

Специальные пособия для Высшей Школы

Л. К. ЛАХТИН

Проф. 1-го Московского Государственного Университета

КУРС ТЕОРИИ ВЕРОЯТНОСТЕЙ

Научно-Технической Секцией Государ-
ственного Ученого Совета рекомендовано
в качестве пособия для Высших Учебных
Заведений

Государственное Издательство

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО

МОСКВА — ЛЕНИНГРАД

- Гельмгольц, Г. О сохранении силы. Физическое исследование. Перевод и примечания Лазарева (Серия „Классики естествознания“, № 5). С рисунками и таблицами. 1922 г. 71 стр. Ц. 25 к.
- Глинка, К. Д. Почвы России и прилегающих стран. 1923. Стр. 348. Ц. 3 р.
- Гурса, Э. Курс математического анализа. Том II, ч. I. Теория аналитических функций. Под ред. Б. К. Млодзеевского. С 70 чертежами. 270 стр. Цена 1 р. 75 к.
- Демьянов, Н. Я., проф. Общие приемы анализа растительных веществ. Агрономическая химия. С приложениями А. А. Стольгане. 1923. Стр. 348. Ц. 1 р. 75 к.
- Егоров, Д. Ф., проф. Элементы теории чисел. 119 стр. Ц. 1 р. 10 к.
- Игнатъев, Б. В. Весенняя флора. Определитель весенних растений средней России. 277 стр. Ц. 80 к.
- Каблуков, И. Основные начала физической химии. Вып. II. Электрохимия. С 44 рис. Изд. 2-е, исправл. и дополненн. 317 стр. Ц. 1 р. 70 к.
- Комаров, В. Л., акад. Практический курс ботаники. Ч. I. Строение растений. 1923. Стр. 290. Ц. 3 р.
- Коялович, Б. М., проф. Лекции по высшей математике. Т. I. Дифференциальное исчисление с приложениями к анализу. Изд. 2-е. 244 стр. Ц. 1 р. 50 к.
- Нузнецов, Н. И., проф. Журнал экскурсанта-ботаника. Тетрадь для практических занятий на экскурсии. Флористические экскурсии. 32 стр. Ц. 80 к.
- Лактин, Л. Н. Кривые распределения и построение для них интерполяционных формул по способам Пирсона и Брунса. IV + 152 стр.
- Лучицкий, В. И. Курс петрографии. Изд. 2-е, дополненное и исправленное, 341 стр. Ц. 2 р. 10 к.
- Любименко, В. Н. Курс общей ботаники. Руководство для слушателей высших учебных заведений и самообразования. Берлин. 1923. Стр. 1042.
- Меншуткин, Н. А. Аналитическая химия. Двенадцатое издание, переработанное и дополненное. XVI + 439 стр. С 66 рис. Ц. 2 р. 20 к.
- Млодзеевский, Б. К. Основы высшей алгебры. 111 стр. Ц. 85 к.
- Млодзеевский, Б. К. Основы аналитической геометрии в пространстве. Изд. 4-е. 151 стр. Ц. 1 р.
- Наумов, Н. А., проф. Курс фитопатологии. VI + 392 стр. С 53 рис. Цена 5 руб.
- Никитинский, Я. Я. Практич. занятия по микробиологии. 92 стр. Ц. 70 к.
- Ньюкомб, С. Звезды. Под редакцией проф. В. В. Серафимсва. 240 стр. Ц. 1 р. 50 к.
- Резерфорд, Э. Строение атома и искусственное разложение элементов. 177 стр. Ц. 1 р. 25 к.
- Ростовцев, С. И., проф. Фитопатология. Болезни и повреждения растений. Руковод. для студентов, агрономов, лесоводов, садоводов и огородников. 4-е изд. Под ред. и с дополн. проф. Л. И. Курсанова. 415 стр. Ц. 5 р.

СПЕЦИАЛЬНЫЕ ПОСОБИЯ ДЛЯ ВЫСШЕЙ ШКОЛЫ

Депозитарий

Л. К. ЛАХТИН

Проф. 1-го Московского Государственного Университета

КУРС ТЕОРИИ ВЕРОЯТНОСТЕЙ

Научно-Технической Секцией Государ-
ственного Ученого Совета рекомендовано
в качестве пособия для Высших Учебных
Заведений

Очищено

РЕСПУБЛИКАНСКАЯ
НАУЧНО-ТЕХНИЧЕСКАЯ
БИБЛИОТЕКА

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
МОСКВА 1924 ГОДА ПЕТРОГРАД

И. М. Ливин

Информационный

— ПЕТРОГР. ОТД. —
ГЛАВНОЙ К-РЫ «ИЗВЕСТИЯ Ц.И.К.
С.С.С.Р. И В.Д.И.К.»
Тип. «КРАСНЫЙ АГИТАТОР»
Пгг., 7-я рота, 25

ОГЛАВЛЕНИЕ

Предисловие	Стр. III
-----------------------	----------

ВВЕДЕНИЕ.

§ 1. О случайных явлениях	1
2. Определение величины вероятности случайного явления	3
3. Вероятность субъективная и объективная	2
4. Вероятность a priori и a posteriori	8
5. Исторический очерк	11

ЧАСТЬ I.

Основы теории вероятностей.

ГЛАВА I.

Приемы вычисления вероятностей.

§ 1. Примеры непосредственного вычисления вероятностей	17
2. Теорема сложения вероятностей	22
3. Теорема умножения вероятностей	24
4. Примеры приложения теорем сложения и умножения вероятностей	27
5. Полная вероятность	30
6. Различные косвенные приемы вычисления вероятностей	34
7. О геометрических вероятностях	40
8. Примеры на вычисление геометрических вероятностей	43

ГЛАВА II.

О математическом ожидании.

§ 1. Понятие о математическом ожидании случайной суммы	47
2. О безобидных играх	49
3. Петербургская игра	51
4. О нравственном ожидании	52

ГЛАВА III.

О вероятностях сложных событий, имеющих место при большом числе испытаний.

§ 1. Выражение вероятности сложного события при большом числе испытаний над простыми событиями, между собою независимыми	57
2. Наивероятнейшее сложное событие при большом числе испытаний	60
3. Графическое изображение членов биномиального ряда	61
4. Интерполяционная формула вероятностей сложных событий при многократном повторении испытаний	64
5. Вид интерполяционной кривой	69
6. Теорема Якова Бернулли	72
7. Вероятное отклонение	77
8. Среднее отклонение	77
9. Среднее квадратическое отклонение	80

§ 10.	Примеры приложения теоремы Якова Бернулли	80
§ 11.	Рассмотрение одного примера сложного события, состоящего из большого числа простых событий, между собою зависимых	83
§ 12.	Случай, когда сложное событие состоит из простых событий, зависящих от двух причин с различными вероятностями	89
§ 13.	Случай, когда сложное событие состоит из простых событий, зависящих от многих причин с различными вероятностями	93
§ 14.	Закон больших чисел, теорема Пуассона	97
§ 15.	Теорема Чебышева	100
§ 16.	Вывод теорем Пуассона и Бернулли из теоремы Чебышева	108

ГЛАВА IV.

О вероятностях на основании опыта.

§ 1.	Постановка вопроса	112
§ 2.	Теорема Бейеса	113
§ 3.	Приложение теоремы Бейеса к случаю бесконечно большого числа гипотез	118
§ 4.	Теорема, обратная теореме Якова Бернулли	122

ЧАСТЬ II.

Приложение теории вероятностей к обработке результатов измерений.

ГЛАВА I.

Теория ошибок при непосредственном измерении.

§ 1.	Происхождение и характер ошибок измерения	131
§ 2.	Свойства случайных ошибок измерения	132
§ 3.	Начало арифметической середины. Вывод формулы Гаусса вероятности ошибки измерения	135
§ 4.	Предел возможных ошибок, вероятная ошибка, средняя и средняя квадратическая ошибка	140
§ 5.	Нахождение меры точности произведенных измерений	143
§ 6.	Мера точности результата, выведенного из системы равноточных измерений	146
§ 7.	О весах неравноточных измерений	150
§ 8.	Наивероятнейший вывод из системы неравноточных измерений	152
§ 9.	Мера точности результата, выведенного из системы неравноточных измерений	153
§ 10.	Вычисление величины нормальной меры точности	154
§ 11.	Пример вычисления в случае измерений неравноточных	155

ГЛАВА II.

О функциях величин, полученных измерением.

§ 1.	Постановка вопроса	157
§ 2.	Вычисление величины функции: $w = ax + b$	157
§ 3.	Вычисление величины функции: $w = x + y + z + \dots + u$	158
§ 4.	Вычисление величины: $w = ax + by + cz + \dots + fu + g$	161
§ 5.	Вычисление величины: $w = f(x, y, z, \dots, u)$	162

ГЛАВА III.

Способ наименьших квадратов.

§ 1.	Задача способа наименьших квадратов. Начальные уравнения	163
§ 2.	Нормальная система уравнений	166
§ 3.	Нормальная система уравнений, преобразованная по способу Гаусса	168
§ 4.	Вес результатов, даваемых способом наименьших квадратов	173
§ 5.	Выражение весов в способе наименьших квадратов помощью определителей	177
§ 6.	Нахождение величины нормальной меры точности h в способе наименьших квадратов	178
§ 7.	Пример применения способа наименьших квадратов	181

ЧАСТЬ III.

Приложение теории вероятностей и изучению массовых явлений.

ГЛАВА I.

О законах распределения.

	Стр.
§ 1. Понятие о массовых явлениях	185
§ 2. Понятие о законе распределения	186
§ 3. Средняя величина признака, понятие о моментах	193
§ 4. Понятие о моде, о радиусе асимметрии, о среднем квадратическом отклонении и о скошенности кривой распределения	202
§ 5. Построение интерполяционной формулы по способу моментов	204
§ 6. Формула нормального распределения	207
§ 7. Формула биномиального распределения	210
§ 8. Понятие о формулах Пирсона гипергеометрического закона распределения	217

ГЛАВА II.

О корреляциях.

§ 1. Распределение по двум признакам	219
§ 2. Нормальное распределение по двум признакам	226
§ 3. Коэффициент корреляции	232
§ 4. Прямая регрессии. Коэффициент регрессии	234
§ 5. Корреляционный эллипс	240
§ 6. Практические приемы вычисления коэффициентов корреляции и регрессии	243
§ 7. Понятие о криволинейной корреляции и о корреляционном отношении	248

ГЛАВА III.

О статистических рядах.

§ 1. Понятие о статистическом ряде	251
§ 2. Статистическая вероятность	254
§ 3. Сравнение статистических вероятностей в двух совокупностях	259
§ 4. Коэффициент дисперсии	262
§ 5. Схема Лексиса	265
§ 6. Схема Пуассона	269

Приложения:

Таблицы функции $\Phi(x) = \frac{2}{\sqrt{\pi}} e^{-x^2}$	271
---	-----

Таблицы функции $\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$	273
---	-----

ПРЕДИСЛОВИЕ

В последнее время среди учащихся и лиц, имеющих дело с применением статистического метода к различным областям знания, замечается большой интерес к теории вероятностей с ее приложениями. Между тем наша математическая литература очень бедна курсами теории вероятностей, а имеющиеся курсы носят характер чисто теоретический, за редким исключением не затрагивая других применений, кроме теории ошибок и способа наименьших квадратов.

Предлагаемый курс представляет собою попытку пополнить этот пробел. Он рассчитан на подготовку читателя в пределах первых двух курсов математического факультета или общего курса математики, читаемого обыкновенно в специальных технических учебных заведениях и на естественном отделении университета. Доказательства подобраны так, чтобы они, удовлетворяя требованиям строгости, по возможности, были просты и не требовали от читателя большого количества математических формул. Понятно, что этим условиям, почти друг другу противоположным, удовлетворить не легко; но думаю, что насколько возможно, к цели приблизиться удалось. Многие приемы доказательства пришлось составить заново и распределить материал сообразно намеченной цели.

Материал, обычно излагаемый в курсах теории вероятностей, в предлагаемом руководстве сгруппирован около основного понятия о законе распределения в массовых явлениях. Это понятие объединяет теоретические отделы теории вероятностей с различными ее приложениями. Оно появляется почти в начале курса, в главе III части 1-й при доказательстве теоремы Якова Бернулли, обыкновенно излагаемой без ясно выраженного геометрического представления, которое на мой взгляд весьма существенно в этом отделе и служит первым подходом к общей теории кривых распределения. В той же главе III части 1-й рассмотрены различные задачи того же рода и закон больших чисел. Во 2-й части курса изложена Гауссова теория ошибок и способ наименьших квадратов. Имея самостоятельный интерес для лиц, занимающихся обработкою результатов измерения (преимущественно физиков и астрономов), теория ошибок в то же время представляет второй подход к понятию о кривых распределения, осно-

ванный на начале арифметической середины и применении теоремы Бейеса о вероятностях *a posteriori*. Эти два подхода представляют естественный путь, приведший Кетле к созданию математической статистики.

Часть 3-я курса представляет изложение главных вопросов математической статистики. Первая глава выясняет окончательно понятие о законе распределения, опираясь на материал, заготовленный в предшествующих частях, откуда видно, что это понятие развивалось исторически; в этом историческом порядке ведется и все изложение главы I последней части курса. Сначала рассмотрен закон распределения, обычно называемый Гауссовым, так как этот закон лежит в основе его теории ошибок; но так как он получается, как первое грубое приближение для формулы бинома, и обычно это первое приближение применяется во всевозможных вопросах биологии, антропологии, экономики и т. д., то он получил название нормального закона. После этого нормального закона идет закон биномиальный, так как он может быть получен в виде второго приближения к той же формуле бинома. Формула биномиального закона нашла себе неожиданно применение даже в вопросе о подоходном налоге, где она была найдена актуарием Парето эмпирическим путем. Следующее, значительно большее обобщение нормального закона представляет собою содержание обширной теории Пирсона, который вывел ее, рассматривая задачу теории вероятностей, составляющую непосредственное обобщение той задачи, которая приводит к формуле бинома. Эта задача рассмотрена в главе III части 1-й настоящей книги и приводит вместо биномиального ряда к другому ряду, принадлежащему к числу гипергеометрических. Поэтому формулы Пирсона могут быть названы формулами гипергеометрического распределения. О теории Пирсона в настоящем курсе можно было дать лишь беглое понятие ¹⁾.

Следующая глава II части 3-й содержит краткое изложение теории корреляций. Эта теория в последнее время получает большое применение в области биологии, антропологии, экспериментальной психологии и т. д. Обычно ее излагают, подходя от практических вопросов, при чем она остается почти вне связи с теорией вероятностей, и неясно, почему речь идет почти только о так называемой прямолинейной корреляции. В настоящем курсе сделана попытка построить теорию корреляций, как обобщение теории кривых распределения для случая двух переменных и опираясь на теорию вероятностей. Если ограничиться случаем, соответствующим формуле нормального распределения, то сами собой явятся формы прямолинейной корреляции. Такой способ изложения читатель найдет в

¹⁾ Желающие с нею ознакомиться полнее найдут ее изложение в моей отдельной работе: „Кривые распределения и составление для них интерполяционных формул по способам Пирсона и Врунса“. Москва, 1922.

главе II части 3-й этого курса. Корреляция криволинейная является обобщением ненормального закона распределения для двух переменных. О ней дано только небольшое понятие, чтобы не увеличивать сверх меры пределы курса; по той же причине опущено рассмотрение случая корреляции для многих переменных ¹⁾.

Последняя глава курса занята небольшим очерком теории статистических рядов и их устойчивости. Она стоит несколько в стороне от других частей курса и примыкает, главным образом, к теории Лексиса. В ней есть еще вопросы не вполне выясненные. Я ограничился изложением наиболее существенного и наиболее простого.

Так как весьма возможно, что многие читатели интересуются не всеми теоретическими отделами и желали бы скорее перейти к математической статистике, то укажу здесь те отделы, которые они могли бы опустить, не утрачивая связи в понимании дальнейшего. Таким образом могут быть опущены:

1. — В части 1-й, гл. I, § 6. Различные косвенные приемы вычисления вероятностей.
2. — 1, гл. I, § 7. О геометрических вероятностях.
3. — 1, гл. II, § 3. Петербургская игра.
4. — 1, гл. II, § 4. О нравственном ожидании.
5. — 1, гл. III, § 15. Теорема Чебышева.
6. — 2, гл. III вся. Способ наименьших квадратов.

Заканчивая печатание настоящего курса, приношу глубокую благодарность В. М. Турбину за его дружеские советы, которыми я воспользовался при обработке текста для достижения возможной ясности понимания его читателями, а также за содействие тщательности издания книги и за большую помощь при проверке корректур.

¹⁾ Гораздо более полное изложение теории корреляций со включением названных выше вопросов приведено мною в особом сочинении „Теория корреляций“, принятом к напечатанию Государственным Издательством.

ВВЕДЕНИЕ

ВВЕДЕНИЕ

§ 1. О случайных явлениях.

Одна из основных особенностей нашего мышления состоит в том, что мы не можем представить себе никакого события или явления без причины, при чем под причиною мы подразумеваем наличность такого условия или — чаще — совокупности таких условий, без которых данное событие или явление никогда не наступает, а при которых оно непременно совершается.

Пусть, например, речь идет о возникшем пожаре. Для наступления такого события необходима совокупность двух условий: 1) присутствие горючего материала, 2) нагревание хотя бы части этого материала до температуры воспламенения. Наличность этих двух условий — причина данного события — пожара.

Всякое происшедшее событие или само по себе, или в совокупности с другими событиями делается причиною новых событий, которые в свою очередь обуславливают новые следствия, и т. д. без конца. Эта бесконечная цепь причин и следствий составляет сущность всего происходящего в окружающем нас мире: она составляет сущность жизни этого мира.

Однако, несмотря на ясно сознаваемую нами необходимость объяснения всякого события предшествующей ему причиной, мы постоянно употребляем выражения „случайное явление“ или „случайное событие“, указывая тем самым, что мы отказываемся определить его причину. С первого взгляда даже может показаться, что мы считаем случайное явление беспричинным. Например, говорят, что горючий материал, сложенный на дворе, воспламенился от спички, случайно брошенной прохожим по неосторожности, без всякого умысла. Или говорят, что двое знакомых встретились случайно на улице города, расположенного вдали от места их постоянного жительства, прибыв каждый по своему делу и без всякого предварительного соглашения между собою.

Всматриваясь в характер этих так называемых случайных событий, мы замечаем, что они вовсе не беспричинные, но причины их слагаются из таких условий, связь которых между собою очень отдаленная.

ленная и для нас неясная. Так, в первом из приведенных двух примеров причиной возникшего пожара была наличность двух событий: 1) присутствия на дворе горючего материала, 2) присутствия неосторожного прохожего, бросившего спичку. Каждое из этих двух событий есть следствие предшествующих причин, которые имеют между собою отдаленную связь, но проследить ее мы не в состоянии: она как бы выходит из поля нашего зрения.

Французский математик Лаплас, много сделавший как для уяснения основ теории вероятностей, так и для разработки самой теории, высказал свой взгляд на случайные явления в таких словах: „Ум, которому были бы известны для какого-либо данного момента все силы, одушевляющие природу, и относительное положение всех ее составных частей, если бы вдобавок он оказался достаточно обширным, чтобы подчинить эти данные анализу, — обнял бы в одной формуле движения величайших тел вселенной наравне с движениями легчайших атомов; не осталось бы ничего, что было бы для него недостоверно, и будущее, так же как и прошедшее, предстало бы перед его взором“ *).

Наш ум очень далек от такого всеобъемлющего захвата, о котором говорит Лаплас: он в состоянии ввести в рассмотрение лишь ограниченное число данных и найти ближайшие следствия из них; все остальное выходит из его поля зрения. Эта неполнота наших сведений обуславливает для нас впечатление случайности явления.

Если бы случайные явления были беспричинными, то, понятно, к ним нельзя было бы применить математические методы и формулы.

Так как мы заранее знаем, что случайные явления имеют свои причины, только скрытые от нас, то наше желание и их подчинить математическим формулам становится вполне законным, хотя вперед нельзя сказать, удастся ли осуществить такое желание.

В этом отношении много поучительного дает нам статистика. Она, например, показывает, что, когда государство живет установившейся жизнью, то для каждой его области рождаемость, смертность, число браков, число преступлений и т. д. из года в год почти не меняется, особенно если брать процентное отношение к населению. Мало того, оказывается, что число мужских рождений всегда несколько превышает число рождений женских (на 100 девочек рождается около 104 мальчиков); распределив число преступлений по роду орудия, употребленного преступником, заметим и тут значительное постоянство. Близким к постоянству оказывается даже число писем, опускаемых ежегодно в почтовые ящики без адреса или без марки.

*) Laplace. Essai philosophique sur les probabilités (Лаплас. Опыт философии теории вероятностей). Перевод под редакцией А. К. Власова. Москва, 1908 г.

Уже отсюда видно, что ко многим случайным явлениям применима количественная оценка.

Приложение математического метода к изучению случайных явлений составляет предмет теории вероятностей.

§ 2. Определение величины вероятности случайного явления.

Понятие о вероятности одно из элементарных: мы постоянно в разговорном языке говорим, что тот или другой слух вероятен, что одно известие вероятнее другого, что выигрыш в ту или другую игру или лотерею вероятен или невероятен и т. д.

Сущность понятия „вероятность“ мы не можем определить — так же, как в физике нельзя определить сущность материи, силы, тепла, электричества, магнетизма и т. д. Но это не мешает нам искать способов измерения вероятности совершенно так же, как в физике измеряется масса, сила, количество тепла, электричества и т. д.

В науках опытных, желая применить математический метод, прежде всего стараются экспериментально установить законы изучаемого явления в самых простых условиях, а затем, путем математических выкладок, выводят возможно полно следствия из подмеченных законов.

Лаплас в упомянутом выше сочинении говорит, что теория вероятностей есть здравый смысл, сведенный к исчислению.

Поэтому естественно, что в основу ее выводов должна быть положена экспериментальная формула, найденная из наблюдения, как судит здравый смысл о вероятностях случайных явлений.

Начнем с простого примера. Пусть имеется урна, в которую положено 10 шаров одинакового радиуса и одинаковых на осязание; шары в урне тщательно перемешаны; урна имеет сверху отверстие настолько узкое, что может пропустить только руку вынимающего шар, но не позволяет видеть положенных шаров. Пусть из числа 10 положенных шаров 8 белых и 2 черных.

Из такой урны хотят вынуть один шар. Спрашивается, какого цвета шар вероятнее ожидать: белого или черного.

Ответ по здравому смыслу, конечно, будет такой: появление белого шара вероятнее, чем черного потому, что в урне белых шаров больше, чем черных.

Пусть имеются две урны: в обеих по 10 шаров, одинаковых по радиусу и на осязание и тщательно перемешанных; но в первой урне 8 белых шаров и 2 черных, а во второй 4 белых и 6 черных. Спрашивается, для какой из этих двух урн появление белого шара вероятнее.

Конечно, последует ответ, что вероятнее появление белого шара для первой урны, потому что там белых шаров больше, а общее число шаров одинаково в обеих урнах.

На вопрос: во сколько раз появление белого шара вероятнее для первой урны, чем для второй, мы, конечно, ответим: вдвое, потому что в первой урне белых шаров вдвое больше, чем во второй.

Обобщая сказанное, возьмем такую задачу: даны две урны; в обеих по одинаковому числу шаров m , одинаковых по радиусу и на осязание; в первой из них m белых и n черных, во второй m' белых и n' черных шаров.

На вопрос, для которой из этих двух урн появление белого шара вероятнее, — мы, конечно, скажем: для той урны, где белых шаров больше. На вопрос о том, во сколько раз для этой урны вероятность появления белого шара больше, мы ответим: во столько раз, во сколько число белых шаров в этой урне больше, чем в другой.

Короче говоря, мы находим, что при данных условиях вероятности появления белого шара из урны находятся между собою в таком же отношении, как числа белых шаров.

Не входя в определение сущности понятия о вероятности, мы можем обозначить величины вероятностей вынужтия белого шара из данных двух урн буквами p и p' . В таком случае выраженное выше суждение по здравому смыслу представится формулой:

$$\frac{p}{p'} = \frac{m}{m'} \dots \dots \dots (1)$$

В еще более общем случае, чем только-что рассмотренный, нам обыкновенно удастся привести условия к той же схеме: только вместо шаров придется говорить о шансах или статочностях. Статочности, благоприятствующие событию, соответствуют белым шарам в предшествующей задаче; статочности, не благоприятствующие событию, соответствуют черным шарам. Как шары должны быть одинакового радиуса, одинаковы на осязание и тщательно перемешаны, так в общем случае статочности должны быть намечены равно-возможные и при том всевозможные и взаимно исключающие друг друга.

Рассмотрим примеры. Пусть речь идет об игре в орлянку.

Бросается монета, на одной стороне которой орел, а на другой — решетка. Если при бросании этой монеты вскрыется орел, то игрок выигрывает некоторую сумму. Здесь две статочности: 1) выпадение орла, 2) выпадение решетки. Первая из них благоприятствует игроку, вторая ему не благоприятствует. Если монета — правильный цилиндр, сделанный из однородного материала, то обе статочности равновозможны, они взаимно исключают друг друга и суть всевозможные. Еще пример: бросают игорную кость в виде куба; ее шесть граней перенумерованы очками от 1 до 6. Если вскрыется

грань с одним очком, то игрок выигрывает; в противном случае он проигрывает. Здесь 6 статочностей, из которых одна благоприятствует игроку, а пять не благоприятствуют ему. Если кость представляет собою математически правильный куб, сделанный из однородного вещества, то статочности удовлетворяют сказанным выше условиям и т. д.

Рассуждения, которые мы делали при выводе формулы (1) в рассмотренном примере с урнами, остаются справедливыми и в общем случае. Поэтому мы можем сказать: вероятности p и p' двух событий относятся между собою, как числа благоприятствующих им статочностей, при условии, чтобы в обоих случаях число всех статочностей было одно и то же и чтобы статочности были равновозможны.

Формула (1) позволяет нам сравнивать вероятности между собою. Остается сделать выбор единицы вероятности, чтобы быть в состоянии вероятности измерять.

Выбор единицы меры, вообще говоря, произволен: он должен быть сделан лишь целесообразно.

Обратимся к задаче с урнами. Примем за единицу вероятность вынуть белого шара из такой урны, где все шары белые. Иными словами, мы принимаем $p' = 1$, когда $m' = \mu$. В таком случае формула (1) примет вид:

$$\frac{p}{1} = \frac{m}{\mu},$$

откуда

$$p = \frac{m}{\mu} \dots \dots \dots (2)$$

Когда все шары в урне белые, то вынутые из нее белого шара не подлежит сомнению, оно достоверно. Следовательно, мы выбрали за единицу вероятности — вероятность события достоверного.

При таком выборе вероятности имеет место равенство (2). Оно показывает, что вероятность вынуть белого шара из урны, где всего μ шаров, из которых m белых, выражается отношением $\frac{m}{\mu}$ числа белых шаров к числу всех шаров в урне.

Так как формула (1) справедлива для всякого случая, когда удастся установить всевозможные статочности, и эти статочности равновозможны, то и формула (2) будет справедлива во всех таких случаях, если мы примем за единицу вероятность события достоверного.

Итак, вероятность измеряется отношением числа статочностей, благоприятствующих событию, к числу всех статочностей, при условии, чтобы эти статоч-

ности были равновозможные, всевозможные и взаимно исключающие друг друга.

Событие, состоящее в том, что данное событие не состоялось, называется ему противоположным. Так, появление черного шара есть событие, противоположное появлению белого в первом из рассмотренных примеров. Выпадение решетки есть событие, противоположное выпадению орла. В примере с игорной костью событие, противоположное вскрытию грани с одним очком, есть вскрытие какой-либо другой грани, и т. д.

Если мы обозначим буквою q вероятность события, противоположного тому событию, у которого вероятность равна p , то можно вычислить q по формуле, подобной (2). Разница в том, что число статочностей, благоприятствующих этому противоположному событию есть n ; число же всех статочностей по-прежнему есть μ . Следовательно:

$$q = \frac{n}{\mu} \dots \dots \dots (3)$$

Так как $m + n = \mu$, то сложив формулы (2) и (3), получим:

$$p + q = 1, \dots \dots \dots (4)$$

т. е. сумма вероятностей противоположных событий равна 1.

Из формулы (2) видно, что вероятность события выражается положительною дробью, меньшею единицы. Только событие достоверное имеет вероятность, равную 1.

Если число $m = 0$, т. е. если нет ни одной статочности, благоприятствующей событию, то событие невозможно. Формула (2) показывает, что вероятность его равна 0.

Относительно сделанного выше выбора единицы вероятности необходимо сделать некоторое замечание. Говоря строго, такой выбор неправилен, потому что между вероятностью и достоверностью существует различие не только количественное, но и качественное. Чтобы в этом убедиться, возьмем примеры.

Пусть речь идет о лотерее, где на 100 000 билетов есть только один выигрышный с очень большим выигрышем, а остальные билеты — пустые. Вероятность выигрыша по формуле (2) равна $\frac{1}{100\,000}$; она очень мала; и тем не менее участие в лотерее может представлять интерес для участников. Если же все билеты будут пустыми, то участие в лотерее утратит всякий смысл.

Другой пример: пусть в урну положено 99 999 белых шаров и 1 черный. Вероятность появления белого шара будет равна 0,99 999; она очень близка к единице; тем не менее, если от появления черного шара будет зависеть очень значительный ущерб для нашего состояния или даже жизни, то, приступая к опыту с такою урною, мы

будем испытывать душевное волнение. Если же единственный черный шар в урне будет заменен белым, то волнение исчезнет.

Ясно, что в этих условиях различие не только количественное, но и качественное.

Тем не менее есть основание, оправдывающее сделанный выше выбор единицы вероятности.

Цело в том, что в действительности во внешнем мире мы ничего не найдем абсолютно достоверного, если не считать искусственных примеров, как приведенный выше пример с урной, где все шары белые.

Наиболее установленные законы природы, как закон всемирного тяготения, выведены из наблюдения, и не исключена возможность их изменения со временем. Появление солнца в ближайший следующий день тоже нельзя считать достоверным. Многие представления, которые прежде казались бесспорными, например, неподвижность земли, плоский вид ее поверхности и т. п., не только оказались ложными, но теперь мы с трудом могли бы стать на точку зрения древних, которым это казалось достоверным.

Между вероятностью наиболее прочно установленных законов природы и вероятностью других событий качественного различия не окажется.

Обыкновенно, во всякой опытной науке, пользующейся методами теории вероятностей, заранее указывают ту дробь, близкую к 1, напр., 0,99 или 0,999, или 0,9999 и т. д., которую мы приближенно принимаем за 1. Всякую правильную дробь, большую ее, считают за 1, а событие, вероятность которого выражается такою дробью, считают достоверным.

§ 3. Вероятность объективная и субъективная.

В задачах чисто теоретического характера можно предполагать все статистически равновероятными и вычислять вероятность по формуле (2). Такая вероятность соответствует условиям задачи и называется объективной.

На практике осуществить это условие никогда не удастся: как бы тщательно мы ни готовили шар, поверхность его не будет математической сферою; двух шаров совершенно одинакового радиуса приготовить невозможно. При игре в орлянку приходится предполагать, что монета сделана из однородного металла, имеет вид математического цилиндра и бросается на математически правильную горизонтальную плоскость. Не говоря уже о том, что математически правильных поверхностей осуществить невозможно на практике, сама монета заведомо имеет выпуклые изображения на своих сторонах.

Во всех случаях практического применения теории вероятностей нам приходится идеализировать задачу, допуская, что поверхности

математически правильны, материал однороден и т. д. Величина вероятности, вычисленная для такой упрощенной задачи, не будет в точности соответствовать условиям задачи реальной; она служит лишь приближенной величиной истинного решения задачи. Ее мы назовем субъективной вероятностью, в отличие от истинной, объективной вероятности, которую определить мы не в состоянии.

По мере того, как наши сведения о действительных условиях данной задачи увеличиваются, меняется субъективная вероятность, приближаясь к вероятности объективной.

Таким образом, субъективная вероятность есть наша оценка истинной вероятности на основании тех сведений, которые мы имеем об условиях данного опыта. Если таких сведений мало, то и оценка наша может быть лишь грубо приближенная. Бывают крайние случаи, когда об условиях события ничего не известно, кроме того, что оно может состояться, но может и не состояться. В таких случаях, в виде первого приближения, приходится оценивать вероятность события дробью $\frac{1}{2}$ (принцип отсутствующего основания, Prinzip des mangelnden Grundes). Но даже исходя из такого грубого первого приближения, путем постепенных поправок можно иногда получить хороший результат.

§ 4. Вероятность *a priori* и *a posteriori*.

В предшествующих рассуждениях мы предполагали, что условия задачи (реальной или идеализированной) нам заранее даны: исходя из этого предположения, руководствуясь здравым смыслом, мы вывели формулу (2) величины вероятности. Такая вероятность на основании заранее сообщенных нам условий называется вероятностью *a priori*.

Но рядом с этим смыслом слова „вероятность“ в обыденной речи можно подметить употребление того же слова в смысле от него отличном.

Пусть мы производим опыты или наблюдения над некоторым событием, условия которого нам заранее не даны. Если при этих опытах событие появляется часто, то мы ждем его появления и при дальнейших испытаниях и считаем его вероятным; в противоположном случае мы считаем его маловероятным или невероятным.

Пусть, например, при опытах над некоторым явлением замечено, что при каждых 100 испытаниях появление события наблюдается в среднем в 60 случаях. Отношение $\frac{60}{100} = \frac{3}{5}$ показывает, как часто наблюдается появление события при условиях данного опыта, хотя эти условия нам заранее и не даны; они нам неизвестны.

Такую дробь мы называем частотой явления. Мы ожидаем, что при тех же условиях событие и в дальнейших испытаниях будет

наблюдаться с той же частотой. В обыденной речи мы этою частотой измеряем вероятность наблюдаемого события, придавая таким образом тому же слову „вероятность“ смысл совершенно отличный от того, который был рассмотрен выше. В этом новом смысле вероятность вычисляется на основании опыта. Она может быть названа вероятностью *a posteriori* (на основании последующих данных).

Вопрос о том, насколько близки между собою обе величины вероятности одного и того же события в двух различных смыслах, имеет большой теоретический интерес, но он имеет и громадную практическую важность.

Всякое событие должно иметь свою объективную вероятность *a priori*, но чаще всего эта величина нам неизвестна. Производя опыты, мы можем определить частоту этого события. Если бы мы могли судить, насколько частота способна уклониться от объективной вероятности, то были бы в состоянии определять объективную вероятность, хотя и приближенно, но с определеною точностью. В тех случаях, где частота окажется близкою к вероятности, мы можем принять ее за приближенную величину вероятности данного события и ввести в математические вычисления наравне с вероятностями других событий, найденными тем же путем или же вычисленными по начальным данным. Ко всем таким случаям будет применим тот метод исчисления вероятностей, который выводится сначала только для вероятностей *a priori* и который будет изложен дальше в I главе первой части.

На близости величин вероятности, вычисленных 1) по начальным данным и 2) из результатов опыта, основаны все практические применения теории вероятностей к наукам, пользующимся статистическим методом. Если бы эти величины не были между собою близки и если бы нельзя было оценивать степень их близости, то теория вероятностей утратила бы всякий практический интерес.

Решить вопрос о величине возможного уклонения частоты от вероятности можно теоретически на основании некоторой предпосылки. Начало таких исследований положено Яковом Бернулли в 1713 году.

Эти исследования составляют большую область в теории вероятностей; мы займемся ими в III главе первой части.

Но кроме теоретического решения вопроса, может быть сделана и опытная проверка: зная заранее условия какого-нибудь опыта, мы можем вычислить вероятность p некоторого события по формуле (2); а затем, произведя в действительности ряд опытов, вычислить по ним частоту того же события и сравнить ее с вероятностью p .

Такие опыты были много раз произведены и обнаружили, в согласии с тем, что показывает теория, что близость частоты и вероят-

ности увеличивается по мере увеличения числа опытов. Закон, утверждающий эту близость, получил поэтому название „закона больших чисел“.

Для уяснения сказанного рассмотрим результаты некоторых из таких опытов.

Следующая таблица заимствована из сочинения Пирсона*); она составлена на основании результатов, полученных различными исследователями над такими явлениями, где одинаковое число статистических благоприятствует обоим взаимно противоположным событиям:

следовательно, вероятность каждого из них равна $\frac{1}{2}$. В первом

столбце таблицы помещен род испытания, во втором столбце — частость одного из двух взаимно противоположных событий, в третьем столбце — число всех испытаний, в четвертом — лицо, производившее испытания.

Род испытания.	Частость.	Число испытаний.	Лицо, производившее испытания.
Рулетка	0,5015	16.141	Pearson.
„	0,5027	16.019	De Whalley.
Урна с шарами	0,504	4.096	Quetelet.
„	0,5011	10.000	Westergaard.
Монета	0,51	4.040	Buffon.
„	0,5005	4.092	De Morgan's pupil.
„	0,5004	8.178	Griffith.
„	0,5016	12.000	Pearson.
„	0,5005	24.000	„
Лотерея	0,50034	7.275	Westergaard.

Во всех этих случаях частость очень близка к вероятности 0,5.

Чтобы проверить, будет ли частость столь же близка к вероятности, когда вероятность имеет значение, далекое от 0,5, можно воспользоваться результатами подсчета, произведенного Чубером**) над вышедшими номерами лотереи в Праге и Брюнне. При каждой партии игры в лотерею, из колеса, содержавшего 90 номеров, вынимаются 5 номеров. Можно заранее (a priori) вычислить вероятность того, что в числе пяти вынутых номеров однозначных будет: нуль.

*) Pearson. The Chances of Death. London and New York. 1897. Пирсон — современный нам английский математик, очень много подвинувший вперед область приложений теории вероятностей к статистике и биологии.

**) Czuber. Wahrscheinlichkeitsrechnung. Leipzig und Berlin. 1908.

один, два, три, четыре, пять *). На основании опубликованных бюллетеней лотерей можно подсчитать частоту каждого из этих случаев и сравнить ее с соответствующей вероятностью. Результаты такого сравнения для лотерей в Праге и Брюнне приведены в следующей таблице:

Число однознач. номеров.	Вероят- ность.	Частоты для лотерей.	
		В Праге (число испы- таний 2854).	В Брюнне (число испы- таний 2703).
0	0,58293	0,58655	0,57899
1	0,34070	0,32656	0,34591
2	0,06989	0,07919	0,06881
3	0,00619	0,00735	0,00629
4	0,00023	0,00035	0,00000
5	0,00000	0,00000	0,00000
	1,00000	1,00000	1,00000

§ 5. Исторический очерк.

Основные понятия о вероятностях были известны еще в древности; так, в дигестах Юстиниана в 528 г. по Р. Х. есть закон о удовольствии, из которого видно, что римляне имели понятие о вероятности человеческой жизни.

В средние века в Италии существовали морские страховые общества. В начале XVII века Галилей делал попытки исчисления вероятностей ошибок при наблюдениях. В том же веке итальянец Тонтини предложил способ страхования жизни (так называемое тонтинное страхование).

Основание теории вероятностей было положено Паскалем и Ферматом в половине XVII века. Первая задача, решенная Паскалем, была предложена ему кавалером Де-Мере. Она заключалась в том, чтобы найти способ безобидно разделить между игроками ставку, если они хотят разделить ее раньше окончания игры. Современник Паскаля Гюйгенс собрал вопросы, решенные ранее, и издал в 1657 году первый трактат по теории вероятностей под заглавием: „De ratiociniis in aleae ludo“. В 1693 году Галлей составил первую тщательно обработанную таблицу смертности для города Бреславля. Следующее выдающееся явление в теории вероятностей относится к 1713 году, когда появилось в Базеле сочинение Якова Бернулли под заглавием „Ars conjectandi“. Это сочинение было издано спустя 7 лет после

*) Эти вычисления приведены ниже в I главе первой части.

смерти автора его племянником, Николаем Бернулли. В 4-м томе этого сочинения находится знаменитая теорема Якова Бернулли, касающаяся уклонения частоты от вероятности. Там же в Базеле Николай Бернулли издал свою диссертацию „De arte conjectandi in jure“. Это сочинение включает приложение теории вероятностей к судопроизводству. Монтмор и Моавр сделали дальнейшие усовершенствования в теории вероятностей. Первый написал сочинение „Essai d'analyse sur les jeux de hasard“. Сочинение это интересно тем, что в нем находится переписка с Николаем Бернулли, между прочим, по поводу Петербургской игры, решение которой приводило к парадоксу и служило поводом к возражению против верности основ теории вероятностей. Моавр первый стал пользоваться предложенным им самим способом возвратных рядов. В сущности, этот способ есть в скрытой форме приложение исчисления конечных разностей к теории вероятностей. Затем Даниил Бернулли ввел понятие о так называемом нравственном ожидании, в противовес с обычно употребляемым понятием о математическом ожидании. Пользуясь этим понятием, он объяснил происхождение упомянутого выше парадокса Петербургской игры, названной так потому, что переписка по этому вопросу была помещена в Записках Петербургской Академии Наук. Кроме того, Даниил Бернулли первый поставил вопрос о так называемых вероятностях гипотез *a posteriori*, но этот вопрос в первый раз был разрешен Бейесом в 1764 году; решение его выражается так называемой теоремой Бейеса. Из других математиков, занимавшихся в ту же эпоху теорией вероятностей, необходимо отметить Даламбера и Кандорсе, построившего теорию решений по большинству голосов. В 1806 году Лежандр предложил так называемый способ наименьших квадратов; этот способ дает возможность получать наиболее надежные результаты из опытных данных. Но Лежандр предложил этот способ без математических оснований. Гаусс построил полную и стройную теорию ошибок измерений; он дал формулу вероятности ошибок измерений, из которой способ наименьших квадратов естественным образом развивается.

Лаплас написал прекрасно обработанный трактат по теории вероятностей под заглавием „Traité analytique des probabilités“, где он пользуется методом функций-генератрис. Им теория вероятностей была приложена к небесной механике и усовершенствована теория ошибок Гаусса в применении к геодезии. Он же написал сочинение „Essai philosophique sur les probabilités“, в котором все наиболее существенное в теории вероятностей изложено без помощи формул и чертежей. Пуассон написал сочинение „Mémoires sur la probabilité du tir à la cible“, выводы которого блестяще были подтверждены опытами французских артиллеристов. Он же дал теорему, представляющую собою обобщение теоремы Якова Бернулли. Кроме того,

Пуассон обработал математическую теорию судопроизводства. У нас в России значительный успех в теории вероятностей достигнут трудами академика П. Л. Чебышева, который дал носящую его имя теорему; эта теорема содержит в себе, как частные случаи, теоремы Пуассона и Бернулли.

В последние 75 лет развилось и сделало громадные успехи приложение теории вероятностей к статистике и биологии. Эта научная область была основана бельгийским математиком и антропологом Кетле, опубликовавшим систематическое изложение своих мыслей в 1846 году в „Lettres sur la théorie des probabilités“. Он обратил внимание на то, что результаты подсчетов и измерений в области антропологии, биологии и социальных наук имеют такой же характер, как результаты точных измерений в физике, астрономии и т. д. Следовательно, к этой области должна быть применима созданная Гауссом теория ошибок измерения. Естественно, что он и применял в своих работах формулу Гаусса, которая дала на первых порах блестящие результаты. Но уже сам Кетле высказывал мысль, что явления, подлежащие изучению в новой области, могут быть сложнее, чем при простых измерениях, и могут потребовать более сложных формул.

Первая попытка обобщения формулы Гаусса, нахождение так называемых формул распределения, была сделана в Германии Фехнером и его последователем Брунсом. Но главная заслуга в этой области принадлежит английским ученым: Гальтону и, в особенности, Пирсону. Последний основал современную нам школу биометриков, работами которой создавалась обширная литература, содержащая в себе как изложение математических методов, так и их приложений к различным вопросам биологии, статистики, антропологии и обществоведения.

Основной для статистики вопрос об устойчивости статистических рядов был впервые строго поставлен и разрешен Лексисом. В связи с дальнейшей разработкой этого вопроса находится предложенный Борткевичем закон малых чисел, касающийся особенностей явлений с малою вероятностью.

ЧАСТЬ ПЕРВАЯ
ОСНОВЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

ГЛАВА ПЕРВАЯ

Приемы вычисления вероятностей.

§ 1. Примеры непосредственного вычисления вероятностей.

В простейших случаях, сосчитав число всех статочностей ω в данном вопросе и число статочностей m , благоприятствующих рассматриваемому событию E , и убедившись, что эти статочности равно-возможны, всевозможны и взаимно исключают друг друга, мы непосредственно вычисляем вероятность p события E по формуле, указанной во введении:

$$p = \frac{m}{\omega} \dots \dots \dots (1)$$

Вероятность q события F , противоположного E , удовлетворяет условию:

$$p + q = 1, \dots \dots \dots (2)$$

откуда:

$$q = 1 - p. \dots \dots \dots (3)$$

Так, при игре в орлянку вероятность выпадения орла $p = \frac{1}{2}$; при бросании игральной кости выпадение грани с очком 1 имеет вероятность: $p = \frac{1}{6}$.

Рассмотрим несколько более сложных задач, где непосредственно можно подсчитать числа ω и m и затем найти p по формуле (1).

При этих подсчетах часто приносит большую пользу известная из алгебры теория соединений (так называемая комбинаторика).

Задача 1. Бросают одновременно две игральные кости, имеющие форму куба с гранями, перенумерованными очками от 1 до 6. Спрашивается, какова вероятность, что сумма выпавших очков будет равна 5.

Каждая грань 1-ой кости при бросании может выпасть в комбинации с каждой гранью 2-ой кости; следовательно, число всех комбинаций будет равно:

$$p = 6 \cdot 6 = 36.$$

Таково число всех статочностей.

Статочности, благоприятствующие данному событию, представляют комбинации таких граней, для которых сумма выпавших очков равна 5. Их можно расположить в следующую таблицу:

1-я кость.	2-я кость.
1 очко	4 очка
2 "	3 "
3 "	2 "
4 "	1 "

Число статочностей, благоприятствующих данному событию, $m = 4$.

Вероятность события:

$$p = \frac{4}{36} = \frac{1}{9}.$$

Задача 2 (Задача Даламбера)*. Бросают 2 раза подряд монету. Спрашивается, какова вероятность, что хоть один раз выпадет орел.

Все статочности, возможные в данной задаче, располагаем в следующей таблице:

1-е бросание:	2-е бросание:
орел,	орел,
орел,	решетка,
решетка,	орел,
решетка,	решетка.

Всего статочностей 4; из них событию, указанному в условии задачи, благоприятствуют три первые. Вероятность события:

$$p = \frac{3}{4}.$$

Даламбер при решении этой задачи сделал ошибку. Он имел в виду вероятность на выигрыш того игрока, который держал бы пари, что при двукратном бросании монеты хоть один раз выпадет орел.

*) Даламбер (D'Alembert) — французский математик и энциклопедист XVIII века.

Статочности, намеченные Даламбером, можно расположить в следующую таблицу:

1-е бросание.	2-е бросание.
орел,	излишне, так как игрок уже выиграл,
решетка,	орел,
решетка,	решетка.

Всего 3 статочности, из которых две первые благоприятствуют игроку. Отсюда Даламбер заключил, что вероятность на выигрыш равна $\frac{2}{3}$.

Это решение ошибочно потому, что статочности, помещенные в последней таблице не равновозможны: первая из них совмещает в себе две статочности, из которых каждая равновозможна с каждой из двух остальных.

Задача 3. В урну положено α белых и β черных шаров, всего $\gamma = \alpha + \beta$ шаров. Из нее вынимают зараз γ' шаров. Спрашивается, какова вероятность, что в числе вынутых шаров будет α' белых и β' черных, при чем $\alpha' + \beta' = \gamma'$.

Вынутая группа в γ' шаров составляет одну из групп сочетаний, которые можно составить из всех γ элементов по γ' элементов в группе.

Число всех таких сочетаний обозначается символом $C_{\gamma}^{\gamma'}$ и выражается так:

$$C_{\gamma}^{\gamma'} = \frac{\gamma (\gamma - 1) (\gamma - 2) \dots (\gamma - \gamma' + 1)}{1 \cdot 2 \cdot 3 \dots \gamma'}$$

Таково выражение числа μ всех статочностей в данной задаче.

Статочности, благоприятствующие рассматриваемому событию, соответствуют таким группам в γ' шаров, в которые входят α' белых и β' черных шаров. Отдельно взятые α' белых шаров представляют группу сочетаний, составленных из α белых шаров урны. Число таких сочетаний равно $C_{\alpha}^{\alpha'}$. По такой же причине β' черных шаров представляет собою группу сочетаний, составленных из β черных шаров урны. Число таких сочетаний равно $C_{\beta}^{\beta'}$. Число всех групп, состоящих из α' белых и β' черных шаров, выразится произведением

$$C_{\alpha}^{\alpha'} \cdot C_{\beta}^{\beta'}$$

Таково число статочностей, благоприятствующих событию.

Вероятность события выразится формулой:

$$p = \frac{C_{\alpha}^{\alpha'} C_{\beta}^{\beta'}}{C_{\gamma}^{\gamma'}}.$$

Рассмотрим числовой пример: пусть

$$\alpha = 6, \beta = 4, \gamma = 10, \alpha' = 3, \beta' = 0, \gamma' = 3.$$

Иными словами, из урны, куда положено 6 белых шаров и 4 черных шара, вынимают зараз 3 шара. Спрашивается, какова вероятность, что все 3 вынутые шара будут белыми.

Ответ задачи таков:

$$p = \frac{C_6^3}{C_{10}^3} = \frac{6 \cdot 5 \cdot 4}{10 \cdot 9 \cdot 8} = \frac{1}{6}.$$

Задача 4. В колесе лотереи находится 90 номеров: от 1 до 90; в каждую игру из них вынимают зараз 5. Спрашивается, какова вероятность, что в числе 5 вынутых номеров число однозначных номеров будет равно α , при чем α может равняться 0, 1, 2, 3, 4, 5.

Вынутые 5 номеров представляют собою сочетание из 90 элементов по 5. Число таких сочетаний

$$C_{90}^5.$$

В числе 90 номеров, находящихся в колесе лотереи, однозначных всего 9, остальные 81 двузначные.

Если при вынутии 5 номеров из колеса в числе их оказалось α однозначных, а остальные $5 - \alpha$ двузначные, то первые образуют сочетание из 9 элементов по α , а вторые сочетание из 81 элемента по $5 - \alpha$. Числа таких сочетаний соответственно равны: C_9^{α} и $C_{81}^{5-\alpha}$.

Следовательно, число групп номеров, содержащих в себе α однозначных и $5 - \alpha$ двузначных выразится произведением:

$$C_9^{\alpha} C_{81}^{5-\alpha}.$$

Обозначив буквою p_{α} вероятность того, что при вынутии 5 номеров в числе вынутых будет α однозначных, мы найдем p_{α} , разделив $C_9^{\alpha} C_{81}^{5-\alpha}$ на C_{90}^5 :

$$p_{\alpha} = \frac{C_9^{\alpha} C_{81}^{5-\alpha}}{C_{90}^5}.$$

Положив поочередно $\alpha = 0, 1, 2, 3, 4, 5$, найдем $p_0, p_1, p_2, p_3, p_4, p_5$.

Числовые величины этих выражений приведены выше во 2-м столбце 2-й таблицы параграфа 3-го Введения.

Задача 5. В королевской лотерее, существовавшей во Франции до конца XVIII века, условия игры были следующие. В колесе лотереи было 90 номеров, из которых в каждую партию вынимали 5. Каждый игрок перед партией вносил в кассу какую он хочет ставку и записывал по своему желанию один, два, три, четыре или пять номеров. Если все номера, записанные игроком, оказывались вышедшими в числе 5 вынутых номеров, то он получал из кассы свою ставку, увеличенную:

в 15 раз, если им был записан 1 номер.

" 270 " " " " " 2 "

" 5500 " " " " " 3 "

" 7500 " " " " " 4 "

(случаев записи 5 номеров на практике не встречалось).

Спрашивается, какова вероятность на выигрыш игрока, записавшего α номеров.

Обозначим ее через p_α , при чем $\alpha = 1, 2, 3, 4, 5$.

Число всех статочностей так же, как и в предшествующей задаче, равно:

$$C_{90}^5.$$

Статочности, благоприятствующие игроку, записавшему α номеров, соответствуют таким группам из 5 вынутых номеров, в которые входят α номеров, записанных игроком и, кроме того, 5 — α каких-нибудь номеров из числа остальных 90 — α , не записанных. Эти последние составят группу сочетаний из 90 — α элементов по 5 — α ; число их:

$$C_{90-\alpha}^{5-\alpha}.$$

Искомая вероятность равна отношению:

$$p_\alpha = \frac{C_{90-\alpha}^{5-\alpha}}{C_{90}^5}.$$

Давая α последовательно значения: $\alpha = 1, 2, 3, 4, 5$, найдем:

$$p_1 = \frac{C_{89}^4}{C_{90}^5} = \frac{1}{18},$$

$$p_4 = \frac{C_{86}^1}{C_{90}^5} = \frac{1}{511038},$$

$$p_2 = \frac{C_{88}^3}{C_{90}^5} = \frac{2}{801},$$

$$p_5 = \frac{1}{C_{90}^5} = \frac{1}{43949268}.$$

$$p_3 = \frac{C_{87}^2}{C_{90}^5} = \frac{1}{11748}.$$

Нижe мы будем рассматривать условия безобидности игр; тогда обнаружится, что игрок, имеющий вероятность на выигрыш, равную p , при безобидности игры должен был бы в случае выигрыша получить свою ставку, увеличенную в $\frac{1}{p}$ раз.

Поэтому надо сказать, что королевская лотерея была невыгодна для игроков, в особенности для тех из них, которые записывали много номеров.

§ 2. Теорема сложения вероятностей.

Пусть некоторому событию E благоприятствуют статочности $E_1, E_2, E_3, \dots, E_m$ и не благоприятствуют статочности: $F_1, F_2, F_3, \dots, F_n$.

Пусть эти статочности суть всевозможные и взаимно исключающие друг друга в том смысле, что если одно из этих событий имеет место, то ни одно из остальных иметь места не может; вместе с тем непременно одно из событий $E_1, E_2, E_3, \dots, E_m$ или $F_1, F_2, F_3, \dots, F_n$ должно иметь место.

Сами же эти статочности могут быть и неравновозможными.

Пусть нам удалось подразделить все статочности, как благоприятствующие, так и неблагоприятствующие, на статочности более мелкие, между собою равновозможные. Пусть при этом статочность E_1 подразделилась на m_1 мелких статочностей, статочность E_2 на m_2 мелких статочностей, ..., статочность E_m на m_m мелких статочностей; подобным же образом пусть и $F_1, F_2, F_3, \dots, F_n$ подразделяются соответственно на $n_1, n_2, n_3, \dots, n_n$ мелких статочностей.

Пусть:

$$m_1 + m_2 + m_3 + \dots + m_m = M,$$

$$n_1 + n_2 + n_3 + \dots + n_n = N,$$

$$M + N = p.$$

В таком случае можно сказать, что событию E благоприятствуют M статочностей и не благоприятствуют N статочностей, причем эти статочности равновозможны, всевозможны и взаимно исключают друг друга. Поэтому вероятность p события E выразится формулою:

$$p = \frac{M}{p},$$

или

$$p = \frac{m_1}{p} + \frac{m_2}{p} + \frac{m_3}{p} + \dots + \frac{m_m}{p}.$$

Но величины:

$$\frac{m_1}{p}, \frac{m_2}{p}, \frac{m_3}{p}, \dots, \frac{m_m}{p},$$

суть вероятности событий $E_1, E_2, E_3, \dots, E_m$, т.-е. вероятности статочностей, благоприятствующих событию E . Обозначая эти вероятности буквами:

$$p_1, p_2, p_3, \dots, p_m,$$

будем иметь:

$$p = p_1 + p_2 + p_3 + \dots + p_m, \dots \dots \dots (4)$$

т.-е. вероятность события равна сумме вероятностей всех благоприятствующих ему статочностей.

Это — так называемая теорема сложения вероятностей.

Доказательство ее основано на постулате о том, что статочности подразделимы на более мелкие статочности, между собою равновозможные.

Рассмотрим примеры.

Пример 1. В урну положено m_1 красных, m_2 голубых и m_3 белых шаров, при чем:

$$m_1 + m_2 + m_3 = \mu.$$

Спрашивается, какова вероятность, что вынутый шар будет цветной, т.-е. красный или голубой.

Вероятности вынудия шара красного, голубого, белого соответственно суть:

$$p_1 = \frac{m_1}{\mu}, \quad p_2 = \frac{m_2}{\mu}, \quad p_3 = \frac{m_3}{\mu}.$$

По формуле (2) вероятность вынудия цветного шара выразится так:

$$p = p_1 + p_2 = \frac{m_1}{\mu} + \frac{m_2}{\mu}.$$

Справедливость этой формулы видна непосредственно, так как число цветных шаров в урне равно $m_1 + m_2$, а число всех шаров — μ .

Здесь на примере убеждаемся в согласии формулы с обыденным суждением по здравому смыслу.

Пример 2. Возьмем рассмотренную в § 1 задачу Даламбера. Намеченные им статочности таковы:

1-е бросание.

2-е бросание.

Орел.

Не производится.

Решетка.

Орел.

Решетка.

Решетка.

Игроку, державшему пари, что выпадет хоть один раз орел, благоприятствуют две первые статочности. Они не равновозможны: вероятность первой равна $\frac{1}{2}$, а второй — $\frac{1}{4}$. Вероятность, что орел

вскроется хоть один раз по формуле (4) равна:

$$\frac{1}{2} + \frac{1}{4} = \frac{3}{4}.$$

В этой задаче видно, как неравновозможные статочности разбиваются на более мелкие равновозможные: первая статочность разбивается на две:

1-е бросание.

Орел.

Решетка.

2-е бросание.

Орел.

Решетка.

§ 3. Теорема умножения вероятностей.

Сложным событием называется такое событие, которое состоит из совпадения двух или большего числа отдельных событий, называемых простыми.

Различаются два случая: 1) наступление или ненаступление отдельного простого события не зависит от того, произошли или не произошли другие события, составляющие вместе с ним сложное событие; 2) наступление или ненаступление простого события зависит от того, произошли или не произошли предшествующие ему простые события, составляющие вместе с ним сложное событие. Понятно, что в этом случае простые события располагаются в некотором определенном порядке во времени.

В первом случае простые события называются независимыми; во втором случае — зависимыми.

Рассмотрим примеры обоих случаев. Пусть речь идет о появлении шаров из урны, куда положено m белых и n черных шаров, причем $m + n = r$.

Пусть каждый раз, когда шар вынут из урны и отмечен его цвет, он кладется обратно в урну и смешивается с остальными шарами; в таком случае появление шара того или другого цвета при каждом новом испытании не находится в зависимости от того, какого цвета были шары, появившиеся при прежних испытаниях. Это — события независимые.

Если же каждый вынутый шар в урну не возвращается, а откладывается в сторону, то вероятность вынудия шара того или другого цвета при каждом испытании зависит от того, какие шары вышли раньше из урны. Здесь события зависимые между собою.

Начнем со случая независимых событий.

Пусть сложное событие состоит из двух простых, между собою независимых событий E_1 и E_2 . Пусть событию E_1 благоприятствует m_1 статочностей и не благоприятствует n_1 статочностей, причем $m_1 + n_1 = r_1$; пусть событию E_2 благоприятствует m_2 статочностей и не благоприятствует n_2 статочностей, при чем $m_2 + n_2 = r_2$.

Различными статочностями в событии E будут служить комбинации поочередно одной из статочностей события E_1 с одной из статочностей события E_2 . Число всех таких комбинаций выразится произведением:

$$\mu = \mu_1 \mu_2.$$

Различными статочностями, благоприятствующими событию E , будут служить комбинации одной из статочностей, благоприятствующих событию E_1 с одной из статочностей, благоприятствующих событию E_2 . Число всех таких комбинаций выразится произведением:

$$m = m_1 m_2.$$

Искомая вероятность p события E будет равна отношению $\frac{m}{\mu}$:

$$p = \frac{m_1 m_2}{\mu_1 \mu_2},$$

или

$$p = \frac{m_1}{\mu_1} \cdot \frac{m_2}{\mu_2}.$$

Так как $\frac{m_1}{\mu_1}$ есть вероятность p_1 события E_1 , а $\frac{m_2}{\mu_2}$ — вероятность p_2 события E_2 , то:

$$p = p_1 p_2 \dots \dots \dots (5)$$

Иными словами, вероятность сложного события, состоящего из двух независимых между собою простых событий, равна произведению вероятностей этих простых событий.

Поэтому теорема называется теоремой умножения вероятностей.

Если событие E состоит из трех независимых между собою простых событий E_1, E_2, E_3 с вероятностями p_1, p_2, p_3 , то мы можем соединить события E_1 и E_2 в одно сложное событие $(E_1 E_2)$ с вероятностью $p_1 p_2$ по формуле (5). Тогда событие E представится, как сложное событие, состоящее из двух простых: 1) $(E_1 E_2)$ с вероятностью $p_1 p_2$, 2) E_3 с вероятностью p_3 .

Вероятность события E по формуле (5) выразится так:

$$p = p_1 p_2 p_3.$$

Ясно, что таким же образом формула распространится на случай сложного события, состоящего из какого угодно числа независимых между собою простых событий.

Так как вероятности выражаются правильными дробями, то ясно, что вероятность сложного события всегда меньше вероятности каждого из составляющих его простых событий. Такая оценка совершенно со-

гласна с суждением по здравому смыслу: сложное событие есть совпадение нескольких случайных событий; ясно, что такое совпадение представляется нам менее вероятным, чем каждое из составляющих его простых событий в отдельности.

Переходим к случаю, когда сложное событие E состоит из двух событий, зависящих между собою, при чем сначала наступает E_1 , а затем E_2 .

Пусть относительно E_1 возможны μ_1 статочностей, из которых m_1 благоприятствуют событию E_1 , а n_1 ему не благоприятствуют. Пусть относительно E_2 возможны μ_2 статочностей, при чем это число одинаково, как в том случае, когда E_1 состоялось, так и в том случае, когда оно не состоялось; но пусть эти μ_2 статочностей распределяются различно на благоприятствующие и неблагоприятствующие событию E_2 в зависимости от того, состоялось ли событие E_1 или нет. Так, в приведенном выше примере с урною, где вынутый шар откладывается в сторону, число оставшихся в урне шаров будет одно и то же, какого бы цвета ни был шар, вынутый из урны в первый раз, но распределение оставшихся в урне шаров на белые и черные зависит от того, какого цвета шар, вынутый из урны при первом испытании.

Возвращаясь к общим рассуждениям, мы обозначим буквами m_2' и n_2' числа статочностей, благоприятствующих и неблагоприятствующих событию E_2 при условии, что событие E_1 состоялось.

Число всех статочностей относительно сложного события E так же точно, как и в предшествующем случае, выразится произведением:

$$\mu_1 \mu_2.$$

Статочностями, благоприятствующими событию E , будут всевозможные комбинации из статочностей, благоприятствующих событию E_1 со статочностями, благоприятствующими событию E_2 при условии, что E_1 состоялось. Число таких комбинаций:

$$m_1 m_2'.$$

Поэтому вероятность p сложного события E выразится так:

$$p = \frac{m_1 m_2'}{\mu_1 \mu_2},$$

или

$$p = \frac{m_1}{\mu_1} \cdot \frac{m_2'}{\mu_2}.$$

Отношение $\frac{m_1}{\mu_1}$ есть вероятность p_1 события E_1 , отношение же

$\frac{m_2'}{\mu_2}$ есть вероятность события E_2 , вычисленная для того случая,

когда E_1 состоялось. Мы назовем ее условною вероятностью события E_2 и обозначим буквою p_2' . Вероятность события E выразится так:

$$p = p_1 p_2' \dots \dots \dots (6).$$

Итак, вероятность сложного события, состоящего из двух простых зависимых между собою событий, равна произведению вероятности первого события на условную вероятность второго, вычисленную в предположении, что первое событие состоялось.

В случае нескольких простых между собою зависимых событий вероятность сложного события выразится произведением вероятности первого из простых событий на условные вероятности остальных, при чем каждая из них вычисляется в предположении, что предшествующие простые события состоялись:

$$p = p_1 p_2' p_3' \dots$$

Рассмотрим испытания с одной урной, куда положено 6 белых шаров и 4 черных. Спрашивается, какова вероятность, что в первые два раза вынется белый шар, а в третий раз — черный.

1) Если вынутый шар возвращается в урну, то искомая вероятность будет такова:

$$p = \frac{6}{10} \cdot \frac{6}{10} \cdot \frac{4}{10} = 0,144.$$

2) Если вынутый шар откладывается в сторону, то искомая вероятность будет такова:

$$p = \frac{6}{10} \cdot \frac{5}{9} \cdot \frac{4}{8} = \frac{1}{6} = 0,166.$$

§ 4. Примеры применения теорем сложения и умножения вероятностей.

Задача 1. Три игрока A , B и C играют между собою на следующих условиях: в первой партии участвуют A и B , а C остается в стороне; по окончании первой партии проигравший ее выходит из игры, уступая свое место C , по окончании второй партии проигравший ее выходит из игры, уступая свое место неигравшему во второй партии и т. д. Окончательно выигравшим считается тот, кто выиграл две партии под-ряд, при чем в каждой отдельной партии вероятность на выигрыш этой партии одинакова для обоих участников, т. е. равна $\frac{1}{2}$.

Обозначим вероятности на окончательный выигрыш для A , для B и для C соответственно буквами:

$$p_A, p_B, p_C.$$

Будем записывать в последовательном порядке имена игроков, выигравших отдельные партии. Так, например, для того случая, когда первую партию выиграл A , а вторую и третью выиграл C , составитя такая запись:

$$ACC;$$

в этом случае окончательный выигрыш будет принадлежать C .

Мы начнем с вычисления p_A и поэтому расположим в виде таблиц все записи, которые благоприятны выигрышу A . Они расположатся в два столбца, имеющих каждый бесконечное число строк:

I.

- 1) AA ,
- 2) $ACBAA$,
- 3) $ACBACBAA$,
- 4) $ACBACBACBAA$,

и т. д.

II.

- 1) $BCAA$,
- 2) $BCABCAA$,
- 3) $BCABCBACAA$,
- 4) $BCABCBACBACAA$,

и т. д.

По теореме сложения вероятностей мы должны вычислить вероятность каждой из этих статочностей, благоприятствующих A , и составить их сумму.

Каждая из статочностей предшествующей таблицы есть событие сложное, состоящее из нескольких простых событий, между собою независимых с вероятностью, равной $\frac{1}{2}$.

Вероятности этих статочностей расположатся в таблицу:

I

- 1) $\left(\frac{1}{2}\right)^2 = \frac{1}{2^2}$,
- 2) $\left(\frac{1}{2}\right)^5 = \frac{1}{2^5}$,
- 3) $\left(\frac{1}{2}\right)^8 = \frac{1}{2^8}$,
- 4) $\left(\frac{1}{2}\right)^{11} = \frac{1}{2^{11}}$,

II

- 1) $\left(\frac{1}{2}\right)^4 = \frac{1}{2^4}$,
- 2) $\left(\frac{1}{2}\right)^7 = \frac{1}{2^7}$,
- 3) $\left(\frac{1}{2}\right)^{10} = \frac{1}{2^{10}}$,
- 4) $\left(\frac{1}{2}\right)^{13} = \frac{1}{2^{13}}$,

Числа, стоящие в этих столбцах суть члены геометрических прогрессий с знаменателем $\frac{1}{2^3}$. Числа первого столбца дадут сумму:

$$\frac{\frac{1}{2^2}}{1 - \frac{1}{2^3}} = \frac{2}{7}.$$

Числа второго столбца дадут сумму:

$$\frac{\frac{1}{2^4}}{1 - \frac{1}{2^3}} = \frac{1}{14}$$

Сложив эти два числа, найдем вероятность p_A :

$$p_A = \frac{2}{7} + \frac{1}{14} = \frac{5}{14}$$

Так как A и B перед началом игры находятся в одинаковых условиях, то p_A должно равняться p_B ; а потому

$$p_B = \frac{5}{14}$$

Так как из трех лиц A , B , C непременно одно выигрывает, и при том только одно, то

$$p_A + p_B + p_C = 1.$$

Поэтому:

$$p_C = 1 - 2p_A = 1 - \frac{5}{7} = \frac{2}{7}.$$

Задача 2. Некто играет на следующих условиях: из урны, куда положен один белый и один черный шар, он вынимает шар два раза под-ряд, возвращая каждый раз вынутый шар в урну; если игрок вынул оба раза белый шар, то он выиграл; если же нет, то в урну прибавляется один черный шар и испытание повторяется. Если при этом испытании игрок вынет два раза под-ряд белый шар, то он выиграл; если же нет, то в урну прибавляется еще один черный шар и опять производится испытание, и т. д. без конца. Спрашивается, какова вероятность на выигрыш у этого лица.

Мы будем сначала находить вероятность того, что он не выиграет, т.е. что он не вынет двух раз под-ряд белый шар ни при одном из испытаний.

При первом испытании в урне 1 белый и 1 черный шар; вероятность вынуть два раза под-ряд белый шар равна $\frac{1}{2^2}$; вероятность события противоположного: $1 - \frac{1}{2^2}$. Такова вероятность первого простого события: при первом испытании игрок не выигрывает.

При втором испытании в урне 1 белый и 2 черных шара; вероятность вынуть два раза под-ряд белый шар равна $\frac{1}{3^2}$; вероятность события противоположного: $1 - \frac{1}{3^2}$. Такова условная вероятность второго простого события: игрок не выиграл при втором испытании, при чем принято во внимание, что он не выиграл при первом испытании.

Таким же образом найдется условная вероятность того, что игрок не выиграет при третьем испытании $1 - \frac{1}{4^2}$, при чем принято во внимание, что он не выиграл ни при 1-м, ни при 2-м испытании и т. д.

По теореме умножения вероятностей при событиях зависимых, найдем выражение вероятности того, что игрок не выиграет ни при каком испытании.

Она выразится бесконечным произведением:

$$\left(1 - \frac{1}{2^2}\right) \left(1 - \frac{1}{3^2}\right) \left(1 - \frac{1}{4^2}\right) \dots$$

Для вычисления этого выражения берем сначала конечное число множителей:

$$\left(1 - \frac{1}{2^2}\right) \left(1 - \frac{1}{3^2}\right) \left(1 - \frac{1}{4^2}\right) \dots \left(1 - \frac{1}{n^2}\right)$$

Применяя к каждой скобке формулу разности квадратов, находим:

$$\left(\frac{1}{2} \cdot \frac{3}{2}\right) \left(\frac{2}{3} \cdot \frac{4}{3}\right) \left(\frac{3}{4} \cdot \frac{5}{4}\right) \dots \left(\frac{n-1}{n} \cdot \frac{n+1}{n}\right)$$

Второй множитель каждой скобки дает в произведении с первым множителем следующей скобки единицу. Поэтому все выражение равно:

$$\frac{1}{2} \left(1 + \frac{1}{n}\right).$$

Предел этого выражения при n равном ∞ есть $\frac{1}{2}$; следовательно, вероятность того, что игрок не выиграет равна $\frac{1}{2}$.

Поэтому и вероятность его на выигрыш тоже равна $\frac{1}{2}$.

§ 5. Полная вероятность.

Во Введении мы говорили, что причиною в обыкновенном смысле слова называется такое событие или совокупность таких событий, наличие которых с достоверностью обуславливает появление некоторого другого события, называемого следствием.

В теории вероятностей это понятие несколько обобщается: причиною называют такое событие, которое сообщает некоторому другому событию определенную вероятность. Например, если имеется урна с m белыми и n черными шарами, то появление из нее белого шара имеет определенную вероятность $\frac{m}{\mu}$, где $\mu = m + n$. Можно сказать, что данный состав урны есть причина появления белого шара в обобщенном смысле слова.

В этом смысле можно сказать, что событие состоялось под действием одной из нескольких причин.

Назовем причины, под действием которых может состояться событие E буквами $C_1, C_2, C_3, \dots, C_k$. Пусть эти причины в данном вопросе суть всевозможные и взаимно исключающие друг друга. Иными словами, при данных условиях непременно будет действовать одна и только одна из этих причин.

Предположение, что будет действовать та или другая из этих причин, называется гипотезой; будем их называть соответственно 1-ой, 2-ой, 3-ей и т. д. гипотезой. Вероятность того, что будет действовать та или другая причина, называется вероятностью соответствующей гипотезы.

Пусть в данном случае вероятности всех k гипотез суть:

$$P_1, P_2, P_3, \dots, P_k \dots \dots \dots (7).$$

Обозначим вероятности, сообщаемые событию E соответствующими причинами так:

$$p_1, p_2, p_3, \dots, p_k \dots \dots \dots (8).$$

Так как причины $C_1, C_2, C_3, \dots, C_k$ суть события всевозможные и взаимно исключающие друг друга, то сумма вероятностей (7) должна равняться 1:

$$P_1 + P_2 + P_3 + \dots + P_k = 1 \dots \dots \dots (9).$$

Вероятность того, что событие E состоится, безразлично от какой причины, называется полной вероятностью события E . Мы обозначим ее буквою P и займемся ее определением.

Ввиду некоторой сложности и отвлеченности только что установленных понятий, поясним их на конкретном примере.

Пусть имеется k групп урн, совершенно одинаковых по виду и содержащих в себе одинаковое число шаров: по μ шаров в каждой урне; эти μ шаров могут быть белыми или черными, но число белых и черных шаров в различных урнах различно, а именно:

в каждой урне 1-ой группы находится m_1 белых и n_1 черных . . . всего $\mu = m_1 + n_1$;	
" " " 2-ой " " " m_2 " " n_2 " " $\mu = m_2 + n_2$,	
" " " 3-ей " " " m_3 " " n_3 " " $\mu = m_3 + n_3$,	
.....	
" " " k -ой " " " m_k " " n_k " " $\mu = m_k + n_k$.	

Числа урн в группах соответственно равны:

$$M_1, M_2, M_3, \dots, M_k,$$

при чем

$$M_1 + M_2 + M_3 + \dots + M_k = M.$$

Шары во всех урнах тщательно перемешаны, и урны тщательно перемешаны между собою.

После этого наудачу берется одна из урн, и из нее наудачу вынимается один шар.

Вероятность P того, что вынутый шар будет белый, есть полная вероятность события в данном случае. В этом случае k гипотез: случайно взятая урна принадлежит к той или иной из k групп.

Вероятности этих гипотез суть:

$$P_1 = \frac{M_1}{M}, \quad P_2 = \frac{M_2}{M}, \quad P_3 = \frac{M_3}{M}, \dots, \quad P_k = \frac{M_k}{M},$$

при чем

$$P_1 + P_2 + P_3 + \dots + P_k = \frac{M_1 + M_2 + M_3 + \dots + M_k}{M} = \frac{M}{M} = 1.$$

Вероятности, сообщаемые событию (вынугтию белого шара) различными причинами в данном случае таковы:

$$p_1 = \frac{m_1}{u}, \quad p_2 = \frac{m_2}{u}, \quad p_3 = \frac{m_3}{u}, \dots, \quad p_k = \frac{m_k}{u}.$$

Уяснив на приведенном примере смысл введенных понятий и обозначений, вернемся к общему вопросу о вероятности P события E .

Этому событию благоприятствует k статочностей:

- | | | | | | | |
|-------|-------------|------------|-----|-----------|---------|----------|
| 1) | Событие E | состоялось | под | действием | 1-ой | причины: |
| 2) | " | " | " | " | 2-ой | " |
| 3) | " | " | " | " | 3-ей | " |
| | | | | | | |
| k) | " | " | " | " | k -ой | " |

По теореме сложения мы должны вычислить вероятности всех этих статочностей и составить их сумму.

Начнем с 1-ой статочности: „Событие E состоялось под действием 1-ой причины“. Его можно рассматривать как сложное событие, состоящее из двух простых, между собою зависящих событий: 1) подействовала причина 1-я*), 2) под действием этой причины событие E состоялось **).

Вероятность 1-го события есть вероятность 1-ой гипотезы P_1 ; условная вероятность 2-го простого события есть вероятность p_1 , сообщаемая первой причиной событию E .

Следовательно, вероятность первой статочности, благоприятствующей событию E , есть $P_1 p_1$.

*) В приведенном выше примере это значит: „взятая урна принадлежит к 1-ой группе“.

**) В приведенном выше примере это значит: „из взятой урны 1-ой группы появился белый шар“.

На таком же основании вероятности других статочностей, благоприятствующих событию E , суть:

$$P_2 p_2, P_3 p_3, \dots, P_k p_k.$$

Полная вероятность P события E равна сумме вероятностей всех статочностей, благоприятствующих событию E , поэтому:

$$P = P_1 p_1 + P_2 p_2 + P_3 p_3 + \dots + P_k p_k \quad (10)$$

В справедливости этой формулы в приведенном выше примере можно убедиться непосредственно.

В самом деле, непосредственно усматривается, что при равенстве числа шаров во всех урнах вероятность вынудия белого шара будет одинакова в обоих случаях: 1) если взять на удачу одну из урн и из нее вынуть шар; 2) если шары из всех урн переложить в одну общую урну, тщательно перемешать и вынуть из нее один шар.

Вероятность появления белого шара в первом случае есть, как мы говорили, полная вероятность P .

Вероятность появления белого шара из общей урны можно вычислить непосредственно. В общую урну положено по μ шаров, находившихся прежде в отдельных урнах, число которых есть M . Следовательно, число всех шаров в общей урне есть $M\mu$. В том числе белых шаров:

$M_1 m_1$ из 1-ой группы; $M_2 m_2$ из 2-ой группы; $M_3 m_3$ из 3-ей группы...., $M_k m_k$ из k -ой группы; всего:

$$M_1 m_1 + M_2 m_2 + M_3 m_3 + \dots + M_k m_k$$

белых шаров.

Следовательно, вероятность вынудия белого шара из общей урны равна:

$$\begin{aligned} & \frac{M_1 m_1 + M_2 m_2 + M_3 m_3 + \dots + M_k m_k}{M\mu} = \\ & = \frac{M_1}{M} \cdot \frac{m_1}{\mu} + \frac{M_2}{M} \cdot \frac{m_2}{\mu} + \frac{M_3}{M} \cdot \frac{m_3}{\mu} + \dots + \frac{M_k}{M} \cdot \frac{m_k}{\mu} = \\ & = P_1 p_1 + P_2 p_2 + P_3 p_3 + \dots + P_k p_k. \end{aligned}$$

Так как мы видели, что эта вероятность должна равняться полной вероятности P , то:

$$P = P_1 p_1 + P_2 p_2 + P_3 p_3 + \dots + P_k p_k.$$

Рассмотрим числовой пример. Имеются две одинаковые по виду урны: в одной из них 8 белых шаров и 2 черных, в другой 4 белых и 6 черных. Берется на удачу одна из этих урн и из нее вынимается шар. Вероятность P , что этот шар будет белый, такова:

$$P = \frac{1}{2} \cdot \frac{8}{10} + \frac{1}{2} \cdot \frac{4}{10} = \frac{3}{5}.$$

§ 6. Различные косвенные приемы вычисления вероятностей.

При решении некоторых более сложных задач иногда удастся найти решение косвенным путем, пользуясь теми же теоремами, рассмотренными выше, но в различных искусственных комбинациях. Наиболее типичных таких приемов можно указать три: 1) составление алгебраических уравнений, 2) составление уравнений в конечных разностях, 3) применение функций генератрис.

Чтобы получить понятие о каждом из этих приемов, мы рассмотрим по одному простому примеру на применение каждого из них.

Задача 1. Возьмем снова решенную в § 4 задачу о трех игроках A, B, C : обозначим по-прежнему вероятности их на выигрыш буквами p_A, p_B, p_C .

Приведенное выше решение основано на непосредственном рассмотрении статистических, при чем число их оказалось бесконечно большим, а вероятности суть члены двух бесконечно убывающих геометрических прогрессий.

Теперь мы решим ту же задачу косвенным путем, основанным на применении алгебраических уравнений.

Начнем с решения вспомогательной задачи.

Пусть известно, что первая партия, в которой участвовали A и B , окончилась выигрышем A . При этом условии обозначим вероятности на окончательный выигрыш A, B, C соответственно буквами: x, y, z . Следовательно, x есть вероятность на окончательный выигрыш для игрока, только-что выигравшего партию, y — вероятность для игрока, только-что проигравшего партию, z — для игрока, вступающего в партию вновь.

Так как по предположению первую партию выиграл A , то во второй партии участвуют A и C .

Относительно результата второй партии можно сделать две гипотезы: 1) вторую партию выиграет A , 2) вторую партию выиграет C .

Вероятности обеих гипотез равны $\frac{1}{2}$.

При первой гипотезе игрок A после второй партии имеет выигранными две партии под ряд; вероятность его на окончательный выигрыш равна 1.

При второй гипотезе игрок A вторую партию проиграл и выходит из игры; он находится в таком же положении, как B перед второй партией; его вероятность на окончательный выигрыш равна y . По теореме о полной вероятности находим, что перед 2-ой партией вероятность на окончательный выигрыш для A равна:

$$\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot y.$$

Так как она обозначена буквою x , то можно написать:

$$x = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot y.$$

Применяя такие же рассуждения к нахождению полной вероятности для B , находим следующее: при первой сказанной выше гипотезе вероятность на окончательный выигрыш для B равна 0, так как игра уже окончилась в пользу A . При второй гипотезе вторую партию выиграл C , и B вступает в игру подобно тому, как C перед второй партией. Вероятность его на окончательный выигрыш поэтому равна z . Полная вероятность на выигрыш для B перед 2-ой партией равна:

$$\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot z.$$

Так как она обозначена буквою y , то:

$$y = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot z.$$

Рассуждая таким же образом о полной вероятности на окончательный выигрыш для C перед 2-ой партией, найдем:

$$z = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot x.$$

Таким образом, для нахождения трех величин x , y , z мы составили три следующие алгебраические уравнения:

$$\left. \begin{aligned} x &= \frac{1}{2} + \frac{1}{2} y, \\ y &= \frac{1}{2} z, \\ z &= \frac{1}{2} x. \end{aligned} \right\} \dots \dots \dots (11)$$

Решив их, находим:

$$x = \frac{4}{7}, \quad y = \frac{1}{7}, \quad z = \frac{2}{7}.$$

Сумма этих величин равна 1, как и должно быть, потому что x , y , z суть вероятности событий всевозможных и взаимно исключающих друг друга.

Найдя x , y , z , можем вычислить p_A , p_B , p_C опять при помощи формулы полной вероятности.

Относительно первой партии можно сделать две гипотезы: 1) эта партия окончилась в пользу A , 2) партия окончилась в пользу B .

При первой гипотезе игрок A по окончании 1-ой партии будет находиться в тех самых условиях, как в только-что рассмотренной

вспомогательной задаче; вероятность его на окончательный выигрыш равна $\frac{4}{7}$; при второй гипотезе игрок A по окончании 1-ой партии будет находиться в таких же условиях, в каких находится игрок B в рассмотренной вспомогательной задаче. Вероятность для него на окончательный выигрыш равна $\frac{1}{7}$.

По формуле полной вероятности находим, что вероятность для A на окончательный выигрыш перед началом игры, т. е. p_A , такова:

$$p_A = \frac{1}{2} \cdot \frac{4}{7} + \frac{1}{2} \cdot \frac{1}{7} = \frac{5}{14}.$$

Мы уже видели, что $p_A = p_B$, а потому:

$$p_B = \frac{5}{14}.$$

Наконец,

$$p_C = 1 - p_A - p_B = \frac{2}{7}.$$

Задача 2. Два игрока A и B , из которых первый имеет a рублей, а второй — b рублей, играют между собою на следующих условиях. Вероятность на выигрыш в каждой отдельной партии для обоих одинакова, следовательно, равна $\frac{1}{2}$; проигравший партию уплачивает выигравшему 1 рубль. Во время игры никто из них не может ни получить денег со стороны, ни истратить их на что-либо, кроме игры. Эта игра продолжается до тех пор, пока один из них не выиграет всех денег другого. Спрашивается, какова вероятность для A на окончательный выигрыш.

Обозначим буквою s сумму всех денег обоих игроков. Эта сумма

$$s = a + b$$

остается постоянною во все время игры.

Если в какой-либо момент у A всего x рублей, то у B будет $s - x$ рублей.

Обозначим вероятность для A на окончательный выигрыш в тот момент, когда у него x рублей, символом $f(x)$. Ясно, что эта величина есть функция только от x .

Относительно ближайшей следующей партии можно сделать две гипотезы: 1) она может окончиться в пользу A , 2) она может окончиться в пользу B .

При первой гипотезе у A после ближайшей партии будет $x + 1$ рублей, и вероятность на окончательный выигрыш $f(x + 1)$; при

второй гипотезе у A после ближайшей партии будет $x - 1$ рублей, и вероятность на окончательный выигрыш $f(x - 1)$.

Вероятность для A на окончательный выигрыш перед началом партии, когда у него было x рублей, выражается через $f(x)$; та же величина по формуле полной вероятности выразится так:

$$\frac{1}{2} f(x + 1) + \frac{1}{2} f(x - 1).$$

Следовательно:

$$f(x) = \frac{1}{2} f(x + 1) + \frac{1}{2} f(x - 1).$$

Это равенство должно быть справедливо при всяком x . Увеличивая x на 1 и перенеся все члены вправо, найдем:

$$f(x + 2) - 2f(x + 1) + f(x) = 0.$$

Введем символ исчисления конечных разностей, встречающийся и в дифференциальном исчислении:

$$\Delta f(x) = f(x + 1) - f(x);$$

применим то же самое действие к обеим частям последнего равенства; находим:

$$\Delta[\Delta f(x)] = f(x + 2) - 2f(x + 1) + f(x).$$

Выражение в левой части этого равенства называется второю разностью и обозначается символом:

$$\Delta^2 f(x).$$

Следовательно, полученное выше для определения функции $f(x)$ уравнение напишется так:

$$\Delta^2 f(x) = 0 \dots \dots \dots (12)$$

Это — уравнение в конечных разностях, определяющее функции $(f(x))$.

Непосредственной проверкой убеждаемся в том, что ему удовлетворяет функция:

$$f(x) = C_1 x + C_2,$$

где C_1 и C_2 суть произвольные постоянные. Исчисление конечных разностей показывает, что это — полный интеграл полученного разностного уравнения, т. е. что всякая функция $f(x)$, удовлетворяющая полученному уравнению, подходит под написанную формулу.

Остается найти величины постоянных C_1 и C_2 из условий задачи.

Для этого заметим, что при $x = 0$ игрок A уже лишился всех своих денег; для него вероятность на окончательный выигрыш равна 0, потому:

$$f(0) = 0.$$

При $x=s$ игрок A уже получил всю сумму, бывшую у обоих вместе, он выиграл окончательно, а потому:

$$f(s) = 1.$$

Итак, мы имеем:

$$C_1 \cdot 0 + C_2 = 0,$$

$$C_1 s + C_2 = 1.$$

Отсюда:

$$C_2 = 0, C_1 = \frac{1}{s} = \frac{1}{a+b}.$$

Поэтому:

$$f(x) = \frac{x}{a+b}.$$

До начала игры имущество A равнялось a , т. е. $x=a$, а потому вероятность p_A для A на окончательный выигрыш равна:

$$p_A = \frac{a}{a+b}.$$

На таком же основании вероятность p_B для B равна:

$$p_B = \frac{b}{a+b}.$$

Эти формулы можно написать так:

$$p_A = \frac{\frac{a}{b}}{\frac{a}{b} + 1}, \quad p_B = \frac{1}{\frac{a}{b} + 1} \dots \dots \dots (13)$$

Если $\frac{a}{b}$ настолько мало, что его можно принять равным 0, то формулы дадут:

$$p_A = 0, p_B = 1.$$

Это показывает, что из двух игроков, имущества которых очень разнятся между собою, беднейший почти наверно разорится, если они будут продолжать игру неограниченно долго. Такой результат объясняет, почему привычный игрок обыкновенно кончает разорением: совокупность всех его товарищей по игре можно рассматривать, как одно собирательное лицо с очень большим имуществом.

Задача 3. Бросают зараз k игорных костей в виде кубов. Найти вероятность того, что сумма выпавших на гранях очков будет равна s .

Решим задачу, применяя метод функций генератрис *).

*) Этот способ принадлежит Лапласу.

Если какая-либо функция $f(x)$ разлагается в ряд вида:

$$f(x) = A_0 + A_1 x + A_2 x^2 + A_3 x^3 + \dots,$$

то функция $f(x)$ называется генератрисой для чисел $A_0, A_1, A_2, A_3, \dots$

Понятие о функции генератрисе находит применение при решении многих задач по теории вероятностей.

Обозначим числа очков на выпавших k гранях брошенных костей буквами:

$$m_1, m_2, m_3, \dots, m_k; \dots \dots \dots (14)$$

это суть числа из ряда:

$$1, 2, 3, 4, 5, 6, \dots \dots \dots (15)$$

Если сумма выпавших очков равна s , то мы будем иметь:

$$m_1 + m_2 + m_3 + \dots + m_k = s. \dots \dots \dots (16)$$

Решаемая нами задача сводится к следующему вопросу: какова вероятность, что k случайно взятых чисел из ряда (15) удовлетворят уравнению (16).

Число всевозможных комбинаций чисел из ряда (15) равно:

$$n = 6^k.$$

Таково число всех возможных статочностей при данных условиях.

Число m статочностей благоприятствующих—равно числу решений уравнения (16).

Вопрос сводится к подсчету числа решений этого уравнения.

Для решения этого вопроса берем произведение k равных множителей:

$$(x + x^2 + x^3 + x^4 + x^5 + x^6) (x + x^2 + x^3 + x^4 + x^5 + x^6) \dots \dots \dots (17)$$

Совершив раскрытие скобок по правилу умножения многочленов, получим члены такого вида:

$$x^{m_1 + m_2 + m_3 + \dots + m_k}.$$

Если такой член встретится несколько раз, то по приведению получим:

$$A_{m_1 + m_2 + m_3 + \dots + m_k} x^{m_1 + m_2 + m_3 + \dots + m_k}.$$

Если показатель степени $m_1 + m_2 + m_3 + \dots + m_k$ равен s , то коэффициент A_s покажет искомое число решений уравнения (16).

Применяя сказанное выше определение, находим, что выражение (17) есть функция генератриса коэффициентов вида A_s .

Выражение (17) можно представить в таком виде:

$$\begin{aligned}
 (x + x^2 + x^3 + x^4 + x^5 + x^6)^k &= \left(\frac{x - x^6}{1 - x} \right)^k = \\
 &= x^k \cdot \frac{(1 - x^6)^k}{(1 - x)^k} = x^k (1 - x^6)^k (1 - x)^{-k} = \\
 &= x^k \left(1 - kx^5 + \frac{k(k-1)}{1 \cdot 2} x^{10} - \frac{k(k-1)(k-2)}{1 \cdot 2 \cdot 3} x^{15} + \dots \right) \\
 &\left(1 + kx + \frac{k(k+1)}{1 \cdot 2} x^2 + \frac{k(k+1)(k+2)}{1 \cdot 2 \cdot 3} x^3 + \dots \right) \dots \dots (18)
 \end{aligned}$$

Совершив перемножение членов в формуле (18), находим коэффициент A_s при x^s : он равен искомому числу m .

Вероятность, что сумма выпавших очков будет равна s , выразится так:

$$p = \frac{m}{\mu} = \frac{A_s}{6^k}.$$

Возьмем числовой пример: пусть число костей $k=2$, сумма выпавших очков $s=7$.

Выражение (18) примет вид:

$$x^2(1 - 2x^5 + x^{10})(1 + 2x + 3x^2 + 4x^3 + 5x^4 + 6x^5 + \dots).$$

Коэффициент при x^7 будет:

$$A_7 = 6.$$

Следовательно, искомая вероятность равна:

$$p = \frac{6}{36} = \frac{1}{6}.$$

§ 7. О геометрических вероятностях.

В предшествующих рассуждениях мы имели дело с такими случайными явлениями, где число статочностей конечно. Правда, в двух примерах нам встретился случай бесконечно-большого числа статочностей, но эти примеры были такого характера, что искомую вероятность можно было представить как предел выражения, составленного обычным способом.

Кроме задач такого рода существуют задачи совершенно иного характера, где число статочностей по существу бесконечно-большое. Так как задачи этого рода впервые встретились по поводу геометрических вопросов, то соответствующие вероятности получили название геометрических вероятностей.

Прием вычислений вероятностей при конечном числе статочностей мы установили из наблюдения над суждением по здравому смыслу в таких вопросах, где условия можно подвести под схему урны с белыми и черными шарами.

Задачи о геометрических вероятностях подводятся под иную, им соответствующую схему.

Пусть дана определенная площадь S (чертеж 1), внутри которой бросают точку так, что она непременно упадет в площадь S , но в какую именно точку — безразлично. Все точки площади S в этом отношении равновозможны.

Пусть внутри площади S выделена меньшая площадь s ; спрашивается, какова вероятность p , что брошенная точка упадет внутри s .

Руководствуясь здравым смыслом, мы скажем, что p возрастает по мере возрастания

$\frac{s}{S}$. Когда это отношение мало, то попасть, случайно бросив точку, внутрь площади s нам представляется почти невозможным;

когда $\frac{s}{S}$ близко к 1, то можно с уверенностью сказать, что брошенная точка упадет в s .

Поэтому естественно принять отношение $\frac{s}{S}$ за величину вероятности и положить:

$$p = \frac{s}{S} \quad (19)$$

Это равенство служит определением величины p ; оно не может быть доказано, но можно обнаружить целесообразность такого определения.

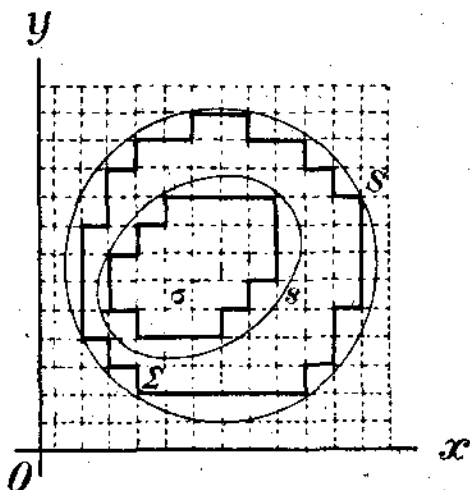
С этой целью на чертеже 1 строим сеть квадратов, со сторонами равными малой величине h , и отбираем те квадраты этой сети, которые укладываются внутри площади S . Число этих квадратов обозначим буквою μ , а ступенчатого вида площадь, составленную из этих квадратов, обозначим буквою Σ .

Так как площадь квадрата равна h^2 , то мы будем иметь:

$$\Sigma = \mu h^2 \quad (20)$$

Таким же образом отберем квадраты, укладываемые внутри площади s ; обозначим число их буквою m , а площадь ими составляемую буквою σ . Будем иметь:

$$\sigma = mh^2 \quad (21)$$



Чертеж 1.

Изменим несколько условия решаемой нами задачи.

Пусть дано, что бросаемая точка непременно должна упасть внутри площади Σ ; требуется определить величину вероятности p того, что точка упадет внутри площади σ .

Согласно новому условию задачи точка непременно должна упасть в один из квадратов, составляющих Σ ; следовательно, существует всего μ статочностей, равновозможных и взаимно исключающих друг друга. Брошенная точка упадет внутри σ , если она ляжет в один из m квадратов, составляющих σ ; это соответствует в данном случае m благоприятствующим статочностям.

Придерживаясь прежнего определения вероятности, находим

$$p' = \frac{m}{\mu}.$$

или

$$p' = \frac{mh^2}{\mu h^2} = \frac{\sigma}{\Sigma} \quad (22)$$

Из курса анализа известно, что

$$\lim_{h \rightarrow 0} \sigma = s, \quad \lim_{h \rightarrow 0} \Sigma = S.$$

Поэтому из равенства (22) следует:

$$\lim_{h \rightarrow 0} p' = \frac{s}{S} \quad (23)$$

Если мы сделаем предпосылку, что

$$\lim_{h \rightarrow 0} p' = p \quad (24)$$

то из равенства (23) будет следовать то равенство, которое мы приняли за определение:

$$p = \frac{s}{S} \quad (19)$$

Следовательно, равенство (19) можно рассматривать, как следствие предпосылки (24), которую мы принимаем в виде постулата.

Геометрические вероятности представляют некоторые особенности сравнительно с теми вероятностями, где число всех статочностей конечно.

1. Площади s и S , входящие в формулу (19), могут выражаться числами иррациональными; в таком случае и вероятность p может оказаться числом иррациональным.

2. Вероятность (19) обратится в 0, если площадь s дегенерирует в конечный отрезок линии или в точку.

Это заключение, с первого взгляда парадоксальное, если в него вдуматься, будет вполне согласно с нашим обычным суждением: бросая точку случайным образом внутри площади конечных размеров, мы считаем невероятным, чтобы она попала в заранее намеченную точку или в заранее намеченный конечный отрезок линии.

§ 8. Примеры вычисления геометрических вероятностей.

1. Задача о ломком стержне. Стержень определенной длины $2a$ случайным образом ломается на три части. Спрашивается, какова вероятность, что эти три части могут быть сторонами треугольника, при чем предполагается, что стержень в одинаковой степени может сломаться в каждой своей точке.

Обозначая буквами x, y, z длины трех частей сломавшегося стержня, находим:

$$x + y + z = 2a \quad (25)$$

где x, y, z числа положительные. Отсюда:

$$x + y < 2a \quad (26)$$

где x и y произвольные положительные числа, удовлетворяющие этому неравенству.

Величины x, y, z будут в состоянии служить сторонами треугольника при условиях:

$$x + y > z, y + z > x, z + x > y,$$

или на основании (25):

$$x + y > 2a - x - y, y + 2a - x - y > x, 2a - x - y + x > y,$$

откуда:

$$x + y > a, x < a, y < a. \quad (27)$$

Величины x, y, z будут в состоянии служить сторонами треугольника, если положительные числа x и y удовлетворяют неравенствам (27).

Будем рассматривать x и y как декартовы координаты точки на плоскости; тогда всякой паре положительных чисел (x, y) будет соответствовать точка в первой четверти плоскости.

Неравенство (26) показывает, что точка (x, y) лежит по ту же сторону от прямой

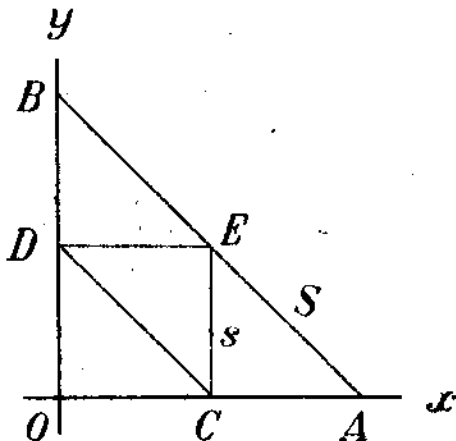
$$x + y = 2a,$$

как и начало координат; иными словами, эта точка лежит внутри треугольника OAB (чертеж 2), имеющего катеты OA и OB , длиною равные $2a$.

Неравенства (27) показывают, что если x, y, z способны составить треугольник, то точка (x, y) лежит следующим образом:

1. Она лежит относительно прямой

$$x + y = a$$



Чертеж 2.

по другую сторону, чем начало координат; иными словами вне прямоугольного треугольника OCD , имеющего катетами OC и OD , длиною равные a .

2. Точка (x, y) имеет абсциссу меньшую, чем a , т. е. она лежит влево от прямой CE , параллельной оси y .

3. Точка (x, y) имеет ординату меньшую, чем a , т. е. она лежит ниже прямой DE , параллельной оси x .

Это значит, что точка (x, y) лежит внутри треугольника CED .

Итак, каковы бы ни были части x, y, z стержня, точка (x, y) будет лежать внутри треугольника OAB ; эти части способны составить треугольник, если точка (x, y) лежит внутри треугольника CED .

В таком случае решаемая нами задача равносильна такой задаче: точка случайным образом бросается внутрь площади OAB ; какова вероятность, что она упадет внутрь CED .

Площадь $OAB = S = 2a^2$,

Площадь $CED = s = \frac{a^2}{2}$.

По формуле (19) находим:

$$p = \frac{a^2}{4a^2} = \frac{1}{4}.$$

II. Задача об игле. На горизонтальную плоскость, с начерченными параллельными прямыми, проведенными друг от друга в расстоянии $2a$, случайным образом бросают тонкую иглу длиною $2l$ при чем $l < a$. Спрашивается, какова вероятность, что эта игла, упав на плоскость, пересечет одну из параллельных прямых.

Прежде всего заметим, что искомая вероятность не зависит от следующих условий:

- 1) в какой полосе между двумя параллельными ляжет середина O иглы после ее падения;
- 2) будет ли эта середина O лежать ближе к верхней или нижней параллельной прямой, между которыми она упала;
- 3) где поместится основание перпендикуляра P , опущенного из O на ближайшую к нему параллельную прямую.

Поэтому, не изменяя результата, можем упростить условия задачи, предполагая, что:

- 1) середина O должна упасть внутри определенной полосы, ограничиваемой параллельными AB и ED (черт. 3);
- 2) точка O лежит ближе к прямой AB ;
- 3) перпендикуляр, опущенный из O на AB есть прямая OP .

Середину отрезка перпендикуляра между параллельными AB и ED обозначим буквою C . Из условия задачи следует, что $CP = a$, при чем $l < a$.

Возможность пересечения иглы с прямою AB определяется величинами двух переменных: 1) углом направления нижней половины иглы с прямою CP ; мы обозначим его буквою x ; 2) длиною PO ; мы обозначим ее буквою y .

Так как мы обозначили буквою x угол, образуемый нижней половиной иглы, то x может получать всевозможные значения

от $-\frac{\pi}{2}$ до $+\frac{\pi}{2}$. Так как мы

принимая, что O лежит ближе к AB , чем к ED , то $y < a$.

Итак переменные x и y должны во всяком случае, удовлетворять неравенствам:

$$\left. \begin{aligned} -\frac{\pi}{2} < x < +\frac{\pi}{2} \\ 0 < y < a. \end{aligned} \right\} \quad (28)$$

Для того, чтобы игла могла пересекать прямую AB , необходимы два условия:

$$1) \quad OP < l, \text{ т.-е. } y < l,$$

2) если мы проведем окружность радиуса l с центром в O , то она пересечет прямую AB в двух точках Q и R ; для того, чтобы игла пересекала прямую AB , надо, чтобы направление ее лежало внутри угла QOR . Если обозначить угол POR буквою α , то будем иметь:

$$-\alpha < x < +\alpha.$$

Из прямоугольного треугольника POR следует: $y = l \cos \alpha$; поэтому предшествующее неравенство равносильно следующему:

$$y < l \cos x. \quad (29)$$

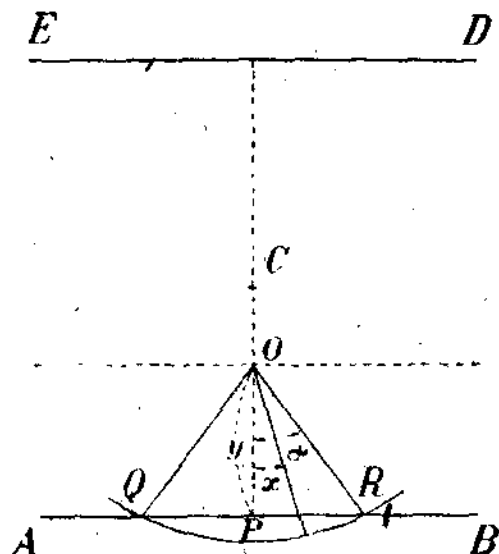
Найденное выше условие $y < l$ уже содержится в неравенстве (29).

Поэтому можно сказать, что для пересечения иглой линии AB нужно только одно условие (29).

Считая x и y за переменные, все значения которых в пределах (28) равновозможны, и принимая их за декартовы координаты точки (x, y) , видим, что точка, координаты которой удовлетворяют неравенствам (28), лежит внутри прямоугольника $ABDC$ (черт. 4), где $OA = OB = \frac{\pi}{2}$,

$AC = a$. Эту площадь мы обозначим буквою S ; величина ее:

$$S = \pi a. \quad (30)$$



Чертеж 3.

Точка (x, y) , координаты которой удовлетворяют условию (29), принадлежит площади s , ограничиваемой осью x и кривой:

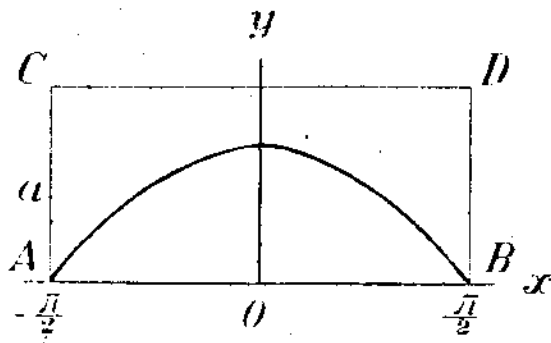
$$y = l \cos x.$$

Величина s выразится так:

$$s = l \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos x \, dx = 2l. \quad (31)$$

Искомая вероятность p представится формулой:

$$p = \frac{s}{S} = \frac{2l}{\pi a}. \quad (32)$$



Чертеж 4.

Примечание. Сделанный вывод основан на выборе переменных x и y , определяющих положение иглы, и на предпосылке о равно-возможности величин этих переменных в найденных пределах. Изменение в выборе переменных может отразиться на результате. Примеры парадоксов, основанных на подобных изменениях, можно найти у Бертрана *). В частности для задачи об игле была произведена опытная проверка, о которой мы будем говорить ниже.

Задача об игле интересна в том отношении, что она приводит к результату, выражающемуся через иррациональное число, зависящее от π .

*) Bertrand. Théorie des probabilités, стр. 4. См. также Czuber. Wahrscheinlichkeitsrechnung, стр. 106; Poincaré. Calcul des probabilités, стр. 94.

ГЛАВА ВТОРАЯ

О математическом ожидании.

§ 1. Понятие о математическом ожидании случайной суммы.

Пусть некоторое лицо имеет право получить некоторую сумму денег или вообще некоторую ценность в A рублей, но лишь при условии, что произойдет событие E , вероятность которого равна p .

Так как получение этой суммы зависит от случая, то ее называют случайной суммой. Пример случайной суммы представляет выигрыш в лотерею (если вынется выигрышный билет), получение товара идущего морем (если в пути не произойдет аварии), получение страховой премии для лица, застраховавшего свою жизнь (если оно доживет до условленного срока) и т. д.

Если бы лицо, ожидающее получения случайной суммы, пожелало продать это право другому лицу, то оно могло бы получить за свою уступку некоторую сумму денег, но эта сумма должна быть меньше полного размера A случайной суммы, так как для покупателя получение ее сопряжено с риском. Стоимость права на получение случайной суммы в A рублей, если произойдет событие E с вероятностью p , называется математическим ожиданием этой суммы при данных условиях. Обозначим его буквою a и займемся его определением.

С этой целью мы опять начнем с рассмотрения простого примера.

Пусть некто, имея ценный предмет в A рублей, хочет разыграть его в лотерею, от которой он не ищет прибыли, а только желает выручить стоимость A разыгрываемого предмета. Он заготовил n лотерейных билетов, которые все распределил между покупателями.

Ясно, что в таком случае каждый билет должен стоить $\frac{A}{n}$ рублей.

Лицо, купившее m билетов в эту лотерею, заплатит $m \cdot \frac{A}{n}$ рублей, или:

$$A \cdot \frac{m}{n} \text{ руб.}$$

Это есть та сумма, которую по справедливости это лицо заплатит за приобретаемое право на получение случайной суммы в A рублей. По условию мы называем ее математическим ожиданием и обозначаем буквою a :

$$a = A \cdot \frac{m}{\mu}.$$

Величина $\frac{m}{\mu}$ есть вероятность на получение выигрыша для лица купившего m билетов в лотерею, где всех билетов μ . Обозначив эту вероятность буквою p , находим:

$$a = A p. \quad (1)$$

Под схему рассмотренной нами задачи можно подвести все вопросы, где речь идет об одной случайной сумме, получение которой зависит от случайного события с определенной вероятностью. Поэтому можно сказать, что математическое ожидание лица на получение случайной суммы равно величине этой суммы, умноженной на вероятность ее получения.

Если какое-либо лицо имеет право на получение одной из случайных сумм $A_1, A_2, A_3, \dots, A_\mu$ в случае, если произойдут соответствующие события $E_1, E_2, E_3, \dots, E_\mu$ с вероятностями $p_1, p_2, p_3, \dots, p_\mu$, то мы назовем математическим ожиданием этого лица сумму математических ожиданий, соответствующих получению $A_1, A_2, A_3, \dots, A_\mu$ в отдельности; обозначив математическое ожидание буквою a , будем иметь:

$$a = A_1 p_1 + A_2 p_2 + A_3 p_3 + \dots + A_\mu p_\mu. \quad (2)$$

Под формулы (1) и (2) подводятся и те случаи, когда некоторые из случайных сумм отрицательны, т. е. соответствуют убытку или проигрышу. Так формула (1) при отрицательном A определит ту сумму, которую по справедливости должно заплатить лицо, находящееся под угрозой потерпеть убыток в A рублей, другому лицу, страхующему его от этого убытка.

Таким образом, на формулах (1) и (2) основаны всевозможные расчеты страховых обществ, эмеритальных касс и других подобных учреждений, где производится страхования от несчастных случаев (пожара, грабительства, на случай смерти и т. д.), а также единовременных пособий, рент и проч.

На этих же формулах основаны и расчеты по поводу азартных игр, давшие первый толчок к развитию теории вероятностей при Паскале и Фермате в XVII веке.

В виде примера приложения формулы (2) рассмотрим задачу:

У ювелира был алмаз ценою в A рублей. В то время как в его магазин пришел покупатель, ювелиру сообщили, что в мастерской,

находящейся вдали от магазина, алмаз разбили на два куска; но каково отношение этих кусков — неизвестно. Покупатель предлагает ему купить заглазно оба куска. Спрашивается, какую плату по справедливости должен назначить ювелир, если известно, что стоимость алмаза прямо пропорциональна квадрату его веса.

Примем вес алмаза за единицу и разделим его мысленно на n равных частей; величина одной части будет $\frac{1}{n}$; мы обозначим ее через dx .

Применяя прием обычный в прикладной математике, будем предполагать, что каждый из двух кусков алмаза содержит в себе целое число таких частей, имея в виду в конечном результате совершить переход к пределу при n равном ∞ .

При таком предположении скажем, что алмаз мог расколоться на два куска только n способами, при чем в 1-м куске было: 1, 2, 3, n частей. Вероятность каждого из этих случаев одна и та же и равна:

$$\frac{1}{n} = dx.$$

В таком случае, когда вес 1-го куска равен x , вес 2-го равен $1 - x$; следовательно, стоимость всего разбитого алмаза:

$$Ax^2 + A(1 - x)^2 = A(2x^2 - 2x + 1).$$

Такова случайная сумма, в действительности получаемая покупателем в том случае, когда 1-ый кусок алмаза весит x .

Применяя формулу (2), мы умножаем эту величину на вероятность такого случая dx и берем сумму выражений вида:

$$A(2x^2 - 2x + 1) dx,$$

для всех возможных значений x между 0 и 1.

В пределе, когда $n = \infty$, эта сумма делается равной интегралу:

$$A \int_0^1 (2x^2 - 2x + 1) dx,$$

или:

$$\frac{2}{3} A.$$

Таково математическое ожидание покупателя, заглазно берущего алмаз; такова сумма, в которую можно оценить безобидно 2 куска алмаза при данных условиях.

§ 2. О безобидных играх.

Выигрыш игрока можно рассматривать, как случайную сумму; поэтому делаемая им перед началом игры ставка есть плата, вносимая за право на получение случайной суммы, т. е. выигрыша.

Из сказанного следует, что в безобидной игре ставка каждого игрока должна равняться его математическому ожиданию.

Рассмотрим пример общего характера. Пусть играют между собою два лица: A и B . Вероятность на выигрыш первого пусть равна p , а второго q , при чем, конечно,

$$p + q = 1.$$

Пусть ставка первого игрока a , а второго — b .

Спрашивается, каковы должны быть эти ставки для безобидности игры, если выигравший получает сумму обеих ставок, т. е. $a + b$.

Математическое ожидание игрока A есть $p(a + b)$; оно должно равняться его ставке a :

$$a = p(a + b).$$

На таком же основании для игрока B мы найдем:

$$b = q(a + b).$$

Отсюда:

$$\left. \begin{aligned} p &= \frac{a}{a+b}, \\ q &= \frac{b}{a+b}. \end{aligned} \right\} \dots \dots \dots (3)$$

Следовательно:

$$\frac{p}{q} = \frac{a}{b} \dots \dots \dots (4)$$

Это показывает, что в безобидной игре ставки должны быть пропорциональны вероятностям на выигрыш.

Формулы (3) определяют вероятности на выигрыш для лиц, сделавших определенные ставки, если известно, что игра безобидна.

Эти равенства (3) можно представить в таком виде:

$$\begin{aligned} p &= \frac{\left(\frac{a}{b}\right)}{\left(\frac{a}{b}\right) + 1} \\ q &= \frac{1}{\left(\frac{a}{b}\right) + 1} \end{aligned}$$

Если величина $\frac{a}{b}$ настолько мала, что ее можно считать равною нулю, то мы будем иметь:

$$p = 0, \quad q = 1.$$

Иными словами, участвуя в безобидной игре, тот игрок, ставка которого очень мала сравнительно со ставкою противника, почти наверно ее потеряет.

Это — результат почти тот же, который мы получили, решая задачу 2 § 6 предшествующей главы. В самом деле, привыч-

ный игрок в сущности ставит все свое имущество, имея против себя собирательное лицо — всю массу публики, посещающей ту же игру. Понятно, что ставка этого собирательного лица громадна сравнительно со ставкою отдельного игрока.

Рассмотрим в виде примера расчета безобидности игры условия игры во французской королевской лотерее, о которой говорили выше (глава I, § 1, задача 5). Мы нашли следующие вероятности для игроков, записавших 1, 2, 3, 4, 5 номеров:

$$p_1 = \frac{1}{18}, p_2 = \frac{2}{801}, p_3 = \frac{1}{11748}, p_4 = \frac{1}{511038}, p_5 = \frac{1}{43949268}.$$

Обозначим буквою p одно из этих чисел, буквою a ставку игрока и буквою A выигрыш, который он должен был бы получить при безобидности игры. Следовательно, $a = Ap$, откуда:

$$A = \frac{1}{p} a.$$

Иными словами, для безобидности королевской лотереи выигрыш должен был превышать ставку в $\frac{1}{p}$ раз, т. е. для игрока, записавшего:

1 номер в	18	раз
2 " "	400,5	"
3 " "	11748	"
4 " "	511038	"
5 " "	43949268	"

В действительности, как мы видели, эти числа были заменены следующими:

15, 270, 5500, 7500;

случаев записи 5 номеров не встречалось.

Отсюда видно, что королевская лотерея не была игрой безобидной.

§ 3. Петербургская игра.

Под именем Петербургской игры известна задача, предложенная Николаем Бернулли и напечатанная впервые в Записках Петербургской Академии. Результат решения этой задачи представляется парадоксальным и вызвал большой спор между математиками начала XVIII века.

Задача состоит в следующем: играют между собою Петр и Павел; сначала Павел ставит 1 дукат, Петр бросает монету; если выпадает орел, то Петр получает 1 дукат; в противном случае Павел удваивает ставку, и Петр опять бросает монету; если этот раз выпадает орел, то Петр получает 2 дуката; в противном случае Павел еще удваивает ставку, и Петр опять бросает монету, и т. д. пока не

вскроется орел; до тех пор после всякого вскрытия решетки Павел удваивает свою ставку; вскрытием орла партия кончается.

Спрашивается, каково математическое ожидание Петра, т. е. какую сумму он должен уплатить Павлу перед началом партии для того, чтобы игра была безобидна.

Случайные суммы, которые может получить Петр суть ставки, делаемые Павлом, т. е.

$$1, 2, 4, 8, 16, \dots;$$

вероятности их получения суть вероятности сложных событий, состоящих в том, что орел появился впервые при 1-м, 2-м, 3-м, 4-м, 5-м, ... бросании; эти вероятности соответственно равны:

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \dots$$

Математическое ожидание Петра по формуле (2) выразится так:

$$1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 4 \cdot \frac{1}{8} + 8 \cdot \frac{1}{16} + 16 \cdot \frac{1}{32} + \dots,$$

или:

$$\frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots$$

Эта сумма равна ∞ .

Такой результат объясняется следующим образом: если бы можно было осуществить очень большое число партий между Петром и Павлом, то какую бы ставку конечной величины ни делал Петр перед началом каждой партии, он в конце всех партий всегда окажется в выигрыше. Тем не менее в практическом смысле результат парадоксален, потому что бесконечной ставки никто сделать не может; кроме того, даже если бы Петр сделал очень большую ставку, то выиграв взамен ее несколько дукатов, он едва ли мог бы продолжать игру.

Парадоксальность ответа задачи вызвала, как упомянуто выше, большой спор между математиками, в котором принял участие Даниил Бернулли (племянник Николая Бернулли, предложившего задачу). Для разрешения спора Даниил Бернулли предложил особое понятие о нравственном значении имущества и нравственном ожидании.

§ 4. О нравственном ожидании.

Основная мысль Даниила Бернулли состоит в том, что приобретенные или потеря какой-либо суммы имеет для каждого человека не абсолютное значение, а значение относительное, в зависимости от имущественного положения каждого человека в отдельности. Напр., приобретение или потеря небольшой суммы для богача почти незаметна, а для бедняка может иметь громадное значение. Так же

точно количество материального блага, которое можно извлечь из данного имущества, не пропорционально ему, а есть лишь некоторая функция его величины.

Обозначая количество материального блага, которое можно извлечь из имущества x через $f(x)$ и не определяя заранее вида этой функции, мы можем однако сделать ряд заключений о ее свойствах, наблюдая над относящимися сюда суждениями из обыденной жизни:

1) функция $f(x)$ положительна при всех значениях x ; 2) функция $f(x)$ растет с увеличением x , а потому величина:

$$\Delta f(x) = f(x+h) - f(x)$$

положительна при всяком положительном h ; 3) разность $\Delta f(x)$ есть приращение блага при увеличении имущества x на величину h ; мы видели, что при малых значениях x это приращение блага имеет большее значение, чем при больших; это значит, что $\Delta f(x)$ уменьшается с возрастанием x , а потому разность от $\Delta f(x)$, т. е. вторая разность $\Delta^2 f(x)$ должна быть отрицательна.

К этому мы присоединим предпосылку, что функция $f(x)$ имеет производные по крайней мере двух первых порядков.

Так как предел $\frac{\Delta f(x)}{h}$ есть первая производная $f'(x)$ от функции $f(x)$, а предел выражения $\frac{\Delta^2 f(x)}{h^2}$ есть вторая производная $f''(x)$, то из свойств 2) и 3) следует, что $f'(x)$ положительна, а $f''(x)$ отрицательна при всяком положительном x .

Из свойства 3) между прочим следует, что нравственное значение проигрыша больше, чем равного ему, выигрыша. В самом деле для лица, имеющего x рублей, нравственное значение проигрыша в h рублей выражается формулой: $f(x) - f(x-h)$, а нравственное значение выигрыша в h рублей формулой: $f(x+h) - f(x)$, разность между этими величинами:

$$\left[f(x) - f(x-h) \right] - \left[f(x+h) - f(x) \right] = - \left\{ f(x+h) - 2f(x) + f(x-h) \right\} = -\Delta^2 f(x-h);$$

она положительна на основании свойства 3).

Отсюда следует, что участие даже в безобидной игре нравственно невыгодно.

Подобным же образом можно сделать и некоторые другие выводы относительно азартных игр и страхований, пользуясь только сказанными выше свойствами функции $f(x)$ и не определяя ближе ее выражения. Но в некоторых вопросах необходимо иметь аналитическое выражение функции $f(x)$.

Для составления формулы, выражающей такую функцию, Даниил Бернулли предложил сделать дальнейшую предпосылку: нравственное значение бесконечно-малого приращения dx имущества x принимается обратно-пропорциональным самому имуществу x , т. е. оно принимается равным:

$$k \frac{dx}{x}, \dots \dots \dots (5)$$

где k постоянный положительный множитель пропорциональности.

В таком случае нравственное значение увеличения имущества от x до $x+h$, т. е. величина $f(x+h) - f(x)$ выразится формулой:

$$f(x+h) - f(x) = k \int_x^{x+h} \frac{dx}{x} = k \lg \left(1 + \frac{h}{x} \right) \dots \dots \dots (6)$$

Отсюда следует, что:

$$f'(x) = \frac{k}{x}, \dots \dots \dots (7)$$

$$f''(x) = -\frac{k}{x^2}, \dots \dots \dots (8)$$

$$f(x) = k \lg x + C \dots \dots \dots (9)$$

Формулы (7) и (8) показывают, что $f'(x)$ положительна, а $f''(x)$ отрицательна при всяком положительном значении x , что согласно с приведенными выше свойствами 2) и 3); формула (9) определяет вид самой функции $f(x)$, при чем значение постоянного C остается неопределенным.

Против правильности выражения (9) могло бы быть сделано следующее возражение: при значениях x , меньших 1, первый член формулы (9) отрицателен, а потому при значениях x близких к 0 все выражение (9) отрицательно, чего быть не должно по свойству 1) функции $f(x)$.

Для устранения сомнения надо вспомнить, что вывод формулы (9) основан на наблюдении над обыденным суждением по поводу имущественного положения людей в обычных условиях. Человека, имущество которого близко к 0, мы не можем себе представить потому, что уже самая жизнь и способность человека к труду и заработку есть имущество весьма высокой цены. Поэтому формулу (9) можно считать пригодною для практических целей, при чем постоянное C остается неопределенным, но настолько большим, чтобы величина (9) не получала отрицательных значений при рассматриваемых значениях x .

Пусть некоторое лицо, владеющее имуществом a , ожидает получить одну из следующих случайных сумм:

$$h_1, h_2, h_3, \dots, h_n,$$

при чем одну из этих сумм оно непременно получит, и при том только одну; получения этих сумм — события всевозможные и взаимно исключающие друг друга; вероятности их мы обозначим буквами

$$p_1, p_2, p_3, \dots, p_n,$$

при чем, понятно:

$$p_1 + p_2 + p_3 + \dots + p_n = 1 \quad (10)$$

Нравственное значение увеличения имущества от a до $a + h_1$ по формуле (6) выразится так:

$$k \lg \left(1 + \frac{h_1}{a} \right).$$

Так как это увеличение имущества произойдет только в случае наступления события с вероятностью p_1 , то для данного лица нравственное значение этого ожидаемого увеличения имущества выразится так:

$$kp_1 \lg \left(1 + \frac{h_1}{a} \right).$$

Подобные же рассуждения можно сделать по поводу остальных ожидаемых случайных сумм: $h_2, h_3, \dots, h_n$.

Нравственное значение увеличения имущества для данного лица при этих условиях выразится формулой:

$$kp_1 \lg \left(1 + \frac{h_1}{a} \right) + kp_2 \lg \left(1 + \frac{h_2}{a} \right) + kp_3 \lg \left(1 + \frac{h_3}{a} \right) + \dots +$$

$$+ kp_n \lg \left(1 + \frac{h_n}{a} \right),$$

или:

$$k \lg \left[\left(1 + \frac{h_1}{a} \right)^{p_1} \left(1 + \frac{h_2}{a} \right)^{p_2} \left(1 + \frac{h_3}{a} \right)^{p_3} \dots \left(1 + \frac{h_n}{a} \right)^{p_n} \right] \quad (11)$$

Обозначим буквою h ту сумму, достоверное получение которой имеет нравственное значение равное величине (11). В таком случае величина (11) будет равна величине: $k \lg \left(1 + \frac{h}{a} \right)$.

Отсюда:

$$\left(1 + \frac{h}{a} \right) = \left(1 + \frac{h_1}{a} \right)^{p_1} \left(1 + \frac{h_2}{a} \right)^{p_2} \left(1 + \frac{h_3}{a} \right)^{p_3} \dots \left(1 + \frac{h_n}{a} \right)^{p_n} \quad (12)$$

Приведя члены к общему знаменателю и приняв во внимание равенство (10), найдем:

$$h = (a + h_1)^{p_1} (a + h_2)^{p_2} (a + h_3)^{p_3} \dots (a + h_n)^{p_n} - a \quad (13)$$

Это выражение Даниил Бернулли назвал нравственным ожиданием при рассмотренных условиях. Лицо, владеющее имуществом a , могло бы нравственно безобидно взять сумму h взамен тех прав, которые оно имеет на случайные суммы $h_1, h_2, h_3, \dots, h_n$.

Применив формулу (12) к нахождению нравственного ожидания Петра при условиях Петербургской игры, найдем:

$$\frac{h}{a} = \left(1 + \frac{1}{a}\right)^{\frac{1}{2}} \left(1 + \frac{2}{a}\right)^{\frac{1}{4}} \left(1 + \frac{4}{a}\right)^{\frac{1}{8}} \dots - 1.$$

Приближенные вычисления по этой формуле дают:

при $a = 100$, $h = 4,36$,

при $a = 1000$, $h = 6$.

ГЛАВА ТРЕТЬЯ

О вероятностях сложных событий, имеющих место при большом числе испытаний.

§ 1. Выражение вероятности сложного события при большом числе испытаний над простыми событиями, между собою независимыми.

Пусть производится испытание над появлением события E или события ему противоположного F . Вероятность первого события обозначим буквою p , а второго — буквою q , при чем

$$p + q = 1 \dots \dots \dots (1)$$

Эти вероятности одни и те же при всех испытаниях.

Пусть сложное событие состоит из последовательного наступления простых событий E и F в таком порядке:

$$\frac{EE \dots E}{m_1} \frac{FF \dots F}{n_1} \frac{EE \dots E}{m_2} \frac{FF \dots F}{n_2} \dots \frac{EE \dots E}{m_k} \frac{FF \dots F}{n_k} \dots \dots (2)$$

при чем:

$$\begin{aligned} m_1 + m_2 + \dots + m_k &= m, \\ n_1 + n_2 + \dots + n_k &= n, \\ m + n &= p \dots \dots \dots (3) \end{aligned}$$

Вероятность сложного события (2) будет такова:

$$\frac{pp \dots p}{m_1} \frac{qq \dots q}{n_1} \frac{pp \dots p}{m_2} \frac{qq \dots q}{n_2} \dots \frac{pp \dots p}{m_k} \frac{qq \dots q}{n_k} = p^m q^n \dots \dots (4)$$

Так выражается вероятность того, что при p испытаниях событие E произойдет m раз, а событие F — n раз, при чем эти события будут чередоваться в определенной последовательности, как показывает схема (2). Для краткости будем называть ее вероятностью схемы (2).

Если мы хотим найти вероятность того, что при p испытаниях событие E случится m раз, а событие F — n раз, независимо от порядка чередования этих событий, то на основании теоремы сложения должны будем составить сумму вероятностей всех

схем, где, подобно (2), буква E встречается m раз, а буква F — n раз. Так как вероятности всех этих схем одинаковы и равны (4), то мы должны будем умножить вероятность схемы (4), т. е. $p^m q^n$ на число различных схем, где m и n имеют те же значения, как и в схеме (2).

Чтобы найти число их, заметим, что в схеме (2) всего μ букв. Совершив всевозможные перемещения этих μ букв, получим $\mu!$ схем; но не все они различны потому, что в группе (2) буква E повторяется m раз, а буква F — n раз; между буквами E можно сделать $m!$ перестановок, которые не изменяют схему (2); вследствие этого число различных схем (2) уменьшится в $m!$ раз; так же точно между буквами F можно сделать $n!$ перестановок, которые не будут менять схему (2), а потому число различных схем еще уменьшится в $n!$ раз.

В результате найдем, что число различных схем вида (2) равно:

$$\frac{\mu!}{m!n!}.$$

Умножив на него величину $p^m q^n$, получим выражение:

$$\frac{\mu!}{m!n!} p^m q^n,$$

определяющее, как мы видели выше, вероятность сложного события, состоящего в том, что при μ испытаниях событие E состоится m раз, а событие F — n раз. Обозначив эту вероятность буквою P_{μ} , можем написать:

$$P_{\mu} = \frac{\mu!}{m!n!} p^m q^n. \dots \dots \dots (5)$$

Если присоединить условия, что $P_0 = q^{\mu}$, $P_{\mu} = p^{\mu}$, то можно сказать, что выражение (5) есть формула общего члена бинома Ньютона:

$$(p + q)^{\mu}.$$

Следовательно,

$$(p + q)^{\mu} = P_0 + P_1 + P_2 + P_3 + \dots + P_{\mu}.$$

На основании равенства:

$$p + q = 1 \dots \dots \dots (1)$$

отсюда следует

$$P_0 + P_1 + P_2 + P_3 + \dots + P_{\mu} = 1, \dots \dots \dots (6)$$

как и должно быть потому, что величины:

$$P_0, P_1, P_2, P_3, \dots, P_{\mu} \dots \dots \dots (7)$$

суть вероятности событий всевозможных и взаимно исключающих друг друга.

Чтобы составить себе конкретное представление о том законе, которому следуют члены формулы (6), или о так называемом биномиальном законе, обратимся к частным примерам: положим $p = \frac{1}{3}$, $q = \frac{2}{3}$ и выполним вычисление членов по формуле (5), положив $\mu = 6$ и $\mu = 12$; затем положим $p = \frac{2}{5}$, $q = \frac{3}{5}$ и выполним вычисление членов по формуле (5), положив $\mu = 20$.

Результаты получатся следующие:

$$\left(\frac{1}{3} + \frac{2}{3}\right)^6 = 0,088 + 0,26 + \underline{0,33} + 0,22 + 0,082 + 0,015 + 0,0014 = 0,99;$$

$$\begin{aligned} \left(\frac{1}{3} + \frac{2}{3}\right)^{12} &= 0,008 + 0,046 + 0,123 + 0,212 + \underline{0,238} + 0,191 + 0,111 + \\ &+ 0,048 + 0,015 + 0,003 + 0,0004 + 0,00004 + 0,000002 = \\ &= 0,99; \end{aligned}$$

$$\begin{aligned} \left(\frac{2}{5} + \frac{3}{5}\right)^{20} &= 0,00004 + 0,00049 + 0,00309 + 0,01235 + 0,03499 + \\ &+ 0,07435 + 0,12441 + 0,16588 + \underline{0,17970} + 0,15970 + \\ &+ 0,11714 + 0,07999 + 0,03550 + 0,01456 + 0,00485 + \\ &+ 0,00129 + 0,00027 + 0,00004 + 0,00000 + 0,00000 + \\ &+ 0,00000 = 0,9996. \end{aligned}$$

Рассматривая эти числовые ряды, мы видим, что в начале и конце каждого из них стоят члены очень малые; по мере удаления от концов к середине каждого из них, члены ряда возрастают; вблизи середины каждого ряда стоит один член самый большой в ряде. Такую же особенность мы могли бы заметить во всех числовых рядах подобных трем рассмотренным. Кроме того, мы видим, что по мере увеличения μ члены ряда делаются меньше; это и понятно: сумма всех членов биномиального ряда равна 1, все члены положительны, число же их $\mu + 1$ увеличивается с возрастанием μ .

Так как члены ряда (6) выражают вероятности различных сложных событий при μ испытаниях, то наибольший из этих членов дает вероятность того сложного события, для которого эта вероятность наибольшая. Мы назовем это событие *наивероятнейшим*. Соответствующее ему значение m есть *наивероятнейшее* число появлений события E при μ испытаниях, а соответствующая ему величина $\frac{m}{\mu}$ — *наивероятнейшая* величина отношения числа появлений события E к числу всех испытаний.

§ 2. Наивероятнейшее сложное событие при большом числе испытаний.

Пусть P_m наибольший член ряда (6); в таком случае он больше чем P_{m-1} и больше, чем P_{m+1} , откуда:

$$\frac{P_m}{P_{m-1}} > 1, \quad \frac{P_m}{P_{m+1}} > 1;$$

но на основании равенства:

$$P_m = \frac{\mu!}{m! n!} p^m q^n \dots \dots \dots (5)$$

имеем:

$$P_{m-1} = \frac{\mu!}{(m-1)! (n+1)!} p^{m-1} q^{n+1}, \quad P_{m+1} = \frac{\mu!}{(m+1)! (n-1)!} p^{m+1} q^{n-1}.$$

Поэтому предшествующие неравенства примут вид:

$$\frac{n+1}{m} \cdot \frac{p}{q} > 1, \quad \frac{m+1}{n} \cdot \frac{q}{p} > 1,$$

или, приняв во внимание, что

$$m + n = \mu, \quad p + q = 1,$$

находим:

$$\left. \begin{aligned} m &< \mu p + p, \\ m &> \mu p - q. \end{aligned} \right\} \dots \dots \dots (8)$$

Эти неравенства устанавливают границы, между которыми лежит наивероятнейшее число m .

Так как μ — число произведенных испытаний — зависит от нашего выбора, то всегда можно взять такое число испытаний μ , чтобы $\mu p + p$ не было числом целым: тогда и $\mu p - q$ будет дробным. Разность между этими двумя границами:

$$(\mu p + p) - (\mu p - q) = p + q = 1.$$

Следовательно, неравенства (8) определяют одно и только одно целое число m : это число m есть наивероятнейшее число появлений события E при μ испытаниях.

Из неравенств (8) следует:

$$\left. \begin{aligned} \frac{m}{\mu} &< p + \frac{p}{\mu}, \\ \frac{m}{\mu} &> p - \frac{q}{\mu}. \end{aligned} \right\} \dots \dots \dots (9)$$

В дальнейших приложениях формул мы будем полагать μ числом весьма большим, а в теоретических рассуждениях будем называть его бесконечно-большим; это надо понимать следующим образом: так как число испытаний μ зависит от нашего выбора, то мы всегда

можем произвести испытаний столько, чтобы число их μ было больше какого угодно постоянного положительного числа.

Если μ весьма велико, то $\frac{p}{\mu}$ и $\frac{q}{\mu}$ весьма малы: обе границы, устанавливаемые неравенствами (9), очень близки между собою; их разность равна $\frac{1}{\mu}$; поэтому $\frac{m}{\mu}$ очень близко к p и приближенно может быть принято равным p . Ошибка которую мы делаем, принимая $\frac{m}{\mu}$ равною p , меньше величины $\frac{1}{\mu}$. Предел величины $\frac{m}{\mu}$ при $\mu = \infty$ есть p .

§ 3. Графическое изображение членов биномиального ряда.

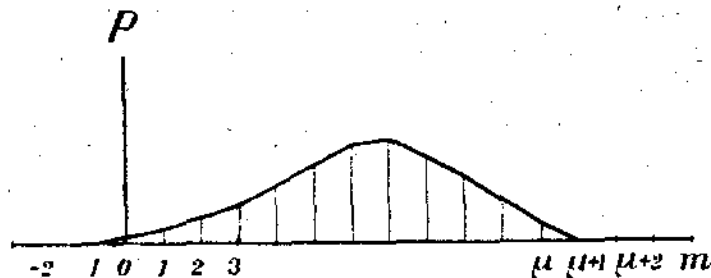
Для уяснения особенностей биномиального ряда целесообразно применить геометрический метод. Прежде всего возьмем прямоугольные оси координат, отметим на оси абсцисс точки с абсциссами:

$$0, 1, 2, 3, \dots, \mu$$

и возставим в них перпендикуляры к оси абсцисс по длине соответственно равные:

$$P_0, P_1, P_2, P_3, \dots, P_\mu.$$

Соединив прямыми линиями конечные точки этих перпендикуляров, получим ломаную, характеризующую закон изменения членов биномиального ряда с изменением номера члена. Для дальнейших рассуждений удобно продолжить эту ломаную безгранично влево и вправо, положив, что ординаты ее вершин равны нулю при всех



Чертеж 5.

отрицательных значениях абсциссы и при всех положительных ее значениях, больших μ . Получившаяся ломаная совпадает с осью абсцисс от $-\infty$ до точки -1 , после чего начинает подниматься, достигает наибольшей высоты над осью абсцисс в том месте, которое соответствует наибольшему члену бинома, затем спускается к оси абсцисс, достигает оси при абсциссе равной $\mu + 1$ и далее до $+\infty$ совпадает с осью абсцисс. Такая ломаная линия схематически изображена на черт. 5.

Площадь s , лежащая между осью абсцисс и построенной ломаной, представляет собою сумму площадей прямолинейных трапеций. Применяя формулу площади трапеции, получим:

$$s = \frac{0 + P_0}{2} \cdot 1 + \frac{P_0 + P_1}{2} \cdot 1 + \frac{P_1 + P_2}{2} \cdot 1 + \dots + \frac{P_{\mu-1} + P_\mu}{2} \cdot 1 + \\ + \frac{P_\mu + 0}{2} \cdot 1 = P_0 + P_1 + P_2 + \dots + P_\mu.$$

На основании формулы (6):

$$s = 1.$$

Следовательно, площадь ломаной линии одна и та же и равна 1 для всевозможных биномов.

Чтобы изучить особенности биномиального закона для различных биномов, мы должны построить ломаные для этих биномов и сравнить их между собой.

Но такое непосредственное сравнение представляет многие неудобства: 1) длина той части ломаной, где она поднимается над осью абсцисс (базис линии) различна при различных значениях μ ; она беспрестанно возрастает с возрастанием μ ; 2) ординаты всех вершин ломаной уменьшаются с увеличением μ ; уменьшается в том числе и наибольшая из ординат, вследствие чего при больших значениях μ ломаная всюду очень близка к оси абсцисс; 3) абсцисса наивысшей точки ломаной меняется вместе с изменением μ .

Для устранения первых двух неудобств мы меняем масштаб по оси абсцисс и по оси ординат: масштаб по оси абсцисс уменьшаем в μ раз, а по оси ординат — в μ раз увеличиваем. Тогда разность между абсциссами соседних вершин ломаной будет $\frac{1}{\mu}$, а вместе с тем и базис уменьшается в μ раз: для всех ломаных базис будет одинаков и равен 1.

Ордината P_m точки ломаной заменится величиной:

$$y = \mu P_m. \quad (10)$$

Точка первоначально взятой ломаной с абсциссой m преобразуется в точку новой ломаной с абсциссой $\frac{m}{\mu}$. Так как мы видели, что абсцисса наивысшей вершины ломаной определяется неравенствами (8), то абсцисса соответствующей наивысшей точки преобразованной ломаной будет определяться неравенствами (9). При μ весьма большом эта величина очень близка к p ; поэтому можно приблизительно сказать, что абсцисса наивысшей точки преобразованной ломаной равна p . Обозначим буквою O основание перпендикуляра, опущенного на ось абсцисс из наивысшей точки ломаной.

Наконец, чтобы устранить последнее из отмеченных выше неудобств прежней ломаной, преобразуем оси координат, перенеся начало в точку O ; если обозначим абсциссу после преобразования буквою x , то сделанное преобразование выразится формулою:

$$\frac{m}{\mu} = p + x \dots \dots \dots (11)$$

При увеличении m на 1 абсцисса x увеличивается на:

$$dx = \frac{1}{\mu} \dots \dots \dots (12)$$

Такова разность между абсциссами двух соседних вершин преобразованной ломаной.

Рассмотрим свойства преобразованной ломаной сравнительно с прежней.

1) Базис преобразованной линии равен 1, как мы уже говорили выше.

2) Ординаты вершин ломаной при большом значении μ уже не будут так малы, как прежде; чтобы иметь понятие о кривой после преобразования, составим числовой ряд, соответствующий третьему из рассмотренных выше рядов.

Умножив на величину $\mu = 20$ написанные выше величины членов бинома $\left(\frac{2}{5} + \frac{3}{5}\right)^{20}$, находим следующие пары значений x и y :

x	y	x	y
-0,40	0,0008	0,15	1,4198
-0,35	0,0098	0,20	0,7100
-0,30	0,0618	0,25	0,2912
-0,25	0,2470	0,30	0,0970
-0,20	0,6998	0,35	0,0258
-0,15	1,4870	0,40	0,0054
-0,10	2,4882	0,45	0,0008
-0,05	3,3176	0,50	0,0000
0,00	3,5940	0,55	0,0000
0,05	3,1940	0,60	0,0000
0,10	2,3428		19,9928

На этом примере видно, что наибольшая ордината преобразованной ломаной имеет величину довольно значительную (3,5940), по мере приближения к концам ряда члены убывают быстро, и при значениях $\pm 0,40$ выражаются уже в тысячных долях.

Подобные же вычисления для μ более значительного, чем 20, показали бы, что наибольшая ордината возрастает вместе с увеличением μ , а убывание ординаты с приближением к концам ряда делается еще более быстрым.

3) Наивысшая точка преобразованной ломаной при всяком μ соответствует $x=0$, это прямо следует из сделанного выше преобразования координат.

4) Площадь преобразованной ломаной равна 1. В самом деле, так как масштаб по оси абсцисс в μ раз уменьшился, а по оси ординат в μ раз увеличился, то площади всех трапеций, составляющих ломаную, не изменились.

Эти свойства преобразованной ломаной делают ее весьма пригодной для изучения биномиального закона вероятностей сложных событий при многократном повторении испытаний.

Для достижения цели остается сделать еще последний шаг: заменить ломаную линию плавной, аналитической кривой, которая возможно близко подходила бы к изучаемой ломаной. Такую аналитическую кривую мы назовем интерполяционной кривой, а соответствующую ей формулу — интерполяционной формулой биномиального закона вероятностей сложных событий при многократном повторении испытаний.

§ 4. Интерполяционная формула для вероятностей сложных событий при многократном повторении испытаний.

Возьмем прямоугольные оси координат и будем откладывать по оси абсцисс величины x , определяемые равенством:

$$\frac{m}{\mu} = p + x, \dots \dots \dots (11)$$

а по перпендикулярам в концах отложенных абсцисс величины:

$$y = \mu P_m \dots \dots \dots (10)$$

Конечные точки этих перпендикуляров будут вершинами преобразованной ломаной. Наша задача — составить уравнение аналитической кривой, возможно близко подходящей к этой ломаной.

Возьмем две последовательные вершины ломаной: одна пусть соответствует некоторому данному значению m , а другая — следующему значению $m+1$. Координаты первой точки на основании равенств (11), (10) и (5) будут:

$$\left. \begin{aligned} x &= \frac{m}{\mu} - p, \\ y &= \mu \cdot \frac{\mu!}{m! n!} p^m q^n, \end{aligned} \right\} \dots \dots \dots (13)$$

при чем:

$$m + n = \mu, \quad p + q = 1.$$

Координаты следующей точки:

$$\left. \begin{aligned} x + \Delta x &= \frac{m+1}{\mu} - p, \\ y + \Delta y &= \mu P_{m+1} = \mu \cdot \frac{\mu!}{(m+1)!(n-1)!} p^{m+1} q^{n-1}, \end{aligned} \right\} \dots (14)$$

при чем на основании равенств (13) и (12):

$$\Delta x = \frac{1}{\mu} = dx.$$

Разделив друг на друга вторые равенства (14) и (13), находим

$$\frac{y + \Delta y}{y} = \frac{np}{(m+1)q},$$

откуда:

$$\frac{\Delta y}{y} = \frac{np - (m+1)q}{(m+1)q},$$

или деля на Δx , заменяя n его значением $\mu - m$ и помня, что $p + q = 1$:

$$\frac{1}{y} \frac{\Delta y}{\Delta x} = \frac{(\mu - m) p - (m+1) q}{(m+1) q} \mu = \frac{\mu p - m - q}{mq + q} \mu = \frac{p - \frac{m}{\mu} - \frac{q}{\mu}}{q \frac{m}{\mu} + \frac{q}{\mu}} \mu.$$

Пользуясь первым из равенств (13), представляем эту формулу в таком виде:

$$\frac{1}{y} \frac{\Delta y}{\Delta x} = \frac{-x - \frac{q}{\mu}}{q(p+x) + \frac{q}{\mu}} \mu \dots (15)$$

Эта формула выведена для всякого целого положительного μ . Теперь посмотрим, какие в ней произойдут упрощения, если принять во внимание, что μ число весьма большое.

Величины p и q мы предполагаем конечными, а потому $\frac{q}{\mu}$, как в числителе так и в знаменателе формулы (15), при μ бесконечно-большом бесконечно-мало и в приближенном вычислении может быть опущено.

Далее, величина x есть разность между $\frac{m}{\mu}$ и p ; мы видели, что при большом значении μ величина y очень быстро убывает с увеличением этой разности x . Величина y имеет конечные значения только при весьма малых значениях x . Поэтому мы в праве при рассмотрении формулы (15) считать x весьма малым и опустить второе слагаемое в выражении $(p+x)$, стоящем в знаменателе.

Тогда правая часть равенства (15) примет вид:

$$-\frac{\mu x}{p q}.$$

В левой части этого равенства отношение $\frac{\Delta y}{\Delta x}$ имеет пределом своим $\frac{dy}{dx}$; поэтому левая часть равенства (15) приближенно равна:

$$\frac{1}{y} \frac{dy}{dx}.$$

Все равенство (15) при весьма большом μ приближенно заменится следующим:

$$\frac{1}{y} \frac{dy}{dx} = -\frac{\mu x}{p q}. \quad (16)$$

Интегрируя уравнение (16), находим:

$$\lg y = -\frac{\mu x^2}{2 p q} + \lg C,$$

где C произвольное постоянное. Обозначив буквою h положительное число:

$$h = \sqrt{\frac{\mu}{2 p q}}, \quad (17)$$

представим найденный интеграл в таком виде:

$$y = C e^{-h^2 x^2}. \quad (18)$$

Такова интерполяционная формула для вероятностей сложных событий при многократном повторении испытаний. Остается только найти C .

Для этого припомним, что площадь, заключенная между осью абсцисс и преобразованной ломаной равна 1. Заметим, что ломаную мы считаем беспрестанно простирающеюся влево и вправо, при чем точки ее, достаточно удаленные от начала координат, лежат на самой оси абсцисс. Поэтому, понятно, что и площадь, лежащая между осью x и интерполяционной кривой (18), должна равняться 1:

$$C \int_{-\infty}^{+\infty} e^{-h^2 x^2} dx = 1.$$

Преобразуя переменное интегрирования подстановкою:

$$hx = t,$$

найдем:

$$\frac{C}{h} \int_{-\infty}^{+\infty} e^{-t^2} dt = 1.$$

Получившийся интеграл есть известный интеграл Пуассона он равен $\sqrt{\pi}$; поэтому находим:

$$\frac{C}{h} \sqrt{\pi} = 1,$$

откуда:

$$C = \frac{h}{\sqrt{\pi}}.$$

Интерполяционная формула для вероятностей сложных событий при большом числе испытаний такова:

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2}, \quad (19)$$

где:

$$h = \sqrt{\frac{\mu}{2pq}}. \quad (17)$$

Мы обозначали буквою P_m вероятность того, что при p испытаниях событие E произойдет m раз, и обозначили буквою y величину:

$$y = p P_m; \quad (10)$$

кроме того мы ввели обозначение:

$$dx = \frac{1}{\mu}. \quad (12)$$

Поэтому теперь приближенно можем сказать, что

$$P_m = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} dx \quad (20)$$

выражает вероятность того, что при p испытаниях событие E состоится m раз, при чем x определяется из равенства (18), т. е.

$$x = \frac{m}{\mu} - p. \quad (21)$$

Величина $\frac{m}{\mu}$ есть отношение числа m всех появлений события E к числу произведенных испытаний p , т. е. та величина, которую мы называли в § 5 Введения частотой; в § 2 настоящей главы мы видели, что наиболее вероятное значение этой частоты есть p . Следовательно, x есть отклонение частоты от ее наиболее вероятного ее значения.

Полученный выше результат можно формулировать так: вероятность того, что действительно наблюдаемая при p испытаниях частота $\frac{m}{\mu}$ будет отклоняться от

своего наимвероятнейшего значения p на величину x , приближенно выражается формулой:

$$P_m = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} dx^*), \dots \dots \dots (20)$$

где:

$$h = \sqrt{\frac{\mu}{2pq}}, \dots \dots \dots (17)$$

$$dx = \frac{1}{\mu} \dots \dots \dots (12)$$

Для того, чтобы составить себе понятие, насколько формула (20) дает близкие к истине результаты, вычислим по ней величины P_m

при $p = \frac{2}{5}$, $q = \frac{3}{5}$, $\mu = 20$ и сравним их с приведенными выше чле-

нами бинома $\left(\frac{2}{5} + \frac{3}{5}\right)^{20}$

x	Величина P_m		x	Величина P_m	
	истинная	по форм. (20)		истинная	по форм. (20)
-0,40	0,0000	0,0002	0,05	0,160	0,163
-0,35	0,0005	0,0009	0,10	0,117	0,118
-0,30	0,003	0,004	0,15	0,074	0,071
-0,25	0,012	0,014	0,20	0,036	0,035
-0,20	0,035	0,035	0,25	0,015	0,014
-0,15	0,074	0,071	0,30	0,005	0,004
-0,10	0,124	0,118	0,35	0,0013	0,0009
-0,05	0,166	0,163	0,40	<u>0,0003</u>	<u>0,0002</u>
-0,00	0,180	0,182		1,000	0,994

Если принять во внимание, что 20 далеко не большое число и что формула (20) выведена в предположении, что μ велико, то надо признать, что приведенная таблица обнаруживает согласие формулы с результатами непосредственного вычисления гораздо большее, чем можно было бы ожидать при таких условиях. При большем значении μ это согласие еще больше.

*) Для вычисления значений этой функции служит табл. I в конце книги.

Чтобы составить себе понятие о том, насколько быстро убывает функция (20) при возрастании x , когда μ велико, приведем некоторые результаты вычисления по этой формуле при $\mu = 2500$.

x	P_m
0,000	0,016287
0,020	0,0020280
0,032	0,000078632
0,040	0,0000039148
0,060	0,0000000019318

Из сказанного в настоящем параграфе следует, что при большом значении μ и при значениях p и q не слишком близких к 0 или 1 интерполяционная формула (20) дает хорошее приближение для вычисления P_m .

Она различается только множителем

$$dx = \frac{1}{\mu} \dots \dots \dots (12)$$

от формулы

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} \dots \dots \dots (19)$$

которую мы называли интерполяционной формулой для закона вероятностей сложных событий при большом числе испытаний.

Ближайшей нашей задачей будет исследование вида кривой (19).

§ 5. Вид интерполяционной кривой.

Из формулы (19) непосредственно видно, что выражаемая ею кривая имеет следующие свойства:

1) наибольшая ордината ее соответствует $x = 0$, т.е. высшая точка кривой лежит на оси y .

2) Ордината этой точки — вершины кривой — выражается формулой:

$$y_0 = \frac{h}{\sqrt{\pi}} \dots \dots \dots (22)$$

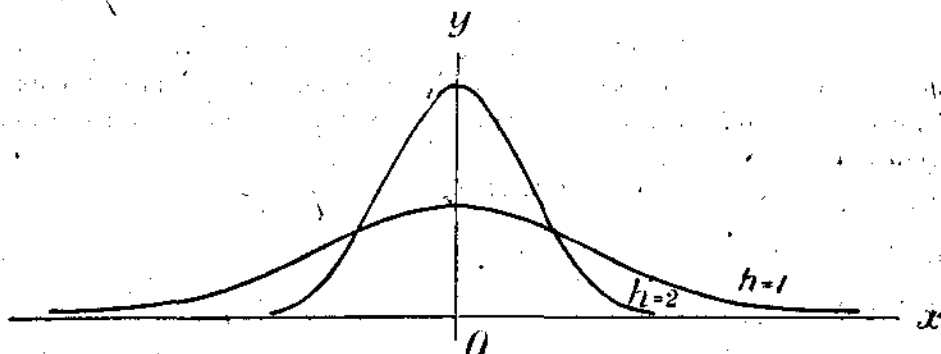
3) Кривая симметрична относительно оси y .

4) Она асимптотична относительно оси x .

В предшествующем параграфе мы видели, что площадь между кривою и ее асимптотой равна 1 при всяком значении h .

Сопоставим это свойство со свойством (22). Возьмем две кривые с параметрами: первую h_1 , вторую h_2 , при чем $h_1 < h_2$. По свойству (22) точки второй кривой при малых значениях x имеют большие ординаты, чем соответствующие точки первой кривой; площадь второй кривой вблизи оси y выходит за пределы первой. Так как площади обеих кривых одинаковы, то должен существовать промежуток, где площадь первой кривой выступает за пределы второй. Отсюда следует, что кривые должны пересекаться. Между началом координат и точками пересечения кривых ордината точки первой кривой меньше, а вне этого промежутка — больше ординаты соответствующей точки второй кривой.

Такое расположение точек кривых видно на черт. 6, где изображены кривые (22): 1) для $h = 1$, 2) для $h = 2$.



Черт. 6.

Это можно выразить такими словами: площадь кривой (19) тем теснее группируется около оси y , чем больше h .

С точки зрения теории вероятностей это значит следующее: с увеличением h вероятности малых значений отклонения

$$x = \frac{m}{n} - p \dots \dots \dots (21)$$

возрастают, а больших уменьшаются. Следовательно, при большом h с большим основанием можно считать частоту $\frac{m}{n}$ близкою к вероятности p , чем при малых значениях этого параметра. Поэтому величина h получила название: мера точности.

Из чертежа 6 видно, что кривая (19) вблизи $x = 0$ обращена к оси x выпуклостью, а при больших значениях x — вогнутостью. Следовательно, она должна иметь точки перегиба. Для нахождения их приравняем нулю вторую производную выражения (19):

$$e^{-h^2 x^2} - 2h^2 x^2 e^{-h^2 x^2} = 0,$$

откуда:

$$x = \pm \frac{1}{h\sqrt{2}}.$$

Величина $\frac{1}{h\sqrt{2}}$ имеет конкретное значение в теории вероятностей и носит название: среднее квадратическое уклонение. Обозначив его буквой σ , будем иметь:

$$\sigma = \frac{1}{h\sqrt{2}} \quad (28)$$

Итак, кривая (19) имеет две точки перегиба с абсциссами $\pm \sigma$.

● Наглядное представление о виде кривой (19) при довольно большом значении h можно получить, пользуясь числовыми величинами примера в конце предшествующего параграфа.

Мера точности в этом случае такова:

$$h = \sqrt{\frac{2500}{2 \cdot \frac{2}{5} \cdot \frac{3}{5}}} = 72,16875.$$

Ординаты точек кривой (19) в данном случае больше соответствующих значений P_m в $\mu = 2500$ раз. Поэтому соответствующие значения x и y расположатся в таблицу:

x	y
0,000	40,72
0,020	5,070
0,032	0,49658
0,040	0,009787
0,060	0,00000048

Мы видим, что кривая имеет вид профиля иглы высотой 40,7, а шириною вблизи основания 0,08. Хотя кривая и имеет точки, бесконечно далеко лежащие от начала координат, но их высота над осью абсцисс очень мала: при x больше, чем 0,04 ординаты точек кривой можно считать равными нулю и считать кривую простирающуюся только от $-0,04$ до $+0,04$, т.е. с базисом длиной 0,08.

Точки перегиба кривой имеют абсциссы:

$$\pm \sigma = \pm 0,0098;$$

они очень близки к началу координат.

Если бы мы пожелали вычислить площадь криволинейной трапеции, ограниченной осью x , рассмотренной кривой и двумя ординатами, проведенными по обе стороны

от нее даже в весьма малом расстоянии, то заранее могли бы сказать, что эта площадь исчерпывает собою почти всю площадь кривой.

Осуществить этот чертеж для рассмотренной кривой было бы затруднительно, так как кривая имела бы вид профиля очень тонкой и длинной иглы. Поэтому ниже на фигуре 7 изображен чертеж для случая небольшого значения $\mu = 20$, а именно для рассмотренного уже нами бинома:

$$\left(\frac{2}{5} + \frac{3}{5}\right)^{20}$$

Площадь криволинейной трапеции для данного случая, ограниченная ординатами, соответствующими абсциссам $-0,2$ и $+0,2$, на чертеже заштрихована. Из чертежа видно, что часть площади, оставшаяся незаштрихованной, составляет очень малую часть всей площади.

Если бы можно было осуществить такой же чертеж для μ более значительного, чем 20, то эта особенность кривой была бы видна еще сильнее. В следующем параграфе мы увидим, что она составляет сущность основной теоремы, которая характеризует отклонение частоты от вероятности.

§ 6. Теорема Якова Бернулли.

Вернемся к формулам § 1 и возьмем ряд величин:

$$P_0, P_1, P_2, P_3, \dots, P_\mu, \dots \quad (7)$$

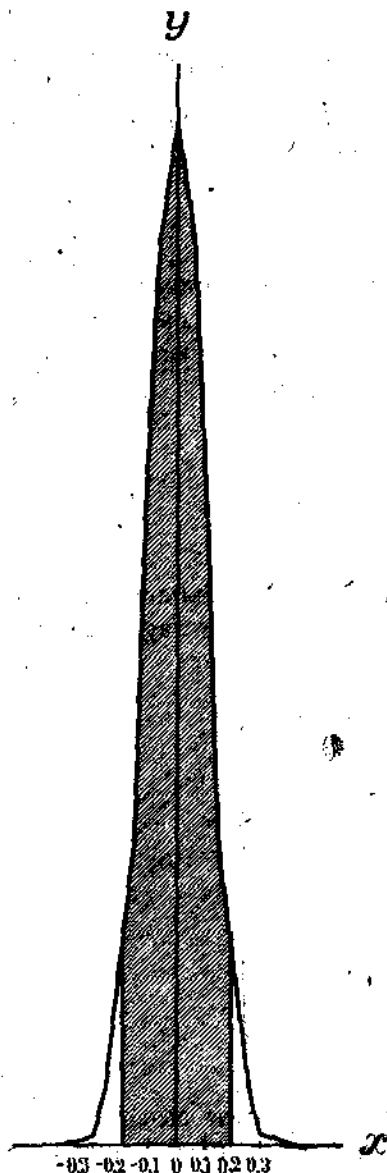
выражающих вероятности всевозможных сложных событий при μ испытаниях.

Пусть k и l два каких-нибудь целых числа из ряда:

$$0, 1, 2, 3, \dots, \mu,$$

при чем

$$k < l.$$



Чертеж 7.

Составим сумму:

$$P_k + P_{k+1} + P_{k+2} + \dots + P_l = \sum_{m=k}^{m=l} P_m \dots \dots \dots (24)$$

По теореме сложения вероятностей она выразит собою вероятность того, что при μ испытаниях событие E случится или k , или $k+1$, или $k+2 \dots$, или l раз, т. е. что E случится не менее k и не более l раз.

Так как каждому значению m соответствует определенная величина $\frac{m}{\mu}$ и определенная величина

$$x = \frac{m}{\mu} - p, \dots \dots \dots (21)$$

то можно сказать, что сумма (24) выражает собою вероятность того, что при μ испытаниях отклонение x будет не меньше, чем

$$a = \frac{k}{\mu} - p$$

и не больше, чем:

$$b = \frac{l}{\mu} - p.$$

Иными словами, величина (24) есть вероятность того, что отклонение x содержится между пределами a и b .

Пользуясь обозначениями:

$$y = \mu P_m \dots \dots \dots (10)$$

$$dx = \frac{1}{\mu} \dots \dots \dots (12)$$

представим сумму (24) в таком виде:

$$\sum_{x=a}^{x=b} y dx, \dots \dots \dots (25)$$

где слагаемые соответствуют значениям переменного x :

$$a, a + dx, a + 2dx, a + 3dx, \dots, b.$$

Здесь y есть та функция, которую мы рассматривали в § 4 и которая приближенно выражается формулою:

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} \dots \dots \dots (19)$$

Из интегрального исчисления известно, что предел суммы (25) выражается определенным интегралом:

$$\int_a^b y dx,$$

т.е.

$$\frac{h}{\sqrt{\pi}} \int_a^b e^{-hx^2} dx \dots \dots \dots (26)$$

При большом значении μ , когда dx близко к 0, величина (25) близка к своему пределу (26). Поэтому (26) приближенно выражает собою величину (25); иными словами, величина (26) приближенно равна вероятности того, что при μ испытаниях отклонение x лежит между пределами a и b .

Положим величину $a = -\alpha$, $b = +\alpha$, где α есть какое угодно положительное число, и положим:

$$P = \frac{h}{\sqrt{\pi}} \int_{-\alpha}^{+\alpha} e^{-hx^2} dx,$$

или

$$P = \frac{2h}{\sqrt{\pi}} \int_0^{+\alpha} e^{-hx^2} dx,$$

или, преобразуя переменное x подстановкой $hx = t$:

$$P = \frac{2}{\sqrt{\pi}} \int_0^{h\alpha} e^{-t^2} dt \dots \dots \dots (27)$$

Такова приближенная величина вероятности того, что при μ испытаниях отклонение

$$x = \frac{m}{\mu} - p \dots \dots \dots (21)$$

будет лежать между $-\alpha$ и $+\alpha$, т.е. по абсолютной величине не превзойдет α .

Мы будем называть эту величину вероятностью предела α отклонения частоты от вероятности.

Если мы обозначим символом $\Phi(x)$ выражение:

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt, \dots \dots \dots (28)$$

то равенство (27) примет такой вид:

$$P = \Phi(h\alpha) \dots \dots \dots (29)$$

Свойства функции $\Phi(x)$ известны из интегрального исчисления.

Главнейшие из них следующие:

- 1) При $x=0$ функция $\Phi(x)$ равна нулю:

$$\Phi(0) = 0;$$

- 2) при x положительном она положительна;

- 3) предел ее при $x=+\infty$ есть 1:

$$\Phi(+\infty) = 1 \dots \dots \dots (30)$$

Величина функции $\Phi(x)$ для всякого x может быть вычислена через разложение функции e^{-x^2} в степенной ряд и почленное интегрирование. Таблицы значений функции $\Phi(x)$ были вычислены с большою точностью. Такие таблицы приложены в конце этой книги *); здесь приведем только некоторые величины из этих таблиц.

x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000 000	2,50	0,9995 930
0,50	0,5204 999	3,00	0,9999 779
1,00	0,8427 008	3,50	0,9999 9925 691
1,50	0,9661 052	4,00	0,9999 9998 458
2,00	0,9953 223	4,80	0,9999 9999 999

Отсюда видно, что при значении hx даже небольшом, как, напр., 2,5 или 3, величина (29) настолько близка к единице, что может быть принята за единицу.

Если мы обозначим буквою γ такое число, что $\Phi(\gamma)$ может быть принято за 1, то будем в праве почти с достоверностью утверждать, что отклонение x по абсолютному значению не превзойдет величины ha , определяемой равенством:

$$ha = \gamma,$$

откуда:

$$a = \frac{\gamma}{h} \dots \dots \dots (31)$$

Эта величина a при большом n , а следовательно, и h , будет весьма мала.

Полученный результат можно формулировать в виде теоремы, называемой теоремой Якова Бернулли: с вероятностью весьма близкой к достоверности можно утверждать, что отклонение частоты от вероятности при большом числе испытаний n не превзойдет по абсолютному значению величины очень малой.

В конце § 2 Введения мы говорили, что в приложениях теории вероятностей к опытным наукам заранее указывают ту дробь близкую

*) См. табл. 2, в конце книги.

к 1, напр., 0,99 или 0,999, или 0,9999 и т. д., которую мы приближенно принимаем равной 1. Всякую правильную дробь, большую ее, считаем за 1, а событие, вероятность которого выражается такою дробью, считаем достоверным.

Применяя это к данному случаю, мы обозначаем буквою γ такое число, что $\Phi(\gamma)$ приближенно принимается за 1. В таком случае с достоверностью скажем, что отклонение x не превзойдет величины

$$\alpha = \frac{\gamma}{h}; \quad (31)$$

это крайний возможный предел отклонений.

Говоря это, мы должны помнить, что в основе такого утверждения лежит предположка опытного происхождения по которой событие с очень малою вероятностью встречается настолько редко, что мы его считаем невозможным; напр., нам представляется невозможным вынуть черный шар из урны, куда положено 1000 шаров, из которых черный один, а остальные белые.

Пусть произведены две серии испытаний над одним и тем же событием E : в первой серии r_1 испытаний, а во второй — r_2 . Соответствующие значения h выразятся так:

$$h_1 = \sqrt{\frac{r_1}{2pq}}, \quad h_2 = \sqrt{\frac{r_2}{2pq}}.$$

Выбрав значение γ , мы из формулы (31) найдем крайние возможные пределы отклонений для той и другой серии:

$$\alpha_1 = \frac{\gamma}{h_1}, \quad \alpha_2 = \frac{\gamma}{h_2}.$$

Ясно, что та серия, где величина h больше, будет иметь меньший возможный предел отклонений. Следовательно, произведя серию опытов и найдя из нее частоту $\frac{m}{n}$ мы меньше рискуем сделать

большую ошибку, принимая $\frac{m}{n}$ равным p в той серии, где h больше.

Величина h характеризует меру точности, с которой мы можем рассматривать частоту, как приближенную величину вероятности.

Это уже отчасти видно из сказанного в § 5. Так же точно теорема Якова Бернулли служит аналитическим выражением того свойства интерполяционной кривой, о котором мы говорили в конце § 5.

В самом деле, величина

$$P = \frac{h}{\sqrt{\pi}} \int_{-\alpha}^{+\alpha} e^{-h^2 x^2} dx,$$

есть площадь криволинейной трапеции, заключенной между осью x , кривою и ординатами ее, соответствующими абсциссам $-\alpha$ и $+\alpha$. Теорема Якова Бернулли утверждает, что эта площадь

$$P = \Phi(h\alpha) \dots \dots \dots (29)$$

при большом значении h исчерпывает собою почти всю площадь кривой при

$$\alpha = \frac{\gamma}{h}, \dots \dots \dots (31)$$

хотя α имеет величину небольшую.

Кроме величины h возможный предел уклонений может быть характеризован некоторыми другими величинами, называемыми вероятным, средним и средним квадратическим уклонениями.

§ 7. Вероятное уклонение.

Пусть r такое число, что с одинаковым основанием можно утверждать как то, что уклонение x больше его, так и то, что уклонение x меньше его. Иначе говоря, вероятность предела r равна $\frac{1}{2}$. В таком случае будем иметь:

$$\Phi(hr) = \frac{1}{2}.$$

Обозначив буквою p число, определяемое из уравнения:

$$\Phi(p) = \frac{1}{2}, \dots \dots \dots (32)$$

мы находим величину p , пользуясь таблицами функции $\Phi(x)$. Оказывается, что:

$$p = 0,4769 \quad 3627 \quad 62.$$

Итак:

$$hr = p,$$

откуда

$$r = \frac{p}{h} \dots \dots \dots (38)$$

Эта величина r называется вероятным уклонением. Мы видим, что она обратно-пропорциональна мере точности h .

§ 8. Среднее уклонение.

Прежде, чем приступить к уяснению нового понятия о среднем уклонении, взглянем с новой точки зрения на известное уже нам понятие о математическом ожидании.

Пусть некоторый игрок принимает участие в игре, где вероятность его на выигрыши равна p , а величина выигрыша A рублей. Пусть он сделал при таких условиях большое число n партий.

Обозначим буквою m число выигранных им партий. На основании теоремы Якова Бернулли, можем сказать, что частость $\frac{m}{n}$ выигрыша очень близка к вероятности p . Поэтому $\frac{m}{n}$ можно принять за приближенную величину p , а следовательно, m можем принять приближенно равным np .

Так как после каждой выигранной партии игрок получает A рублей, то сумма его выигрышей от всех выигранных им m партий составит mA или приближенно npA рублей.

Если такова сумма, выигранная за все n партий, в которых игрок принял участие, то в среднем выигрыш его от одной партии равен:

$$\frac{npA}{n} = Ap \text{ рублей.}$$

Эта величина Ap , как мы знаем, есть математическое ожидание.

Итак, математическое ожидание игрока есть величина среднего выигрыша его на каждую партию.

Если некоторый игрок может получить одну из случайных сумм:

$$A_1, A_2, A_3, \dots, A_k,$$

в зависимости от того, случится ли одно из событий:

$$E_1, E_2, E_3, \dots, E_k,$$

вероятности которых суть:

$$p_1, p_2, p_3, \dots, p_k,$$

то его математическое ожидание, как мы знаем, таково:

$$A_1 p_1 + A_2 p_2 + A_3 p_3 + \dots + A_k p_k.$$

Ясно, что это будет опять средний выигрыш данного лица от одной партии.

В приложениях теории вероятностей под названием случайной суммы приходится подразумевать не только выигрыш игрока, но и вообще значение какого-либо переменного, зависящее от случая. Так как сущность дела остается та же самая, как и в рассмотренном нами конкретном примере, то мы можем сказать, что всегда математическое ожидание случайной величины переменного есть среднее значение этого переменного.

Применим только что сказанное к нахождению средней величины уклонения.

Но предварительно заметим, что x есть величина алгебраическая: уклонение частоты $\frac{m}{n}$ от вероятности p в одну сторону мы

считаем положительным, а в другую — отрицательным. Но всякое отклонение, в какую бы сторону оно ни было, есть удаление от истины; а потому мы будем говорить об абсолютном значении $|x|$ и будем определять среднюю величину его.

Ясно, что x есть случайная величина; вероятность ее выражается формулой:

$$y dx.$$

Математическое ожидание величины $|x|$ представится суммой выражений вида

$$y |x| dx$$

для всевозможных значений x .

Так же, как при выводе теоремы Бернулли, найдем, что эта сумма приближенно выразится интегралом:

$$\int_{-\infty}^{+\infty} y |x| dx,$$

или заменяя y его выражением (19):

$$\frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2 x^2} |x| dx.$$

Это математическое ожидание величины $|x|$ на основании сказанного выше есть средняя величина абсолютного значения отклонения.

Называя ее коротко — среднее отклонение и обозначая буквою d , можем написать:

$$d = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2 x^2} |x| dx.$$

Производя интегрирование, находим:

$$d = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2 x^2} |x| dx = \frac{2h}{\sqrt{\pi}} \int_0^{\infty} e^{-h^2 x^2} x dx = \frac{1}{h\sqrt{\pi}} \left[-e^{-h^2 x^2} \right]_0^{\infty} = \frac{1}{h\sqrt{\pi}}.$$

Итак среднее отклонение выражается формулой:

$$d = \frac{1}{h\sqrt{\pi}} \quad (33')$$

Примечание. Если бы мы взяли не абсолютную величину отклонения $|x|$, а величину алгебраическую x , то математическое ожидание выразилось бы формулой:

$$\frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2 x^2} x dx.$$

Этот интеграл равен нулю вследствие того, что подынтегральная функция нечетная.

Такой результат можно было бы заранее ожидать: так как равные отклонения в ту и другую сторону равно вероятны, то они при составлении суммы сократятся, и сумма всех отклонений, а следовательно, и средняя величина их должна равняться 0.

§ 9. Среднее квадратическое отклонение.

Чтобы избежать влияния знака величины x можно вычислить среднюю величину квадрата x^2 ; она выразится интегралом:

$$\frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2 x^2} x^2 dx.$$

Квадратный корень из этой величины называется средним квадратическим отклонением. Обозначив его буквою σ , будем иметь:

$$\sigma^2 = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2 x^2} x^2 dx.$$

Преобразуя переменные интегрирования подстановкой $hx = t$ и проведя интегрирование по частям, получим:

$$\sigma^2 = \frac{1}{h^2 \sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} t^2 dt = \frac{-1}{2h^2 \sqrt{\pi}} \left[e^{-t^2} t \right]_{-\infty}^{+\infty} + \frac{1}{2h^2 \sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} dt = \frac{1}{2h^2}.$$

Отсюда

$$\sigma = \frac{1}{h\sqrt{2}} \quad (33'')$$

Таково выражение среднего квадратического отклонения. Сравнив его с найденной выше формулой (23), видим, что среднее квадратическое отклонение равно абсциссе точки перегиба интерполяционной кривой (19).

Таков геометрический смысл среднего квадратического отклонения. Соединяя вместе формулы (33), (33') и (33''), находим:

$$h = \frac{\rho}{r} = \frac{1}{d\sqrt{\pi}} = \frac{1}{\sigma\sqrt{2}} \quad (34)$$

Следовательно, все три величины r , d и σ прямо-пропорциональны друг другу и обратно-пропорциональны мере точности h .

§ 10. Примеры приложения теоремы Якова Бернулли.

В виде примеров приложения теоремы Бернулли рассмотрим результат некоторых опытов, при чем два первые, принадлежащие Вестергарду, помещены в 1-й таблице § 4 Введения, а последний

касается геометрической вероятности в задаче об игле, помещенной § 8, главы I.

Пример 1. В урну было положено 20 белых шаров и 20 черных. Из нее было произведено вынутие шара 10000 раз, при чем вынутый шар каждый раз возвращался в урну. Оказалось, что белый шар появился 5011 раз, а черный 4989.

Так как в данном случае $p = q = \frac{1}{2}$, $\mu = 10000$, то $\frac{m}{\mu} = 0,5011$. Уклонение x в данном случае равно:

$$x = 0,5011 - 0,5 = 0,0011.$$

Мера точности равна

$$h = \sqrt{\frac{10000}{2 \cdot \frac{1}{2} \cdot \frac{1}{2}}} = 100 \sqrt{2} = 141,421.$$

Если условимся принимать за единицу всякую правильную дробь, не меньшую 0,9999, то из таблиц $\Phi(x)$ найдем, что величина γ , определяемая равенством:

$$\Phi(\gamma) = 0,9999,$$

такова:

$$\gamma = 2,76.$$

Обозначая буквою α крайний возможный предел уклонения, найдем:

$$h\alpha = 2,76; \alpha = \frac{2,76}{100 \sqrt{2}} = 0,0195.$$

Наблюдаемое в действительности уклонение 0,0011 значительно меньше этого предела.

Если бы оказалось, что при испытаниях, произведенных 1000 раз при тех же условиях, уклонение $\frac{m}{\mu} - p$ оказалось больше, чем 0,0195, то мы были бы в праве объяснить это какими-нибудь специальными причинами; так, например, могла произойти ошибка при счете шаров, положенных в урну, или шары могли быть плохо перемешаны, или неодинакового радиуса, или наконец лицо, вынимавшее шары, различало их на осязание и брало преимущественно шары определенного цвета сознательно или невольно.

Вероятное уклонение в данном случае равно: $r = \frac{p}{h} = 0,0033$; среднее

уклонение $d = \frac{1}{h \sqrt{\pi}} = 0,0039$; среднее квадратич. уклонен. $\sigma = \frac{1}{h \sqrt{2}} = 0,005$.

Пример 2. В копенгагенской лотерее было подсчитано число вышедших билетов: 1) от 1 до 45. 2) от 46 до 90. Оказалось, что к первой категории принадлежало 3640 номеров, а ко второй — остальные 3635 номеров.

В данном случае

$$\mu = 7275, m = \frac{3640}{7275} = 0,50034; p = \frac{1}{2},$$

$$\frac{m}{\mu} - p = 0,00034;$$

$$h = \sqrt{\frac{7275}{2 \cdot \frac{1}{2} \cdot \frac{1}{2}}} = \sqrt{14550} = 120,6.$$

Положив снова $\gamma = 2,76$, находим:

$$\alpha = \frac{\gamma}{h} = 0,022885.$$

Следовательно, возможное отклонение $\frac{m}{\mu} - p$ не больше 0,022885. Наблюденное в действительности отклонение 0,00034 значительно меньше этого предела.

Вероятное отклонение равно

$$r = \frac{\rho}{h} = 0,00895,$$

среднее отклонение

$$d = \frac{1}{h\sqrt{\pi}} = 0,00459,$$

среднее квадратическое отклонение

$$\sigma = \frac{1}{h\sqrt{2}} = 0,00587.$$

Пример 3. Проф. Вольф в Цюрихе сделал опыт бросания иглы на плоскость, разграфленную параллельными прямыми.

Расстояние между параллельными прямыми было $2a = 45$ м/м, длина иглы $2l = 36$ м/м.

Мы видели, что вероятность пересечения иглой какой-либо из параллельных выражается так:

$$p = \frac{2l}{\pi a} = \frac{2 \cdot 36}{\pi \cdot 45} = \frac{8}{5\pi} = \frac{1,6}{\pi}.$$

Вольф произвел опыт 5000 раз, при чем игла пересекла параллельную 2532 раз; следовательно, частота оказалась равною:

$$\frac{2532}{5000} = 0,5064.$$

Зная величину π , можно вычислить вероятность p и сравнить ее с частотой подобно тому, как мы поступали в двух предшествующих примерах; но можно сделать еще иной вывод из полученного результата: воспользоваться им для вычисления приближенного значения π .

На основании теоремы Вернулли вероятность p близка к частоте, которая в данном случае равна 0,5064. Принимая эту величину за приближенное значение p , вычисляем h

$$h = \sqrt{\frac{5000}{2 \cdot 0,5064 \cdot 0,4936}} = 100.$$

Полагая $\gamma = 2,76$, находим

$$\alpha = \frac{\gamma}{h} = 0,0276.$$

Поэтому отклонение частоты от вероятности не может быть больше 0,0276:

$$\left| \frac{1,6}{\pi} - 0,5064 \right| < 0,0276.$$

откуда

$$\left| \frac{1}{\pi} - 0,3165 \right| < 0,01725.$$

Иначе $\frac{1}{\pi}$ разнится от 0,3165 на величину меньшую, чем 0,01725:

$$\frac{1}{\pi} = 0,3165 + \varepsilon,$$

где

$$|\varepsilon| < 0,01725.$$

Поэтому

$$\pi = \frac{1}{0,3165 + \varepsilon} = \frac{1}{0,3165} - \frac{\varepsilon}{0,3165^2},$$

полагая $\frac{\varepsilon}{0,3165^2} = \theta$ находим:

$$\pi = 3,1596 - \theta,$$

где

$$|\theta| < 0,172.$$

Впоследствии такие опыты были повторяемы и привели к следующим результатам*):

М. А. Smith (1855) из 3204 бросаний получил для π 3,1553.

Fox (1894) 1120 " " " π 3,1419

Lazzarini (1901) . . 3408 " " " π 3,1415929.

§ 11. Рассмотрение одного примера, где сложное событие состоит из большого числа простых событий, между собою зависящих.

Пусть в урну положено весьма большое число M шаров, из которых Mr белых и Mq черных, где r и q суть две дроби, сумма которых равна 1:

$$r + q = 1.$$

Из этой урны вынуто друг за другом μ шаров, при чем вынутый шар в урну не возвращался. Предполагая, что μ меньше, чем Mr и Mq , найти вероятность, что в числе μ вынутых шаров будет m белых и n черных при чем:

$$m + n = \mu.$$

В данном случае сложное событие состоит из μ простых событий, между собою зависящих.

Мы обозначим искомую вероятность буквою P_m , при чем m может равняться одному из чисел: 1, 2, 3, ..., μ . Так как в таком случае мы будем иметь вероятности событий всевозможных в данных условиях и взаимно исключающих друг друга, то

$$P_1 + P_2 + P_3 + \dots + P_m = 1. \dots \dots \dots (35)$$

Решение рассматриваемой задачи не изменится, если мы изменим ее условия следующим образом: пусть из урны вынуто μ шаров не друг за другом, а сразу. Спрашивается, какова вероятность, что в числе вынутых шаров будет m белых и n черных.

*) Periodico di matematica. 1901, стр. 140 — 143. См. Czuber. Wahrscheinlichkeitsrechnung, 1908, стр. 152.

Такую задачу мы уже рассматривали в главе I, § 1 (задача 3), где она была сформулирована так: в урну положено α белых и β черных шаров, всего $\gamma = \alpha + \beta$ шаров. Из нее вынимают зараз μ шаров. Спрашивается, какова вероятность, что в числе вынутых шаров будет α' белых и β' черных, при чем

$$\alpha' + \beta' = \gamma'.$$

Решение этой задачи выразилось формулой:

$$\frac{C_{\alpha}^{\alpha'} \cdot C_{\beta}^{\beta'}}{C_{\gamma}^{\gamma'}}.$$

В рассматриваемом теперь случае мы имеем:

$$\gamma = M, \alpha = Mp, \beta = Mq, \alpha' = m, \beta' = n, \gamma' = \mu.$$

Поэтому:

$$P_m = \frac{C_{Mp}^m \cdot C_{Mq}^n}{C_M^{\mu}},$$

или

$$P_m = \frac{\mu! (M - \mu)! (Mp)! (Mq)!}{M! m! (Mp - m)! n! (Mq - n)!} \quad (36)$$

К изучению этой формулы применим тот же прием, который был применен выше при изучении биномиального закона. Применим графический метод, рассматривая m как абсциссу, P_m как ординату. Получится некоторая ломаная. Ее преобразуем, уменьшая абсциссу в μ раз и увеличивая ординату в μ раз. Преобразованную ординату обозначим буквой y :

$$y = \mu P_m; \quad (37)$$

ломаную линию продолжим до бесконечности влево и вправо, считая нулями ординаты ее точек с абсциссами, меньшими нуля и большими, чем 1. Площадь этой ломаной будет равна 1 на основании равенства (35). Тем же способом, который изложен в § 4, строим интерполяционную формулу для данного случая.

Пользуясь формулами (36) и (37), составляем формулу:

$$\frac{y + \Delta y}{y} = \frac{(Mp - m) n}{(m + 1) (Mq - n + 1)},$$

откуда:

$$\frac{\Delta y}{y} = \frac{(Mq - m) n - (Mq - n + 1)(m + 1)}{(m + 1) (Mq - n + 1)}.$$

Величины M и μ мы рассматриваем, как весьма большие; поэтому в целях упрощения приближенной формулы можем опустить конечное

число 1 там, где оно входит слагаемым при этих больших числах, формула принимает вид:

$$\frac{\Delta y}{y} = \frac{Mpn - Mqt}{m(Mq - n)}$$

или на основании равенств

$$\mu = m + n, \quad p + q = 1:$$

$$\frac{\Delta y}{y} = \frac{M\mu \left(p - \frac{m}{\mu} \right)}{m(Mq - \mu + m)} \dots \dots \dots (38)$$

Числитель формулы (38) обращается в 0, когда $\frac{m}{\mu}$ делается равным p ; следовательно, и левая часть обращается в 0 при $\frac{m}{\mu} = p$. Это означает, что при такой величине $\frac{m}{\mu}$ функция y достигает своего наибольшего значения; интерполяционная кривая имеет наивысшую точку при абсциссе, равной p .

Преобразуем оси координат подстановкой:

$$\frac{m}{\mu} = p + x,$$

откуда:

$$x = \frac{m}{\mu} - p. \dots \dots \dots (39)$$

Тогда ось ординат будет проходить через наивысшую точку кривой.

Из равенства (39) следует:

$$\Delta x = dx = \frac{1}{\mu} \dots \dots \dots (40)$$

Совершая преобразование (39) в формуле (38) и деля полученное равенство на (40), находим:

$$\frac{1}{y} \frac{\Delta y}{\Delta x} = \frac{-Mx}{(p+x) \left(\frac{M}{\mu} q - 1 + p + x \right)}$$

Так как для нас представляет интерес та часть кривой, которая находится вблизи наивысшей точки, то мы можем считать x малым и опускать его в тех членах формулы, где оно входит слагаемым при величине конечной. Сделав это упрощение в последней формуле, заменив, кроме того, отношение $\frac{\Delta y}{\Delta x}$ его пределом $\frac{dy}{dx}$ и приняв во вни-

мание, что $1 - p = q$, находим:

$$\frac{1}{y} \frac{dy}{dx} = \frac{-Mx}{pq \left(\frac{M}{p} - 1 \right)}$$

Интегрируя это равенство, находим:

$$y = Ce^{-\frac{M\mu x^2}{2pq(M-\mu)}} \quad (41)$$

где C постоянное число. Величина его определится из условия, чтобы площадь, заключенная между осью абсцисс и самой интерполяционной кривой, равнялась 1:

$$C \int_{-\infty}^{+\infty} e^{-\frac{M\mu x^2}{2pq(M-\mu)}} dx = 1.$$

Выполнив преобразование:

$$\sqrt{\frac{M\mu}{2pq(M-\mu)}} x = t$$

и положив для краткости:

$$\sqrt{\frac{M\mu}{2pq(M-\mu)}} = h, \quad (42)$$

находим:

$$\frac{C}{h} \int_{-\infty}^{+\infty} e^{-t^2} dt = 1,$$

или

$$\frac{C\sqrt{\pi}}{h} = 1,$$

откуда:

$$C = \frac{h}{\sqrt{\pi}}.$$

Формула (41) примет вид:

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} \quad (43)$$

Из равенств (37), (40), (43) следует:

$$P_m = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} dx, \quad (44)$$

где h определяется формулой (42).

Полученные формулы имеют большое сходство с полученными выше формулами для случая независимых между собою простых событий; но между этими формулами есть и существенное различие.

1. Так же, как и в прежнем случае, интерполяционная кривая имеет одну наивысшую точку, так называемую вершину, с абсциссой $\frac{m}{\mu} = p$; но прежде эта величина p имела значение вероятности появления белого шара при всех испытаниях: это была основная вероятность (*grundwahrscheinlichkeit*), постоянная в течение всех опытов; теперь же в каждом испытании вероятность различна, смотря по результатам предшествующих испытаний. Величина p есть вероятность до начала опытов.

2. Величина:

$$x = \frac{m}{\mu} - p \quad \dots \dots \dots (39)$$

есть отклонение частоты от этой величины p .

3. Величины y , соответствующие $-x$ и $+x$, равны между собою, т. е. одинаковые отклонения в ту и другую сторону от p равно вероятны.

4. Интерполяционная кривая выражается уравнением такого же вида, как и в предшествующем случае, но мера точности была иная при независимых простых событиях; она выражалась формулой:

$$h = \sqrt{\frac{\mu}{2pq}}; \quad \dots \dots \dots (17)$$

в рассматриваемом же случае она выражается так:

$$h = \sqrt{\frac{M\mu}{2pq(M-\mu)}} \quad \dots \dots \dots (42)$$

Обозначая для отличия меру точности, даваемую последней формулой, буквою h' и полагая, что p одинаково в обоих случаях, находим:

$$h' = \sqrt{\frac{\mu}{2pq}} \cdot \sqrt{\frac{M}{M-\mu}},$$

или

$$h' = \sqrt{\frac{M}{M-\mu}} \cdot h \quad \dots \dots \dots (45)$$

Так как $\sqrt{\frac{M}{M-\mu}}$ больше 1, то $h' > h$.

Мера точности в рассматриваемом случае больше, чем в случае простых событий, независимых между собой, с основной вероятностью, равной той же величине p .

Припомним сказанное в § 5, находим, что в рассматриваемом теперь случае точки площади интерполяционной кривой теснее группируются около оси y , они менее рассеяны по плоскости, чем в случае с основной вероятностью.

Это выражают словами: кривая в рассматриваемом случае имеет меньшую дисперсию, чем в прежнем случае.

Кривая, соответствующая случаю с постоянной вероятностью, называется кривой с нормальной дисперсией.

Кривую, подобную найденной в настоящем параграфе, мы назовем кривую с поднормальной дисперсией.

Если бы встретился случай, где интерполяционная кривая выражается формулой такого же вида, как (43), но с мерой точности меньшей, чем (17), то мы назвали бы ее кривой с сверхнормальной дисперсией.

Обращаясь к формуле (45), мы видим, что h' разнится от h множителем:

$$\sqrt{\frac{M}{M-\mu}}.$$

Оставляя μ постоянным, станем увеличивать M . По мере увеличения M этот множитель приближается к 1, величина h' приближается к h , дисперсия кривой (43) приближается к норме. При $M = \infty$ будем иметь $h' = h$; кривая (43) делается нормальной. Это и понятно: когда число шаров в урне очень велико, то взятие из нее одного шара почти не влияет на величины вероятностей появления белого и черного шара при следующем испытании.

Уясним теперь, почему дисперсия уменьшается, когда вынутые шары не возвращаются в урну.

Всякий раз, как из урны вышел шар определенного цвета, вероятность вынудия шара того же цвета при следующем испытании уменьшается, а для цвета противоположного — увеличивается. Происходит как бы компенсация в числе появлений шаров того и другого цвета, вследствие чего дисперсия уменьшается. Когда число всех шаров в урне не велико, то влияние такой компенсации очень заметно, а потому и дисперсия значительно ниже нормы.

Объинтегрировав выражение (43) по x в пределах от $-a$ до $+a$, найдем величину:

$$P = \frac{h}{\sqrt{\pi}} \int_{-a}^{+a} e^{-h^2 x^2} dx = \frac{2}{\sqrt{\pi}} \int_0^{ha} e^{-t^2} dt = \Phi(ha),$$

определяющую вероятность того, что отклонение

$$x = \frac{m}{\mu} - p \dots \dots \dots (39)$$

по абсолютной величине не превышает a .

Смысл этого равенства такой же, каков смысл соответствующего равенства в теореме Бернулли. Разница только в том, что в рассматриваемом теперь случае h больше, а потому возможный крайний предел отклонений меньше.

§ 12. Случай, когда сложное событие состоит из простых событий, зависящих от двух причин с различными вероятностями.

Пусть над появлением события E произведены две серии испытаний: в первой серии вероятность события E равна p_1 , события ему противоположного q_1 , всего испытаний произведено n_1 ; во второй серии вероятность события E равна p_2 , события ему противоположного q_2 , всего испытаний произведено n_2 . Число испытаний в обеих сериях

$$n_1 + n_2 = n \dots \dots \dots (46)$$

Спрашивается, какова вероятность, что событие E произойдет всего в обеих сериях n раз.

Для ясности можем взять такой конкретный пример: имеются две урны; для первой из них вероятность вынуть белого шара равна p_1 , черного q_1 ; для второй урны вероятность вынуть белого шара равна p_2 и черного q_2 ; из первой урны вынули шар n_1 раз, из второй n_2 раз, возвращая каждый раз вынутый шар в урну. Спрашивается, какова вероятность, что при всех этих испытаниях белый шар появился n раз.

Обозначим искомую вероятность буквою P_n .

Вероятность того, что при n_1 испытаниях первой серии событие E состоится m_1 раз, обозначим символом $P_{m_1}^I$, а вероятность того, что при n_2 испытаниях второй серии событие E состоится m_2 раз, обозначим символом $P_{m_2}^{II}$.

Положив

$$m_1 + m_2 = n, \dots \dots \dots (47)$$

можем сказать на основании теоремы умножения, что

$$P_{m_1}^I P_{m_2}^{II} \dots \dots \dots (48)$$

есть вероятность одного из тех случаев, когда имеет место то сложное событие, вероятность которого обозначена буквою P_n .

Для нахождения величины P_n мы должны по теореме сложения в формуле (48) дать m_1 и m_2 всевозможные значения, удовлетворяющие уравнению (47) и результаты сложить.

Из равенства (47) следует, что $m_2 = n - m_1$; поэтому формула (48) примет вид:

$$P_{m_1}^I P_{n-m_1}^{II} \dots \dots \dots (49)$$

Свойства и приближенные выражения функций $P_{m_1}^I$ и $P_{n-m_1}^{II}$ нам известны из сказанного в параграфах 1 — 5 настоящей главы: это суть функции, выражающие биномиальный закон. На основании сказанного

в названных параграфах мы можем считать равными нулю значения функции P_m^I при m_1 отрицательном и при m_1 большем μ_1 , а значения P_m^II равным нулю при m_2 отрицательном и при m_2 большем μ_2 . Поэтому составляя сумму всевозможных значений выражения (49), можем в нем давать величине m_1 всевозможные целые значения от $-\infty$ до $+\infty$.

Величина P_m , которая равна сумме всевозможных значений функций (49), выразится так:

$$P_m = \sum_{m_1=-\infty}^{m_1=+\infty} P_{m_1}^I P_{m-m_1}^{II} \dots \dots \dots (50)$$

Вместо $P_{m_1}^I$ и $P_{m_1}^{II}$ вводим их приближенные выражения:

$$P_{m_1}^I = \frac{h_1}{\sqrt{\pi}} e^{-h_1^2 x_1^2} dx_1, \quad P_{m_1}^{II} = \frac{h_2}{\sqrt{\pi}} e^{-h_2^2 x_2^2} dx_2 \dots \dots \dots (51)$$

где:

$$\left. \begin{aligned} x_1 &= \frac{m_1}{\mu_1} - p_1, \quad x_2 = \frac{m_2}{\mu_2} - p_2, \\ h_1 &= \sqrt{\frac{\mu_1}{2p_1q_1}}, \quad h_2 = \sqrt{\frac{\mu_2}{2p_2q_2}}, \\ dx_1 &= \frac{1}{\mu_1}, \quad dx_2 = \frac{1}{\mu_2}. \end{aligned} \right\} \dots \dots \dots (52)$$

Равенство (50) примет вид:

$$P_m = \frac{h_1 h_2}{\pi} \sum_{m_1=-\infty}^{m_1=+\infty} e^{-h_1^2 x_1^2 - h_2^2 x_2^2} dx_1 dx_2 \dots \dots \dots (53)$$

Из равенств (47) и (52) следует:

$$m = \mu_1 p_1 + \mu_2 p_2 + \mu_1 x_1 + \mu_2 x_2,$$

откуда:

$$\frac{m}{\mu} = \frac{\mu_1 p_1 + \mu_2 p_2}{\mu} + \frac{\mu_1}{\mu} x_1 + \frac{\mu_2}{\mu} x_2.$$

Введя обозначения:

$$\frac{\mu_1 p_1 + \mu_2 p_2}{\mu} = p, \dots \dots \dots (54)$$

$$\frac{\mu_1}{\mu} x_1 + \frac{\mu_2}{\mu} x_2 = x, \dots \dots \dots (55)$$

находим:

$$\frac{m}{\mu} = p + x \dots \dots \dots (56)$$

Из сказанного о биномиальном законе мы знаем, что p_1 и p_2 суть наимвероятнейшие значения частот $\frac{m_1}{\mu_1}$ и $\frac{m_2}{\mu_2}$, а потому наимвероятнейшие значения уклонений x_1 и x_2 равны нулю. Отсюда, из равенств (55) и (56) следует, что и наимвероятнейшее значение x равно нулю, а наимвероятнейшее значение $\frac{m}{\mu}$ равно p .

Итак, величина p , определяемая равенством (54), есть наимвероятнейшее значение частоты $\frac{m}{\mu}$ при всех произведенных испытаниях обеих серий.

При вычислении вероятности P_m величина m рассматривается как постоянная; так как p тоже постоянно, то из равенства (56) следует, что и x для данного P_m постоянно.

Из равенства (55) следует:

$$x_2 = \frac{\mu}{\mu_2} x - \frac{\mu_1}{\mu_2} x_1,$$

и формула (53) принимает вид:

$$P_m = \frac{h_1 h_2}{\pi} \sum_{x_1 = -\infty}^{x_1 = +\infty} e^{-h_1^2 x_1^2 - h_2^2 \left(\frac{\mu}{\mu_2} x - \frac{\mu_1}{\mu_2} x_1 \right)^2} dx_1 dx_2, \quad \dots (57)$$

при чем в этой сумме величина x_1 получает всевозможные значения ряда:

$$0, \pm \frac{1}{\mu_1}, \pm \frac{2}{\mu_1}, \pm \frac{3}{\mu_1}, \dots$$

В таком случае, как мы знаем, сумма (57) приближенно может быть заменена интегралом. Сделаем эту замену и кроме того представим показателя степени под знаком суммы в более простом виде:

$$- (ax_1^2 + 2bx_1 + c),$$

где:

$$a = h_1^2 + h_2^2 \frac{\mu_1^2}{\mu_2^2}, \quad b = -h_2^2 \frac{\mu}{\mu_2^2} x, \quad c = h_2^2 \frac{\mu^2}{\mu_2^2} x^2. \quad \dots (58)$$

Получается следующее:

$$P_m = \frac{h_1 h_2}{\pi} dx_2 \int_{-\infty}^{+\infty} e^{-(ax_1^2 + 2bx_1 + c)} dx_1,$$

или

$$P_m = \frac{h_1 h_2}{\pi} dx_2 \int_{-\infty}^{+\infty} e^{-\left(\sqrt{a} x_1 + \frac{b}{\sqrt{a}} \right)^2 + \frac{b^2}{a} - c} dx_1.$$

Вынося постоянный относительно x_1 множитель за знак интеграла и делая постановку

$$\sqrt{a} x_1 + \frac{b}{\sqrt{a}} = t,$$

находим:

$$P_m = \frac{h_1 h_2}{\pi \sqrt{a}} e^{\frac{b^2}{a} - c} dx_2 \int_{-\infty}^{+\infty} e^{-t^2} dt = \frac{h_1 h_2}{\sqrt{a\pi}} e^{\frac{b^2}{a} - c} dx_2.$$

В эту формулу вставим вместо a, b, c их выражения (58) и, кроме того, заметим, что из равенства (55) следует:

$$\frac{\mu_2}{\mu} dx_2 = dx, \quad dx_2 = \frac{\mu}{\mu_2} dx.$$

Получится следующее:

$$P_m = \frac{\mu}{\sqrt{\pi} \sqrt{\frac{\mu_1^2}{h_1^2} + \frac{\mu_2^2}{h_2^2}}} e^{-\frac{\mu^2}{h_1^2 + h_2^2} x^2} dx \quad \dots \quad (59)$$

или, введя обозначение:

$$\frac{\mu}{\sqrt{\frac{\mu_1^2}{h_1^2} + \frac{\mu_2^2}{h_2^2}}} = h, \quad \dots \quad (60)$$

находим:

$$P_m = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} dx \quad \dots \quad (61)$$

- Мы видим, что в рассматриваемом случае частость $\frac{m}{\mu}$ имеет
- наивероятнейшее значение, определяемое формулой:

$$\frac{\mu_1 p_1 + \mu_2 p_2}{\mu} = p, \quad \dots \quad (54)$$

при чем вероятность отклонения x от этой величины выражается формулой (61), подобной формуле вероятности отклонения в случае с одною основной вероятностью, но с иной мерою точности. В данном случае мера точности выражается формулой (60) или, заменяя h_1 и h_2 их выражениями из равенств (52):

$$h = \frac{\mu}{\sqrt{2(\mu_1 p_1 q_1 + \mu_2 p_2 q_2)}} \quad \dots \quad (62)$$

В том частном случае, когда $p_1 = p_2$, мы возвращаемся к опытам с одною вероятностью p , и формула (62) дает выражение меры точности при нормальной дисперсии:

$$h = \sqrt{\frac{\mu}{2pq}} \quad \dots \quad (17)$$

§ 13. Случай, когда сложное событие состоит из простых событий, зависящих от многих причин с различными вероятностями.

Пусть над появлением события Е произведено 3 серии испытаний: в 1-й серии μ_1 испытаний, во 2-й серии μ_2 , в 3-й серии μ_3 испытаний, всего:

$$\mu_1 + \mu_2 + \mu_3 = \mu$$

испытаний. Пусть вероятности события Е и противоположного ему F, в этих сериях соответственно равны:

$$p_1, q_1; \quad p_2, q_2; \quad p_3, q_3.$$

Зададимся вопросом, какова наивероятнейшая величина p частоты $\frac{m}{\mu}$ в результате всех опытов и какова вероятность уклонения x частоты $\frac{m}{\mu}$ от этой наивероятнейшей величины.

Прежде всего заметим, что в рассуждениях предшествующего параграфа, когда мы имели две серии, вначале мы делали только одно допущение, что в той и другой серии вероятности уклонения выражались формулами:

$$P_{m_1}^I = \frac{h_1}{\sqrt{\pi}} e^{-h_1^2 x_1^2} dx_1, \quad P_{m_2}^{II} = \frac{h_2}{\sqrt{\pi}} e^{-h_2^2 x_2^2} dx_2, \dots \dots \dots (51)$$

но не пользовались выражениями параметров h_1, h_2 . Следовательно, для формул (51) дисперсия могла быть и ненормальная. Этим путем мы пришли к равенствам (55), (60) и (61). Только в самом конце параграфа мы преобразовали выражение h , приняв во внимание формулы h_1 и h_2 , даваемые равенствами (52).

Пользуясь этим замечанием, можем свести задачу настоящего параграфа к задаче, решенной в параграфе предшествующем.

Соединим вместе результаты опытов первых двух серий; получится одна вспомогательная серия из

$$\mu' = \mu_1 + \mu_2 \dots \dots \dots (46')$$

опытов.

На основании результатов предшествующего параграфа можем сказать следующее:

1) наивероятнейшее значение частоты в этой серии выражается формулой:

$$p' = \frac{\mu_1 p_1 + \mu_2 p_2}{\mu'} \dots \dots \dots (54')$$

2) вероятность уклонения x' частоты $\frac{m}{\mu'}$ от p' выражается формулой:

$$\frac{h'}{\sqrt{\pi}} e^{-h'^2 x'^2} dx' \dots \dots \dots (61')$$

3) Мера точности h' выражается формулой:

$$h' = \frac{\mu'}{\sqrt{\frac{\mu_1^2}{h_1^2} + \frac{\mu_2^2}{h_2^2}}} \dots \dots \dots (60')$$

Присоединив к результатам испытаний вспомогательной серии, или, что то же, к результатам первых двух серий, результаты испытаний в 3-й серии, мы на том же основании скажем, что:

1) наиболее вероятное значение частоты $\frac{m}{\mu}$ в результате всех опытов выражается формулой:

$$p = \frac{\mu'p' + \mu_2p_2}{\mu} \dots \dots \dots (54'')$$

2) вероятность отклонения x частоты от p выражается формулой:

$$P_m = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} dx; \dots \dots \dots (61'')$$

3) мера точности h выражается формулой:

$$h = \frac{\mu}{\sqrt{\frac{\mu'^2}{h'^2} + \frac{\mu_2^2}{h_2^2}}} \dots \dots \dots (60'')$$

Внеся в равенства (54'') и (60'') выражения (46'), (54'), (60'), найдем:

$$p = \frac{\mu_1 p_1 + \mu_2 p_2 + \mu_3 p_3}{\mu} \dots \dots \dots (54''')$$

$$h = \frac{\mu}{\sqrt{\frac{\mu_1^2}{h_1^2} + \frac{\mu_2^2}{h_2^2} + \frac{\mu_3^2}{h_3^2}}} \dots \dots \dots (60''')$$

Сравнивая формулы (54), (60), (61) и (54'''), (60'''), (61'''), видим, что три последние представляют обобщение трех первых, при чем аналогия настолько ясна, что применение доказательства от n к $n+1$ для распространения ее на случай большего числа серий можно опустить.

Поэтому мы можем сразу формулировать решение такого вопроса:

Над появлением события E произведено n серий опытов, при чем в 1-й серии μ_1 испытаний, во 2-й серии μ_2 испытаний и т. д., наконец, в n -й серии μ_n испытаний, всего:

$$\mu_1 + \mu_2 + \mu_3 + \dots + \mu_n = \mu \dots \dots \dots (46^*)$$

испытаний. Пусть вероятности события Е и ему противоположного F в этих сериях соответственно равны:

$$p_1, q_1; p_2, q_2; \dots; p_k, q_k.$$

Спрашивается, какова наивероятнейшая величина p частости $\frac{m}{\mu}$

в результате всех опытов и какова вероятность уклонения x частости от этой наивероятнейшей величины p .

На основании всего сказанного выше ответ на этот вопрос следующий:

1) величина p выражается формулою:

$$p = \frac{\mu_1 p_1 + \mu_2 p_2 + \mu_3 p_3 + \dots + \mu_k p_k}{\mu} \dots \dots \dots (54^*)$$

2) вероятность уклонения x частости $\frac{m}{\mu}$ от p выражается формулою:

$$P_m = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} dx; \dots \dots \dots (61^*)$$

где

$$h = \frac{\mu}{\sqrt{\frac{\mu_1^2}{h_1^2} + \frac{\mu_2^2}{h_2^2} + \frac{\mu_3^2}{h_3^2} + \dots + \frac{\mu_k^2}{h_k^2}}} \dots \dots \dots (60^*)$$

Так как каждая серия испытаний имеет свою основную вероятность, соответственно равную

$$p_1, p_2, p_3, \dots, p_k$$

то:

$$h_1 = \sqrt{\frac{\mu_1}{2p_1q_1}}, \quad h_2 = \sqrt{\frac{\mu_2}{2p_2q_2}}, \quad h_3 = \sqrt{\frac{\mu_3}{2p_3q_3}}, \dots, \quad h_k = \sqrt{\frac{\mu_k}{2p_kq_k}},$$

и формула (60*) примет вид:

$$h = \frac{\mu}{\sqrt{2(\mu_1 p_1 q_1 + \mu_2 p_2 q_2 + \mu_3 p_3 q_3 + \dots + \mu_k p_k q_k)}} \dots \dots \dots (62^*)$$

Итак, в случае многих причин с различными вероятностями существует наивероятнейшая величина частости $\frac{m}{\mu}$, около которой отдельные наблюдаемые частости группируются подобно тому, как при одной причине с определенной основной вероятностью; но мера точности в случае многих причин выражается иначе: она определяется формулою (62*), и потому дисперсия будет, вообще говоря, не нормальная.

Рассмотрим ближе тот случай, когда во всех сериях число испытаний одинаково:

$$\mu_1 = \mu_2 = \mu_3 = \dots = \mu_k = \frac{\mu}{k}.$$

Формулы (54*) и (62*) примут вид:

$$p = \frac{p_1 + p_2 + p_3 + \dots + p_k}{k}, \dots \dots \dots (63)$$

$$h = \frac{\sqrt{\mu}}{\sqrt{2 \frac{p_1 q_1 + p_2 q_2 + p_3 q_3 + \dots + p_k q_k}{k}}}, \dots \dots \dots (64)$$

Формула (63) показывает, что величина p есть средне-арифметическая из вероятностей $p_1, p_2, p_3, \dots, p_k$ всех групп. Формула (64) показывает, какова мера точности в данном случае.

Выражение, стоящее в знаменателе, есть средне-арифметическое попарных произведений:

$$p_1 q_1; p_2 q_2; p_3 q_3; \dots; p_k q_k;$$

оно не равно произведению

$$pq,$$

где p есть величина (63), а q равно $1 - p$.

Следовательно, величина (64) не равна величине:

$$\sqrt{\frac{\mu}{2pq}}, \dots \dots \dots (65)$$

которую мы получили бы для случая μ испытаний при условии основной вероятности равной p .

Сравним между собою величины (64) и (65).

Пусть:

$$\left. \begin{aligned} p_1 &= p + \varepsilon_1, \\ p_2 &= p + \varepsilon_2, \\ &\dots \dots \dots \\ p_k &= p + \varepsilon_k. \end{aligned} \right\} \dots \dots \dots (66)$$

Отсюда будет следовать:

$$\left. \begin{aligned} q_1 &= q - \varepsilon_1, \\ q_2 &= q - \varepsilon_2, \\ &\dots \dots \dots \\ q_k &= q - \varepsilon_k. \end{aligned} \right\} \dots \dots \dots (67)$$

Сложив между собою равенства (66) и приняв во внимание равенства (63), находим:

$$\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \dots + \varepsilon_k = 0 \dots \dots \dots (68)$$

Перемножив между собою соответствующие равенства (66) и (67) и сложив результаты, находим:

$$p_1 q_1 + p_2 q_2 + \dots + p_k q_k = k p q + (q - p)(\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_k) - (\varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_k^2).$$

На основании (68) находим:

$$\frac{p_1 q_1 + p_2 q_2 + \dots + p_k q_k}{k} = p q - \frac{\varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_k^2}{k}.$$

Так как $\varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_k^2$ существенно положительно и не может быть нулем, если $p_1, p_2, \dots, p_k$ не равны все между собою, то отсюда следует:

$$\frac{p_1 q_1 + p_2 q_2 + \dots + p_k q_k}{k} < p q, \dots \dots \dots (69)$$

а потому величина (64) больше чем (65), т. е. в рассматриваемом случае мы имеем дело с дисперсией поднормальной, так же как и в примере рассмотренном в § 11.

Проверить справедливость этого заключения легко на частном примере.

Пусть у нас имеются две урны с одинаковым числом шаров, но в одной из них шары только белые, а в другой — только черные. Произведя n испытаний над каждой из этих урн, мы получим две серии, из которых в одной записаны только белые шары, а в другой — только черные. Частость белого шара в результате всех испытаний, будет в точности равна $\frac{1}{2}$. Достоверно, что отклонение частоты от

$\frac{1}{2}$ равно 0.

Если же мы переложим шары обеих урн в одну общую урну и произведем над нею n испытаний, то будем иметь обычные условия опытов с урной, где существует основная вероятность, величина которой $\frac{1}{2}$. По теореме Бернулли можно ожидать некоторого отклонения частоты от вероятности, которое будет не велико при большом значении n , но оно все же существовать будет.

§ 14. Закон больших чисел. Теорема Пуассона.

Во всех примерах, рассмотренных выше в настоящей главе, обнаружилось существование наивероятнейшего значения частоты, которое мы обозначали буквою p . Вероятность отклонения x отдельной частоты от величины p выражалась формулою:

$$\frac{h}{\sqrt{\pi}} e^{-h^2 x^2} dx, \dots \dots \dots (A)$$

при чем h более или менее близко, смотря по характеру задачи, к нормальной величине:

$$h = \sqrt{\frac{\mu}{2pq}}, \quad \dots \quad (B)$$

где μ число всех испытаний.

Если событие E при всех произведенных испытаниях имеет одну и ту же вероятность, основную вероятность p , то мера точности строго выражается формулою (B).

Применяя те же рассуждения, как в § 6 при доказательстве теоремы Бернулли, найдем величину вероятности того, что отклонение x по абсолютной величине не превышает α :

$$P = \Phi(h\alpha), \quad \dots \quad (C)$$

где

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx.$$

Найдя такую величину γ , что $\Phi(\gamma)$ приблизительно может быть принята равною 1, можем с достоверностью утверждать, что отклонение x не превзойдет величины:

$$\alpha = \frac{\gamma}{h} \quad \dots \quad (D)$$

В § 10 мы видели, что при $\gamma = 2,76$ величина

$$\Phi(\gamma) = 0,9999;$$

поэтому при h хоть сколько-нибудь значительном величина возможного отклонения α будет очень мала, и полученную из наблюдения частость $\frac{m}{\mu}$ можно принять равною p с ошибкою меньшею, чем α .

Из формулы (B) видно, что при большом μ величина h очень большая, а потому α весьма мало, и ошибка, которую мы делаем, принимая $\frac{m}{\mu}$ равным p чрезвычайно мала.

В других рассмотренных нами примерах с поднормальною дисперсией величина h больше, чем дает формула (B), а потому приближение еще лучше.

На этой особенности сложных событий при большом числе испытаний основывается справедливость утверждения, что частость очень близка к некоторому определенному числу, если число испытаний большое. Это постоянное число p есть вероятность события, если в производимых опытах есть основная вероятность; в более сложных случаях это число называется среднею вероятностью.

Объединяя все случаи рассмотренного характера, можем сказать, что во всех этих случаях при большом числе испытаний час-

тость очень близка к вероятности, при чем эта вероятность может быть или основною или среднею, смотря по условиям задачи.

Это положение носит название закона больших чисел.

На нем основаны все практические приложения теории вероятностей. Без него теория вероятностей имела бы характер только отвлеченной науки без каких-либо приложений к реальному миру.

Самое название „закон больших чисел“ принадлежит Пуассону*), много подвинувшему разработку вопроса, при чем он склонен был придавать закону широкое значение. Надо однако сказать, что закон больших чисел можно считать прочно установленным только для тех случаев, где справедливость формулы (A) теоретически доказана. К числу таких случаев относятся все рассмотренные в предшествующем параграфе и некоторые другие.

В частности, если мы вернемся к случаю, рассмотренному в самом конце предшествующего параграфа и положим $k = \mu$, то найдем:

$$p = \frac{p_1 + p_2 + p_3 + \dots + p_\mu}{\mu} = \frac{\sum p_i}{\mu} \quad (63^*)$$

$$h = \frac{\sqrt{\mu}}{\sqrt{2 \frac{p_1 q_1 + p_2 q_2 + p_3 q_3 + \dots + p_\mu q_\mu}{\mu}}} = \frac{\sqrt{\mu}}{\sqrt{2 \frac{\sum p_i q_i}{\mu}}} \quad (64^*)$$

Эти формулы относятся к следующему случаю:

Произведено μ испытаний над событием E , при чем вероятности события E и ему противоположного F во всех испытаниях различны и соответственно равны:

$$p_1, q_1; p_2, q_2; p_3, q_3; \dots; p_\mu, q_\mu.$$

Формулы (63*) и (64*) показывают, что в таких случаях средняя вероятность есть средне-арифметическое вероятностей

$$p_1, p_2, p_3, \dots, p_\mu,$$

а мера точности выражается формулой (64*). К такому случаю применимы формулы (A) и (C). Это предложение носит название теорема Пуассона. Ее иногда называют законом больших чисел в более узком смысле.

Если в формулах (63*), 64*) положить:

$$p_1 = p_2 = p_3 = \dots = p_\mu = p,$$

то первая из них превратится в тождество, а вторая примет вид:

$$h = \sqrt{\frac{\mu}{2pq}}, \quad (B)$$

*) Poisson. Recherches sur la probabilité des jugements. Paris. 1837.

т. е. мы получим теорему Бернулли как частный случай теоремы Пуассона.

Доказательство этих двух основных теорем Пуассона и Бернулли в иной форме и из иных соображений мы получим еще раз ниже, исходя из теоремы Чебышева.

Академик Чебышев дал теорему большей общности, доказательство которой можно назвать элементарным, так как оно не требует приложения интегрального исчисления. Правда, она не дает выражения искомой вероятности, а только показывает тесные пределы, между которыми эта вероятность заключается; но она имеет то достоинство, что свободна от пользования приближенными вычислениями, какие мы делали в предшествующих параграфах.

§ 15. Теорема Чебышева.

Пусть имеется весьма большое число μ переменных:

$$x, y, z, \dots, u, v, \dots \quad (70)$$

Пусть x должно получить одно, и только одно из значений:

$$x_1, x_2, x_3, \dots, x_a,$$

при чем вероятности этих значений суть:

$$p_1, p_2, p_3, \dots, p_a,$$

а сумма этих вероятностей равна 1, что мы выразим равенством:

$$\sum p_i = 1.$$

Подобные же обозначения введем и для остальных переменных, а именно: для y всевозможные и взаимно исключающие друг друга значения:

$$y_1, y_2, y_3, \dots, y_b,$$

вероятности их:

$$q_1, q_2, q_3, \dots, q_b,$$

при чем:

$$\sum q_i = 1;$$

для z всевозможные и взаимно исключающие друг друга значения:

$$z_1, z_2, z_3, \dots, z_c,$$

вероятности их:

$$r_1, r_2, r_3, \dots, r_c,$$

при чем:

$$\sum r_i = 1,$$

и т. д.

Сумму

$$x_1 p_1 + x_2 p_2 + x_3 p_3 + \dots + x_a p_a$$

мы назовем математическим ожиданием величины x и обозначим буквою m_1 .

Это мы выразим формулой:

$$\sum x_i p_i = a_1.$$

Таким же образом составим выражение:

$$x_1^2 p_1 + x_2^2 p_2 + x_3^2 p_3 + \dots + x_a^2 p_a,$$

назовем его математическим ожиданием квадрата x и обозначим буквою b_1 :

$$\sum x_i^2 p_i = b_1.$$

Так же точно мы обозначим буквою a_2 математическое ожидание y :

$$\sum y_i q_i = a_2,$$

и буквою b_2 математическое ожидание квадрата y :

$$\sum y_i^2 q_i = b_2,$$

и т. д.

Все данные предстоящего рассуждения расположатся в таблицу

Перемен- ное.	Величина пере- менного.	Величина вероятности.		Математич. ожид.	
				Перемен- ного.	Квadrата перемен.
x	$x_1, x_2, x_3, \dots, x_a$	$p_1, p_2, p_3, \dots, p_a$	$\sum p_i = 1$	a_1	b_1
y	$y_1, y_2, y_3, \dots, y_b$	$q_1, q_2, q_3, \dots, q_b$	$\sum q_i = 1$	a_2	b_2
z	$z_1, z_2, z_3, \dots, z_c$	$r_1, r_2, r_3, \dots, r_c$	$\sum r_i = 1$	a_3	b_3
.....					
u	$u_1, u_2, u_3, \dots, u_\mu$	$s_1, s_2, s_3, \dots, s_\mu$	$\sum s_i = 1$	$a_{\mu-1}$	$b_{\mu-1}$
v	$v_1, v_2, v_3, \dots, v_l$	$t_1, t_2, t_3, \dots, t_l$	$\sum t_i = 1$	a_μ	b_μ

Пусть

$$x_\alpha, y_\beta, z_\gamma, \dots, u_\mu, v_l \dots \dots \dots (72)$$

есть одна из возможных систем значений переменных (70), следовательно,

α равно одному из чисел: 1, 2, 3, ..., a ,

β " " " " 1, 2, 3, ..., b ,

γ " " " " 1, 2, 3, ..., c ,

и т. д.

Мы назовем систему значений (72) возможной системой значений переменных (70).

Произведение вероятностей, соответствующих величинам (72):

$$p_\alpha q_\beta z_\gamma \dots s_\mu t_l \dots \dots \dots (73)$$

мы назовем вероятностью случайной системы возможных значений (72) величин (70). Величина (73) есть вероятность того, что переменные (70) окажутся равными числовым значениям (72).

Составим выражение:

$$(x_\alpha + y_\beta + z_\gamma + \dots + u_\lambda + v_\lambda - \Sigma a_i)^2 p_\alpha q_\beta r_\gamma \dots s_\lambda t_\lambda, \dots \quad (74)$$

где

$$\Sigma a_i = a_1 + a_2 + a_3 + \dots + a_\mu$$

есть сумма математических ожиданий всех переменных.

Ясно, что выражение (74) существенно положительно.

Дадим в нем указателям $\alpha, \beta, \gamma, \dots, \lambda$ всевозможные для них системы значений и результаты сложим. Сумму выразим формулой:

$$\Sigma (x_\alpha + y_\beta + z_\gamma + \dots + u_\lambda + v_\lambda - \Sigma a_i)^2 p_\alpha q_\beta r_\gamma \dots s_\lambda t_\lambda \quad (75)$$

Так как все слагаемые этой суммы положительны, то вся сумма положительна.

Раскроем в выражении (75) скобку и произведем суммирование почленно:

$$\left. \begin{aligned} & \Sigma x_\alpha^2 p_\alpha q_\beta \dots t_\lambda + \Sigma y_\beta^2 p_\alpha q_\beta \dots t_\lambda + \dots + \Sigma v_\lambda^2 p_\alpha q_\beta \dots t_\lambda + \\ & + 2 \{ \Sigma x_\alpha y_\beta p_\alpha q_\beta \dots t_\lambda + \Sigma x_\alpha z_\gamma p_\alpha q_\beta \dots t_\lambda + \dots \\ & + \Sigma u_\lambda v_\lambda p_\alpha q_\beta \dots t_\lambda \} - \\ & - 2 \Sigma a_i \{ \Sigma x_\alpha p_\alpha q_\beta \dots t_\lambda + \Sigma y_\beta p_\alpha q_\beta \dots t_\lambda + \dots + \Sigma v_\lambda p_\alpha q_\beta \dots t_\lambda \} + \\ & + (\Sigma a_i)^2 \Sigma p_\alpha q_\beta \dots t_\lambda. \end{aligned} \right\} \quad (76)$$

Вычисление членов этой формулы мы начнем с первой строки и возьмем первый член ее:

$$\Sigma x_\alpha^2 p_\alpha q_\beta \dots t_\lambda.$$

Отберем сначала те члены этой суммы, где $\beta, \gamma, \dots, \lambda$ имеют какие-нибудь определенные значения, а указатель α получает поочередно всевозможные свои значения от 1 до a . Множитель:

$$q_\beta r_\gamma \dots s_\lambda t_\lambda$$

будет общий во всех этих членах и выйдет при составлении суммы за скобку; сумма отобранных членов примет вид:

$$q_\beta r_\gamma \dots s_\lambda t_\lambda (x_1^2 p_1 + x_2^2 p_2 + x_3^2 p_3 + \dots + x_a^2 p_a).$$

Выражение в скобках есть математическое ожидание квадрата величины x , которое мы обозначили буквою b_1 . Получившееся выражение примет вид:

$$b_1 q_\beta r_\gamma \dots s_\lambda t_\lambda.$$

Для вычисления первого члена формулы (76) мы должны дать указателем $\beta, \gamma, \dots, \lambda$ в этом последнем выражении всевозможные значения и результаты сложить.

Отберем сперва те слагаемые этой суммы, где $\gamma, \dots, \lambda$ имеют какие-либо определенные значения, а β получает значения всевозможные: 1, 2, 3 ... b . Множитель

$$b_1 r_\gamma \dots s_x t_\lambda$$

будет общий во всех слагаемых этой суммы и выйдет при составлении суммы отобранных членов за скобку: сумма выразится так:

$$b_1 r_\gamma \dots s_x t_\lambda (q_1 + q_2 + \dots + q_b).$$

Выражение в скобках равно 1, как мы видели выше и как показывает второе равенство 4-го столбца таблицы (71).

Предшествующая формула имеет вид:

$$b_1 r_\gamma \dots s_x t_\lambda.$$

В этом выражении мы должны дать указателям $\gamma, \dots, \lambda$ всевозможные для них значения и результаты сложить. Вычисления мы будем дальше вести тем же способом, какой мы применяли выше, пользуясь равенствами 4-го столбца таблицы (71). Так мы придем, наконец, к выражению:

$$b_1 (t_1 + t_2 + t_3 + \dots + t_l),$$

которое на основании последнего равенства того же столбца 4-го таблицы (71) будет равно b_1 .

Итак мы нашли, что первый член 1-ой строки формулы (76) равен b_1 .

По такой же причине следующие члены 1-ой строки формулы (76) будут равны соответственно $b_2, b_3, \dots, b_\mu$. Все выражение, стоящее в первой строке этой формулы выразится так:

$$b_1 + b_2 + b_3 + \dots + b_\mu$$

или:

$$\Sigma b_i \dots \dots \dots (77)$$

Вычисление членов, стоящих в скобках во второй строке формулы (76), производится тем же способом. Берем первый член:

$$\Sigma x_1 y_\beta p_\alpha q_\beta \dots t_\lambda.$$

Произведя первое суммирование подобно, как в рассмотренном выше случае, найдем:

$$y_\beta q_\beta \dots t_\lambda (x_1 p_1 + x_2 p_2 + x_3 p_3 + \dots + x_\alpha p_\alpha),$$

или:

$$a_1 y_\beta q_\beta \dots t_\lambda.$$

Произведя второе суммирование по β таким же образом найдем:

$$a_1 a_2 r_\gamma \dots t_\lambda.$$

При дальнейших суммированиях по $\gamma, \dots, \lambda$ придется пользоваться равенствами 4-го столбца таблицы (71), и в окончательном результате получится величина:

$$a_1 a_2.$$

Итак, 1-ый член в скобках во второй строке формулы (76) равен $a_1 a_2$.

На таком же основании последующие члены равны:

$$a_1 a_3, a_1 a_4, \dots, a_{\mu-1} a_{\mu}.$$

Следовательно, члены в скобках во второй строке формулы (76) суть попарные произведения математических ожиданий величин переменных.

Всю вторую строку можем выразить формулой:

$$2 \sum a_i a_j, \dots \quad (78)$$

где i и j суть всевозможные неравные между собою числа ряда:

$$1, 2, 3, \dots, \mu.$$

Члены в скобках 3-й строки формулы (76) вычисляются очень сходно с членами 1-ой строки, с той разницею, что войдут математические ожидания самих переменных, а не квадратов их. Получится сумма $\sum a_i$.

Все выражение, стоящее в 3-й строке формулы (76), выразится так:

$$- 2 (\sum a_i)^2. \dots \quad (79)$$

Наконец, в последнем члене формулы (76) сумма

$$\sum p_{\alpha} q_{\beta} \dots t_{\lambda}$$

вычисляется подобным же приемом, который мы применяли выше и пользуясь равенствами 4-го столбца таблицы (71). Получится равенство:

$$\sum p_{\alpha} q_{\beta} \dots t_{\lambda} = 1.$$

Последний член формулы (76) равен:

$$(\sum a_i)^2 \dots \quad (80)$$

Вся величина (76) представится, как сумма выражений (77), (78), (79), (81), т.е. она равна:

$$\sum b_i + 2 \sum a_i a_j - (\sum a_i)^2 \dots \quad (81)$$

Совершая в последнем члене раскрытие скобки, имеем:

$$(\sum a_i)^2 = \sum a_i^2 + 2 \sum a_i a_j.$$

Выражение (81) принимает вид:

$$\sum b_i - \sum a_i^2. \dots \quad (82)$$

Таково окончательное выражение величины (76).

Так как (76) есть та же величина, как и (75), то мы в праве написать:

$$\Sigma(x_\alpha + y_\beta + z_\gamma + \dots + u_\chi + v_\lambda - \Sigma a_i)^2 p_\alpha q_\beta z_\gamma \dots s_\chi t_\lambda = \Sigma b_i - \Sigma a_i^2 \dots (83)$$

Так как левая часть этого равенства существенно положительна, то и правая часть ее положительна:

$$\Sigma b_i - \Sigma a_i^2 > 0.$$

Разделим обе части равенства (83) на правую часть его и, кроме того, на v^2 , где v число действительное и положительное, которое пока мы оставляем неопределенным. Получим:

$$\sum \left\{ \frac{x_\alpha + y_\beta + z_\gamma + \dots + u_\chi + v_\lambda - \Sigma a_i}{v \sqrt{\Sigma b_i - \Sigma a_i^2}} \right\}^2 p_\alpha q_\beta z_\gamma \dots s_\chi t_\lambda = \frac{1}{v^2} \dots (84)$$

Разбиваем слагаемые суммы, стоящей в левой части этого равенства, на две группы: к первой группе отнесем слагаемые, удовлетворяющие условию:

$$\left| \frac{x_\alpha + y_\beta + z_\gamma + \dots + u_\chi + v_\lambda - \Sigma a_i}{v \sqrt{\Sigma b_i - \Sigma a_i^2}} \right| \geq 1, \dots (85)$$

где значком $| |$ обозначена абсолютная величина.

Ко второй группе отнесем все остальные слагаемые, т. е. слагаемые, удовлетворяющие условию:

$$\left| \frac{x_\alpha + y_\beta + z_\gamma + \dots + u_\chi + v_\lambda - \Sigma a_i}{v \sqrt{\Sigma b_i - \Sigma a_i^2}} \right| < 1 \dots (86)$$

Опустим в левой части равенства (84) все слагаемые второй категории, оставив только слагаемые первой категории, т. е. удовлетворяющие условию (85).

Левая часть уменьшится и делается меньше правой. Получится следующее:

$$\sum \left\{ \frac{x_\alpha + y_\beta + z_\gamma + \dots + u_\chi + v_\lambda - \Sigma a_i}{v \sqrt{\Sigma b_i - \Sigma a_i^2}} \right\}^2 p_\alpha q_\beta z_\gamma \dots s_\chi t_\lambda < \frac{1}{v^2},$$

где черта — над знаком суммы Σ означает, что суммирование распространяется только на такие значения указателей, при которых выполняется условие (85).

Это неравенство еще усилится, если мы в каждом члене его первый множитель заменим числом 1. Неравенство примет вид:

$$\sum p_\alpha q_\beta z_\gamma \dots s_\chi t_\lambda < \frac{1}{v^2} \dots (87)$$

Выражение $p_\alpha q_\beta z_\gamma \dots s_\chi t_\lambda$ есть вероятность того, что переменные $x, y, z, \dots, u, v \dots (70)$

получают случайную систему значений:

$$x_\alpha, y_\beta, z_\gamma, \dots, u_\kappa, v_\lambda, \dots \quad (72)$$

или, как мы условились выше говорить, это величина вероятности системы случайных значений (72) переменных (70).

По теореме сложения вероятностей можем сказать:

$$\sum p_\alpha q_\beta r_\gamma \dots s_\kappa t_\lambda$$

есть вероятность того, что переменные (70) получают одну какую-нибудь из систем значений, удовлетворяющих условию (85).

Для краткости назовем ее вероятностью условия (85) и обозначим буквою Q .

Неравенство (87) примет вид:

$$Q < \frac{1}{v^2}.$$

Событие противоположное только-что рассмотренному будет состоять в том, что переменные (70) получают систему значений, удовлетворяющих условию (86). Мы обозначим эту вероятность буквою P и назовем вероятностью условия (86).

Как вероятности противоположных событий, величины P и Q в сумме дают единицу:

$$P + Q = 1,$$

откуда

$$Q = 1 - P,$$

и последнее неравенство примет вид

$$1 - P < \frac{1}{v^2}$$

откуда:

$$P > 1 - \frac{1}{v^2} \quad (I)$$

Такому неравенству удовлетворяет вероятность P условия (86).

Преобразуем несколько вид этого неравенства (86): умножим обе части его на знаменателя левой части и разделим на μ ;

$$\left| \frac{x_\alpha + y_\beta + z_\gamma + \dots + u_\kappa + v_\lambda}{\mu} - \frac{\sum a_i}{\mu} \right| < \frac{v}{\sqrt{\mu}} \sqrt{\frac{\sum b_i}{\mu} - \frac{\sum a_i^2}{\mu}} \quad (II)$$

Полученный результат коротко можно выразить такими словами: вероятность P неравенства (II) к единице ближе, чем $1 - \frac{1}{v^2}$.

Рассмотрим смысл неравенства (II) и величин отдельных его членов.

Величина:

$$\frac{x_\alpha + y_\beta + z_\gamma + \dots + u_\kappa + v_\lambda}{\mu}$$

есть средне-арифметическое системы случайных значений переменных.

Величина:

$$\frac{\sum a_i}{\mu}$$

есть средне-арифметическое математических ожиданий этих переменных; разность, стоящая в левой части неравенства (II) есть уклонение первого из этих средне-арифметических от второго; мы обозначим его буквою X и коротко будем называть его *уклонением*.

Неравенство (II) говорит, что уклонение X по абсолютной величине меньше величины, стоящей в правой части этого неравенства.

Переходим к рассмотрению правой части.

Буквою μ обозначено число всех переменных, входящих в рассмотрение.

В тех вопросах, к которым приходится применять найденные формулы, число переменных зависит от нашего выбора: мы можем взять переменных столько, чтобы число их μ было больше какого угодно данного положительного постоянного числа. Иными словами, мы в праве назвать μ бесконечно-большим числом.

Число ν осталось пока неопределенным. Мы можем его взять тоже бесконечно-большим, но порядка ниже, чем μ , напр., положивши:

$$\nu = \sqrt[3]{\mu}.$$

При таком выборе величины ν , будем иметь:

$$\frac{\nu}{\sqrt{\mu}} = \frac{1}{\sqrt[3]{\mu}};$$

поэтому $\frac{\nu}{\sqrt{\mu}}$ будет величиною бесконечно-малою.

Итак, мы можем выбрать ν таким образом, чтобы оно было бесконечно-велико, а $\frac{\nu}{\sqrt{\mu}}$ — бесконечно-мало.

Далее величины $\frac{\sum b_i}{\mu}$ и $\frac{\sum a_i^2}{\mu}$ суть средне-арифметические: 1) математических ожиданий квадратов переменных, 2) квадратов математических ожиданий самих переменных.

Если значения переменных конечны, то и математические ожидания их самих и квадратов их конечны, а потому конечны и средне-арифметические.

В таком случае выражение:

$$\sqrt{\frac{\sum b_i}{\mu} - \frac{\sum a_i^2}{\mu}}$$

конечно.

Вся правая часть неравенства (II) есть, следовательно, произведение бесконечно-малой на конечную. Оно бесконечно-мало.

Итак, неравенство (II) показывает, что при бесконечно-большом числе μ переменных отклонение X бесконечно-мало.

Неравенство (I) характеризует вероятность P того, что удовлетворится условие (II).

Так как ν бесконечно-велико, то $\frac{1}{\sqrt{2}}$ бесконечно-мало. Формула (I) показывает, что P бесконечно-близко к 1.

Итак, в случае бесконечно-большого числа переменных можно утверждать с вероятностью бесконечно-близкой к достоверности, что отклонение X будет бесконечно-мало, при чем отклонением X мы называем разность между средне-арифметическим системы случайных значений переменных и средне-арифметическим математических ожиданий этих переменных.

В приложениях, где μ определенное весьма большое число, мы можем сказать: вероятность того, что отклонение X весьма мало весьма близка к 1: она разнится от 1 меньше, чем на $\frac{1}{\sqrt{2}}$.

Как мы видим, теорема Чебышева того же характера, как те отдельные формулы, которые мы рассматривали выше и которые объединяются под общим названием закона больших чисел. Но теорема Чебышева имеет гораздо большую общность. Теоремы Пуассона и Бернулли получаются из нее, как частные случаи.

§ 16. Вывод теорем Пуассона и Бернулли из теоремы Чебышева.

Пусть произведено μ испытаний над появлением события E или противоположного ему события F . Пусть вероятности событий E и F при различных испытаниях различны, но не зависят от результатов испытаний предшествующих, а именно при 1-м испытании вероятности E и F суть p_1 и q_1 , при 2-м испытании — p_2 и q_2 , при 3-м — p_3 и q_3 и т. д., наконец, при μ -м — p_μ и q_μ . Мы хотим применить теорему Чебышева к нахождению предела возможных величин частоты $\frac{m}{\mu}$ события E .

С этой целью прибегаем к искусственному приему: вообразим μ переменных: $x, y, z, \dots, u, v$, из которых каждое получает значение 1 или 0 в зависимости от результата соответствующего испытания: x от 1-го, y — от 2-го и т. д. v — от μ -го испытания. Если при каком-либо испытании состоялось событие E , то соответствующее этому испытанию переменное получает значение 1; в противоположном случае — 0.

Найдем математические ожидания переменных. Для x имеем:

$$a_1 = 1 \cdot p_1 + 0 \cdot q_1 = p_1.$$

На таком же основании найдем:

$$a_2 = p_2, \quad a_3 = p_3, \dots$$

Найдем математические ожидания квадратов переменных.

Для x имеем:

$$b_1 = 1^2 \cdot p_1 + 0^2 \cdot q_1 = p_1.$$

На таком же основании:

$$b_2 = p_2, \quad b_3 = p_3, \dots$$

Таблица (71) будет такова:

Переменное.	Величина переменного.	Величина вероятности.		Математическое ожи- дание.	
				Перемен- ного.	Квadrата перемен- ного.
x	1, 0	p_1, q_1	$p_1 + q_1 = 1$	p_1	p_1
y	1, 0	p_2, q_2	$p_2 + q_2 = 1$	p_2	p_2
z	1, 0	p_3, q_3	$p_3 + q_3 = 1$	p_3	p_3
...	...	...	...	...	...
v	1, 0	p_μ, q_μ	$p_\mu + q_\mu = 1$	p_μ	p_μ

Сумма случайных значений переменных имеет слагаемым столько раз единицу, сколько раз событие E состоялось, и столько раз слагаемым нуль, сколько раз событие E не состоялось. Иными словами эта сумма равна числу появления события E при μ испытаниях. Средне-арифметическое для данных случайных значений переменных равно:

$$\frac{m}{\mu},$$

т.е. оно равно частоте события E .

Средне-арифметическое математических ожиданий переменных в данном случае таково:

$$\frac{\sum a_i}{\mu} = \frac{p_1 + p_2 + p_3 + \dots + p_\mu}{\mu},$$

т.е. оно равно средней вероятности p события E .

Уклонение X выразится так:

$$X = \frac{m}{\mu} - p,$$

т.е. это то самое уклонение, о котором мы говорили в предшествующих параграфах.

Остается найти выражение правой части неравенства (II) для данного случая.

Выражение стоящее под радикалом в этой правой части для данного случая таково:

$$\begin{aligned} \frac{\sum b_i}{\mu} - \frac{\sum a_i^2}{\mu} &= \frac{p_1 + p_2 + p_3 + \dots + p_\mu}{\mu} - \frac{p_1^2 + p_2^2 + p_3^2 + \dots + p_\mu^2}{\mu} = \\ &= \frac{p_1(1 - p_1) + p_2(1 - p_2) + p_3(1 - p_3) + \dots + p_\mu(1 - p_\mu)}{\mu} = \\ &= \frac{p_1 q_1 + p_2 q_2 + p_3 q_3 + \dots + p_\mu q_\mu}{\mu} = \frac{\sum p_i q_i}{\mu}. \end{aligned}$$

Неравенство (II) напишется так:

$$\left| \frac{m}{\mu} - p \right| < \frac{\nu}{\sqrt{\mu}} \sqrt{\frac{\sum p_i q_i}{\mu}} \dots \dots \dots (II')$$

Неравенство (I) сохранит свой прежний вид:

$$P > 1 - \frac{1}{\nu^2} \dots \dots \dots (I)$$

В формуле (II') величина $\frac{\sum p_i q_i}{\mu}$ есть средне-арифметическое величин:

$$p_1 q_1, p_2 q_2, p_3 q_3, \dots, p_\mu q_\mu;$$

оно конечно. Величина $\frac{\nu}{\sqrt{\mu}}$ по прежнему бесконечно-мала. Величина $\frac{1}{\nu^2}$ бесконечно-мала.

Поэтому формулы (I) и (II') прочтутся так: с вероятностью P , бесконечно-близкой к 1, можно утверждать, что отклонение $\frac{m}{\mu} - p$ будет бесконечно-мало, если число испытаний бесконечно-велико.

Это — та самая теорема, о которой мы говорили в конце предшествующего параграфа и которая называется теорема Пуассона.

В самом деле, сравним формулы (II') и (I) настоящего параграфа с формулами предшествующего параграфа (D), (63*), (64*), (C). Те и другие формулы определяют вероятность предела отклонения $\frac{m}{\mu} - p$.

В предшествующем параграфе этот предел обозначен буквою α .

Вероятность его приближенно дается формулою (C), где h выражается равенством (64*).

В настоящем параграфе правая часть формулы (II') определяет тот же самый предел α , а формула (I) говорит, что вероятность этого предела ближе к 1, чем $1 - \frac{1}{v^2}$.

Совершенно так же, как в конце предшествующего параграфа, мы можем рассмотреть частный случай теоремы Пуассона, полагая:

$$p_1 + p_2 + p_3 + \dots + p_\mu = p.$$

Полученный результат будет выражать теорему Бернулли, но в новых обозначениях: равенства (I) и (II') примут вид:

$$P > 1 - \frac{1}{v^2}, \dots \dots \dots (I)$$

$$\left| \frac{m}{\mu} - p \right| < \frac{v}{\sqrt{\mu}} \sqrt{pq} \dots \dots \dots (II')$$

Чтобы сравнить результаты, даваемые применением теоремы Бернулли к числовому примеру при обеих формулировках этой теоремы, обратимся к рассмотренному в § 10 примеру 1.

В этом примере мы имели:

$$\mu = 10.000, \quad p = q = \frac{1}{2},$$

и нашли, что с вероятностью $P = 0,9999$ отклонение $\frac{m}{\mu} - p$ по абсолютной величине не превзойдет величины:

$$\alpha = 0,0195,$$

или, ограничиваясь двумя десятичными знаками,

$$0,02.$$

Применяя формулы настоящего параграфа, полагаем:

$$\frac{v}{\sqrt{\mu}} \sqrt{pq} = \frac{v}{200} = 0,02,$$

откуда:

$$v = 4.$$

Неравенство (I) будет таково:

$$P > 1 - \frac{1}{16}, \quad P > 0,9375.$$

Конечно, эта последняя оценка величины P дает результат не такой отчетливый, как полученный выше $P = 0,9999$. Это объясняется тем, что при доказательстве теоремы Чебышева приходится значительно усиливать неравенства, вследствие чего и получается результат несколько грубый. Но, достоинство его в том, что утверждая $P > 0,9375$ мы совершенно свободны от ошибки, тогда как говоря $P = 0,9999$ мы имеем результат только приближенный.

ГЛАВА ЧЕТВЕРТАЯ

О вероятностях на основании опыта.

§ 1. Постановка вопроса.

Закон больших чисел позволяет нам найти вероятность события на основании опыта, и при том с произвольной степенью точности. Такую вероятность *a posteriori* мы можем ввести в вычисления наравне с известными нам вероятностями *a priori* других событий.

Но такой способ определения вероятности из опыта, основанный на законе больших чисел, требует производства большого числа испытаний. Между тем иногда уже результаты нескольких испытаний значительно расширяют наши знания о вероятностях событий, подлежащих исследованию.

Чтобы получить понятие о таких случаях, возьмем пример.

Пусть известно, что в урне 3 шара, которые могут быть только белыми или черными; но сколько белых и сколько черных — неизвестно.

Заранее можно сделать 4 гипотезы: в урне находятся:

- 1) 3 белых шара и 0 черных.
- 2) 2 " " " 1 "
- 3) 1 " " " 2 "
- 4) 0 " " " 3 "

Так как о составе урны ничего не известно, то все 4 гипотезы *a priori* равно вероятны, и мы должны оценить вероятность каждой из них в $\frac{1}{4}$.

Но пусть произведен первый опыт, который дал появление белого шара. После этого опыта вероятность 4-й гипотезы делается равной 0, а вероятности трех первых гипотез увеличиваются.

Пусть при втором опыте появился шар черный. После этого вероятность первой гипотезы делается равной нулю, так же как и вероятность 4-й гипотезы; а вероятности второй и третьей гипотез делаются равными между собою и равными $\frac{1}{2}$.

Пусть затем 3 испытания подряд дали шар белый. Тогда 2-я гипотеза делается значительно вероятнее 3-ей, и мы с большим основанием можем утверждать, что в урне 2 белых шара и 1 черный.

В данном случае 5 испытаний оказалось достаточно для получения надежной оценки вероятности.

Решение подобных задач составляет предмет настоящей главы. Оно основывается на теореме Бейеса *).

§ 2. Теорема Бейеса.

Пусть событие E может состояться под действием k причин, взаимно исключаящих друг друга и всевозможных:
 $C_1, C_2, C_3, \dots, C_k$ **).

Вероятности этих причин до производства опытов а priori пусть соответственно равны:

$$P_1, P_2, P_3, \dots, P_k,$$

при чем:

$$P_1 + P_2 + P_3 + \dots + P_k = 1.$$

Вероятности, сообщаемые этими причинами событию E , обозначим так:

$$p_1, p_2, p_3, \dots, p_k.$$

Пусть произведен опыт, при котором событие E состоялось.

Мы можем сделать k гипотез, предполагая, что событие состоялось под действием 1-ой, 2-ой, 3-ей k -ой причины.

Вероятности этих гипотез после произведенного опыта мы обозначим так:

$$\omega_1, \omega_2, \omega_3, \dots, \omega_k.$$

Наша задача — найти эти величины.

Начнем с первой гипотезы, вероятность которой ω_1 .

Для ее нахождения рассмотрим сложное событие, состоящее из двух простых: 1) подействовала причина C_1 ; 2) под действием этой причины событие E состоялось.

По теореме умножения вероятностей при зависимых между собою простых событиях, эта вероятность выразится так:

$$P_1 p_1.$$

То же самое сложное событие можно разложить на два простых иначе:

1) событие E состоялось; 2) причиной, его вызвавшей, оказалась C_1 .

*) Эта теорема найдена Бейесом (J. Bayes), но напечатана после его смерти Price'ом в 53 томе Philosophical Transactions 1764 г. Развитие теории, основанной на теореме Бейеса, выполнено Лапласом в Théorie analytique des probabilités.

**) См. обозначение параграфа 5, главы I о полной вероятности.

По той же теореме умножения вероятностей эта вероятность выразится произведением:

$$P\omega_1,$$

где P есть полная вероятность события E , выражаемая формулой:

$$P = P_1 p_1 + P_2 p_2 + P_3 p_3 + \dots + P_k p_k = \sum P_i p_i^*), \dots (1)$$

а ω_1 искомая величина вероятности 1-ой гипотезы после произведенного опыта.

Итак:

$$P_1 p_1 = P\omega_1,$$

откуда:

$$\omega_1 = \frac{P_1 p_1}{P}.$$

Таким же образом составим выражения для $\omega_2, \omega_3, \dots, \omega_k$.

Получатся формулы:

$$\left. \begin{aligned} \omega_1 &= \frac{P_1 p_1}{\sum P_i p_i}, \\ \omega_2 &= \frac{P_2 p_2}{\sum P_i p_i}, \\ \omega_3 &= \frac{P_3 p_3}{\sum P_i p_i}, \\ &\dots \\ \omega_k &= \frac{P_k p_k}{\sum P_i p_i}. \end{aligned} \right\} \dots (2)$$

Эта формула составляет теорему Бейеса.

Сложив равенства (2) между собою, получим:

$$\omega_1 + \omega_2 + \omega_3 + \dots + \omega_k = 1 \dots (3)$$

Это равенство при вычислениях может служить проверкой.

Если бы при произведенном опыте состоялось событие F , противоположное E , то вероятности гипотез выразились бы следующими формулами:

$$\left. \begin{aligned} \omega_1 &= \frac{P_1 q_1}{\sum P_i q_i}, \\ \omega_2 &= \frac{P_2 q_2}{\sum P_i q_i}, \\ \omega_3 &= \frac{P_3 q_3}{\sum P_i q_i}, \\ &\dots \\ \omega_k &= \frac{P_k q_k}{\sum P_i q_i}. \end{aligned} \right\} \dots (2)$$

*) См. главу I, § 5, форм. (10).

Будем предполагать по прежнему, что при первом испытании произошло событие E . Пусть после этого хотят произвести второе испытание, при чем известно, что при этом испытании будет действовать непременно та же причина, которая действовала в 1-м.

Вероятность события E при втором испытании выразится по формуле полной вероятности, подобной (1), но с заменой в ней величин

$$P_1, P_2, P_3, \dots, P_k,$$

величинами:

$$\omega_1, \omega_2, \omega_3, \dots, \omega_k,$$

т. е. вероятность события E при втором испытании будет такова:

$$\Omega = \omega_1 p_1 + \omega_2 p_2 + \omega_3 p_3 + \dots + \omega_k p_k \quad (4)$$

Если при этом 2-м испытании опять осуществилось событие E , то вероятности гипотез получают новую оценку по формулам, подобным (2), но с такой же заменой, которую мы только-что делали; они выразятся так:

$$\left. \begin{aligned} \omega'_1 &= \frac{\omega_1 p_1}{\Omega}, \\ \omega'_2 &= \frac{\omega_2 p_2}{\Omega}, \\ \omega'_3 &= \frac{\omega_3 p_3}{\Omega}, \\ &\dots \\ \omega'_k &= \frac{\omega_k p_k}{\Omega}. \end{aligned} \right\} \dots \dots \dots (5)$$

Сложив равенства (5) между собой, получим:

$$\omega'_1 + \omega'_2 + \omega'_3 + \dots + \omega'_k = 1. \quad (6)$$

Это равенство при вычислениях может служить проверкой.

Если бы при втором испытании состоялось событие F , то вероятности гипотез выразились бы формулами, подобными (5), но с заменой величин $p_1, p_2, p_3, \dots, p_k$ величинами $q_1, q_2, q_3, \dots, q_k$:

$$\left. \begin{aligned} \bar{\omega}'_1 &= \frac{\omega_1 q_1}{\Sigma \omega_i q_i}, \\ \bar{\omega}'_2 &= \frac{\omega_2 q_2}{\Sigma \omega_i q_i}, \\ \bar{\omega}'_3 &= \frac{\omega_3 q_3}{\Sigma \omega_i q_i}, \\ &\dots \\ \bar{\omega}'_k &= \frac{\omega_k q_k}{\Sigma \omega_i q_i}. \end{aligned} \right\} \dots \dots \dots (5)$$

и т. д.

Каждое новое испытание приводит к переоценке вероятностей гипотез. Само собою понятно, что все эти оценки субъективны; тем не менее с каждым новым опытом мы приближаемся к истинным величинам вероятностей.

Чтобы показать приложение только-что сказанного, рассмотрим подробно пример, приведенный в предшествующем параграфе.

В урну положено 3 шара, которые могут быть белыми или черными. При первом испытании появился белый шар, при втором черный, при дальнейших трех испытаниях появился шар белый.

Речь идет о вероятностях четырех гипотез, что в урне было: 1) 3 белых шара и 0 черных, 2) 2 белых и 1 черный, 3) 1 белый и 2 черных, 4) 0 белых и 3 черных.

Применяя приведенные выше формулы, мы найдем вероятности сказанных гипотез после каждого из произведенных испытаний.

До опытов все гипотезы равновероятны: поэтому

$$p_1 = p_2 = p_3 = p_4 = \frac{1}{4}.$$

Вероятности, сообщаемые соответствующими причинами событиям: а) появлению белого шара, б) появлению черного шара таковы:

$$1) p_1 = 1, q_1 = 0; 2) p_2 = \frac{2}{3}, q_2 = \frac{1}{3}; 3) p_3 = \frac{1}{3}, q_3 = \frac{2}{3};$$

$$4) p_4 = 0, q_4 = 1.$$

I испытание:

полная вероятность появления белого шара:

$$\frac{1}{4} \cdot 1 + \frac{1}{4} \cdot \frac{2}{3} + \frac{1}{4} \cdot \frac{1}{3} + \frac{1}{4} \cdot 0 = \frac{1}{2};$$

вероятности гипотез после опыта:

$$1) \frac{\frac{1}{4} \cdot 1}{\frac{1}{2}} = \frac{1}{2}; 2) \frac{\frac{1}{4} \cdot \frac{2}{3}}{\frac{1}{2}} = \frac{1}{3}; 3) \frac{\frac{1}{4} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{6}; 4) \frac{\frac{1}{4} \cdot 0}{\frac{1}{2}} = 0;$$

проверка:

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{6} + 0 = 1.$$

II испытание:

полная вероятность появления черного шара:

$$\frac{1}{2} \cdot 0 + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{6} \cdot \frac{2}{3} + 0 \cdot 1 = \frac{2}{9};$$

вероятности гипотез после опыта:

$$1) \frac{\frac{1}{2} \cdot 0}{\frac{2}{9}} = 0; 2) \frac{\frac{1}{3} \cdot \frac{1}{3}}{\frac{2}{9}} = \frac{1}{2}; 3) \frac{\frac{1}{6} \cdot \frac{2}{3}}{\frac{2}{9}} = \frac{1}{2}; 4) \frac{0 \cdot 1}{\frac{2}{9}} = 0;$$

проверка:

$$0 + \frac{1}{2} + \frac{1}{2} + 0 = 1.$$

III испытание:

полная вероятность появления белого шара:

$$0 \cdot 1 + \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} + 0 \cdot 0 = \frac{1}{2};$$

вероятности гипотез после испытания:

$$1) \frac{0 \cdot 1}{\frac{1}{2}} = 0; 2) \frac{\frac{1}{2} \cdot \frac{2}{3}}{\frac{1}{2}} = \frac{2}{3}; 3) \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}; 4) \frac{0 \cdot 0}{\frac{1}{2}} = 0;$$

проверка:

$$0 + \frac{2}{3} + \frac{1}{3} + 0 = 1.$$

IV испытание:

полная вероятность появления белого шара:

$$0 \cdot 1 + \frac{2}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{1}{3} + 0 \cdot 0 = \frac{5}{9};$$

вероятности гипотез после опыта:

$$1) \frac{0 \cdot 1}{\frac{5}{9}} = 0; 2) \frac{\frac{2}{3} \cdot \frac{2}{3}}{\frac{5}{9}} = \frac{4}{5}; 3) \frac{\frac{1}{3} \cdot \frac{1}{3}}{\frac{5}{9}} = \frac{1}{5}; 4) \frac{0 \cdot 0}{\frac{5}{9}} = 0;$$

проверка:

$$0 + \frac{4}{5} + \frac{1}{5} + 0 = 1.$$

V испытание:

полная вероятность появления белого шара:

$$0 \cdot 1 + \frac{4}{5} \cdot \frac{2}{3} + \frac{1}{5} \cdot \frac{1}{3} + 0 \cdot 0 = \frac{9}{15} = \frac{3}{5};$$

вероятности гипотез после опыта:

$$1) \frac{0 \cdot 1}{\frac{3}{5}} = 0; 2) \frac{\frac{4}{5} \cdot \frac{2}{3}}{\frac{9}{15}} = \frac{8}{9}; 3) \frac{\frac{1}{5} \cdot \frac{1}{3}}{\frac{9}{15}} = \frac{1}{9}; 4) \frac{0 \cdot 0}{\frac{9}{15}} = 0;$$

проверка:

$$0 + \frac{8}{9} + \frac{1}{9} + 0 = 1.$$

Мы видим, что после 5 опытов 2-я гипотеза оказывается в 8 раз вероятнее третьей, а 1-я и 4-я совсем невероятны.

В рассмотренном примере до начала опытов все гипотезы были равновероятны. Такой случай встречается очень часто; рассмотрим его особенности.

Если в формулах (2) положить:

$$P_1 = P_2 = P_3 = \dots = P_k,$$

то эта величина выйдет в знаменателе за скобки, как общий множитель, и сократится с первым множителем числителя: формулы (2) примут вид:

$$\omega_1 = C \cdot p_1, \quad \omega_2 = C \cdot p_2, \quad \omega_3 = C \cdot p_3, \dots, \quad \omega_k = C \cdot p_k, \dots \quad (7)$$

где

$$C = \frac{1}{\sum p_i} = \frac{1}{p_1 + p_2 + p_3 + \dots + p_k}.$$

Эти формулы показывают, что когда до опыта все гипотезы равновероятны, то после опыта вероятности гипотез пропорциональны вероятностям, сообщаемым наблюдаемому событию соответствующими причинами.

§ 3. Приложение теоремы Байеса к случаю бесконечно-большого числа гипотез.

Пусть речь идет об испытаниях над событием E , о вероятности которого а priori нам ничего не известно.

Обозначим вероятность события E буквою x . О ней можем сказать только одно: что она заключена между 0 и 1.

Разобьем весь этот промежуток от 0 до 1 на k равных частей; длина каждой части равна $\frac{1}{k}$; мы обозначим ее через dx :

$$dx = \frac{1}{k}.$$

Относительно неизвестной вероятности события E мы можем сделать k гипотез, что эта величина лежит в том или ином из этих k промежутков.

При этом мы будем употреблять несколько сокращенное выражение; вместо того чтобы сказать: „вероятность события E лежит в промежутке между x и $x + dx$ “, будем говорить коротко: „вероятность события E равна x “.

Такое сокращенное выражение соответствует обычным выражениям в приложениях анализа к опытным наукам (механике, физике и т. д.).

Если число k бесконечно-велико, то dx бесконечно-мало; все промежутки бесконечно-малы, а следовательно, все величины x , лежащие в одном и том же промежутке, разнятся между собою бесконечно-мало.

Пусть над появлением события E произведено μ испытаний, и пусть при этих μ испытаниях событие E состоялось m раз и не состоялось n раз, при чем

$$m + n = \mu.$$

Спрашивается, каковы будут на основании этих опытов вероятности тех k гипотез, которые мы можем сделать относительно величины вероятности события E .

Рассмотрим гипотезу, по которой вероятность события E лежит между x и $x + dx$. Пользуясь тем сокращенным выражением, о котором мы говорили выше, назовем ее „гипотеза, что вероятность события E равна x “.

Если вероятность события E равна x , то вероятность противоположного ему события F равна $1 - x$. В таком случае по формуле (5) § 1 главы III мы выразим следующей формулой вероятность наблюдаемого сложного события, что при μ испытаниях состоялось m раз событие E и n раз событие F :

$$\frac{\mu!}{m! n!} x^m (1 - x)^n. \quad \dots \dots \dots (8)$$

Так как при бесконечно-малом dx все значения вероятностей, лежащие между x и $x + dx$, бесконечно близки между собой, то можно сказать, что если вероятность события E лежит в рассматриваемом промежутке, то вероятность наблюдаемого сложного события может разниться от выражения (8) только на бесконечно-малую величину. Опуская эту бесконечно-малую, можем сказать, что при всех значениях вероятности события E , лежащих в рассматриваемом промежутке, вероятность наблюдаемого сложного события равна величине (8).

Величина (8) есть вероятность, сообщаемая наблюдаемому событию рассматриваемой причиной.

Так как до опытов все гипотезы были равно вероятны, то на основании последнего замечания в конце предшествующего параграфа можем сказать, что вероятность рассматриваемой гипотезы пропорциональна величине (8).

Так как dx величина постоянная относительно x , то мы можем представить множителя пропорциональности в виде:

$$C dx.$$

В таком случае, обозначая буквою ω , величину вероятности гипотезы, что величина вероятности события E равна x^*), будем иметь:

$$\omega_x = C \cdot \frac{\mu!}{m! n!} x^m (1-x)^n dx \quad (9)$$

Для нахождения постоянного C припомним, что сумма вероятностей всех гипотез должна равняться 1 **).

При вычислении суммы величин (9) при всевозможных значениях x :

$$0, dx, 2dx, 3dx, \dots, (k-1) dx$$

постоянный множитель

$$C \frac{\mu!}{m! n!}$$

выйдет за скобку, а сумма членов в скобках в пределе выразится определенным интегралом, и мы получим:

$$1 = C \frac{\mu!}{m! n!} \int_0^1 x^m (1-x)^n dx \quad (10)$$

Разделив равенство (9) на (10) мы получим:

$$\omega_x = \frac{x^m (1-x)^n dx}{\int_0^1 x^m (1-x)^n dx} \quad (11)$$

Помощью этой формулы можем решить такой вопрос: при μ испытаниях событие E состоялось m раз и не состоялось n раз; какова вероятность того, что вероятность x события E больше некоторого числа a и меньше b , т.е. что x содержится между a и b .

Мы обозначим эту вероятность символом: ω_a^b .

Пользуясь теоремой сложения вероятностей, для нахождения этой величины мы должны дать в формуле (11) аргументу x всевозможные значения от a до b и результаты сложить. Так как знаменатель формулы (11) от x не зависит, то он будет общий во всех слагаемых;

*) Точнее: вероятность события E лежит в промежутке между x и $x+dx$.

**) См. формулу (3) предшествующего параграфа. Этой формулой мы уже пользовались выше для проверки, рассматривая примеры в конце предшествующего параграфа.

***) Так как m и n суть числа целые, то знаменатель вычислить легко, раскрывая скобки и интегрируя почленно; но, кроме того, из интегрального исчисления известно, что:

$$\int_0^1 x^m (1-x)^n dx = B(m+1, n+1) = \frac{\Gamma(m+1) \Gamma(n+1)}{\Gamma(m+n+2)} = \frac{m! n!}{\mu+1!}.$$

сумма числителей в пределе, когда dx обратится в нуль, выразится интегралом, и мы получим:

$$\omega_a^b = \frac{\int_a^b x^m (1-x)^n dx}{\int_0^1 x^m (1-x)^n dx} \quad (12)$$

В частности, когда $a = \frac{1}{2}$, $b = 1$, формула (12) примет вид:

$$\omega_{\frac{1}{2}}^1 = \frac{\int_{\frac{1}{2}}^1 x^m (1-x)^n dx}{\int_0^1 x^m (1-x)^n dx}$$

Она выражает вероятность того, что x больше половины, т.-е. что событие E может быть названо вероятным.

Возьмем конкретный пример. Пусть в урне шары белые и черные; но сколько тех и других — неизвестно. Пусть произведено 10 испытаний, при чем каждый раз вынутый шар возвращался в урну, и пусть при всех 10 испытаниях шар оказался белым. Какова вероятность того, что в урне белых шаров больше, чем черных?

Если в урне белые шары преобладают, то вероятность вынута белого шара больше половины. Поэтому ответ на поставленный вопрос дается формулой:

$$\frac{\int_{\frac{1}{2}}^1 x^{10} dx}{\int_0^1 x^{10} dx} = \frac{1 - \left(\frac{1}{2}\right)^{11}}{1} = 1 - \frac{1}{2048}$$

Следовательно, до единицы недостает только $\frac{1}{2048}$.

Та же формула (11) позволяет делать заключения о вероятностях для будущих испытаний. Пусть после произведенных μ испытаний, в которых событие E состоялось m раз и не состоялось n раз, имеется в виду произвести μ' новых испытаний. Спрашивается, какова вероятность, что при этих μ' испытаниях событие E состоится m' раз и не состоится n' раз, где $m' + n' = \mu'$.

Сделаем гипотезу, что вероятность события E равна x ; после произведенных μ испытаний эта вероятность выразится формулой (11).

При этой гипотезе вероятность осуществления указанного выше результата испытаний такова:

$$\frac{\mu'}{m'! n'!} x^{m'} (1-x)^{n'} \quad (13)$$

Чтобы найти полную вероятность, мы умножаем вероятность каждой причины (11) на вероятность (13), ею сообщаемую ожидаемому событию, и берем сумму произведений полученного вида:

$$\frac{\mu'!}{m'! n'!} x^{m+m'} (1-x)^{n+n'} dx$$

$$\int_0^1 x^m (1-x)^n dx$$

при всевозможных значениях x от 0 до 1.

В этой сумме постоянные множители выйдут за скобку, а сама сумма в пределе выразится определенным интегралом, и мы получим ответ на поставленный выше вопрос в таком виде:

$$\frac{\mu'!}{m'! n'!} \int_0^1 x^{m+m'} (1-x)^{n+n'} dx$$

$$\int_0^1 x^m (1-x)^n dx \quad *) \dots \dots \dots (14)$$

Возьмем конкретный случай. Пусть при 10 испытаниях подряд появлялся белый шар; какова вероятность, что и при следующем 11-м испытании опять появится белый шар?

Ответ дается формулой:

$$\frac{\int_0^1 x^{11} dx}{\int_0^1 x^{10} dx} = \frac{\frac{1}{12}}{\frac{1}{11}} = \frac{11}{12}.$$

§ 4. Теорема, обратная теореме Якова Бернулли.

Формула (11) предшествующего параграфа дает величину вероятности гипотезы, что вероятность события E равна x . То значение x , при котором величина (11) достигает своего наибольшего значения, можно назвать наимвероятнейшим значением x . Для его нахождения приравняем нулю производную функции (11) по переменному x . Знаменатель выражения (11) постоянный; поэтому приравняем нулю производную числителя, или, еще проще — производную логарифма числителя:

$$\frac{m}{x} - \frac{n}{1-x} = 0,$$

откуда

$$x = \frac{m}{n} \dots \dots \dots (15)$$

*) Так же, как и выше мы можем написать:

$$\int_0^1 x^m (1-x)^n dx = \frac{m! n!}{n+1!}; \quad \int_0^1 x^{m+m'} (1-x)^{n+n'} dx = \frac{m+m'! n+n'!}{n+m'+1!}.$$

Положим в формуле (12)

$$a = \frac{m}{\mu} - \alpha, \quad b = \frac{m}{\mu} + \alpha.$$

Формула (12) примет следующий вид:

$$\omega_{\frac{m}{\mu} - \alpha}^{\frac{m}{\mu} + \alpha} = \frac{\int_{\frac{m}{\mu} - \alpha}^{\frac{m}{\mu} + \alpha} x^m (1-x)^n dx}{\int_0^1 x^m (1-x)^n dx} \dots \dots \dots (16)$$

Она показывает на основании произведенных испытаний вероятность того, что отклонение величины x от наивероятнейшего ее значения $\frac{m}{\mu}$ по абсолютной величине не превышает α .

Займемся преобразованием числителя этого выражения. Положим прежде всего

$$x = \frac{m}{\mu} + t.$$

Числитель выражения (16) примет вид:

$$\int_{-\alpha}^{+\alpha} \left(\frac{m}{\mu} + t \right)^m \left(\frac{n}{\mu} - t \right)^n dt,$$

или

$$\frac{m^m n^n}{\mu^\mu} \int_{-\alpha}^{+\alpha} \left(1 + \frac{\mu t}{m} \right)^m \left(1 - \frac{\mu t}{n} \right)^n dt \dots \dots \dots (17)$$

Переменное t есть отклонение величины x от наивероятнейшего ее значения $\frac{m}{\mu}$; мы можем его рассматривать, как весьма малую величину, не превышающую α по абсолютному значению. Величины m , n и μ мы считаем весьма большими, а величины $\frac{m}{\mu}$ и $\frac{n}{\mu}$ — конечными.

Пользуясь этими замечаниями, разлагаем подынтегральную функцию в ряд по степеням t . Возьмем сначала логарифмы отдельных множителей подынтегральной функции и применим к ним формулу Маклорена, ограничиваясь двумя младшими членами:

$$\lg \left(1 + \frac{\mu t}{m} \right)^m = \mu t - \frac{\mu^2 t^2}{2m} + \dots$$

$$\lg \left(1 - \frac{\mu t}{n} \right)^n = -\mu t - \frac{\mu^2 t^2}{2n} - \dots$$

Сложив эти равенства, находим:

$$\lg \left(1 + \frac{\mu t}{m} \right)^m \left(1 - \frac{\mu t}{n} \right)^n = -\frac{\mu^2 t^2}{2} \left(\frac{1}{m} + \frac{1}{n} \right) + \dots$$

Опуская все члены, кроме первого, приняв кроме того во внимание, что $m + n$ равно μ , и совершив переход от логарифмов к числам, мы представим подынтегральную функцию интеграла (17) в таком виде:

$$e^{-\frac{\mu^2 t^2}{2mn}}$$

Введя для краткости обозначение:

$$\bar{h} = \sqrt{\frac{\mu^2}{2mn}}, \quad (18)$$

мы находим такое выражение подынтегральной функции в интеграле (17):

$$e^{-\bar{h}^2 t^2}$$

а сама функция (17) примет вид:

$$\frac{m^m n^n}{\mu^\mu} \int_{-\alpha}^{+\alpha} e^{-\bar{h}^2 t^2} dt,$$

или, так как подынтегральная функция четная:

$$\frac{2m^m n^n}{\mu^\mu} \int_0^{\alpha} e^{-\bar{h}^2 t^2} dt.$$

Наконец, после подстановки:

$$ht = x,$$

мы представим ее в такой форме:

$$\frac{2m^m n^n}{\mu^\mu \bar{h}} \int_0^{\bar{h}\alpha} e^{-x^2} dx.$$

Таков числитель формулы (16). Сама эта формула примет вид:

$$\omega^{\frac{m}{\mu} + \alpha} = C \int_0^{\bar{h}\alpha} e^{-x^2} dx, \quad (19)$$

где

$$C = \frac{2m^m n^n}{\mu^\mu \bar{h}} \cdot \frac{1}{\int_0^1 x^m (1-x)^n dx} *).$$

*) Эту величину можно вычислить непосредственно, пользуясь приведенною выше формулою

$$\int_0^1 x^m (1-x)^n dx = \frac{m! n!}{\mu + 1!}$$

Для определения постоянного C припомним, что формула (19) дает величину вероятности того, что отклонение вероятности события E от частоты $\frac{m}{\mu}$ не превышает α по абсолютной величине. Отсюда следует, что при $\alpha = \infty$ величина (19) должна равняться 1:

$$1 = C \int_0^{\infty} e^{-x^2} dx = C \cdot \frac{\sqrt{\pi}}{2},$$

откуда

$$C = \frac{2}{\sqrt{\pi}}.$$

Внося эту величину C в формулу (19), найдем:

$$\omega \frac{\frac{m}{\mu} + \alpha}{\frac{m}{\mu} - \alpha} = \frac{2}{\sqrt{\pi}} \int_0^{h\alpha} e^{-x^2} dx.$$

Пользуясь обозначением, введенным выше (см. формулу (28) главы III, § 6), можем писать:

$$\omega \frac{\frac{m}{\mu} + \alpha}{\frac{m}{\mu} - \alpha} = \Phi(h\alpha) \quad (21)$$

Сравним эту формулу с формулой (29) только-что названного параграфа:

$$P = \Phi(h\alpha),$$

где

$$h = \sqrt{\frac{\mu}{2pq}}.$$

В настоящем параграфе частость $\frac{m}{\mu}$ считается данною. Вероятность же x события E неизвестною; формула (21) показывает, какова вероятность того, что отклонение

$$\frac{m}{\mu} - x,$$

по абсолютному значению не превзойдет α ; эта вероятность выражается формулою:

$$\Phi(h\alpha),$$

и приближенною формулою Стирлинга;

$$m! = m^m \sqrt{2\pi m} e^{-m}, \quad n! = n^n \sqrt{2\pi n} e^{-n}, \quad \mu! = \mu^\mu \sqrt{2\pi \mu} e^{-\mu}.$$

Внося эти выражения в найденную формулу C и приняв во внимание, что μ бесконечно-велика, найдем:

$$C = \frac{2}{\sqrt{\pi}}.$$

где

$$\bar{h} = \sqrt{\frac{\mu^3}{2mn}} \dots \dots \dots (18)$$

В предшествующей главе была дана вероятность p события E , а частость $\frac{m}{\mu}$ неизвестна; величина P обозначала вероятность того, что отклонение

$$\frac{m}{\mu} - p,$$

по абсолютному значению не превзойдет α ; эта вероятность выражалась формулой:

$$P = \Phi(h\alpha),$$

где

$$h = \sqrt{\frac{\mu}{2pq}}.$$

Обе формулы: $\Phi(h\alpha)$ и $\Phi(\bar{h}\alpha)$ выражают одну и ту же вероятность отклонения частоты от вероятности. Так как обе формулы приближенные, то они должны быть или в точности равны между собою, или весьма близки одна к другой. Различие в этих формулах состоит только в том, что в одну входит величина h , а в другую вместо ее величина $\bar{h}$. Поэтому надо сравнить h с $\bar{h}$.

Из формулы (18) следует:

$$\bar{h} = \sqrt{\frac{\mu}{2 \frac{m}{\mu} \cdot \frac{n}{\mu}}}.$$

По теореме Бернулли при бесконечно-большом μ величины $\frac{m}{\mu}$ и $\frac{n}{\mu}$ бесконечно-близки к p и q ; поэтому h бесконечно-близко к $\bar{h}$; а потому $\Phi(h\alpha)$ бесконечно-близко к $\Phi(\bar{h}\alpha)$.

Отсюда следует, что теорема Бернулли и теорема, только-что доказанная в настоящем параграфе, утверждают одно и то же: при бесконечно-большом числе испытаний частость разнится от вероятности бесконечно-мало.

Поэтому естественно, что теорема настоящего параграфа носит название: теорема, обратная теореме Бернулли.

Она лежит в основании большинства приложений теории вероятности к статистике: в практических вопросах, за редким исключением вероятность а priori нам неизвестна; приходится путем подсчета находить частость и принимать ее за вероятность на основании теоремы, обратной теореме Бернулли.

В виде примера применения этой теоремы возьмем снова задачу, уже дважды рассмотренную, а именно, пример 1 параграфа 10 главы III.

В урну положено 20 белых шаров и 20 черных. Из нее было произведено вынутие шара 10000 раз, при чем вынутый шар каждый раз возвращался в урну.

Оказалось, что белый шар появился 5011 раз, а черный 4989 раз.

В данном случае:

$$\bar{h} = \sqrt{\frac{10000}{2 \cdot 0,5011 \cdot 0,4989}} = 141,44.$$

Для величины h мы нашли:

$$h = 141,421.$$

Принимая снова за достоверное событие, вероятность которого равна 0.9999 или больше, опять полагаем: $\gamma = 2,76$. В таком случае по теореме, обратной теореме Бернулли, можем сказать: отклонение вероятности вынутия белого шара от наблюдаемой частоты:

$$\frac{m}{p} = 0,5011$$

не превысит предела:

$$\frac{2,76}{141,44} = 0,0195.$$

Раньше, применяя прямую теорему Бернулли, мы нашли для того же предела ту же величину

$$0,0195.$$

ЧАСТЬ ВТОРАЯ.

ПРИЛОЖЕНИЕ ТЕОРИИ ВЕРОЯТНОСТЕЙ
К ОБРАБОТКЕ РЕЗУЛЬТАТОВ
ИЗМЕРЕНИЙ.

ГЛАВА ПЕРВАЯ.

Теория ошибок при непосредственном измерении.

§ 1. Происхождение и характер ошибок измерения.

Всякое измерение сопровождается некоторой погрешностью. Это происходит от весьма многих причин. Укажем лишь на главнейшие из них.

Прежде всего необходимо заметить, что задача всякого измерения несколько неопределенная. Возьмем хотя бы простейший пример, когда измерению подлежит длина сколь угодно тонкого стержня. Во-первых, как бы тонок ни был стержень, он представляет собою материальное тело, имеющее толщину, а потому концы его представляют собою площадки конечных размеров; расстояния между различными парами точек, лежащих на этих площадках, различны, а между тем все эти расстояния с одинаковым правом могут быть названы длиной стержня.

Далее, расстояние между двумя точками может быть иррациональным, а результат измерения всегда выражается рациональным числом.

Затем длина стержня во время производства измерений может меняться, вследствие влияния внешних причин: температуры, прогиба от тяжести и т. д.

Кроме этой неопределенности задачи, источником ошибок служат: несовершенства наших органов чувств, не позволяющие нам различать величины слишком малые; далее, несовершенство инструментов, напр., гнутые зрительной трубы, эксцентричность ее оси, неравенство делений шкалы и т. д.; как бы тщательно ни был изготовлен инструмент, избежать всех этих неточностей нельзя.

Наконец, ошибки происходят и от личных особенностей измеряющих: напр., при определении момента наступления какого-нибудь события по хронометру, один наблюдатель будет запаздывать, другой — опережать наступление наблюдаемого момента.

Ошибки измерения делятся на две категории.

К первой категории относятся так называемые систематические ошибки, искажающие результат в определенную сторону и по более или менее ясному закону. Сюда относятся ошибки от несовершенства инструментов (инструментальные ошибки), вследствие личных особенностей наблюдателей (личные ошибки) и т. д.

Прежде чем приступить к обработке полученного материала измерений, его освобождают от систематических ошибок. С этой целью разрабатывается теория соответствующих инструментов, составляются таблицы поправок отдельных инструментов, располагаются измерения так, чтобы ошибки в противоположные стороны взаимно исключались и т. д.

Подробности этих приемов входят в курсы тех научных отделов, к которым относятся соответствующие измерения.

Прилагая теорию вероятностей, т.-е. теорию явлений случайных, мы предполагаем, что измерения освобождены на сколько возможно от ошибок систематических.

Ошибки, которые останутся в полученном материале по выделении ошибок систематических, будут носить случайный характер. Это будут ошибки второй категории, называемые случайными ошибками.

§ 2. Свойства случайных ошибок измерений.

Относительно случайных ошибок можем сказать следующее, предполагая, что измерения произведены с возможной тщательностью: 1) равные ошибки в ту и другую сторону равно вероятны; 2) малые ошибки более вероятны, чем большие; 3) ошибки, большие некоторого предела, невозможны. Этот предел мы называем пределом возможных ошибок. Он определяется условиями измерения, свойствами инструмента и искусством наблюдателя. Во всяком случае этот предел известен заранее, когда дан тот материал, который подлежит обработке: напр., если измеряемая длина выражается в долях миллиметра, то допустить возможность ошибки в миллиметрах невозможно.

Предел возможной ошибки мы обозначим буквою E .

Вероятность того, что ошибка измерения не превосходит величины ε , мы будем обозначать символом $F(\varepsilon)$. Ясно, что это — функция величины ε .

Вероятность того, что ошибка не превзойдет величину $\varepsilon + d\varepsilon$ будет:

$$F(\varepsilon + d\varepsilon).$$

На основании теоремы сложения вероятностей можем сказать, что разность:

$$F(\varepsilon + d\varepsilon) - F(\varepsilon)$$

выразит вероятность того, что ошибка измерения лежит между ε и $\varepsilon + d\varepsilon$.

Если $d\varepsilon$ бесконечно-мало, то эта разность приближенно выражается так:

$$F(\varepsilon + d\varepsilon) - F(\varepsilon) = F'(\varepsilon)d\varepsilon,$$

при чем $F'(\varepsilon)$ обозначает производную от $F(\varepsilon)$; следовательно, предполагается, что функция $F(\varepsilon)$ имеет производную.

Обозначая для краткости эту производную символом $\varphi(\varepsilon)$, представим вероятность того, что ошибка измерения лежит между ε и $\varepsilon + d\varepsilon$, такою формулою:

$$\varphi(\varepsilon)d\varepsilon.$$

Коротко будем говорить, что $\varphi(\varepsilon)d\varepsilon$ есть вероятность ошибки равной ε .

Сказанные выше три свойства случайной ошибки измерения выражаются следующим образом:

1) Функция $\varphi(\varepsilon)$ четная:

$$\varphi(-\varepsilon) = \varphi(\varepsilon), \dots \dots \dots (1)$$

потому что равные ошибки в ту и другую сторону равно вероятны.

2) Функция $\varphi(\varepsilon)$ убывает с возрастанием ε по абсолютной величине, потому что малые ошибки вероятнее, чем большие:

$$\varphi'(\varepsilon) < 0, \text{ когда } \varepsilon > 0. \dots \dots \dots (2)$$

3) Если E есть предел возможных ошибок, то

$$\varphi(\varepsilon) = 0, \text{ при } \varepsilon \geq E \dots \dots \dots (3)$$

Если мы изобразим линией течение функции:

$$y = \varphi(\varepsilon), \dots \dots \dots (4)$$

то на основании сказанного можем охарактеризовать вид этой кривой следующим образом:

- 1) кривая симметрична относительно оси y ;
- 2) она имеет наибольшее значение ординаты при $x = 0$, а затем справа и слева приближается к оси абсцисс с увеличением ε по абсолютной величине;
- 3) она спускается до оси абсцисс при $\varepsilon = E$, и далее ординаты ее точек равны нулю.

Кривая, таким образом полученная, выражает закон случайных ошибок наблюдения.

Сумма значения функции:

$$\varphi(\varepsilon)d\varepsilon$$

от a до b выражает вероятность того, что ошибка лежит между a и b . Предел этой суммы есть интеграл:

$$\int_a^b \varphi(\varepsilon) d\varepsilon.$$

Поэтому можно сказать, что написанный интеграл выражает вероятность того, что ошибка лежит в пределах между a и b .

Графически он выражает площадь криволинейной трапеции, ограниченной кривою:

$$y = \varphi(\varepsilon), \dots \dots \dots (4)$$

осью абсцисс и ординатами, соответствующими абсциссам a и b .

Вся площадь этой кривой от $-E$ до $+E$ равна 1, так как достоверно, что ошибка лежит в этих пределах:

$$\int_{-E}^{+E} \varphi(\varepsilon) d\varepsilon = 1. \dots \dots \dots (5)$$

Так как величина $\varphi(\varepsilon)$ равна нулю, когда ε по абсолютному значению больше E , то пределы интеграла можно раздвинуть до бесконечности и написать:

$$\int_{-\infty}^{+\infty} \varphi(\varepsilon) d\varepsilon = 1. \dots \dots \dots (5')$$

Рассматривая ошибку, как случайную величину и имея вероятность каждого значения ее, можем вычислить математическое ожидание ошибки, а также математическое ожидание абсолютной величины ошибки и математическое ожидание квадрата ошибки.

При вычислении этих математических ожиданий сумма опять в пределе заменится определенным интегралом, и мы получим следующие выражения *):

1) математическое ожидание ошибки:

$$\int_{-\infty}^{+\infty} \varepsilon \varphi(\varepsilon) d\varepsilon = 0,$$

вследствие равенства (1);

2) математическое ожидание абсолютной величины ошибки:

$$\int_{-\infty}^{+\infty} |\varepsilon| \varphi(\varepsilon) d\varepsilon = 2 \int_0^{+\infty} \varepsilon \varphi(\varepsilon) d\varepsilon$$

*) Ср. формулы параграфов 8 и 9, главы III, ч. I.

3) математическое ожидание квадрата ошибки:

$$\int_{-\infty}^{+\infty} \varepsilon^2 \varphi(\varepsilon) d\varepsilon = 2 \int_0^{+\infty} \varepsilon^2 \varphi(\varepsilon) d\varepsilon.$$

Так как математическое ожидание какого-либо переменного равно среднему значению этого переменного, то мы в праве сказать, что математическое ожидание абсолютной величины ошибки есть среднее значение этой величины. Мы обозначим ее буквою d и назовем среднюю ошибкою в том же смысле, как в главе III, части I, говорили о среднем уклонении. В таком случае будем иметь:

$$d = 2 \int_0^{\infty} \varepsilon \varphi(\varepsilon) d\varepsilon. \quad (6)$$

На таком же основании мы назовем квадратный корень из математического ожидания квадрата ошибки среднюю квадратического ошибкою и обозначим ее буквою σ . В таком случае будем иметь:

$$\sigma^2 = 2 \int_0^{\infty} \varepsilon^2 \varphi(\varepsilon) d\varepsilon. \quad (7)$$

§ 3. Начало арифметической середины. Вывод формулы Гаусса вероятности ошибки измерения.

Гаусс предложил вывод формулы $\varphi(\varepsilon)$, исходя из так называемого начала арифметической середины *), состоящего в следующем: если произведен ряд измерений одного и того же предмета со всею возможною тщательностью одним и тем же лицом и тем же способом, одним и тем же инструментом и при одинаковых условиях, то наимвероятнейший результат, который может быть получен из этих измерений, есть среднее арифметическое всех отдельных полученных результатов.

Это положение есть постулат, а не теорема; доказать справедливость его невозможно тем более, что в него входит такое неопределенное понятие, как измерения, произведенные со всею возможною тщательностью. Понятно, что ряд случайно написанных чисел не подчинился бы никакому закону, а величины, найденные недостаточно тщательно, мало чем разнятся от случайно написанных чисел. Поэтому начало арифметической середины будет тем ближе к истине, чем тщательнее выполнены измерения. Тем не менее, при выводе формулы $\varphi(\varepsilon)$ приходится считать начало арифметической середины абсолютно спра-

*) Gauss. Theoria motus corporum coelestium. 1809.

ведливым. Этот недостаток вывода однако искупается тем, что полученная помощью него формула прекрасно согласуется с действительностью.

Пусть измерен некоторый предмет, величина которого нам неизвестна; мы обозначим ее буквою x . Результаты измерений мы обозначим буквами:

$$x_1, x_2, x_3, \dots, x_\mu \dots \dots \dots (8)$$

Обозначим среднее арифметическое величин (8) буквою ξ :

$$\xi = \frac{x_1 + x_2 + x_3 + \dots + x_\mu}{\mu} = \frac{\sum x_i}{\mu} \dots \dots \dots (9)$$

Согласно началу арифметической середины величина ξ есть наиболее вероятнейшее значение величины x на основании произведенных измерений (8).

Из равенства (9) следует:

$$\mu\xi - \sum x_i = 0,$$

или:

$$\sum (\xi - x_i) = 0 \dots \dots \dots (10)$$

Истинной величины x мы не знаем. Мы можем только делать различные гипотезы относительно этой величины.

Сделаем гипотезу, что истинная величина измеренного предмета равна некоторому числу x . При этой гипотезе ошибки полученных результатов измерения (8) выразятся так:

$$\left. \begin{aligned} x - x_1 &= \varepsilon_1 \\ x - x_2 &= \varepsilon_2 \\ x - x_3 &= \varepsilon_3 \\ x - x_\mu &= \varepsilon_\mu \end{aligned} \right\} \dots \dots \dots (11)$$

При сделанной гипотезе системе результатов измерений (8) соответствует система ошибок (11), и обратно: системе ошибок (11) соответствует система результатов измерения (8). Поэтому вероятность события, состоящего в том, что при μ испытаниях получились результаты (8), равна вероятности события, что при μ измерениях получились ошибки (11). Но первое из этих событий есть то самое, которое наблюдается в действительности, когда произведено μ измерений. Вероятность его при рассматриваемой гипотезе есть вероятность, сообщаемая данной гипотезой наблюдаемому событию. Мы видели, что эта вероятность равна вероятности получения при μ измерениях системы ошибок (11).

Вероятность получения системы ошибок есть событие сложное; по теореме умножения вероятностей она равна произведению вероятностей ошибок:

$$\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_\mu$$

в отдельности, т. е. она равна произведению величин:

$$\varphi(z_1)dz, \varphi(z_2)dz, \varphi(z_3)dz, \dots, \varphi(z_n)dz.$$

Итак, мы можем сказать, что при принятой гипотезе вероятность наблюдаемого события такова:

$$\varphi(z_1) \varphi(z_2) \varphi(z_3) \dots \varphi(z_n) dz^n,$$

или, на основании равенств (11):

$$\varphi(x-x_1) \varphi(x-x_2) \varphi(x-x_3) \dots \varphi(x-x_n) dz^n. \dots (12)$$

Так как о величине x заранее ничего неизвестно, то все гипотезы а priori равно вероятны, а потому, применяя сказанное в конце параграфа 2 главы IV части I, можем утверждать, что вероятность рассматриваемой гипотезы в данном случае пропорциональна величине (12), т. е. она выражается формулой:

$$C. \varphi(x-x_1) \varphi(x-x_2) \varphi(x-x_3) \dots \varphi(x-x_n) dz^n. \dots (13)$$

То значение x , при котором величина (13) имеет наибольшее значение есть наиболее вероятное значение x .

Чтобы найти это наиболее вероятное значение, мы должны приравнять нулю производную выражения (13); или, что в данном случае проще, приравнять нулю производную логарифма выражения (13); получим следующее:

$$\frac{\varphi'(x-x_1)}{\varphi(x-x_1)} + \frac{\varphi'(x-x_2)}{\varphi(x-x_2)} + \frac{\varphi'(x-x_3)}{\varphi(x-x_3)} + \dots + \frac{\varphi'(x-x_n)}{\varphi(x-x_n)} = 0.$$

Положив для краткости:

$$\frac{\varphi'(x)}{\varphi(x)} = \psi(x), \dots (14)$$

представим это равенство в таком виде:

$$\psi(x-x_1) + \psi(x-x_2) + \psi(x-x_3) + \dots + \psi(x-x_n) = 0.$$

Таково равенство, определяющее наиболее вероятное значение x . Так как по началу арифметической середины наиболее вероятное значение x есть величина ξ , определяемая формулой (9), то мы получаем тождество:

$$\psi(\xi-x_1) + \psi(\xi-x_2) + \psi(\xi-x_3) + \dots + \psi(\xi-x_n) = 0. \dots (15)$$

Как тождество, формула (15) должна быть справедлива при всевозможных значениях букв: $x_1, x_2, x_3, \dots, x_n$; а потому ее можно по каждому из этих аргументов дифференцировать.

Произведя такие дифференцирования и принимая во внимание равенство (9), находим:

$$\psi'(\xi - x_1) = \frac{1}{\mu} \left\{ \psi'(\xi - x_1) + \psi'(\xi - x_2) + \psi'(\xi - x_3) + \dots + \psi'(\xi - x_\mu) \right\},$$

$$\psi'(\xi - x_2) = \frac{1}{\mu} \left\{ \psi'(\xi - x_1) + \psi'(\xi - x_2) + \psi'(\xi - x_3) + \dots + \psi'(\xi - x_\mu) \right\},$$

$$\psi'(\xi - x_\mu) = \frac{1}{\mu} \left\{ \psi'(\xi - x_1) + \psi'(\xi - x_2) + \psi'(\xi - x_3) + \dots + \psi'(\xi - x_\mu) \right\}.$$

Так же, как и равенства (15), эти равенства суть тождества. Из них следуют новые тождества:

$$\psi'(\xi - x_1) = \psi'(\xi - x_2) = \psi'(\xi - x_3) = \dots = \psi'(\xi - x_\mu),$$

откуда видно, что функция $\psi(x)$ есть величина постоянная:

$$\psi(x) = C.$$

Отсюда:

$$\psi(x) = Cx + C',$$

где C' тоже постоянное.

Внося это выражение в формулу (14), находим:

$$\frac{\varphi'(x)}{\varphi(x)} = Cx + C',$$

откуда:

$$\lg \varphi(x) = \frac{1}{2} Cx^2 + C'x + C'',$$

и следовательно:

$$\varphi(x) = e^{\frac{1}{2} Cx^2 + C'x + C''} \quad (16)$$

Для определения величин постоянных, припомним свойства функции $\varphi(\varepsilon)$.

Из равенства

$$\varphi(-\varepsilon) = \varphi(\varepsilon) \quad (1)$$

следует, что функция (16) должна быть четной, а потому:

$$C' = 0.$$

Приняв это во внимание, находим:

$$\varphi(x) = e^{\frac{1}{2} Cx^2 + C''}$$

откуда:

$$\varphi'(x) = Cx \cdot \varphi(x).$$

Из равенства:

$$\varphi'(\varepsilon) < 0 \quad \text{когда } \varepsilon > 0,$$

следует, что C отрицательно. Поэтому можем положить:

$$\frac{1}{2} C = -h^2,$$

где h действительная положительная величина.

Формула (16) примет вид:

$$\varphi(x) = e^{C''} \cdot e^{-h^2 x^2} \dots \dots \dots (17)$$

Для нахождения величины C'' припомним формулу:

$$\int_{-\infty}^{+\infty} \varphi(z) dz = 1 \dots \dots \dots (5')$$

Внеся сюда выражение (17) подынтегральной функции, находим:

$$e^{C''} \int_{-\infty}^{+\infty} e^{-h^2 z^2} dz = 1,$$

или после подстановки: $hx = t$:

$$\frac{e^{C''}}{h} \int_{-\infty}^{+\infty} e^{-t^2} dt = 1.$$

Заменяя полученный интеграл Пуассона его выражением $\sqrt{\pi}$, находим:

$$e^{C''} \frac{\sqrt{\pi}}{h} = 1.$$

Разделив формулу (17) на эту величину, равную 1, находим:

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

Поэтому искомое выражение вероятности ошибки таково:

$$\varphi(z) dz = \frac{h}{\sqrt{\pi}} e^{-h^2 z^2} dz \dots \dots \dots (18)$$

Мы видим, что оно тождественно с выражением (20) параграфа 4 главы III части I, т.е. вероятность ошибки выражается той же формулой, как и вероятность отклонения частоты от вероятности. Это делается вполне понятным, если мы вспомним, что частость есть приближенная величина вероятности; мы производим измерение неизвестной вероятности p события E , делая ряд μ испытаний; результат этого измерения есть частость $\frac{m}{\mu}$; отклонение ее от вероятности p , т.е. разность:

$$p - \frac{m}{\mu} = x$$

есть ошибка результата измерения вероятности события E .

Вполне естественно, что в этом частном случае вероятность ошибки выражается той же формулой (18), которая выражает вероятность ошибки в общем случае.

Из тождественности формулы (18) настоящего параграфа с формулой (20) параграфа 4 главы III части I следует, что закон вероятностей ошибок выражается той же кривой, какую мы видели на чертеже 6. Наивысшая ее точка соответствует абсциссе равной 0, кривая симметрична относительно оси ординат, приближается к оси абсцисс по мере удаления от начала координат и имеет ось абсцисс своею асимптотой.

С первого взгляда это последнее обстоятельство может показаться противоречащим приведенному выше свойству 3-у случайных ошибок, по которому должен существовать крайний предел E возможных ошибок. Это сомнение устранилось тем соображением, что функция:

$$y = \varphi(\varepsilon) = \frac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon^2} \dots \dots \dots (19)$$

быстро уменьшается с увеличением ε по абсолютной величине, в особенности при значительной величине h , как это видно из числовых таблиц параграфа 4 главы III части I.

Значение параметра h в настоящем случае то же самое, как и в рассмотренном прежде частном случае: h называется мерой точности произведенных измерений. Чем h более, тем больше точки площади кривой (19) сосредоточены около оси y . Величина h характеризует точность произведенных измерений.

§ 4. Предел возможных ошибок, вероятная ошибка, ошибка средняя и средняя квадратическая ошибка.

В параграфе 1 мы видели, что формула

$$\int_a^b \varphi(\varepsilon) d\varepsilon$$

выражает вероятность того, что ошибка содержится в пределах между a и b .

Внося вместо $\varphi(\varepsilon)$ ее выражение (19) и положив:

$$a = -\alpha, \quad b = +\alpha,$$

находим величину:

$$P_{\alpha} = \frac{h}{\sqrt{\pi}} \int_{-\alpha}^{+\alpha} e^{-h^2 \varepsilon^2} d\varepsilon,$$

выражающую вероятность того, что ошибка заключается между $-\alpha$ и $+\alpha$, т.е. не превосходит α по абсолютной величине.

Мы назовем ее вероятностью предела α ошибок.

Совершив уже знакомые преобразования, найдем:

$$P_{\alpha} = \frac{2}{\sqrt{\pi}} \int_0^{h\alpha} e^{-x^2} dx = \Phi(h\alpha) \quad (20)$$

Выберем такое число γ , чтобы $\Phi(\gamma)$ можно было принять равным 1. Напр., положив по-прежнему γ равным $\gamma = 2,76$, будем иметь:

$$\Phi(\gamma) = 0,9999.$$

Принимая эту дробь приблизительно за 1, мы возьмем величину α , определяемую равенством:

$$h\alpha = \gamma, \quad \alpha = \frac{\gamma}{h} \quad (21)$$

Эта величина есть предел возможных ошибок.

В параграфе 7 главы III часть I мы видели, что величина ρ , удовлетворяющая равенству

$$\Phi(\rho) = \frac{1}{2},$$

такова:

$$\rho = 0,4769362762.$$

Положив:

$$h\alpha = \rho,$$

мы обозначим определяемую из этого равенства величину α буквою r и назовем вероятною ошибкою:

$$r = \frac{\rho}{h} \quad (22)$$

В параграфе 2 мы установили понятие о средней ошибке d и о средней квадратической ошибке σ .

Внеся в формулы (6) и (7) выражение (19) функции $\varphi(x)$ получим следующее:

$$d = \frac{2h}{\sqrt{\pi}} \int_0^{\infty} e^{-h^2 z^2} z dz$$

$$\sigma^2 = \frac{2h}{\sqrt{\pi}} \int_0^{\infty} e^{-h^2 z^2} z^2 dz,$$

или, произведя вычисления, приведенные в параграфах 8 и 9 главы III части I:

$$d = \frac{1}{h\sqrt{\pi}}, \quad (23)$$

$$\sigma = \frac{1}{h\sqrt{2}}. \quad (24)$$

Из формул (22), (23), (24) следует:

$$h = \frac{\rho}{r} = \frac{1}{d \sqrt{\pi}} = \frac{1}{\sigma \sqrt{2}} \dots \dots \dots (25)$$

совершенно так же, как мы видели в главе III части I.

Из формулы (21) видно, что чем больше h , тем меньше возможный предел ошибок, тем лучше сделанные измерения.

Формула (25) показывает, что величины r , d , σ обратно пропорциональны мере точности h ; следовательно, чем они больше, тем ниже качество произведенных измерений.

Выше мы говорили, что формула:

$$\int_a^b \varphi(z) dz$$

выражает вероятность того, что ошибка содержится между a и b .

Пользуясь формулами (19) и (20), имеем:

$$\int_a^b \varphi(z) dz = \int_0^b \varphi(z) dz - \int_0^a \varphi(z) dz = \frac{1}{2} \left\{ \Phi(hb) - \Phi(ha) \right\}.$$

На основании теоремы Бернулли можем сказать: если произведено большое число μ испытаний над событием E , вероятность которого p , то наимвероятнейшее число появления события мало разнится от μp .

Применяя это предложение к данному случаю, можем сказать: если произведено большое число μ измерений величины x , то число

результатов с ошибкою, лежащее между a и b близко к $\mu \int_a^b \varphi(x) dx$,

т.е. оно близко к величине:

$$\frac{\mu}{2} \left\{ \Phi(hb) - \Phi(ha) \right\}.$$

Число результатов, у которых абсолютная величина ошибки лежит между a и b , будет вдвое больше, т.е.

$$\mu \left\{ \Phi(hb) - \Phi(ha) \right\} \dots \dots \dots (26)$$

Этой формулой можно воспользоваться для численной проверки формулы Гаусса. Если мы произведем измерения какой-либо известной величины, то, сравнивая эту известную величину с результатами измерений, мы можем непосредственно подсчитать число ошибок, лежащих в различных промежутках. С другой стороны, зная h или

одну из величин r , d , σ , мы найдем те же числа ошибок для различных промежутков по формуле (26).

Сравнивая те и другие величины, будем в состоянии судить о точности результатов, даваемых формулой Гаусса.

Как найти h , мы рассмотрим в следующем параграфе. Здесь рассмотрим результат сказанного сравнения, приводимый Мейером (Meyer, Vorlesungen über Wahrscheinlichkeitsrechnung, стр. 253) и заимствованный им в Berliner Astronom. Jahrbuch, 1834.

Был измерен некоторый известный угол 470 раз, при чем вероятная ошибка оказалась равною $r = 0,2637$ (следовательно, $h = 1,8085$). Результаты оказались следующие:

Промежуток.	Число ошибок	
	по формуле.	в действительности.
0,0 — 0,1	95	94
0,1 — 0,2	89	88
0,2 — 0,3	78	78
0,3 — 0,4	64	58
0,4 — 0,5	50	51
0,5 — 0,6	36	36
0,6 — 0,7	24	26
0,7 — 0,8	15	14
0,8 — 0,9	9	10
0,9 — 1,0	5	7
свыше 1,0	5	8
	470	470

§ 5. Нахождение меры точности произведенных измерений.

Пусть измерена r раз некоторая неизвестная величина, которую мы обозначим буквою x . Пусть измерения произведены с возможною тщательностью, одним и тем же методом, одним и тем же лицом, при одинаковых условиях, следовательно, с одинаковою мерою точности, которую мы однако не знаем и только обозначаем буквою h .

Наша задача — найти наиболее вероятное значение h .

Пусть полученные результаты измерений таковы:

$$x_1, x_2, x_3, \dots, x_n \dots \dots \dots (8)$$

Сделаем двойную гипотезу: 1) относительно величины x , 2) относительно величины h .

При этой гипотезе, т. е. когда даны величины x и h , вероятность получить при μ измерениях систему результатов (8) выразится по формуле (12) предшествующего параграфа:

$$\varphi(x-x_1) \varphi(x-x_2) \varphi(x-x_3) \dots \varphi(x-x_\mu) d\varepsilon^\mu \dots \quad (12)$$

или, заменяя функцию φ ее выражением:

$$\frac{h^\mu}{\pi^{\frac{\mu}{2}}} e^{-h^2 \sum (x-x_i)^2} d\varepsilon^\mu.$$

Величина вероятности самой гипотезы, по теореме Бейеса, пропорциональна этой величине, так как а priori все значения x и все значения h равно вероятны.

Присоединяя к множителю пропорциональности все те множители, которые не зависят ни от x , ни от h , и, кроме того, умножив и разделив этот множитель пропорциональности на величину dx , тоже не зависящую ни от x , ни от h , мы представим вероятность сделанной двойной гипотезы в таком виде:

$$Ch^\mu e^{-h^2 \sum (x-x_i)^2} dx, \dots \quad (27)$$

где C величина постоянная относительно x и h .

Если мы хотим найти вероятность простой гипотезы, относительно только одной величины h , при каком бы то ни было значении x , то по теореме сложения вероятностей должны взять сумму значений выражения (27) при всевозможных значениях x .

Не повторяя рассуждений, уже много раз встречавшихся, можем сказать, что эта сумма выразится определенным интегралом:

$$Ch^\mu \int_{-\infty}^{+\infty} e^{-h^2 \sum (x-x_i)^2} dx \dots \quad (28)$$

Такова вероятность гипотезы, что мера точности равна h .

Для вычисления интеграла преобразуем переменное интегриации, полагая:

$$x = \xi + \delta,$$

где ξ по-прежнему среднее арифметическое:

$$\xi = \frac{\sum x_i}{\mu},$$

откуда, как мы видели:

$$\sum (\xi - x_i) = 0 \dots \quad (10)$$

Выражение (28) примет вид:

$$C \cdot h^{\mu} \int_{-\infty}^{+\infty} e^{-h^2 \sum (\xi - x_i)^2 - 2h^2 \sum (\xi - x_i) \delta - h^2 \mu \delta^2} d\delta,$$

или, на основании равенства (10):

$$C \cdot h^{\mu} \cdot e^{-h^2 \sum (\xi - x_i)^2} \int_{-\infty}^{+\infty} e^{-h^2 \mu \delta^2} d\delta,$$

или, после подстановки:

$$h \sqrt{\mu} \delta = t;$$

$$\frac{C}{\sqrt{\mu}} \cdot h^{\mu-1} e^{-h^2 \sum (\xi - x_i)^2} \int_{-\infty}^{+\infty} e^{-t^2} dt = \frac{C \cdot \sqrt{\pi}}{\sqrt{\mu}} h^{\mu-1} e^{-h^2 \sum (\xi - x_i)^2} \dots \dots (29)$$

Такова вероятность, что мера точности равна h . То значение h , при котором величина (29) наибольшая, есть наивероятнейшее значение h . Для его нахождения мы приравняем нулю производную выражения (29) по переменному h , или, что проще, приравняем нулю производную логарифма величины (29), получится следующее:

$$\frac{\mu-1}{h} = 2 \cdot h \sum (\xi - x_i)^2,$$

откуда:

$$h = \frac{1}{\sqrt{2 \cdot \frac{\sum (\xi - x_i)^2}{\mu-1}}} \dots \dots \dots (30)$$

Такова наивероятнейшая величина меры точности h ; обыкновенно она принимается за истинную величину меры точности.

Припомним найденную выше формулу:

$$\sigma = \frac{1}{h \sqrt{2}} \dots \dots \dots (24)$$

Сравнив ее с формулой (30), находим:

$$\sigma = \sqrt{\frac{\sum (\xi - x_i)^2}{\mu-1}} \dots \dots \dots (31)$$

Такова наивероятнейшая величина средней квадратической ошибки. Уясним смысл этого результата.

Если бы ξ было истинной величиной измеряемого предмета, то ошибка результата x_i равнялась бы:

$$\xi - x_i = e_i.$$

Корень квадратный из среднего арифметического всех величин ε_i^2 мы были бы в праве назвать средней квадратической ошибкой произведенных измерений и обозначали бы эту величину буквою σ . Она выразилась бы формулою:

$$\sqrt{\frac{\sum \varepsilon_i^2}{\mu}} = \sqrt{\frac{\sum (\xi - x_i)^2}{\mu}} \dots \dots \dots (31')$$

Найденная для σ величина (31) несколько больше величин (31'), что и понятно: так как ξ есть только приближенная величина x , то отклонение полученных результатов от этой приближенной величины в среднем больше, чем отклонение от истинной величины x .

Величина (31) разнится от величины (31') множителем:

$$\sqrt{\frac{\mu}{\mu-1}},$$

который больше единицы; но при большом μ это различие ничтожно.

Поэтому для простоты вычислений часто величину (31') принимают за среднюю квадратическую ошибку измерений.

Найдя величины h и σ , найдем величины r и d по формулам (25).

§ 6. Мера точности результата, выведенного из системы равноточных измерений.

Оставляя обозначения предшествующего параграфа, будем предполагать, что некоторая величина x измерена μ раз с одинаковою мерою точности h , величину которой мы умеем определить помощью приема, изложенного в предшествующем параграфе. Следовательно, величину h мы считаем уже известной.

Результаты μ измерений величины x по-прежнему обозначим буквами

$$x_1, x_2, x_3, \dots, x_\mu \dots \dots \dots (8)$$

На основании только-что сказанного мы назовем эти величины равноточными.

Зная величину h , мы теперь делаем гипотезу относительно только одной величины x . При этой гипотезе вероятность получения при μ измерениях системы результатов (8) выразится прежнею формулою:

$$\frac{h^\mu}{\pi^{\frac{\mu}{2}}} e^{-h^2 \sum (x - x_i)^2} dx^\mu.$$

Вероятность самой гипотезы относительно величины x будет пропорциональна этой величине, при чем множитель пропорциональности будет, понятно, другой, чем в предшествующем параграфе. Мы вклю-

чим в него все множители, не зависящие от x , и, кроме того, умножим его и разделим на величину dx , не зависящую от x .

Вероятность гипотезы относительно величины x выразится формулой:

$$C \cdot e^{-h^2 \sum (x - x_i)^2} dx \dots \dots \dots (32)$$

То значение x , при котором эта функция имеет наибольшую величину, будет наивероятнейшим значением переменного x .

Для его нахождения приравниваем нулю производную по x логарифма функции (32):

$$-2h^2 \sum (x - x_i) = 0,$$

откуда:

$$x = \frac{\sum x_i}{\mu},$$

т. е. наивероятнейшее значение величины x есть среднее арифметическое:

$$\xi = \frac{x_1 + x_2 + x_3 + \dots + x_\mu}{\mu} = \frac{\sum x_i}{\mu} \dots \dots \dots (9)$$

что и понятно, так как начало арифметической середины лежит в основе вывода формулы Гаусса.

Если мы обозначим буквою x истинную величину измеренного предмета, то разность

$$x - \xi = \delta \dots \dots \dots (33)$$

выразит величину ошибки, которую мы сделаем, принимая приближенную величину ξ равною истинной.

Формула (32) есть выражение вероятности того, что истинная величина измеряемого предмета равна x . Если в эту формулу внести вместо x ее выражение из равенства (33):

$$x = \xi + \delta, \dots \dots \dots (33')$$

то формула даст выражение вероятности того, что ошибка при среднем арифметическом ξ равна δ , короче назовем ее: вероятность ошибки δ при величине ξ .

Так как ξ величина постоянная, то из равенства (33') следует:

$$dx = d\delta;$$

формула (32) примет вид:

$$C \cdot e^{-h^2 \sum (\xi - x_i)^2} \cdot e^{-2h^2 \sum (\xi - x_i) \delta - h^2 \mu \delta^2} d\delta.$$

На основании равенства

$$\sum (\xi - x_i) = 0 \dots \dots \dots (10)$$

эта формула примет вид:

$$C \cdot e^{-h^2 \sum (\xi - x_i)^2} \cdot e^{-h^2 \mu \delta^2} d\delta \dots \dots \dots (34)$$

Такова формула вероятности ошибки δ при ξ . Она имеет такой же вид, как и найденная выше формула Гаусса (18) вероятности ошибки ε при измерении величины x .

Для нахождения постоянного C применим обычный способ: припомним, что сумма значений функции (34) при всевозможных значениях x должна равняться 1, как сумма вероятностей событий всевозможных и взаимно исключающих друг друга.

Так как $d\delta$ бесконечно мала, то эта сумма в пределе выразится определенным интегралом и мы получим:

$$C \cdot e^{-h^2 \sum (\xi - x_i)^2} \int_{-\infty}^{+\infty} e^{-h^2 \mu \delta^2} d\delta = 1;$$

после преобразования:

$$h\sqrt{\mu} \delta = t,$$

получим:

$$\frac{C}{h\sqrt{\mu}} e^{-h^2 \sum (\xi - x_i)^2} \int_{-\infty}^{+\infty} e^{-t^2} dt = 1,$$

или:

$$\frac{C \cdot \sqrt{\pi}}{h\sqrt{\mu}} e^{-h^2 \sum (\xi - x_i)^2} = 1.$$

Разделив на это выражение, равное единице, величину (84), найдем:

$$\frac{h\sqrt{\mu}}{\sqrt{\pi}} e^{-h^2 \mu \delta^2} d\delta.$$

Положим:

$$h\sqrt{\mu} = H. \dots \dots \dots (85)$$

Тогда найденное выражение вероятности ошибки δ при среднем арифметическом ξ выразится окончательно формулой:

$$\frac{H}{\sqrt{\pi}} e^{-H^2 \delta^2} d\delta. \dots \dots \dots (86)$$

Это — формула Гаусса, соответствующая мере точности H .

Она показывает, что ξ , т.-е. найвероятнейший вывод из системы произведенных μ равноточных измерений некоторого предмета, дает величину этого предмета с мерою точности H , т.-е. с мерою точности в $\sqrt{\mu}$ раз большею, чем отдельное измерение.

Пример *). Некоторый угол измерен 14 раз одним и тем же наблюдателем, одним и тем же теодолитом, так что результаты измерения можно считать равноточными. Эти результаты помещены в следующей таблице, где вместе с тем приведены и вычисления по формулам предшествующих параграфов.

№	x_i	$\xi - x_i$	$(\xi - x_i)^2$
1	17° 56' 45".00	— 5".37	28.84
2	31".25	+ 8".38	70.22
3	42".50	— 2".87	8.24
4	45".00	— 5".37	28.84
5	37".50	+ 2".13	4.54
6	38".33	+ 1".30	1.69
7	27".50	+ 12".13	147.14
8	43".33	— 3".70	13.69
9	40".63	— 1".00	1.00
10	36".25	+ 3".38	11.42
11	42".50	— 2".87	8.24
12	39".17	+ 0".46	0.21
13	45".00	— 5".37	28.84
14	40".83	— 1".20	1.44
	554".79	+ 27".78	354.35
	$\frac{554".79}{14} = 39".63$	$\frac{-27".75}{+0".03}$	$\sigma = \sqrt{\frac{354.35}{13}}$
			$= 5".22$

Наивероятнейшая величина измеряемого угла:

$$\xi = 17^\circ 56' 39".63.$$

Мера точности отдельного результата измерения:

$$h = \frac{1}{\sigma \sqrt{2}} = 0,135.$$

Мера точности величины ξ :

$$H = h\sqrt{14} = 0,51.$$

Вероятная ошибка величины ξ : 0".94.

Эта величина обыкновенно приписывается с двойным знаком $\pm$ при найденной величине ξ . В данном случае окончательный результат вычисления напишется так:

$$x = 17^\circ 56' 39".63 \pm 0".94.$$

*) См. Мёуер. Vorlesungen über Wahrscheinlichkeitsrechnung. Стр. 272.

§ 7. О весах неравноточных измерений.

Иногда приходится сводить вместе результаты измерений одной и той же величины, полученные при различных обстоятельствах: они могут быть найдены, как средние арифметические из различного числа частных измерений, или произведены различными инструментами, или различными методами, или наблюдателями не одинаково искусными и т. д.

Возникает вопрос, как сделать наивероятнейший вывод из таких неравноточных измерений.

Пусть найденная система неравноточных измерений такова:

$$x_1, x_2, x_3, \dots, x_n \dots \dots \dots (37)$$

Меры точности этих величин мы обозначим так:

$$h_1, h_2, h_3, \dots, h_n \dots \dots \dots (38)$$

Обозначим буквами:

$$g_1, g_2, g_3, \dots, g_n \dots \dots \dots (39)$$

какие угодно числа, удовлетворяющие соотношениям:

$$\frac{h_1^2}{g_1} = \frac{h_2^2}{g_2} = \frac{h_3^2}{g_3} = \dots = \frac{h_n^2}{g_n}.$$

Следовательно, числа (39) суть какие угодно числа, пропорциональные квадратам величин (38). Множитель пропорциональности остается произвольным.

Обозначив его буквою h^2 , будем иметь:

$$\frac{h_1^2}{g_1} = \frac{h_2^2}{g_2} = \frac{h_3^2}{g_3} = \dots = \frac{h_n^2}{g_n} = \frac{h^2}{1} \dots \dots \dots (40)$$

Отсюда следует:

$$h_1 = h \sqrt{g_1}, h_2 = h \sqrt{g_2}, h_3 = h \sqrt{g_3}, \dots, h_n = h \sqrt{g_n} \dots \dots (41)$$

Пользуясь произволом h , мы всегда можем выбрать числа (39) так, чтобы они были целыми. В самом деле, приведа числа

$$h_1^2, h_2^2, h_3^2, \dots, h_n^2$$

к общему знаменателю, мы можем выбрать числа (39) равными соответствующим числителям членов только-что выписанного ряда.

Эти числители суть числа целые, удовлетворяющие условиям (40).

Числа (39) называются весами результатов измерения (37) по причине, которая скоро выяснится.

Как мы только-что видели, веса всегда могут быть выражены в виде целых чисел, хотя в этом нет необходимости. В случае целых чисел веса получают особенно простое конкретное истолкование.

Величина h называется нормальной мерою точности, потому что она служит мерою точности такого результата, у которого вес равен единице.

Уясним конкретный смысл весов в том случае, когда они суть числа целые.

Представим себе, что было сделано g_1 измерений с нормальной мерою точности h , при чем средним арифметическим этих измерений оказалось x_1 . На основании сказанного в предшествующем параграфе мера точности этого среднего арифметического x_1 будет равна $h/\sqrt{g_1} = h_1$. Следовательно, вместо полученного в действительности измерения x_1 с мерою точности h_1 мы могли бы взять g_1 фиктивных величин с нормальной мерою точности h при условии, чтобы их среднее арифметическое было x_1 . Всего проще заменить одну величину x_1 с мерою точности h_1 группою величин, $x_1, x_1, x_1, \dots, x_1$ числом g_1 , с нормальной мерою точности h . Подобные же рассуждения можно применить и к остальным величинам: $x_2, x_2, \dots, x_\mu$.

Таким образом, полученный в действительности ряд результатов (37) с мерами точности (38) заменится равносильным с ним рядом фиктивных величин с одинаковою мерою точности h :

$$\underbrace{x_1, x_1, x_1, \dots, x_1}_{g_1}, \underbrace{x_2, x_2, x_2, \dots, x_2}_{g_2}, \dots, \underbrace{x_\mu, x_\mu, x_\mu, \dots, x_\mu}_{g_\mu}.$$

Уже отсюда можно заранее усмотреть, что наивероятнейший вывод будет среднее арифметическое из написанного ряда фиктивных величин, т.-е.

$$\bar{x} = \frac{g_1 x_1 + g_2 x_2 + \dots + g_\mu x_\mu}{g_1 + g_2 + \dots + g_\mu} = \frac{\sum g_i x_i}{G}, \dots \dots \dots (42)$$

где:

$$G = g_1 + g_2 + \dots + g_\mu, \dots \dots \dots (43)$$

мера же точности величины $\bar{x}$ выразится формулою:

$$H = h \sqrt{G}, \dots \dots \dots (44)$$

В дальнейших параграфах мы выведем эти формулы, не прибегая к понятию о фиктивных измерениях и для весов, хотя бы дробных. Теперь же заметим только, что, на основании сказанного выше, одно измерение с мерой точности h_1 равносильно, т.-е. имеет одинаковый вес с g_1 наблюдениями нормальной меры точности h . Отсюда и произошло название вес результата.

Если меры точности (38) известны, то веса находятся прямо по формулам (40). Если известны вероятные или средние квадратические ошибки величин (37), то веса найдутся по тем же формулам (40), если примем во внимание равенства (25). Если величины (37) получены в действительности как средние арифметические из равноточных измерений, числа которых известны, то эти числа и будут весами. Но бывают случаи, когда вычислителю приходится оценивать веса отдельных измерений по внутреннему убеждению, подобно баллам.

В дальнейших рассуждениях мы будем считать веса числами известными; что касается нормальной меры точности, то ниже мы рассмотрим способ ее вычисления, но в ближайших параграфах будем ее считать величиною данною.

§ 8. Наивероятнейший вывод из системы неравноточных измерений.

Сохраним обозначения предшествующего параграфа. Величина x измеренного предмета нам неизвестна, остальные величины мы считаем данными и, следовательно, известными.

Сделаем гипотезу, что величина измеренного предмета равна x . Тогда ошибки полученных результатов будут:

$$x - x_1 = \varepsilon_1, \quad x - x_2 = \varepsilon_2, \quad x - x_3 = \varepsilon_3, \dots, \quad x - x_\mu = \varepsilon_\mu;$$

вероятности этих ошибок по формуле Гаусса будут:

$$\frac{h_1}{\sqrt{\pi}} e^{-h_1^2 \varepsilon_1^2} d\varepsilon_1, \quad \frac{h_2}{\sqrt{\pi}} e^{-h_2^2 \varepsilon_2^2} d\varepsilon_2, \quad \frac{h_3}{\sqrt{\pi}} e^{-h_3^2 \varepsilon_3^2} d\varepsilon_3, \dots, \quad \frac{h_\mu}{\sqrt{\pi}} e^{-h_\mu^2 \varepsilon_\mu^2} d\varepsilon_\mu;$$

вероятность системы ошибок: $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_\mu$ или, что то же при сделанной гипотезе, вероятность полученной системы результатов измерений, выразится произведением написанных вероятностей. На основании равенств (41) это произведение напишется так:

$$\frac{h^\mu \sqrt{g_1 g_2 g_3 \dots g_\mu}}{\pi^{\frac{\mu}{2}}} e^{-h^2 \sum_{i=1}^{\mu} g_i (x - x_i)^2} d\varepsilon_1 d\varepsilon_2 d\varepsilon_3 \dots d\varepsilon_\mu \dots \quad (45)$$

Такова вероятность полученной в действительности системы результатов измерений при сделанной гипотезе относительно вели-

чины x . Вероятность самой гипотезы на основании теоремы Байеса пропорциональна величине (45). Включив во множитель пропорциональности все множители, не зависящие от x , и умножив и разделив на величину dx , тоже от x не зависящую, получим:

$$C \cdot h^k e^{-h^2 \sum g_i (x - x_i)^2} dx, \dots \dots \dots (46)$$

Такова вероятность на основании полученных результатов измерений, что величина измеренного предмета равна x .

Наивероятнейшее значение x найдется, когда мы приравняем нулю логарифмическую производную величины (46):

$$\sum g_i (x - x_i) = 0, \dots \dots \dots (47)$$

откуда:

$$x = \frac{\sum g_i x_i}{\sum g_i}.$$

Мы видим, что наивероятнейшее значение x есть уже полученная выше величина:

$$\xi = \frac{g_1 x_1 + g_2 x_2 + \dots + g_k x_k}{g_1 + g_2 + \dots + g_k} = \frac{\sum g_i x_i}{G}, \dots \dots \dots (42)$$

где:

$$G = g_1 + g_2 + \dots + g_k, \dots \dots \dots (43)$$

Равенство (47) принимает вид:

$$\sum g_i (\xi - x_i) = 0, \dots \dots \dots (48)$$

Величина ξ называется среднею из взвешенных величин: $x_1, x_2, x_3, \dots, x_k$.

§ 9. Мера точности результата, выведенного из системы неравноточных измерений.

Возьмем формулы предшествующего параграфа и будем рассуждать так же, как мы это делали в параграфе 6.

Пологаем:

$$x = \xi + \delta;$$

тогда δ выразит ошибку величины ξ , если истинная величина измеренного предмета равна x . Внося это выражение в формулу (46), найдем вероятность ошибки δ при найденном среднем ξ .

Формула (46) примет вид:

$$Ch^k e^{-h^2 \sum g_i (\xi - x_i)^2} \cdot e^{-h^2 \sum g_i (\xi - x_i) \delta - h^2 \sum g_i \delta^2} d\delta,$$

или, на основании равенств (48) и (43):

$$Ch^{\mu} e^{-h^2 \sum g_i (\xi - x_i)^2} e^{-h^2 G \delta^2} d\delta. \quad (49)$$

Для вычисления постоянной C припомним, что по той же причине, как и в параграфе 6, интеграл выражения (49) в пределах от $-\infty$ до $+\infty$ должен равняться 1:

$$C \cdot h^{\mu} e^{-h^2 \sum g_i (\xi - x_i)^2} \int_{-\infty}^{+\infty} e^{-h^2 G \delta^2} d\delta = 1.$$

Выполнив подстановку:

$$h \sqrt{G} \delta = t,$$

находим:

$$\frac{Ch^{\mu-1}}{\sqrt{G}} e^{-h^2 \sum g_i (\xi - x_i)^2} \int_{-\infty}^{+\infty} e^{-t^2} dt = 1,$$

откуда:

$$\frac{C \cdot h^{\mu-1} \sqrt{\pi}}{\sqrt{G}} e^{-h^2 \sum g_i (\xi - x_i)^2} = 1.$$

Разделив на эту величину, равную 1, величину (49), находим

$$\frac{h \sqrt{G}}{\sqrt{\pi}} e^{-h^2 G \delta^2} d\delta. \quad (50)$$

Таково выражение вероятности ошибки δ при средней величине ξ .

Она показывает, что средняя из взвешенных результатов неравноточных измерений имеет меру точности:

$$H = h \sqrt{G}. \quad (44)$$

Этот результат был уже заранее намечен в конце параграфа 7.

§ 10. Вычисление величины нормальной меры точности.

Пусть нам неизвестны обе величины: h и x . Рассуждая так же, как и в параграфе 5, делаем двойную гипотезу: относительно h и относительно x . Эта вероятность выразится той же формулой (46), как и в параграфе 8, с той только разницей, что величина постоянного C будет другая. Если мы хотим найти вероятность простой гипотезы относительно одного h , то должны взять сумму выражений функции (46) при всевозможных значениях x .

По той же причине, как и в параграфе 5, эта сумма выразится интегралом:

$$C \cdot h^{\mu} \int_{-\infty}^{+\infty} e^{-h^2 \sum g_i (x - x_i)^2} dx.$$

Преобразуем переменное подстановкою:

$$x = \xi + \delta.$$

После вычислений таких же как в предшествующем параграфе найдем:

$$\begin{aligned} Ch^{\mu} e^{-h^2 \sum g_i (\xi - x_i)^2} \int_{-\infty}^{+\infty} e^{-h^2 G \delta^2} d\delta &= \frac{Ch^{\mu-1}}{\sqrt{G}} e^{-h^2 \sum g_i (\xi - x_i)^2} \int_{-\infty}^{+\infty} e^{-t^2} dt = \\ &= \frac{C\sqrt{\pi} \cdot h^{\mu-1}}{\sqrt{G}} e^{-h^2 \sum g_i (\xi - x_i)^2}. \end{aligned}$$

Такова вероятность гипотезы относительно h .

Чтобы найти наимвероятнейшее значение h , мы приравняем нулю логарифмическую производную последней формулы по h :

$$\frac{\mu-1}{h} = 2h \sum g_i (x - x_i)^2,$$

откуда:

$$h = \frac{1}{\sqrt{2 \frac{\sum g_i (\xi - x_i)^2}{\mu-1}}} \dots \dots \dots (51)$$

Положив:

$$\sigma = \sqrt{\frac{\sum g_i (\xi - x_i)^2}{\mu-1}}, \dots \dots \dots (52)$$

получим обычную формулу:

$$h = \frac{1}{\sigma \sqrt{2}},$$

где σ , выражаемое формулою (52), можно назвать нормальной величиной средней квадратической ошибки.

§ 11. Пример вычислений в случае измерений неравноточных *).

Угол был измерен 14 раз, при чем каждое из этих измерений было получено как среднее арифметическое из различного числа непосредственных равноточных измерений. Следовательно, эти числа суть веса данных измерений. Данные величины и вычисления по формулам предшествующих параграфов приведены в следующей таблице.

*) См. Meyer. Vorlesungen über Wahrscheinlichkeitsrechnung. Стр. 274.

№	x_i	g_i	$g_i x_i$	$\xi - x_i$	$g_i (\xi - x_i)$	$(\xi - x_i)^2$	$g_i (\xi - x_i)^2$
1	17° 56' 45".00	5	225.00	- 5.22	- 26.10	27.248	136.94
2	31.25	4	125.00	+ 8.53	+ 34.12	72.761	291.04
3	42.50	5	212.50	- 2.72	- 13.60	7.398	36.99
4	45.00	3	135.00	- 5.22	- 15.66	27.248	81.74
5	37.50	3	112.50	+ 2.28	+ 6.84	5.198	15.59
6	38.33	3	115.00	+ 1.45	+ 4.35	2.103	6.31
7	27.50	3	82.50	+ 12.28	+ 36.84	150.798	452.39
8	43.33	3	130.00	- 3.55	- 10.65	12.603	37.81
9	40.63	4	162.50	- 0.85	- 3.40	0.723	2.89
10	36.25	2	72.50	+ 3.53	+ 7.06	12.461	24.92
11	42.50	3	127.50	- 2.72	- 8.10	7.398	22.19
12	39.17	3	117.50	+ 0.61	+ 1.83	0.372	1.12
13	45.00	2	90.00	- 5.22	- 10.44	27.248	54.49
14	40.83	3	122.50	- 1.05	- 3.15	1.103	3.31
		46	1830.00			+ 91.04 - 91.16	1167.03
			$\frac{1830}{46} = 39".78$			- 0.12	

$$x = 17^\circ 56' 39".78, \sigma = \sqrt{\frac{1167.03}{13}} = 9".475, h = 0.075 \quad H = 0.506.$$

Вероятная ошибка величины $x: 0".94$.

$$x = 17^\circ 56' 39".78 \pm 0".94.$$

ГЛАВА ВТОРАЯ

О функциях величин, полученных измерением.

§ 1. Постановка вопроса.

Пусть дана определенная функция одного или нескольких аргументов:

$$w = f(x, y, z, \dots, u), \dots \dots \dots (1)$$

при чем аргументы суть величины между собою независимые.

Для определения величин этих аргументов каждый из них был по несколько раз измерен, так что нам известны наивероятнейшие их значения и их меры точности. Спрашивается, какова при таких условиях наивероятнейшая величина функции (1) и ее мера точности.

Мы подойдем к решению общей задачи постепенно, начиная с простейших случаев.

§ 2. Вычисление величины функции: $w = ax + b$.

Пусть имеется функция

$$w = ax + b, \dots \dots \dots (2)$$

где a и b известные числа. Для величины x найдено ее наивероятнейшее значение x_0 с мерою точности h .

Если x_0 содержит в себе ошибку ε , то положив w равным:

$$w_0 = ax_0 + b, \dots \dots \dots (3)$$

мы сделаем ошибку

$$\eta = a\varepsilon. \dots \dots \dots (4)$$

Следовательно, вероятность ошибки η при величине w_0 равна вероятности ошибки ε при величине x_0 .

На основании формулы Гаусса эта вероятность выразится так:

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon^2} d\varepsilon.$$

Введя в эту формулу вместо ϵ ее выражение через η из равенства (4), находим:

$$\frac{\left(\frac{h}{a}\right)}{\sqrt{\pi}} e^{-\left(\frac{h}{a}\right)^2 \eta^2} d\eta. \quad (5)$$

Эта формула показывает, что:

1) Наивероятнейшая величина η есть 0, а потому наивероятнейшая величина w есть w_0 .

2) Мера точности h' величины w_0 выражается формулой:

$$h' = \frac{h}{a}. \quad (6)$$

Если мы обозначим буквами σ и σ' средние квадратические ошибки величин x_0 и w_0 , то будем иметь:

$$\sigma = \frac{1}{h\sqrt{2}}, \quad \sigma' = \frac{1}{h'\sqrt{2}},$$

откуда:

$$\sigma' = a\sigma. \quad (7)$$

Так как вероятные и средние ошибки разнятся от средних квадратических только постоянными множителями, то будем иметь:

$$r' = ar, \quad d' = ad. \quad (8)$$

§ 3. Вычисление величины функции: $w = x + y + z + \dots + n$.

Пусть имеется функция:

$$w = x + y, \quad (9)$$

где согласно условию, выраженному в § 1, величины x и y между собою независимы.

Наивероятнейшее значение x и y обозначим буквами x_0, y_0 , меры точности их h_1 и h_2 .

Если мы обозначим буквами ϵ_1 и ϵ_2 ошибки при x_1 и x_2 , то положив w равным

$$w_0 = x_0 + y_0, \quad (10)$$

сделаем ошибку, равную

$$\eta = \epsilon_1 + \epsilon_2. \quad (11)$$

Вероятности ошибок ϵ_1 и ϵ_2 выражаются формулами:

$$\frac{h_1}{\sqrt{\pi}} e^{-h_1^2 \epsilon_1^2} d\epsilon_1 \text{ и } \frac{h_2}{\sqrt{\pi}} e^{-h_2^2 \epsilon_2^2} d\epsilon_2.$$

Так как величины x и y между собою независимы, то по теореме умножения вероятностей в случае независимых между собою событий можем сказать, что вероятность системы ошибок $\varepsilon_1, \varepsilon_2$ выразится формулою:

$$\frac{h_1 h_2}{\pi} e^{-h_1^2 \varepsilon_1^2 - h_2^2 \varepsilon_2^2} d\varepsilon_1 d\varepsilon_2 \dots \dots \dots (12)$$

Если ε_1 и ε_2 имеют значения, удовлетворяющие равенству (11), то формула (12) выразит вероятность одной из статочностей, когда ошибка величины w равна η . Чтобы найти полную вероятность ошибки η при w_0 надо дать ε_1 и ε_2 в формуле (12) последовательно все значения, удовлетворяющие уравнению:

$$\eta = \varepsilon_1 + \varepsilon_2 \dots \dots \dots (11)$$

при данном η , и все полученные результаты сложить.

На основании равенства (11) выражение (12) примет вид:

$$\frac{h_1 h_2}{\pi} e^{-h_1^2 \varepsilon_1^2 - h_2^2 (\eta - \varepsilon_1)^2} d\varepsilon_1 d\eta.$$

В этом выражении мы должны дать ε_1 всевозможные значения от $-\infty$ до $+\infty$ и взять сумму полученных результатов. В пределе, когда $d\varepsilon_1$ обратится в 0, эта сумма выразится интегралом:

$$\frac{h_1 h_2}{\pi} d\eta \int_{-\infty}^{+\infty} e^{-h_1^2 \varepsilon_1^2 - h_2^2 (\eta - \varepsilon_1)^2} d\varepsilon_1 \dots \dots \dots (13)$$

Такова вероятность ошибки η при величине w_0 . Остается выполнить вычисление интеграла.

Выражение в показателе приводится к виду:

$$-(h_1^2 + h_2^2) \varepsilon_1^2 + 2h_2^2 \eta \varepsilon_1 - h_2^2 \eta^2,$$

или:

$$-(h_1^2 + h_2^2) \left\{ \varepsilon_1 - \frac{h_2^2 \eta}{h_1^2 + h_2^2} \right\}^2 - \frac{h_1^2 h_2^2}{h_1^2 + h_2^2} \eta^2.$$

Внеся это выражение вместо показателя степени в подынтегральной функции формулы (13) и выполнив подстановку:

$$\varepsilon_1 = \frac{h_2^2 \eta}{h_1^2 + h_2^2} + u,$$

приводим выражение (13) к такому виду:

$$\frac{h_1 h_2}{\pi} e^{-\frac{h_1^2 h_2^2}{h_1^2 + h_2^2} \eta^2} d\eta \int_{-\infty}^{+\infty} e^{-(h_1^2 + h_2^2) u^2} du.$$

Совершив еще раз преобразование переменного интегрирования подстановкой:

$$\sqrt{h_1^2 + h_2^2} u = t,$$

находим:

$$\frac{h_1 h_2}{\pi \sqrt{h_1^2 + h_2^2}} e^{-\frac{h_1^2 h_2^2}{h_1^2 + h_2^2} \eta^2} d\eta \int_{-\infty}^{+\infty} e^{-t^2} dt = \frac{\left(\frac{h_1 h_2}{\sqrt{h_1^2 + h_2^2}} \right)}{\sqrt{\pi}} e^{-\frac{h_1^2 h_2^2}{h_1^2 + h_2^2} \eta^2} d\eta.$$

Такова вероятность ошибки η при величине:

$$w_0 = x_0 + y_0 \dots \dots \dots (10)$$

Полученная формула показывает, что:

1) Наивероятнейшая величина η есть 0, а потому наивероятнейшая величина w есть w_0 .

2) Мера точности h' величины w_0 выражается формулой:

$$h' = \frac{h_1 h_2}{\sqrt{h_1^2 + h_2^2}} \dots \dots \dots (14)$$

Обозначив средние квадратические ошибки величин x_0, y_0, w_0 соответственно буквами $\sigma_1, \sigma_2, \sigma'$, найдем:

$$\sigma_1 = \frac{1}{h_1 \sqrt{2}}, \sigma_2 = \frac{1}{h_2 \sqrt{2}}, \sigma' = \frac{1}{h' \sqrt{2}} *).$$

Внеся эти величины в формулу (14), найдем:

$$\sigma' = \sqrt{\sigma_1^2 + \sigma_2^2} \dots \dots \dots (15)$$

Эта формула показывает, что при вычислении средней квадратической ошибки суммы $w_0 = x_0 + y_0$ можно произвести сложение средних квадратических ошибок геометрически: σ' есть гипотенуза прямоугольного треугольника, где σ_1 и σ_2 суть катеты.

Так как вероятные и средние ошибки пропорциональны средним квадратическим, то и для них правило геометрического сложения остается в силе.

Найдя решение вопроса о вероятнейшем значении и средней квадратической ошибки для суммы двух величин, можем тотчас же распространить полученный результат на сумму трех величин. В самом деле, сумму: $x + y + z$ можно представить в виде: $(x + y) + z$.

Если наивероятнейшие значения x, y, z суть x_0, y_0, z_0 , а средние квадратические ошибки этих величин $\sigma_1, \sigma_2, \sigma_3$, то по только что доказанному, наивероятнейшее значение слагаемого $(x + y)$ будет $(x_0 + y_0)$, а средняя квадратическая ошибка этой величины $\sqrt{\sigma_1^2 + \sigma_2^2}$.

*) См. гл. I форм. (24).

Отсюда, на основании того же предложения наиболее вероятное значение суммы

$$(x + y) + z$$

будет: $(x_0 + y_0) + z_0 = x_0 + y_0 + z_0$, а средняя квадратическая ошибка этой величины будет:

$$\sqrt{(\sigma_1^2 + \sigma_2^2) + \sigma_3^2} = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}.$$

Продолжая эти рассуждения, увидим справедливость того же предложения для суммы любого числа слагаемых; если

$$w = x + y + z + \dots + u, \quad (16)$$

при чем наиболее вероятные значения $w, x, y, z, \dots, u$ суть:

$$w_0, x_0, y_0, z_0, \dots, u_0,$$

а средние квадратические ошибки:

$$\sigma', \sigma_1, \sigma_2, \sigma_3, \dots, \sigma_k,$$

то:

$$w_0 = x_0 + y_0 + z_0 + \dots + u_0, \quad (17)$$

$$\sigma' = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \dots + \sigma_k^2}. \quad (18)$$

§ 4. Вычисление величины: $w = ax + by + cz + \dots + fu + g$.

Пусть дано выражение:

$$w = ax + by + cz + \dots + fu + g, \quad (19)$$

при чем вероятнейшие значения аргументов $x, y, z, \dots, u$ суть $x_0, y_0, z_0, \dots, u_0$, средние квадратические ошибки этих величин $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_k$.

Мы найдем наиболее вероятное значение функции w и ее среднюю квадратическую ошибку σ' , пользуясь теоремами двух предшествующих параграфов.

Выражение (19) представим в таком виде:

$$w = (ax) + (by) + (cz) + \dots + (fu + g).$$

На основании теоремы параграфа 2 скажем, что наиболее вероятные значения слагаемых этой суммы суть:

$$ax_0, by_0, cz_0, \dots, fu_0 + g,$$

а средние квадратические ошибки этих величин:

$$a\sigma_1, b\sigma_2, c\sigma_3, \dots, f\sigma_k.$$

На основании теоремы параграфа 3 найдем, что наиболее вероятное значение w есть:

$$w_0 = ax_0 + by_0 + cz_0 + \dots + fu_0 + g, \quad (20)$$

а средняя квадратическая ошибка σ' величины w_0 :

$$\sigma' = \sqrt{a^2 \sigma_1^2 + b^2 \sigma_2^2 + c^2 \sigma_3^2 + \dots + f^2 \sigma_k^2} \dots (21)$$

§ 5. Вычисление величины: $w = f(x, y, z, \dots, u)$.

Пусть дана функция:

$$w = f(x, y, z, \dots, u) \dots (22)$$

при чем наимвероятнейшие значения аргументов $x, y, z, \dots, u$ суть $x_0, y_0, z_0, \dots, u_0$, средние квадратические ошибки этих величин: $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_k$. Спрашивается, каково наимвероятнейшее значение w_0 величины w и какова средняя квадратическая ошибка σ' величины w_0 .

Положим:

$$x = x_0 + \varepsilon_1, y = y_0 + \varepsilon_2, z = z_0 + \varepsilon_3, \dots, u = u_0 + \varepsilon_k \dots (23)$$

На основании теоремы параграфа 2 можем сказать, что наимвероятнейшие значения $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_k$ суть нули, а средние квадратические ошибки этих величин $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_k$. Если эти величины малы, то $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_k$ можно считать малыми. Внося выражение (23) в формулу (22), разложим ее в ряд по формуле Тейлора и ограничимся членами 1-го порядка. Получим:

$$w = f(x_0, y_0, z_0, \dots, u_0) + f_1'' \cdot \varepsilon_1 + f_2'' \cdot \varepsilon_2 + f_3'' \cdot \varepsilon_3 + \dots + f_k'' \cdot \varepsilon_k, \dots (24)$$

где $f_1'', f_2'', f_3'', \dots, f_k''$ суть значения частных производных:

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}, \dots, \frac{\partial f}{\partial u}$$

при значениях $x, y, z, \dots, u$ соответственно равных: $x_0, y_0, z_0, \dots, u_0$.

Так как выражение (24) такого же типа, как (19), то можно приложить к нему теорему предшествующего параграфа.

Результат получится следующий:

наимвероятнейшее значение w величины (24) таково:

$$w_0 = f(x_0, y_0, z_0, u_0) \dots (25)$$

средняя квадратическая ошибка σ' этой величины:

$$\sigma' = \sqrt{(f_1'')^2 \sigma_1^2 + (f_2'')^2 \sigma_2^2 + (f_3'')^2 \sigma_3^2 + \dots + (f_k'')^2 \sigma_k^2} \dots (26)$$

Пример: $w = \frac{a}{x}$.

На основании формул (25) и (26) имеем:

$$w_0 = \frac{a}{x_0}, \sigma' = \frac{a}{x_0^2} \sigma_1.$$

ГЛАВА ТРЕТЬЯ

Способ наименьших квадратов.

§ 1. Задача способа наименьших квадратов. Начальные уравнения.

В предшествующих параграфах мы говорили о тех случаях, когда интересующие нас величины или непосредственно могут быть измерены, или представляют собою функции непосредственно измеряемых величин.

Но встречаются случаи, когда искомые величины не могут быть измерены непосредственно и не могут быть представлены, как явные функции непосредственно измеряемых величин. В таких случаях измеряются величины функций, зависящих от искомым величин, и получается система уравнений, в которые искомые величины входят как неизвестные.

Если число уравнений меньше числа неизвестных, то задача остается неопределенною. Если число уравнений равно числу неизвестных, то мы имеем дело с обыкновенной задачей алгебры. В теории вероятностей рассматривается тот случай, когда число уравнений больше, и даже значительно больше числа неизвестных.

Обозначим неизвестные величины, значение которых мы хотим найти, буквами:

$$x, y, z, \dots, u, v; \dots \dots \dots (1)$$

число их пусть будет n .

Пусть для определения величин этих неизвестных непосредственно измерено m функций этих неизвестных и получились уравнения:

$$\left. \begin{aligned} f_1(x, y, z, \dots, u, v) &= \varphi_1, \\ f_2(x, y, z, \dots, u, v) &= \varphi_2, \\ f_3(x, y, z, \dots, u, v) &= \varphi_3, \\ &\dots \dots \dots \\ f_m(x, y, z, \dots, u, v) &= \varphi_m. \end{aligned} \right\} \dots \dots \dots (2)$$

Функции, стоящие в левых частях этих уравнений суть данные нам выражения, содержащие аргументы (1); коэффициенты этих

Вероятность системы ошибок $\epsilon_1, \epsilon_2, \epsilon_3, \dots, \epsilon_\mu$ равна произведению величин (9), так как результаты измерений (3) между собою независимы.

Если определитель нормальной системы уравнений (13):

$$D = \begin{vmatrix} [gaa] & [gab] & \dots & [gak] & [gal] \\ [gba] & [gbb] & \dots & [gbk] & [gbl] \\ \dots & \dots & \dots & \dots & \dots \\ [gka] & [gkb] & \dots & [gkk] & [gkl] \\ [gla] & [glb] & \dots & [glk] & [gll] \end{vmatrix} \quad (11)$$

не равен 0, то уравнения (13) дадут вполне определенные конечные величины для наиболее вероятных значений аргументов $x, y, z, \dots, u, v$ *).

Подставив обозначение (12) в формулу (11), приведем ее к виду:

$$Ch^2 e^{-2h^2\theta} dx dy dz \dots du dv \quad (15)$$

Это — выражение вероятности определенной системы значений $x, y, z, \dots, u, v$ на основании сделанных измерений величин (3).

Нормальная система уравнений (13) в коротком виде напишется так:

$$\frac{\partial \theta}{\partial x} = 0, \frac{\partial \theta}{\partial y} = 0, \frac{\partial \theta}{\partial z} = 0, \dots, \frac{\partial \theta}{\partial u} = 0, \frac{\partial \theta}{\partial v} = 0. \quad (16)$$

Из формулы (12) непосредственно видно, что величина 2θ существенно положительна: нулем она быть не может потому, что обращение ее в нуль означало бы, что существует система значений величин $x, y, z, \dots, u, v$, удовлетворяющих в точности начальным уравнениям (5); а мы знаем, что между этими уравнениями существуют противоречия, вследствие тех ошибок, которые входят в правые части.

Система нормальных уравнений (16) определяет те значения переменных, при которых функция 2θ имеет наименьшее значение.

Если определитель (14) не равен нулю, то такая система существует, она единственная, вполне определенная. Соответствующее ей значение функции 2θ служит мерой противоречий между уравнениями (5).

§ 3. Нормальная система уравнений, преобразованная по способу Гаусса.

Для нахождения весов тех результатов, которые даются нормальной системой уравнений, Гаусс предложил особый способ пред-

*) Определитель (14) обратится в нуль в исключительных случаях, как напр., если величины x и y в начальной системе (5) встречаются только в определенной комбинации $ax + by$. В этом случае нормальная система даст наиболее вероятное значение величинам $ax + by$, но x и y в отдельности останутся неопределенными.

варительного ее преобразования. Нормальная система в такой преобразованной форме называется Гауссовой нормальной системой.

В числе уравнений (13) есть, по крайней мере, одно, содержащее в себе x ; иначе уравнения (13) не имели бы определенных решений. Возьмем одно из таких уравнений и без всяких изменений переписываем в Гауссову систему.

Пусть это уравнение будет:

$$A_1 x + B_1 y + C_1 z + \dots + K_1 u + L_1 v + M_1 = 0.$$

Решаем это уравнение относительно x , подставляем полученное для x выражение во все прочие уравнения нормальной системы (13), делаем приведение, но не умножаем и не делим уравнений ни на какие числа.

В числе полученных уравнений непременно найдется, по крайней мере, одно, содержащее в себе y ; иначе система (13) не могла бы давать определенных решений. Пусть это уравнение, содержащее в себе y и уже освобожденное от x , будет:

$$B_2 y + C_2 z + \dots + K_2 u + L_2 v + M_2 = 0.$$

Переписываем его без всяких изменений вторым уравнением в Гауссову систему.

Решаем это уравнение относительно y и подставляем полученное для y выражение во все еще не взятые уравнения нормальной системы, уже освобожденные от x , делаем приведения, но не умножаем и не делим полученных уравнений ни на какие числа.

Между полученными уравнениями найдется, по крайней мере, одно, содержащее в себе z ; иначе система (13) не могла бы быть определенной. Пусть это уравнение, содержащее в себе z , но уже освобожденное от x и y , будет:

$$C_3 z + \dots + K_3 u + L_3 v + M_3 = 0.$$

Переписываем его без всяких изменений третьим уравнением в Гауссову систему.

Поступая дальше по тому же плану, дойдем до последних уравнений, которые будут иметь вид:

$$K_{n-1} u + L_{n-1} v + M_{n-1} = 0,$$

$$L_n v + M_n = 0.$$

Положим

$$\omega_1 = \sqrt{\alpha}x + \frac{\beta}{\sqrt{\alpha}}, 2\theta_1 = \gamma - \frac{\beta^2}{\alpha}, \dots \dots \dots (19),$$

где ω_1 многочлен первой степени относительно $x, y, z, \dots, u, v$ с коэффициентом при x существенно положительным, а $2\theta_1$ многочлен 2-ой степени относительно $y, z, \dots, u, v$.

Получим выражение:

$$2\theta = \omega_1^2 + 2\theta_1 \dots \dots \dots (20)$$

Положим в формулах (19) и (20) переменные $y, z, \dots, u, v$ равными каким угодно конечным числам, а x положим равным:

$$x = -\frac{\beta}{\alpha};$$

при таких значениях переменных первое из уравнений (19) дает: $\omega_1 = 0$, а уравнение (20) примет вид:

$$2\theta = 2\theta_1,$$

при чем, как мы сказали, величины $y, z, \dots, u, v$ суть какие угодно конечные числа. Отсюда следует, что $2\theta_1$, так же как 2θ , ни при каких значениях входящих в нее аргументов не может быть ни отрицательным, ни нулем.

В таком случае к выражению $2\theta_1$ можно приложить такие же рассуждения, какие мы приложили к выражению 2θ ; многочлен 2-ой степени $2\theta_1$ с переменными $y, z, \dots, u, v$ можно представить в таком виде:

$$2\theta_1 = \omega_2^2 + 2\theta_2, \dots \dots \dots (21)$$

где ω_2 многочлен первой степени относительно $y, z, \dots, u, v$, при чем коэффициент при y положителен и отличен от нуля, а функция $2\theta_2$ квадратный многочлен с переменными $z, \dots, u, v$, существенно положительный и отличный от нуля при всех значениях входящих в него переменных.

Продолжая такие же рассуждения, получим ряд тождеств такого же вида, как (20) и (21):

$$2\theta_2 = \omega_3^2 + 2\theta_3,$$

$$\dots \dots \dots 2\theta_{r-1} = \omega_r^2 + 2\theta_r,$$

где смысл обозначений такого же характера, как в равенствах (20) и (21); в последнем из написанных равенств ω_{r-1} многочлен 1-ой степени с одним переменным v , имеющий коэффициент при v существенно положительный, а $2\theta_r$ постоянное число, которое не может быть ни отрицательным, ни нулем. Обозначим квадратный корень из

нениях (24) обратятся в нуль все первые члены, а остальные члены, как не содержащие в себе x , останутся без изменения.

Второе уравнение системы (24) после сделанной подстановки, примет вид:

$$\omega_2 \frac{\partial \omega_2}{\partial y} = 0;$$

оно переписывается в Гауссову систему без изменения.

Так как $\frac{\partial \omega_2}{\partial y}$ отлично от нуля, то из этого уравнения следует: $\omega_2 = 0$.

Совершая Гауссово преобразование, выражаем отсюда y через следующие переменные и подставляем в дальнейшие уравнения (24), при этом обратятся в нули вторые члены, а остальные, как не содержащие в себе y , не изменятся.

Продолжая такие рассуждения, найдем Гауссову систему в такой форме:

$$\omega_1 \frac{\partial \omega_1}{\partial x} = 0, \omega_2 \frac{\partial \omega_2}{\partial y} = 0, \omega_3 \frac{\partial \omega_3}{\partial z} = 0, \dots, \omega_v \frac{\partial \omega_v}{\partial v} = 0. \dots (25)$$

Так как при Гауссовом преобразовании не делается ни умножения всех членов уравнений, ни деления их на какие-либо множители, то уравнения (24) должны быть в точности одинаковыми с уравнениями (17).

Последнее из уравнений (17):

$$L_v v + M_v = 0 \dots (17')$$

должно быть одинаково с последним уравнением (25):

$$\omega_v \frac{\partial \omega_v}{\partial v} = 0.$$

Внеся в это уравнение вместо ω_v его выражение из равенств (23), находим:

$$\lambda_v (\lambda_v v + \mu_v) = 0.$$

Поэтому можем написать:

$$L_v = \lambda_v' \dots (26)$$

Это равенство при дальнейших вычислениях будет иметь большое значение.

§ 4. Вес результатов, даваемых способом наименьших квадратов.

Вероятность (15) определенной системы значений $x, y, z, \dots, u, v$ на основании формулы (22) примет вид:

$$Ch^2 e^{-h^2 (\omega_0^2 + \omega_1^2 + \omega_2^2 + \omega_3^2 + \dots + \omega_v^2)} dx dy dz \dots dv. \dots (27)$$

Каждый из интегралов этой формулы равен величине:

$$\int_{-\infty}^{+\infty} e^{-h^2 \omega^2} d\omega,$$

или, после подстановки $h\omega = t$:

$$\frac{1}{h} \int_{-\infty}^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{h}.$$

Поэтому выражение (28) примет такой вид:

$$\frac{Ch^{v-v+1} \pi^{\frac{v-1}{2}}}{\alpha_1 \beta_2 \gamma_3 \dots x_{v-1}} e^{-h^2 \omega_0^2} \cdot e^{-h^2 \omega_v^2} dv.$$

Обозначив буквой C' постоянную относительно v и h величину:

$$C' = \frac{Ch^{v-v+1} \pi^{\frac{v-1}{2}}}{\alpha_1 \beta_2 \gamma_3 \dots x_{v-1}} e^{-h^2 \omega_0^2}$$

напишем выражение (28) в более простой форме:

$$C' e^{-h^2 \omega_v^2} dv. \quad (30)$$

Таково выражение вероятности, что v имеет данное значение.

Припомним, что последнее уравнение Гауссовой системы пишется так:

$$\omega_v \frac{\partial \omega_v}{\partial v} = 0,$$

или:

$$\lambda_v (\lambda_v v + \mu_v) = 0,$$

как мы видели в конце предшествующего параграфа.

Обозначая буквою v_0 наимвероятнейшее значение v , определяемое нормальной системой, находим, что v_0 определяется из уравнения:

$$\lambda_v (\lambda_v v_0 + \mu_v) = 0,$$

откуда:

$$\mu_v = -\lambda_v v_0.$$

Последнее из равенств (23) примет вид:

$$\omega_v = \lambda_v (v - v_0).$$

Внеся это выражение функции ω_v в формулу (30), находим:

$$C' e^{-h^2 \lambda_v^2 (v - v_0)^2} dv,$$

или, приняв во внимание формулу (26):

$$C' e^{-h^2 L_v (v - v_0)^2} dv. \quad (31)$$

Такова вероятность, что v имеет данное значение. В этой формуле C' постоянное относительно v число; L_v коэффициент в последнем уравнении (17) Гауссовой системы нормальных уравнений, v_0 наимвероятнейшее значение, даваемое этими уравнениями для переменной v .

Положим:

$$v - v_0 = \delta_v \dots \dots \dots (32)$$

Если v есть истинное значение этой переменной, то величина (32) есть разность между истинной величиной v и наимвероятнейшим ее значением v_0 . Это значит, что δ_v есть ошибка при наимвероятнейшем значении v_0 величины v .

Вставив выражение (32) в формулу (31), найдем:

$$C' e^{-h^2 L_v \delta_v^2} d\delta_v \dots \dots \dots (33)$$

Такова вероятность ошибки δ_v при величине v_0 .

Для нахождения величины постоянного C' припомним, что сумма значений (33) при всевозможных значениях v от $-\infty$ до $+\infty$ должна равняться 1.

Отсюда следует равенство:

$$C' \int_{-\infty}^{+\infty} e^{-h^2 L_v \delta_v^2} d\delta_v = 1,$$

или, после преобразования:

$$h \sqrt{L_v} \delta_v = t;$$

$$\frac{C'}{h \sqrt{L_v}} \int_{-\infty}^{+\infty} e^{-t^2} dt = 1,$$

откуда:

$$\frac{C' \sqrt{\pi}}{h \sqrt{L_v}} = 1,$$

Разделив величину (33) на это выражение, равное единице найдем:

$$\frac{h \sqrt{L_v}}{\sqrt{\pi}} e^{-h^2 L_v \delta_v^2} d\delta_v \dots \dots \dots (34)$$

Вводя обозначение

$$h \sqrt{L_v} = H_v \dots \dots \dots (35)$$

представим величину (34) в таком виде:

$$\frac{H_v}{\sqrt{\pi}} e^{-H_v^2 \delta_v^2} d\delta_v \dots \dots \dots (36)$$

При этом сложении коэффициенты при $x, y, z, \dots, u$ обратятся в нули, а коэффициент при v будет равен определителю D . Получится уравнение вида:

$$Dv + M'_v = 0 \dots \dots \dots (38)$$

Оно дает то же решение v_v , как и последнее уравнение Гауссовой системы:

$$L_v v + M_v = 0, \dots \dots \dots (17)$$

следовательно, их коэффициенты пропорциональны.

Чтобы найти множителя пропорциональности, припомним, что последнее уравнение Гауссовой системы получилось, когда мы выразили величины $x, y, z, \dots, u$ через v из первых $v-1$ уравнений системы (13) и вставили полученные выражения в последнее уравнение той же системы, не умножая и не деля его ни на какое число.

При составлении уравнения (38) мы умножили последнее уравнение системы (13) на минор D_{vv} . Ясно, что этот минор и есть множитель пропорциональности:

$$\frac{D}{L_v} = D_{vv},$$

откуда:

$$L_v = \frac{D}{D_{vv}},$$

или, на основании равенства (37):

$$G_v = \frac{D}{D_{vv}}, \dots \dots \dots (39)$$

Так как все неизвестные в нормальной системе совершенно равноправны, то формула (39) даст возможность по тому же правилу построить выражения для весов всех неизвестных:

$$G_1 = \frac{D}{D_{11}}, G_2 = \frac{D}{D_{22}}, G_3 = \frac{D}{D_{33}}, \dots, G_v = \frac{D}{D_{vv}} \dots \dots \dots (40)$$

Нормальную величину меры точности h мы считали величиной данной. Имея эту величину h и зная величины весов, мы находим меры точности, а следовательно, и ошибки: вероятные, средние и средние квадратические для всех величин $x, y, z, \dots, u, v$, даваемых нормальной системой уравнений.

Остается рассмотреть способ определения нормальной меры точности h .

§ 6. Нахождение величины нормальной меры точности h в способе наименьших квадратов.

В параграфе 4 мы представили формулу вероятности гипотезы, что величины $x, y, z, \dots, u, v$ имеют данные значения, в такой форме:

$$Ch^v e^{-h^2(\omega_0^2 + \omega_1^2 + \omega_2^2 + \omega_3^2 + \dots + \omega_v^2)} dx dy dz \dots dv. \dots \dots \dots (27)$$

По формуле преобразования переменных в кратных интегралах выражение (41) примет такой вид:

$$\frac{Ch^{\mu}}{\alpha_1 \beta_2 \gamma_3 \dots x_v} e^{-h^2 \omega_0^2} \times \\ \times \int_{-\infty}^{+\infty} e^{-h^2 \omega_1^2} d\omega_1 \times \int_{-\infty}^{+\infty} e^{-h^2 \omega_2^2} d\omega_2 \int_{-\infty}^{+\infty} e^{-h^2 \omega_3^2} d\omega_3 \dots \int_{-\infty}^{+\infty} e^{-h^2 \omega_v^2} d\omega_v \dots \quad (42)$$

Каждый из интегралов этой формулы можно представить в виде:

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

Поэтому выражение (42) напишется так:

$$\frac{Ch^{\mu - \frac{v}{2}} \pi^{\frac{v}{2}}}{\alpha_1 \beta_2 \gamma_3 \dots x_v} e^{-h^2 \omega_0^2} \dots \quad (43)$$

Такова вероятность того, что h имеет определенное, произвольно взятое значение. Множитель

$$\frac{\pi^{\frac{v}{2}}}{\alpha_1 \beta_2 \gamma_3 \dots x_v}$$

от h не зависит.

Чтобы найти наиболее вероятное значение h , приравняем нулю логарифмическую производную по h от функции (43):

$$\frac{\mu - \frac{v}{2}}{h} - 2h\omega_0^2 = 0,$$

откуда:

$$h = \frac{\sqrt{\mu - \frac{v}{2}}}{\sqrt{2} \omega_0} \quad (44)$$

Это наиболее вероятное значение h принимается за истинную величину нормальной меры точности в способе наименьших квадратов.

Уясним смысл этой формулы.

Рассматривая Гауссову систему нормальных уравнений в форме

$$\omega_1 \frac{\partial \omega_1}{\partial x} = 0, \omega_2 \frac{\partial \omega_2}{\partial y} = 0, \omega_3 \frac{\partial \omega_3}{\partial z} = 0, \dots, \omega_v \frac{\partial \omega_v}{\partial v} = 0, \dots \quad (25)$$

мы видим, что решения $x_0, y_0, z_0, \dots, v_0$ нормальной системы уравнений обращают в нуль величины $\omega_1, \omega_2, \omega_3, \dots, \omega_v$. Поэтому функция 2θ при подстановке вместо $x, y, z, \dots, v$ их наиболее вероятных величин $x_0, y_0, z_0, \dots, v_0$ делается равной ω_0^2 .

Следовательно, ω_0 есть квадратный корень из наименьшей величины 2θ .

Выше в конце параграфа 2 мы видели, что эта наименьшая величина 26 есть мера противоречий между уравнениями начальной системы (2).

Отсюда понятно, почему в формуле (44) величина ω_0 стоит в знаменателе: нормальная мера точности тем больше, чем меньше противоречия между уравнениями.

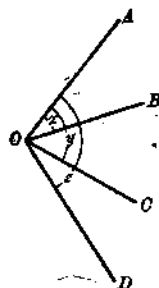
В числителе формулы (44) стоит $\mu - \nu$, что тоже понятно: чем больше избыток числа уравнений над числом неизвестных, тем больше точность полученного результата.

§ 7. Пример применения способа наименьших квадратов.

Для ясности и краткости изложения мы будем пользоваться нумерацией формул, как в теоретической части этой главы, отличая только формулы примера от общих формул чертой — над номером формулы.

Условие задачи*): на местности из точки O проведены 4 прямые и измерены углы:

$$\left. \begin{aligned} AOB = x &= 48^\circ 17' 1''{,}4 \text{ с весом } g_1 = 30, \\ AOC = y &= 96^\circ 52' 16''{,}8 \text{ " " } g_2 = 20, \\ AOD = z &= 152^\circ 54' 6''{,}8 \text{ " " } g_3 = 26, \\ BOC = y - x &= 48^\circ 35' 14''{,}3 \text{ " " } g_4 = 25, \\ BOD = z - x &= 104^\circ 37' 7''{,}8 \text{ " " } g_5 = 28, \\ COD = z - y &= 56^\circ 1' 48''{,}9 \text{ " " } g_6 = 44. \end{aligned} \right\} \dots \dots (2)$$



Эти равенства представляют систему 6 уравнений с 3 неизвестными; ищутся наимвероятнейшие значения x, y, z .

Преобразуем равенства (2) по формулам, подобным (6):

$$x - 48^\circ 17' 1''{,}4 = x', \quad y - 96^\circ 52' 16''{,}8 = y', \quad z - 152^\circ 54' 6''{,}8 = z'; \dots \dots (6)$$

получаем начальную систему уравнений:

$$\left. \begin{aligned} 1 \cdot x' + 0 \cdot y' + 0 \cdot z' &= 0, & -1 \cdot x' + 1 \cdot y' + 0 \cdot z' + 1''{,}1 &= 0, \\ 0 \cdot x' + 1 \cdot y' + 0 \cdot z' &= 0, & -1 \cdot x' + 0 \cdot y' + 1 \cdot z' - 2''{,}4 &= 0, \\ 0 \cdot x' + 0 \cdot y' + 1 \cdot z' &= 0, & 0 \cdot x' - 1 \cdot y' + 1 \cdot z' + 1''{,}1 &= 0, \end{aligned} \right\} \dots \dots (7)$$

Для вычисления коэффициентов нормальной системы уравнений вычисляем таблицу:

g	a	b	c	ga	gab	gac	gbb	gbc	gsc	$ga(m-\varphi)$	$gb(m-\varphi)$	$gc(m-\varphi)$
30	1	0	0	30	0	0	0	0	0	0	0	0
20	0	1	0	0	0	0	20	0	0	0	0	0
26	0	0	1	0	0	0	0	0	26	0	0	0
25	-1	1	0	25	-25	0	25	0	0	-27,5	27,5	0
28	-1	0	1	28	0	-28	0	0	28	+67,2	0	-67,2
44	0	-1	1	0	0	0	44	-44	44	0	-48,4	48,4
—	—	—	—	83	-25	-28	89	-44	98	39,7	-20,9	-18,8

*) См. Meyer. Vorlesungen über Wahrscheinlichkeitrechnung. Сrp. 303.

Нормальная система уравнений будет такова:

$$\left. \begin{aligned} 83 x' - 25 y' - 28 z' + 39,7 &= 0, \\ -25 x' + 89 y' - 44 z' - 20,9 &= 0, \\ -28 x' - 44 y' + 98 z' - 18,8 &= 0. \end{aligned} \right\} \dots \dots \dots (13)$$

Решая эти уравнения, находим:

$$x_0' = -0,338, \quad y_0' = 0,241, \quad z_0' = 0,203,$$

откуда, внося в формулы (6):

$$x_0 = 48^\circ 17' 1,062; \quad y_0 = 96^\circ 52' 17,041; \quad z_0 = 152^\circ 54' 7,003.$$

Таковы искомые наиболее вероятнейшие величины неизвестных x, y, z .

Вставляя найденные величины x_0', y_0', z_0' в левые части уравнений (7), находим величины противоречий:

$$\epsilon_1 = -0,338, \quad \epsilon_2 = 0,241, \quad \epsilon_3 = 0,203, \quad \epsilon_4 = 1,679, \quad \epsilon_5 = -1,859, \quad \epsilon_6 = 1,062.$$

Отсюда:

$$\omega_0^2 = 2 \theta_0 = g_1 \epsilon_1^2 + g_2 \epsilon_2^2 + g_3 \epsilon_3^2 + \dots + g_6 \epsilon_6^2 = 222,53.$$

По формуле (44) найдем:

$$h = \sqrt{\frac{6-3}{222,53 \cdot 2}} = \sqrt{\frac{3}{445,06}} = 0,0821 \dots \dots \dots (44)$$

Подставив коэффициенты уравнений (13) в формулы (40), найдем веса величин x_0, y_0, z_0 :

$$G_x = 54,614; \quad G_y = 50,425; \quad G_z = 54,803 \dots \dots \dots (40)$$

Найдя веса (40) и нормальную меру точности (44), можем вычислить меры точности H_x, H_y, H_z по формуле (3):

$$H_x = 0,607; \quad H_y = 0,583, \quad H_z = 0,608.$$

Отсюда средние квадратические отклонения по общей формуле:

$$\sigma = \frac{1}{h \sqrt{2}}:$$

$$\Sigma_x = 1,17; \quad \Sigma_y = 1,21; \quad \Sigma_z = 1,16.$$

ЧАСТЬ ТРЕТЬЯ.

ПРИЛОЖЕНИЕ ТЕОРИИ ВЕРОЯТНОСТЕЙ
К ИЗУЧЕНИЮ МАССОВЫХ ЯВЛЕНИЙ.

ГЛАВА ПЕРВАЯ.

О законах распределения.

§ 1. Понятие о массовых явлениях.

В вопросах обществоведения, антропологии, биологии, некоторых отделов физики (кинетическая теория) и многих других областей знания приходится иметь дело не с отдельным предметом или явлением, а с совокупностью предметов или явлений, сходных между собой в каком-либо отношении.

Для краткости такие предметы или явления будем называть объектами.

Двух совершенно тождественных объектов нигде во внешнем мире мы не встретим; все они разнятся друг от друга различными свойствами, или, как обыкновенно говорят, признаками. Но по некоторой особенности нашего ума мы отвлекаемся от большинства этих признаков и сосредоточиваем внимание на одном, иногда нескольких признаках, которые считаем в данном вопросе наиболее важными. Все объекты, имеющие эти признаки, мы считаем однородными и принимаем их за единицы при производстве счета объектов. Группу объектов, имеющих общий им признак, мы называем совокупностью, отдельные объекты, входящие в состав совокупности, называем членами ее; число членов совокупности называем ее объемом, а признак, общий всем членам совокупности, называем родовым признаком.

Кроме принадлежащего всем членам совокупности родового признака, эти члены имеют еще множество других признаков, вообще говоря, случайного характера.

В таких науках, как физика, механика, астрономия и др. приходится в большинстве случаев иметь дело с отдельным объектом, признаки которого или непосредственно нам даны, или могут быть непосредственно измерены. Но при решении вопросов, касающихся

совокупностей, где члены имеют признаки случайного характера, такие приемы не приложимы: здесь нужны методы особые, опирающиеся на теорию случайных явлений, т. е. на теорию вероятностей. Явления, происходящие в совокупностях однородных объектов, называются массовыми явлениями; поэтому та область приложений теории вероятностей, которую мы теперь рассматриваем, называется учением о массовых явлениях.

Наблюдения над массовыми явлениями были производимы почти бессознательно и в грубой форме еще в глубокой древности: они касались, главным образом, общества людей и производились преимущественно в интересах государства. Полученные таким путем сведения о государстве, его населении и достопримечательностях получили название статистики (от древне-итальянского слова *statista* — государствовед). В этой области знания вначале применялись приемы простого арифметического счета (политическая арифметика), но по мере накопления материала и возникновения новых, более трудных касающихся его вопросов, стали применяться приемы более тонкого математического анализа, главным образом, теории вероятностей. Эта область математики получила название математическая статистика, а метод ее — статистический метод, хотя область теории массовых явлений гораздо обширнее того круга вопросов, которые вызвали возникновение статистики.

§ 2. Понятие о законе распределения.

Пусть все члены рассматриваемой совокупности обладают каким-либо признаком, допускающим количественную оценку. В таком случае мы можем распределять все члены совокупности на классы, отнеся к одному классу те члены совокупности, у которых величина избранного признака или одинакова, или заключается в одинаковых пределах. Примеры такого распределения приведены в следующей таблице:

№№	Название совокупности.	Члены совокупности.	Основной признак распределения.
1	Определенный полк солдат.	Один солдат полка.	Рост солдата.
2	Группа цветков определенного вида растения.	Один цветок этой группы.	Число лепестков на цветке.
3	Рабочие известного города в известном году.	Один рабочий.	Заработная плата за год.
4	Умершие в известном городе за известный срок.	Один умерший.	Возраст умершего.
5	Младенцы, родившиеся в известном городе в известном году.	Группа в 100 родившихся младенцев.	Число мальчиков в этой группе.
6	Испытания над появлениями шаров из урны.	Серия в n испытаний.	Число появлений белых шаров в этой серии.

Обозначим величину признака, по которому совершается распределение данной совокупности, буквою x , число членов совокупности, имеющих величину x основного признака, — буквою y , а объем совокупности — буквою N .

Величина x в некоторых случаях может меняться только прерывно, скачками, получая, например, лишь целые значения (число лепестков на цветке, число мальчиков в группе 100 новорожденных детей, число появившихся белых шаров в группе при испытании), в других случаях x может меняться через сколь угодно малые промежутки (рост солдата, заработная плата, возраст умершего).

Массовые явления первого рода мы назовем прерывными, а второго — непрерывными.

Результат распределения данной совокупности по избранному основному признаку мы можем представить в виде таблицы, где в 1-м столбце стоят величины признака x , встречающиеся у членов данной совокупности, при чем эти величины идут или в порядке постепенного возрастания, или в порядке постепенного убывания; во втором столбце в каждой строке находится значение y , соответствующее стоящему в той же строке значению x .

Сумма чисел, стоящих во 2-м столбце, равна объему N совокупности. Это может служить проверкой вычисления.

Для тех примеров, которые приведены в предшествующей таблице, значения столбцов таблицы распределения будут следующие:

№№	x .	y .
1	Рост солдата.	Число солдат данного полка, имеющих рост равный x .
2	Число лепестков на цветке.	Число цветков взятой группы, имеющих по x лепестков.
3	Зарботная плата за год.	Число рабочих, получивших заработную плату x в данном городе в данном году.
4	Возраст умершего.	Число умерших в возрасте x в данном городе в данном году.
5	Число мальчиков в группе 100 новорожденных детей в данном году.	Число групп в 100 новорожденных, где было x мальчиков.
6	Число белых шаров в серии испытаний с урной.	Число серий, где появилось x белых шаров.

В виде числового примера рассмотрим таблицу, приводимую Вестергаардом *).

*) Westergaard. Die Grundzüge der Theorie der Statistik. Jena. 1890, стр. 29.

В Датской лотерее в 1887 — 1888 годах на 75.000 билетов было выдано 12.000 выигрышей. Эти билеты распределены на 750 групп по 100 билетов в каждой группе и подсчитано число выигрышей, выпавших на каждую группу.

В данном случае основной признак есть m — число выигрышей на одну группу в 100 билетов; величина y — число групп, в которых было m выигрышей.

В результате подсчета получилась следующая таблица распределения:

Таблица А.

m	y	m	y
5	2	17	68
6	3	18	64
7	2	19	43
8	7	20	41
9	12	21	35
10	28	22	24
11	43	23	20
12	53	24	11
13	59	25	11
14	57	26	2
15	81	27	1
16	83	28	1
			750

В рассматриваемом случае вероятность выигрыша на один билет равна $p = 0,16$. Наиболее вероятное число выигрышей на 100 билетов равно $100 \cdot 0,16 = 16$.

На основании сказанного в параграфе 4 главы III части I, находим, что вероятность получить в таком случае m выигрышей на 100 билетов выразится формулой:

$$P_m = \frac{h}{\sqrt{\pi}} e^{-hx^2} \cdot \frac{1}{\mu},$$

где

$$\mu = 100; x = \frac{m}{\mu} - p = \frac{m}{\mu} - 0,16; h = \sqrt{\frac{100}{2 \cdot 0,16 \cdot 0,84}} = 19,288;$$

$$\frac{h}{\sqrt{\pi}} \cdot \frac{1}{\mu} = 0,10882.$$

По теореме Якова Бернулли, произведение этой вероятности P_m на число всех групп, по 100 билетов в каждой, т.-е. на число 750, весьма близко к числу тех групп, на которые пало m выигрышей.

Так как $750 \cdot \frac{h}{\sqrt{\pi}} \cdot \frac{1}{\mu} = 750 \cdot 0,10882 = 81,61$, то произведя вычисление по формуле:

$$y = 750 \cdot P_m = 81,61 \cdot e^{-19,288 \frac{m^2}{750}}, \dots \dots \dots (B)$$

мы получим таблицу теоретических чисел, которые должны быть близки к числам приведенной выше таблицы (A).

Следовательно, формула (B) есть теоретическая формула, выражающая приближенно закон распределения той же совокупности, опытное распределение которого находится в таблице (A).

Результаты, даваемые формулой (B), расположатся в следующую таблицу *):

Таблица С.

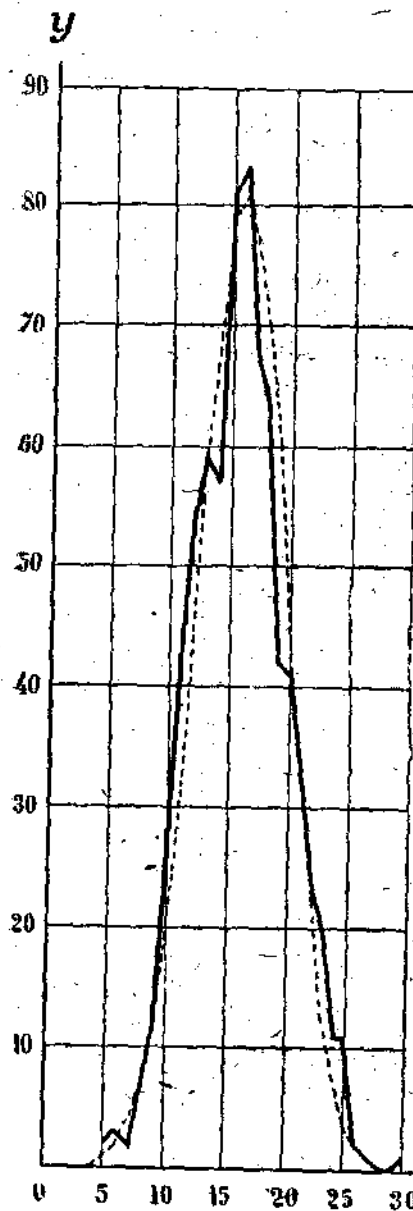
m	y	m	y
5	0,91	17	78,72
6	1,97	18	70,10
7	3,95	19	58,30
8	7,62	20	45,11
9	13,19	21	32,48
10	21,25	22	21,25
11	32,48	23	13,19
12	45,11	24	7,62
13	58,30	25	3,95
14	70,10	26	1,97
15	78,72	27	0,91
16	81,51	30	0,06

Для большей наглядности, материал, даваемый таблицей распределения, представим графически, принимая величину признака x за абсциссу, а y — за ординату точки на плоскости.

В случае прерывного массового явления пары величин (x, y) , даваемые таблицей распределения изобразятся точками на плоскости. Соединяя в последовательном порядке полученные точки прямыми линиями, получим ломанную, которую назовем многоугольником распределения.

*) При вычислениях по формуле (B) удобно пользоваться таблицей I-й в конце книги.

Абсциссы вершин этого многоугольника суть те значения x , которые встречаются в таблице распределения. Для целей вычисления желательно, чтобы последовательные значения x имели между собой



Чертеж 8.

одинаковые промежутки, т.е. были членами арифметической прогрессии. Если какое-либо x , стоящее в ряде членов такой прогрессии, не встречается в таблице распределения, то мы можем сказать, что соответствующее ему значение y равно нулю. При таком условии мы в праве считать все промежутки последовательных значений x равными между собой.

Так как масштаб измерения признака, по которому производилось распределение, может быть избран по нашему желанию, то мы можем выбрать его так, чтобы промежуток между двумя последовательными значениями x равнялся единице.

Многоугольник распределения для приведенного выше примера Вестергаарда представлен на черт. 8. На этом чертеже сплошной линией проведен многоугольник, соответствующий данным таблицы (А), и рядом с ним пунктиром проведен многоугольник, соответствующий теоретическим величинам, даваемым формулой (В) и таблицей (С).

Обозначая зависимость между переменной величиной x и функцией y функциональным знаком $\varphi(x)$, можем написать:

$$y = \varphi(x) \dots \dots \dots (1)$$

Обозначая величины переменного x в таблице распределения буквами:

$$x_0, x_1, x_2, \dots, x_i, \dots, x_{n-1} \dots \dots \dots (2)$$

а соответствующие значения y буквами:

$$y_0, y_1, y_2, \dots, y_i, \dots, y_{n-1} \dots \dots \dots (3)$$

можем написать:

$$y_i = \varphi(x_i), \quad i = 0, 1, 2, 3, \dots, n-1 \dots \dots \dots (4)$$

Сумма всех ординат ломанной распределения, т.е. сумма всех величин (3) даст объем совокупности N :

$$N = \sum_{i=0}^{i=n-1} y_i = \sum_{i=0}^{i=n-1} \varphi(x_i) \dots \dots \dots (5)$$

В случае явления непрерывного поступают несколько иначе. Весь промежуток значений, которые может иметь переменное x в изучаемом массовом явлении, положим от a до b , делится на произвольное число n равных частей, или интервалов, и подсчитывается, сколько членов совокупности имеют величину признака, лежащую соответственно в каждом из интервалов.

Обозначим длину интервала символом Δx ; тогда

$$\Delta x = \frac{b-a}{n}.$$

Обозначим интервалы в последовательном порядке номерами:

$$0, 1, 2, \dots, i, \dots, n-1,$$

а числа членов совокупности, соответствующих этим интервалам, буквами:

$$n_0, n_1, n_2, \dots, n_i, \dots, n_{n-1} \dots \dots \dots (6)$$

Разделив число членов, лежащих в i -том интервале, n_i на длину интервала Δx , найдем величину:

$$y_i = \frac{n_i}{\Delta x},$$

показывающую, сколько объектов пришлось бы на промежуток, равный единице, если бы распределение было всюду равномерно и если бы всюду на интервал приходилось по n_i объектов.

Эту величину y_i можно назвать частотой массового явления для i -того интервала.

Разделив все числа ряда (6) на Δx , найдем частоты для каждого интервала; обозначим их буквами:

$$y_0, y_1, y_2, \dots, y_i, \dots, y_{n-1} \dots \dots \dots (3')$$

Числа (6) изобразятся так:

$$y_0 \Delta x, y_1 \Delta x, y_2 \Delta x, \dots, y_i \Delta x, \dots, y_{n-1} \Delta x \dots \dots \dots (7)$$

Обозначая попрежнему объем совокупности буквой N , будем иметь:

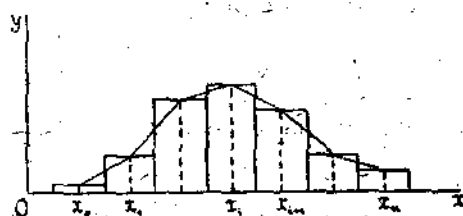
$$N = \Delta x \sum_{i=0}^{i=n-1} y_i$$

Для наглядного изображения закона распределения в случае непрерывности основного признака, мы снова возьмем прямоугольные оси координат и прежде всего отложим длины интервалов в виде последовательных отрезков на оси абсцисс (черт. 9). Обозначим абсциссы средин этих отрезков буквами:

$$x_0, x_1, x_2, \dots, x_i, \dots, x_{n-1}; \quad (2)$$

тогда i -й интервал будет иметь протяжение от $x_i - \frac{1}{2} \Delta x$ до $x_i + \frac{1}{2} \Delta x$.

На каждом интервале, как на основании, строим прямоугольник, площадь которого равна соответствующей величине ряда (6), или, что то же,



Чертеж 9.

ряда (7). Следовательно, высота прямоугольника выражает частоту явления для соответствующего интервала, а площадь выражает число объектов соответствующего промежутка.

Высоту каждого прямоугольника проводим через середину его основания.

Высота каждого прямоугольника есть функция абсциссы x — середины оснований этого прямоугольника.

Обозначая эту функциональную зависимость снова функциональным знаком $\varphi(x)$, будем иметь:

$$y = \varphi(x), \quad (1')$$

откуда:

$$y_i = \varphi(x_i), \quad i = 0, 1, 2, 3, \dots, n-1 \quad (4')$$

Итак, для непрерывного массового явления формула (1') определяет зависимость между частотой явления и величиной признака x .

Совокупность всех построенных прямоугольников образует ступенчатую площадь, верхний край которой графически выражает эту зависимость, иначе говоря, он выражает закон распределения для непрерывного массового явления. Такая ступенчатая фигура носит название гистограмма (название, предложенное Пирсоном).

Площадь построенной гистограммы равна сумме площадей всех входящих в состав ее прямоугольников, иначе говоря, она равна объему совокупности N .

Поэтому можно написать:

$$N = \Delta x \sum_{i=0}^{i=n-1} y_i = \Delta x \sum_{i=0}^{i=n-1} \varphi(x_i) \quad (5')$$

Ступенчатая линия, ограничивающая построенную площадь, как мы видели, выражается уравнением (1'). Когда Δx величина малая, то эта ступенчатая линия близко подходит к виду плавной кривой, а потому естественна мысль заменить уравнение (1') уравнением плавной аналитической кривой, которая возможно близко подходила бы к построенной нами ступенчатой линии.

Мы будем называть такую аналитическую кривую интерполяционной кривой по отношению к ступенчатой, а соответствующее уравнение — интерполяционным уравнением.

Чтобы не вводить новых обозначений, мы будем изображать интерполяционное уравнение той же формулой:

$$y = \varphi(x) \quad \dots \dots \dots (1')$$

и назовем $\varphi(x)$ интерполяционной функцией распределения.

Площадь интерполяционной кривой тем меньше разнится от площади ступенчатой, чем меньше величина Δx и чем ближе ступенчатая линия подходит к виду плавной кривой.

Величина площади интерполяционной кривой выражается интегралом $\int_a^b \varphi(x) dx$; следовательно, можно принять с некоторым приближением, что этот интеграл выражает площадь ступенчатой фигуры, т.е. он равен N , определяемой равенством (5'). На этом основании можем написать:

$$N = \Delta x \sum_{i=0}^{i=n-1} y_i = \Delta x \sum_{i=0}^{i=n-1} \varphi(x_i) = \int_a^b \varphi(x) dx \quad \dots \dots \dots (8)$$

Пользуясь произволом выбора масштаба для величины x , мы можем принять за единицу длины отрезков на оси x величину интервала Δx , т.е. можем положить величину Δx равною единице. Этим значительно упрощаются многие вычисления; нужно только помнить о сделанном условии, особенно, если в конце вычислений хотим вернуться к первоначальному масштабу.

Когда Δx равно 1, формула (8) принимает вид:

$$N = \sum_{i=0}^{i=n-1} y_i = \sum_{i=0}^{i=n-1} \varphi(x_i) = \int_a^b \varphi(x) dx \quad \dots \dots \dots (8')$$

§ 3. Средняя величина признака, понятие о моментах.

Простейший способ характеристики массового явления дает нам вычисление среднего арифметического величины признака, или, как говорят проще, средней величины признака.

Для вычисления средней величины признака в случае явления прерывного, умножаем каждую величину признака на число обладающих ею членов совокупности и сумму всех таких произведений делим на объем совокупности. Получится выражение:

$$\xi = \frac{\sum_{i=0}^{n-1} x_i \varphi(x_i)}{\sum_{i=0}^{n-1} \varphi(x_i)}, \quad \dots \quad (9)$$

или, на основании равенства (5):

$$\xi = \frac{1}{N} \sum_{i=0}^{n-1} x_i \varphi(x_i) \cdot \dots \quad (10)$$

Для явления непрерывного, когда Δx мало, ступенчатая линия близка к плавной, суммы приближенно заменяются интегралами, и мы получаем.

$$\xi = \frac{\int_a^b x \varphi(x) dx}{\int_a^b \varphi(x) dx}, \quad \dots \quad (9')$$

или, на основании равенства (8'):

$$\xi = \frac{1}{N} \int_a^b x \varphi(x) dx, \quad \dots \quad (10')$$

где $\varphi(x)$ есть интерполяционная функция.

Точка оси x с абсциссой ξ называется центром распределения.

В теории массовых явлений рядом со средним значением признака вводятся в рассмотрение и средние величины различных степеней этого признака, т. е. величины:

$$\frac{1}{N} \sum_{i=0}^{n-1} x_i^2 \varphi(x_i), \frac{1}{N} \sum_{i=0}^{n-1} x_i^3 \varphi(x_i), \dots, \frac{1}{N} \sum_{i=0}^{n-1} x_i^k \varphi(x_i), \dots$$

Они носят название моменты соответственно 2-го, 3-го, ... порядков и обозначаются так:

$$\nu_2, \nu_3, \dots, \nu_k, \dots$$

Ясно, что среднюю величину признака ξ можно назвать моментом 1-го порядка и обозначить буквою ν_1 .

В таком случае будем иметь для непрерывного явления:

$$v_1 = \frac{\sum_{i=0}^{n-1} x_i \varphi(x_i)}{\sum_{i=0}^{n-1} \varphi(x_i)}, \quad v_2 = \frac{\sum_{i=0}^{n-1} x_i^2 \varphi(x_i)}{\sum_{i=0}^{n-1} \varphi(x_i)}, \dots, \\ v_k = \frac{\sum_{i=0}^{n-1} x_i^k \varphi(x_i)}{\sum_{i=0}^{n-1} \varphi(x_i)}, \dots, \dots \quad (11)$$

или:

$$v_1 = \frac{1}{N} \sum_{i=0}^{n-1} x_i \varphi(x_i), \quad v_2 = \frac{1}{N} \sum_{i=0}^{n-1} x_i^2 \varphi(x_i), \dots, \\ v_k = \frac{1}{N} \sum_{i=0}^{n-1} x_i^k \varphi(x_i) \quad (12)$$

При этом, как мы говорили выше:

$$v_1 = \xi \quad (13)$$

В случае явления непрерывного суммы приближенно выразятся интегралами, и мы будем иметь:

$$v_1 = \frac{\int_a^b x \varphi(x) dx}{\int_a^b \varphi(x) dx}, \quad v_2 = \frac{\int_a^b x^2 \varphi(x) dx}{\int_a^b \varphi(x) dx}, \dots, \\ v_k = \frac{\int_a^b x^k \varphi(x) dx}{\int_a^b \varphi(x) dx}, \dots, \dots \quad (11')$$

или, на основании равенства (8'):

$$v_1 = \frac{1}{N} \int_a^b x \varphi(x) dx, \quad v_2 = \frac{1}{N} \int_a^b x^2 \varphi(x) dx, \dots, \\ v_k = \frac{1}{N} \int_a^b x^k \varphi(x) dx, \dots, \dots \quad (12')$$

Вычисления по формулам (12) упрощаются, если начало координат помещено в начальной точке первого интервала, а длины всех интервалов одинаковы.

Обозначим общую длину всех интервалов символом Δx , выразим абсциссы:

$$x_0, x_1, x_2, \dots, x, \dots, x_{n-1} \dots \dots \dots (2)$$

в таком виде:

$$0 \cdot \Delta x, 1 \cdot \Delta x, 2 \cdot \Delta x, \dots, i \cdot \Delta x, \dots, (n-1) \cdot \Delta x.$$

Величины моментов по формулам (12) представляются так:

$$\left. \begin{aligned} v_1 &= \frac{\Delta x}{N} (0 \cdot y_0 + 1 \cdot y_1 + 2 \cdot y_2 + \dots + i \cdot y_i + \dots + \\ &\quad + (n-1) \cdot y_{n-1}), \\ v_2 &= \frac{(\Delta x)^2}{N} (0^2 \cdot y_0 + 1^2 \cdot y_1 + 2^2 \cdot y_2 + \dots + i^2 \cdot y_i + \\ &\quad + \dots + (n-1)^2 \cdot y_{n-1}), \\ v_3 &= \frac{(\Delta x)^3}{N} (0^3 \cdot y_0 + 1^3 \cdot y_1 + 2^3 \cdot y_2 + \dots + i^3 \cdot y_i + \dots + \\ &\quad + (n-1)^3 \cdot y_{n-1}), \\ &\dots \dots \dots \end{aligned} \right\} \dots \dots (14)$$

Эти выражения еще упростятся, если, следуя замечанию, сделанному в предшествующем параграфе, примем длину Δx равною единице. В таком случае будем иметь:

$$\left. \begin{aligned} v_1 &= \frac{1}{N} (0 \cdot y_0 + 1 \cdot y_1 + 2 \cdot y_2 + \dots + i \cdot y_i + \dots + \\ &\quad + (n-1) \cdot y_{n-1}), \\ v_2 &= \frac{1}{N} (0^2 \cdot y_0 + 1^2 \cdot y_1 + 2^2 \cdot y_2 + \dots + i^2 \cdot y_i + \dots + \\ &\quad + (n-1)^2 \cdot y_{n-1}), \\ v_3 &= \frac{1}{N} (0^3 \cdot y_0 + 1^3 \cdot y_1 + 2^3 \cdot y_2 + \dots + i^3 \cdot y_i + \dots + \\ &\quad + (n-1)^3 \cdot y_{n-1}), \\ &\dots \dots \dots \end{aligned} \right\} \dots \dots (14')$$

Если по окончании вычислений мы пожелаем вернуться к первоначальному масштабу, то, обозначая моменты в первоначальном масштабе буквами:

$$N_1, N_2, N_3, \dots, N_k, \dots$$

можем написать:

$$N_1 = \Delta x \cdot v_1, N_2 = (\Delta x)^2 v_2, N_3 = (\Delta x)^3 v_3, \dots, N_k = (\Delta x)^k v_k, \dots$$

Случай непрерывного массового явления представляет значительно большие трудности при вычислении моментов. Основные формулы для их вычисления суть выражение (11'), но трудность в том, что вида функции $\varphi(x)$ мы заранее не знаем: нам известны только числовые значения ее, соответствующие значениям (2) независимого переменного x . Поэтому интегралы (11') могут быть вычислены только приближенно. Для уяснения приемов этого вычисления вернемся к геометрическому представлению и возьмем вместо плавной кривой ту ступенчатую, из которой она получается, и которая изображена на черт. 9. Моменты вычисленные для такой ступенчатой площади мы приближенно принимаем за моменты площади, огибаемой плавной кривой. Приближение будет тем больше, чем меньше длина интервала Δx .

Интегралы формул (11') представляются как суммы отдельных интегралов, относящихся последовательно к различным прямоугольникам, входящим в состав ступенчатой площади.

Возьмем один из таких интегралов, относящийся к прямоугольнику, у которого основание простирается от $x_i - \frac{1}{2} \Delta x$ до $x_i + \frac{1}{2} \Delta x$, а именно:

$$\int_{x_i - \frac{1}{2} \Delta x}^{x_i + \frac{1}{2} \Delta x} \varphi(x) x^k dx = \varphi(x_i) \int_{x_i - \frac{1}{2} \Delta x}^{x_i + \frac{1}{2} \Delta x} x^k dx = \varphi(x_i) \frac{\left(x_i + \frac{1}{2} \Delta x\right)^{k+1} - \left(x_i - \frac{1}{2} \Delta x\right)^{k+1}}{k+1} \quad (15)$$

При малом Δx можно разложить величины:

$$\left(x_i + \frac{1}{2} \Delta x\right)^{k+1} \text{ и } \left(x_i - \frac{1}{2} \Delta x\right)^{k+1}$$

по степеням Δx по формуле бинорма Ньютона и ограничиться при этом двумя первыми членами; при вычитании происходит некоторое сокращение, и выражение (15) принимает вид:

$$\varphi(x_i) x_i^k \Delta x.$$

Числитель выражения v_k в формулах (11') приближенно выражается суммой:

$$\Delta x \sum_{i=0}^{i=n-1} x_i^k \varphi(x_i).$$

Все выражение v_k будет следующее:

$$v_k = \frac{\sum_{i=0}^{i=n-1} x_i^k \varphi(x_i)}{\sum_{i=0}^{i=n-1} \varphi(x_i)} \dots \dots \dots (16)$$

Умножая числителя и знаменателя этого выражения на Δx и принимая во внимание равенство (8), найдем:

$$v_1 = \frac{\Delta x}{N} \sum_{i=0}^{i=n-1} x_i \varphi(x_i), \quad v_2 = \frac{\Delta x}{N} \sum_{i=0}^{i=n-1} x_i^2 \varphi(x_i), \dots, \\ v_k = \frac{\Delta x}{N} \sum_{i=0}^{i=n-1} x_i^k \varphi(x_i), \dots \dots \dots (12'')$$

Эти выражения разнятся от правых частей равенств (12) множителем Δx .

Когда Δx принято за единицу, то выражение (12'') совпадут с (12):

$$v_1 = \frac{1}{N} \sum_{i=0}^{i=n-1} x_i \varphi(x_i), \quad v_2 = \frac{1}{N} \sum_{i=0}^{i=n-1} x_i^2 \varphi(x_i), \dots, \\ v_k = \frac{1}{N} \sum_{i=0}^{i=n-1} x_i^k \varphi(x_i), \dots \dots \dots (12)$$

Различие только в том, что в случае непрерывного явления формулы (12) дают лишь приближенные величины моментов, и в них функция $\varphi(x)$ есть частота явления, а не число объектов при данном значении x .

Определяясь свойствами массового явления, моменты зависят и от положения начала координат.

Чтобы видеть, как преобразуются моменты, если передвинуть начало координат по оси абсцисс, совершим преобразование:

$$x = x_0 + x', \text{ откуда } x' = x - x_0, \dots \dots \dots (17)$$

где x_0 есть абсцисса нового начала, а x' — новая величина абсциссы ордината сохраняет свою величину.

Обозначая величины моментов относительно нового начала буквами:

$$v_1', \quad v_2', \quad v_3', \dots$$

будем иметь:

$$\begin{aligned} v'_k = \frac{1}{N} \sum_{i=0}^{k-1} (x_i - x_0)^k \varphi(x_i) = v_k - k_1 x_0 v_{k-1} + k_2 x_0^2 v_{k-2} - \dots + \\ + (-1)^{k-1} k_1 x_0^{k-1} v_1 + (-1)^k x_0^k, \quad \dots \quad (18) \end{aligned}$$

где $k_1, k_2, k_3, \dots$ суть коэффициенты бинома.

Отсюда видно, что, зная величины моментов относительно одного какого-нибудь начала, не трудно найти их величины для всякого другого начала. Поэтому обыкновенно начинают с вычисления моментов для какой-нибудь точки, лежащей в середине промежутка значений переменного в пределах данной таблицы распределения. Это значительно упрощает вычисления потому, что величины $x - x_0$ по абсолютному значению меньше, чем величины x , когда x_0 лежит вблизи середины промежутка распределения, знаки же величин $x - x_0$ то положительны, то отрицательны, вследствие чего при вычислениях происходят многие сокращения.

Если центр распределения принят за начало координат, то отнесенные к нему моменты называются центральными, мы обозначим их буквами: μ_1, μ_2, μ_3 .

Так как на основании равенства (13) абсцисса центра есть v_1 , то мы найдем центральные моменты, положив в формуле (18)

$$x_0 = v_1;$$

получим:

$$\mu_k = v_k - k_1 v_{k-1} v_1 + k_2 v_{k-2} v_1^2 - \dots + (-1)^{k-1} (k-1) v_1^k, \quad (18)$$

откуда:

$$\left. \begin{aligned} \mu_1 &= 0, \\ \mu_2 &= v_2 - v_1^2, \\ \mu_3 &= v_3 - 3 v_2 v_1 + 2 v_1^3, \\ \mu_4 &= v_4 - 4 v_3 v_1 + v_2 v_1^2 - 3 v_1^4. \end{aligned} \right\} \quad \dots \quad (19)$$

В виде примера вычисления моментов вычислим их для той совокупности, распределение которой дается в приведенной выше таблице (А) Вестергаарда.

Для упрощения вычислений перенесем начало координат в точку близкую к середине распределения, как сказано выше. За такую точку можем принять ту, которая имеет абсциссу $m = 16$. Поэтому делаем преобразование:

$$x' = m - 16.$$

Вычисления расположатся так, как показывает следующая таблица:

Таблица D.

m	$x' = m-16$	y	yx'	yx'^2	yx'^3	m	$x' = m-16$	y	yx'	yx'^2	yx'^3
5	-11	2	-22	242	-2662	17	1	68	68	68	68
6	-10	3	-30	300	-3000	18	2	64	128	256	512
7	-9	2	-18	162	-1458	19	3	42	126	378	1134
8	-8	7	-56	448	-3584	20	4	41	164	656	2624
9	-7	12	-84	588	-4116	21	5	35	175	875	4375
10	-6	28	-168	1008	-6048	22	6	24	144	864	5184
11	-5	43	-215	1075	-5375	23	7	20	140	980	6860
12	-4	53	-212	848	-3392	24	8	11	88	704	5632
13	-3	59	-177	531	-1593	25	9	11	99	891	8019
14	-2	57	-114	228	-456	26	10	2	20	200	2000
15	-1	81	-81	81	-81	27	11	1	11	121	1331
16	0	83	0	0	0	30	14	1	14	196	2744
		430	-1177	5511	-31765			820	1177	6189	40483
								430	-1177	5511	-31765
								750	0	11700	+ 8718

Отсюда следует, что:

$$v'_1 = \frac{0}{750} = 0; \quad v'_2 = \frac{11700}{750} = 15,6; \quad v'_3 = \frac{8718}{750} = 11,624.$$

Так как $v'_1 = 0$, то начало координат находится в центре распределения, моменты v'_2 и v'_3 суть центральные моменты, а потому можно написать:

$$\mu_2 = 15,6; \quad \mu_3 = 11,62.$$

Средняя величина равна 16.

Чтобы иметь пример более общего характера, припомним кое-что из сказанного в главе I предшествующей части. Пусть измерен μ раз один и тот же предмет с одинаковой степенью точности. Результаты этих μ измерений представляют собою совокупность. Вероятность, что ошибка одного из этих результатов равна ε , или, точнее говоря, что она лежит в бесконечно-малом промежутке между ε и $\varepsilon + d\varepsilon$, обозначалась в предшествующем отделе символом $\varphi(\varepsilon) d\varepsilon$, при чем:

$$\varphi(\varepsilon) = \frac{h}{\sqrt{\pi}} \cdot e^{-h^2 \varepsilon^2}.$$

По теореме Бернулли число измерений с ошибкой, лежащей в том же промежутке между ε и $\varepsilon + d\varepsilon$ приближенно равно

$$N\varphi(\varepsilon) d\varepsilon.$$

Разделив это число на длину промежутка $d\varepsilon$, найдем частоту рассматриваемой совокупности для промежутка между ε и $\varepsilon + d\varepsilon$. Следовательно, функция распределения в данном случае такова:

$$N\varphi(\varepsilon) = N \frac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon^2} \quad (20)$$

Момент k -того порядка выразится так:

$$\nu_k = \int_{-\infty}^{+\infty} \varepsilon^k \varphi(\varepsilon) d\varepsilon.$$

Как мы видим, это — математическое ожидание k -ой степени ошибки.

Положив $k=1$, находим:

$$\nu_1 = 0.$$

Это значит, что центр распределения лежит в начале координат, что и понятно, так как кривая симметрична относительно оси y .

Поэтому моменты относительно начала координат суть моменты центральные:

$$\mu_k = \int_{-\infty}^{+\infty} \varepsilon^k \varphi(\varepsilon) d\varepsilon.$$

В частности момент 2-го порядка есть математическое ожидание квадрата ошибки, которое, как мы знаем, есть квадрат средней квадратической ошибки, обозначенной нами буквою σ . Следовательно:

$$\sigma = \sqrt{\mu_2} \quad (21)$$

Из сказанного, между прочим, следует, что кривые вероятностей ошибок измерения, определяемые формулой Гаусса, суть кривые распределения ошибок в совокупности результатов измерений.

В частном случае, когда мы говорили в части I о вероятностях сложных событий при многократном повторении испытаний, мы видели, что кривые, выражающие закон этих вероятностей, суть в то же время кривые распределения отклонений отдельных частот от наимвероятнейшего значения частоты. Все эти кривые были кривыми того же Гауссова типа, симметричные относительно оси ординат. Квадратный корень из момента 2-го порядка есть среднее квадратическое отклонение.

Приведенный выше пример Вестергарда есть числовой пример для этого случая.

§ 4. Понятие о моде, о радиусе асимметрии, о среднем квадратическом уклонении и о склошенности кривой распределения.

В тех кривых распределения, которые обычно встречаются в теории массовых явлений, замечаются следующие особенности: кривая имеет конечное протяжение, ординаты ее точек вблизи концов весьма малы, в средней части кривой есть точка с наибольшей ординатой.

Абсцисса этой наивысшей точки кривой называется мода (the mode), ордината ее — модальной ординатой.

Бывают случаи, когда кривая распределения имеет на своем протяжении несколько таких вершин, т. е. имеет несколько мод; но обыкновенно это объясняется неоднородностью материала, сложностью изучаемого явления, которое можно разложить на несколько более простых составляющих его явлений обычного типа: с одномодальной кривой распределения. Так, напр., в биологии кривая с несколькими модами может встретиться при изучении размеров какого-нибудь органа у некоторого вида животных, если среди наблюдаемых животных этого вида встречается несколько различных пород. Рассортировав имеющийся материал по отдельным породам и построив для каждой породы свою кривую распределения, получим кривую одномодальную ¹⁾.

Конечно, такая группировка материала не всегда на практике возможна; тем не менее в теории обыкновенно считают ее выполненной и имеют в виду одномодальную кривую распределения.

Когда кривая распределения симметрична относительно какой-либо прямой, параллельной оси ординат, то наивысшая точка ее, если она единственная, должна лежать на оси симметрии; на той же оси должен лежать и центр. Если мы перенесем начало координат в центр, то в рассматриваемом случае мода будет равна нулю. Итак, в случае симметричной кривой распределения, модальная ордината совпадает со среднюю ординатой.

Такой случай представляет нам кривая Гаусса.

Мы видели, что при любом положении начала координат абсцисса центра равна первому моменту ν_1 относительно оси ординат:

$$\nu_1 = \xi \quad (13)$$

Если мы обозначим моду буквою X , то можем сказать, что для симметричной кривой

$$X = \xi = \nu_1.$$

¹⁾ Изучая весьма сложную кривую смертности, Пирсон разложил ее на 5 слагающих одномодального вида, соответственно пяти родам смерти: младенческой, детской, юношеской, зрелого возраста и старческой. См. Philosophical Transactions, Vol. 186, A, стр. 407

Но случай симметричной кривой есть случай частный, и притом довольно редкий. Если кривая не симметрична, то разность:

$$R = \xi - X \quad (22)$$

не равна нулю. Она называется радиусом асимметрии. Когда наивысшая точка лежит влево от центра, этот радиус положителен; когда она лежит вправо, он отрицателен.

Радиус асимметрии есть один из элементов, характеризующих вид кривой распределения. Другим важным элементом, тоже характеризующим вид этой кривой, служит центральный момент 2-го порядка μ_2 , или лучше, квадратный корень из этого момента:

$$\sigma = \sqrt{\mu_2} \quad (21)$$

Ясно, что выведенная нами в предшествующем параграфе формула (21) есть частный случай последней формулы.

Величина σ называется средним квадратическим отклонением.

Отношение радиуса асимметрии к среднему квадратическому отклонению называется скошенностью кривой:

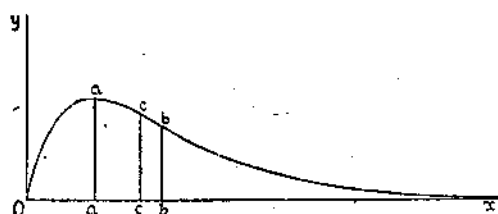
$$S = \frac{\xi - X}{\sigma} \quad (23)$$

Кроме средней величины ξ и моды X , есть еще одна величина, характеризующая кривую распределения. Это — так называемая срединная величина (the median). Если построим вертикальную прямую, делящую пополам площадь кривой распределения, то соответствующая абсцисса называется срединною.

При построении линии распределения мы начинаем, как уже говорили выше, с построения равных интервалов на оси абсцисс и проведения через средину каждого интервала соответствующей ординаты, после чего конечные точки этих ординат в последовательном порядке соединяются прямыми линиями. В обычном случае крайние ординаты справа и слева весьма малы, но они не равны нулю. В таких случаях прибавляют справа и слева еще по одному интервалу, считая для каждого из прибавленных интервалов среднюю ординату равной нулю. Срединны добавленных интервалов служат: одна — начальной точкой, другая — конечной точкой ломанной распределения; расстояние между этими точками называется основанием или базисом кривой распределения.

В некоторых случаях целесообразно считать линию распределения простирающеюся от $-\infty$ до $+\infty$, для чего принимаются равными нулю ординаты всех точек ломанной, лежащих влево от начала и вправо от конца базиса.

Построив ломанную распределения, мы находим плавную кривую возможно близко к ней подходящую. Эту интерполяционную кривую мы назовем кривой распределения.



Чертеж 10.

Этот план мы уже выполнили на частном примере, говоря о кривой, выражающей закон распределения вероятностей сложных событий при большом числе испытаний.

Та кривая, которую мы тогда получили, была симметрична, при чем надо, однако, заметить,

что это произошло отчасти вследствие тех приближений, которые мы делали, составляя уравнение интерполяционной кривой.

В общем случае кривая распределения симметричной не будет. На черт. 10 изображен обычный вид асимметричной кривой распределения и на нем отмечен центр b , середина c и модальная a .

§ 5. Построение интерполяционной формулы по способу моментов.

Пусть $\varphi(x)$ есть некоторая функция, выражение которой нам неизвестно: она определена или графически, или значениями при достаточно близких между собою значениях x . Требуется подыскать выражение аналитической функции $f(x)$ такое, чтобы значения ее возможно близко подходили к значениям функции $\varphi(x)$ в данном промежутке от a до b .

Близость значений функций $\varphi(x)$ и $f(x)$ при данном значении x определяется абсолютным значением разности $\varphi(x) - f(x)$; чтобы избежать влияния знака этой величины, мы берем квадрат ее:

$$(\varphi(x) - f(x))^2 \dots \dots \dots (24)$$

Близость значений обеих функций $\varphi(x)$ и $f(x)$ на всем протяжении от a до b естественно оценить величиною суммы выражений вида (24) при всех значениях x от a до b , иначе говоря, величиною интеграла:

$$\int_a^b (\varphi(x) - f(x))^2 dx \dots \dots \dots (25)$$

Наилучшим выражением интерполяционной формулы мы считаем то выражение функции $f(x)$, при котором величина (25) возможно мала.

Дальнейшие рассуждения мы ведем в предположении, что $f(x)$ может быть разложена в ряд по целым степеням x , сходящийся в пределах от a до b .

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots \quad (26)$$

Внеся это выражение в формулу (25), мы выбираем для коэффициентов:

$$a_0, a_1, a_2, a_3, \dots \quad (27)$$

такие значения, чтобы величина (25) была возможно малой. С этой целью приравняем нулю частные производные выражения (25) по каждому из коэффициентов (27):

$$\int_a^b (\varphi(x) - f(x)) x^k dx = 0, \quad k = 0, 1, 2, 3, \dots$$

откуда:

$$\int_a^b \varphi(x) x^k dx = \int_a^b f(x) x^k dx, \quad k = 0, 1, 2, 3, \dots \quad (28)$$

Интеграл в левой части этих равенств можно вычислить приближенно, пользуясь тем приемом, который изложен в параграфе 3, при выводе формул (12).

Левая часть формулы (28) при $k = 0$ есть величина площади кривой, выражаемой уравнением

$$\varphi(x) = 0, \quad (1)$$

в пределах от a до b . Мы обозначим ее буквою N . Если (1) есть уравнение кривой распределения, то N есть, как мы знаем, объем соответствующей совокупности.

Первое из уравнений (28) показывает, что площадь искомой интерполяционной кривой между a и b

$$f(x) = 0 \quad (29)$$

должна равняться объему совокупности N .

Обозначив площадь кривой (29) буквою S :

$$\int_a^b f(x) dx = S, \quad (30)$$

находим:

$$N = S \quad (31)$$

Разделив на N обе части уравнения (28), находим:

$$\frac{1}{N} \int_a^b \varphi(x) x^k dx = \frac{1}{N} \int_a^b f(x) x^k dx, \quad k = 1, 2, 3, \dots$$

левая часть этого равенства есть момент k -вого порядка кривой (1); правая часть его — такой же момент кривой (29). Обозначая первый буквою v_k , а второй буквою N_k , находим:

$$v_k = N_k, \quad k = 1, 2, 3 \dots \dots \dots (32)$$

Величины левых частей равенств (32) вычисляются только-что упомянутым способом, изложенным в параграфе 3; поэтому приближенные их величины можно считать известными. В таком случае равенства (32) дадут значения всевозможных моментов искомой интерполяционной кривой (29). Эти равенства показывают, что весь бесконечный ряд моментов кривой (29) должен равняться соответствующим моментам кривой распределения (1).

Само собою понятно, что удовлетворить всему бесконечно большому числу равенств (32) на практике невозможно ¹⁾, но мы можем задаться целью удовлетворить возможно большему числу первых по порядку из этих равенств:

$$v_1 = N_1, v_2 = N_2, v_3 = N_3, \dots, v_{k-1} = N_{k-1}, \dots \dots \dots (33)$$

где k — число конечное. К этим равенствам надо присоединить и рассмотренное уже равенство (31):

$$N = S \dots \dots \dots (31)$$

Если функция $f(x)$ выражается некоторой формулой, содержащей в себе k буквенных параметров, то, вообще говоря, можно подобрать числовые величины для этих k параметров так, чтобы удовлетворить k уравнениям (31) и (33).

Выбор вида формулы, выражающей функцию $f(x)$, основывается на изучении свойств данного массового явления и в значительной мере произволен; но в большом числе случаев, к которым приводили вопросы из области различных отделов статистики, антропологии и биологии, получились очень хорошие результаты при взятии для функции $f(x)$ формулы Гаусса и других формул, представляющих ее ближайшие обобщения. Рассмотрением этих видов интерполяционной формулы и способов вычисления входящих в нее буквенных параметров мы займемся в следующих параграфах.

¹⁾ Помимо того, что произвести приближенные вычисления бесконечно-большого числа моментов $v_1, v_2, v_3, \dots$ по существу невозможно, надо заметить, что сами эти вычисления делаются все более и более утомительными в арифметическом смысле с возрастанием порядка момента, а результат приближенного вычисления делается при этом все более и более грубым. Поэтому Пирсов, предложивший способ моментов, на основании очень большой практики применения этого способа, пришел к заключению, что нецелесообразно вводить в вычисление моменты порядка выше 4-го.

§ 6. Формула нормального распределения.

Наиболее давний случай распределения встретился в вопросе о вероятностях появления различного числа белых и черных шаров из урны, а затем при изучении вероятностей ошибок измерения. Как мы видели в конце параграфа 3, в обоих этих случаях, умножив выражение вероятности на число наблюдений n и разделив на длину интервала, получим функцию распределения для данной совокупности. Она выражается формулой (20), или, обозначая аргумент буквою x , а функцию буквою y :

$$y = N \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} \dots \dots \dots (34)$$

Кривая, выражаемая этим уравнением, имеет центр распределения в начале координат. Если мы перенесем начало координат в какую-нибудь точку на расстоянии x_0 от центра, то кривая выразится уравнением такого же вида, но с заменой буквы x выражением $x - x_0$, т.е. мы будем иметь:

$$y = N \frac{h}{\sqrt{\pi}} e^{-h^2 (x - x_0)^2} \dots \dots \dots (35)$$

Такова формула как для распределения ошибок в совокупности n измерений, так и числа появлений белых шаров в совокупности n групп испытаний.

Момент 1-го порядка для кривой (35) выразится формулой:

$$v_1 = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2 (x - x_0)^2} x dx,$$

или, сделав подстановку:

$$x = x_0 + \frac{t}{h};$$

$$v_1 = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} \left(x_0 + \frac{t}{h} \right) dt = x_0,$$

как и должно быть, ибо x_0 есть абсцисса центра распределения; следовательно:

$$x_0 = v_1 = \xi \dots \dots \dots (36)$$

Момент 2-го порядка выразится формулой:

$$v_2 = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-h^2 (x - x_0)^2} x^2 dx.$$

После той же подстановки:

$$x = x_0 + \frac{t}{h}$$

найдем:

$$\begin{aligned} \mu_2 &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} \left(x_0 + \frac{t}{h} \right)^2 dt = \\ &= \frac{x_0^2}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} dt + \frac{1}{h^2 \sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} t^2 dt = x_0^2 + \frac{1}{2h^2}. \end{aligned}$$

На основании второй из формул (19) находим:

$$\mu_2 = \frac{1}{2h^2},$$

откуда, на основании равенства (21):

$$h = \frac{1}{\sqrt{2\mu_2}} = \frac{1}{\sigma\sqrt{2}} \quad (37)$$

Эта формула нам уже встречалась в части I и II ¹⁾.

На основании равенств (36) и (37) формула распределения (35) примет вид:

$$y = \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\xi)^2} \quad (38)$$

Кроме рассмотренных двух случаев (распределения числа шаров и распределения ошибок) формула (38) нашла широкое применение еще в других случаях, преимущественно в антропологии и биологии.

Начало таких применений положено основателем теории массовых явлений — Кетле. Рассматривая таблицы распределения при изучении человеческого роста и размеров различных частей человеческого тела, он заметил большое сходство с теми распределениями, которые встречаются в распределении шаров и распределении ошибок. Это привело его к такой мысли: природа в строении человека и других организмов имела в виду определенные образцы, от которых происходят некоторые отклонения только под влиянием случайных причин. Если эта мысль верна, то, конечно, отклонения в росте человека, в размерах частей его тела и других подобных величинах от нормального образца должны следовать тому же закону, как отклонения результатов измерения от истинной величины измеряемого предмета.

Ряд вычислений, произведенных как самим Кетле, так и его последователями, дали настолько прочное подтверждение верности его

¹⁾ См. ч. I гл. III форм. (34); ч. II глава I форм. (25).

мысли, что формула распределения (38) получила название нормальной формулы распределения и долгое время была единственно применяемая в теории массовых явлений, хотя уже сам Кетле указывал на существование случаев распределения, где кривая асимметрична.

Функция нормального распределения определяется формулой (38), содержащей в себе три буквенных параметра: N , ξ , σ . Величины этих параметров каждый раз при применении формулы (38) определяются по данной таблице распределения на основании того плана, который изложен в предшествующем параграфе: параметр N есть объем совокупности, а ξ и σ выражаются через моменты первых двух порядков по формулам (36) и (37).

В качестве примера применим формулы настоящего параграфа к тому материалу Вестергаарда, который приведен выше в таблице (A). Вычисление моментов находится в таблице (D). Из нее видно, что абсцисса центра $\xi = 16$, момент 2-го порядка $\mu_2 = 15,6$, объем совокупности $N = 750$.

Отсюда следует, что

$$h = \frac{1}{\sqrt{2\mu_2}} = \frac{1}{\sqrt{31,2}} = 0,179.$$

Формула (36) примет вид:

$$y = 750 \frac{0,179}{\sqrt{\pi}} e^{-0,179^2 x^2}.$$

Результаты вычисления по этой формуле находятся в следующей таблице (E), где для сравнения рядом с величинами, вычисленными по этой формуле, приведены и соответствующие величины из таблицы (A), а также и разность между теми и другими.

Таблица E.

m	$x = \xi - 16$	Величина y.		Разность.	m.	$x = \xi - 16$	Величина y.		Разн.
		По формуле	По таблице (A).				По формуле.	По таблице (A).	
5	-11	1,56	2	0	17	1	73,33	68	+ 5
6	-10	3,07	3	0	18	2	66,53	64	+ 3
7	-9	5,67	2	+ 4	19	3	56,59	42	+ 15
8	-8	9,80	7	+ 3	20	4	45,10	41	+ 4
9	-7	15,88	12	+ 4	21	5	33,70	35	- 1
10	-6	24,10	28	- 4	22	6	24,10	24	0
11	-5	33,70	43	- 9	23	7	15,88	20	- 4
12	-4	45,10	53	- 8	24	8	9,80	11	- 1
13	-3	56,59	59	- 2	25	9	5,67	11	- 5
14	-2	66,53	57	+ 10	26	10	3,07	2	1
15	-1	73,33	81	- 8	27	11	1,56	1	1
16	0	75,74	83	- 7	—	—	—	—	—

Самый вид формулы (38) получился исторически тем путем, который мы рассмотрели в I и II части, но тот же самый исторический путь привел к необходимости обобщить вид этой функции, когда обнаружилось, что он дает недостаточно точные результаты особенно в совокупностях явно несимметричных.

После неудачной попытки Фехнера исправить почти искусственным образом формулу Гаусса, явилась естественная мысль вернуться к тем соображениям, на которых основывался вывод этой формулы. Так как при выводе формулы распределения числа шаров, было сделано приближение довольно грубое, то естественно было повторить то же вычисление, но с большей точностью.

Этим путем мы приходим к первому обобщению формулы нормального распределения.

Впервые это обобщение было сделано американским математиком Форест (Forest), но работы его, напечатанные в очень мало распространенном журнале *Analyst*, остались незамеченными не только на континенте Европы, но даже в Англии, где та же формула была найдена заново Пирсоном, который сам об этом говорит в журнале *Nature*.

Так как формула (38) получилась при нахождении интерполяционной формулы закона изменения членов бинома, то естественно назвать то выражение, которое мы получим при нахождении более точной формулы для тех же членов бинома, формулой биномиального распределения.

§ 7. Формула биномиального распределения.

Вернемся к обозначениям параграфа 4 главы III, части I.

Мы положим:

$$P_m = \frac{\mu!}{m!n!} p^m q^n, \quad y = \mu P_m, \quad \frac{m}{\mu} = p + x,$$

откуда следует:

$$\frac{1}{y} \frac{\Delta y}{\Delta x} = \frac{\mu \left(-x - \frac{q}{\mu} \right)}{q(p+x) + \frac{q}{\mu}}, \quad ^1)$$

или:

$$\frac{1}{y} \frac{\Delta y}{\Delta x} = -\frac{\mu}{q} + \frac{\mu p \left(1 + \frac{1}{\mu} \right)}{q} \cdot \frac{1}{p + \frac{1}{\mu} + x}.$$

¹⁾ См. формулу (15) параграфа 4 гл. III, части I.

Введя обозначения:

$$\frac{\mu}{q} = \alpha, \quad \frac{\exp\left(1 + \frac{1}{\mu}\right)}{q} = \beta, \quad p + \frac{1}{\mu} = \gamma. \quad (39)$$

и, совершив переход к пределу при $\Delta x = 0$, найдем:

$$\frac{1}{y} \frac{dy}{dx} = -\alpha + \beta \cdot \frac{1}{\gamma + x} \quad (40)$$

Интегрируя это равенство, мы находим:

$$y = C e^{-\alpha x} \left(1 + \frac{x}{\gamma}\right)^{\beta} \quad (41)$$

Таково уравнение биномиальной кривой распределения.

В него входят 4 параметра:

$$C, \alpha, \beta, \gamma.$$

Для определения их надо найти приближенные величины площади кривой (41) и моментов первых трех порядков: v_1, v_2, v_3 .

Из формул (39) видно, что параметры α, β, γ существенно положительны, а из уравнения (41) видно, что y обращается в 0 при $x = -\gamma$ и при $x = +\infty$. Поэтому площадь кривой выразится интегралом:

$$C \int_{-\gamma}^{+\infty} e^{-\alpha x} \left(1 + \frac{x}{\gamma}\right)^{\beta} dx.$$

Так как она должна равняться объему N совокупности, то:

$$N = C \int_{-\gamma}^{+\infty} e^{-\alpha x} \left(1 + \frac{x}{\gamma}\right)^{\beta} dx. \quad (42)$$

Момент произвольно взятого k -ого порядка выразится формулой:

$$v_k = \frac{1}{N} \int_{-\gamma}^{+\infty} y x^k dx \quad (43)$$

Для вычисления моментов припомним, что ордината y должна обращаться в 0 при обоих пределах интеграла, а потому, произведя интегрирование по частям, будем иметь при всяком положительном k :

$$\int_{-\gamma}^{+\infty} x^k dy = \left[x^k y \right]_{-\gamma}^{+\infty} - k \int_{-\gamma}^{+\infty} x^{k-1} y dx \quad (44)$$

Отсюда:

$$\frac{1}{N} \int_{-\gamma}^{+\infty} x^k dy = -k v_{k-1} \dots \dots \dots (45)$$

Освободив равенство (40) от знаменателя и умножив на $\frac{x^k}{N}$, найдем:

$$\frac{1}{N} \gamma x^k dy + \frac{1}{N} x^{k+1} dy + \frac{\alpha\gamma - \beta}{N} y x^k dx + \frac{\alpha}{N} y x^{k+1} dx = 0.$$

Интегрируя это равенство по x и приняв во внимание (45) и (43), будем иметь:

$$-\gamma k v_{k-1} - (k+1) v_k + (\alpha\gamma - \beta) v_k + \alpha v_{k+1} = 0. \dots \dots (46)$$

Полагая k поочередно равным 0, 1, 2, найдем:

$$\left. \begin{aligned} -1 + (\alpha\gamma - \beta) + \alpha v_1 &= 0, \\ -\gamma - 2v_1 + (\alpha\gamma - \beta)v_1 + \alpha v_2 &= 0, \\ -2\gamma v_1 - 3v_2 + (\alpha\gamma - \beta)v_2 + \alpha v_3 &= 0. \end{aligned} \right\} \dots \dots \dots (47)$$

Если начало координат помещено в центре распределения, то моменты v_1, v_2, v_3 будут центральными и обозначатся буквами μ_1, μ_2, μ_3 при чем $\mu_1 = 0$; равенства (47) примут вид:

$$\left. \begin{aligned} -1 + (\alpha\gamma - \beta) &= 0, \\ -\gamma + \alpha\mu_2 &= 0, \\ (\alpha\gamma - \beta - 3)\mu_2 + \alpha\mu_3 &= 0, \end{aligned} \right\} \dots \dots \dots (48)$$

откуда:

$$\alpha = \frac{2\mu_2}{\mu_3}, \beta = \frac{4\mu_2^2}{\mu_3^2} - 1, \gamma = \frac{2\mu_2^3}{\mu_3^3} \dots \dots \dots (49)$$

Эти равенства определяют величины параметров α, β, γ , когда центр распределения принят за начало координат.

Для определения еще одного оставшегося пока неизвестным параметра C служит равенство (42). Преобразовав в интеграле этой формулы переменное подстановкой

$$x = \frac{t}{\alpha} - \gamma,$$

получаем:

$$N = C \frac{e^{\alpha\gamma}}{\alpha^{\beta+1} \gamma^{\beta}} \int_{-\infty}^{+\infty} e^{-t} t^{\beta} dt.$$

Интеграл этой формулы есть функция параметра β , известный из интегрального исчисления и обозначаемый символом Γ ; он равен $\Gamma(\beta+1)^*$.

*) Для вычисления значений этой функции составлены таблицы.

Поэтому полученное равенство примет вид:

$$N = C \frac{e^{\alpha\gamma}}{\alpha^{\beta+1} \gamma^{\beta}} \Gamma(\beta+1),$$

откуда:

$$C = N \frac{e^{-\alpha\gamma} \alpha^{\beta+1} \gamma^{\beta}}{\Gamma(\beta+1)} \dots \dots \dots (50)$$

В качестве примера применим формулы настоящего параграфа к тому же материалу Вестергаарда, приведенному в таблице (А). Из вычислений, сделанных в таблице (D), мы видим, что в данном случае абсцисса центра $\xi = 16$, момент 2-го порядка $\mu_2 = 15,6$, момент 3-го порядка $\mu_3 = 11.624$, объем совокупности $N = 750$.

Внеся эти величины в формулы (49), найдем:

$$\alpha = 2,68; \beta = 111,39; \gamma = 41,87.$$

Поэтому формула (41) закона распределения в данном случае будет такова:

$$y = C e^{-2,68x} \left(1 + \frac{x}{41,87}\right)^{111,39}.$$

Величину постоянного C можно вычислить по формуле (50). Но это потребовало бы пользования таблицами функции Γ , которые не всегда имеются под руками. В таких простых случаях, как нами рассматриваемый, можно этого избежать следующим образом.

Мы знаем, что сумма значений величин y при всех значениях x , находящихся в таблице, должна равняться объему совокупности $N = 750$. Поэтому мы получим величину C , разделив 750 на сумму S значений выражения

$$u = e^{-2,68x} \left(1 + \frac{x}{41,87}\right)^{111,39}$$

при всех находящихся в таблице значениях x .

Выполнив вычисления значений величины u , и составив их сумму S , в данном случае находим:

$$S = 9,9389;$$

поэтому:

$$C = \frac{N}{S} = \frac{750}{9,9389} = 75,46.$$

Умножая на это число все уже вычисленные значения величины u , находим соответствующие значения y .

Эти значения y , вычисленные по биномиальной формуле распределения (41), приведены в следующей таблице (F), где опять для сравнения помещены величины из основной таблицы (А) и разность между величинами, даваемыми формулой и начальными данными таблицы (А).

Т а б л и ц а F.

m	x = m - 16.	Величина y по:		Разность.	m	x = m - 16.	Величина y по:		Разность.
		формуле.	таблице				формуле.	таблице	
5	11	0,86	2	- 1	17	1	72,63	68	+ 5
6	10	2,07	3	- 1	18	2	65,74	64	+ 2
7	9	4,46	2	+ 2	19	3	56,09	42	+ 14
8	8	8,62	7	+ 2	20	4	45,35	41	+ 4
9	7	14,93	12	+ 3	21	5	31,43	35	- 4
10	6	23,94	28	- 4	22	6	22,94	24	- 1
11	5	35,18	43	- 8	23	7	15,90	20	- 4
12	4	47,62	53	- 5	24	8	10,52	11	0
13	3	59,70	59	+ 1	25	9	6,66	11	- 4
14	2	68,59	57	+ 12	26	10	3,69	2	+ 2
15	1	74,28	81	- 7	27	11	3,34	1	+ 2
16	0	75,46	83	- 8					

Формулы (49) и (50) выражают все четыре параметра формулы биномиального распределения (41), когда центр распределения принят за начало координат. Если же мы хотим поместить начало координат не в центре, то для получения формулы распределения должны так же, как и в предшествующем параграфе заменить аргумент x выражением $x - \xi$, где ξ есть абсцисса центра относительно выбранного начала. Она равна моменту μ_1 первого порядка для этого начала и определяется по данной таблице распределения обычным способом, рассмотренным в параграфе 3.

Для выхождения моды кривой (41), т.е. абсциссы наивысшей точки этой кривой, мы приравняем нулю производную логарифма выражения (41); получим:

$$- \alpha + \frac{\beta}{\gamma + x} = 0,$$

откуда:

$$x = \frac{\beta}{\alpha} - \gamma.$$

Следовательно, мода X равна:

$$X = \frac{\beta}{\alpha} - \gamma = - \frac{\mu_1}{2\mu_2}.$$

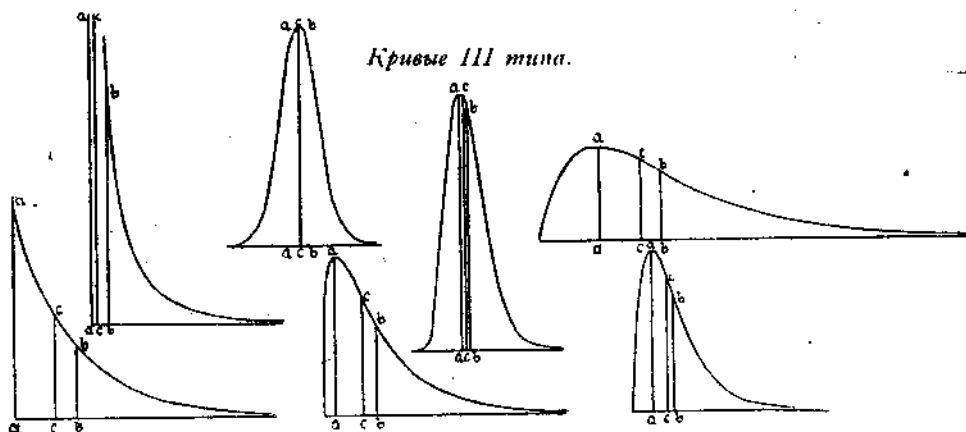
Когда начало координат помещено в центре кривой, то $\xi = 0$, и по формуле (22) радиус асимметрии выразится так:

$$R = -X = \frac{\mu_3}{2\mu_2} \dots \dots \dots (51)$$

Если данное распределение не симметрично, то момент μ_3 не будет равен нулю, а потому и R , определяемое равенством (51), от нуля отлично: кривая (41) асимметрична.

В рассматриваемом случае кривая распределения сравнительно с формулой нормального типа содержит к себе одним параметром больше: в нее входят площадь и три первых момента. Поэтому естественно, что она может ближе подойти в интерполируемой кривой, чем нормальная кривая. Это особенно важно в тех случаях, когда данное распределение асимметрично, вследствие неравенства нулю третьего момента. Чем абсолютная величина его больше, тем значительнее радиус асимметрии (51).

В зависимости от обоих параметров μ_2 и μ_3 вид кривой (41) может быть очень разнообразен, что видно на черт. 11, заимствованном у Пирсона в *Philosophical Transactions*. Vol. 186. A. Табл. 9, фиг. 5.



Черт. 11.

На этом чертеже представлено 7 частных видов биномиальной кривой распределения, при чем буквами *aa* обозначена ордината модальная, буквами *bb* — средняя и *cc* — срединная.

Надо заметить, что, при вычислении коэффициентов α , β , γ по формулам (49) на основании таблицы распределения, показатель β может оказаться отрицательным. Если он по абсолютной величине меньше единицы, то площадь кривой и все ее моменты конечны, и формула (41) дает выражение кривой распределения, несмотря на то, что при $x = -\gamma$ ордината кривой обращается в бесконечность. Такой случай представляют 1-я и 2-я кривая чертежа 11. Когда радиус

асимметрии очень мал, то кривая почти симметрична и мало разнится от нормальной; такой случай представляют кривая 3-я и 4-я черт. 11. Три следующие кривые этого чертежа имеют более или менее значительную асимметрию.

Кривая биномиального типа вследствие этого разнообразия своих видов имеет много приложений к практическим задачам. Между прочим она была найдена экспериментальным путем Лозанским профессором Парето в работе по распределению подоходного налога.

В приложениях к реальным задачам биномиальная кривая встретила Пирсону при упомянутом выше разложении кривой смертности на 5 составляющих по роду смерти. При этом формулой биномиального типа выразились: 1) кривая младенческой смертности, имеющая вид, подобный первой кривой на черт. 11; начало ее с бесконечно большой ординатой лежит на $\frac{3}{4}$ года раньше рождения (случаи смерти до рождения суть случаи мертворожденных детей, и относятся по формуле Пирсона к промежутку за $\frac{3}{4}$ года до рождения); 2) кривая детской смертности, имеющая вид, подобный последней кривой на черт. 11; 3) кривая старческой смертности, которая похожа на нормальную, но имеет небольшую асимметрию отрицательную.

В формулу (41) входит множитель e^{-ax} . Из дифференциального исчисления известно, что e^{-x} есть предел выражения $\left(1 - \frac{x}{\delta}\right)^\delta$ при $\delta = \infty$.

Поэтому при δ бесконечно-большом значение e^{-ax} близко к $\left(1 - \frac{x}{\delta}\right)^{\infty}$, а выражение (41) близко к:

$$C \left(1 - \frac{x}{\delta}\right)^{\infty} \left(1 + \frac{x}{\gamma}\right)^{\beta}.$$

Поэтому, обозначив ∞ одной буквой α , можем сказать, что формула биномиального распределения есть частный случай более общей формулы:

$$C \left(1 - \frac{x}{\delta}\right)^{\alpha} \left(1 + \frac{x}{\gamma}\right)^{\beta}, \dots \dots \dots (52)$$

которая переходит в формулу (41) в пределе при $\delta = \infty$.

Выражение (52) есть формула распределения, предложенная Пирсоном: она называется гипергеометрической формулой распределения,

§ 8. Понятие о формулах Пирсона гипергеометрического закона распределения.

Пирсон применил такой же прием, как изложенный в предшествующем параграфе, к условиям задачи теории вероятностей рассмотренной выше в параграфе 11 главы III, части I. В этой задаче предполагается, что испытания происходят над появлением шаров из урны, когда вынутый шар в урну не возвращается. Обозначим число всех шаров в урне буквою M , вероятность появления белого шара до начала испытаний буквою p ; пусть всего произведено μ испытаний, при чем вынутый шар в урну не возвращался. Спрашивается, какова вероятность, что при произведенных μ испытаниях число появлений белого шара равно m .

Выражение этой вероятности, как мы видели, разнится от выражения вероятности в том случае, когда вынутый шар возвращается в урну, и это различие между обеими формулами тем значительнее, чем ближе к 1 величина отношения $\frac{M}{\mu}$: когда это отношение бесконечно велико, то обе формулы разнятся между собою бесконечно мало.

Если вынутый шар возвращается в урну, то выражения вероятностей всевозможных сложных событий при таких испытаниях служат, как мы знаем, членами бинома. Если вынутый шар откладывается в сторону, то выражения всевозможных вероятностей будут служить членами другого ряда, более сложного, принадлежащего к числу гипергеометрических, почему и распределение, даваемое выражениями вероятностей в этом случае естественно назвать гипергеометрическим.

Из сказанного следует, что гипергеометрическое распределение зависит от величины отношения $\frac{M}{\mu}$; когда это отношение бесконечно велико, то гипергеометрическое распределение очень близко к биномиальному, а когда $\frac{M}{\mu}$ близко к единице, различие должно быть весьма значительно.

Применяя к гипергеометрическому случаю те же рассуждения, как в предшествующем параграфе, мы придем к формуле (52).

Вывод формулы (52) лишь немного сложнее вывода формулы (41), но кривая (52) представляет гораздо более разнообразия случаев, и их изложение потребовало бы гораздо более места, чем сколько можно отвести в этом курсе.

Поэтому ограничимся лишь кратким очерком.

Кривая (52) содержит в себе 5 параметров, т.е. на один более сравнительно с биномиальной и на 2 более сравнительно с нормаль-

ной кривой, чем и объясняется гораздо большая ее гибкость и разнообразие форм.

Для вычисления параметров, необходимо знание площади и моментов первых четырех порядков, вычисляемых по данной таблице распределения известным нам способом.

Пирсон делит виды найденной им кривой распределения на 7 типов (впоследствии один из этих типов он подразделил на 2, так что получилось всего 8 типов). Типы характеризуются особой формулой выражающейся через моменты и называемой критерием Пирсона. В эту классификацию входят, как частные случаи, также кривые нормальная и биномиальная. Нормальная кривая определяется величиной критерия k , равной 0, и кроме того, условиями $v_3 = 0$ и $\mu_4 = 3\mu_2^2$; она составляет тип VII *); биномиальная кривая составляет тип III и определяется величиной критерия $k = \infty$.

Таким образом теория Пирсона охватывает все до сих пор изученные виды кривых распределения **).

*) Пирсон называет этот тип просто нормальным; название кривая типа VII предложено Эльдertonом.

**) Кроме теории Пирсона, дающей классификацию кривых распределения и способ вычисления кривой каждого класса в отдельности, существует другая теория Врунса, основанная на иных началах. Она представляет собою непосредственное обобщение формулы Гаусса и опирается на свойства так называемых многочленов Чебышева и на разложение функций в ряд по этим многочленам. См. Врунс, *Wahrscheinlichkeitsrechnung und Kollektivmasslehre*. Leipzig und Berlin, 1906.

ГЛАВА ВТОРАЯ.

О корреляциях.

§ 1. Распределение по двум признакам.

В предшествующей главе мы говорили о распределении совокупности по одному признаку, например, о распределении группы людей по росту.

Но бывают случаи, когда приходится говорить о распределении совокупности по нескольким признакам. В настоящей главе мы будем говорить только о распределении по двум признакам, например, о распределении группы людей по росту и объему груди.

Говоря о членах совокупности, или об объектах ее, надо иметь в виду, что иногда приходится подразумевать под словом „член совокупности“ сложный объект, как напр., отец и его старший сын, или брат и сестра. Признаками такого сложного объекта в приведенных примерах могут быть: рост отца и его сына, или брата и сестры.

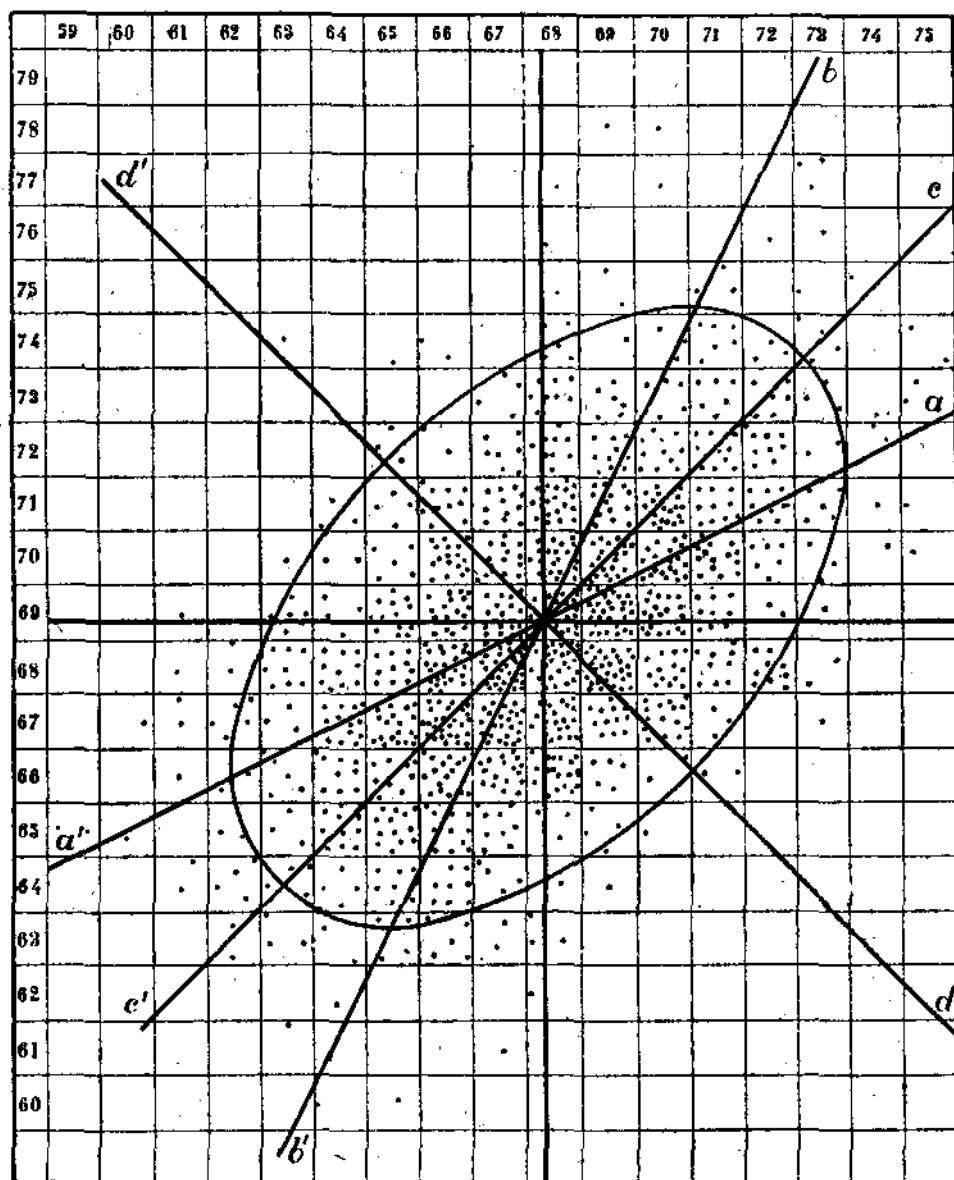
Примем величины двух признаков объекта совокупности x и y за Декартовы координаты точки на плоскости. Тогда каждому объекту совокупности будет соответствовать определенная точка на плоскости. Всем членам совокупности будет соответствовать совокупность точек на плоскости, называемая полем точек. Число точек поля равно числу объектов совокупности; мы его обозначим буквою N и назовем объемом совокупности.

На черт. 12 представлено поле, соответствующее распределению отцов и их сыновей по росту *).

Рассматривая чертеж, мы видим, что он напоминает собою расположение песчинок, насыпанных на горизонтальную плоскость сквозь

*) Этот чертеж составлен по данным распределения, приводимым Пирсоном — K. Pearson and A. Le. On the inheritance in man. Biometrika, Vol. II. Стр. 415. Так как таблица Пирсона дает только число точек поля для каждого из указанных в таблице интервалов, но не расположение точек внутри интервала, то чертеж представляет только приблизительное расположение точек. Подобный же чертеж приведен и у Слудского в сочинении: Теория корреляций. Киев, 1912.

узкое отверстие, помещенное над определенной точкой плоскости. Наибольшее скопление песчинок — около этой точки; по мере удаления от нее песчинки встречаются все реже.



Чертеж 12.

Покроем ту часть плоскости, где встречаются точки поля, сетью прямоугольников со сторонами, параллельными осям координат, длиной равными dx и dy . Площадь каждого прямоугольника будет

равна $dx dy$. Прямоугольник, у которого координаты левого нижнего угла суть x и y , мы будем называть прямоугольником (x, y) .

Число точек поля, лежащих внутри прямоугольника (x, y) , будем обозначать символом n_x^y и назовем его численностью прямоугольника (x, y) *).

Для ясности вводимых обозначений и связанных с ними рассуждений приводим следующую схематическую таблицу на черт. 13.

Числа, стоящие в одном и том же столбце, называются вертикальным строєм. Тот строй, числа которого имеют нижним указателем x , назовем x -м строєм, а сумму его численностей обозначим символом n_x и назовем численностью x -го строя.

В приведенной схематической таблице на черт. 13

численности вертикальных строев подписаны под соответствующими строями. Это суть: $n_1, n_2, n_3, n_4, n_5, n_6$.

Таким же образом числа, стоящие в одной и той же строке, составляют горизонтальный строй. Тот строй, числа которого имеют верхним указателем y , назовем y -м строєм, а сумму чисел этого строя обозначим символом n^y и назовем численностью y -го строя.

На черт. 13 эти числа суть: $n^{(1)}, n^{(2)}, n^{(3)}, n^{(4)}, n^{(5)}$.

Ясно, что сумма численностей вертикальных строев равна сумме численностей горизонтальных строев и равна объему N всей совокупности.

Отношение численности n_x^y прямоугольника (x, y) к его площади $dx dy$ называется частотой для данного прямоугольника.

$x \backslash y$	1	2	3	4	5	6	n^y
5			n_3^5				n^5
4		n_2^4	n_3^4	n_4^4			n^4
3	n_1^3	n_2^3	n_3^3	n_4^3	n_5^3		n^3
2			n_3^2	n_4^2	n_5^2	n_6^2	n^2
1				n_4^1	n_5^1		n^1
n_x	n_1	n_2	n_3	n_4	n_5	n_6	N

$x \backslash y$	1	2	3	4	5	6	n^y
5							
4							
3							
2							
1							
n_x							

Черт. 13.

*) При счете точек, лежащих в прямоугольнике, обыкновенно считают точку, лежащую на стороне, служащей границей между двумя прямоугольниками, за $\frac{1}{2}$ для каждого из этих двух прямоугольников, а точку, лежащую в вершине, где сходятся 4 прямоугольника, за $\frac{1}{4}$ для каждого из этих четырех прямоугольников.

Обозначив ее буквою Z , будем иметь:

$$Z = \frac{n_x^y}{dx dy},$$

откуда:

$$n_x^y = Z dx dy \dots \dots \dots (1)$$

Ясно, что частость для прямоугольника в данном случае имеет такое же значение, как частость для интервала dx в предшествующей главе, в случае распределения по одному признаку.

Величина Z есть функция переменных x и y . Обозначив ее символом $\varphi(x, y)$, можем написать:

$$Z = \varphi(x, y) \dots \dots \dots (2)$$

Эта функция $\varphi(x, y)$ называется функцией распределения по двум признакам.

На прямоугольнике (x, y) , как на основании, строим прямоугольный параллелепипед, имеющий объем, равный численности n_x^y ; высота его вследствие равенства (1) будет равна Z .

Подобное же построение выполним для всех прямоугольников, где встречаются точки поля. Прямоугольники, где нет точек поля (пустые прямоугольники), рассматриваем, как основания параллелепипедов с высотой, равной нулю; следовательно, у таких параллелепипедов верхнее основание совпадает с нижним.

Совокупность всех построенных параллелепипедов составит тело, объем которого равен сумме численностей для всех прямоугольников сети, т.е. он равен объему N совокупности.

Сверху это тело ограничено поверхностью, состоящей из верхних оснований параллелепипедов; это — поверхность ступенчатого вида.

Поступая подобно тому, как мы поступали, говоря о линиях распределения, мы можем предполагать, что полученная ступенчатая поверхность простирается по всем направлениям до бесконечности; для этого, как сказано выше, будем считать каждый пустой прямоугольник за основание параллелепипеда с высотой, равной нулю.

Рассматривая Z , как третью координату, и определяя ее равенством (2), можем сказать, что (2) есть уравнение ступенчатой поверхности.

Для исследования ее свойств построим плавную поверхность, возможно близко подходящую к ступенчатой поверхности.

Координата Z этой плавной поверхности будет приближенным значением функции $\varphi(x, y)$, стоящей в правой части равенства (2). Для избежания излишних обозначений, мы будем предполагать, что приближенная интерполяционная формула выражается уравнением (2).

Это уравнение назовем уравнением поверхности распределения.

На основании сказанного выше, объем, ограничиваемый этой поверхностью, должен равняться N .

Так как поверхность простирается во все стороны до бесконечности, то это свойство выразится формулой:

$$N = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) dx dy \dots \dots \dots (3).$$

Обобщая понятия, введенные в предшествующей главе, мы назовем величины:

$$v_x = \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) x dx dy, \quad v_y = \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) y dx dy \dots \dots (4)$$

моментами 1-го порядка.

Величины:

$$v_{x,x} = \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) x^2 dx dy, \quad v_{x,y} = \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) xy dx dy, \\ v_{y,y} = \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) y^2 dx dy \dots \dots \dots (5)$$

назовем моментами 2-го порядка.

Преобразуем оси координат, перенеся их параллельно самим себе:

$$\left. \begin{aligned} x &= \xi + x', \\ y &= \eta + y', \end{aligned} \right\} \quad \text{откуда: } \left. \begin{aligned} x' &= x - \xi, \\ y' &= y - \eta. \end{aligned} \right\}$$

Обозначая моменты относительно новых осей буквами:

$$v_{x'}, v_{y'}, v_{x',x'}, v_{x',y'}, v_{y',y'},$$

находим:

$$v_{x'} = \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) x dx dy - \frac{\xi}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) dx dy = v_x - \xi \dots (6)$$

$$v_{y'} = \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) y dx dy - \frac{\eta}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) dx dy = v_y - \eta \dots (7)$$

Таким же образом найдем:

$$v_{x',x'} = v_{x,x} - 2\xi v_x + \xi^2, \quad v_{x',y'} = v_{x,y} - \xi v_y - \eta v_x + \xi\eta, \\ v_{y',y'} = v_{y,y} - 2\eta v_y + \eta^2.$$

Выберем ξ и η равными первым моментам:

$$\xi = \nu_x, \eta = \nu_y \dots \dots \dots (8)$$

и назовем точку, определяемую этими равенствами центром распределения, а моменты относительно осей, проходящих через центр, назовем центральными и обозначим их буквами:

$$\mu_1, \mu_2, \mu_{11}, \mu_{12}, \mu_{22}.$$

Из формул (6), (7) и (8) будет следовать:

$$\mu_1 = 0, \mu_2 = 0, \mu_{11} = \nu_{x,x} - \xi^2, \mu_{12} = \nu_{x,y} - \xi\eta, \mu_{22} = \nu_{y,y} - \eta^2. \dots (9)$$

Предполагая, что это преобразование уже выполнено и что после преобразования поверхность распределения выразилась уравнением (2), можем написать:

$$\begin{aligned} \mu_1 &= \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) x \, dx dy = 0, \quad \mu_2 = \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) y \, dx dy = 0, \\ \mu_{11} &= \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) x^2 \, dx dy, \quad \mu_{12} = \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) xy \, dx dy, \\ \mu_{22} &= \frac{1}{N} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) y^2 \, dx dy \dots \dots \dots (10) \end{aligned}$$

Рассмотрим с новой точки зрения ту таблицу, которая схематически изображена на черт. 13.

Возьмем случайным образом один из объектов совокупности и зададимся вопросом, какова вероятность p_x^y , что этот объект соответствует прямоугольнику (x, y) , т. е. что первый признак его лежит между x и $x + dx$, а второй признак — между y и $y + dy$.

Величина этой вероятности а priori нам неизвестна, но мы можем найти ее приближенное значение на основании теоремы, обратной теореме Бернулли, если число N весьма велико.

В самом деле, из числа всех N объектов данной совокупности n_x^y принадлежит к прямоугольнику (x, y) . Если так, то приближенно имеем:

$$p_x^y = \frac{n_x^y}{N} \dots \dots \dots (11)$$

Таким же образом, обозначив буквами p_x и p^y вероятности того, что случайно взятый объект принадлежит к x -вому строю или y -вому строю, будем иметь приближенные равенства:

$$p_x = \frac{n_x}{N}, \quad p^y = \frac{n^y}{N} \dots \dots \dots (12)$$

Далее, если мы отберем те объекты, которые принадлежат к x -вому строю, то можем задаться вопросом, какова вероятность, что случайно взятый объект этой частной совокупности будет принадлежать к прямоугольнику (x, y) . Эту вероятность мы назовем условной вероятностью для x -вого строа и обозначим буквою $(p^y)_x$. На основании того же рассуждения, как и выше, найдем приближенное равенство:

$$(p^y)_x = \frac{n_x^y}{n_x} \dots \dots \dots (13)$$

Наконец, совершенно таким же образом найдем условную вероятность для y -вого строа:

$$(p_x)^y = \frac{n_x^y}{n^y} \dots \dots \dots (14)$$

Событие, состоящее в том, что случайно взятый объект принадлежит к прямоугольнику (x, y) , можно рассматривать, как событие сложное, состоящее из двух простых: 1) взятый объект принадлежит к x -вому строю, 2) в этом строю он принадлежит к прямоугольнику (x, y) . Эти два простых события, вообще говоря, зависимы между собою. Вероятность первого из них есть p_x ; условная вероятность второго есть $(p^y)_x$. По теореме умножения вероятностей будем иметь:

$$p_x^y = p_x \cdot (p^y)_x \dots \dots \dots (15)$$

Справедливость этого равенства проверяется помощью формул (11), (12), (18):

$$p_x^y = \frac{n_x}{N} \cdot \frac{n_x^y}{n_x} = \frac{n_x^y}{N}$$

Подобным же образом, разлагая то же событие иначе, получим равенство:

$$p_x^y = p_y \cdot (p_x)^y \dots \dots \dots (16)$$

Так как обе величины, (15) и (16), выражают одну и ту же вероятность p_x^y , то мы будем иметь тождество:

$$p_x^y = p_x (p^y)_x = p_y (p_x)^y, \dots \dots \dots (17)$$

которое легко проверяется помощью формул (11), (12), (18), (14).

Равенство (17) окажется весьма пригодным в следующем параграфе, но здесь мы остановим внимание на одной его особенности.

Вообще говоря, те два простых события, на которые мы разложили сложное событие, имеющее вероятность p_x^y , зависимы между собою, т.-е. величина признака x зависит от величины y . Но в тех вопросах, к которым приходится применять настоящие рассуждения,

эта зависимость своеобразного характера. Для пояснения сказанного обратимся к упомянутым выше двум примерам.

Рост человека и объем его груди находятся в зависимости между собой: в большинстве случаев человек высокого роста обладает и большим объемом груди, но это не исключает возможности встретить человека с большим ростом и узкогрудого. Так же точно, обыкновенно у отца высокого роста бывает и сын роста большого; но это не исключает возможности случая, когда у отца высокого роста сын низкорослый.

На приведенных примерах видно, что между признаками x и y зависимость существует, но эта зависимость не функционального характера: по величине x нельзя определить величину y .

Зависимость такого рода, к которой приводит изучение таблиц распределения, называется корреляция.

Задачей настоящей главы служит изучение особенностей корреляционной зависимости и нахождение способов оценивать меру этой зависимости.

Рассуждения, которые мы делали при выводе формулы (14), не исключают возможности такого случая, когда x и y друг от друга не зависят. В этих случаях вероятность $(p_x)^y$ не зависит от y ; она равна вероятности x -вого строя p_x ; равенство (17) примет вид:

$$p_x^y = p_x p^y \dots \dots \dots (17')$$

§ 2. Нормальное распределение по двум признакам.

В параграфе 6 предшествующей главы мы установили понятие о нормальном распределении по одному признаку.

Мы обозначили буквою N объем совокупности, величину признака буквою x , среднее значение этого признака буквою ξ , среднее квадратическое отклонение x от этого среднего значения ξ через σ и длину интервала через dx . Тогда функция распределения выражалась формулой:

$$y = \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\xi)^2}.$$

Разделив эту величину на объем N и умножив на длину интервала dx , найдем вероятность p того, что случайно взятый объект совокупности будет принадлежать к интервалу, лежащему между x и $x + dx$:

$$p = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\xi)^2} dx,$$

*) См. формулу (33) предшествующей главы.

или:

$$p = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\xi)^2 - \lg \sigma} dx \quad \dots \quad (18)$$

Обобщая понятие о нормальном распределении на случай двух признаков, естественно сказать: нормальным распределением по двум признакам мы назовем такое распределение, где вся совокупность распределяется по строкам как горизонтальным, так и вертикальным, по нормальному закону, и каждый строй внутри себя составлен по нормальному закону распределения.

В таком случае вероятности, которые выше мы обозначили символами:

$$p_x, p_y, (p^y)_x, (p_x)^y,$$

выразятся по формуле (18) следующим образом:

$$\left. \begin{aligned} p_x &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma_1^2}(x-\xi)^2 - \lg \sigma_1} dx, \\ p_y &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma_2^2}(y-\eta)^2 - \lg \sigma_2} dy, \\ (p^y)_x &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\Sigma_2^2}(y-Y)^2 - \lg \Sigma_2} dy, \\ (p_x)^y &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\Sigma_1^2}(x-X)^2 - \lg \Sigma_1} dx, \end{aligned} \right\} \dots \quad (19)$$

где ξ , η , σ_1 и σ_2 суть постоянные, величины Y и Σ_2 суть, вообще говоря, функции от x , но постоянные относительно y , а X и Σ_1 суть вообще говоря, функции от y , но постоянные относительно x .

Величина ξ есть среднее значение признака x при распределении всей совокупности по одному признаку x , т. е. при распределении всей совокупности по вертикальным строкам. Таким же образом η есть среднее значение y при распределении всей совокупности по горизонтальным строкам.

Внося выражения (19) в равенство (17), находим:

$$\begin{aligned} p^y_x &= \frac{1}{2\pi} e^{-\frac{1}{2\sigma_1^2}(x-\xi)^2 - \lg \sigma_1 - \frac{1}{2\Sigma_2^2}(y-Y)^2 - \lg \Sigma_2} dy dx = \\ &= \frac{1}{2\pi} e^{-\frac{1}{2\sigma_1^2}(y-\eta)^2 - \lg \sigma_2 - \frac{1}{2\Sigma_1^2}(x-X)^2 - \lg \Sigma_1} dx dy \quad \dots \quad (20) \end{aligned}$$

откуда следует:

$$\frac{1}{2\sigma_1^2} (x - \xi)^2 + \lg \Sigma_1 - \frac{1}{2\sigma_2^2} (y - Y)^2 + \lg \Sigma_2 = \frac{1}{2\sigma_2^2} (y - \eta)^2 + \lg \sigma_2 - \\ - \frac{1}{2\sigma_1^2} (x - X)^2 + \lg \Sigma_1 \dots \dots \dots (21)$$

Так как это равенство должно быть тождеством, справедливым при всяких значениях x и y , то мы выводим из него ряд следствий:

1) левая часть этого равенства, так же как и правая, должны быть целым многочленом 2-й степени как относительно x , так и относительно y ;

2) величины Σ_1 и Σ_2 могут быть только постоянными (иначе или $\lg \Sigma_1$ и $\lg \Sigma_2$, или сами величины Σ_1 и Σ_2 были бы трансцендентными).

3) величины X и Y суть или линейные функции величин y и x или постоянные.

Отсюда следует, что как левая, так и тождественная с нею правая часть равенства (21) должна иметь вид:

$$-\frac{1}{2} (ax^2 + 2bxy + cy^2 + 2dx + 2ey + f), \dots \dots (22)$$

где a, b, c, d, e, f суть постоянные числа.

Равенство (20) примет вид:

$$p_x^y = \frac{1}{2\pi} e^{-\frac{1}{2} (ax^2 + 2bxy + cy^2 + 2dx + 2ey + f)} dx dy \dots (23)$$

Приняв во внимание равенства (11), (1) и (2), находим:

$$Z = \varphi(x, y) = \frac{N}{2\pi} e^{-\frac{1}{2} (ax^2 + 2bxy + cy^2 + 2dx + 2ey + f)} \dots (24)$$

Такова функция распределения по двум признакам в нормальном случае.

Она получена впервые французским математиком Бравэ*) из иных соображений. Естественно назвать ее формулой Бравэ.

Найдем величины x и y , при которых частость, выражаемая формулой (24), будет наибольшая. Величины x и y , найденные из этого условия, будут координатами точки наибольшей плотности распределения.

Для нахождения координат этой точки приравняем нулю частные производные по x и по y величины (24). Опуская множители, не-

*) Bravais: Analyse mathématique sur les probabilités des erreurs de situation d'un point. Mémoires présentés par divers savants. T. IX, 1846.

равные нулю, находим:

$$\left. \begin{aligned} ax + by + d &= 0, \\ bx + cy + e &= 0. \end{aligned} \right\} \dots \dots \dots (25)$$

Решения этих уравнений обозначим буквами x_0 и y_0 . Перенесем начало координат, делая преобразование:

$$\left. \begin{aligned} x &= x_0 + x', \\ y &= y_0 + y'. \end{aligned} \right\} \dots \dots \dots (26).$$

Получим:

$$Z = \frac{N}{2\pi} e^{-\frac{1}{2}(ax'^2 + 2bx'y' + cy'^2 + f')},$$

где:

$$f' = ax_0^2 + 2bx_0y_0 + cy_0^2 + 2dx_0 + 2ey_0 + f.$$

Иначе можно написать:

$$Z = Ce^{-\frac{1}{2}(ax'^2 + 2bx'y' + cy'^2)}, \dots \dots \dots (27)$$

где C постоянное.

На основании формул (4) моменты 1-го порядка v_x и v_y выражаются так:

$$\begin{aligned} v_x &= C \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax'^2 + 2bx'y' + cy'^2)} x' dx' dy', \\ v_y &= C \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax'^2 + 2bx'y' + cy'^2)} y' dx' dy'. \end{aligned}$$

В каждом из этих интегралов первый множитель подынтегральной функции четного порядка относительно переменных x' и y' , а второй множитель — функция нечетная; следовательно, все произведение — функция нечетная. Нижние пределы в обоих интегралах разнятся только знаком от верхнего. В таких случаях интеграл равен нулю, а потому имеем:

$$v_x = 0, \quad v_y = 0.$$

Это значит, что, совершив преобразование (26), мы перенесли начало координат в центр распределения. Координаты x_0 и y_0 определяемые уравнениями (25), суть координаты центра ξ , η . Величины x' и y' , определяемые равенствами (26), или, что то же:

$$\begin{aligned} x' &= x - \xi, \\ y' &= y - \eta. \end{aligned}$$

суть отклонения признаков от наивероятнейших значений ξ и η этого признака.

Выше для избежания излишних обозначений мы условились, перенести начало координат в центр, обозначать новые координаты прежними буквами x и y .

Если так, то функция распределения выражается формулой:

$$Z = \varphi(x, y) = C e^{-\frac{1}{2}(ax^2 + 2bxy + cy^2)} \quad (28)$$

при чем начало координат находится в центре распределения, а x и y суть отклонения признаков от их нормальных значений.

Равенства (8) и (10) принимают вид:

$$N = C \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2 + 2bxy + cy^2)} dx dy, \quad (29)$$

$$\mu_1 = 0, \quad \mu_2 = 0,$$

$$\left. \begin{aligned} \mu_{11} &= C \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2 + 2bxy + cy^2)} x^2 dx dy, \\ \mu_{12} &= C \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2 + 2bxy + cy^2)} xy dx dy, \\ \mu_{22} &= C \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2 + 2bxy + cy^2)} y^2 dx dy. \end{aligned} \right\} \quad (30)$$

Займемся вычислением интеграла:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2 + 2bxy + cy^2)} dy dx \quad (31)$$

Преобразуем сначала выражение, стоящее в показателе:

$$ax^2 + 2bxy + cy^2 = \left(\sqrt{a}x + \frac{b}{\sqrt{a}}y \right)^2 + \left(c - \frac{b^2}{a} \right) y^2.$$

Выражение (31) примет вид:

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(c - \frac{b^2}{a}\right)y^2} dy \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\sqrt{a}x + \frac{b}{\sqrt{a}}y\right)^2} dx.$$

Преобразуя в первом интеграле переменное x подстановкой:

$$\sqrt{a}x + \frac{b}{\sqrt{a}}y = \sqrt{2}t,$$

приводим выражение (31) к такому виду:

$$\begin{aligned} & \sqrt{\frac{2}{a}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2} \left(\frac{ac-b^2}{a} \right) y^2} dy \int_{-\infty}^{+\infty} e^{-t^2} dt = \\ & = \sqrt{\frac{2\pi}{a}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2} \left(\frac{ac-b^2}{a} \right) y^2} dy. \end{aligned}$$

Величина:

$$\frac{1}{2} \left(\frac{ac-b^2}{a} \right)$$

должна быть положительною, потому что иначе интеграл (31), а следовательно, и величина N , равнялись бы бесконечности.

Положив:

$$\sqrt{\frac{ac-b^2}{2a}} y = u,$$

приводим интеграл (31) к такому виду:

$$\frac{2\sqrt{\pi}}{\sqrt{ac-b^2}} \int_{-\infty}^{+\infty} e^{-u^2} du = \frac{2\pi}{\sqrt{ac-b^2}}.$$

Итак:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2+2bxy+cy^2)} dx dy = \frac{2\pi}{\sqrt{ac-b^2}} \dots (32)$$

Дифференцируя это равенство по a , b , c , находим:

$$\left. \begin{aligned} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2+2bxy+cy^2)} x^2 dx dy &= \frac{2\pi c}{(ac-b^2)^{\frac{3}{2}}}, \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2+2bxy+cy^2)} xy dx dy &= \frac{-2\pi b}{(ac-b^2)^{\frac{3}{2}}}, \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2+2bxy+cy^2)} y^2 dx dy &= \frac{2\pi a}{(ac-b^2)^{\frac{3}{2}}} \end{aligned} \right\} \dots (33)$$

Из равенств (29) и (32) следует:

$$C = \frac{N\sqrt{ac-b^2}}{2\pi} \dots (34)$$

Из равенств (30) и (33) следует;

$$\mu_{11} = \frac{c}{ac - b^2}, \quad \mu_{12} = \frac{-b}{ac - b^2}, \quad \mu_{22} = \frac{a}{ac - b^2},$$

откуда:

$$\mu_{11} \mu_{22} - \mu_{12}^2 = \frac{1}{ac - b^2}.$$

Отсюда далее:

$$a = \frac{\mu_{22}}{\mu_{11}\mu_{22} - \mu_{12}^2}, \quad b = \frac{-\mu_{12}}{\mu_{11}\mu_{22} - \mu_{12}^2}, \quad c = \frac{\mu_{11}}{\mu_{11}\mu_{22} - \mu_{12}^2} \quad (35)$$

Обозначая буквами σ_1 и σ_2 средние квадратические отклонения x и y от центра, имеем:

$$\sigma_1 = \sqrt{\mu_{11}}, \quad \sigma_2 = \sqrt{\mu_{22}} \quad (36)$$

Кроме того, вводим обозначение:

$$r = \frac{\mu_{12}}{\sigma_1 \sigma_2} \quad (37)$$

Смысл величины r скоро выяснится.

Формулы (34) и (35) примут вид:

$$C = \frac{N}{2\pi \sigma_1 \sigma_2 \sqrt{1-r^2}} \quad (38)$$

$$a = \frac{1}{\sigma_1^2 \sqrt{1-r^2}}, \quad b = -\frac{r}{\sigma_1 \sigma_2 \sqrt{1-r^2}}, \quad c = \frac{1}{\sigma_2^2 \sqrt{1-r^2}} \quad (39)$$

Формула (28) примет вид:

$$Z = \frac{N}{2\pi \sigma_1 \sigma_2 \sqrt{1-r^2}} \cdot e^{-\frac{1}{2(1-r^2)} \left\{ \frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1 \sigma_2} + \frac{y^2}{\sigma_2^2} \right\}} \quad (40)$$

Такова формула нормального распределения по двум признакам в окончательной форме через центральные моменты для данной совокупности.

§ 3. Коэффициент корреляции.

Вернемся к формулам (23) и (24), припомним, что начало координат перенесено нами в центр, и мы условились новые координаты обозначать прежними буквами x и y , чтобы избежать излишних обозначений. Поэтому в координатах центральное уравнение (23) примет вид:

$$p_x^y = \frac{Z}{N} dx dy,$$

или

$$p_x^y = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left\{\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right\}} dx dy, \dots (23')$$

а уравнения (19) примут вид:

$$p_x = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_1^2}} dx, \dots (19')$$

$$p_y = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{y^2}{2\sigma_2^2}} dy \dots (19'')$$

Если переменные x и y не зависимы между собою, то по теореме об умножении вероятностей, вероятность p_x^y должна равняться произведению величин: p_x и p_y , как мы уже говорили в предшествующем параграфе *). Но произведение $p_x p_y$ равно:

$$p_x p_y = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{x^2}{2\sigma_1^2} - \frac{y^2}{2\sigma_2^2}} dx dy.$$

Сравнивая это выражение с выражением (23'), видим, что они равны между собою в том случае, и только в том случае, когда r равно нулю.

Итак, в случае нормального распределения условие $r=0$ есть необходимый и достаточный признак независимости между собою величин x и y .

Если r не равно нулю, то между признаками x и y зависимость существует.

Из формулы (40) видно, что r по абсолютной величине не может быть больше единицы: иначе величина Z была бы мнима. Итак, абсолютная величина r содержится между 0 и 1. Мы уже видели, что равенство $r=0$ есть признак отсутствия всякой связи между x и y . Теперь посмотрим, какова особенность нормального распределения при $r=\pm 1$.

В этом случае функция (40) обращается в 0 при всех значениях x и y , кроме тех, которые обращают в нуль переменный множитель в показателе степени формулы (40), т.е. Z обращается в нуль при всех значениях x и y , за исключением удовлетворяющих уравнению:

$$\frac{x^2}{\sigma_1^2} \mp \frac{2xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2} = 0,$$

или

$$\left(\frac{x}{\sigma_1} \mp \frac{y}{\sigma_2}\right)^2 = 0,$$

*) См. форм. (17').

откуда:

$$y = \pm \frac{\sigma_2}{\sigma_1} x \dots \dots \dots (41)$$

Это значит, что при $r = \pm 1$ переменные x и y связаны между собою функционально и притом линейною зависимостью (41).

Обратно, допустим, что x и y связаны между собою функционально, т.-е. некоторым уравнением:

$$y = f(x). \quad (41')$$

В таком случае выражение (40) должно быть нулем при всех значениях x и y , кроме тех, которые удовлетворяют уравнению (41').

Из формулы (40) видно, что она обратится в нуль только тогда, когда показатель степени равен $-\infty$; а это не может случиться ни при каких значениях x и y , если r^2 не равно 1. В этом случае, как мы только-что видели, x и y связаны между собою уравнением (41). Отсюда следует, что уравнение (41') должно иметь вид (41).

Это значит, что при нормальном распределении условие, необходимое и достаточное для существования функциональной зависимости между x и y , есть $r = \pm 1$, и при этом функциональная зависимость непременно линейная вида:

$$y = \pm \frac{\sigma_2}{\sigma_1} x^* \dots \dots \dots (41'')$$

Мы рассмотрели крайние значения: когда $r = 0$ и когда $r = \pm 1$ и видели, что в первом случае между x и y никакой зависимости нет, а во втором случае x и y связаны между собою функционально. Отсюда следует, что при r^2 , лежащем между 0 и 1, величины x и y связаны между собою, но не функционально, а зависимою особого рода, о которой мы уже говорили выше и которая называется корреляция. Естественно, что величина r называется коэффициент корреляции. Вначале она была названа функцией Гальтона в честь Гальтона, впервые раскрывшего значение этой величины.

§ 4. Прямая регрессии, коэффициент регрессии.

В настоящем параграфе мы будем говорить о тех геометрических представлениях, о которых говорили в параграфе 1, и в частности будем иметь в виду, как пример, чертеж 12.

¹⁾ Необходимо заметить, что полученный вывод справедлив только для нормального распределения. Величины x и y вообще могут быть связаны всевозможными равенствами, даже отличными от (41), но не при нормальном распределении.

Мы видим, что на этом чертеже точки поля расположены наиболее плотно около некоторой определенной точки; такую точку назовем точкой наибольшей плотности поля или модальной точкой. Из формулы распределения:

$$Z = \frac{N}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left\{\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right\}} \dots (40)$$

видно, что точка с наибольшей величиной Z соответствует значениям x и y , равным нулю, т.е. точка наибольшей плотности есть центр распределения, принятый за начало координат.

На черт. 12 мы покрыли поле сетью прямоугольников и обозначили символом p_x^y вероятность того, что случайно взятый объект совокупности принадлежит к прямоугольнику (x, y) ; эта вероятность выражается формулой:

$$p_x^y = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left\{\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right\}} dx dy \dots (23')$$

Символом p_x мы обозначили вероятность того, что случайно взятый объект совокупности принадлежит к x -вому строю; она выражается формулой (19'), а именно:

$$p_x = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_1^2}} dx \dots (19')$$

Формулу (23') можно представить в таком виде:

$$p_x^y = \left\{ \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_1^2}} dx \right\} \left\{ \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2\sigma_2^2(1-r^2)}\left(y - \frac{r\sigma_2}{\sigma_1}x\right)^2} dy \right\} \dots (42)$$

Сравним ее с найденной выше формулой

$$p_x^y = p_x \cdot (p^y)_x \dots (15)$$

Принимая во внимание формулу (19'), мы видим, что первый множитель правой части равенства (42) есть p_x , следовательно, второй множитель есть $(p^y)_x$:

$$(p^y)_x = \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2\sigma_2^2(1-r^2)}\left(y - \frac{r\sigma_2}{\sigma_1}x\right)^2} dy \dots (43)$$

Эта формула выражает вероятность того, что объект, взятый случайно в x -вом строю, будет принадлежать в этом строю к прямоугольнику (x, y) .

Отсюда следует, что функция распределения объектов x -вого строя выражается формулой:

$$\frac{n_x}{\sqrt{2\pi\sigma_2}\sqrt{1-r^2}} e^{-\frac{1}{2\sigma_2^2(1-r^2)}\left(y-\frac{r\sigma_2}{\sigma_1}x\right)^2} \dots\dots\dots (44)$$

Это — нормальная формула распределения объектов по одному признаку y , в данном случае — формула распределения объектов x -вого строя.

Обозначив буквою Σ_2 среднее квадратическое уклонение для этого распределения, находим:

$$\Sigma_2 = \sigma_2 \sqrt{1-r^2} \dots\dots\dots (45)$$

Мода этого распределения выражается формулой:

$$Y = \frac{r\sigma_2}{\sigma_1} x \dots\dots\dots (46)$$

Это — ордината точки наибольшей плотности в x -вом строе, иначе, величина Y есть среднее значение признака y для объектов x -вого строя.

Средняя величина признака y для x -вого строя представляет собою нормальную величину признака y для объектов, имеющих одинаковую величину признака x . Так, в том примере, для которого составлен черт. 12, величина Y есть нормальный рост сына, отец которого имеет рост x .

Среднее квадратическое уклонение Σ_2 характеризует собою рассеяние точек x -вого строя относительно средней точки, т.е. центра этого строя. Мы видим из формулы (45), что это рассеяние для всех вертикальных строев одно и то же, оно не зависит от x . Но оно зависит от величины r . Наибольшее его значение соответствует $r=0$, т.е. когда между x и y корреляции не существует; в этом случае Σ_2 равно среднему квадратическому уклонению σ_2 для y во всем поле. Когда $r=\pm 1$, величина Σ_2 равна 0; это и понятно: при $r=\pm 1$, как мы знаем, существует функциональная зависимость между x и y ; поэтому все точки x -вого строя совпадают с центром; рассеяния никакого нет.

Рассматривая формулу (46), мы видим, что Y есть линейная функция от x .

Это значит, что геометрическое место центров всех вертикальных строев есть прямая линия, выражаемая уравнением (46). Прямая (46) проходит через начало координат, т.е. через центр всего распределения, и наклонена к оси x под углом, тангенс которого равен $\frac{r\sigma_2}{\sigma_1}$. Обозначив угловой коэффи-

циент прямой (46) символом $\rho_{y(x)}$ будем иметь:

$$\rho_{y(x)} = \frac{r\sigma_2}{\sigma_1} \dots \dots \dots (47)$$

Прямая линия (46) называется линией регрессии, а величина (47) коэффициентом регрессии y относительно x .

Как уже упомянуто выше, величина Σ_2 характеризует собою степень рассеяния точек каждого строя относительно соответствующего центра, который лежит на линии регрессии; следовательно, можно сказать, что Σ_2 характеризует степень рассеяния точек поля относительно линии регрессии. Когда $r=0$, это рассеяние наибольшее и равно σ_2 ; когда $r=\pm 1$, величина Σ_2 равна 0, все точки поля сдвинуты на линию регрессии, и эта линия устанавливает функциональную зависимость между x и y ; при $r=\pm 1$ линия регрессии выражается уже рассмотренным выше уравнением (41).

При значении r , промежуточном между 0 и ± 1 , величина Σ_2 имеет промежуточное значение между 0 и σ_2 . Точки поля тем теснее группируются около линии регрессии, чем величина r ближе к 1 по абсолютному значению. Отсюда понятно, почему величина r получила название коэффициента корреляции.

Рассматривая чертеж 12, мы видим, что действительно в каждом строе точки поля вблизи краев строя весьма редки, по мере приближения к линии регрессии они делаются все чаще; наибольшей плотности они достигают вблизи линии регрессии, которая обозначена на чертеже 12 буквами aa' .

Величина Y есть нормальная величина признака y для объектов, имеющих данную величину признака x . Объекты с признаком, отличным от Y представляют собою случайные отклонения от нормы.

Так как мы поместили начало координат в центре распределения, то величины x и y отсчитываются от средних значений признаков для всего поля; это суть отклонения от средних значений признаков для всей совокупности.

В данном примере x есть отклонение роста данного отца от среднего роста всех отцов, вошедших в изучаемую совокупность. Таким же образом y есть отклонение роста данного сына от среднего роста всех сыновей.

Уравнение (46) показывает, что у отцов, представляющих отклонение в росте, равное x , сыновья в среднем будут представлять отклонение в росте равное Y . Величина Y есть линейная функция x , определяемая уравнением (46). Для того материала, для которого построен чертеж 12, величина коэффициента регрессии $\rho_{y(x)}$ равна 0,516. Следовательно, у отцов, представляющих отклонение от нормы, равное x , сыновья в среднем представляют отклонение от нормы в том же

направлении, но не на полную величину x , а только на некоторую долю этой величины, а именно, на 0,516 x .

Природа как бы стремится сгладить отклонение от нормы, уменьшая влияние наследственности. Отсюда и произошло название регрессия.

Из формулы (47) видно, что коэффициент регрессии имеет тот же знак, как и коэффициент корреляции r , так как σ_1 и σ_2 существенно положительны.

Поэтому при положительном r , и $\rho_{y(x)}$ тоже положительно; угловой коэффициент прямой регрессии (46) положителен, прямая регрессии составляет с осью x угол острый. Это значит, что с увеличением x , и Y тоже увеличивается.

Такая корреляция называется положительною.

Обратно, когда r отрицательно, и величина $\rho_{y(x)}$ тоже отрицательна; следовательно, с увеличением x величина Y уменьшается. Такая корреляция называется отрицательною.

Возьмем величину Y для x -вого строя и умножим квадрат ее на численность n_x этого строя; получим выражение:

$$Y^2 n_x.$$

Квадратный корень из среднего арифметического величин такого вида для всех строев обозначим буквою σ_x ; она определится равенством:

$$\sigma_x^2 = \frac{1}{N} \sum Y^2 n_x \dots \dots \dots (48)$$

Так как Y есть ордината точки прямой регрессии, то естественно назвать σ_x средним квадратическим отклонением прямой регрессии от оси x .

Для ее вычисления воспользуемся равенствами (12), (19') и (46). Из этих равенств следует:

$$\frac{1}{N} Y^2 n_x = Y^2 p_x = \left(\frac{r \sigma_2}{\sigma_1} x \right)^2 \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_1^2}} dx.$$

Давая в этом выражении величине x всевозможные значения от $-\infty$ до $+\infty$ и составляя сумму полученных результатов, заметим, что эта сумма приближенно выразится интегралом:

$$\sigma_x^2 = \frac{r^2 \sigma_2^2}{\sigma_1^2} \cdot \frac{1}{\sigma_1 \sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma_1^2}} x^2 dx.$$

Совершая интеграцию по частям, находим:

$$\frac{1}{\sigma_1 \sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma_1^2}} x^2 dx = \sigma_1^2.$$

Поэтому величина σ_x^2 представится в таком виде:

$$\sigma_x^2 = r^2 \sigma_y^2,$$

откуда, обозначая абсолютную величину r символом $|r|$, находим:

$$|r| = \frac{\sigma_x}{\sigma_y} \dots \dots \dots (49)$$

Следовательно, коэффициент корреляции по абсолютной величине равен отношению среднего квадратического отклонения для прямой регрессии к среднему квадратическому отклонению признака y для всей совокупности.

В предшествующих рассуждениях мы рассматривали строй, параллельный оси y . Совершенно так же мы можем поступить со строем, параллельным оси x .

Для такого y -вого строя найдем следующее: среднее квадратическое отклонение в распределении объектов во всех строях одинаково и выражается формулой:

$$\Sigma_1 = \sigma_1 \sqrt{1 - r^2} \dots \dots \dots (45')$$

Среднее значение x для y -вого строя выражается формулой:

$$x = \frac{r \sigma_1}{\sigma_2} y \dots \dots \dots (46')$$

Отсюда следует, что центры распределения во всех горизонтальных строях лежат на одной прямой, выражаемой уравнением (46') и проходящей через центр распределения. Эта прямая называется прямой регрессии: на чертеже 12 она обозначена буквами bb' . Коэффициент ее:

$$\rho_{x(y)} = \frac{r \sigma_1}{\sigma_2} \dots \dots \dots (47')$$

называется коэффициентом регрессии x относительно y .

Перемножив равенства (47) и (47'), находим:

$$\rho_{y(x)} \cdot \rho_{x(y)} = r^2.$$

Это значит, что коэффициент корреляции r есть среднее-пропорциональная между $\rho_{y(x)}$ и $\rho_{x(y)}$.

Изменим масштабы, в которых измеряются признаки. Пусть после изменения масштабов x и y заменятся через hx и ky . Так как численности в прямоугольниках остаются прежними, то из формул (10) будет следовать, что μ_{11} , μ_{12} , μ_{22} заменятся величинами:

$$h^2 \mu_{11}, \quad h k \mu_{12}, \quad k^2 \mu_{22}.$$

Отсюда, из формул (36) и (37) следует, что σ_1 и σ_2 заменятся величинами: $h \sigma_1$ и $k \sigma_2$, а r останется без изменения.

Этот результат имеет большое значение при нахождении величины γ на практике; так как величина r не зависит от масштабов по осям координат, то мы можем выбрать эти масштабы по своему желанию в целях упрощения вычислений при нахождении моментов.

§ 5. Корреляционный эллипс.

Корреляционным эллипсом называется эллипс, выражаемый уравнением:

$$\frac{1}{1-r^2} \left\{ \frac{x^2}{\sigma_1^2} - \frac{2 rxy}{\sigma_1 \sigma_2} + \frac{y^2}{\sigma_2^2} \right\} = \chi^2, \dots \dots \dots (50)$$

где χ величина постоянная, которую мы рассматриваем, как параметр.

Так как функция распределения выражается формулой (40), а численность n_x^y прямоугольника (x, y) связана с величиной этой функции равенством:

$$n_x^y = Z \, dx dy,$$

то мы приходим к выводу, что во всех прямоугольниках, пересекаемых эллипсом (50), численность одинакова и равна:

$$\frac{N}{2 \pi \sigma_1 \sigma_2 \sqrt{1-r^2}} e^{-\frac{1}{2} \chi^2} dx dy \dots \dots \dots (51)$$

Обозначим формулой:

$$\frac{N}{2 \pi \sigma_1 \sigma_2 \sqrt{1-r^2}} \int \int_{\chi} e^{-\frac{1}{2} \chi^2} \left\{ \frac{x^2}{\sigma_1^2} - \frac{2 rxy}{\sigma_1 \sigma_2} + \frac{y^2}{\sigma_2^2} \right\} dx dy \dots \dots \dots (52)$$

двукратный интеграл, распространенный на площадь, лежащую вне эллипса (50).

В этой формуле (52) нижний предел χ может быть каким угодно положительным числом. Если положим $\chi = 0$, то выражение (52) примет вид:

$$\frac{N}{2 \pi \sigma_1 \sigma_2 \sqrt{1-r^2}} \int \int_0 e^{-\frac{1}{2} \chi^2} \left\{ \frac{x^2}{\sigma_1^2} - \frac{2 rxy}{\sigma_1 \sigma_2} + \frac{y^2}{\sigma_2^2} \right\} dx dy \dots \dots \dots (53)$$

и будет выражать интеграл, распространенный на всю плоскость.

Чтобы уяснить себе смысл выражений (52) и (53), составим сначала некоторую площадь s , состоящую из целого числа прямоугольников сети, сплошь наполняющих эту площадь. Число точек поля, находящихся внутри площади, выразится формулой:

$$\sum n_x^y dx dy,$$

где сумма распространена на все прямоугольники, входящие в состав s .

Интеграл (52) есть предел суммы только-что рассмотренного вида и распространен на всю часть плоскости, лежащую вне эллипса (50); следовательно, он выражает собою приближенно число точек поля, лежащих вне эллипса (50).

В таком случае (53) должно приближенно выражать число всех точек поля, которое равно N .

Если так, то выражение:

$$P_{\chi} = \frac{\int_{-\chi}^{+\infty} \iint e^{-\frac{1}{2(1-r^2)} \left\{ \frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2} \right\}} dx dy}{\int_0^{+\infty} \iint e^{-\frac{1}{2(1-r^2)} \left\{ \frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2} \right\}} dx dy} \dots \dots \dots (54)$$

есть приближенная величина отношения числа точек, лежащих вне эллипса (50), к числу всех точек поля. Иными словами, P_{χ} приближенно дает вероятность того, что случайно взятый объект совокупности будет соответствовать точке, лежащей вне эллипса (50).

Величины (52) и (53) мы вычислим несколько искусственным способом.

Будем давать в уравнении (50) параметру χ значения, разнящиеся друг от друга на бесконечно-малую величину $d\chi$, начиная от значения параметра χ , равного 0, до χ , равного $+\infty$. Получатся эллипсы, которые разрежут всю плоскость на бесконечно-узкие эллиптические кольца.

Одним из таких колец служит площадь между двумя эллипсами, из которых один имеет параметр, равный χ , а другой $\chi + d\chi$. Найдем его площадь.

Эллипсы, соответствующие различным значениям параметра χ , имеют полуоси, пропорциональные величинам этого параметра. В таком случае площади их относятся между собою, как квадраты значений параметра *).

Обозначим буквою s площадь эллипса, соответствующего значению параметра, равному 1. Можем сказать, что эллипс, у которого значение параметра есть χ , имеет площадь, равную $s\chi^2$, а эллипс, у которого параметр равен $\chi + d\chi$, имеет площадь $s(\chi + d\chi)^2$.

Эллиптическое кольцо между этими эллипсами будет иметь площадь, равную:

$$s(\chi + d\chi)^2 - s\chi^2,$$

*) Площадь эллипса с полуосями a и b выражается формулою: πab .

или, опуская бесконечно-малые высших порядков:

$$2 s \chi d\chi \dots \dots \dots (55).$$

Рассматривая интеграл (52), как предел суммы, мы сначала отберем те слагаемые, которые относятся к точкам одного и того же кольца, и составим их сумму, а затем сложим все те слагаемые, которые относятся к суммам, лежащим вне эллипса (50).

Так как для всех точек одного и того же эллипса величина χ постоянна, то постоянно и выражение:

$$-\frac{1}{2(1-r^2)} \left\{ \frac{x^2}{\sigma_1^2} - \frac{2 rxy}{\sigma_1 \sigma_2} + \frac{y^2}{\sigma_2^2} \right\}$$

во всех точках этого кольца; оно равно: $-\frac{1}{2} \chi^2$.

Так как ширина кольца на всем его протяжении бесконечно мала, то значение подынтегральной функции можно считать одинаковым во всех точках кольца и вывести величину:

$$\frac{N}{2 \pi \sigma_1 \sigma_2 \sqrt{1-r^2}} e^{-\frac{1}{2} \chi^2}$$

за скобку. Величина в скобках будет равна площади кольца и выразится формулой (55).

Следовательно, сумма всех слагаемых, относящихся к одному и тому же кольцу, выразится так:

$$\frac{2 s N}{2 \pi \sigma_1 \sigma_2 \sqrt{1-r^2}} e^{-\frac{1}{2} \chi^2} \chi d\chi \dots \dots \dots (56)$$

Интеграл (52) равен пределу суммы выражений вида (56) от нижнего предела χ до $+\infty$.

Следовательно, интеграл (52) выразится так:

$$\frac{s N}{\pi \sigma_1 \sigma_2 \sqrt{1-r^2}} \int_{\chi}^{+\infty} e^{-\frac{1}{2} \chi^2} \chi d\chi \dots \dots \dots (57)$$

Положив нижний предел равным 0, найдем величину интеграла (53):

$$\frac{s N}{\pi \sigma_1 \sigma_2 \sqrt{1-r^2}} \int_0^{+\infty} e^{-\frac{1}{2} \chi^2} \chi d\chi \dots \dots \dots (58)$$

Величина (54) представится, как отношение величин (57) и (58)

$$P_z = \frac{\int_0^{+\infty} e^{-\frac{1}{2}\chi^2} \chi d\chi}{\int_0^{+\infty} e^{-\frac{1}{2}\chi^2} \chi d\chi} = e^{-\frac{1}{2}\chi^2} \dots \dots \dots (59)$$

Эта формула дает приближенное выражение вероятности P_z , что случайно-взятый объект совокупности изобразится точкой, лежащей вне корреляционного эллипса (50).

Эллипс на черт. 12 построен для $\chi = 2$. Полагая $\chi = 2$ в формуле (59), находим:

$$P_z = 0,13534.$$

Умножив на число N , равное в данном случае 1078, находим:

$$NP_z = 145,896.$$

Следовательно, приблизительно 146 точек поля из числа 1078 должны лежать вне начерченного эллипса.

§ 6. Практические приемы вычисления коэффициентов корреляции и регрессии.

Пусть дана корреляционная таблица для некоторого распределения и пусть известно, что это распределение нормальное. Спрашивается, как вычислить для него коэффициенты корреляции и регрессии.

Для ясности рассуждений будем предполагать, что данная таблица — та самая, которая схематически изображена на черт. 13.

Припомним те приемы вычисления моментов, которые изложены в параграфе 3 предшествующей главы для случая распределения по одному признаку. Момент k -го порядка выражается интегралом:

$$\nu_k = \frac{1}{N} \int_a^b x^k \varphi(x) dx^*),$$

а приближенная его величина представляется суммой:

$$\nu_k = \frac{\Delta x}{N} \sum_{i=0}^{i=n-1} x_i^k \varphi(x_i),$$

где Δx — длина каждого из интервалов, на которые разбит весь промежуток значений признака, встречающихся в таблице распределения.

*) См. глава I форм. (12').

Если изменять масштаб значений переменного так, чтобы Δx равнялось единице, то приближенное выражение момента будет таково:

$$v_k = \frac{1}{N} \sum_{i=0}^{i=n-1} x_i^k \varphi(x_i) \dots \dots \dots (60)^*$$

В настоящей главе мы имеем дело с распределением по двум признакам, и моменты выражаются приведенными выше формулами (4) и (5), при чем

$$\varphi(x, y) dx dy = Z dx dy = n_x''.$$

Применяя к интегралам (4) и (5) те же рассуждения, какие мы делали при выводе формулы (60), можем сказать, что приближенные величины моментов выражаются формулами:

$$v_x = \frac{1}{N} \sum_x \sum_y n_x'' x, \quad v_y = \frac{1}{N} \sum_x \sum_y n_x'' y \dots \dots \dots (61)$$

$$v_{x,x} = \frac{1}{N} \sum_x \sum_y n_x'' x^2, \quad v_{y,x} = \frac{1}{N} \sum_x \sum_y n_x'' xy, \quad v_{y,y} = \frac{1}{N} \sum_x \sum_y n_x'' y^2 \quad (62)$$

где суммы распространяются на все значения указателей x и y , которые относятся поочередно ко всем прямоугольникам, соответствующим числам данной таблицы распределения. При этом так же, как и формуле (60), мы приняли интервалы Δx и Δy равными единице. Для нахождения моментов при других значениях масштабов, мы можем воспользоваться тем заключением, которое сделали в самом конце параграфа 4.

Итак, для нахождения моментов мы должны вычислить суммы, входящие в формулы (61) и (62).

Имея данную таблицу распределения (см. черт. 13), прежде всего подводим итоги вертикальных столбцов (строев) этой таблицы; эти итоги суть численности n_x'' строев. Они записаны внизу каждого столбца и составляют последнюю строку; сумма чисел этой строки равна объему совокупности N и поставлена справа в конце последней строки.

Далее складываются числа по горизонтальным строкам (строям); эти суммы суть численности n_y'' горизонтальных строев; они записаны в конце каждой строки и составляют последний столбец. Сумма чисел этого столбца равна тому же объему N , уже написанному внизу последнего столбца.

После вычисления численностей вертикальных и горизонтальных строев приступают к вычислению сумм, входящих в формулы (61) и (62).

* См. глава I, формула (12).

Начнем с первой суммы:

$$\sum_x \sum_y n_{xy}^x.$$

Мы можем прежде всего произвести суммирование по y , отбирая те слагаемые, где x равно одному и тому же числу. Эта сумма изобразится так:

$$x \sum_y n_{xy}^y.$$

Так как $\sum_y n_{xy}^y$ есть численность n_x для x -вого строка, то взяв эту сумму изобразится так:

$$x n_x.$$

а потому будем иметь:

$$v_x = \frac{1}{N} \sum_x \sum_y n_{xy}^x = \frac{1}{N} \sum_x n_x^x \dots \dots \dots (63)$$

Отсюда видно, что для вычисления момента v_x мы должны составить сумму попарных произведений чисел последней строки таблицы на соответствующие числа, стоящие в первой строке вверх таблицы; разделив эту сумму на N , получим величину первого момента v_x .

Момент 2-го порядка v_{xx} выразится формулой, подобной (63):

$$v_{xx} = \frac{1}{N} \sum_x n_{xx}^x \dots \dots \dots (64)$$

Для его вычисления составляем сумму попарных произведений чисел последней строки таблицы на квадраты соответствующих чисел в первой строке. Разделив вычисленную сумму на объем N , найдем v_{xx} .

Моменты v_y и v_{yy} вычисляются подобным же образом помощью формул:

$$v_y = \frac{1}{N} \sum_y n_{yy}^y \dots \dots \dots (65)$$

$$v_{yy} = \frac{1}{N} \sum_y n_{yy}^y \dots \dots \dots (66)$$

Для нахождения момента $v_{x,y}$ нам нужно вычислить сумму:

$$v_{x,y} = \frac{1}{N} \sum_x \sum_y n_{xy}^2 \quad (67).$$

Непосредственное вычисление момента $v_{x,y}$ по этой формуле было бы утомительно в арифметическом смысле. Для упрощения вычислений можно воспользоваться так называемым разностным способом, который состоит в следующем.

Проводим диагональные черты, как видно на таблице 13, и заготавливаем место для двух перпендикулярных к ним строк с тем, чтобы на каждой диагональной черте поместилось число в каждой из заготовленных двух строк.

Мы видим, что во всех прямоугольниках одной и той же диагональной черты разность $x - y$ одна и та же. Величину такой разности записываем по соответствующей диагональной черте в верхней из заготовленных двух строк. Сумму численностей, стоящих по каждой диагональной черте, записываем во второй заготовленной строке под соответствующим числом первой. Записанную сумму мы обозначим символом n_{x-y} .

Выполнив эту подготовительную работу, составляем сумму попарных произведений чисел второй строки на квадрат соответствующих чисел первой строки и полученный результат делим на N . Эту величину можем обозначить символом $v_{(x-y)^2}$; она выразится так:

$$v_{(x-y)^2} = \frac{1}{N} \sum_{x-y} n_{x-y} (x-y)^2 \quad (68).$$

Вычисление этой суммы не труднее, чем вычисление сумм (63), (64), (65) и (66).

Раскрыв скобки в выражении (68) и группируя члены, мы без особого труда заметим, что эта сумма выразится так:

$$v_{(x-y)^2} = \frac{1}{N} S_x S_y n_x^2 x^2 - \frac{2}{N} S_x S_y n_{xy}^2 xy + \frac{1}{N} S_x S_y n_y^2 y^2$$

Приняв во внимание равенства (62), находим:

$$v_{(x-y)^2} = v_{x,x} - 2v_{x,y} + v_{y,y}$$

откуда:

$$v_{x,y} = \frac{v_{x,x} + v_{y,y} - v_{(x-y)^2}}{2} \quad (69)$$

Следовательно, совершив вычисление величин $v_{x,x}$, $v_{y,y}$ и $v_{(x-y)^2}$, будем иметь и величину $v_{x,y}$.

Моменты 1-го порядка v_x и v_y суть координаты ξ и η центра распределения. Они находятся рассмотренным выше способом. Вычислив эти величины, переносим начало координат в центр и находим центральные моменты по найденным выше формулам (9):

$$\mu_{11} = v_{x,x} - v_x^2, \quad \mu_{12} = v_{x,y} - v_x v_y, \quad \mu_{22} = v_{y,y} - v_y^2.$$

Отсюда находим величины:

$$\sigma_1 = \sqrt{\mu_{11}}, \quad \sigma_2 = \sqrt{\mu_{22}} \quad (36)$$

$$r = \frac{\mu_{12}}{\sigma_1 \sigma_2} \quad (37)$$

а затем величины:

$$\rho_{y(x)} = \frac{r \sigma_2}{\sigma_1} \quad (47)$$

$$\rho_{x(y)} = \frac{r \sigma_1}{\sigma_2} \quad (47')$$

Предшествующие вычисления все основаны на замене интегралов соответствующими суммами. Такая замена дает лишь приближенные значения интеграла; приближение тем больше, чем меньше интервалы.

Вычислив коэффициент регрессии и найдя положение центра распределения, мы можем на чертеже построить прямые регрессии. Но полезно произвести для них некоторую проверку.

Среднее значение Y для x -вого строа выразится формулой:

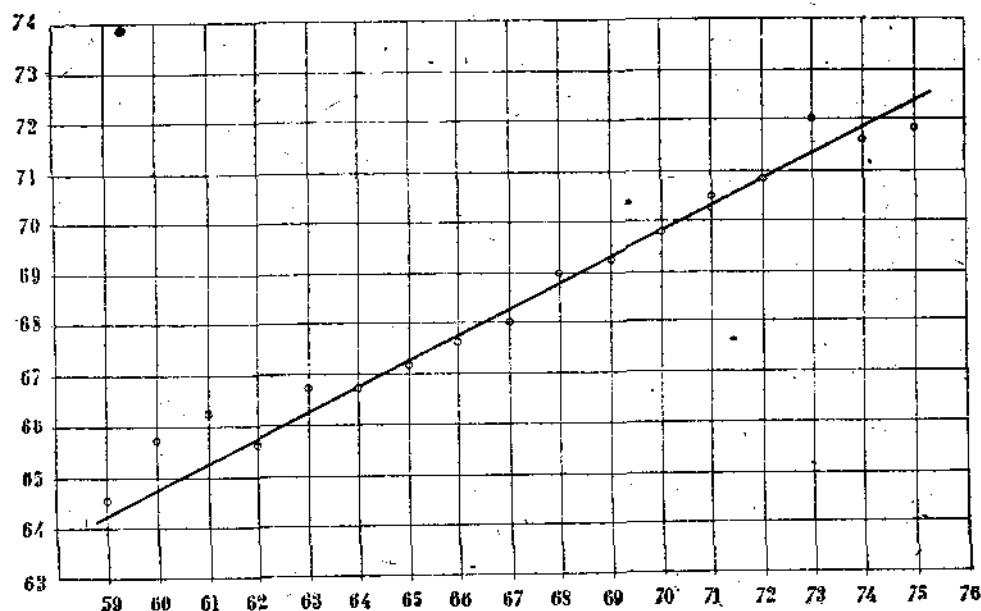
$$Y = \frac{1}{n_x} S_{x,y} \quad (70)$$

Построив по этой формуле величину Y для каждого строа, получим ряд точек, служащих центрами для вертикальных строа; они должны лежать все на одной прямой, — линии регрессии.

Само собой понятно, что в реальной задаче, хотя бы и при нормальном распределении, нельзя ожидать, что все построенные центры строа будут в точности лежать на теоретически построенной прямой регрессии. Отступление же в положениях центров от теоретической линии регрессии происходит оттого, что вычисления нами произведены только приближенные, кроме того, никакое реальное распределение нельзя считать безусловно нормальным. Поэтому весьма важно на чертеже посмотреть, насколько построенные центры строа близки расположены к теоретически построенной прямой регрессии.

Для рассмотренного выше примера корреляции роста отцов и сыновей такую проверку можно видеть на черт. 14, заимствованном у Пирсона *).

*) См. Biometrika II, стр. 362.



Чертеж 14.

Мы замечаем, что в средней части чертежа центры строев близко лежат к прямой регрессии, отступая немного то в ту, то в другую сторону, но вблизи краев чертежа отступления довольно значительны. Это объясняется тем, что число людей с очень малым и очень большим ростом весьма не велико, а потому вычисленные величины $\bar{y}$ при значениях x , близких к границам таблицы, не надежны.

Таким же образом, как мы сделали проверку для линии регрессии y относительно x , можно сделать проверку и для линии регрессии x относительно y .

§ 7. Понятие о криволинейной корреляции и о корреляционном отношении.

Когда таблица распределения получена из наблюдений, заранее нельзя сказать, имеем ли мы дело с нормальным распределением, или нет. Даже чаще всего распределение более или менее уклоняется от нормы.

Во всяком случае, мы можем, применяя приемы предшествующего параграфа, вычислить величины моментов $\nu_{x^2}, \nu_{y^2}, \nu_{x, x^2}, \nu_{x, y^2}, \nu_{y, x^2}, \nu_{y, y^2}$.

По этим величинам мы найдем координаты ξ и η центра распределения, затем центральные моменты 2-го порядка μ_{11}, μ_{12} и μ_{22} , а по ним величины средних квадратических уклонений σ_1 и σ_2 и величину r , которую мы попрежнему назовем коэффициентом корреляции, но которой мы не можем, однако, приписать того простого смысла, какой она имеет в случае распределения нормального.

Чтобы выяснить, будет ли в данном случае распределение нормальным или, в противном случае, насколько оно близко к нормальному, строим центры строев, пользуясь формулой (70) предшествующего параграфа.

Ломаную линию, проходящую через построенные центры строев в последовательном порядке, мы назовем линией регрессии в отличие от прямой регрессии, выражаемой уравнением (46). В случае нормального распределения линия регрессии совпадет с прямой (46), в противном случае она будет от нее более или менее значительно уклоняться.

В этом последнем случае строят интерполяционную кривую, по возможности, близко подходящую к линии регрессии; эту кривую называют кривая регрессии. Если эта кривая значительно уклоняется от прямой, то зависимость между признаками, характеризуемую данным распределением, называют криволинейной корреляцией (*skew correlation*). Чтобы оценить величину корреляционной зависимости, когда она криволинейна, мы обобщим понятие о среднем квадратическом уклонении для линии регрессии, рассмотренном нами в конце параграфа 4 и определяемом формулой (48). Так же, как и в названном параграфе, мы составим произведение численности x -вого строя на квадрат координаты Y центра этого строя, вычисляем среднее арифметическое всех таких попарных произведений и извлекаем квадратный корень из этого среднего арифметического. Результат мы назовем средним квадратическим уклонением для линии регрессии. Обозначив его буквою σ_x , будем иметь:

$$\sigma_x^2 = \frac{1}{N} SY^2n_x \dots \dots \dots (71)$$

Ясно, что в случае нормального распределения выражение (71) тождественно с выражением (48).

Но в общем случае, когда корреляция криволинейна, величина Y уже не выражается формулой (46), а потому и равенство (49) перестает быть справедливым.

В случае криволинейной корреляции мы обозначим буквою η отношение

$$\eta = \frac{\sigma_x}{\sigma_y} \dots \dots \dots (72)$$

и назовем его корреляционным отношением.

Сравнение формул (49) и (72) указывает на аналогию в значении величин $|r|$ и η . В действительности здесь есть связь более глубокая, чем простая аналогия.

Не входя в подробности доказательства, чтобы не увеличивать сверх меры величины курса, ограничусь указанием, что в случае

криволинейной корреляции величина η имеет те самые свойства, какие имеет величина $|r|$ при корреляции прямолинейной, т.е. в случае нормального распределения, а именно:

- 1) когда $\eta = 0$, никакой зависимости между x и y нет;
- 2) когда η лежит между 0 и 1, признаки x и y связаны между собою корреляционной зависимостью, которая тем сильнее, чем η ближе к 1;
- 3) при $\eta = 1$, корреляционная зависимость переходит в функциональную, которая устанавливается кривой регрессии; следовательно, в общем случае эта зависимость может выражаться какою угодно кривой линией, а не исключительно прямою, как в нормальном случае;
- 4) больше 1 величина η быть не может;
- 5) чем распределение ближе к нормальному, тем η ближе к $|r|$: при нормальном распределении η равно $|r|$, что ясно и из сказанного выше, так как в случае нормального распределения правая часть равенства (71) тождественна с правой частью равенства (49), а потому $\eta = |r|$.

ГЛАВА ТРЕТЬЯ.

О статистических рядах.

§ 1. Понятие о статистическом ряде.

Материал, собираемый в статистике, располагается по некоторому плану, смотря по характеру материала и той цели, для которой он собран.

Так, например, при изучении населения можно расположить в таблицу численности населения различных государств или различных городов в известный год, или численность населения какого-либо одного государства или города за целый ряд последовательных лет, при чем эти числа могут быть еще подразделены по полу, возрасту, роду занятий жителей и т. д. При изучении смертности таблицу можно расположить по числу всех умерших в данном государстве или городе погодно, при чем таблица может быть подразделена по полу умерших, по возрасту, по условиям жизни (земледельцы, горожане) и т. д. В вопросах рождаемости таблица может указывать число родившихся младенцев с распределением по годам, по полу, а также с указанием числа родившихся живыми и мертворожденных. При изучении самоубийств и преступлений таблица может указывать число бывших случаев с указанием возраста, пола, орудия совершения самоубийства или преступления и т. д.

Числа одной и той же таблицы называются статистическими числами, а составляемый ими ряд статистическим рядом.

Если числа получились в результате непосредственного подсчета, то они называются полными числами. Если же таблица дает величины отношений чисел, полученных непосредственным подсчетом, то ее члены называются относительными числами.

Так, например, изучая население данного города по годам, мы можем интересоваться постепенным ростом населения и дать отношение численности населения для каждого года к населению, бывшему в определенный год, или дать величину процента прироста населения за ряд последовательных лет. Изучая смертность, мы

можем дать для каждого года отношение числа умерших в этом году к числу живших в начале года в данном месте. Изучая рождаемость, мы можем дать отношение числа родившихся к численности всего населения в начале года в данном месте; кроме того, можно дать число родившихся мальчиков и девочек на 100 родившихся детей, или отношение числа родившихся мальчиков к числу родившихся девочек и т. д.

Такие статистические числа называются относительными числами.

Рассматривая статистические ряды, мы замечаем в большинстве из них большое постоянство, или так называемую устойчивость.

В виде примера приведем следующие таблицы, заимствуемые из книги Вестергаарда *).

I. Число родившихся в Дании.

Год.	Число родившихся		Год.	Число родившихся.	
	Полное	В ‰/о.		Полное.	В ‰/о.
1860	54.797	3.39	1870	56.472	3.15
1861	53.747	3.29	1871	56.407	3.11
1862	53.011	3.21	1872	57.274	3.13
1863	53.939	3.23	1873	58.616	3.17
1864	52.884	3.14	1874	59.324	3.18
1865	55.434	3.25	1875	61.791	3.28
1866	57.353	3.33	1876	63.967	3.36
1867	54.763	3.15	1877	63.772	3.32
1868	56.546	3.22	1878	63.144	3.26
1869	54.056	3.05	1879	64.363	3.29
				1151.660	3.23

II. Число мертворожденных в Дании.

Год.	Всего.	Между ними мальчиков.
1870	2052	1166
1871	2011	1158
1872	2053	1163
1873	2045	1117
1874	2046	1185
1875	2042	1165
1876	2179	1239
1877	1938	1075
1878	1854	1064
1879	1908	1405
	20118	11389

III. В Statistisches Jahrbuch der Stadt Berlin 1900/1902 находятся в числе прочих следующие данные **):

*) См. Westergaard Die Grundzüge der Theorie der Statistik, Jena, 1890. Стр. 40.

**) Эти данные заимствованы из книги А. А. Чупрова. — Очерки теории статистики. Спб. 1909. Стр. 248.

Число переездов с квартиры на квартиру в Берлине было:

год	в октябре	в ноябре
1900	125.627	45.210
1901	133.937	48.493
1902	134.202	46.512

Число извозчиков, отъезжавших с седоками:

	от Потсдамского вокзала.	от Герлицкого вокзала.
1900 г.	5205	1352
1901 .	5738	1306
1902 „	5945	1341

Число несчастных случаев пришлось

	на воскресенье.	на понедельник.
1900 г.	5219	7612
1901 „	5316	7446
1902 „	5250	7702

IV. Относительное число родившихся мальчиков к числу всех родившихся детей представляет пример исключительно большой устойчивости. Чубер*) приводит следующую таблицу для живых законно-рожденных младенцев, построенную по данным немецкой статистики:

1866 г.	0.51637	1886 г.	0.51575
1867 „	51662	1887 „	51323
1868 „	51501	1888 „	51461
1869 „	51575	1889 „	51405
1870 „	51514	1890 „	51495
1871 „	51565	1891 „	51473
1872 „	51594	1892 „	51572
1873 „	51554	1893 „	51562
1874 „	51611	1894 „	51455
1875 „	51397	1895 „	51409
1876 „	51503	1896 „	51358
1877 „	51488	1897 „	51373
1878 „	51398	1898 „	51353
1879 „	51445	1899 „	51351
1880 „	51479	1900 „	51484
1881 „	51469	1901 „	51351
1882 „	51523	1902 „	51379
1883 „	51450	1903 „	51311
1884 „	51378	1904 „	51437
1885 „	51439	1905 „	51272

Средн. 0.51460

*) Czuber Wahrscheinlichkeitsrechnung, 2-te Aufl., Bd. II. Сtp. 51.

§ 2. Статистическая вероятность.

Из приведенных в предшествующем параграфе данных видно, что особенно большая устойчивость замечается в рядах относительных чисел.

Между этими рядами надо отметить те ряды относительных чисел, где знаменатель есть объем некоторой совокупности, а числитель есть объем некоторой части этой совокупности, объекты которой определяются каким-либо признаком. Таков, между прочим, последний из приведенных примеров: разделив число новорожденных младенцев мужского пола на число всех родившихся в данном месте и в данном году младенцев, находим относительное число новорожденных мальчиков. Так же точно разделив число умерших в данном году в данном месте на число всех живых в начале этого года, получим относительное число, называемое смертностью.

Относительные числа этого рода можно назвать статистической частостью по аналогии с той частостью, которую мы наблюдаем, например, производя опыты с извлечением шаров из урны.

Говоря об этих опытах в I-й части курса, мы видели, что здесь приходится различать два случая: 1) когда при условиях опыта есть основная вероятность; в таком случае частость при большом числе испытаний бесконечно близка к этой основной вероятности; 2) когда основной вероятности нет; такие случаи по характеру своему могут быть весьма разнообразны, и мы рассмотрели несколько случаев, где существует так называемая средняя вероятность. При бесконечно большом числе испытаний, частость, как мы видели, бесконечно близка к такой средней вероятности.

Статистическая частость в очень многих случаях похожа на частость в только-что упомянутом тесном смысле слова; тем не менее между ними есть существенное различие. Например, вычисляя смертность, мы имеем дело с группой людей, весьма различных по возрасту, состоянию здоровья, условиям жизни, обычаям и т. д. Такой неоднородный состав совокупности не соответствует требованиям равновозможности всех статочностей, которое действительно близко осуществляется при опытах с извлечением шаров из урны.

Тем не менее постоянство статистических частостей, которое видно на примерах, приведенных в предшествующем параграфе и им подобных, а также согласие результатов применения формул и теорем теории вероятностей к обработке статистического материала, с результатами непосредственной проверки, дают достаточно прочное основание предположке, что во многих случаях статистическая частость обладает такими же свойствами, как частость в узком смысле слова. Поэтому мы считаем ее при большом числе испытаний весьма близкой к вероятности или средней, или, в исключительных случаях, основ-

ной. Поэтому статистическую частоту считают за приближенную величину этой вероятности, называемой статистическою вероятностью.

Чтобы пояснить сказанное, возьмем пример из той области, в которой много работал Лексис, положивший начало разработке вопросов, к которым мы теперь приступаем.

Пусть речь идет о вероятности рождения мальчика для совокупности новорожденных в данном городе или стране.

Мы предполагаем, что существует теоретическое число, которое мы называем вероятностью данного события; когда мы берем некоторую группу новорожденных детей μ , подсчитываем число m , находящихся между ними мальчиков и берем отношение $\frac{m}{\mu}$, то это будет частота рождения мальчика в данной группе. Беря различные группы из числа всех новорожденных в данном месте, мы будем получать различные величины частот, которые группируются около величины вероятности подобно тому, как при опытах с урной частоты появления белого шара группируются около вероятности появления белого шара.

В этом смысле можно говорить об отклонении

$$\frac{m}{\mu} - p = x \dots \dots \dots (1)$$

частоты для данной группы от вероятности p , а также о распределении частот около вероятности. В таком случае частота $\frac{m}{\mu}$ представится, как приближенная величина вероятности p , а x будет ошибкою при этой приближенной величине.

Применяя теорию ошибок наблюдения, рассмотренную во II части, можем сказать, что вероятность отклонения x выражается формулою:

$$\frac{h}{\sqrt{\pi}} e^{-h^2 x^2} dx \quad \dots \dots \dots (2)$$

а умножив эту величину на число N взятых групп и разделив на dx , получим формулу распределения частоты во взятых N группах:

$$y = N \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} \dots \dots \dots (3)$$

где h есть мера точности данного распределения.

Квадратный корень из математического ожидания величины отклонения x попрежнему назовем средним квадратическим отклонением:

*) См. часть II, гл. I, формула (18).

**) См. часть III, гл. I, формула (20).

обозначая его буквою σ , будем иметь:

$$\sigma^2 = \frac{h}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-hx^2} x^2 dx = \frac{1}{2h^2},$$

откуда:

$$h = \frac{1}{\sigma\sqrt{2}} \quad *) \quad \dots \quad (4)$$

Если в данном массовом явлении с вероятностью события, равной p , распределение частот следует тому же закону, как распределение числа появлений белых и черных шаров из урны, для которой вероятность появления белого шара равна тому же числу p , а вынутые шары возвращаются каждый раз в урну обратно, то мера точности h выразится тою же формулой, которая нами получена для случая испытаний с урной, а именно:

$$h = \sqrt{\frac{\mu}{2pq}} \quad **). \quad \dots \quad (5)$$

Такой случай мы назвали нормальным; величина среднего квадратического отклонения в этом случае выразится так:

$$\sigma = \sqrt{\frac{pq}{\mu}} \quad \dots \quad (6)$$

Но кроме случаев с основной вероятностью, существуют случаи распределения более сложные. Некоторые из них мы рассмотрели в главе I и нашли для них соответствующие выражения меры точности h ***). Составить общую формулу меры точности, которая позволяла бы вычислить величину h при всевозможных случаях, встречающихся в статистической практике, понятно, нельзя. Можно только указать прием вычисления меры точности для каждого отдельного случая на основании имеющегося опытного материала.

Пусть произведен подсчет частот в k группах, на которые разбита некоторая совокупность, по возможности, однородного состава. Для простоты вычислений положим, что объем μ во всех группах одинаковый. Пусть получились следующие величины частот:

$$p^{(1)} = \frac{m^{(1)}}{\mu}, \quad p^{(2)} = \frac{m^{(2)}}{\mu}, \quad p^{(3)} = \frac{m^{(3)}}{\mu}, \dots, \quad p^{(k)} = \frac{m^{(k)}}{\mu} \quad \dots \quad (7)$$

Рассматривая их, как результаты k равноточных измерений величины вероятности p , можем сказать, что наимвероятнейшая ве-

*) См. часть I, гл. III, формула (34).

**) См. часть I, гл. III, формула (17).

***) См. часть I, гл. III, формула (42) и (62*).

личина p на основании полученного материала есть средняя арифметическая:

$$p_0 = \frac{p^{(1)} + p^{(2)} + p^{(3)} + \dots + p^{(k)}}{k} \quad (8)$$

В таком случае среднее квадратическое отклонение каждой из величин (7) выразится формулой:

$$\sigma'' = \sqrt{\frac{\sum (p_0 - p^{(i)})^2}{k-1}} \quad (9)$$

По этой величине среднего квадратического отклонения найдется соответствующее значение меры точности:

$$h'' = \frac{1}{\sigma'' \sqrt{2}} = \sqrt{\frac{k-1}{2 \sum (p_0 - p^{(i)})^2}} \quad (10)$$

В нормальном случае, кроме формулы (9), среднее квадратическое отклонение выражается еще формулой (6). Так как точную величину вероятности p мы не знаем, то мы можем заменить величину p ее приближенным значением p_0 , при чем q заменится величиной $q_0 = 1 - p_0$, и мы получим приближенные величины меры точности и среднего квадратического отклонения:

$$h' = \sqrt{\frac{\mu}{2 p_0 q_0}}, \quad \sigma' = \sqrt{\frac{p_0 q_0}{\mu}} \quad (11)$$

Если данное массовое явление удовлетворяет условиям нормального случая, то величины (9) и (11) дают приближенные значения одной и той же величины, а потому они мало разнятся между собою, а отношение их:

$$Q = \frac{\sigma''}{\sigma'} = \frac{h'}{h''} \quad (12)$$

близко к 1.

Когда же данное массовое явление не соответствует условиям нормального случая (т.-е. когда p есть вероятность средняя, а не основная), то величина (9) может разниться от величины (11) более или менее значительно, и тогда величина Q будет разниться от 1 более или менее значительно. Поэтому величина Q может служить мерою отклонения данного массового явления от условий нормального случая. Величина Q может быть названа коэффициентом расхождения***) или коэффициентом дисперсии по причине, которую увидим ниже.

*) См. часть II, гл. I, форм. (31).

**) См. часть II, гл. I, форм. (30).

***) Это понятие введено впервые французским статистиком Дормуа (Dormeu), а затем, независимо от него, снова найдено Лексисом, который назвал его Divergencecoefficient; Ворткевич назвал его Fehlerrelation.

Величина (9) названа Лексисом физической величиной среднего квадратического уклонения, а величина (11) — комбинаторной величиной среднего квадратического уклонения.

Физическая величина всегда дает приближение к истинной величине среднего квадратического уклонения, а комбинаторная показывает, каково было бы среднее квадратическое уклонение, если бы p было основной вероятностью в данном случае.

Внося величины h'' и h' в формулу (3), находим два, вообще говоря, различных выражения:

$$y'' = N \frac{h''}{\sqrt{\pi}} e^{-h''^2 x^2}, \dots \dots \dots (3'')$$

$$y' = N \frac{h'}{\sqrt{\pi}} e^{-h'^2 x^2} \dots \dots \dots (3')$$

Первая формула дает уравнение кривой распределения, близкое к истине в данном случае; а второе показывает вид кривой распределения в нормальном случае при той же величине вероятности p .

Иногда в математической статистике приходится иметь дело с выражением, которое не есть статистическая вероятность, но может быть представлено, как функция статистической вероятности. Например, пусть μ есть число всех родившихся в данном месте за известный промежуток времени детей, из которых m мальчиков и n девочек.

Величина:

$$w = 1000 \frac{m}{n}$$

показывает, сколько мальчиков приходится на 1000 девочек. Это число не есть статистическая вероятность; но отношение:

$$p = \frac{m}{\mu}$$

можно назвать вероятностью рождения мальчика.

Выражение w представляется в виде:

$$w = 1000 \frac{p}{1-p}.$$

Это — функция вероятности p .

Пусть w есть вообще какая-нибудь функция статистической вероятности:

$$w = f(p) \dots \dots \dots (13)$$

и пусть путем подсчета мы нашли наивероятнейшую величину p_0 , вероятности p и, кроме того, вычислили среднее квадратическое уклонение этой величины.

Применяя сказанное во II части, гл. II, форм. (25) и (26), найдем, что наивероятнейшая величина w есть w_0 , определяемая равенством:

$$w_0 = f(p_0), \quad (14)$$

а среднее квадратическое уклонение ее выражается формулой:

$$\Sigma = f'(p_0) \sigma. \quad (15)$$

В частности, когда:

$$w = 1000 \frac{p}{1-p},$$

будем иметь:

$$\Sigma = 1000 \frac{1}{(1-p)^2},$$

где p —вероятность рождения мальчика.

Заменяя p ее приближенной величиною p_0 и пользуясь формулами (11), можем написать:

$$\Sigma' = \frac{1000}{(1-p_0)^2} \sqrt{\frac{p_0(1-p_0)}{p}}, \quad H' = \frac{(1-p_0)^2 \sqrt{p}}{1000 \sqrt{2p_0(1-p_0)}}, \quad (16)$$

где Σ' и H' суть комбинаторные величины среднего квадратического уклонения и меры точности величины w_0 , полученной из наблюдения.

На основании формул (9) и (10) будем иметь:

$$\Sigma'' = \sqrt{\frac{\Sigma(w_0 - w^{(i)})^2}{k-1}}, \quad H'' = \sqrt{\frac{k-1}{2 \Sigma(w_0 - w^{(i)})^2}}. \quad (17)$$

Это—физические величины среднего квадратического уклонения и меры точности того же результата наблюдения w_0 .

Отношение их:

$$Q = \frac{\Sigma''}{\Sigma'} = \frac{H'}{H''} \quad (18)$$

есть коэффициент дисперсии для данного случая.

§ 3. Сравнение статистических вероятностей в двух совокупностях.

Пусть в двух совокупностях одинакового характера найдены статистические вероятности одинакового события: пусть они оказались равными p_0' и p_0'' , при чем средние квадратические ошибки их суть σ_1 и σ_2 . Буквою p_0' мы обозначаем большую из двух найденных вероятностей и полагаем:

$$p_0' - p_0'' = d,$$

где d —положительно.

Спрашивается, какова вероятность p , что эта положительная разность не случайна, т.-е. что вероятность события в первой совокупности в действительности больше, чем во второй.

Обозначая буквами p' и p'' истинные величины вероятностей в 1-й и 2-й совокупности, можем сказать, что на основании формул (10) и (15), § 8 главы II, части II, что вероятнейшая величина разности $p' - p''$ есть $p_0' - p_0'' = d$ и что средняя квадратическая ошибка величины d выражается так:

$$\Sigma = \sqrt{\sigma_1^2 + \sigma_2^2}.$$

Поэтому мера точности величины d будет такова:

$$H = \frac{1}{\Sigma \sqrt{2}} = \frac{1}{\sqrt{2(\sigma_1^2 + \sigma_2^2)}}.$$

Обозначая буквою δ ошибку при величине d , находим, что вероятность этой ошибки выражается формулой:

$$\frac{H}{\sqrt{\pi}} e^{-H^2 \delta^2} d\delta. \quad (19)$$

Вероятность того, что эта ошибка лежит в пределах между a и b , представится в виде интеграла:

$$\frac{H}{\sqrt{\pi}} \int_a^b e^{-H^2 \delta^2} d\delta. \quad (20)$$

Если δ есть ошибка при величине

$$p_0' - p_0'' = d,$$

то имеет место равенство:

$$p' - p'' = d + \delta.$$

Если разность $p' - p''$ положительна, то отсюда будет следовать:

$$d + \delta > 0, \text{ или } \delta > -d,$$

откуда следует, что δ лежит в пределах между $-d$ и $+\infty$.

Полагая в формуле (20): $a = -d$ и $b = +\infty$, найдем:

$$P = \frac{H}{\sqrt{\pi}} \int_{-d}^{+\infty} e^{-H^2 \delta^2} d\delta. \quad (21)$$

Таков ответ на поставленный выше вопрос. Остается вычислить интеграл (21).

С этой целью делаем некоторые простые преобразования:

$$\begin{aligned} P &= \frac{H}{\sqrt{\pi}} \int_{-d}^0 e^{-H^2 \delta^2} d\delta + \frac{H}{\sqrt{\pi}} \int_0^{+\infty} e^{-H^2 \delta^2} d\delta = \\ &= \frac{1}{\sqrt{\pi}} \int_0^{Hd} e^{-x^2} dx + \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-x^2} dx. \end{aligned}$$

Введя известное обозначение:

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx,$$

и помня, что:

$$\frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-x^2} dx = 1,$$

находим:

$$p = \frac{1}{2} \left\{ 1 + \Phi(Hd) \right\}. \quad (22)$$

Если в рассматриваемом случае p' и p'' суть основные вероятности, то величины σ_1 и σ_2 выразятся по формуле (6). Полагая, что:

$$p_0' = \frac{m'}{\mu'}, \quad q_0' = \frac{\mu' - m'}{\mu'} = \frac{n'}{\mu'},$$

$$p_0'' = \frac{m''}{\mu''}, \quad q_0'' = \frac{\mu'' - m''}{\mu''} = \frac{n''}{\mu''},$$

будем иметь:

$$\sigma_1 = \sqrt{\frac{p_0' q_0'}{\mu'}} = \sqrt{\frac{m' n'}{\mu'^3}}, \quad \sigma_2 = \sqrt{\frac{p_0'' q_0''}{\mu''}} = \sqrt{\frac{m'' n''}{\mu''^3}},$$

откуда:

$$\Sigma = \sqrt{\sigma_1^2 + \sigma_2^2} = \sqrt{\frac{m' n' \mu''^3 + m'' n'' \mu'^3}{\mu'^3 \mu''^3}},$$

$$H = \frac{1}{\Sigma \sqrt{2}} = \sqrt{\frac{\mu'^3 \mu''^3}{2 (m' n' \mu''^3 + m'' n'' \mu'^3)}}.$$

Рассмотрим пример, приводимый Чубером *):

В Париже за 40 лет с 1745 по 1784 год родилось:

393.386 мальчиков,

377.555 девочек.

Всего: 770.941 детей.

В Лондоне за 95 лет с 1664 по 1758 год родилось:

737.629 мальчиков,

698.958 девочек.

Всего: 1.436.587 детей.

Следовательно: $p_0' = \frac{m'}{\mu'} = 0.51026$, $p_0'' = \frac{m''}{\mu''} = 0.51353$.

откуда

$$d = p_0'' - p_0' = 0.00327.$$

*) Czuber. Wahrscheinlichkeitsrechnung, zweite Auflage, Bd. II, стр. 11.

Спрашивается, какова вероятность по этим данным, что между новорожденными Лондона вероятность рождения мальчика больше, чем в Париже.

В данном случае:

$$Nd = 3,353.$$

По формуле (22) имеем:

$$p = 0.9999989.$$

Следовательно, до полной достоверности недостает только 0.0000011.

§ 4. Коэффициент дисперсии.

Пусть речь идет о массовом явлении, где статистическая вероятность равна p .

Применяя сказанное в параграфе 2, мы разбиваем имеющийся статистический материал на k групп одинакового объема μ и вычисляем для каждой группы соответствующую ей величину статистической частоты:

$$p^{(1)} = \frac{m^{(1)}}{\mu}, \quad p^{(2)} = \frac{m^{(2)}}{\mu}, \quad p^{(3)} = \frac{m^{(3)}}{\mu}, \dots, \quad p^{(k)} = \frac{m^{(k)}}{\mu}. \quad (7)$$

Средняя арифметическая из этих величин:

$$p_0 = \frac{p^{(1)} + p^{(2)} + p^{(3)} + \dots + p^{(k)}}{k} \quad (8)$$

есть наимвероятнейшее значение величины p .

Физическая величина среднего квадратического уклонения для величин (7) выражается формулою:

$$\sigma'' = \sqrt{\frac{\sum (p_0 - p^{(i)})^2}{k-1}}. \quad (9)$$

Соответствующая физическая величина меры точности:

$$h'' = \frac{1}{\sigma'' \sqrt{2}} = \sqrt{\frac{-k-1}{2 \sum (p_0 - p^{(i)})^2}}.$$

Кроме этих близких к истине физических величин, вычисляем имеющие теоретический смысл комбинаторные величины:

$$h' = \sqrt{\frac{\mu}{2 p_0 q_0}}, \quad \sigma' = \sqrt{\frac{p_0 q_0}{\mu}}. \quad (11)$$

Они показывают, какова была бы мера точности и среднее квадратическое уклонение статистических частостей, если бы распределение их совершалось около основной вероятности.

Для каждой из величин h'' и h' составляем кривую распределения:

$$y'' = N \frac{h''}{\sqrt{\pi}} e^{-h''^2 x^2}, \quad (8'')$$

$$y' = N \frac{h'}{\sqrt{\pi}} e^{-h'^2 x^2}. \quad (8').$$

Первая из этих кривых близка к истине, вторая имеет теоретическое значение в только что упомянутом смысле.

Припомним сказанное в параграфе 11 главы III части первой, мы находим, что в той из кривых (3'') и (3'), где величина параметра h больше, точки площади более сгруппированы около оси ординат; иначе та из этих кривых, где среднее квадратическое отклонение больше, имеет точки площади более рассеянными по плоскости. Кривую (3') мы называем кривою с нормальной дисперсией. Поэтому мы назовем кривую (3'') кривою с поднормальной дисперсией, если $h'' > h'$, и назовем ее кривою с сверхнормальной дисперсией, если $h'' < h'$.

В таком случае естественно назвать величину:

$$Q = \frac{\sigma''}{\sigma'} = \frac{h'}{h''}. \quad (12)$$

коэффициентом дисперсии: если $Q < 1$, то дисперсия поднормальная, если $Q > 1$, то дисперсия сверхнормальная.

Лексис применил эту формулу к исследованию уже упомянутого выше статистического ряда относительных чисел: он взял число мальчиков, приходящихся на 1000 новорожденных девочек в 34 округах Пруссии за 24 месяца в течение 1868 и 1869 годов.

Мера точности и коэффициент дисперсии выразятся по приведенным выше формулам (17) и (18). Величина p_0 есть вероятность рождения мальчика для всех округов за два года: она получается при делении числа родившихся мальчиков на число родившихся детей за 2 года во всех округах, и оказывается равною 0.515.

При вычислении комбинаторной меры точности H' по формуле (16) для каждого округа отдельно, мы принимаем μ равным месячному среднему числу рождений в данном округе за 24 месяца. Следовательно, значения величины H' в различных округах разнятся между собою только величиной множителя $\sqrt{\mu}$.

При вычислении для данного округа физической меры точности H'' по формуле (17) мы принимаем величины:

$$w^{(1)}, w^{(2)}, w^{(3)}, \dots, w^{(24)} \quad (23)$$

соответственно равным значению для этого округа выражения:

$$1000 \frac{m}{n}$$

за каждый из 24 месяцев 1868—1869 г.г. в отдельности. Величина w_0 есть среднее арифметическое из 24 величин (23) для данного округа.

Внося найденные для данного округа величины (23) в формулу (17), и полагая $k=24$, найдем величину физической меры точности H'' .

Внося величины H' и H'' в формулу (18), найдем коэффициент дисперсии Q для каждого округа в отдельности.

Величины, полученные Лексисом, находятся в следующей таблице:

О К Р У Г	μ	H'	H''	Q
Кенигсберг	3426	0.0195	0.0208	0.94
Гумбинден	2275	0.159	0.144	1.10
Давциг	1830	0.142	0.151	0.94
Мариенвердер	2918	0.180	0.249	0.72
Верли	2448	0.165	0.158	1.04
Потсдам	3028	0.183	0.176	1.04
Франкфурт	3211	0.189	0.185	1.02
Штеттин	2167	0.155	0.166	0.93
Кеслин	1844	0.143	0.119	1.20
Штральзунд	639	0.086	0.096	0.90
Познань	3738	0.203	0.205	0.99
Бромберг	2133	0.154	0.145	1.06
Вреславль	4766	0.230	0.205	1.12
Лигниц	2975	0.182	0.163	1.12
Оппельн	4856	0.232	0.214	1.08
Магдебург	3650	0.201	0.174	1.15
Мерзебург	2899	0.179	0.146	1.23
Эрфурт	1235	0.117	0.142	0.82
Шлеизиг	2715	0.173	0.118	1.47
Ганновер	1142	0.112	0.130	0.86
Гильдагейм	1200	0.115	0.114	1.01
Люнебург	975	0.104	0.094	1.11
Штаде	879	0.099	0.093	1.06
Аурих-Оснабрюк	1220	0.116	0.122	0.95
Мюнстер	1118	0.111	0.092	1.21
Минден	1464	0.127	0.141	0.90
Арисберг	2918	0.180	0.177	1.02
Кассель	2441	0.164	0.169	0.87
Висбаден	1837	0.143	0.108	1.32
Кобленц	1700	0.137	0.131	1.05
Трир	1901	0.145	0.148	0.98
Кельн	1936	0.146	0.149	0.98
Дюссельдорф	4305	0.218	0.247	0.88
Аахен	1485	0.128	0.151	0.85

Отсюда видно, что величина Q колеблется между 0.72 и 1.47; средняя величина равна 1.09. Следовательно, она близка к 1.

Отношение числа мужских и женских рождений было разносторонне изучено Лексисом; оно представляет собою статистический ряд с дисперсией, которую назовем нормальной, насколько это возможно в опытных исследованиях.

Для уяснения характера дисперсии в изучаемом явлении мы можем взять теоретическую задачу, условия которой сходны с условиями этого явления, и изучить особенности ее решения. Явление,

соответствующее таким теоретическим условиям, служит как бы моделью того реального явления, которое мы хотим изучить. Эти теоретические условия называются схемами.

Такие схемы нам встретились в главе III первой части, когда мы говорили о вероятностях сложных событий, имеющих место при многократном повторении испытаний.

Первая схема есть простейшая схема испытаний над урной, когда вынутые шары возвращаются в урну. Это — так называемая схема Бернулли, приводящая к нормальной дисперсии. Кроме этой схемы, нами были рассмотрены еще две другие, приводящие к сверхнормальной дисперсии. Первая из них соответствует явлению распределения шаров, вынимаемых из урны при условии, чтобы вынутый шар в урну не возвращался. Вторая схема соответствует условиям, при которых имеет место теорема Пуассона: поэтому она называется схема Пуассона.

В задачах, встречающихся в статистической практике, мы постоянно встречаем дисперсию сверхнормальную. Для уяснения этой особенности Лексис предложил особую схему, называемую схемой Лексиса. Займемся ее рассмотрением.

§ 5. Схема Лексиса.

Пусть имеется обширный статистический материал для определения вероятности некоторого явления. Пусть весь материал распадается на k серий, при чем в каждой серии имеется своя основная вероятность. Каждая серия содержит в себе μ величин частоты изучаемого явления, при чем это число μ членов одинаково во всех сериях.

Из всех полученных μk частот составляется среднее арифметическое, которое мы называем средней вероятностью явления при данных условиях.

Задача состоит в нахождении среднего квадратического отклонения для этой средней вероятности p .

Для пояснения мысли возьмем приведенные выше вычисления Лексиса, касающиеся вероятности рождения мальчика.

Частоты для одного и того же округа составляют отдельную серию. Числа одной серии дают величины частот в данном округе за каждый из 24 месяцев отчетного периода.

Можно допустить, что частоты для одного и того же округа в 2 года изменялись только случайно, им соответствует основная вероятность для данного округа, но вероятности для различных округов различны.

В примере Лексиса число серий $k \approx 34$, а число членов в каждой серий $\mu = 24$.

Сообразно этому мы можем распределить все полученные частоты в таблицу: частоты, стоящие в одной и той же строке, соответствуют 24 месяцам для одного и того же округа, а различные строки соответствуют 34 округам. Число величин во всех строках одинаково.

Схема Лексиса и основана на такой таблице. Число строк мы обозначим буквою k , а число величин в каждой строке — буквою μ . Такую таблицу можем представить в следующем виде:

$p_1^{(1)}$	$p_1^{(2)}$	$p_1^{(3)}$	...	$p_1^{(\mu)}$	$p_1 = \frac{p_1^{(1)} + p_1^{(2)} + \dots + p_1^{(\mu)}}{\mu} = \frac{\sum_j p_1^{(j)}}{\mu}$
$p_2^{(1)}$	$p_2^{(2)}$	$p_2^{(3)}$	...	$p_2^{(\mu)}$	$p_2 = \frac{p_2^{(1)} + p_2^{(2)} + \dots + p_2^{(\mu)}}{\mu} = \frac{\sum_j p_2^{(j)}}{\mu}$
$p_3^{(1)}$	$p_3^{(2)}$	$p_3^{(3)}$	...	$p_3^{(\mu)}$	$p_3 = \frac{p_3^{(1)} + p_3^{(2)} + \dots + p_3^{(\mu)}}{\mu} = \frac{\sum_j p_3^{(j)}}{\mu}$
...	...	...	...	...	...
$p_k^{(1)}$	$p_k^{(2)}$	$p_k^{(3)}$	...	$p_k^{(\mu)}$	$p_k = \frac{p_k^{(1)} + p_k^{(2)} + \dots + p_k^{(\mu)}}{\mu} = \frac{\sum_j p_k^{(j)}}{\mu}$

Отдельные частоты мы обозначаем символом $p_i^{(j)}$, где нижний указатель i есть номер строки, а верхний j есть место, занимаемое числом в строке.

Средние арифметические чисел в отдельных строках мы обозначаем символом p_i , а среднее арифметическое всех чисел p есть средняя величина частоты для всех данных таблицы.

Ясно, что указатель i меняется от 1 до k , а указатель j меняется от 1 до μ .

Величины p_i мы принимаем за статистические вероятности отдельных серий; это суть основные вероятности для этих серий. Величина p есть средняя вероятность для всего материала, даваемого таблицей.

Суммирование по обоим указателям: i и j будем обозначать символом: $\Sigma_i \Sigma_j$, а суммирование по одному указателю i и j будем обозначать соответственно символами Σ_i и Σ_j .

Обозначая физическое среднее квадратическое отклонение для всего материала символом σ'' , будем иметь:

$$\mu k \sigma''^2 = \Sigma_i \Sigma_j (p_i^{(j)} - p)^2, \dots \dots \dots (25)$$

или, раскрывая квадрат и суммируя почленно:

$$\mu k \sigma''^2 = \Sigma_i \Sigma_j (p_i^{(j)})^2 - 2p \Sigma_i \Sigma_j (p_i^{(j)}) + \mu k p^2.$$

Так как:

$$\Sigma_i \Sigma_j (p_i^{(j)}) = \Sigma_i (\mu p_i) = \mu k p,$$

то мы будем иметь:

$$\mu k \sigma'^2 \Rightarrow \Sigma_i \Sigma_j (p_i^{(j)})^2 - \mu k p^2 \quad (26).$$

Для нахождения величины суммы

$$\Sigma_i \Sigma_j (p_i^{(j)})^2$$

раскрываем выражение:

$$\Sigma_i \Sigma_j (p_i^{(j)} - p_i)^2 = \Sigma_i \Sigma_j (p_i^{(j)})^2 - 2 \Sigma_i \Sigma_j p_i^{(j)} p_i + \Sigma_i \Sigma_j (p_i)^2 \quad (27)$$

В результате суммирования по j в левой части найдем величину:

$$\mu \cdot \frac{\Sigma_j (p_i^{(j)} - p_i)^2}{\mu},$$

где второй множитель есть величина квадрата среднего квадратического отклонения частостей $p_i^{(j)}$ i -той строки от их средней величины p_i : так как по предположению p_i есть основная вероятность для чисел i -той строки, то среднее квадратическое отклонение выразится

формулой $\sqrt{\frac{p_i q_i}{\mu}}$ ^{*)}, а потому вся левая часть равенства (27) примет вид:

$$\Sigma_i p_i q_i.$$

При вычислении второго члена правой части равенства (27), произведя суммирование по j , можем вынести p_i за знак суммы и получим:

$$\Sigma_i \Sigma_j p_i^{(j)} p_i = \Sigma_i p_i \Sigma_j (p_i^{(j)}) = \mu \Sigma_i (p_i)^2.$$

Наконец, последний член формулы (27) можем представить так:

$$\Sigma_i \Sigma_j (p_i)^2 = \mu \Sigma_i (p_i)^2.$$

Поэтому все равенство (27) примет вид:

$$\Sigma_i p_i q_i = \Sigma_i \Sigma_j (p_i^{(j)})^2 - \mu \Sigma_i (p_i)^2,$$

откуда:

$$\Sigma_i \Sigma_j (p_i^{(j)})^2 = \Sigma_i p_i q_i + \mu \Sigma_i (p_i)^2. \quad (28)$$

Внеся это выражение в равенство (26), находим:

$$\mu k \sigma'^2 = \Sigma_i p_i q_i + \mu \Sigma_i (p_i)^2 - \mu k p^2. \quad (29)$$

Так как $q_i = 1 - p_i$, то первый член правой части принимает вид:

$$\Sigma_i p_i q_i = \Sigma_i p_i - \Sigma_i (p_i)^2 = k p - \Sigma_i (p_i)^2.$$

*) См. выше форм. (6).

Внеся это выражение в (29), найдем:

$$\mu k \sigma'^2 = kp + (\mu - 1) \Sigma_i (p_i)^2 - \mu kp^2 \quad (30).$$

Наконец, для нахождения суммы в правой части этого равенства раскрываем выражение:

$$\Sigma_i (p_i - p)^2 = \Sigma_i (p_i)^2 - 2p \Sigma_i p_i + kp^2 = \Sigma_i (p_i)^2 - kp^2,$$

откуда:

$$\Sigma_i (p_i)^2 = \Sigma_i (p_i - p)^2 + kp^2.$$

Внеся это выражение в формулу (30), находим:

$$\begin{aligned} \mu k \sigma'^2 &= kp + (\mu - 1) \Sigma_i (p_i - p)^2 + (\mu - 1) kp^2 - \mu kp^2 = \\ &= kp + (\mu - 1) \Sigma_i (p_i - p)^2 - kp^2 = kpq + (\mu - 1) \Sigma_i (p_i - p)^2, \end{aligned}$$

откуда окончательно:

$$\sigma' = \sqrt{\frac{pq}{\mu} + \frac{\mu - 1}{\mu} \cdot \frac{\Sigma_i (p_i - p)^2}{k}} \quad (31).$$

Это — формула Лексиса физического среднего квадратического уклонения для массового явления, соответствующего условиям его схемы.

Согласно прежнему обозначению находим:

$$\sigma' = \sqrt{\frac{pq}{\mu}}^* \quad (32);$$

это — комбинаторная величина среднего квадратического уклонения для чисел одной серии от средней величины, при чем эта средняя величина принята равной p . Это так называемая нормальная ошибка (Normalfehler).

Второе слагаемое под радикалом можно представить в виде M^2 , где:

$$M = \sqrt{\frac{\mu - 1}{\mu} \cdot \frac{\Sigma_i (p_i - p)^2}{k}} \quad (33).$$

Оно характеризует уклонение вероятностей отдельных серий от средней вероятности p для всего материала (absoluter Fehlerexcess).

Формула (31) примет вид:

$$\sigma' = \sqrt{\sigma'^2 + M^2} \quad (34);$$

она напоминает собою выражение среднего квадратического уклонения суммы двух результатов измерения:

$$w = x + y,$$

*) Сравни. форм. (11).

которое представляется формулой:

$$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 *},$$

где σ , σ_1 и σ_2 суть средние квадратические отклонения соответственно для w , x и y .

Величины σ' и M можно назвать слагающими для σ'' . Величина σ' называется несущественной слагающей (unwesentliche Schwankungskomponente), а M — физической слагающей (physische Schwankungskomponente).

Коэффициент дисперсии выразится так:

$$Q = \frac{\sigma''}{\sigma'} = \sqrt{1 + \frac{M^2}{\sigma'^2}} = \sqrt{1 + (\mu - 1) \frac{\sum_i (p_i - p)^2}{k p q}} \quad (35).$$

Ясно, что он больше, чем единица, т.е. дисперсия — сверхнормальная.

Величина M тем больше, чем сильнее разнятся вероятности отдельных серий от общей средней вероятности p ; от числа измерений μ в каждой серии она почти не зависит, потому что величина $\frac{\mu - 1}{\mu}$ с увеличением μ меняется в слабой степени. Между тем величина σ' , выражаемая формулой (32), обратно пропорциональна величине $\sqrt{\mu}$, с увеличением μ она быстро уменьшается. Поэтому величина Q с увеличением μ увеличивается.

Отсюда следует, что с увеличением числа наблюдений в сериях коэффициент дисперсии Q увеличивается; с уменьшением же числа наблюдений этот коэффициент приближается к 1.

§ 6. Схема Пуассона.

Сравним только что рассмотренную формулу коэффициента дисперсии в схеме Лексиса с коэффициентом дисперсии в схеме Пуассона, где, как мы знаем, дисперсия поднормальная.

Возьмем выражение меры точности в схеме Пуассона, приведенное в § 14, главы III, части I:

$$h'' = \frac{\sqrt{\mu}}{\sqrt{2 \frac{\sum_i p_i q_i}{\mu}}} **);$$

отсюда следует такое выражение для среднего квадратического отклонения:

$$\sigma'' = \sqrt{\frac{1}{\mu} \cdot \frac{\sum_i p_i q_i}{\mu}} \quad (36)$$

*) См. часть II, гл. II, форм. (16) и (18).

**) См. формулу (64*), параграфа 14, гл. III, ч. I.

Совершая преобразование, подобное рассмотренному в предыдущем параграфе, найдем:

$$\Sigma_i p_i q_i = \Sigma_i p_i - \Sigma_i (p_i)^2 = \mu p - \Sigma_i (p_i)^2 \quad (37)$$

Для нахождения суммы в правой части раскроем выражение:

$$\Sigma_i (p_i - p)^2 = \mu p^2 - 2p \Sigma_i p_i + \Sigma_i (p_i)^2 = \Sigma_i (p_i)^2 - \mu p^2,$$

откуда:

$$\Sigma_i (p_i)^2 = \Sigma (p_i - p)^2 + \mu p^2.$$

Внеся это выражение в формулу (37), найдем:

$$-\Sigma_i p_i q_i = \mu p - \mu p^2 - \Sigma_i (p_i - p)^2 = \mu p q - \Sigma_i (p_i - p)^2.$$

Формула (36) принимает вид:

$$\sigma = \sqrt{\frac{pq}{\mu} - \frac{1}{\mu} \cdot \frac{\Sigma_i (p_i - p)^2}{\mu}} \quad (38)$$

Коэффициент дисперсии выразится формулой:

$$Q = \sqrt{1 - \frac{\Sigma_i (p_i - p)^2}{\mu p q}} \quad (39)$$

Из этой формулы непосредственно видно, что Q меньше 1. Величина 2-го члена под радикалом зависит, главным образом, от величины:

$$\sqrt{\frac{\Sigma_i (p_i - p)^2}{\mu}}$$

Это — среднее квадратическое отклонение отдельных вероятностей p_i от средней вероятности p .

Таблица 1-я.

Таблицы функции $\Phi(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}$
(Czuberg, Wahrscheinlichkeitsrechnung, стр. 352.)

x	$\Phi(x)$	Раз- ность.	x	$\Phi(x)$	Раз- ность.	x	$\Phi(x)$	Раз- ность.
0.00	+1.1284	1	0.50	+0.8788	88	1.00	+0.4151	88
.01	+1.1283	4	.51	+0.8700	90	.01	+0.4068	81
.02	1.1279	5	.52	0.8610	90	.02	0.3987	81
.03	1.1274	8	.53	0.8520	90	.03	0.3906	80
.04	+1.1266	10	.54	+0.8430	92	.04	+0.3826	79
.05	1.1256	13	.55	0.8338	92	.05	0.3747	79
.06	1.1243	14	.56	0.8246	92	.06	0.3668	77
.07	+1.1229	17	.57	+0.8154	94	.07	+0.3591	76
.08	1.1212	19	.58	0.8060	93	.08	0.3515	76
.09	1.1193	21	.59	0.7967	95	.09	0.3439	74
0.10	+1.1172	24	0.60	+0.7872	94	1.10	+0.3365	74
.11	+1.1148	26	.61	+0.7778	95	.11	+0.3291	72
.12	1.1122	27	.62	0.7683	96	.12	0.3219	72
.13	1.1095	30	.63	0.7587	96	.13	0.3147	71
.14	+1.1065	32	.64	+0.7491	96	.14	+0.3076	69
.15	1.1033	34	.65	0.7395	96	.15	0.3007	69
.16	1.0999	37	.66	0.7299	96	.16	0.2938	68
.17	+1.0962	38	.67	+0.7203	97	.17	+0.2870	66
.18	1.0924	40	.68	0.7106	96	.18	0.2804	66
.19	1.0884	43	.69	0.7010	97	.19	0.2738	65
0.20	+1.0841	44	0.70	+0.6913	97	1.20	+0.2673	63
.21	+1.0797	46	.71	+0.6816	97	.21	+0.2610	63
.22	1.0751	49	.72	0.6719	97	.22	0.2547	62
.23	1.0702	50	.73	0.6622	96	.23	0.2485	60
.24	+1.0652	52	.74	+0.6526	97	.24	+0.2425	60
.25	1.0600	54	.75	0.6429	96	.25	0.2365	58
.26	1.0546	56	.76	0.6333	96	.26	0.2307	58
.27	+1.0490	57	.77	+0.6237	96	.27	+0.2249	57
.28	1.0433	59	.78	0.6141	96	.28	0.2192	55
.29	1.0374	61	.79	0.6045	95	.29	0.2137	55
0.30	+1.0313	63	0.80	+0.5950	95	1.30	+0.2082	54
.31	+1.0250	64	.81	+0.5855	95	.31	+0.2028	52
.32	1.0186	66	.82	0.5760	94	.32	0.1976	52
.33	1.0120	68	.83	0.5666	94	.33	0.1924	51
.34	+1.0052	69	.84	+0.5572	93	.34	+0.1873	49
.35	0.9983	71	.85	0.5479	93	.35	0.1824	49
.36	0.9912	72	.86	0.5386	93	.36	0.1775	48
.37	+0.9840	73	.87	+0.5293	91	.37	+0.1727	47
.38	0.9767	75	.88	0.5202	92	.38	0.1680	46
.39	0.9692	77	.89	0.5110	90	.39	0.1634	45
0.40	+0.9615	77	0.90	+0.5020	90	1.40	+0.1589	44
.41	+0.9538	79	.91	+0.4930	90	.41	+0.1545	43
.42	0.9459	80	.92	0.4840	88	.42	0.1502	42
.43	0.9379	81	.93	0.4752	88	.43	0.1460	41
.44	+0.9298	83	.94	+0.4664	88	.44	+0.1419	41
.45	0.9215	83	.95	0.4576	86	.45	0.1378	39
.46	0.9132	85	.96	0.4490	86	.46	0.1339	39
.47	+0.9047	85	.97	+0.4404	85	.47	+0.1300	38
.48	0.8962	87	.98	0.4319	84	.48	0.1262	37
.49	0.8875	87	.99	0.4235	84	.49	0.1225	36
0.50	+0.8788		1.00	+0.4151		1.50	+0.1189	

(продолжение)

x	$\Phi(x)_1$	Раз- ность	x	$\Phi(x)_1$	x	$\Phi(x)_1$	x	$\Phi(x)_1$
1.50	+ 0.1189	35	2.00	+ 0.0207	2.50	+ 0.0022	3.00	+ 0.0001
.51	+ 0.1154	34	.01	+ 0.0199	.51	+ 0.0021	.01	+ 0.0001
.52	0.1120	34	.02	0.0191	.52	0.0020	.02	0.0001
.53	0.1086	33	.03	0.0183	.53	0.0019	.03	0.0001
.54	+ 0.1053	32	.04	+ 0.0176	.54	+ 0.0018	.04	+ 0.0001
.55	0.1021	31	.05	0.0169	.55	0.0017	.05	0.0001
.56	0.0990	31	.06	0.0162	.56	0.0016	.06	0.0001
.57	+ 0.0959	29	.07	+ 0.0155	.57	+ 0.0015	.07	+ 0.0001
.58	0.0930	29	.08	0.0149	.58	0.0015	.08	0.0001
.59	0.0901	29	.09	0.0143	.59	0.0014	.09	0.0001
1.60	+ 0.0872	27	2.10	+ 0.0137	2.60	+ 0.0013	3.10	+ 0.0001
.61	+ 0.0845	27	.11	+ 0.0132	.61	+ 0.0012	.11	+ 0.0001
.62	0.0818	26	.12	0.0126	.62	0.0012	.12	0.0001
.63	0.0792	26	.13	0.0121	.63	0.0011	.13	0.0001
.64	+ 0.0766	25	.14	+ 0.0116	.64	+ 0.0011	.14	+ 0.0001
.65	0.0741	24	.15	0.0111	.65	0.0010	.15	0.0001
.66	0.0717	23	.16	0.0106	.66	0.0010	.16	0.0001
.67	+ 0.0694	23	.17	+ 0.0102	.67	+ 0.0009	.17	0.0000
.68	0.0671	22	.18	0.0097	.68	0.0009	.18	
.69	0.0649	22	.19	0.0093	.69	0.0008	.19	
1.70	+ 0.0627	21	2.20	+ 0.0089	2.70	+ 0.0008	3.20	
.71	+ 0.0606	20	.21	+ 0.0085	.71	+ 0.0007	.21	
.72	0.0596	20	.22	0.0082	.72	0.0007	.22	
.73	0.0566	20	.23	0.0078	.73	0.0007	.23	
.74	+ 0.0546	18	.24	+ 0.0075	.74	+ 0.0006	.24	
.75	0.0528	18	.25	0.0071	.75	0.0006	.25	
.76	0.0510	18	.26	0.0068	.76	0.0006	.26	
.77	+ 0.0492	17	.27	+ 0.0065	.77	+ 0.0005	.27	
.78	0.0475	17	.28	0.0062	.78	0.0005	.28	
.79	0.0458	16	.29	0.0060	.79	0.0005	.29	
1.80	+ 0.0442	16	2.30	+ 0.0057	2.80	+ 0.0004	3.30	
.81	+ 0.0426	15	.31	+ 0.0054	.81	+ 0.0004	.31	
.82	0.0411	15	.32	0.0052	.82	0.0004	.32	
.83	0.0396	14	.33	0.0050	.83	0.0004	.33	
.84	+ 0.0382	14	.34	+ 0.0047	.84	+ 0.0004	.34	
.85	0.0368	13	.35	0.0045	.85	0.0003	.35	
.86	0.0355	13	.36	0.0043	.86	0.0003	.36	
.87	+ 0.0342	13	.37	+ 0.0041	.87	+ 0.0003	.37	
.88	0.0329	12	.38	0.0039	.88	0.0003	.38	
.89	0.0317	12	.39	0.0037	.89	0.0003	.39	
1.90	+ 0.0305	11	2.40	+ 0.0036	2.90	+ 0.0003	3.40	
.91	+ 0.0294	11	.41	+ 0.0034	.91	+ 0.0002	.41	
.92	0.0283	11	.42	0.0032	.92	0.0002	.42	
.93	0.0272	10	.43	0.0031	.93	0.0002	.43	
.94	+ 0.0262	10	.44	+ 0.0029	.94	+ 0.0002	.44	
.95	0.0252	10	.45	0.0028	.95	0.0002	.45	
.96	0.0242	9	.46	0.0027	.96	0.0002	.46	
.97	+ 0.0233	9	.47	+ 0.0025	.97	+ 0.0002	.47	
.98	0.0224	9	.48	0.0024	.98	0.0002	.48	
.99	0.0215	8	.49	0.0023	.99	0.0001	.49	
2.00	+ 0.0207		2.50	+ 0.0022	3.00	+ 0.0001	3.50	

Таблица 2-я.

Таблицы функции $\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

(Czuber. Wahrscheinlichkeitsrechnung, стр. 385.)

x	Φ(x)	раз- ность	x	Φ(x)	раз- ность	x	Φ(x)	разность
0,00	0,0000000	112833	0,40	0,4283922	95768	0,80	0,7421010	59023
0,01	0,0112833	112811	0,41	0,4379690	94986	0,81	0,7480033	58075
0,02	0,0225644	112766	0,42	0,4474678	94191	0,82	0,7538108	57130
0,03	0,0338410	112699	0,43	0,4568867	93384	0,83	0,7595238	56189
0,04	0,0451109	112609	0,44	0,4662251	92567	0,84	0,7651427	55253
0,05	0,0563712	112497	0,45	0,4754818	91737	0,85	0,7706680	54322
0,06	0,0676215	112362	0,46	0,4846555	90897	0,86	0,7761022	53396
0,07	0,0788577	112204	0,47	0,4937452	90046	0,87	0,7814398	52475
0,08	0,0900781	112025	0,48	0,5027498	89185	0,88	0,7866873	51559
0,09	0,1012906	111824	0,49	0,5116683	88316	0,89	0,7918432	50650
0,10	0,1124630	111600	0,50	0,5204999	87438	0,90	0,7969082	49746
0,11	0,1236230	111354	0,51	0,5292437	86550	0,91	0,8018828	48849
0,12	0,1347584	111087	0,52	0,5378987	85654	0,92	0,8067677	47958
0,13	0,1458671	110799	0,53	0,5464641	84751	0,93	0,8115635	47075
0,14	0,1569470	110489	0,54	0,5549392	83841	0,94	0,8162710	46198
0,15	0,1679959	110158	0,55	0,5633233	82924	0,95	0,8208908	45328
0,16	0,1790117	109806	0,56	0,5716157	82001	0,96	0,8254236	44467
0,17	0,1899923	109434	0,57	0,5798158	81071	0,97	0,8298703	43612
0,18	0,2009357	109041	0,58	0,5879229	80136	0,98	0,8342315	42766
0,19	0,2118398	108627	0,59	0,5959365	79196	0,99	0,8385081	41927
0,20	0,2227202	108193	0,60	0,6038561	78251	1,00	0,8427008	41097
0,21	0,2335218	107740	0,61	0,6116812	77302	1,01	0,8468105	40275
0,22	0,2442958	107267	0,62	0,6194114	76349	1,02	0,8508380	39462
0,23	0,2550225	106775	0,63	0,6270463	75394	1,03	0,8547842	38657
0,24	0,2657000	106263	0,64	0,6345857	74435	1,04	0,8586499	37861
0,25	0,2763263	105734	0,65	0,6420292	73473	1,05	0,8624360	37075
0,26	0,2868997	105185	0,66	0,6493765	72510	1,06	0,8661435	36297
0,27	0,2974182	104618	0,67	0,6566275	71545	1,07	0,8697732	35529
0,28	0,3078890	104034	0,68	0,6637820	70579	1,08	0,8733261	34769
0,29	0,3182834	103433	0,69	0,6708399	69611	1,09	0,8768030	34020
0,30	0,3286267	102814	0,70	0,6778010	68644	1,10	0,8802050	33280
0,31	0,3389081	102178	0,71	0,6846654	67676	1,11	0,8835330	32549
0,32	0,3491259	101526	0,72	0,6914330	66708	1,12	0,8867879	31828
0,33	0,3592785	100859	0,73	0,6981038	65742	1,13	0,8899707	31116
0,34	0,3693644	100175	0,74	0,7046780	64776	1,14	0,8930823	30415
0,35	0,3793819	99477	0,75	0,7111556	63811	1,15	0,8961238	29724
0,36	0,3893296	98763	0,76	0,7175367	62849	1,16	0,8990962	29042
0,37	0,3992059	98034	0,77	0,7238216	61888	1,17	0,9020004	28370
0,38	0,4090093	97292	0,78	0,7300104	60931	1,18	0,9048374	27709
0,39	0,4187385	96537	0,79	0,7361035	59975	1,19	0,9076068	27057

(продолжение)

x	$\Phi(x)$	Разность	x	$\Phi(x)$	Разность	x	$\Phi(x)$	Разность
1.20	0,9103140	26415	1.70	0,9837904	6166	2.20	0,9981372	872
1.21	0,9129555	25784	1.71	0,9844070	5958	2.21	0,9982244	835
1.22	0,9155339	25162	1.72	0,9850028	5757	2.22	0,9983079	799
1.23	0,9180501	24551	1.73	0,9855785	5561	2.23	0,9983878	764
1.24	0,9205052	23940	1.74	0,9861346	5371	2.24	0,9984642	731
1.25	0,9229001	23358	1.75	0,9866717	5186	2.25	0,9985373	698
1.26	0,9252359	22777	1.76	0,9871903	5007	2.26	0,9986071	668
1.27	0,9275136	22206	1.77	0,9876910	4832	2.27	0,9986739	638
1.28	0,9297342	21645	1.78	0,9881742	4664	2.28	0,9987377	609
1.29	0,9318987	21093	1.79	0,9886406	4499	2.29	0,9987986	582
1.30	0,9340080	20552	1.80	0,9890905	4340	2.30	0,9988568	556
1.31	0,9360632	20020	1.81	0,9895245	4186	2.31	0,9989124	531
1.32	0,9380652	19498	1.82	0,9899481	4036	2.32	0,9989655	507
1.33	0,9400150	18987	1.83	0,9903467	3892	2.33	0,9990162	484
1.34	0,9419137	18485	1.84	0,9907359	3751	2.34	0,9990646	461
1.35	0,9437622	17992	1.85	0,9911110	3615	2.35	0,9991107	441
1.36	0,9455614	17510	1.86	0,9914725	3482	2.36	0,9991548	420
1.37	0,9473124	17036	1.87	0,9918207	3355	2.37	0,9991968	401
1.38	0,9490160	16573	1.88	0,9921562	3231	2.38	0,9992369	382
1.39	0,9506733	16118	1.89	0,9924793	3111	2.39	0,9992751	364
1.40	0,9522851	15673	1.90	0,9927904	2995	2.40	0,9993115	347
1.41	0,9538524	15238	1.91	0,9930899	2883	2.41	0,9993462	331
1.42	0,9553762	14811	1.92	0,9933782	2775	2.42	0,9993793	315
1.43	0,9568578	14393	1.93	0,9936557	2664	2.43	0,9994108	300
1.44	0,9582966	13984	1.94	0,9939226	2568	2.44	0,9994408	286
1.45	0,9596950	13585	1.95	0,9941794	2469	2.45	0,9994694	272
1.46	0,9610535	13194	1.96	0,9944263	2374	2.46	0,9994966	260
1.47	0,9623729	12812	1.97	0,9946637	2283	2.47	0,9995226	246
1.48	0,9636541	12433	1.98	0,9948926	2194	2.48	0,9995472	235
1.49	0,9648979	12073	1.99	0,9951114	2109	2.49	0,9995707	223
1.50	0,9661052	11716	2.00	0,9953223	2025	2.50	0,9995930	213
1.51	0,9672768	11367	2.01	0,9955243	1947	2.51	0,9996143	202
1.52	0,9684135	11027	2.02	0,9957195	1868	2.52	0,9996345	192
1.53	0,9695102	10695	2.03	0,9959063	1795	2.53	0,9996537	183
1.54	0,9705857	10370	2.04	0,9960858	1733	2.54	0,9996720	173
1.55	0,9716227	10054	2.05	0,9962581	1654	2.55	0,9996893	165
1.56	0,9726281	9745	2.06	0,9964235	1587	2.56	0,9997058	157
1.57	0,9736026	9444	2.07	0,9965822	1522	2.57	0,9997215	149
1.58	0,9745470	9150	2.08	0,9967344	1461	2.58	0,9997364	141
1.59	0,9754620	8864	2.09	0,9968805	1400	2.59	0,9997505	135
1.60	0,9763484	8585	2.10	0,9970205	1343	2.60	0,9997640	127
1.61	0,9772069	8312	2.11	0,9971548	1288	2.61	0,9997767	121
1.62	0,9780381	8048	2.12	0,9972836	1234	2.62	0,9997888	115
1.63	0,9788429	7789	2.13	0,9974070	1183	2.63	0,9998003	109
1.64	0,9796218	7538	2.14	0,9975253	1133	2.64	0,9998112	103
1.65	0,9803756	7293	2.15	0,9976386	1086	2.65	0,9998215	98
1.66	0,9811049	7055	2.16	0,9977472	1039	2.66	0,9998313	93
1.67	0,9818104	6824	2.17	0,9978511	994	2.67	0,9998406	88
1.68	0,9824928	6598	2.18	0,9979505	954	2.68	0,9998494	84
1.69	0,9831526	6378	2.19	0,9980459	913	2.69	0,9998578	79

(продолжение)

x	Ф(x)	Разность	-x	Ф(x)	Разность	x	Ф(x)
2,70	0,9998657	75	3,20	0,9999940	4	3,70	0,99999983285
2,71	0,9998732	71	3,21	0,9999944	3	3,71	0,99999984517
2,72	0,9998802	67	3,22	0,9999947	4	3,72	0,99999985663
2,73	0,9998870	63	3,23	0,9999951	3	3,73	0,99999986726
2,74	0,9998933	61	3,24	0,9999954	3	3,74	0,99999987712
2,75	0,9998994	57	3,25	0,9999957	3	3,75	0,99999988629
2,76	0,9999051	54	3,26	0,9999960	2	3,76	0,99999989477
2,77	0,9999105	51	3,27	0,9999962	3	3,77	0,99999990265
2,78	0,9999156	48	3,28	0,9999965	2	3,78	0,99999990995
2,79	0,9999204	46	3,29	0,9999967	2	3,79	0,99999991672
2,80	0,9999250	43	3,30	0,9999969	2	3,80	0,99999992300
2,81	0,9999293	41	3,31	0,9999971	2	3,81	0,99999992881
2,82	0,9999334	38	3,32	0,9999973	2	3,82	0,99999993421
2,83	0,9999372	37	3,33	0,9999975	2	3,83	0,99999993921
2,84	0,9999409	34	3,34	0,9999977	1	3,84	0,99999994383
2,85	0,9999443	33	3,35	0,9999978	2	3,85	0,99999994812
2,86	0,9999476	31	3,36	0,9999980	1	3,86	0,99999995203
2,87	0,9999507	29	3,37	0,9999981	1	3,87	0,99999995575
2,88	0,9999536	27	3,38	0,9999982	2	3,88	0,99999995915
2,89	0,9999563	26	3,39	0,9999984	1	3,89	0,99999996230
2,90	0,9999589	24	3,40	0,9999985	1	3,90	0,99999996521
2,91	0,9999613	23	3,41	0,9999986	1	3,91	0,99999996790
2,92	0,9999636	22	3,42	0,9999987	1	3,92	0,99999997039
2,93	0,9999658	21	3,43	0,9999988	1	3,93	0,99999997260
2,94	0,9999679	19	3,44	0,9999989	0	3,94	0,99999997482
2,95	0,9999698	18	3,45	0,9999989		3,95	0,99999997678
2,96	0,9999716	17	3,46	0,99999900780		3,96	0,99999997860
2,97	0,9999733	17	3,47	0,99999907672		3,97	0,99999998028
2,98	0,9999750	15	3,48	0,99999914101		3,98	0,99999998183
2,99	0,9999765	14	3,49	0,99999920097		3,99	0,99999998327
3,00	0,9999779	14	3,50	0,99999925681		4,0	0,99999998458
3,01	0,9999793	12	3,51	0,99999930905		4,1	0,99999998530
3,02	0,9999805	12	3,52	0,99999935786		4,2	0,99999998571
3,03	0,9999817	12	3,53	0,99999940296		4,3	0,99999998581
3,04	0,9999829	10	3,54	0,99999944519		4,4	0,99999998591
3,05	0,9999839	10	3,55	0,99999948452		4,5	0,99999998580
3,06	0,9999849	10	3,56	0,99999952115		4,6	0,99999998592
3,07	0,9999859	8	3,57	0,99999955527		4,7	0,99999998597
3,08	0,9999867	9	3,58	0,99999958703		4,8	0,99999998599
3,09	0,9999876	8	3,59	0,99999961661			
3,10	0,9999884	7	3,60	0,99999964414			
3,11	0,9999891	7	3,61	0,99999966975			
3,12	0,9999898	6	3,62	0,99999969356			
3,13	0,9999904	6	3,63	0,99999971574			
3,14	0,9999910	6	3,64	0,99999973636			
3,15	0,9999916	5	3,65	0,99999975551			
3,16	0,9999921	5	3,66	0,99999977333			
3,17	0,9999926	5	3,67	0,99999978990			
3,18	0,9999931	5	3,68	0,99999980528			
3,19	0,9999936	4	3,69	0,99999981957			

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО

МОСКВА — ЛЕНИНГРАД

- Сморodinцев, И. А.** Ферменты растительного и животного царства. Ч. I. Общая ферментология. 340 стр. Ц. 1 р. 50 к.
- Таллев, В. И.** Основы ботаники в общебиологическом изложении (эволюционном). 5-е изд. Берлин. 1923. Стр. 698. Ц. 4 р. 50 к.
- Тредвель, Ф.** Курс аналитической химии в двух томах. Т. I. Качественный анализ. Перевод с немецкого А. Комаровского. 3-е изд. 427 стр. Ц. 2 руб.
- Тредвель, Ф.** Курс аналитической химии в двух томах. Том II. Количественный анализ. Перевод с немецкого А. Комаровского. 3-е изд. 572 стр. Ц. 3 р.
- Уокер, Д.** Введение в физическую химию. С предисловием акад. Вальдена. Изд. 2-е. 350 стр. Цена 2 р. 60 к.
- Оаянс, Н.** Радиоактивность и современное учение о химических элементах. Перевод и дополнение Э. Шпольского. 121 стр. (Серия „Современные Проблемы Естествознания“). Ц. 40 к.
- Филиппченко, Ю. А., проф.** Изменчивость и методы ее изучения (основы биологической вариационной статистики). 240 стр. с 50 рис. в тексте. Ц. 1 р. 80 к.
- Шарвин, В. В.** Введение в химию. Краткий курс неорганической химии для высших учебных заведений и самообразования. 4-е изд. 416 стр. Ц. 2 руб.
- Шинкевич, В.** Биологические основы зоологии. Том II. С 179 рис. Издание 4-е, дополненное. 405—723 стр. Ц. 2 р. 25 к.
- Яковлев, Н. Н.** Учебник палеонтологии для высшей школы и для самообразования. Изд. 2-е. 447 стр. Ц. 2 р. 20 к.

Цены обозначены в червонцах по курсу Госбанка.

ТОРГОВЫЙ СЕКТОР ГОСУДАРСТВЕННОГО ИЗДАТЕЛЬСТВА

МОСКВА: Ильинка, Биржевая пл., | ЛЕНИНГРАД: Пресп. 25 Октября, 28.
Богоявленский пер., 4. Тел. 47-35. | Телефон № 5-49-32.

ОТДЕЛЕНИЯ: ВОЛОГДА, Площадь Свободы; ВОРОНЕЖ, Проспект Революции, 1-й дом Совета; КАЗАНЬ, Гостинодворская, Гостинный Двор; КИЕВ, Крещатик, 38; КОСТРОМА, Советская, 11; КРАСНОДАР, Красная, 33; НИЖНИЙ-НОВГОРОД, Б. Покровка, 12; ОДЕССА, ул. Ласкала, 12; ПЕНЗА, Интернациональная, 39-43; РОСТОВ-НА-ДОНУ, ул. Фридриха Энгельса, 106; САРАТОВ, ул. Республики, 48; ТАМБОВ, Коммунальная, 14; ТИФЛИС, Проспект Руставели, 16; ХАРЬКОВ, Московская, 20; ПЯТИГОРСК, книжный магазин, Советский пр., 48.

Цены обозначены в червонцах по курсу Госбанка.

МАГАЗИНЫ ТОРГСЕКТОРА: МОСКВА: 1) Советская площадь (под гост. „Дрезден“). Тел. 1-28-94; 2) Моховая, 17. (под гост. б. „Националь“). Тел. 1-31-50; 3) Ул. Герцена (Б. Никитская), 13, здание Консерватории. Тел. 2-64-96; 4) Никольская ул., 3. Тел. 49-51; 5) Серпуховская площ., 143. Тел. 2-84-82; 6) Кузнецкий мост, 12. Тел. 1-01-35; 7) Лядин пер., 11. 8) М. Харитоньевский пер., 4.

ЛЕНИНГРАД: 1. Проспект 25 Октября, 28; 2. Ул. Володарского, 53; 3. Проспект 25 Октября, 13.

Все новости о книгах — в бюллетенях Торгсектора; очередной номер бюллетеня Торгсектора высылается по получении 10 коп. почтовыми марками.

30 - 326769

