

Н.В. Трухачёва

**МАТЕМАТИЧЕСКАЯ
СТАТИСТИКА
В МЕДИКО-БИОЛОГИЧЕСКИХ
ИССЛЕДОВАНИЯХ
С ПРИМЕНЕНИЕМ ПАКЕТА**
Statistica



Москва
ИЗДАТЕЛЬСКАЯ ГРУППА
«ГЭОТАР-Медиа»
2012

ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ ИЗУЧЕНИЯ ОСНОВ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ В УСЛОВИЯХ МЕДИЦИНСКОГО ВУЗА С ПРИМЕНЕНИЕМ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

1.1. ИСТОРИЯ РАЗВИТИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Возникновение статистической науки относят к середине XVIII в., хотя практические операции по сбору данных о населении, его составе, имущественном положении и других сведений были известны задолго до возникновения статистической науки. Термин «статистика» произошел от латинского слова *status*, что означает «состояние, положение вещей». Первоначально статистика давала словесное описание «достопримечательностей» государства, и только с XIX в. статистические сведения стали сообщать в количественной форме. Сейчас слово «статистика» имеет несколько значений. Под статистикой понимают прежде всего отрасль практической деятельности, в задачи которой входят сбор, обработка и анализ статистических данных о различных явлениях общественной жизни. Под статистикой понимают и специальную научную дисциплину, занимающуюся разработкой теоретических положений и методов, используемых статистической практикой. Иногда применяя методы статистической обработки данных для разных отраслей знаний, различают статистику экономическую, медицинскую и т.д.

Возникновение статистической науки связано с именами английского экономиста У. Петти (1623–1687) и Д. Граунта (1629–1674), статистические и демографические идеи которых были развиты их последователями — немецким пастором И. Зюсмильхом (1707–1767) и виднейшим бельгийским ученым XIX в. А. Кетле (1796–1874). Работы Адольфа Кетле показали значение статистики в познании закономер-

ностей общественной жизни, определив, что эти закономерности четко проявляются лишь в массе явлений, т.е. при изучении данных по большому числу случаев. А. Кетле заложил и основы биометрии. Его учение о статистической закономерности получило развитие в работах немецкого статистика и экономиста В. Лексиса (1837–1914). Лексис был сторонником применения математических методов в экономике. Это направление развития статистической науки нашло отражение в работах Ф. Гальтона (1822–1911), К. Пирсона (1857–1936), Р. Фишера (1890–1962), В. Госсета (1876–1937) и других западных ученых. Их именами назван целый ряд показателей и критериев. Ф. Гальтоном в 1886 г. введен термин «регрессия». Он обнаружил, что в среднем сыновья высоких отцов имеют не такой большой рост, а сыновья отцов с небольшим ростом выше своих отцов. Это было интерпретировано им как «регрессия к посредственности». К. Пирсон усовершенствовал предложенные Ф. Гальтоном методы корреляции и регрессии. К. Пирсон ввел в биометрию такие понятия, как среднее квадратичное отклонение и вариация. Ему принадлежит разработка критерия согласия, он ввел термин «нормальное распределение», который сейчас стал общепринятым.

В российской статистической школе развитие предмета и метода статистической науки шло по трудному и порой противоречивому пути, и до сих пор эти проблемы остаются дискуссионными. Известны работы К.Ф. Германа (1767–1838), в которых обосновывалась необходимость выделения особой науки — теории статистики, Д.П. Журавского (1810–1856), в которых он называет статистику наукой «категорического вычисления». Широкую известность приобрели во второй половине XIX в. труды профессора Петербургского университета Ю.Э. Янсона (1835–1893) — видного теоретика в области статистики и руководителя практических статистических работ. В 1881 и 1890 г. он организовал переписи населения Петербурга.

В конце XIX и начале XX в. практические потребности пореформенной России стали мощным толчком развития статистической науки. Эти годы ознаменовались крупными успехами Петербургской школы. П.Л. Чебышев (1821–1894), а также его ученики А.М. Ляпунов (1875–1918) и А.А. Марков (1856–1922) создали русскую школу теории вероятностей, достижения которой стали использоваться первоначально в демографии, а затем и в различных областях социально-экономической статистики.

1.2. ПРИМЕНЕНИЕ МЕТОДОВ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ В БИОЛОГИИ И МЕДИЦИНЕ

Идея контролируемых клинических испытаний возникла уже в XII в., когда римский император Фредерик II изучал влияние физических упражнений на пищеварение. Двум рыцарям он дал одинаковую пищу. После этого одного из них отправил спать, а другого на охоту. Спустя несколько часов, умертвив обоих, он изучил содержимое их пищеварительного тракта. Пищеварение в желудке у спящего оказалось более интенсивным. В XVII в. Ж.Б. Ван Хельмонт решил оспорить практику кровопускания и предложил первое рандомизированное клиническое испытание с большим числом участников и статистическим анализом. Предполагалось около 500 людей случайным жребием разделить на две группы. В одной группе не применяли бы кровопускание, а в другой врачи делали бы его столько раз, сколько нужно. Эффективность кровопусканий предлагалось оценивать по числу похорон в каждой группе. По неизвестным для нас причинам эксперимент не был проведен, а неэффективность кровопусканий была доказана позже П. Луи [36].

Однако систематическое применение статистических данных для медицинских исследований началось только в XIX в. Наука о применении математики в биологии непосредственно переплетается с развитием генетики. Именно генетика, и в первую очередь законы Менделя, стали основной областью применения статистических методов исследования в биологии. О своих опытах по гибридизации гороха Грегор Мендель сообщил на заседании Общества естествоиспытателей в 1865 г. Однако сообщение было без интереса встречено слушателями.

В России труд Г. Менделя получил научный отклик, но несмотря на это именно в России статистика развивалась драматично. В 1928 г. Т.Д. Лысенко изучал влияние температурного режима на продолжительность вегетативного периода у растений. Он вывел формулу, по которой можно было определить количество дней, необходимых для предварительной обработки семян: $N = A_1 / (B_1 - t_0)$, где B_1 — максимальная температура, которая может существовать «без предварительной обработки»; A_1 — количество дней, необходимых для завершения фазы развития растения, t_0 — средняя дневная температура. Труд Т.Д. Лысенко раскритиковали за попытку свести сложные отношения между растением и средой только к статистической закономерности и за то, что работа основана на слишком малом количестве экспериментов и нуждается в дальнейшей проверке. После этой критики Т.Д. Лысенко с недоверием