)^2$. Так мы и получим искомую вероятность:

(120) $$p=\frac{1}{a^2}\left[a^2-(a-b)^2\right]=2\,\frac{b}{a}-\frac{b^2}{a^2}\,.$$

З а д а ч а 10. (*Задача о ломком стержне*).

На отрезке AB, равном a, берут наудачу две точки P и Q. Какова вероятность, что из полученных таким путем трех отрезков можно построить треугольник?

Необходимо и достаточно, чтобы ни один из отрезков не превышал их общую полусумму $\frac{AB}{2} = \frac{a}{2}$.

Обозначим через x, y и z длины трех отрезков; это будут три положительных числа, связанных соотношением

$$x+y+z=a.$$

Точка M с пространственными координатами x, y, z в пространстве $Oxyz$ может служить отображением разбиения стержня AB на три куска, длиной x, y и z.

Всем способам возможного разбиения соответствует на плоскости

$$x+y+z=a.$$

Все точки, лежащие внутри треугольника $\alpha\beta\gamma$, построенного так, что его вершинами служат концы отложенных на осях Ox, Oy и Oz отрезков $O\alpha$, $O\beta$ и $O\gamma$, равных каждый a.

Элементарной вероятностью будет служить $dxdy$ пропорциональное площади в плоскости $\alpha\beta\gamma$.

Благоприятные же случаи соответствуют всем точкам треугольника $\alpha'\beta'\gamma'$ (рис. 13), стороны которого соединяют середины сторон треугольника $\alpha\beta\gamma$. Искомая вероятность составляет, следовательно,

$$p = \frac{\text{площадь } \alpha'\beta'\gamma'}{\text{площадь } \alpha\beta\gamma} = \frac{1}{4}.$$

57. ЗАДАЧА НА БРОСАНИЕ ИГЛЫ

Прославленная задача об игле легко поддается экспериментальной проверке, которая часто и предпринималась, так как сама задача известна с давних времен[1]. Задача состоит в следующем: *на горизонтальную плоскость с начерченными на ней параллельными равноотстоящими линиями бросается наугад игла правильной цилиндрической формы. Спрашивается, какова вероятность, что игла пересечет одну из начерченных прямых?*

Обозначим через $2a$ расстояние между двумя параллельными прямыми AB и CD и через $2l$ длину иглы PQ. Можно представить себе, что середина иглы M находится на некотором перпендикуляре к параллелям AB и CD. Обозначим через X расстояние EM, предполагая его меньшим, чем FM, и считая также, что $l < a$, т. е. что игла не может пересечь больше одной параллели. Мы увидим на самом деле, что если бы игла могла пересечь несколько параллелей, то, для того чтобы получить простое решение, необходимо заменить отыскание вероят-

[1] Ее *правильное* решение первым дал Бюффон (Buffon).

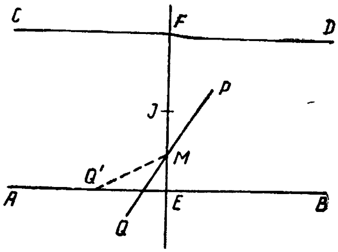

Рис. 14

ности вычислением математического ожидания лица, которому обещано столько франков, сколько будет всего пересечений иглой параллелей.

Для того чтобы игла могла пересечь AB (рис. 14) необходимо, чтобы X был меньше l. Если точкой I расстояние EF делится пополам, то ясно, что вероятность того, что для точки M, находящейся на EI, расстояние EM лежало бы между x и $x+dx$, равна $\frac{dx}{a}$. Если положение M зафиксировано, то нужно еще, чтобы острый угол EMQ был меньше угла EMQ', при котором конец иглы Q' еще попадает на AB. Имеем:

$$\cos EMQ' = \frac{EM}{MQ'},$$

т. е.

$$EMQ' = \arccos\frac{x}{l}.$$

Вероятность того, чтобы острый угол EMQ (имеющий величину между 0 и $\frac{\pi}{2}$) находился бы внутри угла EMQ', равна:

$$\frac{2}{\pi}\arccos\frac{x}{l}, \tag{121}$$

и вероятность того, что игла пересечет AB, будет поэтому

$$\frac{2}{\pi}\int_0^l \arccos\frac{x}{l}\,\frac{dx}{a}.$$

Если в интеграле положить $x = ly$, то интеграл преобразуется

$$\frac{2l}{a\pi}\int_0^1 \arccos y\,dy.$$

Полагая же $y = \cos t$, получим

$$\int_0^1 \arccos y\,dy = -\int_{\pi/2}^0 t\sin t\,dt = \left| t\cos t - \sin t \right|_{\pi/2}^0 = 1,$$

и, следовательно, искомая вероятность окажется равной

$$p = \frac{2l}{\pi a} \tag{122}$$

Если предположить, как частный случай, что $2l = a$, т. е. взять длину иглы, равной половине расстояния между параллелями, то получим

$$p = \frac{1}{\pi}.$$

Эти результаты замечательны тем, что их можно экспериментально проверить. Если вероятность, указанная в (121) может быть экспериментально получена, то в силу эмпирического закона случайности можно получить приближенное значение π, деля отношение $\frac{2l}{a}$ на частость пересечения иглой параллелей, установленной опытным путем.

Серия в 5000 испытаний была произведена в 1850 г. в Цюрихе Вольфом (Wolff), принявшим $l = 36$ *мм*, $a = 45$ *мм*, так что вероятность пересечения равнялась $\frac{8}{5\pi}$. Полученное приближенное значение для π оказалось равным 3,1596. Этот результат удовлетворителен в такой же мере, каким было бы определение дроби $\frac{8}{5\pi} = 0{,}50929568$ в относительном числе белых шаров, извлеченных при 5000 испытаниях, если бы урна заключала в себе 100 000 000 шаров, из коих белых было бы 50 929 568.

Для $n = 5000$ извлечений при $p = \frac{8}{5\pi}$, $q = \frac{5\pi - 8}{5\pi}$ типичное отклонение равнялось бы:

$$\sigma = \sqrt{npq} = \frac{200}{5\pi}\sqrt{5\pi - 8},$$

что приблизительно дает 35; откуда для соответственной относительной ошибки получаем величину $\frac{\sigma}{n} = 0{,}007$.

Для отклонения такого размера, который по § 35 соответствует вероятности быть превзойденным (по абсолютной величине) в 35%, соответствовало бы в оценке вероятности, равной $\frac{8}{5\pi}$ и, следовательно, ошибка порядка

$$\frac{\sigma}{np} = 0{,}014 \text{ (приблизительно).}$$

Между тем результат опыта Вольфа искажен всего лишь относительной ошибкой в 0,006. Его, таким образом, следует признать вполне удовлетворительным, пригодным к тому, чтобы подтвердить правильность сделанных предположений.

Очень остроумное решение задачи с иглой дал Барбье (E. Barbier) около 1860 г.

Пусть какая-нибудь ломаная линия, замкнутая или незамкнутая, брошена наудачу на лист бумаги, на котором нанесены параллельные прямые. Каково математическое ожидание игрока, получающего по франку за каждую точку пересечения?

Ответ получается сложением математических ожиданий для каждой из сторон, а эти последние пропорциональны длине каждой из них. В конечном счете математическое ожидание равно kL, где L — длина всей ломаной линии, а k — множитель, подлежащий определению. Этот результат распространяется, естественно, и на предельный случай, когда брошенная линия становится кривой.

Чтобы определить k, рассмотрим тот случай, когда брошенная линия представляет собой окружность радиуса a; в этом случае мы имеем обязательно две точки пересечения и

$$2\pi ak=2,$$

откуда

$$k=\frac{1}{\pi a}.$$

Установив это, мы найдем, что математическое ожидание для иглы длиною $2l$ составляет

$$2kl=\frac{2l}{\pi a},$$

а в том случае, когда длина иглы $2l$ меньше, чем расстояние $2a$ между параллельными прямыми, это математическое ожидание принимает величину, одинаковую с вероятностью пересечения.

58. ЗАДАЧИ О ПРЯМЫХ НА ПЛОСКОСТИ

Мы видели выше, как относительно прямых на плоскости, определяемых уравнением

$$x\cos\theta+y\sin\theta-p=0,$$

мы приходим к необходимости принимать в качестве элементарной вероятности выражение $dpd\theta$.

Каждому множеству прямых на плоскости соответствует, таким образом, некоторое положительное число, которое представляет собой меру данного множества и получается вычислением двойного интеграла

$$\iint dp\,d\theta,$$

для всех прямых множества. Мы не развертываем здесь аналитического исследования этих множеств, которое относится к области интегрального исчисления. Одним из наиболее замечательных положений такого исследования является следующее.

«*Если* (C) *есть замкнутый выпуклый контур длиной* L *и если* (C') *есть также замкнутый выпуклый контур длиной* L', *лежащий внутри* (C), *то вероятность того, что секущая контура* (C) *пересечет также и контур* (C'), *равна*

$$p=\frac{L'}{L}.$$

Применим этот результат к тому случаю, когда (C) есть контур окружности радиуса a, а (C') — сегмент, длиною $2l$, тогда мы имеем:

$$p = \frac{4l}{2\pi a} = \frac{2l}{\pi a},$$

так как надо брать длину сегмента вдвойне, если его рассматривать как предельную форму выпуклого замкнутого контура. Небольшого размышления достаточно, чтобы убедиться, что поставленная так задача адекватна задаче об игле, в предположении, что $l < a$.

Задача Жозефа Бертрана. *В некотором круге берем наудачу одну из возможных в ней хорд. Какова вероятность, что длина взятой хорды окажется больше стороны вписанного в круг равностороннего треугольника?*

Мы исследуем эту задачу ради той важной роли, которую в ней играет разумный выбор элементарной вероятности. Бертран в своей книге «Исчисление вероятностей» указывает на неудовлетворительную постановку самой задачи и в доказательство этого дает три решения, которые приводят к трем различным результатам.

1) Если известно положение одного из концов хорды, то такое знание не меняет величины вероятности. Направление хорды берется наугад и хорды длиной больше, чем сторона вписанного правильного треугольника, должны находиться внутри угла в 60°, в то время как множество всех хорд охватывает углы в пределах 180°. Вероятность, следовательно, равна $\frac{1}{3}$.

2) Можно заранее задать направление хорды, что не изменит величины вероятности. Расстояние от центра круга должно быть менее половины радиуса, чтобы хорда была больше стороны правильного вписанного треугольника. Вероятность получается равной $\frac{1}{2}$.

3) Чтобы наугад задать хорду, достаточно наудачу выбрать положение ее срединной точки. Середина хорды должна лежать внутри круга концентричного с заданным и имеющего радиус, равный половине радиуса данной окружности. Вероятность определяется как отношение площадей обоих кругов и равна $\frac{1}{4}$.

Ввиду такого разноречия результатов не следует ли согласиться со скепсисом Бертрана? В этом случае надо пойти в анализе дальше Ж. Бертрана, ибо его скептицизм — одна лишь видимость.

Три решения получаются от принятия для прямой с уравнением

$$x \cos\theta + y \sin\theta - p = 0$$

трех различных *элементарных вероятностей*:
для третьего случая — это $p\,dp\,d\theta$,
во втором случае — $dp\,d\theta$.

Чтобы сделать выбор между тремя разными решениями или даже бесконечным их числом, которое получилось бы от введения произвольной функции, необходимо выяснить условия реальной проверки решения экспериментом, таким, например, как бросание диска радиуса $\frac{a}{2}$ на параллельные прямые в задаче с иглой. Для того чтобы диск вырезал на покрытой им параллели отрезок, больший, чем сторона равностороннего треугольника, вписанного в диск, необходимо и достаточно, чтобы этот отрезок (хорда диска) пересекал круг, имеющий радиус в $\frac{a}{2}$ и концентричный диску. Согласно положения доказанного Барбье, вероятность этого равна отношению длин обеих окружностей, т. е. $\frac{1}{2}$.

К тому же числу приводит второе решение Бертрана. Должны ли мы заключать отсюда, что первое и третье решения ошибочны? Нет! В действительности все три решения правильны, но это решение *трех различных задач.*

Отметим, что условие экспериментальной проверки возвращают нас в § 55, который привел нас к элементарной вероятности. Отношение значений интеграла

$$\iint dp\,d\theta,$$

распространенного на прямые, пересекающие два концентрических круга, может быть найдено по соображениям симметрии безотносительно к $dpd\theta$; искомое отношение есть отношение радиусов обоих кругов.

59. ОБОБЩЕНИЕ ЗАДАЧИ Ж. БЕРТРАНА

Рассмотрим следующую задачу, представляющую обобщение задачи Бертрана.

З а д а ч а 11. *Какова вероятность, чтобы длина какой-либо взятой наудачу хорды в окружности радиуса R, оказалась между a и b?*

Если задаться диаметром перпендикулярным хорде, то мы найдем, что расстояния от центра для хорд, удовлетворяющих заданию, будут

$$\sqrt{R^2-\frac{a^2}{4}} \text{ и } \sqrt{R^2-\frac{b^2}{4}}.$$

Искомая вероятность в предположении, что $a < b$, будет, следовательно, равна:

$$\frac{1}{R}\left[\sqrt{R^2-\frac{a^2}{4}}-\sqrt{R^2-\frac{b^2}{4}}\right].$$

Вероятность того, что длина хорды будет заключаться между a и $a+da$, окажется равной:

$$-\frac{1}{R}d\left(\sqrt{R^2-\frac{a^2}{4}}\right)=\frac{ada}{2R\sqrt{4R^2-a^2}}. \quad (123)$$

Этот результат можно применить к задаче об игле. Пусть AB и CD— две соседние параллели, EF — их общий перпендикуляр, I — середина EF; мы ограничимся случаем, когда $EF = l$ (рис. 15).

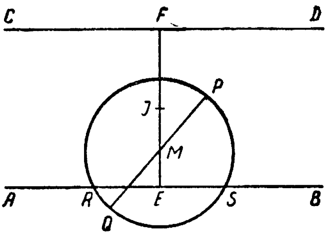

Рис. 15

Мы можем предположить, что центр иглы находится на сегменте EI и искать вероятность того, что в этих условиях игла пересечет AB.

Опишем окружность с центром в точке M и с радиусом MQ, равным l. Пусть точки R и S будут точками, в которых эта окружность пересекает AB, тогда искомая вероятность равна отношению дуги RS к полуокружности; если принять длину хорды RS, равной $2c$, то это отношение будет равняться

$$\frac{2}{\pi}\arcsin\frac{c}{l}.$$

Если, следовательно, обозначить через $\varphi(c)dc$ вероятность того, что длина хорды в окружности радиусом l будет заключаться между $2c$ и $2c + 2dc$, то общая искомая вероятность окажется равной

$$\int_0^l \frac{2}{\pi}\arcsin\frac{c}{l}\,\varphi(c)\,dc.$$

Исходя же из предшествующего и полагая в формуле (123) $a = 2c$, $R = l$, получим

$$\varphi(c)\,dc = \frac{4cdc}{2l\sqrt{4l^2-4c^2}} = \frac{cdc}{l\sqrt{R^2-c^2}},$$

а полагая $c = l\sin\theta$, приведем интеграл к виду

$$\int_0^{\pi/2} \frac{2}{\pi}\,\theta\sin\theta d\theta = \frac{2}{\pi},$$

что совпадает с уже ранее найденным результатом. Если принять первое или третье из решений, указанных Бертраном, для задачи, исследованной в предыдущем параграфе, то из них последует другое значение для $\varphi(c)$ и, следовательно, иное решение задачи об игле. Так как для последней легко осуществить экспериментальную проверку, то из этого вытекает очевидность произвольного характера решений; они соответствуют совершенно особым гипотезам, которые редко поддаются реализации в порядке экспериментальной проверки решений.

60. ЗАДАЧИ, ОТНОСЯЩИЕСЯ К ТОЧКАМ, ВЗЯТЫМ НА ПОВЕРХНОСТИ СФЕРЫ

Мы видели выше, как соображения о неизменности вероятности по отношению к произвольному перемещению приводят к выбору в качестве элементарной вероятности для задач, относящихся к положению точки, определяемой уравнениями:

$$x = R\cos\varphi\sin\theta,$$

$$y = R\sin\varphi\sin\theta,$$

$$z = R\cos\theta$$

элемента площади

$$d\omega = R^2\sin\theta d\varphi\, d\theta.$$

Пусть даны наудачу две точки на сфере, найдем вероятность того, что угол, образованный радиусами, проведенными из центра сферы к этим точкам, меньше некоторого данного угла α.

Симметрия сферы относительно всех осей, проходящих через ее центр, приводит нас к очевидному допущению, что *вероятность остается неизменной, если задать заранее положение одной из точек*. Тогда другая точка должна находиться в зоне, полюсом которой служит первая точка, а высотой — $R\ (1 - \cos\alpha)$. Отношение площади зоны к площади всей сферы равно:

$$\frac{1-\cos\alpha}{2} = \sin^2\frac{\alpha}{2}.$$

Заметим относительно элементарной вероятности $\sin\theta d\varphi d\theta$, что все долготы имеют одинаковую вероятность, но что относительно широт дело обстоит совсем иначе. Здесь мы встречаемся с фактом, который может показаться удивительным, если бы вопрос не был исследован достаточно внимательно. Если рассматривать все значения долгот и все значения широт как равновероятные, то вероятность того, чтобы радиус сферы, направленный в точку сферы M, составлял бы с осью Oz угол меньший, чем α, равнялась бы $\frac{\alpha}{2}$. Разница между этим результатом и правильной величиной $\sin^2\frac{\alpha}{2}$ оказывается очень значительной в случае малых углов. Когда $\alpha = 1°$, величина $\frac{\alpha}{\pi}$ приблизительно в 200 раз больше, чем $\sin^2\frac{\alpha}{2}$.

Задача 12. *На сфере размещены наудачу n точек. Каково математическое ожидание лица, получающего по франку за каждую пару точек, угловое расстояние которых меньше α?*

Для каждой пары искомая вероятность составляет (если α мало) величину, близкую к $\frac{\alpha}{4}$ следовательно, математическое ожидание будет:

$$\frac{n(n-1)}{2}\,\frac{\alpha^2}{4}=\frac{n(n-1)\alpha^2}{8}.$$

В дальнейшем мы будем иметь случай применить этот результат к задаче на вероятность причин, относящихся к распределению звезд на небесной сфере.

61. ЗАДАЧА, ОТНОСЯЩАЯСЯ К ДВУМ ТОЧКАМ В ПРОСТРАНСТВЕ

Раз форма элементарной вероятности установлена, то решение данной задачи сводится к вычислению простых или кратных интегралов. Такие вычисления часто бывают трудными даже в тех случаях, когда по внешнему виду интеграл представляется элементарным.

Мы покажем на одном примере, как можно иногда избегнуть сложных вычислений кратных интегралов.

Пусть (D) — область обычного пространства, и обозначим объем области (D) через V. Требуется вычислить вероятность того, что n точек, взятых наудачу в (D), обладает свойством П, которое зависит только от размеров и от относительного положения элементов фигуры F, образованной этими точками, и ни в коем случае не зависит от (D).

Если dV_1 — элемент объема, взятый внутри (D), то вероятность для точки M_1 попасть в этот элемент равна $\frac{dU_1}{V}$. Точки M_1, M_2, ..., M_n, определяющие фигуру F, взяты в объеме (D) независимо одна от другой; по правилу перемножения вероятностей мы можем принять для фигуры F элементарную вероятность:

$$dV_1 \cdot dV_2 \ldots dV_n,$$

т. е. произведение объемных элементов, составляющих окрестности точек M_1, M_2, ... , M_n. В таких условиях вероятность равна:

$$P=\frac{U}{V^n}, \tag{124}$$

где U — некоторый числитель, соответствующий множеству благоприятных случаев.

Пусть (D') — другая область, содержащая в себе или на своих границах все точки области (D); обозначим через ΔV объем области (D''), образованной точками из (D'), не принадлежащими (D).

Если вместо (D) рассматривать (D'), то вероятность P изменится и станет равной:

$$P+\Delta P=\frac{U+\Delta U}{(V+\Delta V)^n}. \tag{125}$$

Обозначим через P_k вероятность выполнения условий задачи фигурой F, образованной $n-k$ точками, взятыми наугад в (D), и k точками из (D''); U_k выражает для такой фигуры меру множества благоприятных статочностей, и мы получаем

$$P_k=\frac{U_k}{C_n^m V^{n-k}(\Delta V)^k}. \tag{126}$$

Если в выражении, взятом из (125) для суммы

$$U+\Delta U=U+U_1+U_2+\ldots+U_k+\ldots+U_n,$$

мы заменим каждое U_k его значением, взятым из (126), то мы получим соотношение

$$(P+\Delta P)(V+\Delta V)^n=PV^n+C_n^1 P_1 V^{n-1}\Delta V+\ldots+P_n(\Delta V)^n$$

и, пользуясь разложением $(V+\Delta V)^n$,

$$(127)\quad (V+\Delta V)^n\Delta P=C_n^1(P_1-P)V^{n-1}\Delta V+\ldots+(P_n-P)(\Delta V)^n.$$

Теперь представим себе, что (D'') есть бесконечно малое приращение области (D). Соотношение (127) при этом дает главную часть δP приращения ΔP и если ограничиться бесконечно малыми первого порядка, то $\Delta V=\delta V$.

$$(128)\qquad \delta P=n(P_1-P)\frac{\delta V}{V}.$$

Этой формулой мы обязаны английскому математику Крофтону (Crofton). Она позволяет вычисление P свести к вычислению P_1, которое представляет собой вероятность осуществить П с помощью $(n-1)$ точек, взятых наугад в области (D), и одной точкой, взятой в области (D''); вычислять P_1, как правило, легче, чем непосредственно вычислять P. Этот метод мы применим к следующему примеру.

З а д а ч а 13. *Найти вероятность того, чтобы расстояние между двумя точками, взятыми наугад во внутреннем пространстве сферы радиуса R, было бы меньше данной длины a.*

Назовем эти две точки буквами A и B. Вероятность того, что AB заключено между x и $x+dx$, имеет форму $p(x)dx$, и искомую вероятность мы найдем исходя из $p(x)$ с помощью формулы

$$(129)\qquad P=\int_0^a p(x)\,dx.$$

Обозначим через δp вариацию $p(x)$, соответствующую бесконечно малому приращению R, равному δR, мы имеем здесь $n=2$, а также то, что $V=\frac{4}{3}\pi R^3$, $\frac{\delta V}{V}=3\frac{\delta R}{R}$. Поэтому формулу Крофтона мы напишем в таком виде:

$$(130)\qquad \delta p=6(p_1-p)\frac{\delta R}{R}.$$

Вероятность $p_1(x)dx$ есть вероятность события

$$x<AB<x+dx$$

в том случае, когда A взято наудачу в (D''), т. е. в объеме между концентрическими сферами с радиусами R и $R+\delta R$, а B взято наудачу внутри шара (S), образующего объем (D).

Учитывая симметрию и тот факт, что δR бесконечно малая величина, мы можем считать A взятым на самой сфере (S); точка же B должна находиться в бесконечно малом объеме, входящем во внутреннюю область сферы (S), ограниченном двумя шаровыми сегментами с радиусами x и $x + dx$.

Этот бесконечно малый объем имеет в качестве своей главной части

$$d\omega = \sigma dx,$$

где σ есть поверхность шарового сегмента радиуса x с центром, находящимся на (S) и взятым в той своей части, которая расположена внутри (S) (рис. 16).

Таким образом мы находим

$$\sigma = 2\pi x \cdot ED,$$

принимая же во внимание, что

$$\overline{AH}^2 = x^2 = 2R\, AE = 2R\,(x - ED)$$

получим для σ

$$\sigma = \frac{\pi x^2}{R}(2R - x).$$

В этих условиях мы имеем

$$p_1(x)\, dx = \frac{d\omega}{V} = \frac{3}{4}\frac{x^2(2R-x)}{R^3}\, dx,$$

откуда, внося в (130) найденное значение для p_1, найдем

(131) $$\frac{\delta p}{\delta R} + 6\frac{p}{R} = \frac{9}{2}\frac{x^2(2R-x)}{R^5}.$$

Мы имеем здесь дифференциальное уравнение, в котором R играет роль независимой переменной, а p — неизвестная функция. Это уравнение — линейное. Уравнение без второго члена имеет свой общий интеграл

$$p = \frac{C}{R^6},$$

а применяя метод вариации, постоянной C, получим и общий интеграл

$$p = \frac{3x^3}{R^3} - \frac{9}{4}\frac{x^2}{R^4} + \frac{\lambda}{R^6},$$

в котором λ означает произвольную постоянную (не зависящую от R, но могущую оказаться зависящей от x).

Для дифференциального уравнения (131) мы должны сохранить решение, которое обращается в нуль при $x = 2R$, т. е. когда p рассматривается как функция от R, исчезающая для $R = \frac{x}{2}$; таким путем мы получим λ равной $\frac{3}{16}x^5$, откуда

$$p=\frac{3x^2}{R^3}-\frac{9x^3}{4R^4}+\frac{3x^5}{16R^6}$$

и, наконец, согласно (129) имеем:

$$P=\frac{a^3}{R^3}-\frac{9a^4}{16R^4}+\frac{a^6}{32R^6}. \tag{132}$$

Заметим, что P становится равным 1 при $a=2R$. Ясно, что для $a>2R$ формула (132) перестает показывать вероятность $Pr(AB>a)$, которая имеет значение единицы.

62. ВЫЧИСЛЕНИЕ НЕКОТОРЫХ ВЕРОЯТНЫХ ЗНАЧЕНИЙ

Вспомним, что *вероятное значение* означает то же самое, что *среднее значение* или *математическое ожидание* (см. § 10 и 20)[1].

Задача 14. *Найти математическое ожидание длины хорды, взятой наудачу среди всех хорд окружности.*

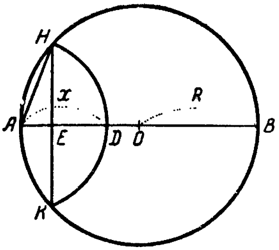

Рис. 16

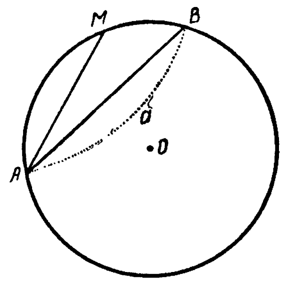

Рис. 17

Мы показали в § 59, что вероятность для хорды, взятой в окружности с радиусом R, иметь длину, лежащую между α и $\alpha+d\alpha$, равна

$$P=\frac{ada}{2R\sqrt{4R^2-a^2}}.$$

Вероятная величина длины хорды в этих условиях равна:

$$l=\int_0^{2R}\frac{a^2\,da}{2R\sqrt{4R^2-a^2}}.$$

Положим $a=2R\cos\varphi$ и тогда получим

[1] Терминология французских авторов не должна смущать читателя: в русской литературе принято *средней* величиной называть сумму *произведений* возможных и наблюденных *значений* переменной, умноженных *на их частости*, полученные из наблюдений или экспериментально, а *математическим ожиданием* (или теоретической средней) — сумму произведений всех возможных *значений*, умноженных *на вероятности их*, т. е. на теоретические величины, лежащие в основании частостей. — *Прим. перев.*

$$l=\int_0^{\pi/2} 2R\cos^2\varphi d\varphi=\frac{\pi R}{2}.$$

Вероятная длина хорды в окружности равна, следовательно, четверти от длины всей окружности.

Необходимо подчеркнуть, что вероятное значение какой-либо величины, *как и вероятность вообще*, зависит от того, какие статочности мы условимся считать равновозможными, иначе говоря, считать *элементарными вероятностями.*

Найдем вероятное значение *расстояния между двумя точками, взятыми наугад на окружности.*

Мы можем, учитывая симметрию, рассматривать одну из точек, скажем точку A, как данную (рис. 17); тогда, вероятность того, что AM окажется заключенной между 0 и a, будет равна отношению дуги AB, соответствующей хорде $AB = a$, к длине полуокружности; поэтому

$$p=\frac{2}{\pi}\arcsin\frac{a}{2R}.$$

Вероятность для $\overline{AM}$ заключаться между a и $a+da$ равна

$$\frac{2}{\pi}d\left(\arcsin\frac{a}{2R}\right),$$

а математическое ожидание AM равно

$$\int_0^{2R}\frac{2}{\pi}ad\left(\arcsin\frac{a}{2R}\right).$$

Вставляя вместо a его выражение через угол θ, равное $2R\sin\theta$, мы для математического ожидания найдем:

$$\frac{2}{\pi}\int_0^{\pi/2}2R\sin\theta d\theta=\frac{4R}{\pi}.$$

Результат, отличающийся от предыдущего. Мы вновь сталкиваемся здесь с таким различием, которое мы выше встречали при рассмотрении задачи Ж. Бертрана.

Задача 15. *Берем наудачу две точки M и M' внутри квадрата со стороной a. Чему равно математическое ожидание квадрата расстояния MM'?*

Если координаты точек M и M' обозначить через x и y и через x' и y' соответственно, то решение задачи осуществится вычислением четырехкратного интеграла:

$$\frac{1}{a^4}\int_0^a\int_0^a\int_0^a\int_0^a[(x-x')^2+(y-y')^2]\,dx\,dy\,dx'\,dy'.$$

Последовательно интегрируя, получим

$$\int_0^a [(x-x')^2+(y-y')^2]\,dx=\frac{(a-x')^3+x'^3}{3}+a(y-y')^2,$$

$$\int_0^a \left[\frac{(a-x')^2+x'^2}{3}+a(y-y')^2\right]dx'=\frac{a^4}{6}+a^2(y-y')^2,$$

$$\int_0^a\int_0^a \left[\frac{a^4}{6}\,a^2(y-y')^2\right]dy\,dy'=\frac{a^6}{6}+\frac{a^6}{6}=\frac{a^6}{3}.$$

Следовательно, искомое математическое ожидание равно $\frac{a^6}{3}$.

63. СНОВА О ПРОИЗВОЛЬНЫХ ФУНКЦИЯХ ПУАНКАРЕ

То обстоятельство, что существуют задачи, *решения которых остаются неизменными*, каков бы ни был тип принятых элементарных вероятностей, порождает соблазн обратиться к позиции Пуанкаре, вводя произвольную функцию в определение элементарной вероятности. Это, само собой разумеется, делается с оговоркой, что выбранная положительная функция реализует условия очень общего характера, такие, например, как условия для переменной быть дискретной или иметь конечную производную и т. п. Теория Пуанкаре придает задачам такого рода большую значимость, увеличивая в некотором смысле определенность результата. И так как сам А. Пуанкаре указал некоторые примеры, то мы произведем разбор одного из них.

Задача 16. *Какова вероятность того, что какое-нибудь число, взятое в пределах между 0 и 1, окажется рациональным?*

Если мы примем в качестве элементарной вероятности функцию $f(x)dx$, которая предполагается положительной или равной нулю, и имеющей конечное значение для значений x, взятых между 0 и 1, то общее количество всех случаев измеряется интегралом:

$$Q=\int_0^1 f(x)\,dx,$$

который не может быть отрицательным и который, как мы полагаем, не должен равняться нулю. Легко показать, что числитель меньше любого данного числа.

Доказательство основывается на возможности пересчитать все рациональные числа, лежащие между 0 и 1. Принцип, делающий такой пересчет возможным, состоит, например, в том, что берутся все дроби, у которых сумма числителя и знаменателя равна числу K (таких чисел, взятых между 0 и 1, для каждого значения N имеется всегда конечное количество), и затем числу N придаются последовательно значения $N=1, 2, 3, \ldots$, не заботясь о том, что некоторые величины могут при этом повторяться.

Пусть μ есть верхняя граница $f(x)$; на отрезке $OA = 1$ окружим точку, абсцисса которой занимает n-е место в указанном выше перечислении, сегментом протяженностью в $\frac{l}{2^n}$, где l — некоторая длина, которую мы сейчас уточним.

Совокупность благоприятных статочностей, очевидно, меньше интеграла

$$\int f(x)\,dx,$$

распространенного на все помеченные сегменты, а сам этот интеграл меньше, чем

$$\mu l\left[\frac{1}{2}+\frac{1}{2^2}+\ldots+\frac{1}{2^n}+\ldots\right], \text{ или } \mu l.$$

Если взять $l=\frac{\varepsilon}{\mu}$, то мы увидим, что числитель вероятности меньше, чем ε, какова бы ни была величина ε, т. е. вероятность равна нулю.

УПРАЖНЕНИЯ К ГЛАВЕ ТРЕТЬЕЙ

1. Задача о разбитом алмазе

Неграненый алмаз стоимостью в a франков разбит на два куска. Каково математическое ожидание ценности разбитого алмаза, если принять, что ценность алмаза пропорциональна его весу?

Допустим, что алмаз весом P разбивается на два куска так, что вес одного из кусков есть x, а другого куска $P - x$, причем x заключается в пределах α и $\alpha + d\alpha$ — вероятность такого разбиения равна $\frac{d\alpha}{P}$. Тогда математическое ожидание стоимости определится в $\frac{2}{3}a$.

2. Та же задача в случае, если алмаз расколот на три куска. При тех же условиях найдем, что математическое ожидание стоимости всех трех кусков равно $\frac{a}{2}$.

3. На поверхности сферы берутся наугад три точки: A, B и C. Каково математическое ожидание площади сферического треугольника ABC?

Ответ: $\frac{1}{8}$ общей площади сферы.

4. Решить задачу 3, заменив заданную сферу кругом того же радиуса, данным на плоскости.

Метод решения тот же, лишь несколько более простой. Интегрирование дает

$$p=\frac{\pi a^2+2\alpha\,(R^2-a^2)-(2R^2+a^2)\sin\alpha\cos\alpha}{\pi R^2},$$

где $a=2R\sin\alpha$.

5. Положение твердого тела, отнесенного к трем твердым осям, определяется координатами одной его точки a, b, c и углами Эйлера θ, φ, ψ, какими пользуются в механике. Какова элементарная вероятность положения, данного этому телу?

Ответ:

$$dp = \sin\theta \, d\varphi \, d\psi \, d\theta \, da \, db \, dc.$$

6. Две точки B и C взяты наугад внутри определенного круга. Каково математическое ожидание площади треугольника ABC, если A — точка, взятая на периферии круга?

Ответ: $\frac{35}{36}\frac{S}{\pi^2}$, где S — площадь круга.

4. ВЕРОЯТНОСТЬ ПРИЧИН ПРОБЛЕМА ОЦЕНКИ

I. ФОРМУЛА БАЙЕСА (BAYES) И ЕЁ ПРИМЕНЕНИЕ

64. ОБЩИЕ ЗАМЕЧАНИЯ

В тех проблемах вероятности, которые рассматривались до сих пор, вопрос ставился следующим образом: если известно, что произошло событие C, то какова вероятность, что осуществится событие E? Событие C *предшествует* событию E; его можно рассматривать как *причину*.

Встанем теперь на другую точку зрения: событие E уже произошло; мы знаем, что оно может быть следствием одной из многих причин: $C_1, C_2, \ldots, C_n$, взаимно исключающих друг друга. Если известно, что событие E произошло, то ставится вопрос о вероятности, что оно произошло именно от причины C_n.

Теперь уместно здесь оговориться о том, что причины $C_1, C_2, \ldots, C_n$, имеют заранее нам известную вероятность способствовать осуществлению события E. Предполагается, что эти *вероятности* даны á priori, но вопрос стоит о вероятности a posteriori, той, *когда событие E уже осуществилось.*

Такова общая постановка проблемы, называемой *проблемой вероятности причин.*

65. ФОРМУЛА БАЙЕСА

Пусть $\omega_1, \omega_2, \ldots, \omega_n$ означают *вероятности причин* $C_1, C_2, \ldots, C_n$ и пусть также $p_1, p_2, \ldots, p_n$ — вероятности осуществления E в зависимости от одной из причин $C_1, C_2, \ldots, C_n$. Мы можем обозначить через

$$P_k = Pr\,(C_n/E)$$

искомую вероятность того, что причиной E была именно причина C_k при условии, что E уже осуществилось.

Рассмотрим событие (C_k/E), т. е. осуществление E причиной C_k; ее вероятность можно рассчитать двумя способами.

Прежде всего эта вероятность, согласно правилу перемножения вероятностей, равна:

$$Pr[C_k/E] = Pr(C_k)\, Pr[E/C_k] = \omega_k p_k.$$

Но вместе с тем и по тому же правилу перемножения вероятностей мы можем написать:

$$Pr[E/C_k] = Pr(E)\, Pr[C_k/E] = P_k\, Pr(E). \tag{133}$$

Остается лишь вычислить $Pr(E)$. Так как причины C_i несовместимы, то по правилу сложения вероятностей мы можем написать:

$$Pr(E) = Pr[C_1 E] + Pr[C_2 E] + \ldots + \\ + Pr[C_n E] = \omega_1 p_1 + \omega_2 p_2 + \ldots + \omega_n p_n.$$

Внося это выражение в формулу (133), мы и получим *формулу Байеса*:

$$P_k = \frac{\omega_k p_k}{\omega_1 p_1 + \omega_2 p_2 + \ldots + \omega_n p_n}, \tag{134}$$

которая была опубликована в 1764 г. в посмертно изданной статье упомянутого автора. Пользоваться этой формулой несложно, если только установлены априорные вероятности $\omega_1, \omega_2, \ldots, \omega_n$, а именно это и является слабым местом данной теории.

66. ПРИМЕРЫ

Задача 17. *Две урны содержат шары: одна девять белых и один черный, а другая пять черных и двадцать белых. Извлекают один шар из какой-либо урны, взятой наудачу, и этот шар оказывается белым. Какова вероятность, что извлечение было сделано из первой урны?*

Мы предполагаем, что обе урны по своему внешнему виду тождественны и вследствие этого допускаем, что вероятности взять для испытания ту или другую из урн одинаковы:

$$\omega_1 = \omega_2 = \frac{1}{2}.$$

С другой стороны, учитывая состав урн, мы знаем, что

$$p_1 = \frac{9}{10}, \quad p_2 = \frac{5}{25}.$$

Формула Байеса дает при таких исходных данных

$$P_1 = \frac{\omega_1 p_1}{\omega_1 p_1 + \omega_2 p_2} = \frac{9}{11}.$$

З а д а ч а 18. *В двух классах, в каждом из которых учится по 20 учеников, имеется: в первом 10 хороших учеников, 5 посредственных и 5 плохих; во втором же классе 5 хороших учеников, 5 посредственных и 10 плохих. Инспектор спрашивает по одному ученику, взятому наугад из каждого класса и находит, что ученик из класса A ответил лучше, чем ученик из класса B. Какова вероятность того, что класс A был первым из перечисленных выше классов?*

В испытании, состоявшем из опроса одного ученика, взятого в одном из классов, а затем другого ученика из другого класса, событие E состоит в том, что первый из опрошенных учеников оказался лучше второго.

Назовем группы, образованные в каждом классе из учеников хороших, посредственных и плохих, — «наборами 1, 2 и 3».

Событие (E/C_1) может осуществиться *либо*, если взять из первого класса C_1 ученика из набора № 1, а из класса C_2 ученика из наборов № 2 и 3 безразлично, *или же*, беря ученика в классе C_1 из набора № 2, а ученика в классе C_2 из набора № 3, мы имеем, следовательно, благоприятных статочностей $10 \times 15 + 5 \times 10 = 200$ на $20 \times 20 = 400$ всех возможных статочностей и таким образом

$$p_1 = \frac{1}{2}.$$

Событие (E/C_2) может осуществиться *либо*, если в классе C_1 взят ученик из набора № 1, а в классе C_1 — из наборов № 2 и 3; *либо же* беря из класса C_2 ученика из набора № 2, а в классе C_1 — ученика из набора № 3; откуда имеем

$$5 \times 10 + 5 \times 5 = 75$$

благоприятных статочностей на 400 возможных. Таким образом,

$$p_2 = \frac{3}{16}.$$

Так как вероятности á priori равны $\omega_1 = \omega_2 = \frac{1}{2}$, то по формуле Байеса мы получим:

$$P_1 = \frac{\omega_1 p_1}{\omega_1 p_1 + \omega_2 p_2} = \frac{8}{11} = 0{,}7272\ldots$$

Представим себе теперь, что инспектор вызывает для опроса по два ученика, взятых наугад, как из класса A, так и из класса B. При этом оказывается, что ученики, вызванные из класса A оба ответили лучше, чем оба ученика из класса B.

Событие (E/C_1) может осуществиться следующими способами, взаимно исключающими друг друга:

а) оба ученика из класса C_1 были взяты из набора № 1, а оба ученика из класса C_2 — из наборов № 2 и 3 безразлично;

б) один ученик класса C_1 был вызван из набора № 1, а другой — из набора № 2; из класса же C_2 оба ученика взяты из набора № 3;

в) наконец, ученики из класса C_1 принадлежали к набору № 2, а оба ученика из класса C_2 — к набору № 3.

Все три возможности, вместе взятые, дают число благоприятных статочностей для (E/C_1), равное

$$C_{10}^2 C_{15}^2 + 10 \times 5C_{10}^2 + C_5^2 + C_{10}^2 = 7425.$$

Таким же путем получим число благоприятных статочностей и для (E/C_2):

$$C_5^2 C_{10}^2 + 5 \times 5 C_5^2 + C_5^2 C_5^2 = 800.$$

Отсюда, даже не вычисляя до конца значений p_1 и p_2, мы получим, предполагая, как раньше, $\omega_1 = \omega_2 = \frac{1}{2}$,

$$P_1 = \frac{7\,425}{7\,425 + 800} = \frac{297}{329} = 0{,}9027\ldots$$

Мы видим, как увеличение числа испытаний увеличивает шансы правильного умозаключения, если, конечно, мы не совершаем ошибок в наших суждениях об отдельных учениках, как мы это молчаливо предполагали при самой постановке задачи.

67. ИЗВЛЕЧЕНИЯ ИЗ УРНЫ, СОСТАВ КОТОРОЙ НЕИЗВЕСТЕН

Задача 19. *Урна содержит 6 шаров, среди которых могут быть и белые, и черные. Производится 6 извлечений, причем вынутый шар каждый раз возвращается обратно в урну; при каждом извлечении появлялся белый шар. Что можно сказать о составе урны?*

Можно составить семь различных гипотез о составе урны, предполагая в урне 0 белых и 6 черных шаров, 1 белый шар, и 5 черных и т. д., наконец, 6 белых шаров и ни одного черного. Если рассматривать эти гипотезы как равновозможные, то мы получим:

$$\omega_0 = \omega_1 = \omega_2 = \omega_3 = \omega_4 = \omega_5 = \omega_6 = \frac{1}{7}.$$

Вероятности извлечения белого шара, соответствующие перечисленным гипотезам, равны

$$0,\ \frac{1}{6},\ \frac{2}{6},\ \frac{3}{6},\ \frac{4}{6},\ \frac{5}{6},\ \frac{6}{6}$$

а вероятности извлечения шести белых шаров соответственно равны:

$$0,\ \left(\frac{1}{6}\right)^6,\ \left(\frac{2}{6}\right)^6,\ \left(\frac{3}{6}\right)^6,\ \left(\frac{4}{6}\right)^6,\ \left(\frac{5}{6}\right)^6,\ \left(\frac{6}{6}\right)^6.$$

При этих условиях вероятность того, что в испытании участвовала урна с шестью белыми шарами, равна

$$\frac{\omega_6 p_6}{\Sigma\, \omega_i p_i},$$

или, так как все ω_i равновероятны, то получим

$$\frac{p_6}{\Sigma p_i},$$

т. е.

$$P_6 = \frac{6^6}{1^6+2^6+3^6+4^6+5^6+6^6} = \frac{46\,656}{67\,171}$$

Эта наиболее вероятная гипотеза из семи возможных относительно состава урны; она даже более вероятна, чем все остальные гипотезы, вместе взятые. Отметим, что сделанное выше предположение о равновозможности всех гипотез совершенно произвольно; мы могли бы принять иные схемы, логически столь же приемлемые. Так, например, можно вообразить, что, имея урну U, содержащую одинаковое число белых и черных шаров, мы производим из нее шесть извлечений (заменяя вынутый шар новым того же цвета) и эти вынутые шары кладем в одну из урн из числа тех, которые входят в задачу.

Ясно, что в этом случае ω_i пропорциональны биномиальным коэффициентам ряда

$$(a+b)^6 = a^6 + 6a^5 b + 15a^4 b^2 + 20a^3 b^3 + 15a^2 b^4 + 6ab^5 + b^6.$$

Вероятность того, что урна содержит 6 белых шаров после того, как в шести испытаниях появились одни белые шары, равна:

$$P_6 = \frac{6^6}{6.1^6+15.2^6+20.3^6+15.4^6+6.5^6+6^6} = \frac{46\,656}{217\,392}.$$

Вероятность того, что в урне, взятой для испытания, все шары были белыми, теперь уже не является самой большой: она занимает лишь третье место, так как мы имеем

$$P_5 = \frac{93\,750}{217\,392}, \quad P_4 = \frac{61\,440}{217\,392}.$$

На этом примере мы еще раз убеждаемся в том, какое важное значение имеют априорные вероятности гипотез.

68. СЛУЧАЙ НЕПРЕРЫВНО ИЗМЕНЯЮЩИХСЯ ВЕРОЯТНОСТЕЙ

Речь идет об игре, аналогичной *орлянке*, при которой подбрасывается многогранник с одной или несколькими гранями, окрашенными в черный цвет, тогда как остальные оставлены белыми.

Когда подброшенный многогранник придет в состояние покоя, отмечают, какой из своих граней он лег на плоскость: белой или черной; эти два случая соответствуют орлу и решке при игре в орлянку. В данном случае вероятность орла неизвестна, так как мы ничего не знаем ни о числе, ни о площадях граней, закрашенных черным цветом, еще менее того знаем мы о самом многограннике; им могла бы быть игральная кость в форме кубика, у которой пять граней выкрашены черным; это мог бы быть деревянный кубик, у которого один из углов срезан, и образованный маленький треугольник выкрашен черным, тогда как все остальные грани оставлены белыми.

Сейчас мы хотим рассмотреть следующую задачу.

Событие E, состоящее в том, что $n+p$ партий дали p раз орла и n раз решку, осуществилось. Что можно из этого заключить относительно вероятности X выпадения орла в подобной игре?

По предположению мы ничего не знаем á priori относительно этой вероятности: *в качестве первого приближения мы принимаем, что наше незнание дает нам право рассматривать все возможные значения X между 0 и 1 как равно правдоподобные.* Исходя из этого, вероятность а priori для X заключаться между x_1 и x_2 мы полагаем равной $x_2 - x_1$, а вероятность для X лежать между x и $x+dx$ равна dx. Если X имеет значение x, то вероятность выкинуть орла p раз, а решку n раз равна:

$$\frac{(n+p)!}{n!\,p!}x^p(1-x)^n, \tag{135}$$

так как поскольку вероятность выпасть орлу равна x, постольку вероятность решки составляет $1-x$. Если мы воспользуемся обозначениями § 61, то величины ω_i (вероятность á priori) все равны dx, а величины p_k даны выражением (135).

Применяя формулу Байеса

$$P=\frac{p_k\,\omega_k}{\Sigma\, p_k\,\omega_k},$$

можно множитель, не зависящий от x, который имеется при всех p_k, отбросить. Таким образом, мы получим

$$P=\frac{x^p(1-x)^n\,dx}{\int\limits_0^1 x^p(1-x)^n\,dx}. \tag{136}$$

Такова вероятность a posteriori, т. е. та, которую мы определяем учитывая наблюденное событие, что X содержится между x и $x+dx$.

Мы могли бы исходя из более общих соображений ввести элементарную вероятность á priori в виде

$$d\omega=f(x)\,dx,$$

причем плотность вероятности $f(x)$ выражалась бы функцией, выбранной в соответствии с тем, что нам известно было бы до проведения испытания и удовлетворяющей, естественно, условию

$$\int\limits_0^1 f(x)\,dx=1;$$

Выражение для P получило бы в этом случае вид:

$$P=\frac{x^p(1-x)^n f(x)\,dx}{\int\limits_0^1 x^p(1-x)^n f(x)\,dx}.$$

На каких бы гипотезах мы ни остановили свой выбор, но величина p бесконечно мала, как бесконечно мало и dx; каково бы ни было значение x между 0 и 1, но $X = x$ всегда представляет собой событие почти невозможное и необходимо рассматривать интервал для x конечной величины, чтобы вероятность для X заключаться в конечном интервале сама имела бы конечную величину. Мы применим теперь полученную формулу к некоторым случаям.

69. ПРИМЕРЫ

З а д а ч а 20. *При игре в орлянку в условиях, указанных в предыдущем параграфе, две партии обе дали орла. Какую сумму можно поставить против одного франка за то, что этот результат является следствием того, что многогранник сделан так, что вероятность выпадения орла больше вероятности выпадения решки?*

Здесь мы имеем $p = 2$, $n = 0$, откуда для элементарной вероятности получаем значение

$$P = \frac{x^2\,dx}{\int\limits_0^1 x^2\,dx} = 3x^2\,dx.$$

Чтобы найти вероятность того, что X больше $\frac{1}{2}$, нам нужно проинтегрировать элементарную вероятность от $\frac{1}{2}$ до 1, что дает

$$\int\limits_{1/2}^{1} 3x^2\,dx = 1 - \frac{1}{8} = \frac{7}{8};$$

следовательно, можно поставить 7 франков против одного за то, что дело обстоит так, как это указано в задаче.

З а д а ч а 21. *Решить ту же задачу в предположении, что было сыграно k партий, в которых во всех выпадал орел.*

Мы знаем, что $p = k$, $n = 0$, следовательно,

$$P = (k+1)\,x^k\,dx;$$

вероятность, что $\left(X > \frac{1}{2}\right)$ выразится интегралом

$$\int\limits_{1/2}^{1} (1+k)\,x^k\,dx = 1 - \frac{1}{2^{k+1}}.$$

Вероятность противоположного события равна $\frac{1}{2^{k+1}}$; она очень быстро уменьшается по мере увеличения числа сыгранных партий.

З а д а ч а 22. *Решить ту же задачу в предположении, что было сыграно $h + 1$ партий, и только в одной из них выпала решка.*

На этот раз мы имеем $p = h$, $n = 1$, следовательно,

$$P = \frac{x^h(1-x)\,dx}{\int\limits_0^1 x^h(1-x)\,dx} = (h+1)(h+2)\,x^h(1-x)\,dx.$$

Вероятность, что X превышает $\frac{1}{2}$, находится вычислением интеграла

$$(h+1)(h+2)\int\limits_{1/2}^{1} x^h(1-x)\,dx = 1 - \frac{h+3}{2^{h+2}}.$$

В случае $h = 1$ эта вероятность равна $1 - \frac{1}{2} = \frac{1}{2}$, как это и следовало ожидать, для $h = 2$ ее значение равно $1 - \frac{5}{16} = \frac{11}{16}$, т. е. мы получаем меньшую величину, чем найденная в задаче 20.

70. СЛУЧАИ, КОГДА ЧИСЛО ИСПЫТАНИЙ ОЧЕНЬ ВЕЛИКО

Посмотрим теперь, что будет с элементарной вероятностью (136), если числа n и p очень велики.

Сперва вычислим интеграл

$$I_{p,n} = \int\limits_0^1 x^p(1-x)^n\,dx,$$

который стоит в знаменателе. Интегрируя по частям, получим

$$\int x^p(1-x)^n\,dx = \frac{1}{p+1}\,x^{p+1}(1-x)^n + \frac{n}{p+1}\int x^{p+1}(1-x)^{n-1}\,dx,$$

откуда, вводя пределы от 0 до 1,

$$I_{p,\,n} = \frac{n}{p+1}\,I_{p+1,\,n+1}$$

Эта рекуррентная формула позволяет написать

$$I_{p,\,n} = \frac{n}{p+1}\cdot\frac{n-1}{p+2}\cdots\frac{2}{p+n-1}\cdot\frac{1}{p+n}\,I_{p+n,\,0}$$

$$I_{p+n,\,0} = \int\limits_0^1 x^{p+n}\,dx = \frac{1}{p+n+1},$$

то мы окончательно получим

$$I_{p,\,n} = \frac{n!\,p!}{(n+p+1)!}. \tag{137}$$

Та плотность вероятности, из которой находится (136), равна, таким образом,

$$\delta x = \frac{(n+p+1)!}{n!\,p!} x^p (1-x)^n. \tag{138}$$

Вычислим исходя из этого математическое ожидание и типическое отклонение X; на основании § 20 мы имеем

$$m = E(X) = \int_0^1 x\,\delta x dx,$$

т. е.

$$m = \frac{(n+p+1)!}{n!\,p!} \int_0^1 x^{p+1} (1-x)^n\,dx;$$

из этого выражения, меняя в формуле (137) p на $p+1$ и производя сокращения, получим

$$m = \frac{p+1}{n+p+2}.$$

Дисперсия σ^2 со своей стороны дается выражением

$$\sigma^2 = E(X^2) - m^2 = \int_0^1 x^2\,\delta x dx - m^2,$$

т. е. после вычислений, подобных предыдущим,

$$\sigma^2 = \frac{(p+1)(n+1)}{(n+p+2)^2 (n+p+3)},$$

что можно написать и в таком виде:

$$\sigma^2 = \frac{m(1-m)}{n+p+3}.$$

Применение к неравенству Бьенэмэ—Чебышева

Так как вычисленное выше математическое ожидание m не может выйти за пределы сегмента от 0 до 1, то мы имеем неравенство

$$m(1-m) \leqslant \frac{1}{4}, \text{ откуда}$$

$$\sigma^2 \leqslant \frac{1}{4(n+p+3)}.$$

Вероятность для X отклониться от m на величину большую, чем ε, меньше, согласно неравенству Бьенэмэ—Чебышева, чем величина

$$\frac{1}{4\varepsilon^2 (n+p+3)}.$$

Поскольку величина ε может быть взята сколь угодно малой, но она все же постоянна, то можно выбрать $n+p$ настолько большим, чтобы вероятность

$$Pr(|X-m|>\varepsilon)$$

оказалась меньше данного числа α, которое само может быть взято произвольно малым. Отсюда получается «почти полная достоверность» при достаточно большом количестве испытаний, что их результат дает величину крайне близкую к неизвестной нам вероятности.

Предельный закон отношений для вероятности X.

Будем искать наиболее вероятную величину ξ для X; она будет такой, какая придает максимальное значение элементарной вероятности, пропорциональной $x^p(1-x)^n$.

Приравнивая логарифмическую производную этого произведения нулю, мы получаем соотношение

$$\frac{p}{\xi}=\frac{n}{1-\xi}\text{, откуда } \xi=\frac{p}{n+p};$$

говоря попросту, наиболее вероятной величиной для X является та, которая обусловливает равенство между *частостью* и *вероятностью* (выпадения орла).

Если число испытаний $p+n$ очень велико, то значение ξ очень мало отличается от математического ожидания m, это позволяет вместо предельного закона для отклонений X относительно m, строить предельный закон отклонений относительно ξ, а эта замена весьма облегчает вычисления.

Таким образом, исходя из формулы (138) и полагая

$$x=\xi+y=\frac{p}{n+p}+y,$$

получим

$$\log\delta(x)=p\log(\xi+y)-n\log(1-\xi-y)+\log(n+p+1)!-$$

$$-\log n!-\log p!.$$

Расчет, аналогичный произведенному в § 31, приводит, если ввести переменную λ такую, чтобы

(139) $$y=\frac{1}{n+p}\sqrt{\frac{np}{n+p}}\lambda,$$

к асимптотическому решению:

(140) $$\delta(z)\,dz\cong\frac{1}{\sqrt{2\pi}}e^{-\frac{\lambda^2}{2}}d\lambda.$$

Таким образом, отклонение $X-\xi$ следует закону вероятностей, сходящемуся к закону Лапласа—Гаусса для средней, равной 0 и типическому отклонению:

$$\frac{1}{n+p}\sqrt{\frac{np}{n+p}}. \tag{141}$$

Это типическое отклонение стремится *к нулю*, когда $n+p$ беспредельно возрастает; так как для нас «почти достоверно», что y не превзойдет пятикратного типического отклонения, то мы можем считать $\frac{p}{n+p}$ дающей достаточно точную величину для вероятности выкинуть орла в рассматриваемой нами игре.

Дадим теперь числовой пример, не забывая, что предшествующие расчеты делались в предположении равенства априорных вероятностей.

Положим, что мы имеем

$$p = 800, n = 100,$$

откуда

$$\frac{p}{n+p} = \frac{8}{9} = 0{,}888\ldots$$

Формула (139) дает в этом случае

$$\lambda = 95{,}4 \cdot y,$$

и мы видим, что вероятность для отклонения y превзойти 0,032 равна вероятности для λ оказаться большей примерно 3 единиц, которая меньше, чем 0,003. Мы имеем поэтому 997 шансов из 1000, что число орлов окажется в пределах

$$[0{,}888 - 0{,}032;\ 0{,}888 + 0{,}032],$$

иначе говоря, в пределах

$$[0{,}856;\ 0{,}920[.$$

Если бы мы предположили, что $p = 80\,000$, а $n = 10\,000$, то имели бы 997 шансов на 1000 за то, что та же вероятность выкинуть орла была бы заключена в пределы

$$[0{,}888 - 0{,}003;\ 0{,}888 + 0{,}003],$$

т. е. $[0{,}885;\ 0{,}891]$.

Мы видим, что увеличение количества испытаний ведет к сужению пределов: разность между пределами, как это вытекает из приведенных расчетов, изменяется обратно пропорционально корням квадратным из чисел испытаний.

Если условиться считать точностью величину обратно пропорциональную расхождению между пределами, то можно выразить вывод в таких словах: *точность пропорциональна квадратному корню из числа наблюдений.*

II. ОБЩИЙ ОБЗОР НЕКОТОРЫХ ПРОБЛЕМ, СВЯЗАННЫХ С ОЦЕНКОЙ

71. ОБЩИЕ ЗАМЕЧАНИЯ

Проблема, которая подлежит рассмотрению, относится к категории *проблем оценки*, среди которых она занимает особенно важное место.

Серия из $n+p$ испытаний дает нам значение $\xi = \frac{p}{n+p}$ для частости «орла». Таким образом, эта частость, установленная для $n+p$ испытаний, представляет собой случайную переменную, подчиняющуюся биномиальному закону, согласно которому $Pr\left(\xi = \frac{p}{n+p}\right)$ равна для $p = 0, 1, 2, \ldots, n+p$,

$$C^p_{n+p}\, x^p (1-x)^n,$$

где x *означает неизвестную вероятность.*

Невозможно в результате одной серии (или нескольких серий) подобных испытаний найти *точную величину* x; мы можем получить лишь приближенную величину, которую называют оценкой x. Это и есть то, что мы сделали в предыдущем параграфе, когда мы установили, ссылаясь на закон больших чисел, что $\xi = \frac{p}{n+p}$, представляет собой достаточно хорошее приближение к неизвестной вероятности. Рассмотрим типическое отклонение

$$\frac{1}{n+p}\sqrt{\frac{np}{n+p}}.$$

Это позволяет нам определить некоторую окрестность значения ξ, в которой находится точное значение x с известной вероятностью (например 997 шансов из 1000 для окрестности вправо и влево от ξ на утроенное типическое отклонение).

Эту оценку величины x отношением $\frac{p}{n+p}$, так же как отличную от нее оценку отношением $\frac{p+1}{n+p+2}$, полученную в том же параграфе, находят с помощью формулы Байеса и, опираясь на следующие допущения.

1) принимают, что x можно принципиально рассматривать как случайную переменную; 2) задаются произвольным законом вероятности a priori для этой переменной (предполагая, что все значения внутри сегмента (0, 1) имеют до всякого эксперимента равную вероятность).

Некоторые математики отбрасывают эти гипотезы, ибо, как они говорят, хотя x и неизвестен, но он все же величина определенная, а не случайная, а такими, т. е. случайными, являются результаты испытаний (опирающиеся на произвольные, а потому и спорные допущения), на основании которых устанавливаются принимаемые оценки. Мы не станем подробно разбирать эти возражения и ограничимся обзором некоторых правил, называемых *правилами непосредственной оценки*, которые можно формулировать, не прибегая к формуле Байеса.

72. ПРАВИЛА НЕПОСРЕДСТВЕННОЙ ОЦЕНКИ ПАРАМЕТРОВ

Мы ограничимся некоторыми указаниями, касающимися случая, когда речь идет об одном единственном параметре.

Пусть X — случайная переменная, закон вероятностей которой связан некоторым образом с неизвестной величиной θ (статистики говорят: «с неизвестным параметром θ»). Проводят серию n независимых между собой испытаний, из которых для X получаются значения x_1, x_2, ..., x_n, и затем принимают некоторую функцию

$$\theta' = \varphi(x_1, x_2, \ldots, x_n) \tag{142}$$

в качестве «оценки» θ. Поскольку каждое испытание призвано иметь одинаково важное значение, функция φ должна быть симметричной относительно x_1, x_2, ..., x_n. Такую функцию часто называют *статистическим выражением* (expression statistique) относительно испытаний или наблюдений, результатами которых и являются x_1, x_2, ..., x_n.

Статистическое выражение (142) считается определенно установленным, если каждая возможная серия из n испытаний, выполненная по данному типу, связывается с соответствующим значением θ'. В этом случае θ' представляет собой случайную переменную Y.

Про θ' говорят, что она является правильной оценкой θ, если случайная переменная Y выполняет следующие два условия:

$$\lim_{n \leftarrow \infty} E(Y) = \theta, \quad \lim_{n \to \infty} \sigma_Y = 0. \tag{143}$$

θ' называют *совершенно правильной оценкой* θ (estime absolument correcte), если Y выполняет два следующих условия:

$$E(Y) = \theta$$

каково бы ни было n,

$$\lim_{n \to \infty} \sigma_Y = 0. \tag{144}$$

Из этих условий второе, одинаковое со вторым условием (143), имеет в виду дисперсию, получаемую в форме

$$\sigma_y^2 = E\left[Y - \overline{Y}\right]^2.$$

В обоих случаях указанные условия имеют своим следствием на основании рассуждений, приведенных в § 36 и 37 то, что *Y стремится к θ как к вероятностному пределу* при безграничном возрастании n.

С каждой оценкой на практике связаны *доверительные границы*, которые сами зависят от определенного *коэффициента доверия*. В случае, например, задачи, рассмотренной в § 68, частость представляет собой случайную переменную, подчиняющуюся биномиальному закону, приведенному в § 71. Мы в этом случае уверены, согласно § 27 в том, что

$$E(\xi)=x, \tag{145}$$

и, кроме того, мы знаем, что типическое отклонение ξ равно

$$\sigma_\xi=\sqrt{\frac{x(1-x)}{n+p}}.$$

Это типическое отклонение нам неизвестно, но так как произведение $x(1-x)$ не может превысить $\frac{1}{4}$, мы с уверенностью принимаем, что

$$\sigma_\xi \leqslant \frac{1}{2\sqrt{n+p}}=\sigma_\theta \tag{146}$$

Отсюда следует, что x заключен в *доверительные границы* ($\xi-3\sigma_\theta$; $\xi+3\sigma_\theta$) с *коэффициентом доверия*, равным 0,997.

Оценка ξ для x, которая, таким образом, выполняет согласно (145) первое условие (144), выполняет также согласно (146) и второе условие, следовательно, эта *оценка — совершенно правильна.*

Вернемся к общей проблеме, поставленной в самом начале этого параграфа, чтобы связать с ней некоторые обычные формы определений. Множество всех возможных значений X называют генеральной совокупностью (population parente), множество значений $x_1, x_2, \ldots, x_n$, данных наблюдениями, именуют *выборкой*, взятой из этой совокупности; а количество n — *размером выборки*. Эту численность предполагают здесь *не ведущей к исчерпанию* всей генеральной совокупности.

Все эти выражения возникли в силу того важного значения, какую имеет оценка в *демографической статистике.*

73. ПРИМЕРЫ НЕПОСРЕДСТВЕННОЙ ОЦЕНКИ

Предположим, что имеется некоторая генеральная совокупность (P) и допустим, что мы все же знаем четыре первых момента, характеризующих ее закон распределения, т. е. моменты, обозначенные символами M, M_2, M_3 и M_4.

Произведена выборка (по схеме возвращенного шара), которая дала n величин $x_1, x_2, \ldots, x_n$ — значений X; поставим себе задачей дать оценку *математического ожидания*, а также *дисперсии* для X (иногда это выражают так: «дать оценку средней и дисперсии для генеральной совокупности»).

1) *Оценка средней для* (P). Пусть

$$m=\frac{x_1+x_2+\ldots+x_n}{n}$$

есть средняя для выборки. Обозначим через Y случайную переменную, которая может принимать всевозможные значения m, в зависимости от всех возможных составов выборки, численностью n, взятой из генеральной совокупности (P); символами же $X_1, X_2, \ldots, X_n$ мы обозначим случайные переменные, которые дают для выборок значения $x_1, x_2, \ldots, x_n$, при этом $X_1, X_2, \ldots, X_n$ подчиняются тому же закону вероятностей, как и X.

Мы имеем соотношение

$$Y = \frac{X_1 + X_2 + \ldots + X_n}{n},$$

а из него вытекает, что

$$E(Y) = \frac{1}{n}[E(X_1) + \ldots + E(X_n)] = \frac{n}{n}E(X) = M. \tag{147}$$

Таким образом математическое ожидание Y равно по величине математическому ожиданию X, и это имеет место независимо от того, как велико n; оценка M с помощью m удовлетворяет первому условию (144).

Теперь надо вычислить дисперсию Y; случайные переменные $X_1, X_2, \ldots, X_n$ очевидно взаимно независимы; поэтому согласно теореме из § 17 о сложении дисперсий мы получаем соотношение:

$$\sigma_Y^2 = \frac{1}{n^2}\sigma_{X_1}^2 + \ldots + \frac{1}{n^2}\sigma_{X_n}^2 = \frac{1}{n}\sigma_X^2 = \frac{1}{n}(M_2 - M^2). \tag{148}$$

Мы видим, что σ_Y стремится к нулю, когда n беспредельно возрастает, и таким образом, соблюдается и второе условие (144), следовательно, m есть *совершенно правильная оценка* для M.

2) *Оценка дисперсии* (P). Вычисления, которые мы проводим, несколько кропотливы, но дают интересный пример определения математического ожидания.

Рассмотрим *дисперсию, полученную из наблюдений выборки.*

$$z = \frac{1}{n}\sum (x_i - m)^2; \tag{149}$$

мы связываем с ней, как делали выше, случайную переменную

$$Z = \frac{1}{n}\sum (X_i - Y)^2 \tag{150}$$

и ставим задачу: найти математическое ожидание Z.

Положим, что для облегчения расчетов,

$$X = M + \xi,\ X_1 = M + \xi_1, \ldots, X_n = M + \xi_n, \tag{151}$$

где ξ_i — случайная переменная, подчиненная одному и тому же закону вероятности, а присущее им всем математическое ожидание равно:

$$E(\xi) = E(X) - M = 0,$$

Обозначим через μ_2, μ_3, μ_4 моменты для ξ второго, третьего и четвертого порядков; в частности:

$$\mu_2 = E(\xi^2) = M_2 - M^2.$$

Согласно (151) мы имеем

$$Y=M+\frac{1}{n}\sum \xi_i$$

и мы обозначим через ξ_0 арифметическую среднюю:

$$\xi_0=\frac{1}{n}\sum \xi_i.$$

Тогда мы получим, применяя к ξ_0 те же рассуждения, какие применялись в 1) относительно Y, соотношение:

$$E(\xi_0)=0,\ E(\xi_0^2)=\frac{\mu_2}{n}=\sigma_Y^2. \tag{152}$$

Вычислим $E(\xi_0^4)$, которое нам понадобится в дальнейшем ходе исследования. Мы имеем:

$$\xi_0^4=\frac{1}{n^4}(\Sigma\xi_i)^4=\frac{1}{n^4}\left[\xi_i^4+6\sum\xi_i^2\xi_j^2+\dots\right],$$

Члены, содержащие, хотя бы один множитель ξ_i в первой степени, не выписаны, так как их математическое ожидание равно нулю в силу взаимонезависимости переменных, и так как при любом i $E(\xi_i^4)=\mu_4$, и так как при любых i и j $E(\xi_i^2\xi_j^2)=\mu^2{}_2$, то отсюда следует, что выражение

$$E(\xi_0^4)=\frac{1}{n^4}\left[n\mu_4+6\frac{n(n-1)}{2}\mu_2^2\right]=\frac{\mu_4+3(n-1)\mu_2^2}{n^3}. \tag{153}$$

Теперь вернемся к случайной переменной Z, как она дана в формуле (150). Мы можем написать, замещая каждое X_i через $M+\xi_i$ и Y через $M+\xi_0$:

$$Z=\frac{1}{n}\sum(\xi_i-\xi_0)^2;$$

или, если принять во внимание, что $\Sigma\xi_i=n\xi_0$:

$$Z=\frac{1}{n}\sum\xi_i^2-\xi_0^2. \tag{154}$$

Отсюда получается следующее выражение для $E(Z)$:

$$E(Z)=\frac{1}{n}\cdot n\mu_2-\frac{\mu_2}{n}=\frac{n-1}{n}\mu_2,$$

и мы видим, что $E(Z)$ стремится к μ_2, когда n беспредельно растет. При данных условиях дисперсия z для выборки, полученной из наблюдений, *может притязать самое большее на наименование правильной оценки для* μ_2 (но еще нужно удостовериться в том, что σ_ξ стремиться к нулю вместе с $\frac{1}{n}$).

Для того чтобы получить оценку, относительно которой можно было бы при благоприятных обстоятельствах надеяться, что она будет совершенно правильной оценкой, мы должны принять выражение

$$\frac{1}{n-1}\sum (x_i-m)^2,$$

связав с ним случайную переменную

$$T=\frac{1}{n-1}\sum (X_i-Y)^2=\frac{n}{n-1}Z. \tag{155}$$

В этом случае мы, конечно, получим равенство:

$$E(T)=\mu_2.$$

Остается удостовериться в том, что

$$\sigma_T^2=E\left[(T-\mu_2)^2\right]$$

стремится к нулю при беспредельном возрастании n.

На основании (154) мы можем написать

$$T-\mu_2=\frac{1}{n-1}\left[\sum_i (\xi_i^2-\mu_2)-(n\xi_0^2-\mu_2)\right],$$

т. е.

$$\frac{1}{n-1}(A-B),$$

полагая

$$A=\sum_i (\xi_i^2-\mu_2), \quad B=n\xi_0^2-\mu_2. \tag{156}$$

Дисперсия σ_T^2 при этом получает такое выражение:

$$\frac{1}{(n-1)^2}\left[E(A^2)+E(B^2)-2E(AB)\right]. \tag{157}$$

Вычисление $E(A^2)$. Мы имеем, развертывая A^2,

$$E(A^2)=\sum_i E\left[(\xi_i^2-\mu_2)^2\right]+2\sum_{ij} E\left[(\xi_i^2-\mu_2)(\xi_j^2-\mu_2)\right].$$

Случайные переменные ξ_i^2 взаимно независимы, поэтому то же самое можно сказать и относительно величин $\xi_i^2-\mu_2$; каждый из ортогональных членов в приведенном выражении является, следовательно, произведением математических ожиданий

$$E(\xi_i^2-\mu_2); \quad E(\xi_j^2-\mu_2)$$

и соответственно равен нулю, так как $E(\xi_i^2)=\mu_2$.

Поэтому остаются только члены

$$E(A^2)=nE[(\xi^2-\mu_2)^2]=nE\left[\xi^4-2\mu_2\xi^2+\mu_2^2\right],$$

что, после приведения подобных членов, дает

$$(158)\qquad E(A^2)=n\left(\mu_4-\mu_2^2\right).$$

Вычисление $E(B^2)$. Мы получаем, применяя те же соображения,

$$E(B^2)=E\left(n^2\xi_0^4-2n\mu_2\xi_0^2+\mu_2^2\right),$$

что согласно (152) и (153) после некоторых сокращений, дает

$$(159)\qquad E(B^2)=\frac{\mu_4}{n}+\frac{2n-3}{n}\mu_2^2.$$

Вычисление $E(AB)$. Мы можем B придать такую форму:

$$n\xi_0^2-\mu_2=\frac{(\xi_1+\xi_2+\ldots+\xi_n)^2-n\mu_2}{n}=\frac{1}{n}\left[\sum(\xi_j^2-\mu_2)+2\sum\xi_j\xi_k\right].$$

В этом случае $E(AB)$ можно написать в виде:

$$\frac{1}{n}E\left\{\left[\sum_i(\xi_i^2-\mu_2)\right]\left[\sum(\xi_j^2-\mu_2)+2\sum\xi_j\xi_k\right]\right\},$$

что дает, после исключения всех тех членов, математическое ожидание которых равно нулю,

$$\frac{1}{n}E\left[\sum(\xi^2-\mu_2)^2\right]=E(\xi^2-\mu_2)^2,$$

откуда получим окончательно

$$(160)\qquad E(AB)=\mu_4-\mu_2^2.$$

Внося значения (158), (159) и (160) в формулу (157), мы в конце концов находим

$$(161)\qquad \sigma_T^2=\frac{\mu_4}{n}-\frac{n-3}{n(n-1)}\mu_2^2.$$

Дисперсия σ_T^2 стремится, очевидно, к нулю при беспредельном возрастании n и выражение

$$\frac{1}{n-1}\sum(x_i-m)^2$$

вполне заслуживает наименования совершенно правильной оценки μ_2.

Делаем вывод: *если* (P) — *совокупность, имеющая какой-либо закон распределения, характеризуемый моментами первых четырех порядков, то мы получим совершенно правильные оценки неизвестных нам математических ожиданий и дисперсий совокупности* (P), *если исходя из* *выборки (по схеме возвращенного шара) численностью в n мы распространим на генеральную совокупность следующие величины:*

(162) $$m=\frac{1}{n}\Sigma x_i,\quad \mu_2'=\frac{1}{n-1}\Sigma(x_i-m)^2.$$

74. ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ, СВЯЗАННЫЙ С ОЦЕНКОЙ МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ

Примем, как допущение, что генеральная совокупность (P) *имеет нормальное распределение.* Плотность вероятности X в таком случае равна:

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

Все переменные X_i следуют тому же закону и согласно § 37 переменная

$$Y=\frac{1}{n}(X_1+X_2+\ldots+X_n)$$

подчиняется нормальному закону со средней M и типическим отклонением $\frac{\sigma}{\sqrt{n}}$.

Отсюда следует, например, что m имеет 95% шансов за то, что она (см. § 35) отклонится от M на величину, не превышающую по своей абсолютной величине $\frac{2\sigma}{\sqrt{n}}$. В таком случае мы говорим, что доверительным интервалом является интервал

(163) $$m-\frac{2\sigma}{\sqrt{n}},\quad m+\frac{2\sigma}{\sqrt{n}},$$

в котором лежит M с доверительным коэффициентом в 95%.

Типическое отклонение нам, правда, неизвестно, но мы пользуемся на практике его оценкой $\sqrt{\frac{\Sigma(x_i-m)^2}{n-1}}$, которая с большой вероятностью близка к подлинному значению σ если n достаточно велико.

Важные замечания. Можно доказать, что *при большом n, Y приближенно следует нормальному закону даже в тех случаях, когда распределение (P) имеет другой закон, и это справедливо при весьма широких допущениях:* существования M и μ_2. Эту основоположную теорему часто называют *центральной предельной теоремой.*

Поэтому вышесказанные замечания, например, относительно доверительного интервала (163) практически всегда справедливы. Типическое отклонение $\frac{\sigma}{\sqrt{n}}$ (или его оценка) обычно называется *стандартной ошибкой средней величины.*

75. ПРИМЕНЕНИЯ

В июне 1953 г. во время экзаменов на степень бакалавра в Тулузе два экзаменатора C_1 и C_2 проверили каждый по 63 сочинения по французскому языку из числа 1400, принадлежавших к серии «новые» (moderne). Сочинениям давалась оценка по 31-балльной системе (от 0 до 30, теоретическая средняя 15). Средние, выведенные из отметок этих двух экзаменаторов, были

$$m_1 = 9{,}80, \quad m_2 = 12{,}95,$$

а дисперсия отметок

$$v_1 = 10{,}80, \quad v_2 = 33{,}13.$$

Можно считать, что выборки, взятые в порядке алфавита, попавшие в руки экзаменаторов C_1 и C_2, были получены в случайном порядке, с другой стороны, учитывая значительность генеральной совокупности, состоявшей из 1400 сочинений, выборку можно признать сделанной по схеме возвращенного шара.

Дадим оценку средней M и дисперсии μ_2 на основании выборки тех сочинений, которые проверялись экзаменатором C_1: оценка средней равняется $m_1 = 9{,}80$, а для дисперсии имеем:

$$\mu_2' = \frac{n}{n-1} v_1,$$

оценка же стандартной ошибки средней равна:

$$\frac{\sigma'}{\sqrt{n}} = \sqrt{\frac{v_1}{n-1}} = \sqrt{\frac{10{,}80}{62}} = 0{,}41,$$

доверительный интервал на уровне 95% равен, следовательно, согласно (163) интервалу

$$(9{,}80 - 0{,}82; \quad 9{,}80 + 0{,}82), \quad \text{или} \quad (8{,}98; \quad 10{,}62),$$

и мы имеем, таким образом; приближенно 95 шансов из 100 за то, что средняя M находится внутри интервала (8,98; 10,62).

Сходные расчеты, но исходящие от сочинений, проверявшихся экзаменатором C_2, дают в качестве доверительного интервала [11,49; 14,41] при уровне 95%.

Таким образом интервалы не пересекаются. Это заставляет нас отказаться от предположения, что обе проверенные выборки взяты из одной и той же генеральной совокупности (P). Однако в рассматриваемом примере случайный характер отбора гарантирован тем, что сочинения брались в порядке алфавита. Это вынуждает нас признать, что *оба экзаменатора по-разному расценивали работы.*

Более точным и более полным методом, который мы здесь не станем излагать, были для всех 1400 сочинений установлены оценки: 12,88 для среднего балла (M), и, с другой стороны, оказалось возможным утверждать с вероятностью в 0,95, что средние по выборкам в 63 работы, отобранные по жребию, должны находиться в пределах

$$(11{,}29; \ 14{,}57).$$

Средняя оценка работ у экзаменатора C_2 равна 12,95 и лежит в указанных границах; средняя же у экзаменатора C_1 вышла далеко за указанные пределы. Поэтому мы вправе с «почти полной достоверностью» утверждать, что экзаменатор C_1 оказался гораздо более строгим, чем большинство его коллег, принимавших участие в поверке 1400 сочинений, образующих генеральную совокупность (P).

76. МЕТОД НАИБОЛЬШЕГО ПРАВДОПОДОБИЯ

Выше мы указали некоторые правила, пригодные к тому, чтобы оценить степень доверия, с которым можно относиться к непосредственной оценке, каким бы способом эта оценка ни была получена; затем мы применили эти правила к особенно важному случаю оценки средней.

Мы закончим эту главу указанием на один общий метод непосредственных оценок, носящий наименование *метода наибольшего правдоподобия.*

Примем условия § 68 и рассмотрим случайную переменную X, закон вероятностей которой зависит от некоторого параметра θ, которому и надлежит дать оценку на основании результатов n взаимно независимых испытаний. Обозначим через

$$f(x_i, \theta)$$

вероятность того, что $X = x_i$, тогда вероятность получить выборку, состоящую из $(x_1, x_2, \ldots, x_n)$, выразится произведением:

$$f(x_1, \theta) f(x_2, \theta) \ldots f(x_n, \theta) = \mathcal{L}(x_1, x_2, \ldots, x_n, \theta). \tag{164}$$

Эта вероятность называется *правдоподобием выборки, и* метод состоит в том, чтобы *выбрать такую оценку* θ, *которая обусловила бы максимум этого правдоподобия.*

В том случае, когда случайная переменная X принимает непрерывные значения

$$Pr[x_i < X < x_i + dx_i] = f(x_i, \theta)\, dx_i$$

правдоподобие пропорционально выражению (164).

Чтобы вычислить θ', нужно приравнять нулю логарифмическую производную (164); тогда θ' будет корнем уравнения

$$\sum \frac{1}{f(x_i, \theta)} \frac{\partial}{\partial \theta} f(x_i, \theta) = 0. \tag{165}$$

Если имеется несколько параметров $\theta_1, \theta_2, \ldots$, подлежащих совместной оценке, то метод состоит в том, чтобы решить систему уравнений:

$$\frac{\partial \mathcal{L}}{\partial \theta_1} = 0, \qquad \frac{\partial \mathcal{L}}{\partial \theta_2} = 0, \ldots$$

Пример I. *Случай биномиального закона.*

X может принимать значения 0 или 1 с вероятностями θ и $1 - \theta$. x_i, следовательно, равняются или 0, или 1.

Если обозначить через k количество x_i, равных 0 в выборке численностью в n наблюдений, то получим:

$$\mathscr{L} = \theta^{k}(1-\theta)^{n-k};$$

уравнение для наибольшего правдоподобия будет таким:

$$\frac{k}{\theta} - \frac{n-k}{1-\theta} = 0,$$

что для оценки дает значение

$$\theta' = \frac{k}{n};$$

этот результат мы уже получали с несколько иными обозначениями в § 66.

Пример II. *Случай нормального закона вероятностей.* Элементарная вероятность составляет:

$$f(x, m, \sigma)\,dx = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-m)^2}{2\sigma^2}}\, dx,$$

правдоподобие выборки $(x_1, x_2, \ldots, x_n)$ пропорционально

$$\mathscr{L} = \frac{1}{\sigma^n (2\pi)^{n/2}}\, e^{-\frac{1}{2\sigma^2}\Sigma(x_i-m)^2}.$$

Оценку параметров m и σ методом наибольшего правдоподобия мы получим из решения системы уравнений:

$$\frac{1}{\mathscr{L}}\frac{\partial \mathscr{L}}{\partial m} = \frac{1}{2\sigma^2}\Sigma(x_i - m) = 0,$$

$$\frac{1}{\mathscr{L}}\frac{\partial \mathscr{L}}{\partial \sigma} = \frac{n}{\sigma} + \frac{1}{\sigma^3}\Sigma(x_i - m)^2 = 0.$$

Отсюда сразу же получаем

$$m' = \frac{1}{n}\sum x_i; \quad \sigma'^2 = \frac{1}{n}\sum(x_i - m')^2;$$

и мы убеждаемся в том, что полученные оценки равны соответственно *средней и дисперсии выборки.* Вторая из оценок, как мы видели в § 69, является оценкой *правильной* но не *совершенно правильной.*

Можно показать, что при очень общих допущениях метод наибольшего правдоподобия всегда приводит к *правильным оценкам.* В случае одного единственного параметра и для *бесконечно большого* числа наблюдений получаемой этим методом оценке присуще *типическое отклонение наименьшей величины.*

77. ОБ ОПРЕДЕЛЕНИИ ПРИЧИН

Изложенные выше объяснения и особенно пример, разобранный в § 75, показывают, что задачи, с которыми сталкиваются на практике, часто ставят вопрос в такой форме: *обусловлен ли полученный результат лишь игрой случая, или же он вызывается какой-либо существенной причиной?*

На этот вопрос теория вероятностей чаще всего дает такой ответ, который нельзя рассматривать как окончательно определенный, но который нельзя по справедливости рассматривать и как лишенный всякого значения, подобно тому как это делает Ж. Бертран в своем классическом трактате.

Один из вопросов, рассмотренных Бертраном, — это вопрос о *двойных звездах*. Задача состоит в следующем: для двух звезд наблюдение дает очень близкие направления — очень близкие *на небесной сфере*; можно ли отсюда заключить, что эти звезды *близкие соседи и в пространстве?*

Можно получить очень важное указание, если вычислить вероятность того, что угловое расстояние двух точек, *взятых наугад* на поверхности полушария, окажется меньше величины α. Из расчета, приведенного в § 60, следует, что эта вероятность равна:

$$1 - \cos\alpha,$$

или приближается к $\frac{\alpha^2}{2}$, если α мало. Так, для $\alpha = 10'$ получим:

$$p = \frac{\pi^2}{2\,(1\,080)^2}, \text{ т. е. приближенно } \frac{1}{236\,000}.$$

Число звезд первых трех степеней яркости, согласно данным, приведенным у Бертрана, достигает числа 230; они образуют

$$\mu = \frac{230 \times 229}{2} = 26\,335$$

комбинаций по две звезды. Вероятность того, что ни для одной пары звезд угловые расстояния между ними не окажутся меньше $10'$, равна, если применить правило перемножения вероятностей,

$$(1 - p)^{\mu} = 0{,}897 \text{ (приближенно)},$$

если приравнять наблюдаемые звезды к точкам на небесной сфере, взятым наугад.

Такая справка не лишена некоторого значения и подкрепляет достаточно точно презумпции того, что наблюдаемые *двойные звезды* являются на самом деле близко расположенными в пространстве; à fortiori можно предполагать, что звездная группа Плеяд составлена из звезд, действительно расположенных близко одна от другой.

Теория вероятностей показывает вероятность того действия, которое может оказать *случай сам по себе*. Она не может дать большего, но, несмотря на это, она указывает путь для исследования в данном случае и в аналогичных, приводящих практически к полной достоверности, которая может быть достигнута в каждой науке только такими *методами, которые присущи* данной отрасли знаний.

1. Петр играет в экартэ с незнакомым ему партнером, который с первой партии открывает короля. Какова вероятность, что партнер Петра шулер? (Пуанкаре).

Ответ: необходимо приписать какое-либо определенное значение вероятности à priori ω, допуская, кроме того, что шулер вскрывает короля один раз на каждые две игры (допущение вполне произвольное).

$$\omega_1 = \omega,$$
$$\omega_2 = 1 - \omega$$

и

$$p_1 = \frac{1}{2}, \quad p_2 = \frac{1}{8}.$$

Тогда

$$p = \frac{p_1 \omega_1}{p_1 \omega_1 + p_2 \omega_2} = \frac{4\omega}{3\omega + 1}.$$

Таким образом, если $\omega = 0$, то и $p = 0$; если же $\omega = 1$, то и $p = 1$.

2. Урна содержит шесть шаров: белых и (или) черных. Производят четыре извлечения (без возвращения шаров обратно в урну) и получают выход четырех белых шаров. Какова вероятность того, что следующие два вынутых шара будут также белыми?

Ответ: Примем, что априорные вероятности $\omega_1, \omega_2, \ldots, \omega_7$ пропорциональны числам 1, 6, 15, 20, 15, 6, 1. Результат испытаний, т. е. выход четырех белых шаров, имеет при таких предположениях вероятности (соответственно), равные 0,0, 0, 0, $\frac{1}{15}$, $\frac{1}{3}$, 1. При таких предположениях вероятность того, что все шары в урне были белыми, составляет:

$$\frac{1}{1 + \frac{6}{3} + \frac{15}{15}} = \frac{1}{4}.$$

Выявленный результат ничего нам не говорит об оставшихся в урне шарах, но легко отдать себе отчет о том, что так обстоит дело всегда, когда состав урны, послужившей для назначения априорных вероятностей, определяется случайными извлечениями, произведенными независимо для каждого шара.

3. Дайте оценку для λ в случае закона Пуассона (§ 18) методом наибольшего правдоподобия.

Ответ: Оценка равна $\lambda' = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$.

4. Пусть x — неизвестная пропорция белых шаров в урне, из которой произведено n извлечений, сделанных наугад по типу возвращенного шара; предполагается, что n велико. Пусть y — частость белых шаров в этой выборке и M — точка с прямоугольными координатами x и y.

Показать, что если x не слишком мало, то условие

$$|y-x|<2\sqrt{\frac{x(1-x)}{n}}$$

можно выразить словами: точка M должна находиться внутри некоторого эллипса, который можно построить. Вывести оценку x на доверительном уровне в 95% для случая, когда $n=900$, $y=\frac{1}{9}$.

5. Мы имеем урну с m шарами, помеченными каждый очередным номером x_i. Положим

$$\bar{x}=\frac{1}{m}\sum x_i, \qquad \sigma^2=\frac{1}{m}\sum (x_i-\bar{x})^2.$$

С каждым шаром связываем случайную переменную A_i, равную единице, если шар извлечен, и равную 0 — в противоположном случае.

Из урны делается n извлечений без возврата шара. Пусть $\bar{X}$ есть средняя из извлеченных таким образом номеров.

Показать, что вероятность для $A_i=1$ равна $\frac{n}{m}$ и что математическое ожидание случайной переменной $\bar{X}$ равно:

$$E(\bar{X})=\sum_{i=1}^{m}\frac{x_i}{n}E(A_i)=\bar{x},$$

а для дисперсии получается:

$$\text{дисперсия } \bar{X}=\frac{m-n}{m-1}\,\frac{\sigma^2}{n}.$$

5. ОШИБКИ НАБЛЮДЕНИЯ. ЗАКОН ГАУССА

78. ПОСТАНОВКА ПРОБЛЕМЫ

Требуется измерить какую-либо величину, скажем, длину, истинное значение которой равно a; если измерение дает величину a', то ошибка x равна $a' - a$.

Несколько раз повторенное измерение дало:

$$a_1, \ a_2, \ \ldots, \ a_n;$$

что следует взять за оценку неизвестной величины a, и в какой мере можно доверять принятой оценке?

Такова основная проблема теории ошибок наблюдений. Темой настоящей главы служит рассмотрение закона вероятностей, общепризнанного в области случайной переменной X, значениями x которой являются ошибки, совершаемые при измерении a, производимом в неизменных условиях. Этот закон называется *законом Гаусса*.

79. СИСТЕМАТИЧЕСКИЕ ОШИБКИ. СЛУЧАЙНЫЕ ОШИБКИ

Какое-либо физическое измерение может быть искажено ошибками двоякого рода: систематическими и случайными.

Систематические ошибки вызываются недостатками измерительных приборов; они в почти неизменном размере отражаются на многочисленных измерениях одного и того же объекта и они изменяют их все в одном и том же направлении.

Случайные ошибки возникают от многочисленных причин, которые искажают результаты последовательно повторенных измерений какого-либо определенного предмета; эти искажения тем меньше, чем точнее инструмент, с помощью которого производятся измерения. Они образуют две группы, действующие в противоположных направлениях.

В дальнейшем мы будем предполагать, что систематические ошибки исключены из всех произведенных измерений путем введения «поправок» применительно к каждому из них; известно, как многочисленны и значительны такие исправления, хотя бы, скажем, при измерении температур.

Однако в действительности это требование соблюдается лишь номинально и никогда в точности не выполняется.

80. СОДЕРЖАНИЕ ЗАКОНА ГАУССА

Закон Гаусса состоит в том, что вероятность для ошибки $a' - a$ иметь величину, лежащую в пределах x и $x + dx$ равнялась бы:

$$Pr\,[x < X < x + dx] = \frac{k}{\sqrt{\pi}}\, e^{-k^2 x^2}\, dx, \tag{166}$$

где k — некоторая постоянная, которую называют точностью, присущей данному способу измерения.

Уже сама формулировка этого закона предполагает некоторое множество наблюдений.

Прежде всего спрашивается, что следует понимать под истинным значением измеряемой величины. Часто бывает так, что это значение выясняется с некоторой точностью лишь в результате многих измерений; не вращаемся ли мы здесь в порочном кругу, определяя ошибки измерений с помощью числа, находимого из этих же самых измерений?

Мы не будем настаивать здесь на этом затруднении, возникающем в теории вероятностей причин, укажем лишь, что *бывают случаи в которых это затруднение отсутствует*: можно допустить что мы ограничиваемся, по крайней мере предварительно, именно такими случаями; это имеет место, скажем, тогда, когда измеряемая величина нам известна на основании предшествовавших измерений, в сочетании или независимо от теоретических выводов, *с точностью, значительно превышающей точность вновь предпринимаемых измерений*. Так бывает при физических или химических опытах, в которых большое число учеников должны определить какую-либо константу, заранее хорошо известную.

То же можно сказать и о постоянном k, входящем в формулу Гаусса, которая определяется в результате наблюдений; далее будет указано, как нужно производить это определение.

Согласно закону Гаусса вероятность положительных и отрицательных ошибок одинакова, если они равны по своей абсолютной величине, а их симметрию можно предполагать оправданной лишь в тех случаях, когда она фактически проявилась.

Заметим, наконец, что при замене x какой-либо его функцией математическая форма закона может оказаться измененной; это служит поводом для одного возражения против закона Гаусса, аналогичного тем, с которыми мы уже встречались, и на него следует ответить так же, как мы отвечали раньше, т. е. обращаясь к фактам; если мы для измерения какой-либо длины пользовались линейкой с нанесенными на ней делениями, то эти деления в разных частях линейки должны быть однородны и мы измеряем не квадраты длин, а их протяженность и именно к ней, а не к квадратам относятся совершаемые ошибки.

Вместе с тем, если ошибки малы, то формула Гаусса сохраняет свой вид в первом приближении. Положим

$$b=\varphi(a), \qquad b'=\varphi(a'),$$
$$x=a'-a, \qquad y=b'-b.$$

Тогда получим

$$y=\varphi'(\alpha)\,x,$$

обозначая через α некоторое число, лежащее между a и a', и так как все измерения a' лежат в достаточно узких пределах, то с большим приближением мы получаем $\varphi'(\alpha)=\varphi'(a)=c$, откуда следует

$$\frac{k}{\sqrt{\pi}}e^{-k^2x^2}dx=\frac{k}{c\sqrt{\pi}}e^{-\frac{k^2}{c^2}y^2}dy.$$

Мера точности k оказывается просто поделенной на c.

Предположим, например, что измеряется расстояние в 100 *м* и что результаты отдельных измерений различаются между собой не менее чем на метр, если вместо измеренной величины введем ее квадрат, то получим:

$$\varphi(a)=a^2, \qquad \varphi'(a)=2a,$$

и мы видим, что значение c содержится между 199 и 201 *м*. Принимая c равным 200 *м*, мы совершаем относительную ошибку, лежащую в пределах $\pm\frac{1}{200}$.

81. ОСНОВЫ ЗАКОНА ГАУССА

Можно было бы ограничиться указанием на то, что закон Гаусса оправдывается на опыте, и дать название *нормальных серий* тем сериям наблюдений, которые ему подчиняются, тогда как серии, к которым этот закон не применим, называть *анормальными сериями*.

Установив эти определения, можно построить теорию нормальных рядов измерений, правомерно основывая ее на законе Гаусса. Некоторые, поступая так, полагают, что это есть право или, как они утверждают, даже обязанность каждого ученого определять употребляемые им термины наиболее ясным и для него удобным образом, другие же, более требовательные считают, что этим лишь перемещается та трудность, которую нельзя разрешить условной терминологией, как бы она ни была остроумна, и что дело, по существу, состоит в том, чтобы выяснить, почему серии измерений так часто оказываются нормальными. Вот это мы и попытаемся кратко пояснить.

Что происходит, когда мы производим измерение?

Причины ошибок чрезвычайно многочисленны и мы можем принять, что итоговая ошибка есть алгебраическая сумма частных ошибок. **Эти** частные ошибки распределены в некотором поле (по ту и по другую сторону от нуля, и мы можем разбить это поле на достаточно узкие интервалы, центры которых имеют своими абсциссами $\alpha_1, \alpha_2, \ldots, \alpha_l$. Можно затем признать за всеми частными ошибками, приходящимися на каждый интервал, вероятность, равную абсциссе центра этого интервала. Тогда окончательную ошибку можно рассматривать как сумму некоторого количества n_1 ошибок размера α_1, некотороо количества n_2 ошибок размера α_2 и т. д. При этом вероятность ошибки размера α_1 равна p_1, вероятность ошибки размера α_2 равна p_2 и т. д.

Все эти допущения приемлемы; они направлены в конечном счете на то, чтобы *схематизировать происходящее в действительности в виде опыта с урнами.*

Мы предполагаем ряд урн

$$U_1, U_2, \ldots, U_l.$$

Урна U_1 содержит в пропорции p_1 шары, помеченные числом α_1, тогда как другие шары в ней имеют цифру 0; урна U_2 содержит в пропорции p_2 шары, помеченные числом α_2, остальные же имеют на себе цифру 0 и т. д. Мы производим n_1 извлечений из урны U_1; n_2 извлечений из урны $U_2, \ldots, n_l$ извлечений из U_l. Итоговая ошибка равна сумме номеров, которыми помечены извлеченные шары.

Пусть X_i случайная переменная, представляющая собой сумму номеров, которую можно получить при n_i извлечениях из урны U_i. Случайная переменная, показывающая общую сумму, равна:

$$X = X_1 + X_2 + \ldots + X_l. \tag{167}$$

Введем для каждого значения указателя i случайную переменную Y_i, показывающую число шаров с пометкой α_i, которые могут оказаться вынутыми из урны U_i при n_i извлечениях. Обе случайные переменные X_i и Y_i связаны точным соотношением

$$X_i = \alpha_i Y_i.$$

И вот мы знаем из теории повторных испытаний в случае одной урны с шарами двух сортов (§ 26), что математическое ожидание $E(Y_i)$ равно $n_i p_i$ и что дисперсия Y_i равна $n_i p_i q_i$; кроме того, мы знаем, что при очень больших n_i (вследствие того, что число частных ошибок очень велико) приведенное отклонение

$$\frac{Y_i - n_i p_i}{\sqrt{n_i p_i q_i}}$$

может быть приравнено случайной переменной с непрерывными значениями, подчиняющимися приведенному нормальному закону.

Соотношение $X_i = \alpha_i Y_i$ в этих условиях приводит к тому, что

$$E(X_i) = \alpha_i E(Y_i) = \alpha_i n_i p_i,$$

и дисперсия σ_i^2 переменной X_i выразится равенством:
$\sigma_i^2 = \alpha_i^2$ дисп. $(Y_i) = \alpha_i^2 n_i p_i q_i$.

Отношение

$$\lambda_i = \frac{X_i - E(X_i)}{\sigma_i} = \frac{\alpha_i Y_i - \alpha_i n_i p_i}{\alpha_i \sqrt{n_i p_i q_i}} = \frac{Y_i - n_i p_i}{\sqrt{n_i p_i q_i}} \tag{168}$$

подчиняется, следовательно, *приведенному нормальному закону*.

Обратимся теперь к математическому ожиданию X:

$$E(X) = E(X_1) + \ldots + E(X_l) = \sum \alpha_i n_i p_i.$$

Это математическое ожидание *равно нулю по условию*, согласно которому отрицательные ошибки столь же вероятны, как равновеликие с ними положительные. Из выражения (168) вытекает, что сумма

$$\sigma_1 \lambda_1 + \sigma_2 \lambda_2 + \ldots + \sigma_l \lambda_l$$

имеет значение X. Теорема сложения дисперсий дает поэтому для дисперсии σ^2 для X выражение:

$$\sigma^2 = \alpha_1^2 n_1 p_1 q_1 + \ldots + \alpha_l^2 n_l p_l q_l,$$

благодаря чему отношение

$$\lambda = \frac{X - E(X)}{\sigma}$$

можно написать в виде:

$$\lambda = \frac{X}{\sigma} = \frac{\sigma_1}{\sigma} \lambda_1 + \ldots + \frac{\sigma_l}{\sigma} \lambda_l.$$

Так как λ_i следует приведенному нормальному закону, то из § 40 вытекает, что λ следует нормальному закону для величин, математическое ожидание которых равно нулю, а типическое отклонение

$$\frac{\sigma_1^2}{\sigma^2} + \ldots + \frac{\sigma_l^2}{\sigma^2} = 1,$$

т. е. следует *приведенному нормальному закону*.

В этих условиях элементарная вероятность для λ равна:

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{\lambda^2}{2}} d\lambda,$$

а элементарная вероятность для $X = \sigma\lambda$ напишется вследствие этого так:

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx.$$

Это выражение тождественно равно (166) со значением

$$k = \frac{1}{\sigma\sqrt{2}} \tag{169}$$

параметра, характеризующего точность.

Сводя все изложенное, скажем: закон Гаусса оправдывается тем фактом, что возможные ошибки многочисленны, и их сочетания подчинены законам, известным нам из комбинаторики для случая многократно повторенных испытаний.

82. ДОКАЗАТЕЛЬСТВО, ПРИВЕДЕННОЕ ГАУССОМ

При наличии многих измерений, заслуживающих одинаковой степени доверия, все наблюдения должны приниматься во внимание в одинаковой мере, и тогда наиболее правдоподобной оценкой измеряемой величины a надо считать среднюю арифметическую

$$\frac{z_1 + z_2 + \ldots + z_n}{n}$$

из результатов измерений. Гаусс показал, что если эта средняя арифметическая во всех случаях является *наиболее вероятным значением* измеряемой величины, рассматриваемой как одно из возможных значений случайной переменной Z, то закон ошибок непременно должен быть таким, каким его дает формула (166).

Пусть $f(z)dz$ есть априорная вероятность того, что Z имеет значение, заключенное между z, и $z + dz$; Гаусс принципиально считает, что при полном нашем незнании распределения á priori значении Z надо $f(z)$ принять за постоянную величину. С другой стороны, если предположить, что Z приняло значение z, вероятность того, что измерение дает результат z_1, искаженный ошибкой $x_1 = z_1 - z$, лежащей в пределах x_1 и $x_1 + dx_1$, имеет вид: $\varphi(x_1)dx_1$, или $\varphi(z_1 - z)dz_1$.

Предполагая n измерений взаимно независимыми, а Z приняло значение z, мы будем иметь для вероятности того, что эти измерения дадут результаты, лежащие между z_1 и $z_1 + dz_1$, между z_2 и $z_2 + dz_2$ и т. д. согласно правилу перемножения вероятностей:

$$\varphi(z_1 - z) \cdot \varphi(z_2 - z) \ldots \varphi(z_n - z)\, dz_1\, dz_2 \ldots dz_n.$$

Вероятность же à posteriori того, что значение Z будет находиться между z и $z + dz$, если известно, что уже выполненные измерения n дали результаты $z_1, z_2, \ldots, z_n$, выражается, следовательно, и по формуле Байеса и после сокращения наверху и внизу произведений $dz_1\, dz_2 \ldots dz_n$, а также и постоянного множителя $f(z)$ в таком виде:

$$\frac{\varphi(z_1 - z)\, \varphi(z_2 - z) \ldots \varphi(z_n - z)}{\int\limits_{-\infty}^{+\infty} \varphi(z_1 - z)\, \varphi(z_2 - z) \ldots \varphi(z_n - z)\, dz}. \tag{170}$$

Величина z, придающая наибольшее значение вероятности à posteriori, получается из уравнения:

$$\frac{\varphi'(z_1 - z)}{\varphi(z_1 - z)} + \ldots + \frac{\varphi'(z_n - z)}{\varphi(z_n - z)} = 0, \tag{171}$$

и наша задача состоит в том, чтобы найти для φ такую форму, при которой значение z было бы всегда средним арифметическим из результатов измерений.

Положим для $i = 1, 2, \ldots, n$

$$x_i = z_i - z, \quad F(x_i) = \frac{\varphi'(z_i - z)}{\varphi(z_i - z)};$$

тогда мы должны получить

$$F(x_1) + F(x_2) + \ldots + F(x_n) = 0$$

во всех тех случаях, когда

$$x_1 + x_2 + \ldots + x_n = 0,$$

и это условие должно удовлетворяться, каким бы ни было n и какими бы ни были $x_1, x_2, \ldots, x_{n-1}$.

Это неизбежно влечет за собой $F(x) = hx$, причем h *означает собой величину постоянную*. На самом деле, заметим для начала, что поставленное условие должно быть удовлетворено в случае двух величин x_1 и x_2, и, следовательно, так как равенство $x_1 + x_2 = 0$ влечет за собой $F(x_1) + F(x_2) = 0$, а это означает, что $F(x)$ *нечетная*. Такие условия должны быть соблюдены также и для случая трех величин x_1, x_2, x_3, а если

$$-x_3 = x_1 + x_2,$$

то мы должны положить

$$F(x_1) + F(x_2) = -F(x_3),$$

а так как функция $F(x)$ нечетная, то

$$F(x_1 + x_2) = F(x_1) + F(x_2).$$

Дадим x_1 переменной значение x, а x_2 — значение постоянное, но произвольное α, и ради простоты предположим, что $F(x)$ имеет производную, тогда, беря эту производную по x, мы получим тождество:

$$F(x + \alpha) = F(x) + F(\alpha)$$

и равенство:

$$F'(x + \alpha) = F'(x),$$

которое должно удовлетворяться для всех значений x и α; следовательно, производная $F'(x)$ — величина постоянная и так же, как и $F(x)$, должна быть функцией нечетной. Поэтому она с очевидной необходимостью имеет вид $F(x) = hx$. Ясно, что функция, таким образом определенная, должна несомненно обладать вышеуказанным свойством, каким бы ни было n и какими бы ни были $x_1, x_2, \ldots, x_{n-1}$.

Таким образом, обозначая через x совершенную ошибку, мы получаем для элементарной вероятности $\varphi(x)dx$

$$Log\,\varphi = h\frac{x^2}{2} + \text{const},$$

откуда $$\varphi = Ae^{h\frac{x^2}{2}}.$$

Коэффициент h должен быть отрицательным для того, чтобы интеграл

$$J = \int_{-\infty}^{+\infty} \varphi(x)\, dx$$

имел смысл. Пусть $h = -2k^2$. Для того чтобы $J = 1$, нужно, чтобы

$$A = \frac{k}{\sqrt{\pi}}, \text{ откуда } \varphi = \frac{k}{\sqrt{\pi}} e^{-k^2x^2}.$$

Таково доказательство, данное Гауссом. Ж. Бертран выдвинул возражение против того, что φ, которое Гаусс предположил à priori, имеющим вид функции от $z_1 - z$, должно в действительности считаться как $\varphi(z, z_1)$; Бертран относится критически и к выбору средней как наиболее вероятной величины, а не как величины просто вероятной, что не одно и то же. Мы не станем подробно останавливаться ни на возражениях Бертрана, ни на аналогичном разборе А. Пуанкаре, считаясь с объемом настоящей книжки.

83. ИНВАРИАНТНОСТЬ ЗАКОНА ГАУССА

Можно à posteriori оправдать свойство инвариантности закона Гаусса, выражающееся в следующем.

Если ошибки измерений, произведенных над несколькими величинами, подчиняются закону Гаусса, то последний действителен также и для ошибок суммы этих измерений, если под нею понимать сумму значений измерений.

Это утверждение непосредственно следует из теоремы, доказанной в § 40. Если k — точность суммы измерений, а $k_1, k_2, \ldots$ — точности отдельных слагаемых этой суммы, то мы на основании теоремы о сложении случайных переменных имеем равенство

$$\frac{1}{k^2} = \frac{1}{k_1^2} + \frac{1}{k_2^2} + \ldots$$

84. ТОЧНОСТЬ СЕРИИ ИЗМЕРЕНИЙ

Здесь уместно показать конкретный смысл коэффициента k, названного нами «точностью». Вероятность для ошибки лежать в пределах

$$\frac{1}{k}\lambda_1 \quad \text{и} \quad \frac{1}{k}\lambda_2$$

равна

$$\frac{k}{\sqrt{\pi}}\int_{\lambda_1/k}^{\lambda_2/k} e^{-k^2x^2}\,dx=\frac{1}{\sqrt{\pi}}\int_{\lambda_1}^{\lambda_2} e^{-\lambda^2}\,d\lambda,$$

т. е. оказывается независящей от k. Таким образом, если одна система в десять раз точнее другой, то это значит, что вероятность ошибок в первой системе, заключенной в известных пределах, такая же, как вероятность ошибок для второй системы, заключенной в пределах в десять раз больших. Так, например, ошибки измерения порядка миллиметров в первой системе будут столь же вероятны, как ошибки порядка сантиметров во второй системе.

Число k очевидно дает меру точности в общеупотребительном смысле этого слова, подсказанном здравым смыслом.

Вероятная ошибка, по определению, представляет собой сумму произведений каждого возможного значения ошибки на ее вероятность, т. е. она равна математическому ожиданию ошибки.

Если принимать во внимание знаки ошибок, то вероятная ошибка равна нулю, ибо положительные ошибки столь же вероятны, как равновеликие отрицательные ошибки; если же брать ошибки по их абсолютной величине, то значение вероятной ошибки выразится интегралом:

$$\frac{2k}{\sqrt{\pi}}\int_{0}^{\infty} e^{-k^2x^2}\,x\,dx=\frac{1}{k\sqrt{\pi}}\,.$$

Таким образом, вычисленная ошибка, как и следует ожидать, оказывается тем меньшей, чем больше точность. Если мы имеем n наблюдений и обозначим их ошибки (взятые по их абсолютной величине) через e_1, e_2, ..., e_n, то мы получим приближенное значение вероятной ошибки, вычисляя среднюю величину ошибок

$$\frac{1}{k\sqrt{\pi}}=\frac{e_1+e_2+\ldots+e_n}{n},$$

что дает возможность практически удобно вычислить величину точности k[1].

---

[1] «Вероятная ошибка» как она определена авторами, отличается от того, что нередко встречается в нашей литературе, где под «вероятной ошибкой» подразумевается отклонение от истинной величины, делящее все ошибки на две равночисленные группы: не достигающие ее и превышающие ее, т. е. то, что авторы называют «медианной ошибкой». В терминах «перцентилей» Фр. Гальтона (Fr. Galton) «вероятная ошибка» соответствует квартилям.

Против неудачного выбора термина «вероятная ошибка» протестовал еще Ог. Курно (Aug. Cournot). См. его «Основы теории шансов и вероятностей», п. 34 стр. 77 русского перевода. «Наука», 1970; однако лишенный смысла термин и до сих пор удерживается в литературе. — *Прим. перев.*

Подобным же образом и вероятное или среднее значение квадрата ошибок (дисперсия) дается выражением

$$\frac{2k}{\sqrt{\pi}} \int_0^\infty e^{-k^2 z^2} z^2 \, dz = \frac{1}{2k^2}$$

и также приближенно

$$\frac{1}{2k^2} = \frac{e_1^2 + e_2^2 + \ldots + e_n^2}{n}.$$

Наконец, можно назвать *медианной ошибкой* ту, которая соответствует вероятности, равной $\frac{1}{2}$ того, что отдельные ошибки по своей абсолютной величине ее не превзойдут. Она соответствует медиане отклонений $\lambda = 0{,}4769$ и ее величина равна $\frac{0{,}4769}{k}$.

Ошибка, квадрат которой равнялся бы математическому ожиданию квадрата ошибок, равна

$$Q = \frac{1}{k\sqrt{2}} = \frac{0{,}707}{k}.$$

Вероятность для какой-либо ошибки оказаться меньшей по абсолютной величине, чем Q, равна

$$\Theta \frac{1}{\sqrt{2}} = \frac{2}{3} \text{ (приближенно)},$$

а оказаться меньшей, чем $2Q$:

$$\Theta \sqrt{2} = \frac{19}{20} \text{ (приближенно)}.$$

Когда на практике дается одно единственное измерение, сделанное с помощью инструмента, точность которого известна, то часто говорят о «возможной ошибке» (l'erreur à craindre), понимая под этим среднюю квадратическую ошибку, а также пользуются в качестве *«возможной ошибки»*[1] величиной $\frac{1}{k}$, для которой $\Theta(1) = \frac{17}{20}$ (приближенно) указывает вероятность не быть превзойденной по абсолютному значению; вероятность эта равна $Q\sqrt{2}$.

[1] «Возможная ошибка», принятая в нашей и в англо-американской литературе мера, обычно обозначается греческой буквой σ (синонимы: среднее квадратическое отклонение, standard deviation, écart-type); $\frac{1}{k}$ соответствует вышедшей из употребления мере «модуль» (В. Лексис — W. Lexis), равной $\sigma\sqrt{2} = 1{,}41\sigma$. (См. § 38 и 39 этой книги). — *Прим. перев.*

Весом измерения называют квадрат (k^2) *точности* (k). Такое наименование оправдывается решением следующей задачи.

Для определения длины a сделана первая серия измерений с точностью k_1 и вторая серия — с точностью k_2; предположим, что нам известен результат одного только измерения m_1 из первой серии и одного — m_2 — из второй серии. Какую оценку величины a следует принять?

Эта оценка должна удовлетворять следующим условиям.

1) Если умножить m_1 и m_2 на одну и ту же постоянную величину, то и сама оценка должна оказаться умноженной на ту же величину; она, следовательно, должна быть линейной и однородной функцией от m_1 и m_2.

2) Если к m_1 и к m_2 прибавить одну и ту же постоянную величину, то и оценка должна увеличиться на ту же величину.

Поэтому мы должны для оценки принять форму

$$m = \frac{h_1 m_1 + h_2 m_2}{h_1 + h_2}. \tag{172}$$

Обозначая через x_1 и x_2 ошибки $m_1 - a$ и $m_2 - a$, мы для ошибки $m - a$ получим:

$$x = \frac{h_1 x_1 + h_2 x_2}{h_1 + h_2}.$$

Пусть X_1, X_2, X — случайные переменные, соответствующие ошибкам x_1, x_2, и x. Мы знаем, что X_1 и X_2 следуют нормальному закону с математическим ожиданием ошибки, равным 0, и с соответственными дисперсиями:

$$\sigma_1^2 = \frac{1}{2k_1^2}, \quad \sigma_2^2 = \frac{1}{2k_2^2}.$$

Случайная переменная X подчиняется, следовательно, в согласии с § 40 нормальному закону с математическим ожиданием, равным нулю, и с дисперсиями:

$$\sigma^2 = \frac{1}{(h_1 + h_2)^2}\left(h_1^2 \sigma_1^2 + h_2^2 \sigma_2^2\right),$$

иначе говоря

$$\sigma^2 = \frac{1}{2(h_1 + h_2)^2}\left(\frac{h_1^2}{k_1^2} + \frac{h_2^2}{k_2^2}\right). \tag{173}$$

Наша цель состоит в том, чтобы получить наибольший шанс того, что ошибка относительно a будет возможно малой. Попытаемся же получить наименьшую величину для дисперсии σ^2, рассматривая ее в форме отношения

$$\omega = \frac{h_2}{h_1}.$$

Мы имеем

$$\sigma^2 = \frac{1}{2(1+\omega)^2}\left(\frac{1}{k_1^2} + \frac{\omega^2}{k_2^2}\right),$$

откуда следует, что

$$\frac{d\sigma^2}{d\omega} = \frac{\omega}{k_2^2} \cdot \frac{1}{(1+\omega)^2} - \frac{1}{(1+\omega)^3}\left[\frac{1}{k_1^2} + \frac{\omega^2}{k_2^2}\right].$$

Условие $\frac{d\sigma^2}{d\omega} = 0$ сразу приводит к результату

$$\omega = \frac{k_2^2}{k_1^2},$$

и искомая оценка равна, следовательно,

$$m = \frac{k_1^2 m_1 + k_2^2 m_2}{k_1^2 + k_2^2}. \tag{174}$$

Таким образом, мы видим, что m есть *взвешенная средняя* из результатов измерений m_1 и m_2, умноженных соответственно на *веса* k_1^2 и k_2^2, что и оправдывает определение, данное в начале § 85 понятию «веса наблюдения».

Так как несколько наблюдений принадлежат одной и той же серии и обладают одной и той же точностью, то это наводит на мысль брать среднюю арифметическую из них, кроме того, как мы уже указывали, правило средней арифметической послужило для Гаусса отправным началом для вывода его закона ошибок.

Беря проблему в общем виде, мы предположим n_1 наблюдений, сделанных с точностью k_1, средняя величина которых равна α_1; n_2 наблюдений с точностью k_2 и средней α_2; тогда значение, которое следует принять для z, должно быть:

$$z = \frac{n_1 k_1^2 \alpha_1 + \ldots + n_p k_p^2 \alpha_p}{n_1 k_1^2 + \ldots + n_p k_p^2},$$

так что вес для совокупности наблюдений получается умножением их обычного веса на число наблюдений. Отсюда заключаем, что *точность средней из многих наблюдений получается умножением точности отдельного наблюдения на квадратный корень из числа наблюдений.*

Таким образом, беря среднюю из 100 наблюдений, мы получаем результат в 10 раз более точный, чем тот, который дается одним наблюдением; беря среднюю из 10 000 измерений, мы повышаем точность в 100 раз по сравнению с точностью отдельного измерения. Мы не будем останавливаться на вопросе о пределе точности, достижимой практически путем увеличения числа измерений. Наибольшая трудность, возникающая в этом случае, состоит в том, чтобы обеспечить после значительного числа измерений нашу уверенность в том, что мы продолжаем измерять объект вполне тождественный с тем, с которым мы начали наши измерения.

86. ОБОСНОВАНИЕ ЗАКОНА ГАУССА

Какие бы ни были теоретические соображения относительно того, чтобы предложить экспоненциальный закон для ошибок наблюдений, все же его универсальная применимость обусловлена тем, что он превосходно согласуется с фактами. Всякий раз, когда мы встречаемся с многочисленными измерениями, выполненными в одинаковых условиях, общая совокупность результатов показывает не только превосходное согласие группировок ошибок с законом Гаусса, но и значения для k, которые можно на основании этого закона вывести из серии наблюдений различными способами, также очень хорошо согласуются между собой.

Мы видим, что средняя из абсолютных ошибок равна:

$$\frac{2k}{\sqrt{\pi}}\int_0^\infty e^{-k^2x^2}x\,dx=\frac{1}{k\sqrt{\pi}}$$

и что средняя из квадратов ошибок равна:

$$\frac{2k}{\sqrt{\pi}}\int_0^\infty e^{-k^2x^2}x^2\,dx=\frac{1}{2k^2}.$$

Также и средняя из абсолютных кубов ошибок выражается через

$$\frac{2k}{\sqrt{\pi}}\int_0^\infty e^{-k^2x^2}x^3\,dx=\frac{1}{k^3\sqrt{\pi}},$$

а для значения средней из четвертых степеней ошибок мы имеем

$$\frac{2k}{\sqrt{\pi}}\int_0^\infty e^{-k^2x^2}x^4\,dx=\frac{3}{4k^4}.$$

Если, как выше, обозначить через e_1, e_2, ..., e_n абсолютные значения ошибок наблюдений, то мы получим приближенно:

$$S_1=e_1+e_2+\ldots+e_n=\frac{n}{k\sqrt{\pi}},$$

$$S_2=e_1^2+e_2^2+\ldots+e_n^2=\frac{n}{2k^2},$$

$$S_3 = e_1^3 + e_2^3 + \ldots + e_n^3 = \frac{n}{k^3 \sqrt{\pi}},$$

$$S_4 = e_1^4 + e_2^4 + \ldots + e_n^4 = \frac{3n}{4k^4}.$$

Величины S_1, S_2, S_3, S_4 обосновывают, таким образом, ряд соотношений, подобных следующему:

$$S_1^2 = \frac{2n}{\pi} S_2.$$

Из подобных соотношений только три являются независимыми. «Этим своеобразным формулам,—говорит Бертран,—можно в такой мере довериться, что вычислитель, которому даны эти наблюдения и который находит, что эти соотношения не соблюдаются, вправе с полной уверенностью утверждать, что сообщенные ему данные — фальшивы и результаты наблюдений подвержены искажениям».

Наиболее известным примером, показывающим согласие между наблюдениями и законом ошибок, является сопоставление, сделанное Бесселем (Bessel) для серии из 470 наблюдений Брэдли (Bradley) для периода более чем сто лет. Числа Брэдли измеряют прямое восхождение и склонение одной определенной звезды. Что касается отклонений, то наибольшее отклонение измерений от общей средней достигает 4 дуговых секунд. Если распределить измерения по классам с интервалами в одну десятую этого наибольшего отклонения, то мы получим следующее процентное распределение, против которого выписаны соответствующие теоретические проценты, рассчитанные с помощью закона Гаусса.

Интервалы		Проценты наблюдений	Теоретические проценты
От 0″,0	до 0″,4	22,0	19,5
0″,4	0″,8	19,3	18,3
0″,8	1″,2	18,3	16,2
1″,2	1″,6	9,3	13,6
1″,6	2″,0	9,0	10,6
2″,0	2″,4	7,7	7,9
2″,4	2″,8	5,0	5,5
2″,8	3″,2	3,3	3,6
3″,2	3″,6	2,7	2,2
3″,6	4″,0	1,3	1,3
больше	4″,0	2,0	1,4

Многочисленные аналогичные исследования были произведены над самыми разнообразными сериями измерений: старыми и современными; и всюду закон Гаусса подтверждался весьма убедительно.

87. СЛУЧАЙНЫЕ ОТКЛОНЕНИЯ ТОЧКИ

Рассмотрим серию измерений, каждое из которых связано со значениями координат x и y какой-либо точки на плоскости. Точка, соответствующая истинному положению M_0, имеет своими координатами

x_0 и y_0; отклонения точки, соответствующей наблюдениям, от точки, соответствующей истинному положению, определяются двумя величинами:

$$x - x_0 = \xi, \quad y - y_0 = \eta.$$

Требуется выяснить закон отклонений в этом случае. Задача состоит в определении элементарной вероятности

$$\Phi(\xi, \ \eta)d\xi d\eta,$$

соответствующей значениям переменных ξ и η. Известно лишь то, что x и y суть функции определенных результатов измерений u и v, полученных непосредственно, и что каждое из них подчиняется закону Гаусса.

Величины u и v должны иметься в количестве не менее двух, без чего точка не может получить своего места (x, y) в двумерном пространстве, но число их может сколь угодно превышать (или быть равным) два.

Таким образом, проблема может рассматриваться как состоящая в определении точки попадания какого-либо снаряда и зависит от его веса, начальной скорости, угла прицела и т. д., следовательно, от элементов, подверженных мелким колебаниям, приводящим к определенным отклонениям.

Теоретическое исследование этой проблемы выполнил Бравэ (Bravais)[1], вычисления которого очень усложняются, если число переменных превышает два. Вопрос был снова подвергнут исследованию в более позднее время Гаагом (J. Haag) и рассмотрен более простым способом в более общей постановке. Укажем важнейшие результаты, не вдаваясь в подробности.

Элементарная вероятность имеет вид

$$Ae^{-(\alpha\xi^2 + 2\beta\xi\eta + \gamma\eta^2)}\, d\xi d\eta,$$

квадратичная форма $\alpha\xi^2 + 2\beta\xi\eta + \gamma\eta^2 = H$ всегда положительна, а постоянная A выбрана так, чтобы интеграл

$$A\int_{-\infty}^{+\infty} e^{-(\alpha\xi^2 + 2\beta\xi\eta + \gamma\eta^2)}\, d\xi d\eta,$$

распространенный на всю плоскость, равнялся единице. Кривая $H = \text{const}$ представляет собой гомотетические (т. е. подобные и подобно расположенные) эллипсы с центром в M_0; их называют *эллипсами равной вероятности* (ср. § 25).

---

[1] Исследование кристаллографа Бравэ было опубликовано в 1846 г. A. Bravais, Analyse mathématique sur les probabilités de situation d'un point. Memoires présentes par divers savants à l'Académie Royal des Sciences de l'Institut de France, 1846. Sciences mathématiques et physiques. Tome neuvième, pp. 255—332. — *Прим. перев.*

88. ЧАСТНАЯ ЗАДАЧА РАССЕЯНИЯ ПОПАДАНИЯ ПРИ СТРЕЛЬБЕ В ЦЕЛЬ

Если рассматривать задачу о рассеянии снарядов при их попадании на горизонтально расположенный участок или в вертикально поставленную мишень, то эллипсы равных вероятностей имеют оси, весьма близко расположенные к пересечению плоскости стрельбы с плокостью наблюдаемых результатов. Бертран приводит пример стрельбы по стоячей мишени из ручного оружия; значения α, β, γ таковы, что

$$\frac{\alpha}{568} = \frac{\beta}{33} = \frac{\gamma}{648},$$

что дает для угла i указанной оси с плоскостью стрельбы

$$\operatorname{tg} 2i = \frac{2\beta}{\alpha - \gamma} = -\frac{66}{80} \quad (i = 20^\circ \text{ приближенно})$$

Условия $i = 0$ на практике могут быть выполнены лишь приближенно; все зависит от того, в какой мере соблюдаются различные условия симметрии, которая здесь предполагается. Мы вправе принять, что i окажется действительно очень малым в случае установки артиллерийского орудия на симметричном лафете, как это и имеет место в современном вооружении.

В таком случае закон Бравэ принимает вид:

$$\frac{kk'}{\pi} e^{-(k^2\xi^2 + k'^2\eta^2)} d\xi d\eta,$$

так что отклонения в стороны и по дальности можно считать *независимыми* друг от друга. Каждая категория отклонений, взятая отдельно, подчиняется закону Гаусса.

Удобно и вместе с тем принято в артиллерии пользоваться преимущественно *медианным отклонением*, т. е. таким, для которого вероятность быть превзойденным составляет 50 шансов из 100. Мы знаем, что его значение получается из приведенного отклонения умножением на $\mu = 0{,}4769$.

Согласно таблице значений функции $\Theta(\mu)$, мы имеем:

$$\Theta(\mu) = \frac{1}{2}$$

$$\Theta(2\mu) = 0{,}82 \quad \text{(приближенно)}$$

$$\Theta(3\mu) = 0{,}96 \quad \text{»}$$

$$\Theta(4\mu) = 0{,}99 \quad \text{»}$$

Отсюда следует, что если мы остановимся, к примеру, на отклонениях по дальности и если разметим участок на восемь полос симметричных относительно линии фронта, соответствующей средней линии попаданий, и если ширину полос выберем равной медианному отклонению по дальности, то процент снарядов, которые лягут в пределах этих восьми полос, будут приближенно равны: 2, 7, 16, 25, 25, 16, 7, 2.

Эти числа хорошо известны всем артиллеристам.

6. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ ТОЧНОСТЬ РЕЗУЛЬТАТОВ

89. ОСНОВА МЕТОДА

Как мы уже указывали в § 78, основная проблема с практической точки зрения состоит в следующем: зная результаты серии измерений, можно ли из них вывести значение (или значения) величины (или величин), которые измерялись?

Ясно, что эта проблема по своему существу не может быть решена с полной точностью. Все, что можно сделать,—это определить правила для оценки каждой из измеряемых величин и связать с этой оценкой некоторую зону, внутри которой с определенной вероятностью содержится точное значение неизвестной нам величины.

Представим себе очень простой случай, когда мы имеем n непосредственных измерений, давших значения

$$x_1, x_2, \dots, x_n$$

для одной и той же неизвестной нам величины θ. Допустим, что соответственные им случайные ошибки следуют закону Гаусса с коэффициентом точности, равным k. Каждое значение x_i представлет собой одно из возможных значений случайной переменной X_i, для которой можно написать вероятность неравенств:

$$Pr\,[x_i < X_i < x_i + dx_i] = \frac{k}{\sqrt{\pi}} e^{-k^2(\theta - x_i)^2}\, dx_i.$$

Так как измерения взаимно независимы, то вероятность получить выборку численностью n, удовлетворяющую условиям

$$x_i < X_i < x_i + dx_i \quad (i = 1, 2, \dots, n),$$

равна

(175) $$dP = Ae^{-k^2\sum(\theta-x_i)^2}dx_1dx_2\ldots dx_n,$$

где A — некоторая постоянная.

Если мы применим *правило наибольшего правдоподобия*, то для θ мы должны взять значение, при котором e^{-S} получает свое наибольшее значение, где S означает:

$$S = k^2\sum(\theta - x_i)^2;$$

величина θ определяется из условия

$$\frac{\partial S}{\partial\theta} = 2k^2\sum(\theta - x_i) = 0,$$

а это приводит к тому, что

$$\theta' = \frac{1}{n}(x_1 + x_2 + \ldots + x_n).$$

Мы вновь получаем среднюю арифметическую; это значение θ' величины θ является тем, при котором данная выборка оказывается наиболее вероятной среди других возможных выборок такого же объема n.

Если измерения имели различные веса

$$k_1^2,\ k_2^2,\ \ldots,\ k_n^2,$$

то мы имели бы

$$S = \sum k_i^2(\theta - x_i)^2$$

и

$$\theta' = \frac{\sum k_i^2 x_i}{\sum k_i^2}.$$

Этот анализ, несмотря на свою элементарность, уже содержит в себе основу метода наименьших квадратов в его общем виде.

Рассмотрим теперь следующую задачу. *Даны n измерений, непосредственными результатами которых являются величины y_1, y_2, ..., y_n, представляющие собой функции x-ов.*

(176) $$\begin{cases} y_1 = f_1(x_1, x_2, \ldots, x_m) \\ y_2 = f_2(x_1, x_2, \ldots, x_m) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ y_n = f_n(x_1, x_2, \ldots, x_m) \end{cases}$$

причем $n > m$; какие значения было бы целесообразно принять для x_1, x_2, ..., x_n?

Обозначим через η_1, η_2, ..., η_n точные значения y_i; а через θ_1, θ_2, ..., θ_m точные значения для x_i; равенства (176) представляют собой твердые (не случайные) соотношения и мы поэтому имеем:

$$(177)\qquad \begin{cases} \eta_1 = f_1(\theta_1,\ \theta_2,\ \ldots,\ \theta_m) \\ \eta_2 = f_2(\theta_1,\ \theta_2,\ \ldots,\ \theta_m) \\ \ldots\ldots\ldots\ldots \\ \eta_n = f_n(\theta_1,\ \theta_2,\ \ldots,\ \theta_m) \end{cases}$$

С другой стороны, мы принимаем, что ошибки

$$z_i = y_i - \eta_i \quad (i = 1,\ 2,\ \ldots,\ n),$$

совершенные при непосредственных измерениях $y_1,\ y_2,\ \ldots,\ y_n$, следуют закону Гаусса. Поэтому эти ошибки являются значениями случайных переменных $Z_1,\ Z_2,\ \ldots,\ Z_n$, плотности вероятностей для которых при $i = 1,\ 2,\ \ldots,\ n$ соответственно равны:

$$\frac{k_i}{\sqrt{\pi}} e^{-k_i^2 [f_i(\theta_1, \theta_2,\ \ldots,\ \theta_m) - y_i]^2}.$$

Элементарная вероятность выборки, тождественно одинаковой с выборкой $(y_1,\ y_2,\ \ldots,\ y_n)$, пропорциональна, следовательно, e^{-S}, где на этот раз S означает:

$$(178)\qquad S = \sum k_i^2 [f_i(\theta_1,\ \theta_2, \ldots, \theta_m) - y_i]^2.$$

Оценки $\theta_1',\ \theta_2',\ \ldots,\ \theta_m'$, к которым приводит метод наибольшего правдоподобия, даются поэтому решением системы уравнений:

$$(179)\qquad \frac{\partial S}{\partial x_1} = 0,\ \frac{\partial S}{\partial x_2} = 0,\ \ldots,\ \frac{\partial S}{\partial x_m} = 0.$$

Число уравнений системы (179) равно числу неизвестных. Таким образом, система значений, которые мы принимаем для x_i, — это те значения, для которых сумма квадратов непосредственных ошибок наблюдений, умноженных на соответственные им веса, получают наименьшее значение. Отсюда и наименование: *метод наименьших квадратов.*

Заметим, что если измерения не искажены никакими ошибками, то точные значения (177) для x_k превращают S *в нуль.* Поэтому метод, придающий S наименьшее значение, представляется вполне естественным.

90. ОБЩАЯ ПРОБЛЕМА ИНТЕРПОЛЯЦИИ

Укажем, в какого рода исследованиях нам чаще всего приходится применять метод наименьших квадратов.

Рассмотрим серию физических измерений, назначение которых состоит в том, чтобы выявить закон зависимости результатов измерения явления Y от изменений, происходящих с явлением X. Искомый закон проявляется на m результатах измерения $Y_1,\ Y_2,\ \ldots,\ Y_m$, соответствующих значениям $X_1,\ X_2,\ \ldots,\ X_m$ переменной X. Если в системе прямоугольных координат OX, OY отметить точки P_1, P_2, ..., P_m с координатами X_1Y_1, X_2Y_2, ..., X_mY_m, то кривая (γ) с уравнением

$$Y = f(X),$$

изображающая изучаемое явление, должна проходить через все точки P, поскольку измерения не искажены никакими ошибками.

Остановимся на минуту на этом теоретическом допущении; задача аналитического определения вида функции $f(X)$ содержит в себе немалую долю произвола, поскольку от неизвестной функции требуется только, чтобы она принимала при абсциссах $X = X_1, X_2, \ldots, X_m$ значения ординат Y, равные $Y_1, Y_2, \ldots, Y_m$. Эта проблема является *общей проблемой интерполяции*, постоянно возникающей в экспериментальных науках.

Неопределенность исчезает, если решиться выбрать функцию $f(X)$ из числа функций, аналитическое выражение которых принадлежит к определенному удобному типу; так, например, если для $f(X)$ взять форму полинома

$$a_0 + a_1X + \ldots + a_{m-1}X^{m-1}$$

от X степени $m - 1$; коэффициенты $a_0, a_1, \ldots, a_{m-1}$ являются неизвестными в системе из $m - 1$ линейных уравнений, относительно которых можно утверждать, что они дают однозначное решение, если X_i все различны между собой.

Так ограниченная задача является задачей *параболической интерполяции*; форма, выбранная для функции $f(X)$, представляется наиболее естественной, ибо если принять, что физическое явление изменяется непрерывно, то возможно $f(X)$ разложить в ряд Маклорена (Mac Laurin), который на практике обрывают на целесообразно выбранном члене ряда, а именно так, примерно, и происходит, если хотят представить изменение длины l металлического стержня, как функцию температуры t. В этом простом случае в качестве первого приближения пользуются «биномом растяжения»

$$l = \alpha_0 + \alpha_1 t.$$

Если такое приближение признается достаточным, то коэффициент считают постоянным; кривой (γ) соответствует в этом случае прямая, полностью определяемая двумя измерениями. Если исследование провести более тщательно, то мы убеждаемся в том, что третья точка P_3, полученная в результате третьего измерения, не ложится в точности на прямую P_1P_2, тогда прибегают к трехчлену

$$l = a_0 + a_1 t + a_2 t^2,$$

которому соответствует парабола, полностью определяемая тремя точками P_1, P_2, P_3 и которая очень точно воспроизводит явление[1].

[1] К этим соображениям надо, однако, относиться с осторожностью. Основанные на формуле Маклорена, они могут перестать быть справедливыми при слишком больших интервалах, принятых для изменений t. Такая простая кривая, как

На практике, когда выбрано надлежащее число членов, в нашем распоряжении для определения постоянных оказывается не соответствующее их количеству число точек, полученных точными измерениями, а *избыточное количество измерений, искаженных случайными ошибками*. И теперь дело уже не в том, чтобы провести кривую (γ) параболу m — 1 порядка через m точек, а надо заставить ее пройти между гораздо бо́льшим числом точек. Уравнения, содержащие a_0, a_1, ..., a_{m-1}, оказываются сходными с типом уравнений (176), и значениями этих коэффициентов, которые удовлетворяют требованию «наибольшего правдоподобия», являются значения, придающие наименьшую величину сумме

$$S=\sum k_i^2 (a_0+a_1 X_i+\ldots+a_{m-1} X_i^{m-1}-Y_i)^2.$$

91. ЧАСТНЫЙ СЛУЧАЙ. ПРЯМАЯ РЕГРЕССИИ

Предположим, что «скопление» точек $M_i(X_i, Y_i)$ подсказывает, что в качестве кривой, характеризующей изучаемое явление, следует выбрать прямую (D) с уравнением

$$Y = a + bX.$$

Предположим для простоты, что все измерения Y имеют одинаковый вес, тогда следует определять значения a и b так, чтобы

$$S=\sum (Y_i-bX_i-a)^2$$

получила свое наименьшее значение x. Следствием этого будет то, что прямая (D), полученная при условии минимума для S, пройдет через *центральную* (*среднюю*) *точку скопления точек* с координатами:

$$X_0=\frac{1}{n}\sum X_i, \quad Y_0=\frac{1}{n}\sum Y_i. \tag{180}$$

Уравнения, решающие эту задачу: $\frac{\partial S}{\partial a}=0$; $\frac{\partial S}{\partial b}=0$. Первое из них в данном случае имеет вид:

$$\sum (Y_i-bX_i-a)=\sum Y_i-b\sum X_i-na=0.$$

Теперь уравнение прямой (D) мы можем написать в таком виде:

$$Y-Y_0=b(X-X_0),$$

$y=\frac{1}{1+x^2}$, представленная в форме полинома, с возрастающими степенями, начиная с равноудаленных точек, например в интервале (—5, +5), может и в действительности приводит в этом частном случае к *расходящемуся* ряду, так что интерполирующие полиномы по мере возрастания степеней все дальше и дальше отходят от функции.

См. Ém. Borel, «Leçons sur les fonctions de variables réelles» под редакцией Maurice Fréchet (Gauthier Villars), а также Montel. Leçons sur les séries de polynômes (Gauthier Villars).

а S будет выражено как функция одного лишь параметра b в виде

$$S=\sum [Y_i-Y_0-b(X_i-X_0)]^2.$$

Минимум для S получается из уравнения $\frac{\partial S}{\partial b}=0$, т. е.

$$\sum (X_i-X_0)[Y_i-Y_0-b(X_i-X_0)]=0,$$

которым определяется значение b. Мы получим

$$b=\frac{\Sigma (X_i-X_0)(Y_i-Y_0)}{\Sigma (X_i-X_0)^2}. \tag{181}$$

Если предположить, что значения $X_1, X_2, \ldots, X_n$ определены точно, а значения $Y_1, Y_2, \ldots, Y_n$ суть значения случайной переменной, то прямая (D) называется *прямой регрессии* Y по X.

Численный пример. Конкретное вычисление b облегчается в силу следующих двух замечаний:

во-первых, знаменатель в (181) может быть написан в таком виде:

$$\sum (X_i-X_0)^2=\sum X_i^2-2X_0\sum X_i+nX_0^2,$$

или, принимая во внимание (180),

$$\sum (X_i-X_0)^2=\sum X_i^2-\frac{1}{n}(\sum X_i)^2;$$

и числителю можно также сразу же придать форму

$$\sum (X_i-X_0)(Y_i-Y_0)=\sum X_iY_i-Y_0\sum X_i-X_0\sum Y_i+X_0Y_0,$$

что, согласно (180) дает:

$$\sum (X_i-X_0)(Y_i-Y_0)=\sum X_iY_i-\frac{1}{n}\sum X_i\sum Y_i,$$

Предположим, что мы, экспериментируя, получили такие результаты:

X_i:	0	3	10	19	30	41	63	112
Y_i:	9	12	18	28	42	57	85	132

Если мы эти данные нанесем на график, то увидим, что можно в качестве уравнения регрессии взять прямую (D), проведенную сквозь «скопление» точек, определенных указанными измерениями — координатами. Читатель без труда вычислит следующие значения:

$$\sum X_i=278,\quad \sum Y_i=383, \sum X_iY_i=24\,484, \sum X_i^2=19\,564,$$
$$\sum Y_i^2=30\,995.$$

По ним найдем координаты центральной точки

$$X_0=34{,}75,\quad Y_0=47{,}88,$$

а уравнение регрессии напишется в виде:

$$Y=8{,}68+1{,}13X,$$

В случае, когда для серии наблюдений подбирается по методу наименьших квадратов параболический закон

$$Y = a_0 + a_1 X + a_2 X^2 + \ldots + a_{m-1} X^{m-1},$$

то уравнения (179), служащие для определения коэффициентов, получаются первой степени относительно этих коэффициентов.

В общем случае это условие, очевидно, не осуществляется, но можно этот общий случай легко привести к решению *системы линейных уравнений* при помощи удобных приближений, и к этому на практике постоянно и прибегают.

Положим, что требуется привести к минимуму выражение

$$S = \sum k_i^2 (f_i - y_i)^2,$$

где под f_i подразумевается определенная функция от $x_1, x_2, \ldots, x_m$. Всегда возможно поступить хотя бы так: из системы (176) выбрать произвольно m уравнений и решить их точно относительно $x_1, x_2, \ldots, x_m$, найдя значения для них

$$\xi_1, \xi_2, \ldots, \xi_m,$$

которые будут приближенными относительно требуемых значений. Полагая затем

$$\begin{cases} x_1 = \xi_1 + \alpha_1, \\ x_2 = \xi_2 + \alpha_2, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ x_m = \xi_m + \alpha_m, \end{cases}$$

где α весьма малы, можно ограничиться в разложениях f_i членами, содержащими α в первых степенях.

Если положить

$$\eta_i = f_i(\xi_1, \xi_2, \ldots, \xi_m),$$

то будем иметь:

$$f_i(x_1, x_2, \ldots, x_m) = \eta_i + A_i^1 \alpha_1 + A_i^2 \alpha_2 + \ldots + A_i^m \alpha_m,$$

следовательно,

$$S = \sum k_i^2 \left(A_i^2 x_1 + A_i^2 x_2 + \ldots + A_i^m x_m + \eta_i - y_i\right)^2.$$

Поэтому уравнения (179) получают следующий вид:

$$(182) \qquad \begin{cases} \dfrac{1}{2} \dfrac{\partial S}{\partial \alpha_1} = \sum\limits_i A_i^1 k_i^2 \left[\sum\limits_j A_i^j \alpha_j + \eta_i - y_i\right] = 0, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \dfrac{1}{2} \dfrac{\partial S}{\partial \alpha_m} = \sum\limits_i A_i^m k_i^2 \left[\sum\limits_j A_i^j \alpha_j + \eta_i - y_i\right] = 0. \end{cases}$$

Эти уравнения легко выводятся из одного выражения для S. Мы вернемся к этому вопросу с более простыми обозначениями и с числом неизвестных, равным трем.

93. НОРМАЛЬНЫЕ УРАВНЕНИЯ

Согласно методу наименьших квадратов надо решать систему следующих n уравнений:

$$(183)\qquad \begin{cases} ax+by+cz+l=0, \\ a'x+b'y+c'z+l'=0. \\ \ldots\ldots\ldots \end{cases}$$

Дело идет при этом о том, чтобы сделать минимальной сумму S равную

$$(ax+by+cz+l)^2+\ldots,$$

в которой все члены по отдельности равнялись бы нулю, если бы система (183) решалась точно.

Уравнения

$$\frac{1}{2}\frac{\partial S}{\partial x}=0,\quad \frac{1}{2}\frac{\partial S}{\partial y}=0,\quad \frac{1}{2}\frac{\partial S}{\partial z}=0$$

в развернутом виде пишутся:

$$(184)\qquad \begin{cases} (aa)x+(ab)y+(ac)z+(al)=0 \\ (ba)x+(bb)y+(bc)z+(bl)=0 \\ (ca)x+(cb)y+(cc)z+(cl)=0 \end{cases}$$

где

$$(aa)=a^2+a'^2+a''^2+\ldots.$$
$$(ab)=ab+a'b'+a''b''+\ldots$$

Три уравнения системы (184) называются *нормальными уравнениями* системы (183); для того чтобы решить систему (183) методом наименьших квадратов, надо построить, а затем и решить систему нормальных уравнений.

Из первого уравнения системы (184) получаем.

$$(185)\qquad x=-\frac{(ab)}{(aa)}y-\frac{(ac)}{(aa)}z-\frac{(al)}{(aa)}.$$

Внесем это значение x в два остальных уравнения, введя символы предложенные Гауссом

$$(bb\cdot 1)=(bb)-\frac{(ab)}{(aa)}(ab)$$

$$(bc\cdot 1)=(bc)-\frac{(ac)}{(aa)}(ab)$$

$$(bl\cdot 1)=(bl)-\frac{(al)}{(aa)}(ab) \text{ и т д.}$$

Теперь система для определения y и z напишется так:

$$(186)\qquad \begin{aligned}(bb\cdot 1)\,y+(bc\cdot 1)\,z+(bl\cdot 1)=0\\ (bc\cdot 1)\,y+(cc\cdot 1)\,z+(cl\cdot 1)=0.\end{aligned}$$

Из первого уравнения (186) находим:

$$(187)\qquad y=-\frac{(bc\cdot 1)}{(bb\cdot 1)}z-\frac{(bl\cdot 1)}{(bb\cdot 1)}$$

и, внеся это выражение для y во второе уравнение системы (186), получим:

$$\left[(cc\cdot 1)-\frac{(bc\cdot 1)}{(bb\cdot 1)}(bc\cdot 1)\right]z+(cl\cdot 1)-\frac{(bl\cdot 1)}{(bb\cdot 1)}(bc\cdot 1)=0,$$

что Гаусс в своей символике пишет так

$$(cc\cdot 2)\,z+(cl\cdot 2)=0.$$

Уравнение (188) дает значение для z, уравнение (187)— для y, а из (185) получим тогда решение и для x.

94. РАСЧЕТ ЧИСЛОВОГО ПРИМЕРА

Следующий числовой пример мы заимствуем из сборника Грюи (Gruey «Упражнения по астрономии»). Вычисления были упрощены Ж. Шюардом (Jules Chuard), за что мы выражаем ему благодарность.

Требуется решить по методу наименьших квадратов систему из следующих 9 уравнений, в которых постоянные члены — результат непосредственных измерений.

$$\begin{aligned}
3{,}6\,x+2{,}5\,y-4{,}2\,z-5{,}2&=0\\
4{,}8\,x-1{,}7\,y+5{,}3\,z+2{,}1&=0\\
6{,}4\,x+0{,}9\,y+3{,}6\,z-0{,}8&=0\\
1{,}7\,x-6{,}4\,y+8{,}1\,z+7{,}9&=0\\
3{,}6\,x-3{,}2\,y+5{,}9\,z+3{,}7&=0\\
4{,}9\,x+1{,}8\,y-2{,}7\,z-4{,}4&=0\\
5{,}6\,x+3{,}9\,y-4{,}2\,z-6{,}7&=0\\
4{,}0\,x+5{,}2\,y-6{,}7\,z-8{,}2&=0\\
2{,}5\,x+7{,}6\,y-3{,}2\,z-6{,}7&=0
\end{aligned}$$

Чтобы составить «уравнение для x», нужно произвести 36 умножений и 4 сложения каждых 9 произведений; тогда мы получим:

$$170{,}43x+54{,}66y-3{,}18z-96{,}66=0.$$

Что касается «уравнений для y», то нам уже известен коэффициент при x в этом уравнении, равный 54,66; о нем нам не надо уже заботиться, и остается сделать 27 умножений и 3 сложения по 9 слагаемых в каждом. Тогда мы получим:

$$54{,}66x + 164{,}40y - 167{,}39z - 203{,}46 = 0.$$

Теперь мы уже знаем в «уравнении для z» коэффициенты при x и при y; их значения соответственно равны: $-3{,}18$ и $-167{,}39$. Остается сделать еще 18 умножений и 2 сложения по 9 слагаемых в каждом. Получим уравнение:

$$-3{,}18x - 167{,}39y + 239{,}17z + 227{,}45 = 0.$$

Три нормальных уравнения, если их решать по методике Крамера (Cramer), дают решения:

$$x = \frac{\Delta x}{\Delta} = \frac{502\,887{,}04}{1\,267\,838{,}80} = 0{,}3966$$

$$y = \frac{\Delta y}{\Delta} = \frac{629\,917{,}95}{1\,267\,838{,}80} = 0{,}495,$$

$$z = \frac{\Delta z}{\Delta} = \frac{-758\,158{,}58}{1\,267\,838{,}80} = -0{,}5980.$$

II. ТОЧНОСТЬ РЕЗУЛЬТАТОВ

95. О РАЗЛИЧНЫХ ВЫРАЖЕНИЯХ ДЛЯ ЭЛЕМЕНТА ОБЪЕМА В ЭВКЛИДОВОМ ПРОСТРАНСТВЕ n ИЗМЕРЕНИЙ

Прежде чем браться за проблему точности, на которую можно рассчитывать для результатов, получаемых по методу наименьших квадратов, мы укажем на различные выражения для элемента объема, которыми мы воспользуемся для рассмотрения проблемы точности.

Мы имеем в виду элемент объема

$$dV = dx_1\, dx_2 \ldots dx_n$$

в n-мерном эвклидовом пространстве ($\mathscr{E}$).

Мы предполагаем, что читатель знаком с основоположной теоремой, относящейся к исчислению кратных интегралов путем замены переменных и согласно которой, если положить

$$x_1 = f_1(u_1, u_2, \ldots, u_n)$$
$$\cdots\cdots\cdots$$
$$x_n = f_n(u_1, u_2, \ldots, u_n),$$

элемент объема dV выражается в системе новых переменных по формуле:

$$dV = |\Delta|\, du_1\, du_2 \ldots du_n,$$

в которой Δ означает абсолютную величину якобиана:

$$\Delta = \begin{vmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \dots & \frac{\partial x_1}{\partial u_n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \dots & \frac{\partial x_n}{\partial u_n} \end{vmatrix} \tag{188}$$

Выражение dV в полярных координатах

Положим для $0 < \rho < +\infty$; $0 < \varphi_1 < \pi$, ...; $0 < \varphi_{n-2} < \pi$; $0 < \varphi_{n-1} < 2\pi$

$$\begin{aligned} x_1 &= \rho \cos \varphi_1, \\ x_2 &= \rho \sin \varphi_1 \cos \varphi_2, \\ x_3 &= \rho \sin \varphi_1 \sin \varphi_2 \cos \varphi_3, \\ &\dots\dots\dots\dots\dots \\ x_{n-1} &= \rho \sin \varphi_1 \sin \varphi_2 \dots \sin \varphi_{n-2} \cos \varphi_{n-1}, \\ x_n &= \rho \sin \varphi_1 \sin \varphi_2 \dots \sin \varphi_{n-1}. \end{aligned} \tag{189}$$

Якобиан напишется так:

$$\Delta = \begin{vmatrix} \frac{x_1}{\rho} & -x_1 \operatorname{tg} \varphi_1 & 0 & \dots & 0 \\ \frac{x_2}{\rho} & x_2 \operatorname{ctg} \varphi_1 & -x_2 \operatorname{tg} \varphi_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{x_n}{\rho} & x_n \operatorname{ctg} \varphi_1 & x_n \operatorname{ctg} \varphi_2 & \dots & -x_n \operatorname{ctg} \varphi_{n-1} \end{vmatrix}.$$

Его, таким образом, можно представить в виде произведения $\frac{x_1, x_2, \dots, x_n}{\rho}$ на детерминант, который легко вычислить, если к элементам второй колонки прибавить элементы первой, умноженные на $\operatorname{tg} \varphi_1$, а к элементам третьей колонки — элементы первой, умноженные на $\operatorname{tg} \varphi_2$, и т. д. Если учесть, что $\operatorname{tg} \varphi + \operatorname{cotg} \varphi = \frac{1}{\sin \varphi \cos \varphi}$, то мы получим

$$\Delta = \frac{x_1 x_2 \dots x_n}{\rho (\sin \varphi_1 \cos \varphi_1)(\sin \varphi_2 \cos \varphi_2) \dots (\sin \varphi_{n-1} \cos \varphi_{n-1})},$$

или иначе

$$\Delta = \rho^{n-1} \sin^{n-2} \varphi_1 \sin^{n-3} \varphi_2 \dots \sin \varphi_{n-2}.$$

Отсюда следует, что элемент объема в полярных координатах будет иметь такой вид:

$$dV = \rho^{n-1} d\rho \left(\sin^{n-2} \varphi_1 \sin^{n-3} \varphi_2 \dots \sin \varphi_{n-2}\right).$$

Множитель в скобках измеряет *элементарный телесный угол* $d\Omega$ и мы имеем окончательно

$$dV = \rho^{n-1}\, d\rho\, d\Omega. \tag{190}$$

Выражение для dV в полуполярных координатах. Предположим, что систему координат $(x_1, x_2, \ldots, x_n)$ точки M в пространстве $(\mathscr{E})$ мы разбиваем на две частные системы: в одну из них войдут первые m координат, а в другую — остальные $n - m = p$ координат. Если для $k = 1, 2, \ldots, p$ положить

$$x_{m+k} = y_k,$$

то точка M будет иметь в качестве своих координат

$$(x_1, x_2, \ldots, x_m), \quad (y_1, y_2, \ldots, y_p).$$

Введем для каждой из частных систем новые переменные

$$\lambda,\ \varphi_1,\ \varphi_2, \ldots, \varphi_{m-1} \quad \text{и} \quad \mu,\ \psi_1,\ \psi_2, \ldots, \psi_{p-1},$$

которые, удовлетворяя условиям $0 < \lambda < +\infty$, $0 < \varphi_1 < \pi, \ldots,$ $0 < \varphi_{n-1} < 2\pi$, находятся с прежними переменными в соотношениях:

$$\begin{aligned} x_1 &= \lambda \cos\varphi_1, \\ x_2 &= \lambda \sin\varphi_1 \cos\varphi_2, \\ &\cdots\cdots\cdots \\ x_{m-1} &= \lambda \sin\varphi_1 \sin\varphi_2 \ldots \sin\varphi_{m-2} \cos\varphi_{m-1}, \\ x_m &= \lambda \sin\varphi_1 \sin\varphi_2 \ldots \sin\varphi_{m-2} \sin\varphi_{m-1}, \end{aligned} \tag{191}$$

и таким же образом, полагая $0 < \mu < +\infty$; $0 < \psi_1 < \pi$;

$$0 < \psi_2 < \pi, \ldots, 0 < \psi_{p-1} < 2\pi,$$

причем новые переменные связываются с прежними соотношениями:

$$\begin{aligned} y_1 &= \mu \cos\psi_1, \\ y_2 &= \mu \sin\psi_1 \cos\psi_2, \\ &\cdots\cdots\cdots \\ y_p &= \mu \sin\psi_1 \sin\psi_2 \ldots \sin\psi_{p-1}. \end{aligned} \tag{192}$$

Систему $\lambda, \varphi_1, \varphi_2, \ldots, \varphi_{m-1}$ и $\mu, \psi_1, \psi_2, \ldots, \psi_{p-1}$ называют системой *полуполярных координат* точки M.

Якобиан переменных $x_1, x_2, \ldots, x_m$; $y_1, y_2, \ldots, y_p$, взятый относительно новых переменных, введенных выше, имеет вид

$$\Delta = \left| \begin{array}{c|c} \Delta_1 & 0 \\ \hline 0 & \Delta_2 \end{array} \right|,$$

т. е. представляет собой произведение якобианов, взятых по отдельности при заменах переменных (191) и (192). Основная теорема приводит нас, следовательно в согласии с формулой (190), к тому, чтобы элемент объема представить в виде:

$$dV = \lambda^{m-1}\, d\lambda\, d\Omega_1\, \mu^{p-1}\, d\mu\, d\Omega_2. \tag{193}$$

Здесь $d\Omega_1$ и $d\Omega_2$ означают элементарные телесные углы.

$$d\Omega_1 = \sin^{m-1}\varphi_1 \sin \ldots \varphi_{m-2}\, d\varphi_1 \ldots d\varphi_{m-2},$$

$$d\Omega_2 = \sin^{p-1}\psi_1 \ldots \sin\psi_{p-2}\, d\psi_1 \ldots d\psi_{p-2}.$$

Выражения (190) и (193) являются обобщением хорошо известных обычных выражений

$$dV = \rho^2 \sin\theta\, d\rho\, d\theta\, d\varphi, \quad dV = r\, dr\, d\varphi\, dz$$

для элемента объема в обычном пространстве, выраженным в полярных и цилиндрических координатах.

96. ГЕОМЕТРИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ

Вернемся теперь к общей проблеме непрямого определения некоторого числа m неизвестных параметров, пользуясь результатами более многочисленных непосредственных наблюдений, произведенных над величинами, связанными с неизвестными параметрами данными точными соотношениями.

Припоминая сказанное в § 88, мы предполагаем, что эти последние соотношения приведены к *линейной* форме. Пусть же мы имеем систему связей:

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 + \ldots + a_{1m} x_m &= y_1, \\ a_{21} x_1 + a_{22} x_2 + \ldots + a_{2m} x_m &= y_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1} x_1 + a_{n2} x_2 + \ldots + a_{nm} x_m &= y_n. \end{aligned} \tag{194}$$

В этой системе мы предполагаем, что $n > m$ и что в ней y_i представляет собой результаты измерений величин, обозначенных нами как

$$\eta_1, \eta_2, \ldots, \eta_n,$$

величин точных, но нам неизвестных, измерения которых искажены ошибками

$$\zeta_i = y_i - \eta_i \quad (i = 1, 2, \ldots, n).$$

Предполагается, что все ошибки следуют одинаковому для всех закону Гаусса с элементарной вероятностью

$$\frac{k}{\sqrt{\pi}}\, e^{-k^2\zeta^2}\, d\zeta,$$

точность k, одинаковая для всех измерений, нам неизвестна, Пусть $\theta_1, \theta_2, \ldots, \theta_m$ означают точные значения неизвестных нам $x_1, x_2, \ldots, x_m$; их величины связаны с величинами $\eta_1, \eta_2, \ldots, \eta_m$ равенствами

$$a_{i1}\theta_1 + a_{i2}\theta_2 + \ldots + a_{im}\theta_m = \eta_i \quad (i = 1, 2, \ldots, n).$$

С другой стороны, пусть $\theta_1', \theta_2', \ldots, \theta_m'$ — произведенные нами оценки $\theta_1, \theta_2, \ldots, \theta_m$ по методу наименьших квадратов.

Обычно принимают (для i — 1, 2, ..., n),

$$u_i = y_i - a_{i1}\theta_1' - a_{i2}\theta_2' - \ldots - a_{im}\theta_m'$$

и эти u_i называются *погрешностями* в наблюдениях, результатами которых являются y_i. Возникает такая задача: *как на основании погрешностей u_i оценить точность k?*

Для решения этой задачи мы используем геометрические соображения, изложенные в § 91.

Рассмотрим в n-мерном эвклидовом пространстве (E_n) точку M, определяемую прямоугольными координатами, которыми служат результаты измерений $y_1, y_2, \ldots, y_n$, эта точка носит название точки выборки. Через ω обозначим точку в (E_n), координатами которой служат точные значения $\eta_1, \eta_2, \ldots, \eta_n$ измеренных величин: ω называется *постоянной точкой* (*point fixe*).

Теперь обратимся к точке Q, координаты которой определяются выражениями:

$$(195) \qquad \begin{aligned} q_1 &= a_{11}x_1 + \ldots + a_{1m}x_m, \\ &\ldots\ldots\ldots \\ q_n &= a_{n1}x_1 + \ldots + a_{nm}x_m, \end{aligned}$$

в которых x_i рассматриваются как *произвольные параметры*.

Когда эти параметры принимают всевозможные для них значения, то точка Q пробегает в (E_n) некоторое *линейное множество* (π), расположенное в m измерениях. К этому множеству принадлежит и точка ω, когда $x_1 = \theta_1, \ldots, x_m = \theta_m$.

Изложенный в первой части этой главы метод наименьших квадратов сводится к выбору таких значений $\theta_1', \theta_2', \ldots, \theta_m'$ для x_i, которые *придают наименьшее значение величине*

$$\overline{MQ}^2 = \sum (q_i - y_i)^2,$$

т. е. *квадрату расстояния между точкой выборки M и точкой Q множества* (π).

Пусть H то особое значение Q, при котором осуществляется этот минимум; можно, просто по интуиции, утверждать, что *эта точка является основанием для перпендикуляра, опущенного из точки M на множество* (π). На самом деле, нетрудно доказать, что вектор $\overrightarrow{HM}$ ортогонален по отношению ко всем векторам $\overrightarrow{HQ}$, конечными точками которых служат произвольно взятые точки из множества (π). Ибо если через $y_1', y_2', \ldots, y_n'$ мы обозначим координаты точки H, то нормальные уравнения могут быть написаны в виде:

$$a_{11}(y_1'-y_1)+a_{21}(y_2'-y_2)+\ldots+a_{n1}(y_n'-y_n)=0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{1m}(y_1'-y_1)+a_{2m}(y_2'-y_2)+\ldots+a_{nm}(y_n'-y_n)=0,$$

и достаточно их сложить почленно, умножив предварительно на $\theta_1'-x_1, \ldots, \theta_m'-x_m$, чтобы получить условие перпендикулярности:

$$(y_1'-y_1)(y_1'-q_1)+\ldots+(y_n'-y_n)(y_n'-q_n)=0.$$

Отметим, что эта ортогональность влечет за собой в качестве своего следствия, что $\overrightarrow{HM}$, в частности, перпендикулярен и к $\overrightarrow{\omega H}$.

$$\overrightarrow{\omega M}^2=\overrightarrow{\omega H}^2+\overrightarrow{HM}^2.$$

97. ОЦЕНКА ДОПУЩЕННЫХ ОШИБОК

Обратимся теперь с целью установить точность полученных результатов, к рассмотрению *ошибок наблюдений*, допущенных при измерении величин $\eta_1, \eta_2, \ldots, \eta_n$, принимая вместо них результаты $y_1, y_2, \ldots, y_n$.

Из наших гипотез вытекает, что *M есть случайная точка, законы вероятностей которой определяются элементарной вероятностью*

$$\frac{k^n}{(\sqrt{\pi})^n}\, e^{-k^2\Sigma\zeta_i^2}\, d\zeta_1\, d\zeta_2 \ldots d\zeta_n, \tag{196}$$

которую можно считать определенной в пространстве (E_n), рассмотренном ранее и взятом относительно прямоугольных координатных осей с началом в ω.

1) Выразим прежде всего эту элементарную вероятность в полярных координатах с началом в точке ω. Полагаем $\omega M = \rho$, тогда выражение (196) преобразуется в

$$Ae^{-k^2\rho^2}\rho^{n-1}\, d\rho\, d\Omega, \tag{197}$$

в котором A представляет собой некоторую постоянную, а Ω есть элементарный телесный угол, для которого не стоит давать более подробное объяснение.

2) С другой стороны, заметим, что точку M можно получить, беря векторную сумму из векторов $\overrightarrow{\omega H}$ на множестве (π), зависящих от параметров x_i, в числе m и, поскольку H уже таким образом выбрано, также и от вектора, перпендикулярного $\overrightarrow{HM}$, зависящего от $n-m$ параметров. Отсюда следует, что точку M можно определить системой *полуполярных координат*, в которую входят векторы-лучи $\omega H = \lambda \cdot HM = \mu$, сопровождаемые первый $m-1$, а второй $n-m-1$ полярными углами, описывать которые также нет надобности.

Элементарная вероятность (196) в этой системе принимает вид:

(198) $$Ae^{-k^2(\lambda^2+\mu^2)}\lambda^{m-1}\,d\lambda\,d\Omega_1\,\mu^{n-m-1}\,d\mu\,d\Omega_2,$$

где $d\Omega_1$ и $d\Omega_2$ означают элементарные телесные углы, объяснять которые здесь не стоит.

Полученные здесь выражения дают право утверждать, во-первых, что случайная переменная $\omega M = \rho$ следует закону вероятности, не зависящему от полярных углов $\varphi_1, \varphi_2, \ldots, \varphi_{n-1}$, которыми определяется $d\Omega_1$; элементарная вероятность, характеризующая этот закон, имеет такой вид:

(199) $$A_1 e^{-k^2\rho^2}\rho^{n-1}\,d\rho,$$

где A_1 является величиной постоянной.

Во-вторых, случайные переменные $\omega H = \lambda$ и $HM = \mu$ также следуют законам, независимым между собой и характеризуемым элементарными вероятностями

(200) $$A_2 e^{-k^2\lambda^2}\lambda^{m-1}\,d\lambda \quad \text{и} \quad A_3 e^{-k^2\mu^2}\mu^{n-m-1}\,d\mu,$$

где A_2 и A_3 — постоянные величины.

Опытные данные дают нам возможность оценить случайную переменную

$$\mu = HM = \sqrt{(y_1'-y_1)^2+\ldots+(y_n'-y_n)^2}.$$

Вместе с тем, если мы положим во второй из формул (200)

$$k\mu = \frac{\chi}{\sqrt{2}},$$

то увидим, что χ следует закону, характеризуемому элементарной вероятностью

(201) $$Ce^{-\frac{\chi^2}{2}}\chi^{\frac{n-m}{2}-1}\,d\chi^2;$$

а это согласно § 43 представляет собой *закон для* χ^2 при $n-m$ *степенях свободы*.

Мы знаем, если принять $\nu = n - m$, то математическое ожидание χ^2 равно:

$$E(\chi^2) = \nu,$$

при этом дисперсия χ^2 составляет $\mu_2 = 2\nu$. Это означает, что математическое ожидание и дисперсия случайной переменной $2k^2\overline{HM}^2$ равны соответственно $n-m$ и $2(n-m)$.

Выражая математическое ожидание в других обозначениях, мы для $\frac{\overline{HM}^2}{n-m}$ напишем:

(202) $$E\left(\frac{\overline{HM}^2}{n-m}\right) = \frac{1}{2k^2},$$

а дисперсия этой случайной величины равна:

$$\text{var}\left(\frac{\overline{HM}^2}{n-m}\right)=\frac{1}{2(n-m)k^4}. \tag{203}$$

Из (202) мы заключаем, что согласно § 68 *величина* $\frac{\overline{HM}^2}{n-m}$, *получаемая из наблюдений, если для* M *взять точку выборки, имеет вид*:

$$\frac{\sum(a_{i1}\theta_1'+\ldots+a_{im}\theta_m'-y_i)^2}{n-m}=\frac{\sum u_i^2}{n-m}$$

и является совершенно правильной оценкой для $\frac{1}{2k^2}$.

98. ЧИСЛОВОЙ ПРИМЕР

Вернемся к обозначениям § 93, в котором мы остановились для ясности на случае трех неизвестных, применяя символику Гаусса. Коэффициенты a, b, c, a', $в'$, c', ... известны в точности, тогда как l, l', ... представляют собой результаты измерений, точность которых k мы предполагаем одинаковой.

Сумма квадратов погрешностей равна:

$$\sum u_i^2=(uu)=\sum(ax+by+cz+l)^2,$$

предполагая здесь, что x, y, z — значения неизвестных, определяемых из нормальных уравнений.

Обозначим черех X, X', ... величины, принимаемые при произвольных x, y, z, правыми частями уравнений (183), тогда мы получим тождества:

$$X^2=(ax+by+cz+l)\,X=axX+byX+czX+lX$$

откуда, складывая почленно, найдем

$$\sum X^2=x\sum aX+y\sum bX+z\sum cX+\sum lX.$$

Если x, y. z удовлетворяют нормальным уравнениям

$$\sum aX=0,\quad \sum bX=0,\quad \sum cX=0,$$

то вышеприведенное тождество дает для Σu^2 величину

$$(uu)=\sum lX-(al)\,x+(bl)\,y+(cl)\,z+(ll). \tag{204}$$

Для числового примера в § 94 мы, таким образом, получим: $(uu)=96{,}66x-203{,}46y+227{,}45z+275{,}45$, или, сделав все нужные вычисления, $(uu)=0{,}0086$.

Так как в разобранном примере n и m равны соответственно 9 и 3, то полученная оценка для $\frac{1}{2k^2}$ равна:

$$\frac{1}{2k^2}\simeq\frac{0{,}0086}{6},$$

а средняя квадратическая ошибка для результатов измерений получает оценку

$$\sqrt{\frac{0,0086}{6}} = 0,0378.$$

99. ВЕСА ДЛЯ НЕИЗВЕСТНЫХ

Если на основании выводов § 98 произвести оценку общей точности k тех измерений, с помощью которых получена величина l, то нетрудно поставить их в соответствие с теми ошибками, которые можно ожидать в значениях неизвестных x, y, z, получаемых из нормальных уравнений.

Действительно, решение нормальных уравнений, найденных в том виде, который указан в (184),дают, если пользоваться формулами Крамера вместо метода Гаусса, значения

$$x = -(al)\lambda - (bl)\mu - (cl)\nu, \tag{205}$$

где λ, μ, γ представляют собой миноры, взятые по первой колонке детерминанта из коэффициентов при неизвестных, поделенных на значение самого детерминанта. Эти величины λ, μ, γ представляют собой, следовательно, решения системы уравнений:

$$\begin{aligned} (aa)\lambda + (ab)\mu + (ac)\nu &= 1, \\ (ab)\lambda + (bb)\mu + (bc)\nu &= 0, \\ (ac)\lambda + (bc)\mu + (cc)\nu &= 0, \end{aligned} \tag{206}$$

легко поддающихся решению хотя бы способом, указанным Гауссом для решения нормальных уравнений.

Из (205) мы, раскрывая правую часть, получим:

$$x = -(\alpha l + \alpha' l' + \alpha'' l'' + \ldots), \tag{207}$$

где

$$\alpha = a\lambda + b\mu + c\nu,$$

$$\alpha' = a'\lambda + b'\mu + c'\nu,$$

$$\alpha'' = a''\lambda + b''\mu + c''\nu,$$

$$\cdots\cdots\cdots\cdots$$

Таким образом, если средняя квадратическая ошибка для l равна $\frac{1}{k\sqrt{2}}$, то средняя квадратическая ошибка для αl равна $\sqrt{\frac{\alpha^2}{2k^2}}$, а после сложения дисперсий средняя квадратическая ошибка для x согласно (207) составит:

$$Q = \sqrt{\frac{\alpha^2 + \alpha'^2 + \alpha''^2 + \ldots}{2k^2}}.$$

Но, принимая во внимание выражения для α, α', α'', ..., мы получим:

$$\alpha^2 = \alpha(a\lambda + b\mu + c\nu) = \lambda a\alpha + b\mu\alpha + c\nu\alpha,$$

откуда

$$\alpha^2+\alpha'^2+\alpha''^2+\ldots=\lambda(a\alpha+a'\alpha'+a''\alpha''+\ldots)$$
$$+\mu(b\alpha+b'\alpha'+b''\alpha''+\ldots)$$
$$+\nu(c\alpha+c'\alpha'+c''\alpha''+\ldots)$$

или еще иначе

$$\lambda[(aa)\lambda+(ab)\mu+(ac)\nu]+\mu[(ab)\lambda+(bb)\mu+(bc)\nu]+$$
$$+\nu[(ac)\lambda+(bc)\mu+(cc)\nu],$$

и в конце концов, согласно (206)

$$\alpha^2+\alpha'^2+\alpha''^2+\ldots=\lambda, \tag{208}$$

откуда

$$Q^2=\frac{\lambda}{2k^2}.$$

Величину $\frac{1}{\lambda}$ *принято называть весом* x; веса эти получаются через решение, как указано выше, уравнений (206).

Если мы воспользуемся числовым примером § 94, то вычисление нам даст для величин, обратных весам x, y, z, следующие значения:

$$0{,}0089,\quad 0{,}0321,\quad 0{,}0197,$$

а средние квадратические ошибки значений x, y, z могут быть оценены как

$$0{,}0378\sqrt{0{,}0089}=0{,}0036,$$
$$0{,}0378\sqrt{0{,}0321}=0{,}0067,$$
$$0{,}0378\sqrt{0{,}0197}=0{,}0053.$$

100. ТОЧНОСТЬ СРЕДНЕЙ, ПОЛУЧЕННОЙ ИЗ РЯДА ИЗМЕРЕНИЙ

Исследование, проведенное в § 97, охватывает, как частный случай, оценку точности серии измерений, произведенных над одной и той же неизвестной нам величиной.

Если $x_1, x_2, \ldots, x_n$ — результаты n измерений величины θ и если ошибки всех этих измерений следуют закону Гаусса с мерой точности k, т. е. с одинаковой средней квадратической ошибкой, то «статистика»

$$\theta'=\frac{x_1+x_2+\ldots+x_n}{n}$$

искажена ошибкой, подчиняющейся закону Гаусса, в котором точность имеет величину $k\sqrt{n}$. Мы не знаем k, но совершенно правильная оценка для $\frac{1}{2k^2}$ дается формулой:

$$s^2=\frac{(x_1-\theta')^2+\ldots+(x_n-\theta')^2}{n-1}. \tag{209}$$

Дисперсия в законе Гаусса, создаваемая тем, что мы принимаем θ' за истинные, и которая измеряется величиной $\frac{1}{2nk^2}$, может получить свою оценку в виде:

$$\frac{s^2}{n}=\frac{\sum(x_i-\theta')^2}{n(n-1)}.$$

Оценка (209) была получена уже раньше другим путем в § 73, причем относительно закона ошибок непосредственных измерений можно ограничиться лишь очень широкими допущениями.

101. ОБЩИЕ ВЫВОДЫ

Указания, данные в этой последней главе, достаточны, чтобы читатель мог ознакомиться с методом наименьших квадратов и при случае его применить. Различные возражения выдвигались против этого метода; но мы не станем их рассматривать, так как метод от них не погиб. Этим мы не хотим сказать, что метод наименьших квадратов следует применять всюду, без разбора, и приписывать ему способность извлекать хорошие результаты из серии плохих опытных данных.

Но все же, после того как наблюдатель приложит все старания, чтобы улучшить точность каждого сделанного им измерения, метод наименьших квадратов и вся теория ошибок в целом дает способ обсудить и проконтролировать полученные результаты. И это средство тем более надо оценить, когда дело идет об измерениях, предназначенных для подтверждения какой-либо важной научной теории или определения фундаментальных постоянных.

ПРИЛОЖЕНИЯ

I. *Вычисление интегралов*

$$I_m = \int_0^{\pi/2} \sin^m x\, dx \text{ и } J_m = \int_0^{\pi/2} \cos^m x\, dx.$$

Оба интеграла одинаковы: кривая $y = \sin^m x$ и кривая $y = \cos^m x$ обе симметричны по отношению к прямой $x = \frac{\pi}{4}$.

Вычислим хотя бы I_m. Применяя интегрирование по частям, мы напишем:

$$I_m = \int_0^{\pi/2} \sin^{m-1} x\, d(-\cos x) = -\left| \sin^{m-1} x \cos x \right|_0^{\pi/2} +$$

$$+ \int_0^{\pi/2} \cos x\, d(\sin^{m-1} x),$$

откуда сразу же находим

$$I_m = (m-1) \int_0^{\pi/2} \sin^{m-2} x \cos^2 x\, dx,$$

что дает

$$mI_m = (m-1) I_{m-2}.$$

Эта формула показывает, что надо различать два случая: когда m четно и когда m нечетно.

1-й случай: m — четно.

Пусть $m = 2p$, тогда мы имеем:

$$2pI_{2p} = (2p-1) I_{2p-2},$$

$$(2p-2) I_{2p-2} = (2p-3) I_{2p-4},$$

$$\cdots\cdots\cdots\cdots$$

$$4I_4 = 3I_1,$$

$$2I_2 = I_0.$$

Но

$$I_0 = \int_0^{\pi/2} dx = \frac{\pi}{2},$$

откуда

(1) $$I_{2p}=\frac{\pi}{2}\cdot\frac{1.2.3\ldots(2p-1)}{2.4.6\ldots 2p}.$$

2-й случай: m — нечетно.

Пусть $m=2p+1$, тогда мы имеем:

$$(2p+1)I_{2p+1}=2pI_{2p-1},$$
$$(2p-1)I_{2p-1}=(2p-2)I_{2p-3},$$
$$\cdots\cdots\cdots\cdots$$
$$3I_3=2I_1.$$

Но

$$I_1=\int_0^{\pi/2}\sin x\,dx=1.$$

Следовательно,

(2) $$I_{2p+1}=\frac{2.4.6\ldots 2p}{1.3.5\ldots(2p+1)}.$$

Сопоставление формул для I_{2p} и для I_{2p+1} приводит к любопытному выводу, на который мы должны обратить внимание в связи с доказательством формулы Стирлинга.

В интервале $(0, \frac{\pi}{2})$ функция $\sin x$ изменяется от 0 до 1; поэтому мы имеем:

$$\sin^{2p}x>\sin^{2p+1}x>\sin^{2p+2}x,$$

откуда получаются неравенства

$$I_{2p}>I_{2p+1}>I_{2p+2},$$

которые можно писать и в таком виде:

$$1>\frac{I_{2p+1}}{I_{2p}}>\frac{I_{2p+2}}{I_{2p}}.$$

Таким образом, отношение $\frac{I_{2p+2}}{I_{2p}}=\frac{2p+1}{2p+2}$ стремится к единице, когда p беспредельно возрастает. Следовательно, отношение

$$\frac{I_{2p+1}}{I_{2p}}$$

также при этом же условии стремится к единице. Если написать

$$\frac{I_{2p+1}}{I_{2p}}=1+\varepsilon,$$

то отсюда вытекает, принимая во внимание формулы (1) и (2), что

$$\frac{\pi}{2}=\frac{(2.4.6\ldots 2p)^2}{(1.3.5.\ \ldots\ (2p-1))^2\,(2p+1)}\ (1+\varepsilon). \tag{3}$$

Эта формула, дающая возможность приближенно вычислять π, приписывается Валлису (Wallis).

II. *Формула Стирлинга*

Положим

$$n!=n^n e^{-n}\sqrt{n}\,f(n).$$

Сперва мы покажем, что при безграничном возрастании целых чисел $n, f(n)$ *стремится к некоторому определенному пределу.*

Замещая n через $n+1$ мы получим:

$$(n+1)!=(n+1)^{n+1}e^{-(n+1)}\sqrt{n+1}\,f(n+1);$$

откуда

$$\frac{f(n+1)}{f(n)}=\left(\frac{n}{n+1)}\right)^{n+\frac{1}{2}}e.$$

Отсюда вытекает, что

$$\operatorname{Log}\frac{f(n+1)}{f(n)}=1-\left(n+\frac{1}{2}\right)\operatorname{Log}\left(1+\frac{1}{n}\right);$$

а затем, если воспользоваться разложением в ряд Log $(1+n)$,

$$\operatorname{Log}(1+x)=x-\frac{x^2}{2}+\ldots+(-1)^{n+1}\frac{x^n}{n}+\ldots,$$

мы выводим следующую формулу:

$$\operatorname{Log}f(n+1)-\operatorname{Log}f(n)=1-\left(n+\frac{1}{2}\right)\left(\frac{1}{n}-\frac{1}{2n^2}+\frac{1}{3n^3}-\ldots\right).$$

Ряд, в котором общий член u_n равен

$$u_n=\operatorname{Log}f(n+1)-\operatorname{Log}f(n)$$

есть ряд сходящийся, так как

$$u_n=-\frac{1}{12n^2}+\frac{1}{12n^3}+\ldots$$

Произведение $n^2\,|u_n|$ действительно имеет при безпредельно возрастающем n определенный предел.

Отсюда следует, что Log $f(n+1)$ и, следовательно, также $f(n+1)$ или $f(n)$ имеют предел при бесконечном n. Обозначим этот предел буквой C. Мы видим, что

$$n!\to Cn^n e^{-n}\sqrt{n}.$$

Чтобы найти чему равна постоянная величина C, мы обратимся к выведенной выше формуле Валлиса.

$$\Phi(n)=\frac{2.4.6\ldots 2n}{1.3.5\ldots(2n-1)\sqrt{2n+1}}$$

стремится к $\sqrt{\frac{\pi}{2}}$ при $n\to\infty$.

Если умножить числитель и знаменатель на $2.4.6\ldots 2n$, то в числителе окажется

$$(2.4.6\ldots 2n)^2=2^n(n!)^2,$$

а в знаменателе

$$(2n)!\sqrt{2n+1}.$$

Заменяя $n!$ на $Cn^ne^{-n}\sqrt{n}$ и $(2n)!$ на $C(2n)^{2n}e^{-2n}\sqrt{2n}$, мы умножим $\Phi(n)$ на множитель, стремящийся к единице, когда n беспредельно возрастает. После всех сокращений мы получим выражение

$$\frac{nC}{\sqrt{2n(2n+1)}}.$$

Оно стремится к $\sqrt{\frac{\pi}{2}}$, а так как его предел при другом подходе равен $\sqrt{\frac{\pi}{2}}$, то получается, что

$$C=\sqrt{2\pi}.$$

Если при тех же условиях мы напишем

$$n!=n^ne^{-n}\sqrt{2\pi n}\,\varphi(n),$$

то $\varphi(n)$ будет величиной, стремящейся к единице при безграничном возрастании n.

Заметим, что $\varphi(n)$ стремится к единице, проходя через значения, большие единицы, ибо имеем

$$\mathrm{Log}\,\varphi(n+1)-\mathrm{Log}\,\varphi(n)=u_n=-\frac{1}{12n^2}+\ldots,$$

откуда видно, что $\varphi(n+1)$ меньше, чем $\varphi(n)$.

Поэтому мы вправе представить $n!$ в виде:

$$n!=n^ne^{-n}\sqrt{2\pi n}(1+\varepsilon_n),$$

т. е. в той форме, которую мы приняли в § 29.

Замечание. Можно уточнить форму, найденную для $\varphi(n)$. Например:

$$\varphi(n)=1+\frac{1}{\alpha n}+\frac{1+\theta_n}{\beta n^2},$$

где α и β — коэффициенты, подлежащие определению, а θ_n стремится к нулю, когда n беспредельно возрастает.

Мы получим:

$$\operatorname{Log}\varphi(n)=\frac{1}{\alpha n}+\frac{1+\theta_n}{\beta n^2}-\frac{1}{2}\left[\frac{1}{\alpha n}+\frac{1+\theta_n}{\beta n^2}\right]+\ldots,$$

откуда

$$\operatorname{Log}\varphi(n+1)-\operatorname{Log}\varphi(n)=-\frac{1}{\alpha n^2}+\left[\frac{1}{\alpha}-\frac{2}{\beta}+\frac{1}{\alpha^2}\right]\frac{1}{n^3}+\ldots,$$

и путем сопоставления с разложением в ряд общего члена u_n получим

$$\alpha=12,$$
$$\beta=288.$$

Поэтому можно в окончательном виде написать новое приближенное значение для $n!$ в следующем виде:

$$n!\doteq n^n e^{-n}\sqrt{2\pi n}\left[1+\frac{1}{12n}+\frac{1+\theta_n}{288\,n^2}\right].$$

Таблица 1

$$\psi(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{x^2}{2}} dx$$

(О том, как пользоваться этой таблицей, см. § 34)

t	$\psi(t)$	$1-2\psi(t)$	t	$\psi(t)$	$1-2\psi(t)$
0,0	0,0000	1	2,1	0,4821	0,0358
0,1	0,0398	0,9204	2,2	0,4861	0,0278
0,2	0,0793	0,8414	2,3	0,4893	0,0214
0,3	0,1179	0,7642	2,4	0,4918	0,0164
0,4	0,1555	0,6890	2,5	0,4938	0,0124
0,5	0,1915	0,6170	2,6	0,4953	0,0094
0,6	0,2258	0,5484	2,7	0,4965	0,0070
0,7	0,2580	0,4840	2,8	0,4974	0,0052
0,8	0,2881	0,4238	2,9	0,4981	0,0038
0,9	0,3159	0,3682	3,0	0,4987	0,0026
1,0	0,3413	0,3174	3,5	0,4997674	0,0004652
1,1	0,3643	0,2714	4,0	0,4999683	0,0000634
1,2	0,3849	0,2302	4,5	0,4999966	0,0000068
1,3	0,4032	0,1936	5,0	0,499999713	0,000000574
1,4	0,4192	0,1616			
1,5	0,4332	0,1336			
1,6	0,4452	0,1096			
1,7	0,4554	0,0892			
1,8	0,4641	0,0718			
1,9	0,4713	0,0574			
2,0	0,4773	0,0454			

Таблица II

$$\theta(x) = \frac{1}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt$$

(О том, как пользоваться этой таблицей, см. § 38)

x	$\theta(x)$	$1-\theta(x)$
0,1	0,1125	0,8875
0,2	0,2227	0,7773
0,3	0,3286	0,6714
0,4	0,4284	0,5716
0,4769	0,5	0,5
0,5	0,5205	0,4795
0,6	0,6039	0,3961
0,7	0,6778	0,3222
0,8	0,7421	0,2579
0,9	0,7969	0,2031
1,0	0,8427	0,1573
1,1	0,8802	0,1198
1,163	0,9	0,1
1,5	0,9661	0,0339
1,821	0,99	0,01
2,0	0,9953	0,0047
2,327	0,999	0,001
2,751	0,9999	0,0001
3,0	0,999979	0,000021
3,123	0,99999	0,00001
3,46	0,999999	0,000001
3,763	0,9999999	0,0000001
4,0	0,999999985	0,000000015

Таблица III

Если положить

$$P = C_\lambda \int_{a^2}^{+\infty} e^{-\frac{x}{2}} \frac{\nu - 2}{x^2} \, dx,$$

то нижеследующая таблица дает в значениях функции от ν (ν — число степеней свободы) те значения a^2, для которых

колонка I : $P = 0{,}05$,
колонка II : $P = 0{,}01$.

ν	Значения a^2, для которых $P=0{,}05$	Значения a^2, для которых $P=0{,}01$
1	3,84	6,64
2	6,99	9,21
3	7,81	11,34
4	9,49	13,28
5	11,07	15,09
6	12,59	16,81
7	14,07	18,48
8	15,51	20,09
9	16,92	21,67
10	18,31	23,21
11	19,68	24,72
12	21,03	26,22
13	22,36	27,69
14	23,68	29,14
15	25,00	30,58
16	26,30	32,00
17	27,59	33,41
18	28,87	34,80
19	30,14	36,19
20	31,41	37,57
21	32,67	38,93
22	33,92	40,29
23	35,17	41,64
24	36,41	42,98
25	37,65	44,31
26	38,88	45,64
27	40,11	46,96
28	41,34	48,28
29	42,56	49,59
30	43,77	50,89

ОГЛАВЛЕНИЕ

Эм. Борель, Р. Дельтейль, Роже Юрон

ВЕРОЯТНОСТИ, ОШИБКИ

* * *

Редактор *З. А. Сумник*
Техн. редактор *Р. Н. Феоктистова*
Корректор *Т. М. Васильева*
Худ. редактор *Т. В. Стихно*

Сдано в набор 12/VII 1971 г. Подп. к печ. 20/XII 1971 г.
Формат бумаги 60×90/16 Бумага № 2 Объем печ. л. 11
Уч.-изд. л. 9,84 Тираж 13 500 экз.
(Тематич. план 1971 г. № 61) Зак. № 416 Цена 69 коп.
Издательство «Статистика», Москва, ул. Кирова, 39.

Московская типография № 4 Главполиграфпрома
Комитета по печати при Совете Министров СССР

69 коп.

СТАТИСТИКА 1972