

Министерство образования и науки Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Московский государственный технический университет имени Н.Э. Баумана»  
(национальный исследовательский университет)

---

*Кафедра «Прикладная математика»*



## *Реферат*

по дисциплине «История математики»

## **История развития математической статистики и теории случайных процессов**

*Выполнил* студент группы ФН2-81Б

*Разумов Т.Е.*

*Принял* к.ф-м.н., доцент кафедры ФН-2

*Новожилова О.В.*

# Оглавление

<b>Введение</b> . . . . .	3
<b>1. Математическая статистика</b> . . . . .	3
1.1. Зарождение и формирование статистики . . . . .	3
1.2. Метод наименьших квадратов . . . . .	7
1.3. Метод Монте-Карло . . . . .	9
1.4. Корреляция и регрессия . . . . .	10
1.5. Точечные оценки параметров . . . . .	12
1.6. Интервальное оценивание . . . . .	13
1.7. Проверка статистических гипотез . . . . .	13
1.8. Демография и смежные вопросы . . . . .	15
<b>2. Случайные процессы</b> . . . . .	15
2.1. Ветвящиеся процессы . . . . .	17
2.2. Марковские процессы . . . . .	18
2.3. Винеровские процессы . . . . .	18
<b>3. Вклад А.Н. Колмогорова</b> . . . . .	19
<b>Заключение</b> . . . . .	21
<b>Список литературы</b> . . . . .	22

## Введение

Статистика – это наука о том, как обрабатывать данные. В современном обществе статистика выполняет важную роль в механизме управления экономикой. Независимо от уровня и стадии экономического развития, характера политической системы, статистика на протяжении сотен лет своего существования всегда выступала как необходимый и эффективный инструмент государственного управления и одновременно как наука, исследующая количественную сторону массовых явлений.

Статистические методы основаны на вероятностных моделях. Они активно применяются в технических исследованиях, экономике, теории и практике управления (менеджмента). А также в социологии, медицине, геологии, истории и т.д. С обработкой результатов наблюдений, измерений, испытаний, опытов, анализов имеют дело специалисты во всех отраслях практической деятельности, почти во всех областях научных исследований.

Развитие наукоемких технологий, как правило, основано на применении высоких статистических технологий организации и управления производством. Особенно активно они используются в высокотехнологичных отраслях промышленности. Без вероятностно-статистических методов немыслимы оценка и анализ риска, страхование, финансовая деятельность. Инженеры, менеджеры, экономисты, социологи, врачи, психологи, историки успешно применяют интеллектуальные инструменты принятия решений, основанные на вероятности и статистике.

Особенность статистики заключается в том, что статистические данные сообщаются в количественной форме, т.е. статистика говорит языком цифр, отображающих общественную жизнь во всем многообразии ее проявлений. При этом статистику прежде всего интересуют те выводы, которые можно сделать на основе анализа надлежащим образом собранных и обработанных цифровых данных.

Выполняя самые разнообразные функции сбора, систематизации и анализа сведений, характеризующих экономическое и социальное развитие общества, она всегда играла роль главного поставщика факторов для научно-исследовательских, управленческих и прикладных практических нужд различного рода структур, организаций и населения.

Понятие случайного процесса введено в XX столетии и связано с именами А.Н. Колмогорова (1903-1987), А.Я. Хинчина (1894-1959), Е.Е. Слуцкого (1880-1948), Н. Винера (1894-1965). Это понятие в наши дни является одним из центральных не только в теории вероятностей, но также в естествознании, инженерном деле, экономике, организации производства, теории связи.

Теория случайных процессов принадлежит к категории наиболее быстро развивающихся математических дисциплин. Несомненно, что это обстоятельство в значительной мере определяется ее глубокими связями с практикой.

## 1. Математическая статистика

### 1.1. Зарождение и формирование статистики

Статистика имеет древние корни и многовековую историю развития. Она зародилась как результат обобщения уже достаточно развитой статистической практики, вызванной потребностями развития общества, например: подсчет населения, скота,

учет земельных угодий, имущества и т.д. Сам термин «статистика» произошел от латинского слова «статус», что означает «определенное положение вещей», а употреблялся он первоначально в значении слова «государствование».

Некоторые сведения применения числовых данных относящихся к тем или иным явлениям проявлялись ещё в глубокой древности. В Китае более чем за 2 тысячи лет до н.э. производились исчисления населения по полу и возрасту, а также собирались сведения о состоянии промышленности и сельского хозяйства. Упоминания о статистических обследованиях встречаются и в библейские времена. В Древнем Риме велась статистика численности населения и имущественного положения граждан, проводились цензы (учеты) свободных граждан и их имущества. В Европе в конце IX в. проводились первые учетные операции: инвентаризация королевских имений, учет жителей, пригодных к военной службе. Первыми и основными учетно-статистическими источниками на Руси были летописи, в которых уже в IX-XI вв. упоминается о сборе различной информации. Так, приводятся учетные данные о возникновении и развитии городских поселений, расположенных на водных путях, о наличии в них храмов, церквей, монастырей, жилых строений. Однако сбор числовых данных в государствах древнего мира был несовершенен. В этот период статистические операции, как правило, проводились в исключительных случаях и в основном в военной и финансовой сферах. Позднее потребность в статистических операциях порождалась необходимостью стимулирования роста народонаселения, производительных сил страны, регулирования потребления.

Развитие общественного производства, внутренней и внешней торговли, торговых и международных товарно-денежных отношений увеличило потребность в статистической информации. Это расширило сферу деятельности статистики, вело к совершенствованию ее приемов и методов, явилось стимулом для дальнейшего формирования учета и статистики.

Однако если сбор статистических данных начался в самой глубокой древности, то их обработка и анализ, т.е. зарождение статистики как науки, относятся к более позднему периоду. Так со временем многообразная практика учетно-статистических работ стала подвергаться теоретическим обобщениям. Началось формирование статистической науки.

Первой работой, с которой начинается история математической статистики, следует назвать книгу Джона Граунта (1620-1674 гг.) «Естественные и политические наблюдения, перечисленные в прилагаемом оглавлении и сделанные над бюллетенями смертности. По отношению к управлению, религии, торговле, росту, воздуху, болезням и разным изменениям означенного города» 1662 г. Основная задача, которая заинтересовала Граунта, состояла в указании метода, который позволял бы установить с достаточной точностью возрастной состав населения города в результате наблюдений за возрастом умерших. С этой целью им были проанализированы результаты 229 250 регистрации смертей в Лондоне, происшедших за 20 лет. В этой книге ученый ввел представление о частоте события. Для развития математической статистики это обстоятельство сыграло огромную роль, как, впрочем, и его замечание: «... мы хотели бы отметить, что некоторые из случайностей имеют постоянное отношение к числу всех смертей». Здесь Граунт вплотную подошел к представлению о статистической устойчивости средних. Им была составлена первая таблица смертности (см. таблица 1). В этой таблице поражает огромная детская и юношеская смертность: только 64% в ту пору доживали до 6 лет и только 40% — до 16

лет. Граунт прекрасно понимал, что точность его выводов тем больше, чем больше наблюдений имеется для обработки. Именно в связи с этим он отметил, что недостаточно ограничиваться обработкой бюллетеней смертности только за одну неделю для получения полноценных выводов о составе населения.

Идеи Граунта развивали и другие ученые, например, Вильям Петти (1623-1687 гг.), наиболее значимыми работами которого являются: «Политическая арифметика» 1676 г. и «Замечания относительно Дублинских бюллетеней смертности» 1683 г. В этих работах он подсчитывает необходимое количество людей различных профессий, как в настоящее время, так и в будущем, величину необходимых налогов, величину народного богатства и доходов, количество населения Лондона и т.п. Конечно ни Граунт, ни Петти не пользовались теорией вероятностей, но применявшиеся ими понятия и методы, по существу, были тесно связаны с ней, а поставленные в них вопросы стимулировали развитие этой науки. Необходимая связь между статистикой и теорией вероятностей была установлена только Яковом Бернулли, который впоследствии ввёл значительную часть современных понятий теории вероятностей и сформулировал первый вариант закона больших чисел. Якоб Бернулли подготовил монографию в этой области, однако издать её не успел. Она была напечатана посмертно, в 1713 году, его братом Николаем, под названием «Искусство предположений» (*Ars coniectandi*). Это содержательный трактат по теории вероятностей, статистике и их практическому применению, итог комбинаторики и теории вероятностей XVII века.

Возраст	Процент доживших людей
6 лет	64%
16 лет	40%
26 лет	25%
36 лет	16%
46 лет	10%
56 лет	6%
66 лет	3%
76 лет	1%
86 лет	0%

Таблица 1

Математическая статистика как наука начинается с работ знаменитого немецкого математика Карла Фридриха Гаусса, который на основе теории вероятностей исследовал и обосновал метод наименьших квадратов, созданный им в 1795 г. и примененный для обработки астрономических данных (с целью уточнения орбиты малой планеты Церера). Его именем часто называют одно из наиболее популярных распределений вероятностей – нормальное, а в теории случайных процессов основной объект изучения – гауссовские процессы.

К последней трети XIX века, некоторые науки (например, молекулярная физика) достигли такого уровня, что использование в них теории вероятности и математической статистики стало необходимым. Важную роль в развитии статистики в XIX веке сыграли Адольф Кетле, Фрэнсис Гальтон и Людвиг Больцман.

Особая роль в развитии социальной статистики принадлежит бельгийскому ученому-естествоиспытателю и математику Адольфу Кетле. Он является создателем вероятностных методов обработки социальной информации, инициатором создания наци-

ональных статистических обществ в Англии и Франции и Международной статистической ассоциации. Свои идеи он изложил в работе «Социальная физика, или опыт исследования о развитии человеческих способностей» (1835). С помощью теории вероятности он хотел разработать науку об обществе, имеющую статистическое и естественнонаучное обоснование, которую он назвал «социальной физикой». Кетле устанавливал определенные статистические закономерности событий и действий и соотносил их с состоянием общества таким образом, что становились понятны законы развития общества. Из статистического факта наличия устойчивых числовых корреляций между видами преступлений, полом, происхождением, возрастом, местом проживания и другими характеристиками преступника Кетле делал вывод о том, что некоторые количество и некоторые виды преступлений сопровождают общество с необходимостью закона природы. Стало афоризмом утверждение Кетле, прозвучавшее в докладе 1831 г., о том, что «общество подготавливает преступления, а преступник есть только орудие».

В 1859 г. книга Дарвина произвела революцию в биологии, и вскоре английский статистик, психолог, антрополог и метеоролог Фрэнсис Гальтон, двоюродный брат Чарльза Дарвина, заложил основ генетики и разработал методы статистической обработки результатов наблюдений (в частности, метод вычисления корреляций между случайными величинами). Он ввел в статистику понятия средней регрессии, коэффициента корреляции, создал биометрическую шкалу, заложил основы генетики человека. Хотя термин «генетика» вошел в обиход только в 1905 г., результаты Гальтона привлекли всеобщее внимание еще в XIX веке.

Австралийский физик-теоретик Людвиг Больцман является основателем статистической механики и молекулярно-кинетической теории. Он показал статистический характер второго начала термодинамики, связав энтропию замкнутой системы с числом возможных микросостояний, реализующих данное макросостояние.

В конце XIX в. – начале XX в. крупный вклад в математическую статистику внесли английские исследователи, прежде всего К. Пирсон (1857-1936) и Р.А. Фишер (1890-1962). В частности, Пирсон разработал критерий «хи-квадрат» проверки статистических гипотез, а Фишер — дисперсионный анализ, теорию планирования эксперимента, метод максимального правдоподобия оценки параметров.

В 30-е годы XX в. поляк Ежи Нейман (1894-1977) и англичанин Э. Пирсон развили общую теорию проверки статистических гипотез, а советские математики академик А.Н. Колмогоров (1903-1987) и член-корреспондент АН СССР Н.В. Смирнов (1900-1966) заложили основы непараметрической статистики. В сороковые годы XX в. русский А. Вальд (1902-1950) построил теорию последовательного статистического анализа.

Математическая статистика бурно развивается и в настоящее время. Так, за последние 40 лет можно выделить четыре принципиально новых направления исследований:

- разработка и внедрение математических методов планирования экспериментов;
- развитие статистики объектов нечисловой природы как самостоятельного направления в прикладной математической статистике;
- развитие статистических методов, устойчивых по отношению к малым отклонениям от используемой вероятностной модели;
- широкое развертывание работ по созданию компьютерных пакетов программ, предназначенных для проведения статистического анализа данных.

В современном мире теория вероятностей и математическая статистика – основа вероятностно-статистических методов обработки данных. А данные мы обрабатываем и анализируем прежде всего для принятия решений. Чтобы воспользоваться современным математическим аппаратом, необходимо рассматриваемые задачи выразить в терминах вероятностно-статистических моделей.

Применение конкретного вероятностно-статистического метода для реальных моделей состоит из трех этапов:

- 1) переход от экономической, управленческой, технологической реальности к абстрактной математико-статистической схеме, т.е. построение вероятностной модели системы управления, технологического процесса, процедуры принятия решений, в частности по результатам статистического контроля, и т.п.
- 2) проведение расчетов и получение выводов чисто математическими средствами в рамках вероятностной модели;
- 3) интерпретация математико-статистических выводов применительно к реальной ситуации и принятие соответствующего решения (например, о соответствии или несоответствии качества продукции установленным требованиям, необходимости наладки технологического процесса и т.п.), в частности, заключения (о доле дефектных единиц продукции в партии, о конкретном виде законов распределения контролируемых параметров технологического процесса и др.).

Расскажем более подробно про работы великих ученых, внесшие наибольший вклад в развитие и становление современной математической статистики.

## 1.2. Метод наименьших квадратов

До начала XIX в. учёные не имели определённых правил для решения системы уравнений, в которой число неизвестных меньше, чем число уравнений; до этого времени употреблялись частные приёмы, зависевшие от вида уравнений и от остроумия вычислителей, и потому разные вычислители, исходя из тех же данных наблюдений, приходили к различным выводам. Помимо этого из-за неизбежных ошибок измерений часто кажется, что теоретические формулы и эмпирические данные часто противоречат друг другу. Метод наименьших квадратов (МНК) был изобретен в первом десятилетии XIX века практически одновременно тремя учеными, два из которых – это математики А.М. Лежандр (1752-1833) и К.Ф. Гаусс (1777-1855). В том же десятилетии на другом берегу Атлантики третий автор изобретения Р.А. Эдрейн (1775-1843) напечатал свой вывод нормального закона распределения вероятностей ошибок измерений и применил его «к установлению принципа наименьших квадратов», позволив уменьшить влияние ошибок измерений.

В 1806 году А.М. Лежандр опубликовал алгоритм МНК и небольшой пример его реализации, не дав обоснования возникновению идеи метода, а привел алгебраическое решение задачи и пример «о применении способа к меридианному измерению во Франции», наложив на искомые поправки ограничение в форме

$$\sum_i (f(x_i) - y_i)^2 = \sum_i v_i^2 \rightarrow \min, \quad (1)$$

К.Ф. Гаусс в 1809 году публикует свою работу «Теория движения небесных тел», в которой он вначале дает вероятностное обоснование МНК, а далее указывает, что

«принцип может рассматриваться независимо от теории вероятностей». Несколько ниже, в том же разделе, Гаусс отдает дань своему коллеге, отмечая, что «данный принцип, которым он пользуется ещё с 1795 г., также недавно был изложен в трудах известного Лежандра, где приведено много других его свойств».

Один из недостатков МНК заключался в разной точности нахождения значений неизвестной функции в различных узлах, что могло быть следствием резкой осцилляции исследуемой функции или же других её особенностей. Для решения данного недостатка Гауссом была введена мера точности  $h = \frac{1}{\sqrt{2}\sigma}$ , где  $\sigma$  – среднеквадратичное отклонение. Термин «мера точности» заимствован из теории ошибок измерений: чем точнее измерение, тем больше мера точности. Однако с её помощью удавалось регулировать лишь среднюю погрешность изменения.

Учет неравноточности в отдельных узлах с помощью «весов», введение которых впервые было предложено Р. Котсом (1682-1716) до разработки МНК, проводился графически. «Веса» Котса отличались, по сути, от «меры точности»  $h$ , использовавшейся Гауссом. В последствие, мера неравноточности наблюдений  $p$ , как величина обратно пропорциональная квадратам этих ошибок, сохранила за собой название «вес». Это привело к усложнению ограничения (1) до вида

$$\sum_i p_i v_i^2 \rightarrow \min,$$

причем, как было показано позднее, для весов  $p_i$  должно выполняться следующее соотношение

$$\sum_i p_i = 1.$$

П.С. Лаплас (1749-1827) оценил достоинство МНК и занимался вопросом вероятностного обоснования метода на основе нормального закона, который называют законом Лапласа-Гаусса. В 1812 г. он опубликовал свой основной труд по теории вероятности, посвятив его «великому Наполеону». На протяжении всей четвертой главы его книги излагается метод исчисления ошибок. С того времени метод наименьших квадратов развился как новый раздел математики. Однако возможности метода попросту переоценивают и часто его используют тогда, когда более подходящими были бы другие методы. На эту проблему обращал внимание еще Коши во время «дебатов» в Бьенэме (в ходе диспута Коши использовал плотность вероятности  $1/(\pi(10+x^2))$ ), названную позднее его именем, хотя он и не был первым ученым, применившим плотность Коши. Позднее Лаплас предположил, что при неограниченном увеличении числа наблюдений МНК дает наилучшие результаты даже при законе распределения, отличном от нормального.

Важно подчеркнуть, что изначально МНК не использовался для подбора функциональной основы математической модели параллельно с оцениванием её параметров. В 1821 году Гаусс опубликовал «Сообщения (Anzeigen)», в которых он «исходил из такой же точки зрения, как и Лаплас, но, в отличие от него, не приблизительно, а со всей математической строгостью доказал, что функция вероятности может быть какой угодно, и число наблюдений может быть большим или малым». Этим Гаусс вновь подтвердил свою мысль о самодостаточности МНК «независимо от теории вероятностей».

Двухвековая история применения МНК в астрономии и геодезии, использовавших этот метод для аппроксимации (точностной оптимизации) данных наблюдений в



моделях с известной функциональной структурой, демонстрируют его вычислительную простоту и универсальность. Метод успешно применяется как для обработки и анализа фиксированных в пространстве данных, так и для свободных геодезических построений. Технологии спутникового позиционирования (GPS, ГЛОНАС и другие) так же успешно опираются на алгоритм МНК-оптимизации данных.

Благодаря работам А.А. Маркова в XIX-XX веках МНК внедрился в математическую статистику, или, лучше сказать математическая статистика взяла на вооружение МНК. С тех пор и по сей день он является важной и естественной частью математической статистики. Оценки, а правильнее «оценивающие функции» (ОФ) генеральной совокупности, как это определяется в математической статистике, должны быть состоятельными и несмещёнными. Для одного и того же параметра можно построить несколько несмещённых ОФ. Предпочтение отдаётся той ОФ, дисперсия которой минимальна. При этом ОФ должна быть эффективной, то есть использовать всю информацию (согласно условию Р. Фишера), имеющуюся в выборке, а это происходит лишь для тех МНК-оценок, которые получены по выборке из нормальной генеральной совокупности. Усилиями Ю. Неймана, Ф.Дэвида, А. Эйткена, С. Рао было получено множество немаловажных результатов в этой области.

В настоящее время МНК является одним из базовых методов регрессионного анализа для оценки неизвестных параметров регрессионных моделей по выборочным данным.

### 1.3. Метод Монте-Карло

Идея метода Монте-Карло впервые появилась в 1777 г. в работе французского естествоиспытателя Бюффона, где излагается метод оценки числа  $\pi$  путем бросания иглы наугад. Хотя идея метода довольно стара, его широкое применение началось в 1949 г., когда Е. Нейман, С. Улам и Э. Ферми использовали метод Монте-Карло для приближенного решения трудных вычислительных задач, связанных с созданием атомных реакторов.

Это численный метод, основанный на моделировании случайных величин и построении статистических оценок для искомых величин. Вместо того, чтобы описывать случайное явление с помощью аналитических зависимостей, производят его моделирование с помощью процедуры, которая дает результат, т.е. одну реализацию случайного явления. После неоднократного проведения такой процедуры, получают статистический материал, который можно обработать обычными методами математической статистики. Нередко такой прием оказывается проще, чем попытки построения аналитической модели явления и исследования зависимостей между его параметрами на этой модели. Для сложных операций, в которых задействовано большое число элементов, а случайные факторы сложным образом взаимодействуют между собой, метод статистических испытаний, как правило, оказывается проще аналитического.

Название метода объясняется тем, что в нем применяются последовательности случайных чисел, в качестве которых могли бы выступать регулярно объявляемые результаты игр, проходимых в казино, например в Монте-Карло. Однако на практике случайные числа, необходимые для реализации метода, выдает сам компьютер. Следовательно, данное название метода, впервые использованное Н. Метрополисом и С. Уламом, скорее вводит в заблуждение непосвященных и вряд ли поможет вы-

играть в казино Монте-Карло.

Последовательности псевдослучайных чисел, генерируемые на ЭВМ специальными алгоритмами, применяются довольно широко. Их используют для численного интегрирования и решения дифференциальных уравнений, но главное – при моделировании на ЭВМ физических, химических, биологических, технических и экономических процессов, а также для решения задач уличного движения, транспортных и других оптимальных задач и создания астрономических моделей.

#### 1.4. Корреляция и регрессия

Исторически человеческие характеристики служили основой не только для изучения статистических частотных распределений, но и для формулировки математической корреляции. Корреляция – это степень, в которой изменения значения одной переменной приводят к изменению другой. Например, чем выше женщина, тем больше у нее должен быть размер обуви. Подобным же образом психологи обнаружили корреляцию между интеллектом родителей и школьной успеваемостью детей.

Понятие корреляции особенно полезно в ситуациях, когда между двумя переменными нет точной функциональной взаимозависимости. Например, представим себе, что одна переменная – максимальная дневная температура на юге Сибири, а другая – количество лесных пожаров в том регионе. Невозможно предсказать, какое количество лесных пожаров возникает при данной температуре, поскольку количество пожаров зависит и от других переменных, в частности, от влажности воздуха и от количества костров, которые расжигают люди. Иначе говоря, любому значению температуры соответствует разное количество лесных пожаров и наоборот. И все же математическое понятие коэффициента корреляции позволяет нам количественно изменить прочность отношений между двумя подобными переменными.

Сэр Фрэнсис Гальтон (1822-1911), введший в арсенал математиков коэффициент корреляции, был человеком сугубо практического склада и обычно предоставлял другим математикам доводить свои новаторские понятия до совершенства; особенно ему помогал в этом статистик Карл Пирсон (1857-1936). Вот как Гальтон объяснил понятие корреляции.

*«Длина локтя коррелирует с телостроением, поскольку длинный локоть обычно предполагает высокий рост. Если корреляция между ними очень тесная, то очень длинный локоть обычно предполагает очень высокий рост, однако, если бы она была не очень тесная, то очень длинный локоть в среднем связывался бы всего лишь с высоким, но не с очень высоким ростом, тогда как если бы она была нулевая, то очень длинный локоть не был бы связан ни с какими особенностями роста, а следовательно, в среднем, был бы связан с заурядным ростом.»*

В дальнейшем Пирсон дал более точное математическое определение коэффициента корреляции. Этот коэффициент определяется таким образом, что когда корреляция очень высока – то есть когда колебания одной переменной очень точно следуют за взлетами и падениями другой, – коэффициент приобретает значение 1. Если же две величины антикоррелированы, то есть одна величина возрастает, когда другая уменьшается, и наоборот, коэффициент равен -1. Если две переменные ведут себя так, будто другой и вовсе не существует, коэффициент корреляции равен 0 (например, поведение иных правительств, к сожалению, демонстрирует практически нулевую

корреляцию с пожеланиями народа, который они якобы представляют).

От выявления и вычисления корреляций в наши дни зависят и медицинские исследования, и экономические прогнозы. Например, связь между курением и раком кожи изначально была выявлена благодаря обнаружению и вычислению корреляций. Биржевые аналитики постоянно пытаются найти и вычислить корреляции между поведением рынка и другими переменными – и любое подобное открытие приносит фантастические прибыли.

Запишем коэффициент корреляции в общем виде. Пусть  $X$  и  $Y$  – две случайные величины, определенные на одном вероятностном пространстве. Тогда их коэффициент корреляции задается формулой:

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sqrt{\mathbb{D}[X] \cdot \mathbb{D}[Y]}} = \frac{\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]}{\sqrt{(\mathbb{E}[X^2] - (\mathbb{E}[X])^2) \cdot (\mathbb{E}[Y^2] - (\mathbb{E}[Y])^2)}},$$

где  $\text{cov}$  обозначает ковариацию,  $\mathbb{D}$  – дисперсию, а  $\mathbb{E}$  – математическое ожидание.

После знакомства с книгой Чарльза Дарвина «Происхождение видов» в 1859 г. Ф. Гальтона стала занимать мысль о том, почему люди из поколения в поколение не сильно различаются по внешнему виду и природным способностям. Это привело его к изучению наследственности. В частности, он занялся выяснением зависимости роста детей от роста родителей. По логике дети должны быть каждый раз очень похожи на своих родителей. Высокие родители должны иметь высоких детей, а низкорослые родители – детей низкого роста. При таком положении вещей через несколько поколений мы имели бы, с одной стороны, род великанов, а с другой – род карликов. Но вскоре в результате обширных статистических исследований и опытов над животными Ф. Гальтон убедился, что такой тенденции нет, а, скорее, напротив, дети очень высоких или очень низких родителей в среднем имеют менее высокий или соответственно менее низкий рост. Кроме того, отклонение роста детей не так велико, как отклонение роста их родителей от среднего роста исследованных лиц. Это движение назад в направлении к среднему Ф. Гальтон назвал регрессией (to regress – двигаться в обратном направлении).

В 1885 г. была издана известная работа Ф. Гальтона «Регрессия в направлении к общему среднему размеру при наследовании роста», где он приходит к выводу, что, в общем, признаки родителей не полностью наследуются детьми, и чем отдаленнее предок, тем в меньшей мере сказывается его свойства на потомке. Закон регрессии веско свидетельствует против полного наследования какого-либо признака. «Из большого числа детей только немногие будут уклоняться от среднего уровня по сравнению с отклонением одного из родителей, отличающегося своими природными качествами. Чем ярче талант одного из родителей, тем реже родители имеют счастье видеть, что природа также щедро одарила их сына, и еще реже бывает, чтобы одаренность передавалась в последующие поколения. Закон беспристрастен и объективен. Он равномерно распределяет наследование хороших и плохих признаков. Он разрушает чрезмерные иллюзии одного одаренного родителя, лелеющего мечту, что его дети унаследуют все его способности. Закон устраняет также преувеличенные опасения относительно того, что детям передадутся все слабости, недостатки и болезни родителей. Разумеется, эти утверждения не находятся в противоречии с общей теорией, согласно которой дети талантливых родителей имеют большую вероятность обладать какими-либо дарованиями, чем дети родителей со средними способностями. Наши рассуждения выражают только тот факт, что самый одаренный из всех детей

немногих высокоодаренных родительских пар не так будет талантлив, как самый одаренный из всех детей очень многих родительских пар со средними способностями». Понятие регрессии, применяемое вначале только для процессов с тенденцией сдвигаться в направлении к среднему, с течением времени все более обобщалось и сегодня служит для характеристики односторонней стохастической зависимости.

Коэффициент корреляции описывает зависимость между двумя случайными величинами одним числом, а регрессия выражает эту зависимость в виде функционального соотношения и поэтому дает более полную информацию. Вначале регрессионный анализ применялся в биологии. Важнейшим научным журналом, в котором освещалась эта тема, был журнал «Биометрика», выходявший с октября 1901 года. В 1920-1930 гг. большое значение приобрело использование регрессионного анализа в экономике, и возникла новая область науки – эконометрика (автор термина Р. Фриш).

От изучения частных регрессионных задач исследователи постепенно перешли к регрессионному анализу структуры, присущей глобальным экономическим системам. Исследования в этой области производили Л. Клейн, которому в 1980 г. была присуждена Нобелевская премия по экономике, Д. Кейнс, Я. Тинберген и многие другие ученые.

### 1.5. Точечные оценки параметров

Разработанный Р. Фишером метод максимального правдоподобия, который восходит ещё к Даниилу Бернулли и К.Ф. Гауссу, предполагает максимум плотности вероятности вектора выборки:

$$L(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i) \rightarrow \max,$$

где  $L$  – функция правдоподобия,  $f$  – плотность вероятности заданного распределения.

Метод дал новый импульс теории оценивания в математической статистике. Принцип максимума плотности вероятности для многомерного нормального распределения привёл к тому, что МНК стал рассматриваться как частный случай метода максимального правдоподобия.

Данный метод применим в ряде управленческих, производственных, экономических и народнохозяйственных задачах оценки характеристик и параметров распределений вероятностей.

Рассмотрим типичный пример из повседневной жизни. Пусть на контроль поступила партия из  $N$  электроламп. Из этой партии случайным образом отобрана выборка объемом  $n$  электроламп. Возникает ряд естественных вопросов. Как по результатам испытаний элементов выборки определить средний срок службы электроламп, с какой точностью можно оценить эту характеристику? Как изменится точность, если взять выборку большего объема? При каком числе часов  $T$  можно гарантировать, что не менее 90 процентов электроламп прослужат  $T$  и более часов?

Предположим, что при испытании выборки объемом  $n$  электроламп дефектными оказались  $d$  электроламп. Какие границы можно указать для числа  $D$  дефектных электроламп в партии, для уровня дефектности  $D/N$  и т.п.?

Существует множество способов получения качественных оценок заданных параметров по имеющейся выборке. Наиболее распространенными из них является метод моментов и описанный ранее метод максимального правдоподобия. Стоит подчеркнуть тот факт, что последний из них асимптотически совпадает с эффективной оценкой для неизвестного параметра. Так, например, для случайной величины  $X$  оценка математического ожидания по выборке  $\{X_i\}_{i=1}^n$ , полученной по результатам  $n$ , экспериментов равна

$$\mathbb{E}[X] \approx \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$$

Отметим, что данная оценка обладает всеми необходимыми качествами для любого распределения.

### 1.6. Интервальное оценивание

Теория интервального оценивания была разработана преимущественно Р. Фишером и Е. Нейманом в 1925-1935 гг. Доверительный интервал Неймана содержит неизвестный параметр с заданной вероятностью. В отличие от выборки экспериментальных данных этот параметр не случаен. При другом подходе к интервальному оцениванию случайным считают, наоборот, неизвестный параметр. Такой вид интервальных оценок, называемых фидуциальными интервалами, ввел Фишер. Он начал заниматься интервальными оценками намного раньше чем Нейман, которого Фишер обвинил в присвоении своих идей.

В случае нормального распределения доверительные и фидуциальные интервалы формально совпадают; различается только их «философия». В течении некоторого времени считали, что эти два вида интервалов практически совпадают и что споры о различии между доверительными и фидуциальными интервалами являются чисто теоретическими. Однако вскоре обнаружили парадоксы, имеющие практическое значение. Разные подходы Неймана и Фишера привели к различным подходам в приложениях. В 1959 г. К. Стейн указал на чрезвычайно парадоксальный случай, когда доверительный и фидуциальный интервал могут сильно различаться.

Толкование доверительного интервала, основанное на интуиции, с уровнем доверия, скажем, 95% состоит в следующем: если провести очень большое количество независимых экспериментов с аналогичным построением доверительного интервала, то в 95% экспериментов доверительный интервал будет содержать оцениваемый параметр, а в оставшихся 5% экспериментов нет.

### 1.7. Проверка статистических гипотез

Трудно сказать что-то определенное о том, когда предпринимались первые попытки проверить статистические гипотезы. Джон Арбунтот, английский математик, врач и писатель, был первым, кто заметил в 1710 г., что гипотеза о равном соотношении родившихся мальчиков и девочек должна быть отвергнута, так как согласно демографическим данным за 82 года (доступным в то время) мальчиков каждый год рождалось больше чем девочек. Если бы вероятность рождения мальчиков была 0.5, то итог за 82 года был бы настолько маловероятен ( $1/2^{82}$ ), что его можно было бы считать практически невозможным.

Этот парадокс заинтересовал Лапласа. В 1784 г. он с удивлением обнаружил, что в нескольких различных районах доля родившихся мальчиков приблизительно равнялась  $22/43$ , а в Париже это отношение было равно  $25/49$ . Лаплас был заинтригован таким различием, но вскоре нашел для него разумное объяснение. В общее число родившихся в Париже включались также все подкидыши, а население из пригородов предпочитало подкидывать в основном девочек. Когда Лаплас исключил подкидышей из общего числа родившихся, доля новорожденных мальчиков и в Париже стала близкой к  $22/43$ .

В 1734 г. Парижская Академия наук присудила Даниилу Бернулли премию за исследование орбит планет. С помощью некоторого критерия проверки гипотез он пытался показать, что схожесть орбит планет является далеко не случайной. В 1812 г. Лаплас пытался применить статистические выводы для решения вопроса о том, какую из гипотез следует принять: считать кометы обычными элементами Солнечной системы или «незванными гостями». Он пришел к выводу, что кометы не являются обычными элементами Солнечной системы.

Основоположниками современной теории проверки статистических гипотез были К. Пирсон, Э. Пирсон, Р. Фишер и Е. Нейман. Для проверки гипотезы о виде закона распределения случайной величины К. Пирсон, Х. Крамер, Р. Мизес, А.Н. Колмогоров, Н.В. Смирнов и другие предложили несколько различных критериев. Вскоре возникла необходимость сравнить эффективность этих критериев.

При определении интервальных оценок и проверке статистических гипотез широко используются распределение Стьюдента. Оно введено в рассмотрение ирландцем Уильямом Госсетом в 1908 году. С 1899 г. он работал в Дублине на пивоваренном заводе Гиннеса, и его начальник настоял на том, чтобы Госсет писал под псевдонимом Стьюдент.

Э. Пирсон и Е. Нейман сделали первые шаги в работе по исследованию теоретических задач на нахождение лучших методов принятия решений. Они ввели понятие альтернативной гипотезы, которая в общем случае не является полным отрицанием проверяемой («нулевой») гипотезы и понятие ошибок первого и второго рода.

При проверке гипотез теория Неймана и Пирсона стала основной, однако, не лишённой парадоксов. В 1950 г. Герберт Роббинс показал, что существует критерий, который в некотором смысле является более мощным, чем критерий Неймана-Пирсона.

В классической теории математической статистики предполагается, что элементы выборки заранее известны. В основе одного из важнейших направлений современной статистики лежит понимание того, что не нужно фиксировать заранее объем выборки, его следует определять в зависимости от результатов более ранних наблюдений. Таким образом, объем выборки случаен. Эта идея последовательного выбора постепенно развивалась в работах многих ученых, но основателем теории последовательного анализа в математической статистике является американский математик Абрахам Вальд. Его последовательный критерий отношения правдоподобия, разработанный в 1943 г., стал важным открытием, позволившим в типичных ситуациях на 50% уменьшить среднее число наблюдений (при тех же вероятностях ошибок). Неудивительно, что в годы Второй мировой войны открытие Вальда было объявлено секретным. Его основная работа «Последовательный анализ» опубликована лишь в 1947 г. Спустя год он и Дж. Вольфовиц доказали, что методы, отличные от последовательного критерия отношения правдоподобия, не дают такого уменьшения числа элементов выборки.

### 1.8. Демография и смежные вопросы

Математические исследования смертности населения и продолжительности жизни, начавшиеся на раннем этапе развития капитализма, обуславливались потребностями страховых компаний. Галлей опубликовал в 1693 г. статью о таблицах смертности, которая положила начало математической теории страхования жизни.

Работа Д. Бернулли «О средней продолжительности браков при всяком возрасте супругов и о других смежных вопросах» является примером применения вероятностных идей к статистике народонаселения. Другая его работа – «Опыт нового анализа смертности, вызванной оспой, и преимущество предотвращающей ее инокуляции» – содержит исследования о влиянии прививок оспы на продолжительность жизни. Д. Бернулли поставил перед собой цель математически показать пользу прививок, вызвавших в народе недовольство. Классическая проблема статистики народонаселения – соотношение рождаемости мальчиков и девочек – была им исследована им в работе, вышедшей в Петербурге в 1771 г.

Понятие смертности нетрудно расширить. Если рассматривать амортизацию промышленных изделий или распад атомов как смерть, получим широко применимую математическую теорию, возникающую из исследований смертности населения. Понятие периода полураспада стало фундаментальным в некоторых областях науки. Радиоуглеродный метод, разработанный американским физикохимиком Уиллардом Фрэнком Либби, до сих пор является самым распространенным методом датирования в области археологической хронологии. В 1960 г. за это открытие ему была присуждена Нобелевская премия. Продолжительность существования радиоактивных частиц описывается показательным распределением.

В 1950 г., следуя идеям Либби, М. Свадеш применил его метод в лингвистике, предполагая, что не только радиоактивные атомы, но и атомы речи, т. е. слова, можно считать распадающимися. Предполагая, что период полураспада древнего базового словаря языков составляет 2000 лет, можно определить дату, когда два родственных языка разделились. Для этого достаточно знать, какая часть базового словаря до сих пор существует в обоих языках. Этот метод применяется часто и известен как лексикостатистика, или глоттохронология.

## 2. Случайные процессы

XX век не мог удовлетвориться тем идейным наследием, которое было получено от прошлого. Действительно, в то время, как физика, биолога, инженера интересовал процесс, т.е. изменение изучаемого явления во времени, теория вероятностей предлагала им в качестве математического аппарата лишь средства, изучавшие стационарные состояния.

Для исследования изменения во времени теория вероятностей конца XIX - начала XX века не имела ни разработанных частных схем, ни тем более общих приемов, а необходимость их создания буквально стучала в окна и двери математической науки. Изучение броуновского движения в физике подвело математику к порогу создания теории случайных процессов. В исследованиях датского ученого А. К. Эрланга (1878-1929) была открыта новая важная область, связанная с изучением загрузки телефонных сетей. Число абонентов изменяется во времени случайно, а длительность каждого разговора обладает большой индивидуальностью. И вот в этих-то

условиях двойной случайности следует производить расчет пропускной способности телефонных сетей, коммуникационной аппаратуры и управляющих связью систем. Несомненно, что работы Эрланга оказали значительное влияние не только на решение чисто телефонных задач, но и на формирование элементов теории случайных процессов, в частности процессов гибели и размножения.

Во втором десятилетии XX века начались исследования динамики биологических популяций. Итальянский математик Вито Вольтерра (1860-1940) разработал математическую теорию этого процесса на базе чисто детерминистских соображений. Позднее ряд биологов и математиков развивали его идеи уже на основе стохастических представлений. Первоначально и в этой теории применялись исключительно идеи процессов гибели и размножения. Собственно именно от задач биологии и пошло наименование этого очень частного типа случайных процессов.

Представим себе, что мы задались целью проследить за движением какой-нибудь молекулы газа или жидкости. Эта молекула в случайные моменты сталкивается с другими молекулами и меняет при этом направление движения и скорость. Состояние молекулы, таким образом, подвержено случайным изменениям и представляет собой нечто иное, как случайный процесс. Этот процесс определяется шестью параметрами – тремя координатами и тремя компонентами скорости. Многие физические явления для своего изучения требуют умения вычислять вероятность того, что определенная доля молекул успеет за заданный промежуток времени перейти из одной области пространства в другую. Происходит диффузия. Как быстро происходит процесс диффузии, по каким законам и когда образующаяся смесь становится практически однородной? На эти и многие другие вопросы дает ответы статистическая теория диффузии, базирующаяся на использовании теории случайных процессов. Очевидно, что подобные же задачи возникают в химии, когда приступают к изучению химических реакций. Какая часть молекул уже вступила в реакцию, какая особенность протекания реакции со временем, когда реакция практически уже закончилась?

Весьма важный круг явлений протекает по принципу радиоактивного распада. Суть его состоит в том, что атомы радиоактивного вещества распадаются, превращаясь в атомы другого элемента. Распад каждого происходит мгновенно, подобно взрыву, с выделением некоторого количества энергии. Многочисленные наблюдения показывают, что распад отдельных атомов происходит в случайно взятые моменты времени и расположение этих моментов, если количество распадающегося вещества не превосходит некоторого определенного критического предела, не зависит друг от друга. Для изучения процесса радиоактивного распада весьма важно определить вероятность того, что за определенный промежуток времени распадается то или иное число атомов. Формально, если задаться целью выяснения только математической стороны явления, оказывается, что аналогично происходят многие другие процессы: обрывы нитей в прядильной машине, число броуновских частиц, оказавшихся в данный момент в определенной области пространства, вызовы от абонентов, поступающие на телефонную станцию и т.д.

Теория броуновского движения, исходящая из теоретико-вероятностных предположений, была разработана в 1905 г. двумя известными физиками М. Смолуховским (1872-1917) и А. Эйнштейном (1879-1955). Позднее высказанные ими идеи использовались неоднократно как при изучении физических явлений, так и в различных инженерных задачах. В частности, именно с этих работ, как, впрочем, и с работ Эрланга, проявился широкий интерес к процессу Пуассона. Впрочем, сам Пуассон ввел в рас-



смотрение только распределение Пуассона, но он заслужил, чтобы его имя произносилось и при рассмотрении случайных процессов, связанных с его распределением. Это не единственный случай, когда в честь того или другого исследователя новым понятиям присваиваются их имена, хотя до этих понятий они и не доходили. Теперь широко распространены гауссовские случайные процессы, хотя сам Гаусс о них не имел никакого представления, да и само исходное распределение задолго до его рождения было получено Муавром и Лапласом. Попытка изучения средствами теории вероятностей явления диффузии была предпринята в 1914 г. двумя известными физиками – М. Планком (1858-1847) и А. Фоккером (1887-1972). Н. Винер в середине двадцатых годов при изучении броуновского движения ввел в рассмотрение процесс, получивший название винеровского процесса (процесса броуновского движения).

Следует упомянуть еще о двух важных группах исследований, начатых в разное время и по разным поводам. Во-первых, эта работы А. А. Маркова (1856-1922) по изучению цепных зависимостей. Во-вторых, работы Е. Е. Слуцкого (1880-1948) по теории случайных функций. Оба этих направления играли очень существенную роль в формировании общей теории случайных процессов. Для этой цели уже был накоплен значительный исходный материал и необходимость построения теории как бы носились в воздухе. Оставалось осуществить глубокий анализ имеющихся работ, высказанных в них идей и результатов и на его базе осуществить необходимый синтез.

В 1931 г. была опубликована большая статья А.Н. Колмогорова «Об аналитических методах в теории вероятностей», а через три года – работа А. Я. Хинчина «Теория корреляции стационарных стохастических процессов», которые следует считать началом построения общей теории случайных процессов. В первой из этих были заложены основы марковских процессов, а во второй – основы стационарных процессов. Они были источником огромного числа последующих исследований, среди которых следует отметить статью В. Феллера «К теории стохастических процессов» (1936), давшую интегро-дифференциальные уравнения для скачкообразных марковских процессов.

Обе только что упомянутые основополагающие работы содержат не только математические результаты, но и глубокий философский анализ причин, послуживших исходным пунктом для построения основ теории случайных процессов.

### 2.1. Ветвящиеся процессы

В первой половине XIX в. было замечено следующее интересное явление: некоторые знаменитые аристократические и простые фамилии постепенно исчезали. Эту проблему с математической точки зрения изучали Бьенэме в 1845 г. и де Кондолье в 1873 г. В 1874 г. Гальтон и Ватсон опубликовали важнейшую статью, посвященную этому вопросу. Ветвящиеся цепочки фамилий стали первым примером случайного ветвящегося процесса. Процессы такого типа наблюдаются в химии, физике и некоторых других областях. Например, процесс деления ядер или цепная реакция в ядерной физике хорошо моделируются случайными процессами. Поколения нейтронов сменяют друг друга чаще, чем поколения людей, однако в обоих случаях главный вопрос остается одним и тем же: при каких условиях процесс затухнет (фамилия исчезнет) или разовьется до бесконечности (бомба взорвется)? Понятие ветвящегося процесса введено в 1947 г. А.Н. Колмогоровым и Н.А. Дмитриевым.

## 2.2. Марковские процессы

Понятие марковской цепи принадлежит русскому математику А.А. Маркову, чьи первые статьи по этому вопросу были опубликованы в 1906-1908 гг. Он использовал новое понятие для статистического анализа распределения букв в романе А.С. Пушкина «Евгений Онегин». «Цепь Маркова» – важнейшее математическое понятие, возникшее при решении лингвистических проблем. Последовательности такого типа появляются во многих областях, например в классической физике, где будущее развитие системы полностью определяется ее настоящим состоянием и не зависит от того, как система оказалась в этом состоянии. В наши дни цепи Маркова и их обобщение на случай непрерывного времени и непрерывного фазового пространства (марковские процессы) играют в естественных и технических науках намного большую роль, чем в лингвистике, где они первоначально применялись.

## 2.3. Винеровские процессы

Во время экспериментов в 1827 г. с использованием микроскопа английский ботаник Роберт Броун не только обнаружил существование ядер у клеток, но и наблюдал интересное, хотя в то время необъяснимое явление: случайное движение взвешенных частиц, известное в настоящее время как броуновское движение. Поскольку он проводил свои эксперименты с цветочной пылью, предположили, что движение имеет биологические причины. Великой заслугой Броуна является экспериментальное доказательство исключительно физической природы этого явления.

На одной из лекций в Париже в 1904 г. Пуанкаре утверждал, что когда большие частицы размером примерно 0.1 мм со всех сторон много раз ударяются движущимися атомами, то эти частицы остаются на месте, так как по закону больших чисел случайные столкновения нейтрализуют друг друга. Однако в случае более мелких частиц воздействия толчков не достаточно для их общей нейтрализации, поэтому частицы двигаются зигзагообразно. Количественное объяснение явления было дано в 1905 г. Эйнштейном и польским ученым Смолуховским независимо друг от друга. По теореме Эйнштейна, средняя длина пути частиц пропорциональна корню квадратному из времени движения  $t$ . Значит, их средняя скорость пропорциональна  $1/\sqrt{t}$ . Отсюда вытекает, что мгновенная скорость частиц в любой момент времени должна равняться бесконечности, следовательно, при определении мгновенной скорости для броуновского движения возникают проблемы. Для их решения требовался более глубокий математический анализ, который провел Норберт Винер. В знак признания его заслуг в этой области математическая модель броуновского движения была названа винеровским процессом. Этот процесс можно интерпретировать как движение с непрерывной траекторией, нигде не дифференцируемой с вероятностью, равной единице. Это означает, что мгновенную скорость нельзя определить ни в одной точке.

Непрерывные, но нигде не дифференцируемые функции были известны задолго до Винера. В 1875 г. немецкий математик Пауль Дюбуа-Реймон впервые опубликовал пример непрерывной, но нигде не дифференцируемой функции, открытой Вейерштрассом в 1872 г. До этого подобная функция была придумана Больцано. Открытие таких функций многими выдающимися математиками было встречено без энтузиазма. В 1909 г. Пуанкаре писал: «Некогда при нахождении новых функций имелась в виду какая-нибудь практическая цель. Теперь функции изобретаются специально

для того, чтобы обнаружить недостаточность рассуждений наших отцов; никакого иного вывода, кроме этого, из них нельзя извлечь». В письме нидерландскому математику Томасу Стилтьесу Эрмит высказывался аналогичным образом: «С чувством непреодолимого отвращения я отшатываюсь от достойного всякого сожаления зла – непрерывных функций, не имеющих производных».

Винеровский процесс, очевидно, опровергал приведенные выше обвинения, так как никто не мог утверждать, что броуновское движение введено лишь для придумывания патологических контрпримеров.

### 3. Вклад А.Н. Колмогорова

Стоит отметить колоссальный вклад выдающегося советского математика Андрея Николаевича Колмогорова в развитие вероятностно-статистических наук в которых, благодаря своим достижениям, он был признан главой во всем мире.



Рис. 1. А.Н. Колмогоров

В теории вероятностей Колмогоров сделал исключительно много, получив важные результаты в различных областях этой обширной в наше время науки. Его имя обычно связывают с созданием аксиоматики теории вероятностей. Только после выхода в свет его монографии «Основные понятия теории вероятностей» (в 1933 г. – на немецком языке, в 1936 г. – на русском) стало возможным говорить о теории вероятностей как о математической науке в современном понимании, т.е. основанной на системе аксиом. Теперь теория вероятностей, занимающая видное положение как наука о реальных стохастических связях между явлениями, в рамках абстрактной математики воспринимается как частная глава теории меры или теории булевских алгебр. Даже Гильберт, формулируя в 1900 г. свои знаменитые проблемы, фактически относил теорию вероятности к разделу физики.

Когда в 1930-х годах начались преследования генетики, Колмогоров применил к обработке статистических данных, которые считали опровергающими законы Менделя, разработанный им и представленный в статье «Об эмпирическом определении закона распределения» (1933) аппарат математической статистики. Колмогоров пришел к выводу, что опубликованные статистические данные являются блестящим новым подтверждением законов Менделя. Публикация подобного материала в те годы

требовала от автора большой смелости. Сейчас в математической статистике этот метод общепризнан и называется критерием Колмогорова.

Теория случайных процессов стала центральным направлением развития теории вероятностей. При работе в данном направлении Колмогоров тесно контактировал с А.Я. Хинчиным, но Хинчин исследовал в основном стационарные случайные процессы, а Колмогоров занимался описанием явлений диффузии и броуновского движения. За рубежом решением этих вопросов занимался Норберт Винер, будущий «отец кибернетики». Винер отмечал, что в течении 20 лет он и Колмогоров занимались в одной и той же области и постоянно наступали друг другу на пятки.

В начале Великой Отечественной войны Колмогоров записался в дивизию народного ополчения, но на фронт его не взяли, а предложили работать в тылу. Андрей Николаевич стал заниматься исследованием в теории стрельбы. Вскоре им был предложен метод искусственного рассеивания для увеличения вероятности попадания торпед в корабли противника.

А.Н. Колмогоров и его ученики И.М. Гельфанд и А.М. Яглом заложили фундамент теории информации. Их работы определили высокий стандарт уровня математической строгости, которого с тех пор неизменно придерживаются математики и инженеры. С появлением в 1965 г. статьи Колмогорова «Три подхода к определению понятия количества информации» родилась алгоритмическая теория информации, в основе которой лежало понятие колмогоровской сложности конечного объекта. Продолжая исследования, Колмогоров приходит к выводу, что теория информации должна предшествовать теории вероятности и математической статистике, а не опираться на них. Строгое построение теории вероятности на базе теории информации – дело будущего.

Также Колмогоров занимался знаменитой «задачей трех тел», которая в общем виде не решена до сих пор. Работы Колмогорова и Арнольда по этим вопросам были отмечены Ленинской премией в 1965 г.

Много времени у Колмогорова занимала организационная работа. С 1954 по 1956 г. и с 1978 и до 1987 г. он был заведующим Отделением математики механико-математического факультета Московского университета, с 1954 по 1958 г. – деканом этого факультета, с 1938 по 1946 г. и с 1948 по 1960 г. – заведующим отделом теории вероятности Математического института им. В.А. Стеклова АН СССР, а с 1983 – заведующим отделом математической статистики и теории информации этого института.

В 1960 г. Колмогоров организовал при кафедре теории вероятностей Московского университета статистическую лабораторию. Одним из направлений исследования стало применение математических методов в языкознании. При статистической лаборатории Колмогоров организовал семинар. Его участниками стали математики, филологи, литературоведы. На семинаре изучали русскую поэзию. Решено было «поверить алгеброй гармонию». Не эмоции, а факты, не субъективная оценка, а объективный анализ поэтических произведений – такую задачу ставила перед собой группа Колмогорова. Исходная позиция состояла в том, что в произведениях поэзии имеются количественные закономерности, которые могут быть восприняты и в отрыве от содержания. Математический аппарат для изучения этих закономерностей включал в себя теорию вероятностей, математическую статистику, теорию информации.

Работы Колмогорова богаты глубокими идеями, оказавшими влияние на дальнейшее развитие науки. Разработанные им методы сыграли важную роль в решении

самых разнообразных прикладных задач. Исследования прикладного характера Колмогорова относились к математической геологии, теории стрельбы, гидро- и аэромеханике, генетике, физике, биологии, океанологии, метеорологии, кристаллографии.

В Московском университете Андрей Николаевич не только читал курсы лекций и вел семинары, но и был учредителем новых дисциплин учебного плана, которые сам наполнял содержанием. Много времени Колмогоров уделял преподаванию в физико-математической школе-интернате при МГУ, которая с 1989 г. носит его имя. Колмогоров является организатором и первым главным редактором журнала «Теория вероятностей и ее применения». На механико-математическом факультете Московского университета он создал и первым возглавил кафедру теории вероятностей в декабре 1935 г. и кафедру математической статистики в феврале 1976 г.

Колмогоров имел высокие понятия об этике ученого и всегда следовал им. Его отличали предельная научная честность и объективность, скромность, отзывчивость и щедрость.

Трудно перечислить все иностранные академии, членом которых являлся Колмогоров. Он был награжден семью орденами Ленина и другими орденами и медалями. В 1963 г. получил звание Героя Социалистического Труда, в 1965 г. стал лауреатом Ленинской премии.

В полдние 20 лет жизни Андрей Николаевич принимал активное участие в разработке вопросов преподавания математики в школе и в написании многих школьных учебников. Его выступления перед одаренными школьниками на олимпиадах и математических кружках способствовали раскрытию новых математических талантов.

## Заключение

Таким образом, история развития статистики показывает, что статистическая наука сложилась в результате теоретического обогащения накопленного человечеством передового опыта учетно-статистических работ, обусловленных прежде всего потребностями управления жизни общества, государственной арифметикой и обработкой экспериментальных данных. Сегодня статистика имеет огромное количество новых применений, но при всем этом, интересно отметить, что несмотря на доказанную практикой эффективность статистических методов, роль случайности в природе и границы статистической устойчивости остаются предметом дискуссий.

История математической статистики и случайных процессов еще далека от совершенства и требуется систематическая работа для того, чтобы восстановить пройденный путь и воздать должное ее создателям.

Мы видим, как человечество переходило от первичных догадок к более полному и совершенному знанию, как создание теории вероятностей и математической статистики позволяло переходить от строгих детерминистических представлений к более широким стохастическим концепциям, тем самым открывая новые возможности для глубоких заключений о природе вещей.

Теория случайных процессов продолжает бурно развиваться, в ней появляются новые направления исследований – оптимальное управление случайными процессами, теории мартингалов, теории просачивания, случайные операторы, вероятностные закономерности на алгебраических и топологических структурах. Эти направления представляют значительный общетеоретический и прикладной интерес.

## Список литературы

1. *Панов В.Ф.* А Современная математика и её творцы. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011, 648 с.
2. *Фишер Р. А.* Статистические методы для исследователей. М.: Изд-во Госстатиздат, 1958. 267 с.
3. *Орлов А.И.* Прикладная статистика. М.: Изд-во Экзамен, 2004. 483 с.
4. *Гнеденко Б.В.* Очерк по истории теории вероятностей. М.: Эдиториал УРСС, 2001. 88 с.
5. *Ливио, Марио* Был ли Бог математиком? М.: Изд-во АСТ, 2016. 384 с.