

Двенадцатая хрестоматия

по истории теории вероятностей и статистики

Составитель и переводчик О. Б. Шейнин

На правах рукописи

Берлин
2014

Оглавление

От составителя

- I.** Оскар Шейнин, Иоганн Кеплер
- II.** Филлис Дийн, Уильям Петти
- III.** Хилари Л. Сил, Абрахам де Муавр
- IV.** Уильям Краскл, Ричард Прайс
- V.** Черчилль Эйзенхарт, Карл Фридрих Гаусс
- VI.** Дейвид Ландау, Пол Ф. Лазарсфельд, Адольф Кетле
- VII.** [Чарлз Биббидж], Заметка о происхождении Статистического общества
- VIII.** Клаус-Петер Хейс, Вильгельм Лексис
- IX.** Элен Уокер, Карл Пирсон
- X.** М. С. Бартлетт, Р. А. Фишер
- XI.** Томассо Сальвенини, Коррадо Джини
- XII.** Деннис В. Линдли, Джон Мейнард Кейнс. Работы по статистике
- XIII.** Мервин Муллер, Случайные числа
- XIV.** Натан Кейфиц, Государственная статистика
- XV.** Бернар Лекуе, Энтони Р. Обершаль, Ранняя история социологических исследований
- XVI.** Эдгар Боргатти, Социометрия
- XVII.** Уильям Краскл, Статистика, её сфера

От составителя

Мы перевели некоторые статьи из *International Encyclopedia of the Social Sciences* (сокращённо IESS). New York, 1968. Редактор David L. Sills. Впрочем, её статистические статьи были перепечатаны в энциклопедии *International Encyclopedia of Statistics* (сокращённо IES), vols 1 – 2. New York – London.

Редакторы William H. Kruskal, Judith M. Tanur.

Некоторые достаточно длинные статьи либо не были разбиты на параграфы, либо эти параграфы не были пронумерованы. Эту работу мы выполнили сами. Качество статей оказалось подчас совсем неважным; в нужных случаях мы указывали на допущенные ошибки в своих примечаниях, а немногие места выпустили из перевода. В то же время полезны были и приводимые факты, и почти всегда обильная библиография. Источники, на которые не были даны ссылки в текстах, мы как правило, исключали.

Общие замечания к некоторым статьям

[iii] Муавр был одним из основных предшественников Лапласа (другими были Бейес и Даниил Бернулли), но его предельная теорема, частный случай центральной предельной теоремы, длительное время оставалась неизвестной (Шейнин 2013, § 5.4). Он, как и Якоб Бернулли, не заметил, что при одном и том же числе испытаний определение неизвестной вероятности события по его статистической вероятности, т. е. обратный закон больших чисел, менее точно, чем определение его статистической вероятности по теоретической. Установил это Бейес (там же, § 6.2).

[vi] Изложение во многом ошибочно и/или поверхностно. О Кетле см. Sheynin (1986).

[viii] В статье нет даже намёка ни на роль Лексиса в возникновении континентального направления статистики, ни на разгром его первоначального рассуждения (1879а), который в основном осуществил Чупров (Шейнин 2013, §§ 16.1, 16.1.3). Само это рассуждение описано только на ином языке последующих авторов.

[ix] Автор сообщает интересные сведения о Пирсоне, однако к ним мы могли бы прибавить многое другое из нашей статьи 2010 г. и книги (2013, § 16.2). Кроме того, в § 1.3 автор допустил ошибки и неточности, которые также можно обнаружить по нашим работам и которые мы указали в примечаниях. Утверждение о том, что никто до Пирсона не помышлял о проверке близости эмпирических и теоретических данных, явно ошибочно. Уже Якоб Бернулли мог бы проверять, относится ли исследуемый ряд к типу *бернуллиевых* или нет. Наконец, в последней фразе статьи было бы уместно упомянуть (да и описать) пирсоновскую *Грамматику науки*.

[xi] E. Regazzini, см. Johnson & Kotz (1997), сообщает, что Джини опубликовал две статьи в *J. Roy. Stat. Soc.*, в томе 89 за 1926 г. и томе A30 за 1966 г. и кроме того указывает на посмертное издание сочинений Джини (1968). G. M. Giorgi, см. Heyde & Seneta (2001), характеризует существенную деятельность Джини в международном масштабе. Мы сами заметим, что в 1931 г. в журнале *Bull. Inst. Stat. Intern.*, была помещена статья Борткевича о распределении доходов и последовавшая дискуссия между Джини и им. Оказалось, что Борткевич не упомянул соответствующей статьи Джини (1912b), которой он не видел, но о которой знал из писем А. А. Чупрова.

[xii] О статистических работах Кейнса см. также Bortkiewicz (1923).

[xv] Описание материала поверхностно (некоторую часть очерка мы и переводить не стали), что лишь частично оправдывается широтой темы. Упущены, например, социальное законодательство Бисмарка и борьба с инфекционными заболеваниями в Германии, есть и грубые ошибки, см., например, Примечания 3 и 33, а земская статистика России как бы не существовала, якобы не было переписей ни в Германии, ни в

США, и ни в Италии, ни в Индии будто бы ничего интересного вообще не происходило. Вместе с тем, безусловно интересны сведения о некоторых исследованиях и богатая библиография.

[xvi] Малознакомая социометрия изучает межличностные отношения в группах и связи в них. Автор не указал, что подобные исследования необходимы для обеспечения действий небольших групп, изолированных от общества (например, зимовщикам на полярных станциях и, недавно, космонавтам).

I

Оскар Шейнин

Иоганн Кеплер

O. B. Sheynin, Kepler, Johannes. IES, pp. 487 – 488

Иоганн Кеплер (1571 – 1630) был великим астрономом, но, конечно, не статистиком в современном понимании. Тем не менее, статистические идеи появляются в его астрономии при обработке наблюдений. Он (1609/1929, с. 114) заметил, что погрешности неизбежны, пытался свести к минимуму влияние случайных и систематических ошибок (с. 209 и гл. 51) и заявил, что для исключения систематических ошибок требуется симметричная программа наблюдений. Подобные высказывания относятся к предыстории планирования эксперимента.

Кеплер неоднократно уравнивал прямые наблюдения, обычно применяя среднее арифметическое. В одном случае он (1609/1992, с. 200) применил среднее арифметическое с апостериорными весами как *среднее по добру и справедливости* (*medium ex aequo et bono*). Это латинское выражение восходит к Цицерону, стало широко известно, и молчаливо подразумевало: *а не в соответствии с буквой закона*. Буквой закона Кеплер, видимо, посчитал обычное среднее арифметическое, астрономическую историю которого поэтому можно начинать со второй половины XVI в.

Для уравнивания косвенных наблюдений он применял итерации, начиная с некоторой части наблюдений, и на каждом шагу сверял результаты с остальными наблюдениями.

При вычислении орбит планет Кеплер намеренно искажал наблюдения малыми произвольными поправками, видимо выбирая их с учётом свойств *обычных* случайных ошибок (там же, с. 197), т. е. ввёл элементы метода Монте Карло. Он также применял метод минимакса. Птолемееву систему мира Кеплер отверг, потому что (даже помимо физических соображений) примерно вычисленная минимаксная ошибка подгонки наблюдений превысила возможную погрешность наблюдений Тихо Браге (там же, с. 166).

Кеплеру (1609; 1618/1952, с. 932) пришлось признать влияние случайных причин и в связи с эксцентриситетами планетных орбит, и его рассуждение повторил Кант (1775/1910, с. 337) и, как

ни странно, Лаплас (1796/1982, с. 328). Странно, потому что Ньютон доказал, что эти эксцентриситеты определялись скоростями обращения планет вокруг Солнца.

Далее, приняв вначале ошибочную форму третьего кинематического закона обращения планет, Кеплер (1596/1936, с. 114) посчитал невозможным одновременное возвращение всех планет в их исходное на момент создания Вселенной положение, а потому, в соответствии с древними верованиями, и невозможность конца света. Во втором издании, однако, он (1621/1936, с. 145) исправил свою ошибку и допустил крохотную вероятность указанного события. Если два числа взяты *случайно*, рассудил Кеплер, они по всей вероятности окажутся несоизмеримыми.

Этот же довод встретился у Орема (Oresme 1351/1966, с. 247, 422) и был первым стохастическим рассуждением о математическом понятии, притом относящимся к естествознанию. Впрочем, непонятно, каким образом космические расстояния могут быть рациональными или иррациональными, они всегда приближённые.

Кеплер (1610/1941, с. 248) считал себя основателем научной астрологии, науки об общем влиянии небесных тел на земную жизнь. Он (1619/1939, с. 256, 263) заявил, что существует как бы корреляционная связь через *промежуточные причины* между примечательным расположением планет (аспектами) и метеорологическими явлениями. Метеорология не была ещё количественной наукой, и проверить своё утверждение Кеплер не сумел. Той же *научной* астрологии придерживались Тихо Браге и ведущие учёные древности (например, Бируни). Возникает, правда, существенный вопрос: какие расположения планет считать аспектами? А среди промежуточных причин не было и не могло быть количества населения.

Заметим, что Кеплер (1610/1941, с. 238) отказался отвечать спросившему его, жив или нет безвестно отсутствовавший, но вот Якоб Бернулли, а ещё раньше Николай Бернулли предложили решать подобные вопросы в вероятностном смысле.

Библиография

И. Кеплер

1596, 1621, латин., *Das Weltgeheimnis*. München – Berlin, 1936.

1609, латин., *Neue Astronomie*. München, 1929. *New Astronomy*. Cambridge, 1992.

1610, Tertius interveniens. Das ist Warnung an etliche Theologos, Medicos und Philosophos. *Ges. Werke*, Bd. 4. München, 1941, pp. 145 – 258.

1618, латин., *Epitome of Copernican Astronomy*, NNo. 4 – 5. Chicago. (*Great Books of the Western World*, vol. 16, 1952.)

1619, латин., *Weltharmonik*. München – Berlin, 1939. *Harmony of the World*. Philadelphia, 1997.

Другие авторы

Kant I., Кант И. (1755, нем.), *Allgemeine Naturgeschichte und Theorie des Himmels*. *Ges. Schriften*, Bd. 1. Berlin, 1910, pp. 215 – 358. *Всеобщая естественная история и теория неба*. Соч., т. 1. М., 1963, с. 393 – 508.

Laplace P. S., Лаплас П. С. (1796, франц.), *Изложение системы мира*. Л., 1982.

Oresme N. (1351, латин.), *De proportionibus proportionum* и *Ad pauca respicientes*. Madison, 1966. Латино-англ. текст. Перевод с Введением и комментариями E. Grant.

Sheynin O. (1973), Mathematical treatment of astronomical observations. *Arch. Hist. Ex. Sci.*, vol. 11, pp. 97 – 126.

--- (1974), On the prehistory of the theory of probability. Там же, vol. 12, pp. 97 – 141.

--- (1977), Kepler as a statistician. *Bull. Intern. Stat. Inst.*, vol. 46, pp. 341 – 354.

II

Филлис Дийн

Уильям Петти

Phyllis Deane, Petty, William. IES, pp. 702 – 704.

Уильям Петти (1623 – 1687), английский экономист, был первым, кто систематически истолковывал *искусство политической арифметики*, кто был обязан всем самому себе и обладал талантами во многих областях знания и громадной умственной энергией. Он был сыном clothier (портного или торговца мануфактурой), которого биографы обычно называют *бедным*, но представляется правдоподобным, что отец оставил ему хороший дом и 8 акров земли [около 3,2 га] в родном Romsey, так что семья находилась намного выше черты бедности.

Ставши юнгой в возрасте около 14 лет, Петти достаточно знал латинский и греческий, чтобы попасть в иезуитский колледж в Кане¹, в том месте, в котором его списали на берег из-за сломанной ноги. В колледже он изучал латинский, греческий, французский языки и математику. После краткой службы на королевском флоте, он отправился изучать медицину в университетах Утрехта, Амстердама, Парижа и Оксфорда. В 1651 г. он стал доктором физики Оксфордского университета и вступил в лондонский колледж врачей. В 28 лет Петти оказался заместителем директора Брейзноус-Колледжа и профессором анатомии в Оксфорде.

Будучи одним из ведущих интеллектуалов своего времени и членом-основателем Королевского общества, Петти всё же был скорее светским человеком, а не учёным. Он писал больше, чем читал, был упорным изобретателем. До получения медицинской степени он запатентовал устройство для двойных записей² и построил четыре прототипа корабля с двойным дном. После Великого лондонского пожара 1666 г. он составил схему перестройки города.

В 1652 г. Петти отправился в Ирландию в качестве главного врача армии. В этой стране, опустошённой чумой и другими смертельными эпидемическими заболеваниями, эта должность не была синекурой. Не более, чем через два года после своего прибытия в страну, он перенял сложную задачу кадастровой съёмки конфискованных имений ирландских мятежников для их последующей передачи английским землевладельцам. И эту задачу он выполнил, как обычно, умело и энергично, а также решительно и драчливо, что восстановило многих против него и привело к бесконечным судебным процессам.

Большую часть активного периода своей жизни он провёл, защищая свою ирландскую съёмку и заботясь о своих ирландских имениях, но сумел прочесть некоторое число сообщений Королевскому обществу по темам от окраски тканей до исследования минеральных вод и написать огромный объём

серьёзной литературы по экономике, демографии, военно-морской, медицинской и научной темам³.

О Петти было написано много, и опубликованные мнения его современников не оставляют сомнения в том, что его уважали и почитали. Вряд ли, вопреки сделанным заявлениям, он был автором *Наблюдений ...*, подписанных его другом Джоном Граунтом⁴, но он вполне мог бы их написать. А его собственные *Наблюдения* (1683) явно связывают его с этим предшествующим трудом, проложившим путь демографическим исследованиям. Его первый биограф, Aubrey (1898)⁵, авторы дневников Джон Эвелин и Сэмюэль Пепис и экономист и государственный деятель Давенант, будучи сами известными людьми, считали его выдающейся личностью.

Притязания на славу Петти за достижения в экономике основаны не столько на его оригинальности, сколько на теоретических способностях и аналитическом умении. Он настаивал на измерениях, и его ясный обобщённый взгляд на экономику означал, что он был первым эконометристом. Он всё время разрабатывал и применял понятия и аналитические методы, опережающие своё время. Его оценка выгоды внешней торговли (Петти 1682) была основана на утверждении о преимуществах разделения труда и специализации, которое он сформулировал на столетие раньше знаменитого мнения Адама Смита.

Труду Петти придавал такое большое значение в создании богатства, что, например, Маркс считал его ранним выразителем трудовой теории стоимости. Но его характерное и часто цитируемое выражение (1662), *Труд – это отец и активный принцип богатства, а Земля – мать*, означает, что его теория производства и стоимости была основана на двух первоначальных факторах производства ранних экономистов. Он (1665; 1676; 1687) был автором первой известной оценки национального дохода, хотя, правда, не удосужился определять или развивать свои идеи и грубо, иногда даже небрежно обращался с числами. Некоторые его вычисления, например, в (1662), являются по существу упражнениями в анализе *cost – benefit* (затраты – прибыль). Насколько известно, Петти (1665) первым понял идею о скорости обращения денег, хотя в дальнейшем (1695) не сказал об этом ни слова.

Петти не гнушался передергиванием своих данных для обоснования спорных доводов, и было бы опрометчиво слепо воспринимать его статистику. Но он никогда не был рабом политических предубеждений, а его анализ экономики Англии и Ирландии (1671 – 1676) и многочисленные эскизы о *политической арифметике* умны и проницательны.

Примечания

1. Кан, населённый пункт в Нормандии (северо-запад Франции).
2. Это устройство иначе называется полиграфом. Впрочем, гораздо известнее другое значение этого термина: детектор лжи, одновременно записывающий показания различных физиологических функций организма.
3. Отделять медицину от наук не следовало бы.

4. Этому вопросу было посвящено много литературы. Сам Петти (1674, Посвящение Лорду Брукнеру) назвал Граунта автором *Наблюдений*. Вообще о Петти см. также Шейнин (2013, §§ 3.1.1 и 3.1.4).

5. В Библиографию, составленную автором, включены и более ранние биографы.

Библиография

W. Petty, В. Петти

1899, *Economic Writings*, vols 1 – 2. Editor, C. H. Hull. London, 1997.

Автор ссылается (иногда косвенно) на следующие сочинения из указанного источника:

Treatise of taxes and contributions, 1662; vol. 1, pp. 1 – 97.

Verbum sapienti, написано 1664, опубликовано 1691; vol. 1, pp. 99 – 120.

The political anatomy of Ireland, написано 1671 – 1676, опубликовано 1691; vol. 1, pp. 121 – 231.

Political Arithmetic, написано в 1676, пиратское издание 1683, опубликовано в 1690 г.; vol. 1, pp. 233 – 313.

Another essay in political arithmetic concerning the growth of the City of London, 1682; vol. 2, pp. 451 – 478.

Observations upon the Dublin bill of mortality, 1681; vol. 2, pp. 479 – 491.

Sir William Petty's *Quantulumcunque* concerning money (1682), 1695; vol. 2, pp. 437 – 448. В другом издании той же книги: New York, 1963.

A treatise of Ireland, 1687; vol. 2, pp. 545 – 621.

Мы добавили:

Appendix [Extract from Discourse made before the Royal Society concerning the use of the duplicate proportion, 1674]; vol. 2, pp. 622 – 632.

Другие сочинения

1851, *History of the Cromwellian Survey of Ireland*. Dublin.

1927, *Petty Papers*, vols 1 – 2. London, 1997.

1928, *Petty – Southwell Correspondence 1676 – 1687*. London.

1931, *The Double-Bottom or Twin-Hulled Ship*. Oxford.

В книге

Петти В. (1940), *Экономические и статистические работы*. М.

из указанных выше сочинений содержатся:

Трактат о налогах и сборах, с. 3 – 78.

Verbum sapienti, с. 79 – 89.

Другой опыт по политической арифметике, рассматривающий рост г. Лондона, с. 218 – 234.

Другие авторы

Шейнин О. Б. (2013), *Теория вероятностей. Исторический очерк*. Берлин. Также Google, Oscar Sheynin, Download area.

Aubrey J. (1898), Sir William Petty. В книге автора *Brief Lives*. Ann Arbor, 1957, pp. 237 – 241.

Bevan W. L. (1894), Sir William Petty. *Publ. Amer. Econ. Assoc.*, vol. 9, pp. 370 – 472.

Fitzmaurice E. G. P. (1895), *Life of Sir William Petty*. London.

Greenwood M. (1928), Graunt and Petty. *J. Roy. Stat. Soc.*, vol. 91, pp. 79 – 85.

Pasquier M. (1903), *Sir William Petty*. Paris.

Strauss E. (1954), *Sir William Petty*. London – Glencoe, Ill.

III

Хилари Л. Сил

Абрахам де Муавр

Hilary L. Seal, Moivre, Abraham de. IES, pp. 601 – 604

Родители Абрахама де Муавра (1667 – 1754) были французскими протестантами по фамилии Муавр. Его также звали Демуавр, сам же он написал свою фамилию в обратном адресе письма Иоганну Бернулли как деМуавр. Математику и физику он изучал в Париже под руководством Озанама, а в возрасте 21 года эмигрировал в Англию, чтобы избежать религиозных преследований (Walker 1934).

Муавр был математическим гением и обладал исключительной аналитической мощью. Он переписывался со многими ведущими математиками своего времени и общался с ними в Королевском обществе, но так и не сумел получить университетскую должность. Ему пришлось обучать сыновей знати и давать советы игрокам и спекулянтам пожизненных рент, – распространённой формой вложения капитала в первой половине XVIII в. (Walford 1871)

Злополучие Муавра обернулось выигрышем для потомства, потому что проблемы, с которыми он сталкивался в своих консультациях и их удачное решение послужили материалом для двух его знаменитых книг. В свои последние годы Муавр видимо должен был существенно полагаться на продажи позднейших изданий своей книги о вычислении пожизненных рент.

Практически полезная книга Муавра по теории вероятностей впервые появилась в 1718 г. как перевод и переработка латинской статьи 1711 г. Он посвятил её Ньютону и благодарил его за его сочинения и беседы с собой¹. В своём окончательном виде в издании 1756 г. эта книга примечательна за оригинальное исследование следующих тем, каждая из которых находится в центре современной теории вероятностей:

1. Общие законы сложения (David & Barton 1962, гл. 2) и умножения вероятностей (Montucla 1758/1802, часть 5-я, кн. 1, гл. 39).
2. Биномиальное распределение (Cantor 1898, гл. 96)².
3. Производящие функции (Seal 1949a).
4. Разностные уравнения, включающие вероятности и их решение при помощи возвратных последовательностей (Czuber 1900).
5. Новые и общие решения проблем о продолжительности игры или разорения игрока (Todhunter 1865, гл. 9).
6. Предельная форма члена бинома

$$C_n^x p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n, \quad 0 < p < 1 \quad (1)$$

при $n \rightarrow \infty$ и постоянном np (только при $x = 0$, David 1962) и при $n \rightarrow \infty$ и $np \rightarrow \infty$.

Результат во втором случае³

$$\frac{1}{\sqrt{2\pi np(1-p)}} \exp\left[-\frac{(x-np)^2}{2np(1-p)}\right],$$

а именно ордината нормального распределения, был получен в явном виде (в других обозначениях).

В книге содержалась тригонометрическая теорема, названная его именем, и аппроксимация логарифма факториала, улучшенная открытием Стирлинга, который в том же году обнаружил, что константа найденного закона была равна $\sqrt{2\pi}$. Некоторые математические выкладки этой книги были опубликованы для более широкого круга читателей (1730)⁴.

Вторая книга Муавра заложила основы математики случайностей в жизни (Saar 1923). Её первое издание распродавалось медленно, но плагиаторский текст Симпсона (1742) побудил Муавра опубликовать в 1743 г. новое, полностью пересмотренное издание (Young 1908). Его успех был очевиден ввиду появления двух других лишь слегка изменённых изданий на протяжении девяти лет.

В 1756 г. появилось его окончательное, тщательно переработанное издание, добавленное в конце его *Учения о случае*. В приложении Муавр сослался на мемуар 1755 г. Джеймса Додсона, отца научного страхования жизни (Ogborn 1962)⁵, который возможно и был тем *другом*, отредактировавшим это посмертное издание. Оригинальность этого учебника о случайностях в жизни очевидна ввиду включения в него следующих собственных результатов:

1. Рекуррентные формулы для вычисления стоимости пожизненных рент для возраста x при её заданной стоимости для возраста $x + 1$. Но Муавр вряд ли имел в виду её вычисление для всех возрастов начиная с самого преклонного (Young 1908).

2. Общие соотношения для вычисления стоимости пожизненных рент в терминах рент на одно или несколько лиц.

3. Применение исчисления для подсчёта определённого вида пожизненных рент.

4. Закон смертности, а именно равномерное убывание числа остающихся в живых. Он существенно согласовывался с таблицами для Бреслау, опубликованными в 1693 г. его другом Галлеем. Как следствие,

5. Легко вычисляемые значения стоимости ренты на одного или на нескольких человек, на ограниченный срок или пожизненно.

6. Выражения для вычисления вероятностей одному человеку пережить другого.

7. Стоимость пожизненной ренты с пропорциональной выплатой в год смерти.

Самые ранние опубликованные мемуары Муавра 1690-х годов были написаны под влиянием ньютонова метода флюксий и теории рядов (Cantor 1898, гл. 86), а его интерес к теории вероятностей возник только после выхода первого издания книги Монмора (Montmort 1708). Возможно важнейшая работа Муавра была впервые отпечатана в 1733 г. в качестве дополнения к *Misc. Anal.* (1730)⁶. В ней, используя полученный им результат (1), он сузил широкие пределы, полученные Якобом Бернулли и, применив приближённую квадратуру нормальных ординат, вычислил сумму биномиальных коэффициентов от $x = np - l$ до $x = np + l$ при $p = 1/2$ и $l = k\sqrt{n/4}$, $k = 1, 2, 3$.

Это оказалось первым табулированием нормальных площадей с отходом от среднего на 1, 2 и 3 стандартных отклонений. Нет, однако, никаких свидетельств того, что Муавр представлял себе непрерывное распределение вероятностей (Seal 1954, 1957)⁷, и всё же его мемуар явно послужил основой для последующего доказательства Лапласа (1812/1820, с. 275 – 284; 1886, гл. 3) того, что при большом n биномиальное распределение стремится к нормальному (Pearson 1924).

Муавр был чистым математиком, мало заинтересованным в практическом приложении своей теории. Хотя он написал о случайностях в жизни, но лишь в заключительном приложении к посмертному изданию своей книги (1718/1756) кратко сослался на данные о смертности, более поздние, чем указанные в таблицах г. Бреслау⁸.

В начале и середине XVIII в. было опубликовано несколько сборников статистики смертности, которые могли бы послужить плодородным полем для приложения исправленного Муавром варианта теоремы Бернулли (Seal 1949b). Эти данные привели к распространённому убеждению в божественной закономерности демографических отношений, и в нескольких параграфах *Учения о случае*, в изданиях 1738 и 1756 гг., имеются ссылки на связь этих убеждений и теоремы Бернулли. К сожалению, ни Муавр, ни его современники не исследовали этой темы (Westergaard 1932, гл. 7 и 10), и нельзя считать, что он повлиял на демографов XVIII и раннего XIX вв. (Pearson 1926).

Добавление. Квадратура нормальной кривой у Муавра распространяется от 0 до l и аппроксимирует сумму биномиальных членов от $x = np$ до $x = np + l\sqrt{npq}$. Он удвоил полученный результат и вывел сумму членов в пределах $x = np \pm l\sqrt{npq}$ и Czuber (1898) ошибочно посчитал, что Муавр дважды включил средний биномиальный член. Schneider (1968) опубликовал лучший обзор жизни и трудов Муавра. Интересный поиск оставшихся экземпляров (второго) приложения к *Misc. Anal.* [т. е. мемуар 1733 г.] описали Daw & E. S. Pearson (1972).

Примечания

1. В посвящении Ньютону (перепечатанном в третьем издании книги, с. 329) Муавр косвенно сформулировал цель теории вероятностей, отделение *шанса* от предначертания, чтобы тем самым облегчить изучение божественного замысла в сотворении природы. Вопреки автору, это указывает на

непосредственный интерес Муавра в богословии (пусть в его приложении). Далее, исследование биномиального распределения (косвенно, в связи с мужскими и женскими рожденьями) у него означало, что это распределение было преднамеренным, случайными же были лишь эмпирические отклонения от него. Наконец, действительно ли Муавр не думал о приложении математики (фактически, теории вероятностей), как заявила ещё Уокер (1934, с. 320)? Мы (1970, с. 209) решительно отрицаем это мнение.

2. Биномиальное распределение конечно же было у Бернулли.

3. Общий случай $a \neq b$ не был рассмотрен явно, но фактически Муавр имел его в виду (см. хотя бы название его мемуара 1733 г.) и указал, что обобщение на этот случай не представляет труда.

4. Автор явно принизила значимость этой книги.

5. Мату, см. Bellhouse (2007), сообщил, что Додсон был учеником Муавра.

6. Мемуар 1733 г. действительно принято называть вторым дополнением к его книге 1730 г., однако только потому, что его сохранившиеся экземпляры в нескольких библиотеках были приплетены к экземплярам книги.

7. Муавр мог представлять себе непрерывное распределение, хоть об этом и нет свидетельств. Кроме того, он же ввёл непрерывный закон смертности.

8. См. Примечание 1.

Библиография

A. De Moivre

1711, De mensura sortis, or the measurement of chance. *Intern. Stat. Rev.*, vol. 52, 1984, pp. 236 – 262.

1718, *Doctrine of Chances*. London, 1738, 1756. New York, 1967.

1725, *Annuities upon Lives*. London, 1752.

1730, *Miscellanea Analytica*. London. Франц. перевод: Paris, 2009.

1733, A Method of Approximating the Sum of the Terms of the Binomial $(a + b)^n$ etc. Перепечатано в авторском английском переводе в его книге (1718/1738 и 1756), на с. 243 – 254 в издании 1756 г.

Другие авторы

Шейнин О. Б. (1970), К истории предельных теорем Муавра – Лапласа. *История и методология естеств. наук*, вып. 9, с. 199 – 211.

Bellhouse Chr. G. (2007), Maty's [1755] biography of A. De Moivre translated, annotated and augmented. *Stat. Sci.*, vol. 22, No. 1, pp. 109 – 136.

Cantor M. (1898), *Vorlesungen über Geschichte der Mathematik*, Bd. 3. New York – Stuttgart, 1965.

Czuber E. (1898), Die Entwicklung der Wahrscheinlichkeitstheorie und ihrer Anwendungen. *Jahresber. der Deutschen Mathematiker-Vereinigung*, Bd. 7, No. 2, pp. 1 – 279.

--- (1900), Wahrscheinlichkeitsrechnung. *Enz. Math. Wiss.*, Bd. 1, Tl. 2, pp. 733 – 767.

David F. N. (1962), *Games, Gods and Gambling*. London.

David F. N., Barton D. E. (1962), *Combinatorial Chance*. New York.

Daw R. H., Pearson E. S. (1972), A. De Moivre's 1733 derivation of the normal curve. *Biometrika*, vol. 59, pp. 677 – 680. Перепечатка: *Studies in the History of Statistics and Probability*, vol. 2, 1977, pp. 63 – 66. Редакторы M. G. Kendall & R. L. Plackett.

Laplace P. S. (1812), *Théorie analytique des probabilités*. *Œuvr. Compl.*, t. 7. Paris, 1886.

Montmort P. R. (1708), *Essay d'analyse sur les jeux de hazard*. Paris, 1713; New York, 1980.

Montucla J. E. (1758), *Histoire des mathématiques*. Paris, 1802, t. 3.

Ogborn M. E. (1962), *Equitable Assurances*. London.

Pearson K. (1924), Historical note on the origin of the normal curve of errors. *Biometrika*, vol. 16, pp. 402 – 404.

--- (1926), A. de Moivre. *Nature*, vol. 117, pp. 551 – 552.

Saar J. du (1923), De beteekenis van De Moivre's werk over lijfrenten etc. *Verzekerings archief*, Bd. 4, pp. 28 – 45.

Schneider I. (1968), Der Mathematiker A. de Moivre. *Arch. Hist. Ex. Sci.*, vol. 5, pp. 177 – 317.

Seal Hilary L. (1949a), Historical development of the use of generating functions in probability theory. *Mitt. Vereinigung Schweiz. Versicherungsmathematiker*, Bd. 49, pp. 209 – 228.

--- (1949b), Mortality data and the binomial probability law. *Skand. Aktuarietidskrift*, Bd. 32, pp. 188 – 216.

--- (1954), Budget of paradoxes. *J. Inst. Actuaries Students' Soc.*, vol. 13, pp. 60 – 65. Correction: там же, vol. 14, 1957, pp. 210 – 211.

Simpson T. (1742), *Doctrine of Annuities and Reversions*. London, 1775.

Todhunter I. (1865), *History of Mathematical Theory of Probability*. New York, 1949, 1965.

Walford C. (1871), Annuities. В книге автора *Insurance Cyclopaedia*, vol. 1. London, pp. 98 – 169.

Walker H. M. (1929), *Studies in the History of Statistical Method*. Baltimore.

--- (1934), A. de Moivre. *Scripta Math.*, vol. 2, pp. 316 – 333. Перепечатка: De Moivre (1718/1756, с. 349 – 368).

Westergaard H. L. (1932), *Contributions to History of Statistics*. New York, 1968.

Young T. E. (1908), Historical notes relating to the discovery of the formula etc. *J. Inst. Actuaries*, vol. 42, pp. 188 – 205.

IV

Уильям Краскл

Ричард Прайс

William H. Kruskal, Price, Richard. IES, pp. 733 – 734

Ричард Прайс (1723 – 1791) существенно способствовал успеху количественного направления [в науке] в Англии XVIII в. Будучи священником-диссидентом, он опубликовал важные сочинения по теории вероятностей, демографии, актуарным методам, социальной помощи, философии, богословию и политике.

В статистических кругах он лучше всего известен как редактор, комментатор и продолжатель мемуара Бейеса. Он применил идеи Бейеса при исследовании вероятности свидетельств, особенно относящихся к чудесам. По моему мнению, он забавно ошибался, перепутывая условные вероятности, а именно вероятность А при данном В и вероятность В при данном А.

Совсем в другой области, оценивая население Англии до того, как там начали проводить переписи, Прайс не применял идеи Бейеса. В 1780 г. он сообщил одну из самых низких существовавших оценок, от 5 до 11 миллионов, о которых происходили резкие споры. Сравнивая его мнение с ранними переписями, которые ведут начало с 1801 г., скажем, что его оценка действительно была намного занижена.

Прайс основал актуарную науку в Англии. В его время работа актуария не имела твёрдого обоснования ни в имеющихся данных, ни в теории. Он подошёл к этой теме, будучи озабочен пенсиями для священников и страхованием их вдов, а позже, беспокоясь о поддержке стариков вообще. Он поощрял улучшения в статистике выживания, прояснил многое в актуарной математике, опубликовал ранний и влиятельный учебник по страхованию (1771) и оказывал сильное воздействие на Equitable Life Assurance Soc. в начальный период существования этого общества.

Прайс рьяно поддерживал американских колонистов в их борьбе за свободу от произвольного Британского владычества, активно переписывался с Бенджамином Франклиным и длительное время был в дружбе с ним и с другими американскими руководителями. В конце жизни Прайс стал пылким сторонником французской революции, которая в то время только началась. Его проповедь (1789) привела к ставшему широко известному опровержению (Burke 1789).

В некотором смысле смерть Прайса в 1791 г. была своевременной, наступившей до кровавых крайностей во Франции. Все последующие авторы согласно указывают, что он имел мягкий характер, был прост в своих вкусах, исключительно искренен и скромн. В богословии и философии Прайс писал о свободе, молитвах, предусмотрительности и этике. В области этики его считали предшественником Канта, а его соответствующая работа комментируется до сего времени.

Библиография

R. Price

1771, *Observations on Reversionary Payments ...*, vols 1 – 2. London, 1783.

Большое число последующих изданий.

1789, *Discourse on the Love of Our Country ...* London, 1790. Шестое издание.

Другие авторы

Burke E. (1789), *Reflexions on the Revolution in France ...* London, 1790.

Cone C. B. (1952), *Torchbearer of Freedom. The Influence of R. Price on 18th Century Thought.* Lexington.

Glass D. V. (1973), *Numbering of the People.* Farnborough, England.

Holland J. D. (1968), An 18th century pioneer: R. Price. *Notes and Records Roy. Soc.*, vol. 23, pp. 43 – 64.

Laboucheix H. (1970), *R. Price, théoricien de la révolution américaine, le philosophe et le sociologue, le pamphlétaire et l'orateur.* Paris.

Morgan W. (1815), *Memoirs of the Life of the Revd R. Price.* London.

Приведенные подробности ненадежны.

Ogborn M. E. (1962). *Equitable Assurances.* London. Гл. 7-я посвящена Прайсу.

Thomas R. (1924), *R. Price, Philosopher and Apostle of Liberty.* London.

Черчилль Эйзенхарт

Карл Фридрих Гаусс

Churchill Eisenhart, Gauss Carl Friedrich. IES, pp. 378 – 386

[1. Биография]

Карл Фридрих Гаусс (1777 – 1855), величайший из немецких математиков, родился в Брауншвейге 30 апреля 1777 г. При крещении его назвали Иоганн Фридрих Карл, но от первого имени он отказался, а второе и третье поменял местами. Наряду с Архимедом и Ньютоном он считается величайшим математиком всех времён и народов. Впрочем, в нём весьма необычно сочетались интерес чистого математика к отвлечённым идеям и логической строгости; интерес физика-теоретика к созданию математических моделей физического мира; талант астронома в умении наблюдать, искусство экспериментатора в их приложении и в изобретении новых методов их проведения. Он был одарён изумительной способностью вычислений в уме, что позволяло ему экспериментально исследовать количественные соотношения и быстро и точно производить сложные или длительные вычисления¹. И, наконец, у него был талант и к изучению древних и современных языков, что оказалось его любимым занятием.

Гаусс внёс существенный вклад во все известные в то время ветви чистой и прикладной математики, а некоторые из них он сам и основал, и он также серьёзно способствовал успехам астрономии, геодезии, физики [, геофизики] и метрологии. Основные интересы Гаусса менялись во времени. До 1800 г. им владели филология и теория чисел; в 1800 – 1820 гг., астрономия; в 1820 – 1830 гг., геодезия, дифференциальная геометрия, конформное проектирование; в 1830 – 1840 гг., геомагнетизм, электромагнетизм и общая теория сил, обратно пропорциональных квадрату расстояний; в 1840 – 1855 гг., топология и геометрия функций комплексного переменного. Он умер в Гёттингене 23 февраля 1855 г. Из многих его сочинений наиболее полезными для физических, биологических и социальных наук являются те, которые относятся к методу наименьших квадратов (МНКв), первая формулировка которого произошла в 1795 – 1798 гг., вторая же, в 1821 – 1823 гг.

Гаусс был единственным ребёнком Гебхарда Дитриха Гаусса, каменщика и садовника, и его второй жены, урождённой Доротей Бенце, дочери каменотёса. Гебхард умел писать и считать, был счетоводом в местной страховой фирме и пользовался уважением горожан. Дома, однако, он был суров и груб, что отталкивало от него его выдающегося сына. Мать Гаусса, необразованная, не умевшая писать и читавшая с трудом, была жизнерадостна, умна и обладала сильным характером.

Её младший брат, Иоганн Фридрих, сердечный и необычно смысленный, был искусным ткачом художественной камки. Он быстро подметил необычную одарённость и возможности своего

племянника и любил оттачивать своё мышление разумом молодого гения. В раннем детском возрасте Гаусс был высокого мнения о своём дяде, а позже, в 1809 г., скорбя о его преждевременной смерти, часто заявлял, что в нём был потерян прирождённый гений.

В истории математики невероятно ранние способности Гаусса были неповторимы. Будучи ещё моложе трёх лет, он заметил ошибку в отцовском подсчёте выручки и указал верный результат. В возрасте 10 лет его приняли в начальный арифметический класс народной школы в Брауншвейге, и скорость и точность его вычислений в уме настолько поразила учителя, что тот купил для Гаусса лучший из существовавших учебник по арифметике. Гаусс быстро одолел его и тем самым убедил учителя, что превзошёл его. К счастью, тот имел помощника, заинтересованного в математике [Бартельса]², и Гаусс вместе с ним вечерами изучал алгебру и начала анализа. Общими усилиями они преодолевали трудности и обобщали доказательства.

К 11 годам Гаусс уже ознакомился с биномиальной теоремой. Заметив, что *доказательство* в учебнике годилось только для целого положительного показателя степени бинома, Гаусс [рассмотрел общий случай и] придумал своё собственное доказательство сходимости соответствующего бесконечного ряда. Тем самым он стал одним из первых ригористов, что вдохновило его к некоторым позднейшим самым великим из его трудов.

В начале 1791 г. о поразительных способностях Гаусса узнал герцог Брауншвейгский Карл Вильгельм Фердинанд. Он стал его покровителем и уплатил взносы за обучение Гаусса в колледже Collegium Carolinum в 1792 – 1795 гг., а затем, в 1795 – 1798 гг., в Гёттингенском университете. До своей смерти в 1806 г. герцог кроме того существенно помогал Гауссу деньгами.

В колледже Гаусс с равным успехом отдался классической литературе, философии и повышенным отделам математики. Он тщательно изучал труды Ньютона, Эйлера и Лагранжа. В марте 1795 г. (Dunnington 1955, с. 391) он заново установил квадратичный закон взаимности, этот бесценный принцип теории чисел, который Лежандр опубликовал в 1785 г., и впервые строго доказал его в 1801 г. И всё же, поступив в университет в октябре 1795 г., он ещё не решил, остановиться ли на математике или на филологии.

Математика победила 30 марта 1796 г., когда Гаусс доказал, что при помощи циркуля и линейки можно построить правильный 17-тиугольник. Две тысячи лет эта возможность ускользала от математиков. 1 июня 1796 г. он установил много больше: правильный многоугольник с нечётным числом сторон мог быть построен тем же образом, если и только если число его сторон являлось простым числом Ферма, т. е. было равно 2 в степени 2^n плюс единица³, или же выражалось произведением различных чисел указанного вида.

В том же году Гаусс впервые строго доказал фундаментальную теорему алгебры (каждый многочлен, не сводящийся к константе, имеет корень). Это доказательство стало предметом его

докторской диссертации, успешно защищённой заочно в 1799 г. в университете Хельмштедта.

Осенью 1798 г. Гаусс закончил последний вариант своих великих *Арифметических исследований*. Они начали печататься в 1799 г., но вышли в свет лишь в сентябре 1801 г. ввиду перехода типографии в другие руки. В этом сочинении Гаусс свёл воедино свои собственные оригинальные труды по теории целых и рациональных чисел, равно как и все основные соответствующие, но недостаточно связанные друг с другом результаты своих предшественников. Мало того, Гаусс настолько обогатил эти последние, строго сформулировав их заново, и создал объединённое целое, так что его великая книга считается началом теории чисел как отдельной систематической ветви математики.

Среди ранних достижений Гаусса было его упрощение исключительно сложных вычислений даты Пасхи (*Enc. Brit.*, 11-е издание, т. 4, с. 991 – 999), которые он свёл к нескольким простым формулам. Для любого года эти даты оказались возможным вычислить за несколько минут. Гаусс обратился к указанной задаче, поскольку его мать не смогла вспомнить точной даты его рождения, зная только, что оно произошло в среду, за 8 дней до Вознесения. Его вычисления были опубликованы в 1800 г., но день Пасхи для 1734 г. оказался ошибочным, и также неверным он был бы для 1886 г. В 1807 г. Гаусс подправил свои формулы, а позже указал ещё одну поправку, которая потребовалась бы, начиная с 4200 г. Не слишком трудное изложение метода Гаусса см. у Howe (1954).

[2. Обработка наблюдений и метод наименьших квадратов]

[2. Начальные исследования различных авторов.] МНКв Гаусс впервые изобрёл в свои студенческие годы. Осенью 1794 г. он (Galle 1924, с. 5) прочёл, что Ламберт (1765, с. 428 – 488), исходя из множества n ($n > 2$) точек наблюдения (Y_i, x_i) , вычислял коэффициенты линейного выражения

$$y = \alpha + \beta x$$

по методу средних⁴. Далее, в 1795 г. он придумал более простой и менее произвольный метод определения α и β по условию минимума суммы квадратов остаточных свободных членов v_i :

$$\alpha + \beta x_i - Y_i = v_i$$

и указал необходимые подробности вычислений за исключением определения весов наблюдений различной точности (Galle 1924, с. 7).

В 1797 г. Гаусс попытался теоретико-вероятностно обосновать свой метод, но (1821/1957, с. 143)

скоро убедился, что разыскание “вероятнейших” значений неизвестных было бы невозможно, если не будет известна функция, которая представляет собой вероятности ошибок.

Он продолжал:

Ему казалось, что наиболее естественно сперва избрать обратный путь и поискать такое распределение вероятностей ошибок, из которого должно вытекать общепризнанное удовлетворительное правило для всех простейших случаев на практике, именно, что среднее арифметическое из многих найденных значений одной и той же неизвестной величины, полученных из наблюдений одинаковой достоверности, должно считаться за вероятнейшее.

Понятие о распределении вероятностей или законе ошибок впервые появилось у Симпсона (1756)⁵. Он исследовал выборочное распределение среднего арифметического из независимых и одинаково распределённых ошибок, подчиняющихся дискретным равномерному и треугольному законам, и заключил (с. 93), что

чем больше при тех же самых обстоятельствах сделано наблюдений или экспериментов, тем меньше будет вывод ошибочен.

В 1757 г. он распространил своё исследование на непрерывное треугольное распределение. Иные непрерывные законы ошибок исследовали: Лагранж, равномерный закон, в 1774 г.;⁶ Лаплас, двойной экспоненциальный закон $Ce^{-m|x|}$, $-\infty < x < +\infty$; в 1774 г.; закон в виде полуокружности $C\sqrt{r^2 - x^2}$, $-r \leq x \leq r$, Даниил Бернулли в 1778 г.; и двойной логарифмический закон $C \ln a/|x|$, $-a \leq x \leq a$, Лаплас в 1781 г. Подробности см. Eisenhart (1964), Todhunter (1865), Merriman (1877).

Применяя апостериорную вероятность, Гаусс (1809, §§ 175 – 177) установил, что искомое распределение вероятностей ошибок было

$$\frac{h}{\sqrt{\pi}} \exp(-h^2 v^2) \quad (1)$$

и заметил, что (в современных обозначениях) $h = 1/\sigma\sqrt{2}$, где σ было *средней квадратической ошибкой*⁷ или *стандартным отклонением* распределения и *может рассматриваться как мера точности наблюдений* (§ 178).

Муавр в 1733 г. ввёл функцию $\exp(-t^2)$ для аппроксимирования сумм последовательных членов разложения бинома $(a + b)^n$ при больших n , а Лаплас в нескольких мемуарах 1774 – 1786 гг. весьма подробно исследовал применение этой функции и её производных для аппроксимирования различных распределений вероятности, возникающих в азартных играх, и особенно для биномиального и гипергеометрического распределений. Он также рассматривал соответствующую неполную бета-функцию, полученную при помощи теоремы Бейеса, и предложил табулировать её интеграл для применения при решении подобных задач (Todhunter 1865/1949, §§ 890 – 911). И всё же ни Муавр, ни Лаплас, видимо, не рассматривали функцию (1) или какое-либо равносильное ей выражение как

закон ошибок или распределение вероятностей в собственном смысле.

Затем Гаусс вывел необходимые следствия из указанной функции по методу, который мы сегодня назвали бы методом наибольшего правдоподобия, и обнаружил, что, если точность наблюдений различна, до вычисления квадрата суммы каждый соответствующий остаточный свободный член должен быть умножен на свою меру точности (§ 179), вывел формулу для точности $h(\bar{Y})$ любого взвешенного среднего арифметического из независимых наблюдений (§ 181) и обнаружил, что точность h обычного среднего арифметического равнозначных независимых наблюдений пропорциональна корню квадратному из их числа (§ 173) и вывел правило для образования *нормальных уравнений*, как мы их называем⁸ (§ 180). Наконец, он схематично пояснил свой знаменитый метод исключения неизвестных (§ 182), весьма широко известный в видоизменении Дулиттля (Doolittle 1878). Оно одновременно указывает легко оцениваемое выражение (§ 182, пункт 5 [?]) для соответствующего минимума квадрата остаточных свободных членов через попутно определяемые величины и привёл аналогичное выражение (§ 183) для оценки точности оценок неизвестных.

Запись в Дневнике Гаусса (W-10, с. 533) показывает, что сказанное выше он закончил 17 июня 1798 г. Был найден отрывок его первого применения всего сказанного весной 1799 г. (W-12, с. 64 – 68) и краткая заметка 24 августа 1799 г. о другом подобном применении была опубликована в октябре 1799 г. (W-8, с. 136 – 141, см. с. 136).

В 1801 г. Гаусс всерьёз занялся разработкой и применением новых астрономических методов. Их сочетание с его МНКв привело к удивительно точным предсказаниям, притом, что данных об орбитах новонайденных [малых] планет было немного. Успех его новых методов немедленно вывел Гаусса в ряд астрономов высшего класса (Bell 1937; Dunnington 1955, с. 49 – 57).

В 1805 г. Гаусс начал готовить (на немецком языке) свой второй шедевр, *Теорию движения* (1809), который содержал полный набор формул и методов для вычисления движения тела по коническому сечению и общий метод определения орбит комет или планет лишь по трём наблюдениям. Третий раздел (§§ 172 – 189) он посвятил подробному изложению теории и приложения своих методов (?) наименьших квадратов.

Немецкий текст был готов в 1806 г. (Galle 1924, с. 10) и предложен издателю в 1807 г. Ввиду весьма неопределённого политического положения, вызванного страшным поражением прусской армии (под командованием патрона Гаусса, герцога Брауншвейгского) силами Наполеона в битве при [Йене и] Ауэрштедте в октябре 1806 г., тот, однако, согласился публиковать его только на латинском языке.

По этой причине все подробности первой гауссовой формулировки МНКв были опубликованы лишь в 1809 г., а тем временем (?) независимая формулировка Лежандра появилась в

печати в 1805 г. Он пошёл лишь чуть дальше разработанного Гауссом в 1795 г. Никаких вероятностных рассуждений он не привёл, и не обсуждал точности или взвешивания наблюдений. Последовала полемика, ожесточённая со стороны Лежандра. Многие [но не сам Гаусс] приняли в ней участие, и она *крайне неблагоприятно отразилась на последующем развитии математики* (Bell 1937/1956, с. 331)⁹.

[2.2. Гаусс: точность наблюдений.] В 1809 г. Гаусс считал, что меры точности h_1, h_2, \dots соответствующих наблюдений были заданы, и ничего не сказал об их практическом определении. Первым в этом направлении оказался Бессель (1815, с. 234). Он ввёл *вероятную ошибку* как меру неопределённости, а позднее (1816, с. 142) определил её как такую величину r , которую погрешность может без учёта знака с одним и тем же шансом либо превзойти, либо оказаться меньше её.

Из 48 определений прямого восхождения Полярной он (1815, с. 234) вывел их вероятную ошибку по формуле

$$R = 0,8453 \sum_i \frac{|Y_i - \bar{Y}|}{n}.$$

Численный результат соответствует формуле Гаусса (1816, § 8). Бессель (1816, с. 141 – 142) показал, что для распределения (1) вероятная ошибка равна

$$0,476936/h = 0,8453\varepsilon = 0,6745\sigma.$$

Здесь $\varepsilon = 1/h\sqrt{\pi}$, средняя абсолютная ошибка распределения. Гаусс (1816, § 3) тут же заметил, что при m независимых [нормально] распределённых ошибках V_1, V_2, \dots, V_m , *будет ли m большим или меньшим, вероятнейшие значения* (т. е. оценки наибольшего правдоподобия) величин h и r равны

$$\hat{H} = \sqrt{\frac{m}{2 \sum V_i^2}}, \quad \hat{R} = 0,6744897 \sqrt{\frac{\sum V_i^2}{m}}.$$

Он (§ 4) показал, что при большом m величины $\hat{H} - h$ и $\hat{R} - r$ распределены примерно [нормально] с $h_H = \sqrt{m}/h$ и $h_R = \sqrt{m}/r$ и там же привёл явные выражения *вероятных пределов* (50%

доверительных пределов) истинных значений h и r через \hat{H} и \hat{R} .

Далее, он (§ 5) рассмотрел оценивание r по

$$R_p = C_p \sqrt[p]{\sum |V_i|^p / m},$$

где $C_p(m, h)$ выбирается так, чтобы r оказалось средним распределения R_p при больших выборках и установил, что наиболее точным является R_2 .

Для $m = 100$ R_2 оказывается настолько же точной оценкой для r , насколько R_1 при $m = 114$ и насколько R_3, R_4, R_5, R_6 при $m = 109$,

133, 178 и 251 (§ 6). Кроме того, Гаусс (§ 7) исследовал оценивание r по *самому среднему* (по медиане) из абсолютных ошибок $|V_i|$ при нечётном m . В этом случае требовалось, соответственно, $m = 249$. Оценивая в § 8 R_1, R_2, R_3, R_4 для 48 определений Бесселя (1815), Гаусс вычислил $(Y_i - \bar{Y})$ в качестве меры ошибки V_i и заключил, что Бессель применил формулу для R_1 .

[2.3. Гаусс: основной мемуар.] 15 февраля 1821 г. Гаусс представил Королевской академии наук в Гёттингене совершенно новую формулировку МНКв, полностью независимую от какого бы то ни было распределения вероятностей ошибок. Его мемуар (1823) исходил из предшествовавших работ Лапласа. Тот (1774/1891, с. 41 – 48) заявил, что при заданном распределении вероятностей независимых и одинаково распределённых ошибок параметр θ следует оценивать такой функцией $T = T(Y_1, Y_2, \dots)$ наблюдений Y_1, Y_2, \dots , для которой среднее (или ожидание) абсолютной погрешности оценивания $E|T - \theta|$ минимально.

Лаплас указал метод определения подобной функции при оценке параметра сдвига их общего распределения $f(y - \tau)$, если оно полностью известно за исключением значения τ . Наконец, он (1812, § 23) показал, что при независимых наблюдениях той же самой или различной точности метод Гаусса [1809] сведения к минимуму суммы квадратов остаточных свободных членов приводится к той же оценке, как и его собственный метод, если и только если погрешности наблюдений $X_i = Y_i - \tau$ распределены в соответствии с формулой (1) с заменой h и x на h_i и x_i , $i = 1, 2, \dots$

Гаусс (1823а, §§ 6 – 7) предложил, однако, применять среднюю *квадратическую* ошибку наблюдений $E(Y - \eta)^2$ или статистики $E(T - \theta)^2$ в качестве лучшей меры *ненадёжности*, и в § 9 вывел своё примечательное неравенство (Savage 1961, уравнение (9)) для вероятности того, что случайная величина Z с непрерывным одновершинным распределением вероятностей будет отличаться в ту или другую сторону от своей моды z_0 более, чем в λ раз своей средней квадратической ошибки, измеряемой от z_0 .

После этого Гаусс (§§ 18 – 23) показал, что если средние значения $EY_i = \eta_i$ множества независимых наблюдений Y_i являются линейными функциями k ($k < n$) неизвестных параметров θ_j , $j = 1, 2, \dots, k$ и (в современной терминологии) *дисперсии* $E(Y_i - \eta_i)^2 = \sigma_i^2$ величин Y_i конечны, МНКв приводит к оценкам T_j параметров θ_j , обладающими наименьшими средними квадратическими ошибками. Распределение (распределения) ошибок здесь может (могут) быть любым(и).

Во второй части мемуара Гаусс распространил этот результат на оценки L_r линейных функций $\Lambda_r(\theta_1, \theta_2, \dots, \theta_k)$, $r = 1, 2, \dots, m \leq k$, и показал (§§ 37 – 40), что минимум суммы квадратов остаточных свободных членов в точности равносильен сумме $(n - m)$ независимых ошибок (т. е. обладает $n - m$ степенями свободы). Многие современные авторы (Graybill 1961; Scheffé 1959; Zelen 1962) вывели и обсудили эти результаты, одно время ошибочно приписываемые Маркову (Neuman 1938/1952, p. 228), назвав их теоремой Гаусса – Маркова¹⁰. МНКв таким образом утвердился

как метод вывода *линейных несмещённых оценок с наименьшей дисперсией*.

Гаусс определённо предпочитал свою вторую формулировку МНКв, см. его автореферат (1821). В письме Бесселю 28 февраля 1839 г. (W/Erg-1; W-8, с. 146 – 147) он указал, что никогда не разъяснял, по какой причине отбросил метафизический подход своей первой формулировки. Решающей причиной послужило его убеждение в том, что приведение к максимуму ничтожной ошибки менее важно, чем достижение минимума вероятности крупных погрешностей¹¹.

В свои поздние годы Гаусс особенно гордился своим вкладом в развитие МНКв, и, несмотря на своё неизменное в течение всей жизни отвращение к преподаванию, ежегодно с 1835 г. до своей смерти читал курс по этому предмету в Гёттингенском университете. При его жизни МНКв стал основным инструментом в астрономии и геодезии по всему миру, и таковым он и остался. И когда в 1890-е годы Карл Пирсон и Юл начали заниматься математической теорией корреляции, они обнаружили, что, хотя цели корреляционного анализа прямо противоположны целям теории ошибок, можно было сразу же применить многое из математического инструментария, разработанного Гауссом.

Гальтон (1908, с. 305) заметил, что

*основные цели гауссовского закона ошибок были в некотором смысле в точности противоположны моим. Этими целями было избавление от ошибок или их правильный учёт. Но эти ошибки или уклонения были как раз тем, что я желал сохранять и понимать*¹².

И поэтому вклад Гаусса в МНКв включает математические рассуждения, существенные для статистической теории и её приложения почти к каждой отрасли современной науки.

[2. Геомагнитные наблюдения.] В 1829 г. Гаусса посетил бельгийский астроном, физик и статистик Адольф Кетле, который в то время проводил геомагнитные измерения в Голландии, Германии, Италии и Швейцарии. Геомагнетизмом Гаусс начал интересоваться примерно на рубеже XIX в. и в письме Ольберсу 1 марта 1803 г. (W/Erg-4) заметил: *Полагаю, что здесь открывается более широкое поле для приложения математики, чем думали раньше*. Несмотря на энергичные усилия Гумбольдта в 1804, а затем в 1828 г. убедить его заняться исследованиями в этой области, активная деятельность Гаусса в других отраслях науки воспрепятствовала этому.

Прибыв к Гауссу, Кетле застал его в передышке за изучением русского языка:

Я очень устал от занятий достаточно хорошо мне известными астрономией, геодезией и другими предметами, и захотел обратить внимание на язык, который совсем не знал, и вот теперь я читаю по-русски (Quetelet 1866, с. 646).

Измерение интенсивности земного магнетизма было новым для Гаусса, и он страстно желал узнать, как её измерять и какую точность можно достичь (с. 645). Кетле установил свой прибор во дворе у Гаусса, и они совместно проделали ряд экспериментов,

измеряя одновременно, но применяя несколько отличные методы. Согласие результатов поразило Гаусса, который воскликнул (с. 646): *Но ведь по своей точности эти наблюдения соответствуют астрономическим!*

С января 1831 г. геомагнитные наблюдения стали регулярно проводиться в Гёттингене. К февралю 1832 г. Гаусс был уже серьёзно вовлечён в геомагнитные исследования и обнаружил, что может выражать интенсивность геомагнетизма в *абсолютных единицах*, как он назвал их, т. е. в единицах трёх основных физических величин, длины, массы и времени. 15 декабря 1832 г. он представил свои результаты (1832) Гёттингенской королевской академии наук. Вскоре они были признаны одной из важнейших работ столетия.

Первые электромагнитные измерения Гаусс провел в октябре 1832 г. К Пасхе 1833 г. он, и его молодой коллега, физик Вильгельм Вебер, установили электромагнитный телеграф между обсерваторией Гаусса и физической лаборатории Вебера (Dunnington 1955, с. 147 – 148, 395). Вначале они посылали друг другу отдельные слова, затем полные предложения. В 1835 – 1836 гг. появились планы применения телеграфа на железной дороге Лейпциг – Дрезден, но от них пришлось отказаться, поскольку управление железной дороги заявило, что провода должны были бы быть проложены под землей. В городке Гёттингенского университета был установлен памятник: Гаусс и Вебер обсуждают свой телеграф.

В 1832 г. Гаусс начал подготавливать монографию о земном магнетизме, но недостаток экспериментального материала задержал её окончание, и она появилась только в 1839 г. По предложению Гаусса в Гёттингенском университете была построена магнитная обсерватория, и к 1836 г. Гёттинген стал основным европейским центром исследования геомагнетизма. Для координации одновременных наблюдений геомагнитных явлений в масштабе Европы была учреждена ассоциация магнитных обсерваторий, – Гёттингенский магнитный союз.

В 1837 г. Гаусс и Вебер совместно изобрели прибор типа гальванометра, бифилярный магнетометр, для измерения интенсивности магнитных полей. Этим прибором Гаусс проверял свой теоретический закон магнитного притяжения, обратно пропорционального расстоянию. В 1840 г. появился его и Вебера *Атлас земного магнетизма*, и тогда же его великое сочинение о том, что сейчас мы называем *теорией потенциала*. Оно оказалась вершиной работы Гаусса по физике и окончанием занятий магнетизмом.

К концу жизни Гаусс, как и Лаплас, Фурье и Пуассон, обратил внимание на социальные науки и на возможности приложения к ним физических наук. В частности, он живо заинтересовался применением теории вероятностей к социальным закономерностям. В 1847 г. он переписывался с датским астрономом Шумахером о законах смертности и составлении таблиц смертности (Quetelet 1866, с. 653 – 655).

[4. Король математики]

Уже при жизни Гаусса называли *королём математиков*, и он был осыпан почестями. Dunnington (1955, Приложение В) перечисляет их в хронологическом порядке, начиная с избрания членом-корреспондентом Императорской Петербургской академии наук в 1802 г. В 1804 г. его избрали в Королевское общество и в 1838 г. он был награждён медалью Копли. В 1810 г. он стал действительным членом Королевской академии наук в Берлине, членом Американской академии искусств и наук (Бостон) в 1822 г. и Американского философского общества (Филадельфия) в 1853 г., а в 1842 г. его наградили высшим орденом Прусского королевства *За заслуги*. 16 июля 1848 г., в точности через 50 лет после присуждения докторской степени, юбилей Гаусса был отпразднован в Гёттингене. Последовали новые почести, и в частности самое ценное для него присуждение звания почётного гражданина Брауншвейга и Гёттингена.

[5.] Добавление

Утверждение о том, что в свои школьные годы Гаусс доказал сходимость бесконечного ряда, возможно преувеличенное, исходит от сообщения Белла (Bell 1937/1956, p. 299), который в свою очередь несомненно имел в виду Сарториуса фон Вальтерсхаузена Sartorius von Waltershausen (1856/1966, p. 5), профессора геологии в Гёттингенском университете и близкого друга Гаусса в его поздние годы:

Бартельс [школьный учитель] смог достать полезные книги по математике, которые эти молодые люди изучали совместно. Таким образом Гаусс вполне ознакомился с биномиальной теоремой в её полной общности, а вскоре и с теорией бесконечных рядов, что открыло ему путь к высшему анализу.

Равносильность математических форм выражений, выведенных по методам *наибольшего правдоподобия* и *апостериорной вероятности* в случае равномерного априорного распределения оказалась источником постоянной путаницы. Нет сомнения в том, что Гаусс (1809, § 181) действительно вывел плотность распределения вероятностей

$$\frac{h\sqrt{n}}{\sqrt{\pi}} \exp[-nh^2(\bar{x} - \mu)^2].$$

Присоединив дифференциал $d\mu$, он истолковал полученное как апостериорное распределение μ при *заданном* наблюдаемом значении \bar{x} выборочного среднего и заранее принятом равномерном априорном распределении μ . При этих условиях \bar{x} очевидно оказывалось *вероятнейшим* значением μ . Сегодня, поступая примерно тем же самым образом, мы выводим ту же функцию, присоединяем к ней дифференциал $d\bar{x}$ и называем полученный результат *выборочным распределением выборочного среднего \bar{X} в повторных случайных выборках из нормального распределения с установленным средним μ и (установленным) стандартным отклонением $\sigma = 1/h\sqrt{2}$, а \bar{X} называем оценкой наибольшего правдоподобия для μ .*

Таким образом, грубо говоря, мы склонны приписывать Гауссу первоначальный вывод этого *выборочного распределения*. С другой стороны, я вполне уверен, что в 1809 г. Гаусс не представлял себе понятия *выборочного распределения* \bar{X} в *повторных выборках* при заданном μ . Аналогично, *стандартные ошибки* (Гаусс 1816)¹³ были несомненно выведены методом *апостериорной вероятности*, и, строго говоря, являются стандартными отклонениями апостериорных распределений вероятностей для соответствующих параметров при заданных наблюденных значениях соответствующих функций рассматриваемых выборочных значений.

С другой стороны, ввиду указанной выше равносильности форм, формулы Гаусса 1816 г. можно признать как приводящие к *стандартным ошибкам*, т. е. к стандартным отклонениям выборочных распределений функций выборочных значений при *установленных* значениях соответствующих параметров совокупностей. Снова грубо говоря, сегодня мы склонны приписывать Гауссу первый вывод формул этих *стандартных ошибок*, хоть в 1816 г. он и не владел понятием *выборочного распределения* для установленного значения параметра совокупности, т. е. о выборочной функции, оценивающей значение этого параметра.

В противоположность этому, в 1823 г. он недвусмысленно изучает выборочное распределение s , т. е. стандартного отклонения выборки независимых значений, как оценку стандартного отклонения σ апостериорного распределения совокупности, а не как апостериорное распределение σ при заданном s .

Biermann сообщил мне, что в своих заметках о посещении Гёттингена Кетле (Quetelet 1830, pp. 175 – 178) не упоминал, что Гаусс читал русские тексты. Эти заметки он написал вскоре после своего *научного путешествия* по Германии летом 1829 г. И действительно, Гаусс начал изучать русский язык лишь через 10 лет. В письме Шумахеру 17 авг. 1839 г. Гаусс указывал:

В начале прошедшей весны, рассматривая овладение новым мастерством как своего рода омолаживание, я стал заниматься русским языком [...] и нашёл, что он весьма интересен (W/Erg-5, № 3, с. 242 – 243), см. также Biermann (1964, с. 44). 22 августа Шумахер послал Гауссу русский астрономический календарь. Он был поражён, что Гаусс, в возрасте 62 лет, решил изучить новый язык как средство омолаживания. Он также спросил, не будет ли полезнее поискать отвлекающее средство в шахматах. 8 сентября Гаусс поблагодарил Шумахера за календарь, и ответил, что шахматы слишком похожи на его другие занятия и не смогут дать ему никакого облегчения, см. там же, с. 247 – 248, 269, и также Biermann (1964, с. 44).

В своей биографии Шумахера Кетле (1866, с. 639) упоминает 6 томов переписки Гаусса и Шумахера, называет её *очень интересной* и цитирует её том 5 по поводу биографии Гаусса (с. 653). Следовательно, как Бирман указал мне, Кетле несомненно узнал о том, что Гаусс изучал русский язык при чтении

некоторых писем из указанной переписки, затем видимо забыл об этом обстоятельстве и в возрасте 70 лет, составляя биографию Гаусса, поверил [в то, что автор описал выше].

Примечания

1. Гаусс вычислял необычно быстро и нередко ошибался (Maennchen 1918/1930, p. 65).
2. Баргельс впоследствии стал учителем Лобачевского (Депман 1950).
3. По ошибочному мнению Ферма все подобные числа были простыми.
4. Метод средних применялся для уравнивания косвенных измерений при избыточном числе наблюдений. Пусть при двух неизвестных имеются уравнения

$$a_i x + b_i y + l_i = v_i$$

с измеренными свободными членами и коэффициентами, заданными соответствующей теорией. Ненулевые остаточные члены v_i неизбежны, и в методе средних их сумма принималась равной нулю. Подробнее см. Идельсон (1947, § 21).

5. До Симпсона непрерывный равномерный закон смертности вводили Гюйгенс, Ник. Бернулли и, в 1725 г., Муавр. Термина *закон ошибок* у Симпсона не было, он и не вводил подобного закона.

6. Мемуар Лагранжа был опубликован в 1776 г. См. о нём Шейнин (2013, § 7.3.1). Неясно, как этот двойной логарифмический закон мог применяться при небольших по абсолютной величине значениях x .

7. Термин средняя квадратическая ошибка появился, видимо, в конце XIX в. Как заметил David (p. 221), Эджуорт (Edgeworth 1885, с. 188/1996, с. 31) ввёл среднюю квадратическую ошибок (mean square of errors).

8. Термин *нормальные уравнения* ввёл сам Гаусс (1823b/1957, с. 30).

9. Вместо письма Гауссу Лежандр вполне мог бы просто указать впоследствии, что Гаусс, считая МНКв, *своим*, нарушает установившийся порядок, и что он, Лежандр, уверен, что все согласятся с ним.

Белл говорит о влиянии на математику в целом, мы же можем только добавить, что плохими оказались последствия для теории ошибок. Французские математики включая Пуассона (но не Лаплас) не обращались к сочинениям Гаусса по теории ошибок, и их научные результаты оказались гораздо слабее возможного. Пуассон (1833) в своём выступлении на похоронах Лежандра обошёл Гаусса молчанием.

10. Нейман (Neuman 1934, с. 595) ошибочно приписал Маркову вторую формулировку МНКв, и далее F. N. David & Neuman (1938) повторили эту ошибку. В том же 1938 г., см. основной текст, Нейман признал свою ошибку.

11. Гораздо лучше сказать, что Гаусс предпочёл интегральную меру точности дифференциальной.

12. Мнение Гальтона сомнительно. Чтобы избавиться от ошибок надо было вначале исследовать их источники, возможную величину и т. д., а *закон ошибок* Гальтон упомянул напрасно.

13. Термина *стандартная ошибка* у Гаусса не было.

Библиография

Gauss C. F., Гаусс К. Ф.

Обозначения: W-i = *Werke*, том i; W/Erg-i = *Werke, Ergänzungsreihe*, том i

1800, Berechnung des Osterfestes. W-6, pp. 73 – 79.

1801, латин, *Арифметические исследования*. В книге автора *Труды по теории чисел*. М., 1959.

1807, Noch etwas über die Bestimmung des Osterfestes. W-6, pp. 82 – 86.

1809, латин, *Теория движения небесных тел ...*, кн. 2, раздел 3. В книге автора (1957, с. 89 – 109).

1816, нем. Определение точности наблюдений. Там же, с. 121 – 128

1821, нем., Авторское сообщение о мемуаре 1823 г. Там же, с. 141 – 144.

1823a, латин., Теория комбинаций наблюдений ..., части 1 – 2. Там же, с. 17 – 57.

1823b, нем., Приложение теории вероятностей к одной задаче практической геометрии. Там же, с. 129 – 133.

1825 – 1827, латин., *Общие исследования о кривых поверхностях*. В книге *Об основаниях геометрии*, 2-е издание. Казань, 1895.

1832 – 1839, латин. и нем., *Избранные труды по земному магнетизму*. М., 1952.

1840, нем., General propositions relating to attractive and repulsive forces acting in the inverse ratio of the square of the distance. В книге Taylor R., редактор, *Scientific Memoirs*, vol. 3. New York, 1988, pp. 153 – 196.

1840, *Atlas des Erdmagnetismus etc.* W-12, pp. 335 – 408. Соавтор W. Weber.

1863 – 1930, W-1 – 12. Göttingen. Перепечатка Hildesheim, 1973 – 1981.

1887, *Abhandlungen zur Methode der kleinsten Quadrate*. Hrsg. A. Börsch, P. Simon. Vaduz, 1998.

1957, *Избр. геодезич. соч.*, т. 1. М.

1975 – 1987, W/Erg-1 – 5. Hildesheim. Переписка Гаусса. Ссылки на переписку с Бесселем, Ольберсом и Шумахером: W/Erg-1, 1880 (1975); W/Erg-4, 1890 – 1899/1987; W/Erg-5, 1860 – 1865/1975.

1995, латин. и англ., *Theory of combination of observations etc.* Переводчик и автор Послесловия W. Stewart. Philadelphia.

Другие авторы

Депман И. Я. (1950). М. Ф. Бартельс – учитель Н. И. Лобачевского. *Историко-математич. исследования*, вып. 3, с. 475 – 485.

Идельсон Н. И. (1947), *Способ наименьших квадратов ...* М.

Шейнин О. Б. (2013), *Теория вероятностей. Исторический очерк*. Берлин. Также Google, Oscar Sheynin, Download area.

Bell E. T. (1937), Gauss: the prince of mathematicians. В книге Newman J.R., редактор, *The World of Mathematics*, vol. 1. New York, 1956, 1960, pp. 295 – 299.

Bessel F. W. (1815), Über den Ort des Polarsterns. *Astron. Jahrb.* за 1818, pp. 233 – 241.

--- (1816), Untersuchungen über die Bahn des Olbersschen Kometen. *Abh. Akad. Wiss. Berlin* за 1812 – 1813, Math. Kl., pp. 119 – 160.

Biermann K.-R. (1964), Einige Episoden aus den russischen Sprachstudien der Mathematikers C. F. Gauss. *Forschungen u. Fortschritte*, Bd. 38, pp. 44 – 46.

David F. N., Neyman J. (1938), Extension of the Markoff theorem on least squares. *Stat. Res. Mem.*, vol. 2, pp. 105 – 117.

David H. A. (2001), First (?) occurrence of common terms in statistics and probability. В книге David H. A., Edwards A. W. F., *Annotated Readings in the History of Statistics*. New York, pp. 209 – 246.

De Moivre A. (1733 латин.), Method of approximating the sum of the terms of the binomial etc. В последующих изданиях книги автора *Doctrine of Chances*. London, 1738 и 1756, с. 243 – 254. Перепечатка последнего издания: Нью-Йорк, 1967.

Doolittle M. H. (1878), [Solution of normal equations etc.] *Rept Superintendent U. S. Coast and Geod. Survey*, Appendix 8, Paper 3, pp. 115 – 120.

Dunnington G. W. (1955), *C. F. Gauss, Titan of Science*. New York. [Washington, 2004.]

Edgeworth F. Y. (1885), Methods of Statistics. *Jubilee Volume of the Royal Statistical Society*, pp. 181 – 217. Перепечатка в сочинениях автора *Writings in Probability, Statistics and Economics*, vol. 2. Cheltenham, 1996, pp. 24 – 60.

Eisenhart C. (1964), The meaning of “least” in least squares. *J. Wash. Acad. Sci.*, vol. 54, pp. 24 – 33.

Galle A. (1924), Über die geodätischen Arbeiten von Gauss. W-11, часть 2, первая статья, отдельная пагинация.

Galton Fr. (1908), *Memories of My Life*. London.

Graybill F. A. (1961), *Introduction to Linear Statistical Models*, vol. 1. New York.

Hänselmann L. (1878), *K. F. Gauss*. Leipzig.

Howe H. H. (1954), How to find the date of Easter. *Sky and Telescope*, vol. 13, p. 196.

- Lambert J. H.** (1765), *Beiträge zum Gebrauche der Mathematik und deren Anwendung*, Bd. 1. Berlin.
- Laplace P. S.** (1774), Sur la probabilité des causes par les événements. *Œuvr. Compl.*, t. 8. Paris, 1891, pp. 27 – 65.
- (1812), *Théorie analytique des probabilités*. *Œuvr. Compl.*, t. 7. Paris, 1886.
- Maennchen Ph.** (1918), Gauss als Zahlenrechner. W–10, часть 2, шестая статья, отдельная пагинация.
- May K. O.** (1972), C. F. Gauss. *Dict. Scient. Biogr.*, vol. 5, pp. 208 – 215.
- Merriman M.** (1877), List of writings relating to the method of least squares etc. В томе 1 сборника перепечаток без сквозной нумерации страниц Stigler S. M., *American Contributions to Mathematical Statistics in the 19th Century*, vols 1 – 2. New York, 1980.
- Neuman J.** (1934), Two different aspects of the representative method. *J. Roy. Stat. Soc.*, vol. 97, pp. 558 – 625.
- (1938), *Lectures and Conferences on Mathematical Statistics and Probability*. Washington, 1952.
- Poisson S. D.** (1833), Discours prononcé aux funérailles de Legendre. *J. f. d. reine u. angew. Math.*, Bd. 10, pp. 360 – 363.
- Quetelet A.** (1830), Notes extraites d'un voyage scientifique etc., pt. 2. *Corr. math. et phys.*, t. 6, pp. 161 – 178.
- (1866), C. F. Gauss. В книге автора *Sciences mathématiques et physiques etc.* Bruxelles, pp. 643 – 655.
- Sartorius von Waltershausen W.** (1856), *Gauss zum Gedächtniss*. Leipzig, 2012. Colorado Springs, 1966: английский перевод.
- Savage I. R.** (1961), Probability inequalities of the Chebyshev type. U. S. Nat. Bureau of Standards, *J. of Res.*, Math. and Math. Phys., vol. 65B, pp. 211 – 222.
- Scheffé H.** (1959), *The Analysis of Variance*. New York.
- Simpson T.** (1756), On the advantage of taking the mean of a number of observations. *Phil. Trans. Roy. Soc.*, vol. 49, pp. 82 – 93.
- Todhunter I.** (1865), *History of the Mathematical Theory of Probability*. New York, 1949, 1965.
- Whittaker E. T., Robinson G.** (1924), *Calculus of Observations*. Princeton, 1944. 4-е издание. *Математическая обработка результатов наблюдений*. М., 1935.
- Zelen M.** (1962), Linear estimation and related topics. В книге Todd J., редактор, *Survey of Numerical Analysis*. New York, pp. 558 – 584.

VI

Дейвид Ландау, Пол Ф. Лазарсфельд

Адольф Кетле

David Landau, Paul F. Lazarsfeld, Quetelet, Adolphe. IES, pp. 824 – 834

[1. Биография]

Ламберт Адольф Жак Кетле (1796 – 1874), лучше всего известный своими трудами по статистике, родился в бельгийском городе Генте. Когда ему было семь лет, умер его отец, и после окончания средней школы ему пришлось в 17-тилетнем возрасте зарабатывать себе на жизнь. Кетле согласился преподавать математику в средней школе в Генте, хотя в то время склонялся он не к науке, а к искусству. Некоторое время он пробыл учеником художника, затем нарисовал несколько хорошо встреченных картин. Он также сочинял довольно неплохие стихи и вместе со своим школьным товарищем Данделеном работал над оперой.

Лишь под влиянием Жана Гильома Гарнье, профессора математики в недавно учреждённом Гентском университете, Кетле, наконец, убедился в том, что ему следовало отойти от артистических попыток и полностью посвятить себя изучению математики. Впрочем, примерно до 30-и лет он продолжал заниматься поэзией как любитель.

Его докторская диссертация, в которой он заявил об открытии новой кривой, *la focale*¹ была первой, представленной в университете (24 июля 1819 г.), и её приветствовали как оригинальный вклад в аналитическую геометрию, а в 23 года Кетле призвали в Брюссель заведывать кафедрой элементарной математики в свободном университете *Athaeum*. Всего лишь через несколько месяцев, в начале 1820 г., его избрали членом брюссельской Королевской академии наук [, словесности] и изящных искусств.

Кетле так и не снизил скорости первоначального продвижения своей научной карьеры, и его научная производительность в течение последующих 50 лет была необыкновенной. Его деятельность в течение примерно 10 лет после прибытия в Брюссель хорошо иллюстрирует его невероятную работоспособность. Он сразу же начал публиковать длинную серию очерков, в основном по математике и физике. Вначале они появлялись в академических *Nouveaux Mémoires*, затем в *Correspondances Mathématiques et Physiques*. Этот журнал он основал в 1825 г. и редактировал его (первые два года совместно с Гарнье) до конца существования в 1839 г. Какое-то время *Correspondances* оставался главным европейским периодическим изданием в своей области; свои рукописи посылали туда самые выдающиеся учёные континента Европы.

В 1824 г. Кетле кроме того взялся за чтение публичных лекций в Брюссельском музее, вначале по геометрии, теории вероятностей, физике и астрономии, а затем по истории наук. Там,

как в Athenaeum, он вскоре стал известен как выдающийся учитель, и его аудитория была всегда переполнена студентами, слушателями вообще и известными учёными из всех европейских стран. Лекции он читал до 1834 г., после чего музей стал частью Брюссельского университета.

Кетле предложили в нём кафедру математики, но он отказался, посвятив себя своим многочисленным исследованиям. Впрочем, в 1836 г. лекции возобновились в Брюссельском военном училище, которое было учреждено два года ранее. В тот же период Кетле опубликовал несколько элементарных работ (1826; 1827a) по естествознанию и математике для широкого круга читателей. Некоторые биографы, особенно Reichesberg (1896) и Hankins (1908) заявили, что вторая книга Кетле была *почётно* включена в Индекс (список запрещённых), но Lottin (1912, с. 34 – 37) доказал, что это утверждение было вымышленным.

Тогда же Кетле (1827b) опубликовал сводку своих лекций по физике. В 1835 г. она была переведена на английский язык, и Robert Wallace в Предисловии к переводу отметил её значимость:

Ни одна другая работа по физике на английском языке не содержала такого широкого и сжатого отчёта о различных ветвях физики и не охватывала всего подобного материала в столь небольшом объёме.

Через год Кетле (1828) опубликовал книгу, назвав её сводкой своих вводных лекций по физике и астрономии. Она ознаменовала его отход от сосредоточения на математике и естествознании к изучению статистики² и, впоследствии, к исследованию социальных явлений.

Интересно, однако, что Кетле не считал указанные занятия своим основным делом или основным достижением десятилетнего периода 1823 – 1832 гг. Больше всего он заботился об устройстве обсерватории в Брюсселе. Неизвестно, правда, почему именно он начал считать это своей основной целью, но ясно, что таким образом он совершенно случайно серьёзно изменил направление своей научной карьеры.

Согласившись с предложением об учреждении обсерватории, министр образования тут же послал Кетле в Париж для ознакомления с новейшими астрономическими методами и инструментами. Там его тепло встретили Араго и Бювар³, и последний ввёл его в кружок интеллектуалов, в центре которого были знаменитые математики Пуассон, Лаплас и Фурье. Какое-то время эти учёные занимались закладкой фундамента современной теории вероятностей, а некоторые из них исследовали эмпирические социальные данные.

Наставления этих математиков, особенно Лапласа, равно как и постоянное непринуждённое общение с их группой в Политехнической школе возбудили у Кетле живой интерес к статистическим исследованиям и теории, основанной на теории вероятностей. Эта тема стала средоточием всей его научной работы. В последующих воспоминаниях он указал, что после ознакомления со статистическими идеями своих французских наставников немедленно подумал о применении этих идей к

измерению человеческого тела, о чём он начал любопытствовать ещё в свою бытность художником.

Одним из непосредственных следствий изучения теории вероятностей стало его понимание *необходимости исследования земных явлений дополнительно к изучению небесных, что до сих пор не было возможным* (1870). Решающее воздействие парижского опыта на его мысли стало очевидным (там же, через несколько строк): *Так, именно среди учёных статистиков и экономистов того времени я начал свои изыскания ...*

После возвращения Кетле из Парижа в 1823 г. строительство обсерватории продолжалось урывками. Вначале произошли затруднения с финансированием и некоторое разногласие между Кетле и архитектором, а позднее строительство прервалось ввиду бельгийской революции 1830 г.: недостроенная обсерватория использовалась как наскоро оборудованная крепость и была повреждена. Наконец, в 1832 г. Кетле поселился в почти построенной обсерватории, возникшей при его поездке в Париж.

[2.] Сбор данных

Пока строилась обсерватория, интерес Кетле к статистике, равно как и его явные организационные способности, привели его ко всё более активному участию в исследованиях, требующих сбора эмпирических социологических данных. В 1826 г. была учреждена Королевская статистическая комиссия, и Кетле стал её корреспондентом в провинции Брабант. С 1814 по 1830 гг. Бельгия находилась под владычеством Голландии, так что *Королевская* в данном случае относилась к оранскому дому.

Первые публикации Кетле содержали количественные сведения о Бельгии, которые можно было практически использовать, например, для составления таблиц смертности и, стало быть, особенно для решения проблем страхования. В 1827 г. он исследовал уголовную статистику, имея в виду усовершенствовать отправление правосудия. В 1828 г. Кетле отредактировал общестатистический справочник по Бельгии с обширным объёмом сравнительного материала, полученного от французских коллег и из Англии. По его настоянию в 1829 г. была проведена перепись, и после революции 1830 г. её труды были опубликованы отдельно для Голландии и Бельгии.

В 1841 г., в основном ввиду усилий Кетле, была учреждена Центральная статистическая комиссия, которая вскоре стала центральным агентством по сбору статистики в Бельгии. Он же вплоть до своей смерти оставался её председателем, и под его руководством она с замечательной тщательностью действительно исполняла свои обязанности и тем самым установила стандарт для подобных организаций по всей Европе.

В 1833 г. Кетле стал официальным представителем [правительства] на собрании Британской ассоциации продвижения науки и сыграл ключевую роль в образовании её статистической секции. Однако, разочаровавшись узостью области её деятельности, он убедил её председателя, Беббиджа, в необходимости учреждения Лондонского статистического

общества⁴. Оно было создано в 1834 г. и с 1877 г. называется Королевским.

В своих статистических сочинениях, равно как и в естествознании, Кетле особенно настаивал на необходимости единообразия в методах сбора данных, табулирования и представления результатов. Основной целью всех его организационных усилий было проведение этих мер в жизнь. В 1851 г. Кетле предложил группе учёных, собравшихся в Лондоне на Всемирную выставку, план международного сотрудничества в сборе статистических данных. Его идея была горячо одобрена, и дело продвинулось так быстро, что в 1853 г. в Брюсселе собрался первый Международный статистический конгресс. На предварительном заседании Кетле был избран президентом, и он, естественно, посвятил свой вступительный доклад значимости единообразия методов и терминологии в официальных статистических публикациях.

В течение последующих 25 лет Конгресс⁵ исключительно эффективно побуждал развивать официальную статистику в Европе, устанавливая постоянные связи между статистиками и улучшая сравнимость результатов. Но в течение 1870-х годов внутренние раздоры и разногласия ослабили Конгресс и в конце концов привели к его роспуску в 1880 г. К тому времени, однако, первоначальные предложения Кетле о необходимости некоторой международной организации для сохранения единообразия и способствования сотрудничеству в сборе и анализе официальной статистики стали полностью восприниматься. Неудивительно, что всего лишь через пять лет была учреждена новая организация, Международный статистический институт, который должен был продолжать деятельность Конгресса.

Чтобы должным образом закончить этот рассказ об успехах Кетле как организатора, остаётся лишь упомянуть его выдающуюся роль в Королевской академии. В 1820 г., когда он был избран её членом, Академия почти умерла. Лишь с полдюжины престарелых членов посещало её сессии, и по существу она ничего не публиковала. Кетле вдохнул в Академию новую жизнь и энергию, быстро принял на себя основную ответственность за её деятельность, пополнил её ряды многими своими молодыми коллегами и укрепил её публикации своими собственными многочисленными трудами. В 1832 и 1833 гг. Кетле был директором Академии, а в 1834 г. его избрали неперменным секретарём, и эту должность он занимал 40 лет, в течение которых его считали *руководящей душой Академии*.

[3.] Исследования и теория социальных явлений

Два мемуара, ставшие основанием всех его последующих исследований социальных явлений, Кетле опубликовал в 1831 г. К тому времени он решил, что хотел бы выделять из общих статистических совокупностей специальные части, относящиеся к людям. В первом мемуаре (1831a) он использовал большое число измерений физических размеров человека, а через несколько месяцев появилась статистика преступлений (1831b). Упор в этих мемуарах был на том, что мы назвали бы циклом жизни. Оба они

содержали большое число многомерных таблиц, например, различия преступности мужчин и женщин в разных возрастных группах в различных странах, притом для разных социальных групп. Hanks (1908, с. 55) и Lottin (1912, с. 128 – 138) заметили, что основную идею второго мемуара, постоянство годового *бюджета* преступлений для каждой возрастной группы, можно проследить до мемуара, опубликованного в начале 1829 г. Кетле (1835, английский перевод 1842 г., с. 96) был, видимо, прав, утверждая здесь свой приоритет вопреки Guerry (1832).

Кетле (1833а) опубликовал третий мемуар об изменении веса человека с возрастом. К тому времени он сформулировал идею о социальной физике и свёл воедино свои ранние мемуары (1835), а много позже (1869) выпустил исправленное издание этой сводки, включив в качестве одного из основных разделов весьма положительную анонимную рецензию, написанную Дж. Гершелем (1850) на другие труды Кетле, в первую очередь на его *Письма* (1846).

К ранее замеченному постоянству преступности он добавил доказательства закономерности в самоубийствах и в относительных количествах женитьб для мужчин и женщин и разных возрастных групп. Он вообще был убеждён, что существуют и многие другие закономерности, но указанные три были единственными *моральными* характеристиками человека (зависящими от выбора), которые он доказал. С публикацией этой книги (1835) все основные идеи Кетле стали доступными широкому кругу читателей.

[3.1.] Основные принципы. В 1830-е годы Кетле установил теоретические основания своих трудов по моральной статистике (современный термин, социология⁶). Во-первых, он высказал идею об исключительной закономерности социальных явлений вообще и о том, что эмпирические закономерности могут быть выявлены статистическими методами. Далее, существовали причины этих закономерностей. Кетле полагал, что его средние аналогичны *физическим фактам* и таким образом указал на связь физических и социальных законов. Однако, в отличие от Зюссмильха и других авторов предшествовавшего века, он объяснял эти законы не божественным провидением, а социальными условиями, зависящими от времени и места.

Из этих мыслей вытекали два следствия. Появилось много этических проблем, ставящих под сомнение свободную волю человека, и, к примеру, индивидуальную ответственность за преступления, а в практическом смысле они заложили основу для убеждения в возможности ослабления преступности и относительного числа самоубийств улучшением законодательства. В методологическом смысле в работах Кетле очень рано появилось два ключевых принципа. Первый состоял в том, что *причины пропорциональны своим действиям* (1831b, с. 7). Его легко признать по отношению к физическим характеристикам человека и он позволяет нам заключить, к примеру, что один человек *вдвое сильнее* другого (причина) просто потому, что мы *видим*, что он может поднять что-то вдвое тяжелее (действие).

Кетле предположил, что научное исследование моральных и умственных качеств человека возможно лишь, если этот же принцип удастся приложить к ним. О роли этого принципа в теориях Кетле см. ниже.

Второй ключевой принцип состоял в том, что для вывода надёжных заключений необходимы большие числа. Его можно объяснить влиянием Лапласа (1812), Фурье (1826) и Пуассона (1837)⁷. Вот иллюстрация переплетения этих принципов с теоретическими идеями, описанными выше (1831b, с. 80 – 81):

Мне представляется, что то, что относится к роду человеческого, рассмотренному в больших числах, аналогично физическим фактам. Чем больше число индивидуумов, тем более стирается влияние индивидуальности и заменяется рядом общих фактов, которые зависят от общих причин, в соответствии с которыми существует и поддерживается общество. Именно эти причины мы хотим понять, и тогда мы смогли бы выявить их влияние на социальные явления в том же порядке, в котором мы устанавливаем действия под влиянием причин в физических науках.

Кетле был серьёзно озабочен тем, чтобы принятые им методы изучения человека во всех отношениях были столь же *научными*, как те, которые применялись в любой физической науке. Решение этой задачи он видел в выработке методологии, которая позволила бы полностью применять теорию вероятностей. Поразительно противореча своему современнику Конту, Кетле считал, что применение математики является не только единственным необходимым условием любой точной науки, но и мерой её значимости⁸ (1828, с. 230):

Чем дальше продвигаются науки, тем сильнее они стремятся к сфере математики как к некоторому центру. Мы можем судить о мере совершенства, к которому подошла наука, по меньшим или большим усилиям, которые требуются для приближения к наблюдениям при помощи вычислений.

[3.2.] Схема работы. До того, как подробно описывать работу Кетле в области моральной статистики, следует указать на его метод публикации своих трудов. Понять этот метод помогает его литературная квалификация и тот факт, что его друзья-гуманитарии оставались для него важным ориентиром.

Получив новые данные или разработав новый метод или идею, он вначале сообщал о своём открытии в кратких заметках, обычно в отчётах Академии или в своём журнале *Correspondances* ..., а иногда во французских или английских журналах. После подобной публикации он описывал тот же материал в более подробных статьях и истолковывал его в социальном и философском направлении. Наконец, он объединял эти статьи в книги, надеясь на их привлекательность для многих читателей. Он безусловно очень сильно чувствовал, что эмпирические находки должны быть истолкованы как можно лучше и представлены в виде, интересном для читателей с широкими социальными и гуманными заботами.

Кетле усиливал своё влияние обширной перепиской с учёными, государственными деятелями и литераторами Европы и Америки. Wellens-De Donder (1964) установила примерно 2500 его корреспондентов, включая Гаусса, Ампера, Фарадея, Ал. Гумбольдта, James A. Garfield, в то время члена Конгресса США, который попросил у Кетле совета о мерах усовершенствования переписей, Джозефа Генри, Lemuel Shattuck, Чарлза Уитстона, Виллерме и Гёте, который дружески отнёсся к Кетле при его поездке в Германию в 1829 г. Но быть может самыми знаменитыми корреспондентами были князья Эрнст и Альберт, которых он с 1836 г. обучал математике по просьбе их дяди Леопольда I, *короля бельгийцев*. Они, правда, вскоре покинули Бельгию, чтобы посещать школу в Германии, но уроки продолжались много лет по переписке. Во второй основной работе Кетле по моральной статистике (1846) опубликованы его письма из этой переписки. В полном названии книги упомянут только Эрнст, который в качестве правящего герцога был главой Кобургского дома.

Социальную систему (1848) Кетле посвятил Альберту, с которым он установил особо тесную дружбу. Сильное влияние Кетле на мысли Альберта ясно видно во вступительном обращении Альберта к четвёртой сессии Международного статистического конгресса в Лондоне 16 июля 1860 г. (Schoen 1938).

[3.3.] Средний человек. Понимание социальной реальности у Кетле подчинено его идее о *среднем человеке*. Он сам (1848, с. vii – ix) указал, что таковым было его центральное понятие и проследил за его развитием в ряде своих трудов. Так, он заметил, что в 1835 г. считал, что характеристики среднего человека можно только представить, если учесть среднее и верхнюю и нижнюю границы отклонений от него; в 1846 г. он показал, что *индивидуальные значения роста человека какой-либо нации группируются симметрично около среднего в соответствии с [...] законом случайных причин*⁹ (с. viii) и что для нации *средний человек фактически является типом или стандартом и что люди более или менее отличаются от него только ввиду влияния случайных причин, чьё действие можно вычислить, если число испытаний достаточно велико [...]. Я показываю, что закон случайных причин является всеобщим и относится к индивидуумам и народам, что он управляет нашими моральными и умственными качествами так же, как и физическими, и то, что считается случайным, перестает быть таковым, когда наблюдения распространяются на большое число случаев* (с. ix).

Неудивительно, что обсуждение теорий и исследование общества у Кетле неизменно начинаются с его понятия *среднего человека*. К этому понятию Кетле (1831a) впервые подошёл при измерении физических характеристик человека, в частности роста и веса. Он представлял средний рост группы людей примерно одного возраста как среднее, около которого *колеблется* рост всех людей этого возраста, хотя и не смог сказать как именно колеблется.

Уже в этом первом описании своего понятия Кетле предположил, что аналогичные средние и колебания можно наблюдать при измерении моральных и умственных качеств. Сам термин *средний человек* впервые появился в мемуаре (1831b). Там же впервые сформулирована идея о том, что, будь средний человек определён для некоторой нации, он оказался бы её *типом*, и аналогично произошло бы при исследовании человека в целом.

Следующий шаг в развитии понятия среднего человека произошёл в 1844 г., когда Кетле впервые заметил, что его наблюдения были симметрично расположены около среднего, почти так же, как следовало ожидать, если исходить из биномиального и нормального распределений. Он здесь размышлял о вероятности того, что и все физические характеристики могут быть распределены таким же образом. Применяв теорию вероятностей, он смог вывести теоретические распределения роста, веса или окружности грудной клетки, замечательно совпадавшие с эмпирическими распределениями его данных.

В том же мемуаре содержался интересный факт, который он установил при помощи того же метода, а именно уклонение от призыва во французскую армию. Определив расхождение между распределением роста 100 000 французских призывников и его предсказанием (т. е. теоретическим распределением, вычисленным в предположении вероятной ошибки, равной 49 мм), он заключил, что примерно 2 тысячи человек уклонилось от военной службы каким-то образом чуть уменьшив свой рост относительно наименьшего допустимого для призыва. Так чисто случайно Кетле впервые указал полезное, пусть несколько элементарное, приложение своего статистического метода.

Обсуждая среднего человека, Кетле до того времени ограничивался вычислением средних и распределений лишь для некоторых физических характеристик. Теперь же он захотел расширить приложение нового понятия ко всем физическим характеристикам человека (и сформулировать основу *социальной физики*, как он её назвал¹⁰), а затем и ко всем моральным и умственным качествам (*моральная статистика*). Кроме того, он имел в виду приложить это понятие к коллективам любого размера от небольших групп до всего человечества и ожидал, что результаты окажутся такими же удовлетворительными для любого периода истории человечества.

Подобное же расширение исследований Кетле предложил в более ранних трудах, но полного обобщения не произошло до 1848 г. Кетле изложил суть этой книги на с. 16 – 17:

Существует общий закон, который управляет нашей вселенной [...]. Он придаёт каждому живому существу бесконечное разнообразие. [...] Наука долгое время неверно понимала этот закон, и до сих пор он оставался практически бесполезным. Я назову его законом случайных причин.

Все элементы организованных существ изменяются в ту и другую сторону от среднего значения, и [...] вариации, происходящие от случайных причин, отрегулированы столь

гармонично и точно, что мы можем заранее расположить их в пределах их границ численно и по порядку величины.

Часть этой книги посвящена доказательству закона случайных причин и для физического, и для морального и умственного человека, рассмотренного и самим по себе, и в группе.

[3.4.] Понятие причинности. Ясно, что для понимания среднего человека совершенно необходимо объяснить, как Кетле представлял себе случайные причины и закон. Он предположил, что каждое представленное им среднее было следствием действия постоянных причин, тогда как отклонения от среднего вызывались *пертурбациями* или случайными причинами. *Постоянными причинами* являлись те, которые действовали непрерывно и с той же самой интенсивностью и в том же направлении (1846, английский перевод 1849 г., с. 107). Среди них он назвал пол, возраст, профессию, географическую широту и экономические и религиозные институты. Иногда Кетле упоминал и *переменные* причины, действовавшие непрерывно с переменными энергией и интенсивностью (там же). Примером служили сезоны, но Кетле имел в виду все периодические явления вообще. Наконец, *случайные причины проявляются только беспорядочно и действуют безразлично в любом направлении* (там же)¹¹.

Кетле часто относил свободную волю человека к случайным причинам (хотя иногда утверждал, что она вообще не играет никакой роли), однако настаивал, что её действие заключено в очень тесных границах. Суть его теории заключалась в том, что, при наличии достаточных данных, распределённых во времени, форма и размах отклонений от среднего состояния, происходящих от случайных величин, могут быть с высокой точностью классифицированы заранее при помощи теории вероятностей независимых событий¹².

Понятие закона у Кетле зависело от того, рассуждал ли он о физических, моральных или вообще обо всех характеристиках человека. Так, в 1848 г. мы находим все три различных применений этого термина. В начале книги он сослался на тренд в серии средних, распределённых во времени как на закон (с. 11):

*Будь нам известен [средний] рост [человека] от одного столетия к другому, мы имели бы ряд размеров, который выразил бы закон развития человечества по росту*¹³.

Ниже (с. 86), представляя закон наклонности к преступлению, он применил тот же термин, чтобы обозначить законом схему корреляций:

Наклонность к преступлению весьма быстро возрастает к достижению совершеннолетия, достигает наибольшего значения, а затем убывает вплоть до самого конца жизни.

Этот закон представляется неизменным, изменчивым лишь в отношении величины и момента максимума. Напротив, закон наклонности к самоубийству возрастает *вплоть до самого преклонного возраста* (с. 88).

Здесь важно отметить, что при изучении человека Кетле пошёл намного дальше столь простых корреляций между двумя

переменными. Его книга (1835, т. 2) в особенности включает большое число таблиц с тремя и четырьмя переменными. В одном случае он (с. 91) представил таблицу, которая указывала соотношение среднего веса к возрасту, полу и занятиям. Аналогичные таблицы подразделяли преступность в различных группах по полу и уровню образования (с. 297), по возрасту и полу (с. 302) и по возрасту, полу и типу судов, рассматривавших преступления (с. 308). Это примечательное предвосхищение современных методов¹⁴ почти не было замечено современниками Кетле, и лишь недавно социологи заново обнаружили его и полностью исследовали возможности многомерного анализа. Поразительный разрыв в истории эмпирических социальных исследований конечно же заслуживает дальнейшего изучения и объяснения.

Оба указанных применения термина *закон* аналогичны друг другу, поскольку имеют какое-то отношение к корреляции, – в одном случае, между ростом и столетием, в другом, – между возрастом и проявлением какого-либо социального действия. Третий тип закона, закон случайных причин, совсем другого рода. Это просто утверждение о том, что каждая характеристика человека распределена нормально около среднего, и что чем больше число наблюдений, тем ближе окажется эмпирическое распределение к теоретическому. Короче, термин *закон* применяется или для указания тенденции в ряду определённых эмпирических данных, или как эмпирическое обобщение или утверждение (фактически как теория) о том, что определённый тип закономерностей существует во всех явлениях, связанных с человеком.

[3.5.] Измерения. Возможно, что одной из причин, почему Кетле не ограничился единым понятием *закон*, состояла в том, что типов мер, которые он применял для обоснования законов, было немного, притом не подходящих его цели. Обработывая данные лишь из доступных официальных статистических публикаций, он вынужден был наскоро придумывать новые и к тому же различные методы в случаях физических и моральных качеств и таким образом пришёл к различным типам законов. Несколько удивляет, что во всех исследованиях о человеке он фактически применял только три типа измерений. Он, правда, предложил и иные, но не применил их к своим данным и не попытался собрать материалы, для которых они были бы пригодны. Lazarsfeld (1961) исследовал типы мер, которые Кетле упомянул в 1869 г., в свете современных понятий о введении количественных мер.

Вначале Кетле определял эмпирические распределения некоторых характеристик человека данной группы и вычислял средние, которые объявлял характеристикой среднего человека этой группы. Он неоднократно утверждал, что аналогичные распределения и средние могут быть установлены для моральных и умственных характеристик, приводил некоторые грубые условные кривые, но никогда не проделывал подобных вычислений с использованием набора эмпирических наблюдений.

Второй тип измерений включал подсчёт числа определённых социальных действий, например, преступлений или женитьб, происшедших за ряд лет в определённых группах. Среднее из этих ежегодных подсчётов принималось за вероятное число подобных событий, которые произойдут в каждой группе в течение следующего года.

В качестве третьего типа измерений Кетле использовал коэффициенты, – относительные числа преступлений или иных событий в каждой возрастной группе. Эти результаты он рассматривал как соответствующие вероятности совершения преступлений в различных возрастах и называл их *наклонностями* (или *тенденциями*) к преступлению в каждом возрасте. В одном случае Кетле (1848, с. 93) предложил заменять *наклонности* *возможностями*, но затем вернулся к прежнему термину.

Снова и снова Кетле заявлял, что рассматривает группы, а не отдельных лиц, но во многих случаях он ясно применял установленную таким образом *наклонность к преступлению* для характеристики *каждого* члена данной возрастной группы.

[3.6.] Видимые и истинные наклонности. Понятие *наклонность* Кетле разработал, чтобы преодолеть встреченные им методологические и теоретические проблемы становления науки, которая исследовала бы все аспекты человека. При ограничении его физическими характеристиками он не столкнулся ни с какими трудностями. Метод индивидуальных измерений был очевиден, а кроме того были доступны обильные материалы, уже собранные различными агентствами. Но как только Кетле перешёл к моральным и умственным характеристикам, единственными данными оказались относительные числа (преступлений, самоубийств, женитьб) для различных совокупностей. Чтобы исследовать эти характеристики подобным же образом, потребовались бы измерения во времени каждого индивидуума в группе, но сводок подобных измерений нет и сегодня. Иногда, в предположительных разделах своих трудов, Кетле смог предлагать метод таких измерений. К примеру, для измерения продуктивности учёного он имел в виду подсчёт числа его публикаций¹⁵.

К сожалению, он так и не применил подобных идей в своих статистических работах, а довольствовался имевшимися данными и установлением критического различия между *видимыми* и *истинными наклонностями*. Первые вычисляются, как кратко указано выше, по официальным статистическим публикациям, необходимо было лишь знать количество действий (преступлений, самоубийств, женитьб) в соответствующих возрастах и общее население, опять же в каждом возрасте. Истинные же наклонности это те, которые служат причинами наблюдаемых закономерностей, и, как утверждал Кетле, их можно было устанавливать не по непосредственным наблюдениям, а только по их действию.

Вот как он (1848, с. 77) комментировал вычисление вероятности женитьб для некоторых горожан:

Можно считать, что эта вероятность определяет меру видимой склонности к женитьбе бельгийца в возрасте от 25 до 30 лет. Я намеренно говорю видимой склонности, чтобы не смешивать её с истинной, которая может быть совсем иной. Кто-то может всю жизнь обладать истинной склонностью жениться, но так и не женится, другой же, напротив, может при благоприятных обстоятельствах жениться, хоть и не имел к этому никакой склонности. Это отличие существенно.

Намёк на представление Кетле о соотношении тех и других склонностей можно найти в его повторных утверждениях о том, что *причины пропорциональны действиям*. Так, если видимые склонности вызваны подспудными истинными (там же, с. 78), то

Возможную ошибку, возникающую при замене видимой на истинную, можно вычислить при помощи теории вероятностей.

Это утверждение лишь формулирует проблему, но не решает её. Трудность у Кетле возникла потому, что он так и не подразделил чётко свою задачу на две её составляющие. Одна из них интересна и сегодня: как реальные данные относятся к скрытым склонностям? Вторая составляющая является методологической проблемой о возможности разработки системы измерений моральных и умственных характеристик, аналогичной тем, которые применяются к индивидуальным физическим характеристикам.

В теоретическом смысле Кетле не представлял себе, что его понятие *склонности* могло бы с равным успехом применяться и к физическим характеристикам. Можно, например, иметь в виду *склонность к полноте* в населении. В методологическом смысле Кетле не понял, что некоторые из его собственных методов измерения индивидуальной личности или интеллекта в точности соответствовали его *непосредственным* измерениям физических характеристик.

Особенно удивляет, что, неоднократно применяя статистику преступлений, он совершенно забыл, к примеру, об исследовании рецидивистов. Он мог бы таким образом найти количественные индивидуальные *меры* уголовного поведения, соответствующие измерениям тела или роста. Представляется вероятным, что эти упущения были вызваны не просто недостаточной интуицией, а по меньшей мере частично отсутствием данных. Только задним числом мы можем установить подобные пробелы, и нельзя отрицать заслуги Кетле, который впервые попытался исследовать то, что остаётся критическими проблемами в анализе эмпирических социальных явлений.

[4.] После Кетле

В 1855 г. Кетле перенёс инсульт, от которого полностью так и не оправился. Он вскоре возобновил свою работу, но новых идей уже не высказал. Его последующие публикации, хоть и многочисленны, в основном являлись сводками прежних очерков или новых исследований, обосновывающими более ранние идеи. И его сын Эрнст фактически взял на себя работу обсерватории¹⁶. Кетле умер 17 февраля 1874 г. и (Hankins 1908) *был похоронен с почестями, достойными лиц высокой социальной значимости.*

Его статуя, установленная по общенародной подписке, была торжественно открыта в Брюсселе в 1880 г.

Заботам Кетле о распределении характеристик человека было уготовано интересное будущее. Его основная идея состояла в том, что определённые социальные процессы (соответствовавшие его представлениям о взаимодействии причин) могут объяснить окончательное распределение некоторых доступных наблюдению данных. Эта идея была достаточно обоснована современными математическими моделями о распределении, скажем, доходов, слов или размеров городов.

Кетле, однако, был сосредоточен только на биномиальном и нормальном распределениях¹⁷, которые предполагали независимость изучаемых событий. Сегодня известны и многие другие распределения, основанные [соответствующие и т. д.] на более сложных процессах. В центре внимания, особенно в социологии, находятся *инфекционные события*, зависящие друг от друга. О других математических возможностях Кетле не знал, но его основная идея была не только верна, но вероятно непосредственно повлияла на авторов, подобных Пуассону и Лексису, которые начали расширять всё это поле¹⁸.

Вообще же современники Кетле сосредоточили своё внимание в основном на понятии среднего человека и на утверждении Кетле о том, что социальные явления воспроизводятся исключительно закономерно. По разным причинам эти идеи быстро стали предметом наиболее живых и распространённых споров статистиков, философов и социологов XIX в., большинство же остальных идей Кетле почти не были замечены.

Споры о *среднем человеке* оказались следствием предположения Кетле о том, что средние различных характеристик можно объединить, чтобы образовать единого образцового человека, который представлял бы *тип* группы, города, нации и даже всего человечества.

Характерным для ранней критики было рассуждение Курно (1843) по математической аналогии: средние сторон многих прямоугольных треугольников не образуют прямоугольного треугольника, и средние физических характеристик никак не будут соответствовать друг другу. Их объединение, вопреки Кетле, приведёт не к *типу человеческой красоты* или *физического совершенства*, а к чудовищу.

Настойчивые утверждения Кетле о том, что средний человек являлся лишь *вымышленным существом*, были восприняты как уклонение от сути дела, а его попытки непосредственно ответить Курно оказались неубедительными. Guilbaud (1952) обсудил замечание Курно, поскольку оно относится к современным статистическим понятиям, подобным *составлению*.

Кроме того, неспособность Кетле опровергнуть критику Курно воодушевили других авторов, которые опубликовали аналогичные нападки, а Bertillon (1876) сформулировал то, что обычно считается окончательным утверждением, по существу прикрывшим споры. Он критиковал не только понятие Кетле об

объединении средних физических характеристик, но и его идеи о моральных и умственных качествах.

Конечно же, заявил Бертильон, Кетле ошибался, полагая, что его средний человек представляет идеал моральных достоинств или умственное совершенство. Напротив, подобное существо воплотит посредственность или (1876, с. 311) [страница неверна] *тип вульгарности*. После этого среднего человека посчитали понятием, недостойным серьёзного внимания, хотя время от времени пытались оживить его¹⁹. Последнюю попытку сделал Fréchet (1955), который предложил *восстановить в правах* среднего человека, считая его понятием *типичного человека*, – *определённого* человека из группы, характерные черты которого в целом находятся *ближе всего* к среднему. Он сумел избежать большинства критических замечаний по поводу среднего человека.

Второй, и даже более распространённый спор сосредоточился на том, какие выводы должны быть сделаны из поразительных закономерностей, существование которых Кетле доказал в своих исследованиях социальных явлений²⁰. Вот утверждение Кетле (1835/1865, т. 1, с. 10) о том, что

общество включает в себе зародыши всех совершаемых преступлений, равно как и средства, необходимые для их осуществления. Это оно в некотором смысле подготавливает преступления, а виновные являются лишь его исполнителями.

Но означало ли это, что человек лишён свободной воли? Соответствующие философские рассуждения Кетле действительно допускали подобное истолкование. В результате начался пылкий и длительный спор между теми, кто поддерживал *детерминированное* объяснение социальных закономерностей по Кетле, и теми, кто заявлял, что лишь начиная исследование с отдельного лица можно объяснить поведение человека. Свободную волю, они утверждали, следует считать тем основным, которое определяет действие, а не практически пренебрегаемой *случайной величиной*. Одним из побочных результатов этого спора было рождение *немецкой школы* моральных статистиков под руководством Дробиша, одного из самых страстных противников Кетле (Drobisch 1867).

Споры окончились не опровержением идей Кетле, а их обсуждением вплоть до забвения. По мнению Fahlbeck (1900), самый значительный результат споров состоял в том, что до конца XIX в. статистики были настолько вовлечены в обсуждение последствий предложений Кетле, что мало озаботились эмпирически подтвердить суть и охват обнаруженных Кетле закономерностей. Подробное обсуждение спора по поводу детерминизма Кетле и понятия среднего человека см. Lottin (1912, с. 413 – 458).

Лишь в последние годы социологические труды Кетле начали восприниматься должным образом. Его убеждение в том, что научное изучение социальной жизни основано на количественных и математических методах, предвосхитило руководящий принцип современных социологических исследований²¹. Некоторые

конкретные методы, которые он применял и рекомендовал, например, замену одноразового наблюдения населения повторными наблюдениями отдельных лиц, и его ранние попытки многомерного анализа столь же важны сегодня, сколь они были новы в его время.

То же можно сказать про его попытки преобразовать статистику, перейти от простой канцелярской задачи сбора важных фактов о государстве (отсюда и сам термин *статистика*) к точному методу наблюдения, измерения, табулирования и исправления результатов, что послужило бы лесами для построения его науки моральной статистики. Уже поэтому трудно оспорить описание книги (1835) у Сартона (Sarton 1935/1962, с. 229) как *одной из величайших книг XIX в.*²², или, если на то пошло, его мнения о том, что Кетле, а не Конт был *основателем социологии*.

Краткие сведения об упомянутых лицах

Dandelin, Germain P., Данделен, Жерминаль П., 1794 – 1887, математик

Shattuck, Lemuel, 1793 – 1859, политик, историк, издатель. Один из основателей Американской статистической ассоциации

Villermé, L. R., Виллерме Л. Р., 1782 – 1863, врач, экономист, статистик

Wheatstone, Charles, Уитстон, Чарлз, 1802 – 1875, физик, изобретатель

Примечания

1. Кривая третьего порядка, см. *Ciel et Terre*, t. 44, 1928, p. 60.
2. Неясно, как здесь появилась статистика.
3. Астроном А. Бювар (1767 – 1843), неутомимый вычислитель. Ниже Лаплас почему-то упомянут после Пуассона, а современная теория вероятностей была всё-таки заложена лишь в начале XX в.
4. Ср. [vii].
5. Следует различать сессию Международного статистического конгресса и Международный статистический конгресс вообще. Во втором случае мы применяем заглавное к: Конгресс.
6. Со времён Кетле сфера моральной статистики невообразимо расширилась, но всё же никак нельзя отождествлять её с социологией.
7. Необходимость больших чисел для вывода надёжных следствий была ясна уже Якобу Бернулли. Впрочем, математическая статистика вынуждена обходиться и малыми числами.
8. Примерно в 1819 г. на необходимость основывать статистику математикой (т. е. теорией вероятностей) указывал Фурье в письме Кетле (Sheynin 1986, с. 306). Косвенно о том же писал Пуассон (там же, с. 307), а в 1834 – 1835 гг. он уточнил это утверждение (Шейнин 2013, § 9.9).
9. Закон случайных причин остался у Кетле совершенно неопределённым (Шейнин 2013, § 11.5).
10. Термин *социальная физика* ввёл Конт. Он обвинил Кетле в том, что тот воспользовался указанным термином, притом в другом смысле, см. Sheynin (1986, с. 295).
11. Кетле ничего не сказал о направлении действия переменных причин, а случайные причины понимал в ограниченном смысле. Это лишнее свидетельство о том, что он как-то совсем забыл о своём математическом образовании. Вообще же во многих его сочинениях чувствуется безалаберность.

12. Упоминание теории вероятностей независимых событий непонятно: в то время иной теории не было.

13. Из текста не следовало, что рост человека монотонно изменялся во времени.

14. Многомерные таблицы можно отыскать в работах по медицинской и звёздной статистике ещё до Кетле, например у Гершеля в 1806 г., см. Sheynin (1984a, с. 171).

15. Измерение продуктивности учёного по количеству его публикаций крайне ненадёжно, достаточно сослаться на Бейеса.

16. Никаких астрономических открытий Кетле так и не сделал, хотя производил восхитившие Фарадея наблюдения атмосферного электричества (Sheynin 1984b, с. 78 – 79). См. там же о заслугах Кетле в метеорологии. Наконец, известно [v, § 3], что Кетле измерял интенсивность геомагнетизма.

17. Это неверно, см. Шейнин (2013, § 11.5).

18. Эта идея представляется весьма простой, а потому вряд ли заслуживающей особого внимания. Влияние Кетле на Пуассона сомнительно, а вот обратное влияние несомненно. По поводу Лексиса мы можем только сказать, что он обозначил свой знаменитый коэффициент дисперсии буквой Q .

19. И всё-таки средний человек это средний производитель и средний потребитель.

20. Фактически Кетле этого не доказал, что было замечено еще в 1876 г. (Шейнин 2013, § 11.5).

21. См. Прим. 8. По поводу преобразования статистики и т. д. (см. чуть ниже) следовало бы добавить, что уже Граунт обработал собранные им же, казалось бы ненужные (и недостоверные) факты.

22. Эта оценка непомерно завышена. Можно сказать (Фрейденталь), что до Кетле статистической науки не было, но следует упомянуть и Кнаппа, который заметил, что дух Кетле богат идеями, но не является методичным или философским, см. Шейнин (2013, § 11.5).

Библиография

1826, *Astronomie élémentaire*. Bruxelles, 1834. Третье издание.

1827a, *Astronomie populaire*. Bruxelles, 1832.

1827b, *Positions de physique*. Bruxelles, 1834.

1828, *Instructions populaires sur le calcul des probabilités*. Bruxelles.

1831a, *Recherches sur la loi de croissance de l'homme*. Bruxelles. Также в *Nouv. Mém. Acad. Roy. Sci., Lettres, Beaux-Arts de Belgique*, t. 7.

1831b, *Recherches sur le penchant au crime aux différens âges*. Bruxelles, 1833.

Также в *Nouv. Mém. Acad. Roy. Sci., Lettres, Beaux-Arts de Belgique*, t. 7.

1833a, *Recherches sur le poids de l'homme aux différens âges*. Bruxelles. Также в *Nouv. Mém. Acad. Roy. Sci., Lettres, Beaux-Arts de Belgique*, t. 7.

1833b, Lettre ... sur la possibilité de mesurer l'influence des causes ... *Annales d'hygiène publique et de médecine légale*, t. 9, p. 309.

1835, *Sur l'homme*, tt. 1 – 2. Bruxelles. *Человек и развитие его способностей*, тт. 1 – 2. СПб, 1865.

1844, Sur l'appréciation des documents statistiques ... *Bull. Comm. Centr. Statistique Belgium*, t. 2, pp. 205 – 286.

1846, *Lettres ... sur la théorie des probabilités*. Bruxelles.

1848, *Du système social*. Paris. *Социальная система*. СПб, 1866.

1869, *Physique sociale*, tt. 1 – 2. Bruxelles. *Социальная физика*, тт. 1 – 2. Киев, 1911 – 1913.

1870, Des lois concernant le développement de l'homme. *Bull. Acad. Roy. Sci., Lettres, Beaux-Arts de Belgique*, t. 29, pp. 669 – 680.

Другие авторы

Шейнин О. Б., Sheynin O. B. (1984a), On the history of the statistical method in astronomy. *Arch. Hist. Ex. Sci.*, vol. 29, pp. 151 – 199.

--- (1984b), On the history of the statistical method in meteorology. Там же, vol. 31, pp. 53 – 93.

--- (1986), Quetelet as a statistician. Там же, vol. 36, pp. 286 – 325.

--- (2013), *Теория вероятностей. Исторический очерк*. Берлин. Также Google, Oscar Sheynin, Download area

Bertillon Ad. (1876), La théorie des moyennes en statistique. *J. Soc. Stat. Paris*, t. 17, pp. 265 – 271, 286 – 308.

Cournot A. A., Курно О. (1843, франц.), *Основы теории шансов и вероятностей*. М., 1970.

Drobisch M. W. (1867), *Die moralische Statistik etc.* Leipzig.

Fahlbeck P. E. (1900), La régularité dans les choses humaines. *J. Soc. Stat. Paris*, t. 41, pp. 188 – 201.

Fourier J. B. J. (1826), Sur les résultats moyens déduits d'un grand nombre d'observations. *Œuvres*, t. 2. Paris, 1890, pp. 525 – 545.

Fréchet M. (1955), Rehabilitation de la notion statistique de l'homme moyenne. В книге автора *Les mathématiques et le concret*. Paris, pp. 310 – 341.

Guerry A. M. (1832), Statistique comparée de l'état l'instruction et du nombre des crimes. *Rev. Enc.*, t. 55, pp. 414 – 424.

Guilbaud G. Th. (1952, франц.), Theories of the general interest ... В книге P. F. Lazarsfeld & N. W. Henry, редакторы (1966), *Readings in Mathematical Social Science*. Chicago, pp. 271 – 292.

Hankins F. H. (1908), *Adolphe Quetelet As Statistician*. New York.

[**Herschel J.**] (1850), Обзор книги Quetelet (1846). *Edinb. Rev.*, vol. 92, pp. 1 – 57.

Kendall M. G., Plackett R.L. (1977), *Studies in History of Statistics and Probability*, vol. 2. London.

Laplace P. S. (1812), *Théorie analytique des probabilités*. *Œuvr. Compl.*, t. 7. Paris, 1886.

Lazarsfeld P. F. (1961), Notes on the history of quantification in sociology. *Isis*, vol. 52, pp. 277 – 333. Перепечатка: Kendall & Plackett (1977, pp. 213 – 269).

Lottin J. (1912), *Quetelet, statisticien et sociologue*. Paris.

Poisson S.-D. (1837), *Recherches sur la probabilité des jugements ...* Paris, 2003. *Исследования о вероятности приговоров ...* Берлин. Также Google, Oscar Sheynin, Download area.

Reichesberg N. (1896), Der berühmte Statistiker Adolphe Quetelet. *Z. f. schweiz. Statistik*, Bd. 32, pp. 418 – 460.

Sarton G. (1935), Quetelet. В книге автора *On the History of Science*. Cambridge, Mass., 1962, pp. 229 – 242.

Schoen H. H. (1938), Prince Albert and the application of statistics to the problems of government. *Osiris*, vol. 5, pp. 276 – 318.

Wellens-De Donder Liliane (1964), La correspondance d'Adolphe Quetelet. *Archives et Bibl. de Belgique*, t. 35, pp. 49 – 66.

VII

[Чарлз Беббидж]

Заметка о происхождении Статистического общества

[Charles Babbage,] Note respecting the origin of the Statistical Society.
J. Roy. Stat. Soc., vol. A124, 1961, p. 546

В 1828 г. в Берлине я присутствовал на торжественном заседании немецких философов под председательством Александра Гумбольдта. Я был единственным присутствовавшим английским членом, а [их] ассоциация не была известна английской общественности, и поэтому я написал отчёт о заседании. Он появился в *Journal of Science* [Эдинбург], будучи возможно первой публикацией, которая обратила внимание нашей общественности на значимость подобных ассоциаций.

Через несколько лет была образована Британская ассоциация продвижения науки. Она добилась существенного успеха, но вскоре я заметил недостаток в её структуре, который следовало устранить. Её секции ограничивались естественными и чистыми науками, и поэтому для весьма влиятельных лиц, занимавших высшие положения в правительстве, для членов парламента и других, связанных с общественной жизнью, не было такой секции, в которой их знания могли бы стать общедоступными.

Я безуспешно отыскивал какое-нибудь средство для исправления этого положения, но затем немедленно воспользовался представившимся счастливым случаем. В период заседаний Британской ассоц. в Кембридже в 18 ... г. [пропуск в тексте] я зашёл к мистеру Джонсу, главе городского отдела налогов (Comm[issioner] of Tythes). Он сообщил мне, что бельгийское правительство послало на конференцию нашего общего друга, Кетле, с ценнейшими статистическими материалами, но что, к великому сожалению, сама Ассоц. не может воспользоваться ими.

Джонс попросил меня зайти к нему вечером, чтобы встретиться с Кетле, Сэром Чарлзом Лемоном и некоторыми другими, заинтересованными в этом деле. Я с большой охотой согласился и ушёл, но уже через несколько минут мне пришло в голову, что предложение Джонса предоставляло мне ту самую возможность, которую я отыскивал. Я сейчас же вернулся к нему и предложил каждому из нас пригласить всех тех, чей интерес в статистических исследованиях был нам известен, чтобы объединиться и образовать статистическую секцию, притом на следующем же общем заседании Ассоц. я объявил бы о существовании новой секции и попросил бы простить нас за внеуставные действия. Я хорошо представлял себе, что, будь наша просьба подана предварительно, а не задним числом, она оказалась бы безуспешной.

И действительно, общее собрание отчитало нас, утвердило существование новой секции, хоть и ограничило нас довольно требовательными правилами, и заставило нас обещать никогда

впредь не поступать подобным образом (с чем мы охотно согласились). Ещё до окончания этого собрания статистическая секция стала самой многочисленной и самой популярной, поскольку её исследования оказались [бы] интересными во многих отношениях.

Вскоре, однако, стало ясно, что для достижения статистических целей необходима более постоянная организация и её более частые встречи. Поэтому секция решила, что в Лондоне следовало создать Статистическое общество, и меня попросили заняться его оформлением. Затем было созвано открытое заседание под председательством Лорда Ландоуна, и Лондонское статистическое общество родилось на нём.

Примечание редакции. В марте 1961 г. Королевское статистическое общество получило фотокопию этой заметки и её французского перевода, который выполнил Кетле, от Президента Биометрического общества.

VIII

Клаус-Петер Хейсс

Вильгельм Лексис

Klaus-Peter Heiss, Lexis, Wilhelm. IES, pp. 507 – 512

Вильгельм Лексис (1837 – 1914), немецкий статистик и экономист, внёс существенный вклад в статистику и её приложения, особенно в исследование населения и экономических временных рядов. Как математик, Лексис очень скептически относился к состоянию математической экономики своего времени. Критика некоторых современных ему работ по этой дисциплине привела его к основополагающим соображениям об экономических явлениях и их взаимосвязи.

Лексис родился в Эшвайлере возле Ахена. Он изучал многие дисциплины, и круг его интересов включал юриспруденцию, естественные науки и математику. Он окончил Боннский университет в 1859 г., написав диссертацию по аналитической механике, но получил степень и по математике. Некоторое время Лексис работал исследователем в химических лабораториях Бунзена в Гейдельберге. В 1861 г. он отправился в Париж изучать социальные науки, в результате чего появилась его первая большая публикация (1870), трактат об экспортной политике Франции.

В нём проявились черты, которые характеризовали его позднейшие экономические работы: скептицизм по поводу *чистой экономики* и применения будто бы описательных математических моделей, не имевших отношения к экономической реальности. Уже в этой ранней работе Лексис настаивал на том, что экономическая теория должна быть основана на количественных экономических данных. По его мнению, подробный анализ общего экономического равновесия, основная цель которого состояла в установлении уравнений по числу неизвестных, не способствовал пониманию или решению экономических проблем и потому не должен был восприниматься слишком серьёзно. Он оказался одним из математически образованных исследователей экономики, которые во второй половине XIX в. отошли от экономики. Другим, более известным, оказался Макс Планк. Попытавшись прочесть работу Маршалла, он отбросил её в сторону и навсегда изменил направление своих исследований (Schumpeter 1954, с. 957 – 958).

В 1872 г. Лексис был назначен в университет г. Страсбург. Там он составил введение в теорию населения (1875), а в 1874 г. перешёл в Дерпт [Тарту] профессором географии, этнологии и статистики, затем, в 1876 г., – во Фрайбург в качестве профессора экономики. Именно в то время Лексис написал свои основные работы по статистике (1876; 1877; 1879а) и экономике (1879b; 1881; 1882а; 1882b). После периода работы в Бреслау в 1884 – 1887 гг. его назначили профессором политической науки в Гёттингенский университет.

В свои позднейшие годы работы Лексиса оказались в высшей степени разнообразны. Он был редактором и активным автором главной немецкой экономической энциклопедии, *Handwörterbuch der Staatswissenschaft* и директором первого института актуарных наук в Германии. В 1890-е годы он написал или отредактировал несколько работ об образовании и особо по университетской системе (1893; 1901; 1902; 1904a; 1904b)¹. Его исследования населения, экономики и статистики в последующее десятилетие (1903; 1906a; 1906b; 1908; 1914) объединяют и уточняют некоторые прежние доводы. Он умер в Гёттингене в первые дни первой мировой войны.

Статистика. Главные труды Лексиса были посвящены статистике. Первыми из них были проблемы, с которыми он столкнулся при исследовании населения (1875; 1879b; 1891; 1903), социологии (1877) и экономики (1870; 1879a; 1908; 1914). В связи с изучением массовых социальных явлений (1877) и временными рядами, которые встретились в нескольких социальных науках (1879b), он подступил к проблемам статистической однородности, которыми до того времени видимо пренебрегали, хотя Bortkiewicz (1918) указал, что аналогичные идеи примерно в то же время разработал Дормуа (1874; 1878)². Другими возможными предшественниками были Bienaumé (1855), Cournot (1843) и R. Campbell (1859). Лексис признал, что Дормуа предвосхитил некоторые его идеи, но именно он более или менее независимо придумал новое направление исследованию статистических рядов и сдвинул основной интерес статистиков с чисто математического подхода, с которым был связан Лаплас, к эмпирическому или индуктивному подходу (Keynes 1921, с. 392 и след.). Он также начал исследовать анализ рассеяния и дисперсионный анализ в попытке разработать статистики для оценивания качественных изменений населения во времени (Keynes 1921; Polya 1919).

Лексис (1877; 1879a; 1879b) показал, что в генеральной совокупности массовых социальных явлений условия статистической однородности (случай случайных выборок из устойчивого распределения³) если и выполняются, то редко. Вероятностные структуры вполне могут отличаться друг от друга в пределах выборки ввиду особых обстоятельств, относящихся к рассеиванию в пространстве или времени либо по иным причинам. Генеральные совокупности, отдельные выборки из которой отобраны из возможно отличающихся друг от друга совокупностей, называются по его имени, см., например, Herdan (1966).

В некоторой мере работа Лексиса была реакцией на не критические предположения однородности, вводимые в прежних статистических изысканиях, например, у Кетле. Keynes (1921) указал, что и Кетле, и другие авторы, имея лишь немного оснований для этого, просто утверждали, что различные социальные статистики вероятно устойчивы из года в год⁴.

Лексис в основном интересовался рассеиванием наблюдений около их локальных средних и поведением средних и рассеивания во времени. Он придумал статистики для изменения степени

устойчивости подобных временных рядов и указал в качестве полезного обобщения, что они либо подтвердят статистическую однородность, указывая на существование *бернуллиевого* ряда, либо при уклонении от него, *пуассонова* или *лексисова* ряда. Эту терминологию предложил С. V. L. Charlier (A. Fisher 1915, с. 117).

Лексис рассматривал только дихотомические случайные величины (мужчины – женщины, живые – мёртвые и т. д.), но его доводы в равной мере относятся к числовым случайным величинам в общем смысле (Polya 1919). Ниже мы приводим его идеи в обобщённой форме.

Пусть x_{ij} , $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$ являются набором n выборок по m наблюдений в каждой, \bar{x}_i – среднее из x_{ij} в выборке i и \bar{x} – среднее по всем выборкам. Аналогично, пусть $a_{ij} = Ex_{ij}$, $\bar{a}_i = E\bar{x}_i$, и $\bar{a} = E\bar{x}$.

Лексис рассматривал квадратичные формы, которые измеряли рассеивание трёх различных видов,

$$s_w^2 = \frac{1}{nm} \sum_i \sum_j (x_{ij} - \bar{x}_i)^2, \quad s_b^2 = \frac{1}{n} \sum_i (\bar{x}_i - \bar{x})^2,$$

$$s^2 = \frac{1}{nm} \sum_i \sum_j (x_{ij} - \bar{x})^2, \quad s^2 = s_b^2 + s_w^2. \quad (1a, b, c, d)$$

Ранги (степени свободы) левых частей были равны

$$r_b = n - 1, \quad r_w = n(m - 1), \quad r = r_b + r_w = nm - 1.$$

Если эти n выборок отобраны случайно из одной и той же совокупности, ожидаемое значение каждого наблюдения будет равно

$$a_{ij} = \bar{a}_i = \bar{a}$$

и статистическая однородность имеет место. Примерами могут служить измерения расстояний или nm извлечений шариков из урны с возвращением. Указанные квадратичные формы, умноженные на подходящие константы, будут иметь одни и те же ожидания

$$E\left[\frac{n}{r} s^2\right] = E\left[\frac{n}{r_b} s_b^2\right] = E\left[\frac{n}{r_w} s_w^2\right]$$

и их общее значение окажется равным $1/m$ доли дисперсии совокупности. Набор выборок, отобранных при этих условиях, известен как ряд Бернулли, см., например, A. Fisher (1915).

Может, однако, случиться, что статистическая однородность существует в выборках ($a_{i1} = a_{i2} = \dots = a_{im}$), но не между ними, т. е. что n выборок относятся к n различным совокупностям. В этом случае порождается лексисов ряд (сверхнормальный в его терминологии). Подобный случай может произойти, например,

если наборы шариков (выборки) извлечены из различных урн. Другие примеры таких рядов проясняют значение лексисовых рядов. Так, m наблюдений, сделанных в момент или период t_0 , m других – в момент t_1 , и т. д. вплоть до t_n образуют *временной ряд* mn наблюдений, причём i -я выборка, соответствующая периоду t_i , вполне может относиться к единой совокупности, однако различные выборки происходят от различных совокупностей.

Аналогично, социальные или экономические выборки по m наблюдений, отобранных из n различных географических районов или наций, вероятно происходят от различных статистических совокупностей, хотя в каждом районе (каждой нации) все m наблюдений в выборке относятся к одному и тому же распределению (*межрегиональный* или *международный ряд*).

Короче говоря, если полное рассеивание обусловлено не только случайными изменениями в ту и другую сторону от константы, но и трендами и другими систематическими факторами, которые изменяются от выборки к выборке, то возникают лексисовы ряды. Их подробное и элементарное рассмотрение см. Polya (1919). В таких случаях можно считать, что дисперсия между выборками окажется более значительной составной частью полной дисперсии всех mn наблюдений, чем дисперсия внутри выборок и что ожидаемое значение наблюдения будет равно среднему выборочному, отличному от среднего из всех наблюдений. Далее, хотя $a_{ij} = \bar{a}_i$, по крайней мере для некоторых i $\bar{a}_i \neq \bar{a}$ и

$$E\left[\frac{n}{r_w} s_w^2\right] < E\left[\frac{n}{r} s^2\right] < E\left[\frac{n}{r_b} s_b^2\right],$$

так что статистической однородности не существует.

Гораздо менее реалистичным, но формально дополняющим ряды Бернулли и Лексиса, является ряд Пуассона (поднормальный ряд по терминологии Лексиса). Модель Пуассона рассматривалась как порождающая более значительные вариации внутри выборок, а не между ними. В этом случае j -е наблюдение в каждой выборке отобрано из своей j -й совокупности и $a_{1j} = a_{2j} = \dots = a_{nj}$, но при фиксированном i не все величины a_{ij} совпадают. Поэтому $\bar{a}_i = \bar{a}$ и между выборками нет изменчивости, зависящей от величин a_i , так что

$$E\left[\frac{n}{r_b} s_b^2\right] < E\left[\frac{n}{r} s^2\right] < E\left[\frac{n}{r_w} s_w^2\right].$$

Лексис предложил статистику, основанную на указанных квадратических формах, для описания степени, в которой данный ряд однороден, сверх- или поднормален. Это лексисово отношение равно

$$L = \frac{s_b^2/r_b}{s/r}, \quad (2)$$

монотонно возрастающей функции другой статистики⁵, $s_b^2/r_b \div s_w^2/r_w$, которая также может быть использована.

Chuprov (1922) показал, что при статистической однородности

$EL = 1$, и дисперсия L примерно равна $2/(n - 1)$.

Дальнейшие исследования приводят к критерию значимости. Прибавляя и вычитая стандартную ошибку отношения L к (из) его ожидаемого значения, мы получим

$$1 \pm \sqrt{2/(n-1)}.$$

Если значение L оказывается внутри этих границ, мы можем заключить при уровне значимости примерно равном $2/3$, что статистическая масса однородна. Если L значимо превышает 1, можно считать, что ряд был извлечён из лексисовой совокупности, и из пуассоновой, если L значимо меньше 1.

Позднее многие авторы указали на отношение и связи, существующие между лексисовым рядом и L с дисперсионным анализом. Фишер подробно исследовал формальную связь между L и статистикой χ^2 Пирсона и показал, что в случае классификации $2 \times n$ $\chi^2 = nL$.

Между статистиками L и F существует непосредственное отношение, и можно сказать, что Лексис предвосхитил её, см. Coolidge (1921), Rietz (1932), Geiringer (1942a; 1942b), Gini (1956), Herdan (1966). При использовании статистики L (2) дисперсия между выборками сравнивается с дисперсией всех nm наблюдений, а при вычислении статистики F

$$F = \frac{s_b^2/r_b}{s_w^2/r_w}$$

первая дисперсия сравнивается с дисперсией внутри выборок. Кроме того, L больше, равно или меньше единицы одновременно с F , что основано на указанном выше равенстве (1d).

R. A. Fisher (1925, с. 97) и Cochran (1934, с. 178) установили асимптотическое распределение статистики F при $m \rightarrow \infty$, см. также обобщение (Madow 1940). Впрочем, то же распределение вывел Гельмерт (1876), который применил метод характеристических функций.

Борткевич расширил применение теории дисперсии Лексиса, а Чупров (1922) обобщил и самым подробным образом исследовал её. Влияние Лексиса испытали также J. von Kries, H. Westergaard и Эджуорт, – единственный в то время учёный из англоязычных стран, хорошо знакомый со статистическими исследованиями на континенте Европы (Keynes 1921).

Экономика. [...]

Примечания

1. Уже не только в 1890-е годы.
2. Изложение здесь беспомощно. Сослаться следовало на статью Лексиса (1879а), а несколько ниже указание на сдвиг статистиков с чисто математического подхода неверен. В то время статистики не владели математикой и чурались её; короче, такого похода у них не могло быть. Наконец, о Дормуа сообщил сам Лексис (Чупров 1909/1959, с. 236).
3. Устойчивое распределение понимается здесь только как почти неизменное.
4. Резкую критику Кетле по указанному поводу задолго до Кейнса опубликовал Rehnisch (1876).
5. Эта другая статистика (*F*) появляется ниже.

Библиография

В. Лексис

- 1870, *Die französischen Ausfuhrprämien etc.* Bonn.
- 1875, *Einleitung in die Theorie der Bevölkerungsstatistik.* Straßburg.
- 1876, Das Geschlechtsverhältnis der Geborenen und die Wahrscheinlichkeitsrechnung. *Jahrbücher f. Nationalökonomie u. Statistik*, Bd. 27, pp. 209 – 245.
- 1877, *Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft.* Freiburg i/B.
- 1879а нем., О теории стабильности статистических рядов. В книге Четвериков (1968, с. 5 – 38).
- 1879b, Gewerksvereine und Unternehmenverbände in Frankreich. *Schriften Verein f. Socialpolitik*, Bd. 17, pp. 1 – 280.
- 1881, *Erörterungen über die Währungsfrage.* Leipzig.
- 1882a, Die Volkswirtschaftliche Konsumtion. *Handbuch der politischen Ökonomie*, 3-е издание, Bd. 1. Tübingen, 1890, pp. 685 – 722.
- 1882b, Handel. Там же, Bd. 2, pp. 811 – 938.
- 1886, Über die Wahrscheinlichkeitsrechnung und deren Anwendung auf die Statistik. *Jahrbücher f. Nationalökonomie u. Statistik*, Bd. 47, pp. 433 – 450.
- 1891, Bevölkerungswechsel. Allgemeine Theorie ... *Handbuch der Staatswissenschaften*, Bd. 2. Jena, pp. 456 – 463.
- 1893, *Die deutschen Universitäten*, Bde 1 – 2. Berlin.
- 1895a, The present monetary situation. В книге *Der gegenwärtige Stand der Währungsfrage. Econ. Studies Amer. Econ. Assoc.*, vol. 1, No. 4, 1896.
- 1895b, Grenznutzen. *Handbuch der Staatswissenschaften*, Suppl. Bd. 1. Jena, pp. 422 – 432.
- 1901, *Die neuen französischen Universitäten.* München.
- 1902, редактор, *Die Reform des höheren Schulwesens in Preußen.* Halle.
- 1903, *Abhandlungen zur Theorie der Bevölkerungs- und Moralstatistik.* Jena.
- 1904a, редактор, *Das Unterrichtswesen im Deutschen Reich*, Bde 1 – 4. Berlin.
- 1904b, *A General View of the History and Organisation of Public Education in the German Empire.* Berlin.
- 1906a, Das Wesen der Kultur. В книге *Die allgemeinen Grundlagen der Kultur der Gegenwart.* Berlin, pp. 1 – 53.
- 1906b, *Das Handelswesen*, Bde 1 – 2. Berlin.
- 1908, Systematisierung, Richtungen und Methoden der Volkswirtschaftslehre. В книге *Die Entwicklung der deutschen Volkswirtschaftslehre im 19. Jahrhundert.* Leipzig, Bd. 1, pp. 1 – 45.
- 1910, Allgemeine Volkswirtschaftslehre. В книге *Die Kultur der Gegenwart*, 3-е изд., Bd. 2, pt. 10, § 1. Berlin – Leipzig.
- 1914, *Das Kredit- und Bankwesen.* Berlin, 1929.

Другие авторы

- Четвериков Н. С., составитель и переводчик (1968), *О теории дисперсии.* М.
- Чупров А. А. (1909), *Очерки по теории статистики.* М., 1959.
- (1922, нем.), Можно ли на основании эмпирических данных доказать, что устойчивость ряда нормальна? В книге автора *Вопросы статистики.* М., 1960, с. 239 – 258.

- Шейнин О. Б.** (2013), *Теория вероятностей. Исторический обзор*. Берлин. Также Google, Oscar Sheynin, Download area
- Bauer R. K., Bayur P. K.**, нем. (1955), Теория дисперсии Лексиса в её отношениях к новым течениям статистической методологии, в особенности к анализу рассеяния. В книге Четвериков (1968, с. 225 – 238).
- Bienaymé I. J.** (1855), Sur un principe que M. Poisson avait cru découvrir etc. *J. Soc. Statistique Paris*, t. 17, 1876, pp. 199 – 204.
- Bortkiewicz L. von** (1918), Der mittlere Fehler des zum Quadrat erhobenen Divergenzkoeffizienten. *Deutsche Mathematischer-Vereinigung Jahresbericht*, Bd. 27, pp. 71 – 126.
- (1930), Lexis und Dormoy. *Nordic Statistical J.*, vol. 2, pp. 37 – 54.
- Campbell R.** (1859), On the probability of uniformity in statistical tables. *Lond., Edinb. and Dublin Phil. Mag.*, vol. 18, pp. 359 – 368.
- Cochran W. G.** (1934), The distribution of quadratic forms in a normal system. *Proc. Camb. Phil. Soc.*, vol. 30, pp. 178 – 191.
- Coolidge J. L.** (1921), The dispersion of observations. *Bull. Amer. Math. Soc.*, vol. 27, pp. 439 – 442.
- Cournot A. A., Курно О.**, франц. (1843), *Основы теории шансов и вероятностей*. М., 1970.
- Dormoy E.** (1874), *Théorie mathématique des paris de courses*. Paris. Также в *J. des actuaires français*, t. 3.
- (1878), *Théorie mathématique des assurances sur la vie*, tt. 1 – 2. Paris.
- Edgeworth F. Y.** (1885), Methods of statistics. *Jubilee Volume, Roy. Stat. Soc.*, pp. 181 – 217; *Writings in Probability, Statistics and Economics*, vol. 2, pp. 24 – 60. Cheltenham, 1996.
- Fisher Arne** (1915), *Mathematical Theory of Probabilities* etc., vol. 1. London, 1922.
- Fisher R. A.** (1925), Applications of Student's distribution. *Metron*, t. 5, No. 3, pp. 90 – 104.
- (1928), On a distribution yielding the error functions etc. *Proc. Second Intern. Congr. of Mathematicians Toronto 1924*, Toronto, vol. 2, pp. 805 – 813.
- Geiringer H.** (1942a), New explanation of non-normal dispersion in the Lexis theory. *Econometrica*, vol. 10, pp. 53 – 60.
- (1942b), Observations on analysis of variance theory. *Annals Math. Stat.*, vol. 13, pp. 350 – 369.
- Gini C.** (1956), Généralisations et applications de la théorie de la dispersion. *Metron*, t. 18, No. 1/2, pp. 1 – 75.
- Helmert F. R.** (1876), Über die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler etc. *Z. f. Math. u. Phys.*, Bd. 21, pp. 192 – 218.
- Herdan G.** (1966), *Advanced Theory of Language As Choice and Chance*. New York.
- Keynes J. M.** (1921), *Treatise on Probability. Coll. Writings*, vol. 8. London, 1988.
- Kries J.** (1886), *Die Principien der Wahrscheinlichkeitsrechnung*. Tübingen, 1927.
- Madow W. G.** (1940), Limiting distributions of quadratic and bilinear forms. *Annals Math. Stat.*, vol. 11, pp. 125 – 146.
- Polya G.** (1919), Anschauliche und elementare Darstellung der Lexisschen Dispersionstheorie. *Z. f. schweiz. Statistik u. Volkswirtschaft*, Bd. 55, pp. 121 – 140.
- Rehnsch E.** (1876), Zur Orientierung über die Untersuchungen und Ergebnisse der Moralstatistik. *Z. Philos. u. philos. Kritik*, Bd. 69, pp. 43 – 115.
- Rietz H. L.** (1932), On the Lexis theory and the analysis of variance. *Bull. Amer. Math. Soc.*, vol. 38, pp. 731 – 735.
- Schumpeter J. A.** (1954), *History of Economic Analysis*. New York, 1960.

IX

Элен Уокер

Карл Пирсон

Helen M. Walker, Pearson, Karl. IES, pp. 691 – 698

[1. Введение]

[1.1. Биография.] Карл Пирсон, *основатель научной статистики*, родился в Лондоне в 1857 г. и умер в Колдхарборе, в Англии, в 1936 г. Его отец, Уильям Пирсон, был королевским адвокатом [по назначению правительства] и председателем суда Лорда-Канцлера. Это был человек выдающихся способностей, исключительно энергичный и в умственном, и в физическом смысле, сильно заинтересованный в исторических изысканиях. Все эти особенности выказывал и его сын.

Сотрудница Пирсона, Юлия Белл, сообщила об эпизоде из его раннего детства, который в зачаточном состоянии содержал многие отличительные черты его дальнейшей жизни. Самым первым воспоминанием в жизни, как он сказал в ответ на её вопрос, было сосание своего большого пальца. Он сидел тогда на высоком детском стульчике, и какая-то женщина предупредила его, что от сосания палец засохнет. Он тогда сблизил большие пальцы, долго смотрел на них и сказал самому себе, что *они кажутся одинаковыми. Я не вижу, чтобы тот, который я сосу, был хоть сколько-нибудь меньше другого. Не могла ли она обмануть меня?*

В этом простом рассказе мы замечаем отказ от повиновения законной власти, обращение к эмпирическим свидетельствам, веру в своё собственное истолкование наблюдений, и, наконец, обвинение той, чьё суждение отличалось от его собственного вывода, в нечестности. Эти характерные черты оставались заметными в течение всей его научной карьеры.

Основным источником сведений о ранней жизни Пирсона служит мемуар другого выдающегося статистика, его сына Эгона Пирсона (1938). Первые годы его обучения содержат указания на необыкновенную широту интересов, необычную интеллектуальную энергию, наслаждение спорами, решимость сопротивляться всему, что он считал ошибочным проявлением власти, высокую оценку эрудиции и стремление к выражению своих мнений. Но почти не было никакого предположения о склонности к тем дисциплинам, в связи с которыми он теперь в основном вспоминается.

В 1866 г. его послали в школу при University College в Лондоне, но через несколько лет он выбыл из неё ввиду неважного здоровья. В возрасте 18 лет он получил стипендию в Кингс-Колледже в Кембридже, оказавшись вторым в списке [стипендиатов?]. В автобиографической заметке Пирсон (1936) назвал эти годы одними из счастливейших лет в жизни.

Наслаждение было в дружбе, в потасовках, в преподавании наших инструкторов, в поисках нового видения и в математике, и в философии и религии.

Его преподавателем был Эдвард Джон Раус, которого некоторые считают самым успешным учителем в истории Кембриджа. Пирсон по-настоящему привязался к нему. Он нередко рассказывал, как его преподаватели математики Раус, Burnside и Frost поощряли его, и описывал свои контакты с другими выдающимися лицами. Он оставил забавное описание экзаменов, происшедших в течение четырёх дней на дому у четырёх экзаменаторов: Стокса, которого он чтит как величайшего физика математического направления в Англии и одного из двух лучших среди известных ему лекторов; Максвелла, другого великого физика, но неважного лектора; Кэли, юриста и математика, изобретателя теории матриц и n -мерной геометрии; Тодхантера, который к тому времени уже опубликовал свою *Историю* (1865). Список экзаменационных вопросов, составленных последним, оказался поворотным пунктом в карьере Пирсона. Доказательство, представленное им на этом экзамене, было приложено вместе с одобрительным комментарием Тодхантера к неоконченной рукописи его другой *Истории* (1886/1893).

После смерти Тодхантера его попросили закончить эту рукопись и прокомментировать её, и это задание оказалось началом его близости с издательством Кембриджского университета. В течение последующих 50 лет корректуры из этого издательства почти всегда лежали на письменном столе Пирсона.

Помимо математики и теории упругости Пирсон в течение своих кембриджских лет интересовался философией, особенно работами Спинозы, сочинениями Гёте, Данте и Руссо, которых он прочёл в оригинале, историей религиозной мысли и поисками понятия божества, совместимого с тем, что было ему известно из науки. Серьёзно занимаясь религией, но возмущаясь принуждением, он оспорил администрацию университета своим отказом продолжать посещение обязательных лекций по богословию, а затем возражениями против обязательного посещения церкви. Пирсон победил в обоих случаях, и университетские правила были изменены, но он продолжал посещать церковь уже добровольно.

В 1879 г. Пирсон с отличием сдал экзамены по математике на степень бакалавра. Затем он изучил юриспруденцию в Линкольнс-Инн¹, а в 1881 г. получил право на адвокатскую практику. После этого он уехал в Германию и обучался в университетах Гейдельберга и Берлина, изучая физику, метафизику, римское право и историю Реформации, немецкий фольклор и социализм и дарвинизм.

Вскоре по возвращении в Англию он начал читать лекции и писать о немецкой социальной жизни и мысли, о Мартине Лютере, Карле Марксе, Маймониде и Спинозе; сочинять гимны для

социалистического песенника и статьи по упругости; обучать математике в Кингс-Колледже и схватываться в литературных дуэлях с Мэтью Арнолдом и библиотекарями в Британском музее. Дружба с Генри Бредшоу, библиотекарем Кембриджского университета, которого Пирсон упоминал в своих воспоминаниях (1936), как человека, который в наибольшей степени воздействовал на наше поколение, сильно повлияла на его жизнь. В одном из своих выступлений он описал Бредшоу как идеального библиотекаря, но и как ещё более значимого руководителя молодых и глупых, и добавил, что тот объяснил ему, в чём должно состоять настоящее мастерство. Их дружба была столь тесной, что Бредшоу смог порицать Пирсона за излишнее рвение и недостаток осмотрительности в интеллектуальных спорах.

Первой публикацией 23-х летнего Пирсона оказалась брошюра (1880), в основном должно быть автобиографическая. Написана она была в форме писем молодого человека, Артура, своей невесте и предвещала последующие сочинения (1888; 1892). Вот слова Артура, приведенные Эгоном Пирсоном (1938, с. 8):

Я бросаюсь от науки к философии, а от неё к нашим старым друзьям – поэтам. Затем, утомлённый сверх меры избытком идеализма, я воображаю, что, вернувшись к наукам, стану практиком. Пытались ли вы когда-либо представить себе, сколько в мире есть стоящего узнать? Что любая тема во всей вселенной достойна быть изученной? Гиганты литературы, тайны многомерного пространства, попытки Больцмана и Крукса проникнуть в истинную лабораторию природы, кантианская теория вселенной, последние открытия в эмбриологии с их удивительными рассказами о развитии жизни, – каково же необъятное пространство вне нашей способности восприятия! [...] Человек, видимо, находится на грани нового славного открытия. Чего добился Ньютон для того, чтобы упростить [понимание] различных движений планет, теперь должно быть сделано, чтобы объединить в единое целое различные изолированные теории математической физики.

Молодой автор хотел лишь полностью понять вселенную. Через 30 лет Биометрика опубликовала фото статуи Чарльза Дарвина с надписью *Ignoramus, in hoc signo laboremus*. Эти пять слов звучат как основная тема жизни Пирсона: *Мы невежественны, так приступим же к работе!*

В 1884 г. Пирсон стал профессором прикладной математики и механики в University College и начал обучать математике студентов инженерных специальностей, предлагая также курсы по геометрии. Несколько последующих лет он составлял статьи по упругости, завершал рукопись Годхантера, читал лекции о социализме и свободомыслии, опубликовал свою *Этику* (1888), написал на немецком языке книгу (1887) о легендах о портретах Иисуса, собирал материал для немецкой драмы о его страданиях и смерти, которая позднее оказалась ядром его книги (1897), закончил недописанную книгу Clifford (1885) и принимал активное участие в работе небольшого клуба, нацеленного на

преодоление общепринятого барьера, который препятствовал свободному обсуждению взаимоотношений мужчин и женщин.

В 1890 г. Пирсону поручили читать лекции по геометрии в лондонском Грешем-Колледже, предоставив ему свободный выбор их содержания. В марте 1891 г. он прочёл свой первый курс из четырёх лекций, *Содержание и понятия современной науки*, затем вышло в свет первое издание его *Грамматики науки* (1892), а в 1893 г. он написал статью (1894) об асимметричных кривых плотностей распределения.

Он существенно изменил своё понятие о научном методе, убедился в существовании статистических основ познания и в срочности решения проблем эволюции и наследственности. Короче, его профессиональная жизнь неожиданно и серьёзно изменилась.

[1.2.] Что повлияло на мысли Пирсона. В 1890 г. Уэлдон был назначен заведывать кафедрой биологии в University College. Он уже был знаком с Гальтоном и занимался статистическими исследованиями. В том же году он опубликовал статью о вариациях в креветках, в 1892 г. – о коррелированных вариациях, а в статье 1893 г. заявил, что *Нельзя достаточно сильно указывать, что проблема эволюции животных является по существу статистической.*

В то время подобная точка зрения была еретической, и научную значимость его вскоре образовавшейся дружбы с Пирсоном вряд ли можно преувеличить. Каждому из них было в то время 30 лет с небольшим. Уэлдон задавал вопросы, которые подводили Пирсона к некоторым из его важнейших работ, и его неожиданная смерть от воспаления лёгких в возрасте 46 лет оказалась тяжёлой потерей для науки и серьёзной личной трагедией для Пирсона.

В 1889 г. Гальтон, будучи 67 лет от роду, опубликовал свою книгу *Natural Inheritance*, сведя воедино свои исследования 1877 – 1885 гг. о регрессии. Она побудила Уэлдона изучать регрессию в биологических популяциях, а Пирсона подвинула к арифметическим исследованиям, которые завершились в 1897 г. введением коэффициента корреляции (r) моментов произведений. Разработка понятия корреляции другими авторами привела к таким разнообразным статистическим изобретениям, как факторный и дисперсионный анализы.

Побуждение, которое Пирсон получил от Гальтона, и его преданность этому старшему коллеге, видны на каждой странице составленной им биографии этого учёного (1914 – 1930), одной из самых подробных в мире биографий вообще. В ней более 1300 страниц в четвертую долю листа и около 170 иллюстраций на отдельных страницах. Примерно 20 лет жизни и изрядные личные средства потратил Пирсон на её составление.

[1.3.] Развитие науки статистики. Год 1890-й был переломным в карьере Пирсона и также ознаменовал начало статистики как науки. В течение долгого предварительного периода интерес к статистическому образу мыслей медленно возрастал, и к тому году он был ещё случайным, ограниченным

по охвату и разделялся очень немногими. В основном он проявлялся в сборе данных о населении, о его движении и экономике. Виден он был и в работе по страхованию жизни и в астрономии и метеорологии, главным образом в связи с уравниванием наблюдений по методу наименьших квадратов².

Вне указанных областей развитию статистики препятствовали не только отсутствие интереса, но и скудость данных и отсутствие подходящей теории. Почти вся существовавшая теория была разработана великими астрономами и математиками, которых интересовала математическая вероятность, связанная с ошибками наблюдений и в основном относящаяся к биномиальному и нормальному распределениям одной-единственной переменной. Сбор государственной статистики государственными и частично принадлежащими государствам агентствами был укоренившейся практикой.

Примерно после 1800 г. большинство индустриальных стран установило национальные переписи, и появилось несколько внегосударственных обществ, в основном для повышения качества государственной статистики. Таковыми были Лондонское (ныне Королевское) статистическое общество (в 1834 г.) и, в 1839 г., Американская статистическая ассоциация. Работа актуария стала достаточно разработанной и уважаемой профессией³.

Период 1853 – 1878 гг. был эрой многолюдных международных статистических конгрессов; значительно продвинулась экономическая статистика и заметно улучшились методы сбора данных. Государства начали измерять физические характеристики своих солдат и предоставлять получаемые данные антропометристам.

Среди предшественников Пирсона были учёные, существенно продвинувшие математическую теорию вероятностей по отношению к проблемам азартных игр, но они никогда не испытывали своих теорий по имеющимся данным и не предлагали применять её ни в какой иной области⁴. Теоретиками подобного рода были Ферма, Паскаль, Гюйгенс и Муавр. Другие математики указывали на возможность приложения теории вероятностей к социальным явлениям, но у них не было необходимых данных.

Якоб Бернулли писал о подобных возможных применениях в экономике, Даниил Бернулли исследовал вариацию оспы, а Николай Бернулли, Кондорсе и Пуассон, среди прочих, изучали правдоподобность свидетельств и занимались родственными юридическими проблемами. Ещё до 1800 г. Уильям Плейфейр изобрёл статистический график и опубликовал прекрасные статистические диаграммы, основанные на весьма сомнительных данных. Никто из них не думал о проверке своих теорий по собранным данным⁵, да и не имел надлежащих материалов, а с другой стороны многие другие работали в статистических агентствах табулируя данные, но почти не представляли, как исследовать их.

Две группы, актуарии и математические астрономы, обладали и математической сообразительностью, и необходимыми данными

для проверки теорий, но ни одна из них не предложила общего статистического подхода к проблемам, находящимся вне их собственной сферы деятельности⁶. Великие математические астрономы первой половины XIX в., особенно Лаплас и Гаусс, действительно заложили основы современной статистической теории, разработав понятие погрешности наблюдений⁷ и впечатляющую сопровождающую математическую теорию, а побуждаемые ими идеи распространялись по Европе.

Существенный вклад внесли Бессель и Энке в Германии, Плана в Италии, Лежандр, Пуассон, Фурье, Браве и *гражданин* Крамп⁸ во Франции, Кетле в Бельгии, Эри и Де Морган в Англии и Тиль в Дании. [...]

Пирсон неизменно настаивал на публикации и первоначальных данных, и выведенных из них статистик. В основном он стремился разработать методологию исследования жизни, а не совершенствовать математическую теорию. Разработав новый участок статистической теории, он сразу же применял его к имеющимся данным, а если его математика была неуклюжа, это не заботило его.

[2.] Основные направления трудов

[2.1.] Кривые плотности. Одна из проблем, которой Пирсон посвятил много времени и энергии, была разработка системы обобщённых кривых плотностей, основанных на единственном дифференциальном уравнении, и его параметров, полученных по методу моментов. Кетле, видимо, считал, что при достаточном числе случаев почти все социальные явления примерно подчинены нормальному закону⁹. До 1890 г. Грам и Тиле в Дании предложили теорию асимметричных кривых плотностей, а после публикации пирсоновской (1894; 1895) тщательно разработанной и весьма интересной системы появилось большое число статей о таких родственных темах, как подгонка кривых к усечённым или недостаточно хорошо известным распределениям и таблицы некоторых распределений вероятностей.

[2.2.] Критерий хи-квадрат. Подогнав кривую к точкам наблюдений, Пирсон должен был отыскать критерий добротности этой процедуры и изобрёл критерий хи-квадрат (1900). Кетле и другие авторы, пытаясь доказать хорошую согласованность плотностей распределения наблюдаемых значений и вычисленных в предположении нормального распределения, просто располагали оба ряда друг возле друга и по существу заявляли: *Смотри и убеждайся!* У них не было меры расхождения, и они, видимо, не ощущали от этого никаких неудобств.

Пирсон придумал такую меру и вычислил соответствующее распределение. Но он, кажется, так и не разобрался в понятии степеней свободы ни по отношению к критерию хи-квадрат, ни к своим формулам вероятных ошибок. И всё же его изобретение оказалось исключительно полезным и применимым намного шире, чем только для той конкретной задачи, для которой оно было введено. Оно занимает важное место в современной статистической теории.

[2.3.] Корреляция. Идею корреляции выдвинул Гальтон (1880; 1885) и позже описал её в более известной книге (1889).

Поверхности нормальной корреляции были математически исследованы ещё раньше в связи с погрешностями оценивания положения точки в пространстве. В 1809 г. Эджейн предложил первый известный вывод вероятности совместного появления двух таких погрешностей, но рассматривал только некоррелированные ошибки. Функцию плотности двух связанных друг с другом погрешностей вывел Лаплас в 1810 г., а для n погрешностей – Гаусс в 1823 г. и быть может и раньше¹⁰. В 1812 г. Плана изучал вероятности ошибок в геодезических работах, а Браве в 1856 г. – вероятности ошибок в артиллерийской стрельбе, причём каждый из них вывел уравнение, которое включало член, подобный r . Но, интересуясь вероятностями появления ошибок, а не силой связи между ними, они исследовали выведенную плотность, не обращая внимания на этот член. Свои новшества они применили только к ошибкам наблюдений, а отношение их статей к корреляционной поверхности было замечено лишь намного позже появления важных исследований корреляции. В 1877 г. Боудитч исследовал рост, вес и возраст 24 500 бостонских школьников и опубликовал кривые соотношения роста к весу, но не обнаружил [не изучал] корреляции между ними.

Серьёзным препятствием в корреляционных исследованиях Гальтона был недостаток данных и отсутствие действенных правил вычисления. Данные были у него о душистом горошке и о росте родителей и взрослых детей в 200 семьях, а когда Пирсон начал читать лекции о корреляции, Уэлдон приступил к измерениям креветок с целью изучения корреляции.

В первой основополагающей статье о корреляции Пирсон (1896) обобщил выводы и методы Гальтона, вывел формулу, которая теперь называется корреляцией моментов произведений, и две другие, равносильные формулы, предложил краткие правила вычислений, которыми могли бы воспользоваться исследователи со скромными математическими познаниями, сформулировал общую теорию корреляции для трёх переменных, привёл коэффициенты уравнения множественной регрессии, выраженные через коэффициенты корреляции.

Последовала публикация ряда важнейших мемуаров о различных аспектах корреляции, в том числе написанные такими учениками и сотрудниками Пирсона, как Юлом и Шеппардом, но в основном им самим. Они исследовали корреляции при не нормальных распределениях, тетраэдрические корреляции, корреляции рангов, корреляции, при которых одна или обе переменные не заданы количественно или регрессия нелинейна.

Во многих статьях описывались результаты корреляционного анализа в весьма разнообразных областях, и много труда было потрачено на вывод вероятных ошибок¹¹ каждого из различных коэффициентов и на табулирование различных вероятностей, относящихся к корреляциям, а коэффициент корреляции моментов произведений теперь подобающе называется именем Пирсона.

[2.4.] Индивидуальная изменчивость. Разброс ошибок наблюдения положения небесных тел подробно изучались великими математическими астрономами. Много статей было до 1850 г. посвящено *закону лёгкости ошибок*, выводу формулы нормальной кривой и составлению таблиц вероятности. Термин *вероятная ошибка* начали широко употреблять через несколько лет после его введения Бесселем в 1815 г. в статье о положении Полярной.

Понятие истинной изменчивости от одного индивидуума к другому весьма отлично от понятия вариации случайных ошибок при оценивании единого значения. Идея индивидуальной изменчивости занимает важное место в сочинениях Кетле, Фехнера, Эббингауза, Лексиса, Эджуорта, Гальтона и Уэлдона, но она не была широко воспринята другими учёными XIX в. Настаивание Пирсона на этой идее оказалось одним из его существенных новшеств в понимании жизни.

В своей первой основной работе об асимметричных кривых плотностей он (1894) ввёл термин *стандартное отклонение* и символ σ и постоянно применял и то, и другое при обсуждении вариаций между индивидуумами. Однако, при исследовании вариаций в выборке он неизменно пользовался термином *вероятная ошибка* и таким образом чётко различал вариации, вызванные индивидуальными отличиями, и обусловленные случайными ошибками.

[2.5.] Вероятные ошибки статистик. Сам Пирсон видимо считал, что одним из его важнейших достижений был вывод вероятных ошибок *констант частоты*, как он называл их, и составление различных таблиц для их вычисления. Его метод, бывший уже хорошо известным в силу иных причин, состоял в том, чтобы написать для статистики уравнение, вычислить дифференциалы обеих его частей, возвести в квадрат, суммировать и преобразовывать, часто весьма сложным образом, результат любым алгебраическим способом.

Хоть он и не очень заботился о различии между статистикой и её параметром [её значением?] и часто применял второе вместо первого, эти вероятные ошибки являлись серьёзным продвижением, поскольку до того лишь для немногих статистик существовали какие-либо меры вариации в выборках. Новые статистики начали предлагаться со всех сторон, и громадные усилия тратились на вывод этих вероятных ошибок. После появления статей Стьюдента 1908 г. и Фишера (1915) о выборочном распределении коэффициента корреляции начались успешные поиски точных выборочных распределений. Во многих случаях теперь известны методы, лучшие, чем обеспечиваемые пирсоновыми формулами вероятных ошибок.

[2.6.] Публикация таблиц. Редакционная статья в первом номере *Биометрики*, неизменно приписываемая Пирсону, указывала на срочную потребность в таблицах для облегчения работы статистиков и биометристов и обещала составлять их так быстро, как только было возможно. Существовавшие в то время таблицы были опубликованы в разбросанных источниках, и

некоторые из них оказались почти недоступными. К 1900 г. обычными стали таблицы биномиальных коэффициентов, тригонометрических функций, логарифмов и антилогарифмов. Обширная таблица логарифмов факториалов, которую в 1824 г. опубликовал Ф. С. Degen, была почти неизвестна¹². Существовали таблицы квадратов, кубов, квадратных корней и обратных величин, из которых самой известной была таблица Barlow. Были и таблицы умножений Крелле и Coatsworth. Лежандр опубликовал таблицу логарифмов гамма-функции, но число её экземпляров было очень невелико.

Много раз публиковалась таблица функции нормальной вероятности, но с аргументом вероятной ошибки ($0,6745\sigma$) или модуля ($\sigma/\sqrt{2}$), но не стандартного отклонения. Пуассон не составил таблиц распределения, которое носит его имя, но это сделал Борткевич в 1898 г. в своей брошюре о законе малых чисел. Список таблиц, опубликованных в *Биометрике* начиная с её второго номера 1902 г. до настоящего времени, или тех, которые появились в отдельных томах (1914) или в серии *Drapers' Co. Series of Tracts for Computers*, был бы очень длинным. Некоторые из этих таблиц уже не используются, другие же видимо останутся ценными навсегда, даже после появления электронных компьютеров. Среди последних достижений Пирсона были и таблицы (1934), появившиеся, когда ему было 78 лет.

[3.] Дискуссии

Нельзя обойти молчанием частые научные споры, в которые ввязывался Пирсон. В юности он сражался за такие непривычные радикальные идеи, как социализм, эмансипация женщин и этика свободомыслия. Через несколько лет он втянулся в длительную борьбу за необычную идею о том, что математика должна быть приложена к биологии. По этому вопросу возникли пылкие и неприятные чувства, а Королевское общество, готовое публиковать статьи либо по математике, либо по биологии, отклоняло те, которые обсуждали и то, и другое.

Этот отказ послужил одной из причин, приведших в 1900 г. к появлению *Биометрики*, что в свою очередь серьёзно содействовало молодым дисциплинам биометрии и математической статистики, поскольку появился журнал, в котором можно было публиковать математические статьи о биологических науках. В 1904 г. Гальтон учредил Бюро евгенических записей для поощрения научного изучения евгеники. Через два года он передал своё Бюро Пирсону, чтобы тот мог руководить им вместе со своей Лабораторией биометрии, а Бюро стало известно как Лаборатория евгеники.

Биометрическая лаборатория, чьё существование Пирсон отнёс к 1895 г., была центром обучения лиц, закончивших университет, новой отрасли точных наук. В 1911 г. эти лаборатории были объединены и образовали факультет прикладной статистики в University College, и Пирсон стал его первым профессором. С 1907 г. Лаборатория евгеники начала публиковать многочисленные и весьма важные статистические статьи о трёх из самых важных спорных вопросов того времени, туберкулёзе

лёгких, пьянстве и слабоумии и безумии. В противоположность распространённому тогда мнению о том, что туберкулёз может быть искоренён путём оздоровления окружающей среды, статистические исследования Пирсона указывали, что предрасположение к нему было скорее наследственным. Кроме того, оказалось, что не было явного свидетельства о том, что санатории более успешно вылечивали пациентов, чем другие учреждения.

По общепринятому мнению дети родителей-алкоголиков умственно отстали и физически неполноценны. Первые исследования Лаборатории евгеники по этой проблеме (1910) не обнаружили какого-либо заметного соотношения между пьянством родителей и интеллектом, телосложением или болезненностью потомства. Позднейшие исследования показали, что алкоголизм является скорее следствием, а не причиной слабоумия. Пирсон (Эгон Пирсон 1938, с. 61) заключил, что *подходит время, когда истинное познание должно будет занять место энергичной, но необученной филантропии для указания путей возможной социальной реформы.*

В 1912 г. Американское бюро евгенических записей объявило, что слабоумие почти наверняка является рецессивным менделевским признаком и заявило, что *слабым в какой-либо отличительной черте следует жениться на силе этой черты, силе же допустимо жениться на слабости.* Пирсон и его сотрудники упорядочили статистические свидетельства и опровергли это утверждение.

Каждый раз, когда Пирсон брался за подобную тему, авторитетные врачи и официальные лица возмущались, и их бранные атаки лично на Пирсона продолжались длительное время. Произошёл и открытый конфликт между более традиционным Обществом евгенического образования, почётным президентом которого был Гальтон, и Лабораторией евгеники, которую он основал.

Молодые науки биометрии и статистики быть может выиграли от этих серьёзных битв с организованными группами, потому что у них оказался повод переломить сдерживающие путы апатии, невежества и укоренившихся авторитетов. Пирсон был своего рода крестоносцем, а среди необходимых ему поэтому качеств были самоуверенность, смелость в борьбе за свои убеждения и чуток интеллектуальной нетерпимости¹³. Он был максималистом и вряд ли терпел идеи или работу, которые считал ошибочными. Более того, он получил юридическое образование и с детства в лице своего отца имел перед собой пример удачливого судьи. Но первая его мысль была неизменно направлена на отыскание истины, а убедившись в ошибке, он был готов признать её. Однажды, в 1919 г., он поместил в *Биометрике* статью под названием *Peccavimus!* (Повинились).

Пирсон внёс вклад в статистические методы, которые, как представляется теперь, будут важны длительное время, но существеннее было то, что он пробудил научный мир. Полнейшее безразличие к статистическим исследованиям сменилось

нетерпеливыми усилиями многих хорошо подготовленных лиц, которые разработали новую теорию, собрали и исследовали статистические данные, относящиеся к каждой сфере приложений, составили новые таблицы и переосмыслили основания статистической философии. Это было громадным достижением. Его лаборатория стала международным центром, в котором работали исследователи всех стран. Мало кто во всей истории науки побудил столь многих других культивировать и расширять поле, которые они сами засеяли. Он дал учёным идею общей методологии, лежащей в основе всей науки, т. е. сделал один из громадных вкладов в современное мышление.

Краткие сведения об упомянутых лицах

Arnold, Matthew, Арнолд, Мэтью, 1822 – 1888, поэт, педагог, искусствовед

Bowditch, Nathaniel, Бовдитч, Натаниел, 1773 – 1838, математик и астроном, переводчик *Небесной механики* Лапласа на английский язык

Crookes, William, Крукс, Уильям, 1832 – 1919, физик и химик, Президент Королевского общества в 1913 – 1915 гг.

Ebbinghaus, Herman, Эббингауз, Герман, 1850 – 1909, психолог

Routh, Edw. John, Раус, Эдвард Джон, 1831 – 1907, математик, механик

Примечания

1. Одно из лондонских юридических заведений.
2. К тому времени статистический метод давно уже проник в естествознание (в астрономию, медицину, метеорологию, физику), см. Шейнин (2013, § 11.8).
3. Эту последнюю фразу следовало всё же относить ко второй половине XIX в. (Sheynin 1977, § 2.3.2).
4. Такой областью была смертность (Гюйгенс, Муавр; см. Шейнин (2013, §§ 3.2.2 и 5.2)), не говоря уж о теории ошибок. Также вопреки автору были возможны и приложения к социальным явлениям (см. чуть ниже), достаточно вспомнить об изучении судебной статистики Пуассоном. Наконец, Якоб Бернулли не успел ничего написать о приложениях к экономике.
5. Автор повторно и крайне неправдоподобно утверждает, что никто не думал о проверке теорий по эмпирическим данным.
6. Общими статистическими подходами были методы наибольшего правдоподобия и наименьшей дисперсии.
7. Не разработали понятие, а определили (случайную ошибку).
8. *Гражданин* Крамп, видимо, означает, что он стал гражданином Франции.
9. Вопреки автору, о сочинениях Кетле см. Шейнин (2013, § 11.5).
10. Гаусс (1823, § 12) исследовал интеграл от произведения плотностей n независимых погрешностей. Никакой корреляции он не изучал, и вообще указанный параграф непонятен. Погрешности артиллерийской стрельбы (см. чуть ниже) изучали также Poisson (1837) и Didion (1858).
11. Автор ни слова не сказала о недостатке вероятной ошибки (которая встретилась уже в § 2.2) как меры точности.
12. Первую таблицу этой функции составил Муавр (Шейнин 2013, § 5.4). Вообще же о математических таблицах см. Лебедев и др. (1956).
13. Хорошо известно, что интеллектуальной нетерпимости Пирсону было не занимать.

Библиография

K. Pearson

- 1880, *The New Werther*. London.
- 1887, *Die Fronica: Ein Beitrag zur Geschichte des Christusbildes im Mittelalter*. Strassburg.
- 1888, *The Ethic of Freethought etc.* London, 1901.
- 1892, англ. *Грамматика науки*. СПб, 1911.
- 1894, On the dissection of asymmetrical frequency curves. *Phil. Trans. Roy. Soc.*, vol. A185, pp. 71 – 110. В книге автора (1948, pp. 1 – 40).
- 1895, Skew variation in homogeneous material. *Phil. Trans. Roy. Soc.*, vol. A186, pp. 343 – 414. В книге автора (1948, pp. 41 – 112).
- 1896, Regression, heredity and panmixia. *Phil. Trans. Roy. Soc.*, vol. A187, pp. 253 – 318. В книге автора (1948, pp. 113 – 178).
- 1897, *Chances of Death etc*, vols 1 – 2. New York.
- 1898, On the probable errors of frequency constants etc. *Phil. Trans. Roy. Soc.*, vol. A191, pp. 179 – 261. Соавтор L. N. G. Filon.
- 1900, On the criterion that a given system ... *London, Edinb. and Dublin Phil. Mag.*, vol. 50, pp. 157 – 175. В книге автора (1948, pp. 339 – 357).
- 1906, W. F. R. Weldon, 1860 – 1906. *Biometrika*, vol. 5, pp. 1 – 52.
- 1907, *First Study of the Statistics of Pulmonary Tuberculosis. Drapers' Co. Res. Mem., Studies in Nat. Deterioration* No. 2. London.
- 1910, *First Study of the Influence of Parental Alcoholism etc. Mem. Eugenics Lab.*, vol. 10. London. Соавтор Ethel M. Elderton.
- 1914, редактор, *Tables for Statisticians and Biometricians*, vols 1 – 2. London, 1930 – 1931.
- 1914 – 1930, *Life, Letters and Labours of Fr. Galton*, vols 1 – 3 в четырёх частях. Cambridge.
- 1922, редактор, *Tables of the Incomplete Gamma-Function*. London, 1951.
- 1923, *On the Relationship of Health to the Psychical and Physical Characters in School Children*. Cambridge.
- 1934, редактор, *Tables of the Incomplete Beta-Function*. London.
- 1936, Old tripos days etc. *Math. Gazette*, vol. 20, pp. 27 – 36.
- 1948, *Early Statistical Papers* [статьи 1894 – 1916 гг.]. Cambridge. *Annals of Eugenics*. Редактор с 1925 г. до конца жизни.
- Biometrika*. Редактор с 1900 г. до конца жизни.
- Drapers' Co. Res. Mem. Biometric series*. Редактор с первого выпуска 1904 г. Содержит ряд основных сочинений К. Пирсона и его сотрудников.

Другие авторы

- Лебедев А. В. и др. (1956), *Справочник по математическим таблицам*. М. Дополнение 1959 г.
- Шейнин О. Б., Sheynin O. (1977), Early history of the theory of probability. *Arch. Hist. Ex. Sci.*, vol. 17, pp. 201 – 259.
- (2010), Karl Pearson a century and a half after his birth. *Math. Scientist*, vol. 35, pp. 1 – 9.
- (2013), *Теория вероятностей. Исторический очерк*. Берлин. Также Google, Oscar Sheynin, Download area.
- Clifford W. K. (1885), *Common Sense of the Exact Sciences*. Редакция и предисловие К. Пирсона. Редактор второго издания J. R. Newman. New York, 1946.
- Didion Is. (1858), *Calcul des probabilités etc.* Paris.
- Filon L. N. G. (1936), Karl Pearson as an applied mathematician. *Obituary Notices Fellows Roy. Soc.*, vol. 2, No. 5, pp. 73 – 109.
- Fisher R. A. (1915), Frequency distribution of the value of the correlation coefficient ... *Biometrika*, vol. 10, pp. 507 – 521.
- Galton Fr. (1880), Correlations and their measurement.
- (1885? 1886?), Regression towards mediocrity in hereditary stature. *J. Anthropol. Inst.*, vol. 15, pp. 246 – 263.
- (1888), Correlations and their measurements from anthropometric data. *Proc. Roy. Soc.*, vol. 45, pp. 135 – 145.
- (1889), *Natural Inheritance*. New York, 1973.
- Gauss K. F., Гаусс К. Ф. (1823, латин.), Теория комбинаций наблюдений и т. д. *Избр. геод. соч.*, т. 1. М., 1957, с. 17 – 57.
- Haldane J. B. S. (1957), Karl Pearson. A centenary lecture. *Biometrika*, vol. 44, pp. 303 – 313. Перепечатка: E. S. Pearson & Kendall (1970, pp. 427 – 437).

Morant G. и др., редакторы (1939), *Bibliography of Statistical and Other Writings of Karl Pearson*. Cambridge.

Pearson E. S. (1938), *Karl Pearson*. Cambridge. Также в *Biometrika*, vol. 28, 1936, pp. 193 – 257; vol. 29, 1937, pp. 161 – 248.

Pearson E. S., Kendall M. G., редакторы (1970), *Studies in the History of Statistics and Probability*. London.

Poisson S.-D. (1837), Sur la probabilité du tir à la cible. *Mémorial de l'artill.*, No. 4, pp. 59 – 94.

Todhunter I. (1865), *History of the Mathematical Theory of Probability*. New York, 1949, 1965.

--- (1886), *History of the Theory of Elasticity etc.* vols 1 – 2. Закончил и отредактировал К. Пирсон. Cambridge, 1893.

Walker Helen M. (1958), Contributions of Karl Pearson. *J. Amer. Stat. Assoc.*, vol. 53, pp. 11 – 22.

Wilks S. S. (1941), Karl Pearson. *Scientific Monthly*, vol. 53, pp. 249 – 253.

Х

М. С. Бартлетт

Р. А. Фишер

M. S. Bartlett, Fisher R. A. IES, pp. 352 – 358

[1. Биография]

Уже при жизни Рональд Айлмер Фишер (1890 – 1962) добился всемирного признания как статистик и генетик. В начале XX в., продолжая труды Пирсона, он придал им новые направления. Другие авторы также способствовали громадному всплеску в развитии статистических методов и их приложению к биологии, но эти двое своими энергичными исследованиями и примерами своих трудов по очереди в течение целого поколения продолжали владеть статистической ареной.

Фишер родился в East Finchley возле Лондона [ныне в Лондоне]. Не считая его брата-близнеца, который не прожил долго, он был самым младшим из семи детей. Его отец, Джордж Фишер, был аукционером; никаких особых математических способностей у родственников Ф. не было заметно, кроме быть может у его дяди, который, как и он, кончил Кембриджский университет с отличием по математике.

Фишер учился в школе в Stanmore Park, затем в Харроу [ныне в Лондоне]. Там его интерес к математике одобрили, и он добился стипендии для обучения в колледже Гонвилл-энд-Киз в Кембридже. Его склонность к математике была в духе английской традиции приверженности к натуральной философии.

Свои студенческие годы (1909 – 1913) он закончил, получив диплом с отличием по оптике в 1912 г., а затем закончив аспирантуру по физике. Впрочем, изучив серию работ Пирсона *Математические сочинения по теории эволюции* [1894 – 1912], он уже заметил, что натуральная философия не заканчивалась физикой.

Покинув Кембридж, Ф. недолго пробыл в фирме Mercantile and General Investment Co. В 1914 г. началась Первая мировая война, но в армию его не призвали ввиду сильной близорукости, и в течение четырёх лет он преподавал математику и физику в нескольких школах для мальчиков. В 1917 г. он женился на Рут Эйлин Гиннес, родившей ему восьмерых детей.

По существу Ф. начал свою полноценную статистическую и биологическую карьеру лишь в 1919 г. на Ротемстедской экспериментальной станции, т. е. в сельскохозяйственном исследовательском институте. Впрочем, прежние годы оказались полезным периодом созревания. Ещё в 1912 г., будучи в Кембридже, Ф. предложил метод наибольшего правдоподобия для подбора кривых плотностей. Две другие основательные статьи навсегда закрепили за ним репутацию исследователя.

Первой из них была примечательная статья (1915), опубликованная в *Биометрике*; в ней он впервые полностью проявил свою мощь геометрического рассуждения. Вторая, в 1918 г., исследовала корреляцию между родственниками на основе менделевской наследственности, и в ней он выказал свою способность решать важнейшие задачи статистической генетики. Одновременно с приглашением на Ротемстедскую станцию Пирсон предложил ему работу в Университетском колледже Лондона.

Он разумно предпочёл первое, в основном потому, что тамошний статистик имел бы гораздо более широкое и независимое поле деятельности, но возможно и в связи с тем, что его контакты с Пирсоном не были особо обещающими. Не справившись с задачей о распределении коэффициента корреляции, и, будучи склонен врываться в исследования не особенно заботясь о предшественниках, Пирсон с воодушевлением перенял решение Фишера, но не продолжил тесного сотрудничества с ним, как того видимо требовала профессиональная этика. Ф. был удручён этим, что вполне могло послужить началом длительной и ожесточённой вражды между ними. Он имел основание критиковать многое в трудах Пирсона, однако возникшее личное раздражение не ограничивалось существенными научными разногласиями. Даже в 1950 г. [через 14 лет после смерти Пирсона], когда был опубликован сборник лучших статистических работ Фишера, отсутствие в нём статьи 1915 г. было молчаливым напоминанием о его чувствах.

Ротемстедский период (1919 – 1933) был самым блестящим и плодотворным в карьере Ф. Бригады биологов и благоприятная исследовательская атмосфера представляли собой именно ту среду, в которой он нуждался. Его собственный широкий круг биологических интересов позволил ему понимать проблемы его коллег и творчески обсуждать с ними соответствующие статистические аспекты.

Собственная статистическая деятельность Фишера была лучше всего представлена его известной книгой (1925а), которая выдержала 13 изданий и была переведена на несколько языков. Он сам проделал большой объём вычислений и начал многие свои собственные генетические опыты над домашней птицей, улитками и мышами, но помнить его будут за творческие теоретические способности. В 1929 г. Фишера избрали членом Королевского общества; через год он опубликовал классический труд (1930а), в котором в большой степени примирил дарвиновскую теорию эволюции с генетическими принципами Менделя, о которых Дарвин не знал.

В последних главах книги он обсуждал теорию, впервые предложенную Гальтоном, об эволюции генетической связи бесплодия со способностями. Она могла происходить вследствие женитьб удачливых ввиду врождённых способностей и преуспевающих по причине социальных преимуществ одного ребёнка или двух детей сравнительно малоплодотворных состоятельных родителей.

Энергичные взгляды Фишера привели его к рекомендации семейных пособий, пропорциональных величине семьи и зависящих от её дохода и уравновешивающих неблагоприятное положение детей, наказанных плодovitостью их родителей.

Генетические и евгенические интересы Ф. вскоре послужили причиной его переезда в Лондон в 1933 г., после его назначения преемником Пирсона в качестве гальтоновского профессора в Университетском колледже Лондона, что несомненно разожгло пламя их вражды¹. После отставки Пирсона колледж отделил преподавание статистики на новом факультете статистики под управлением сына Карла Пирсона, Эгона, от евгеники и биометрии, оставленных гальтоновскому профессору.

Эгон Пирсон в большей степени воспринимал новые статистические методы Фишера, которые подчёркивали роль точных методов анализа малых выборок, и выше оценивал их, но Фишер был расстроен, о чём свидетельствует его тогдашнее письмо У. С. Госсету [Стьюденту]. Нейман, который преподавал на факультете статистики в 1935 – 1938 гг., навлек на себя гнев Фишера за публикацию работы (1937), которую тот счёл либо ненужной, либо ошибочной. Их совместная работа в одном и том же здании усиливало у Фишера чувство обиды. Продолжение подобных враждебных чувств начинало быть проявлением собственного характера Фишера в той же мере, как и характера его противников.

Его широкие интересы и властная личность при желании делали его чарующим и оживлённым собеседником, и многие засвидетельствовали, что он был щедрым коллегой тех, кто сочувственно воспринимал его сочинения. Но его чувства и разум слишком сильно определялись работой, и он поэтому не выносил критики, на которую отвечал энергично и иногда совсем несправедливо. Если отвлечься от подобных прегрешений против объективности, можно сказать, что Фишер продолжал укреплять свою научную репутацию изучением генетики (особо, генетики человека) на своём факультете и публикацией новых работ по статистике, подобных книге (1935а), таблиц (1938), составленных в соавторстве и оригинальных статей в факультетском журнале *Annals of Eugenics* и др.

Третий основной период в научной карьере Фишера начался с его назначения в 1943 г. заведующим бальфуоровской кафедрой генетики в Кембридже², которой он руководил до своей отставки в 1957 г. В течение этого времени он написал две книги (1949; 1956) и отредактировал сборник своих статей (1920 – 1945/1950), но большинство его существенных трудов было уже начато.

Награды множилось: три медали Королевского общества (1938 г., королевская; 1948 г., дарвиновская; и 1955 г., медаль Копли) и возведение в рыцарское достоинство в 1952 г. Незадолго до отставки он стал главой своего колледжа в Кембридже, был почётным членом Американской академии искусств и наук, иностранным членом-корреспондентом Национальной академии наук США, членом Папской академии наук, иностранным членом

Королевской датской академии наук и Королевской шведской академии наук.

После отставки Фишер был гостем отделения математической статистики исследовательской организации Содружества наций по науке и технике в Аделаиде и ко времени своей смерти стал его членом.

[2.] Вклад в статистику и генетику

[2.1.] Статистика. Обращаясь несколько подробнее к оригинальным трудам Фишера, заметим, что формальный список его сочинений по статистике может привести к недооценке мощности его влияния. Вызвано оно было бы разнообразием и глубиной этих работ. Более того, такой список не удовлетворит нас, поскольку их теснейшая связь с практическими задачами, происходившими в его профессиональном окружении, иногда приводила к их неполному или запоздавшему научному описанию, или к тому и другому совместно. Обсуждение труда громадной значимости, как, например, метода дисперсионного анализа, в его книге (1925а) было недостаточным, потому что к тому времени он никак не мог считаться завершённым и всё ещё в довольно большой степени сопровождался загромождением идей корреляции внутри класса. Фишер, видимо, никогда не считал нужным улучшать изложение для последующих изданий книги.

Если исключить математическую работу Ф. по генетике, то всё же будет удобно постараться составить список его основных сочинений по теории статистики с разбивкой на две основные части: основополагающие работы по статистическим выводам и статистическая методология и методы.

В первую часть войдут его существенные труды по статистическим оценкам, т. е. в основном (1922) и (1925b). Фишер и до них серьёзно заботился о точных заключениях относительно таких знакомых величин как коэффициент корреляции и хи-квадрат по малым выборкам и опубликовал поток работ об их выборочных распределениях, из которых лучше всего известна статья (1915).

Он очень тщательно различал неизвестный параметр совокупности и его выборочную оценку. Если выборочное распределение оценки было известно в численной форме, становилось возможным проверять значимость любого произвольного значения параметра. Эти точные критерии значимости, к примеру, начальной гипотезы об отсутствии корреляции ($\rho = 0$) по выборочной корреляции, были в то время особенно полезны ввиду склонности биологов и других исследователей не обращать внимания на эти критерии. Но эти усилия Фишера были довольно неустойчивыми ввиду его последовавшего нападения на теорию Неймана – Пирсона исследования гипотез, особенно потому, что подчас необдуманное применение критериев значимости приводило к отрицательному мнению о них. Так, некоторые авторы не представляли себе, что из незначимости не следует, что исследуемая гипотеза верна ($\rho = 0$).

Фишер явно начал обдумывать свою общую теорию оценивания до 1922 г. Помимо его рекомендации применять наибольшее правдоподобие (1912). Понятие достаточности также появилось при рассмотрении особого случая среднего квадратического отклонения в качестве оценки истинного стандартного отклонения для выборок из нормального или гауссова распределения.

И всё же общая теория была впервые систематически разработана в двух указанных выше статьях, в которых также обсуждалось понятие состоятельности и эвристически выводились асимптотические свойства оценок метода наибольшего правдоподобия в больших выборках. Приняло определенную форму и понятие информации параметра θ в формуле

$$I = E[(\partial L / \partial \theta)^2], \quad L = \ln p(s | \theta),$$

где p было вероятностью выборки s при истинном значении θ .

Значимость этого состояла **1)** в исследовании действительных выборочных свойств оценок метода наибольшего правдоподобия, особенно в больших выборках. (Метод наибольшего правдоподобия имеет давнюю историю, по крайней мере поскольку он сходен с максимумом апостериорной вероятности в теореме Бейеса об обращённой вероятности при равномерном априорном распределении.) **2)** в подчёркивании того, что в некотором подходящем статистическом смысле выборка обеспечивает определённое количество информации о параметре. Фишерово понятие информации предшествовало работе Shannon (1948), в которой оно было введено совсем в ином контексте.

Понятие информации особо подходило к большим выборкам (см. выше пункт 1) ввиду возможности упорядочения нормально распределённых оценок в соответствии с их дисперсиями (Pollack 1968). Это понятие Фишер обосновал в более общем виде для малых выборок, имея в виду их большое число, но странно, что он упустил из вида точное неравенство Крамера – Рао, соотносящее функцию информации и дисперсию несмещённой оценки.

Но во всяком случае произвольность дисперсии и несмещённости остаётся, а для малых выборок Фишер ввёл общее понятие достаточной статистики. При независимости условных распределений любых других выборочных величин от неизвестного параметра они исчерпывали *информацию* выборки.

Следующая примечательная статья в этой области появилась позже (1930b). Фишер неизменно надсмехался над оценками, полученными из байесовского подхода к обращённой вероятности и следствиям из него. Он чувствовал, что единственная и более объективная система следствий должна будет оказаться возможной в тех областях, в которых действуют статистические вероятности, и заметил, что точное выборочное распределение, включающее выборочную величину или

статистику T и неизвестный параметр θ , приводит (при надлежащих свойствах регулярности и монотонности) к возможности установления *фидуциального интервала*, как он его назвал, с известной фидуциальной вероятностью того, что θ содержится в нём.

В соответствии с истолкованием, принятым в то время и подразумеваемым в собственных словах Фишера (1930b), этот интервал, необходимо зависящий от T , поэтому является случайным, а фидуциальная вероятность оказывается статистической с обычным частотным значением.

Ссылаясь на случай истинного коэффициента корреляции ρ , он (1930b/1950, с. 22.535) заявил:

Известно, что если взять некоторое количество выборок объёмом 4 единицы из одной и той же или из различных совокупностей, и вычислить для каждой фидуциальное 5%-ное значение ρ , то в 5% случаев истинное значение ρ окажется меньше найденного нами значения.

При некоторых ограничениях (к примеру, для достаточных статистик) это, очевидно, было тождественно с теорией доверительных интервалов, примерно в то же время разработанной Нейманом (1937). По поводу проблем индуктивных выводов Фишер часто не указывал ясно, что именно он заявляет, но следует подчеркнуть, что вне зависимости от его истолкования или отношения этого к решаемой тогда задаче, оно было всё же сформулировано в терминах принятой статистической структуры. И формальный обход теоремы Бейеса был мастерским ходом, который заметили и вне статистического сообщества. Вот замечание Эддингтона (1935a/1960, с. 126)³:

Мы никогда не можем быть уверены в каком-либо конкретном следствии, и поэтому должны стремиться приходить к системе следствий, не более указанной доли выводов из которых окажутся в конце концов ошибочными.

Позднее Фишер попытался обобщить фидуциальную теорию на случай нескольких параметров. В своей первой статье в этом направлении он (1935b) рассмотрел пример вывода разности средних двух выборок из нормальных совокупностей с различными дисперсиями. Он никогда подробно не исследовал логических соотношений достаточных статистик в случае нескольких параметров и поэтому в то время не заметил трудность, состоявшую в том, что влияние неизвестного отношения дисперсий при отсутствии *достаточной* величины, не включающей неизвестных параметров, как, например, средних отдельных совокупностей, не может быть выделено.

Отрицая соответствующую критику (Bartlett 1936), см. также (Bartlett 1965), Фишер явно отказался от общепринятого частотного истолкования фидуциальной вероятности, которое он, видимо, принимал ранее. И он, и другие авторы пытались сформулировать теорию для нескольких параметров, которая была бы и единственна, и не противоречива (self-consistent), но в какой-либо общности это ещё не достигнуто. Многие считают, что поиски подобной теории вводят в заблуждение, потому что

они не смогут исключить её произвольность, которую Фишер так сильно критиковал в байесовском подходе.

Несколько отличной и более формальной проблемой оценивания, которую Фишер решил в 1928 г., был вывод выборочных статистик, являющихся несмещёнными оценками соответствующих величин в совокупности, и выборочных моментов этих статистик. Величинами в совокупностях были семиинварианты (Thiele 1903), и комбинаторные правила вывода соответствующих выборочных статистик и их собственных семиинвариантов оказались поразительным примером интуитивной математической мощи Фишера.

Заслуживает упоминания и оригинальная и независимая статья (1934), относящаяся к теории игр, которую примерно в то же время разрабатывали von Neumann & Morgenstern (1944).

Работа Фишера по планированию эксперимента столь важна и логически, и практически, что её можно считать одним из его важнейших достижений в науке статистических выводов. Впрочем, её удобно рассматривать во второй основной части сочинений Фишера в связи с дисперсионным анализом. Он заметил простоту и действенность сбалансированных и независимых планов эксперимента. Повторение той же агрономической обработки на различных участках земли существенно для любой статистической оценки погрешности, причём желательны формально равные количества одинаково обрабатываемых участков.

Однако, упрощение в статистическом анализе обманчиво, если результат окажется неосновательным. При случайных наблюдениях приходится принимать наиболее разумные предположения о статистической изменчивости, но при контролируемых экспериментах имеется возможность намеренного включения случайности, так что систематическую вариацию можно отделять от чисто случайных ошибок. Таково было первое существенное соображение Фишера, и второе соображение естественно примкнуло к первому.

Анализ был приспособлен к планированию, и никакая вариация, не зависящая от обработки, не должна была включаться в ошибки. При равном числе участков, выделенных на каждый вид обработки, каждое полное повторение можно включать в отдельный блок, так что источником ошибок окажется только изменчивость от участка к участку в пределах одного и того же блока. Изменчивость между блоками может быть тем самым исключена как не имеющая отношения к поставленной цели.

Третье соображение возникло в связи с сочетаниями обработок, например, с применением различных видов удобрения. Если, к примеру, используются азотные (А) и фосфатные (Ф) удобрения, то рекомендуемый набор обработки участков таков:

Контроль (удобрений нет), А, Ф, АФ

Здесь АФ означает совместное применение обоих видов удобрения, каждый из которых вносится в том же количестве, как и при отдельном применении.

Этот план прост и может оказаться более эффективным. [Следует пояснение порядка действия удобрений А и Ф и указан план, который назван примером *факторного эксперимента*.] Фишер (1935а, с. 97) указал, что

Любые видоизменения сложного прибора, машины или промышленного процесса следует неизменно рассматривать как возможно воздействующие друг на друга, и каждое должно оцениваться в соответствии с вероятным результатом подобных взаимодействий. Если видоизменения приходится исследовать по очереди, то не потому, что это соответствует идеальному научному методу, а ввиду того, что их совместное исследование иногда оказывается слишком трудным или слишком дорогим. Во многих случаях [...] это впечатление [однако] весьма преувеличено.

Далее, в факторном анализе естественным образом появилось *смешивание*, при котором при трёх и более факторах некоторые влияния взаимодействий более высоких порядков предполагаются незначительными и намеренно располагаются (?) так, чтобы при анализе они совпадали с определёнными различиями между блоками. Это позволяет уменьшить число участков в блоках и соответственно повысить точность остающихся влияний обработки.

Практическая значимость этих экспериментальных методов в малой степени зависела от статистического анализа, но его простота и ясность существенно способствовали всемирному признанию указанных планов. В принципе этот анализ сводился к классическому методу наименьших квадратов, однако достигнутая независимость приводила к исключительной простоте суждения, а сопутствующая оценка погрешности была систематизирована методом *дисперсионного анализа*.

По сути этот метод состоял в разбивке полной суммы квадратов наблюдений на подходящие аддитивные части, содержащие все систематические члены, приписываемые способам обработки, блокам и т. д. После установления этого метода и при наличии соответствующих критериев значимости (в предположении нормальности и однородности дисперсии ошибки), полученных Фишером при выводе и табулировании распределения *дисперсионного отношения*, его можно было применять при более сложных проблемах, как при более сложных и даже не независимых планах эксперимента или проблемах линейной и криволинейной регрессии.

Одним из полезных обобщений было регулирование наблюждённых экспериментальных величин, например, окончательных [амбарных?] сельскохозяйственных урожаев по некоторой заранее наблюденной величине. Этот метод был назван *анализом ковариаций*, хотя этот термин более подходил в случае одновременного анализа двух или более переменных, т. е. при многомерном анализе. Фишер активно участвовал в разработке

этого анализа. Проблемы, относившиеся к многомерности, безусловно встречались и раньше, но Фишер, исходя из своего геометрически [выведенного] распределения множественного коэффициента корреляции, (1928b), снова блестяще применил геометрический подход. Проблемой, над которой трудились Гарольд Хотеллинг в США и Махаланобис в Индии, равно как и сам Фишер, состояла в действенном применении нескольких коррелированных переменных в задачах методов классификации и регрессии.

Имя Фишера особо связывается с понятием дискриминационной функции, т. е. функции переменных, которая позволяет эффективно различать по их измерениям к какой из двух различных совокупностей относится какое-то одно из них.

[2.2.] Генетика. Вклад Фишера в генетику был сравним по значимости с его чисто статистическими трудами, и в нём в такой же степени отразилась его оригинальность и независимость взглядов. В первом десятилетии XX в. менделевская генетика всё ещё оставалась новой научной дисциплиной, и её количественные следствия ещё не были должным образом оценены. Не было ясно, соответствовала ли она дарвиновской теории эволюции видов естественным отбором или хотя бы наблюдаемой наследственности количественных показателей.

Фишер вначале занялся второй, менее важной задачей и в 1918 г. представил глубокое теоретическое исследование корреляции, разделив её на негенетические влияния, аддитивное действие генов и результат иных осложнений, подобных взаимодействию генов и доминантности. Тем самым он сумел доказать соответствие принципов Менделя с наблюдаемой корреляцией между братом и сестрой или родителями и детьми. Затем, в своей книге о генетической теории естественного отбора, он принялся за более важную задачу. [...]

Обозревая прошлое, можно заметить, что полнейшее погружение в собственные столь широкие и фундаментальные исследования заставило Фишера пропустить некоторые существенные теоретические направления. Его пренебрежение чисто математической теорией вероятностей, строго сформулированной Колмогоровым в 1933 г., видимо распространилось на развитие теории случайных процессов, хоть они и были весьма подходящими к некоторым из его собственных проблем в эволюционной генетике. В Англии А. G. McKendrick опубликовал прекрасные работы о случайных процессах в медицине, а Юл – об анализе временных рядов, но Ф., кажется, так и не оценил их. Так, его собственные статьи о статистическом анализе наблюдений во времени иной раз излишне строго придерживались классических и неоправданно ограничительных предположений.

И, оценивая труды Фишера, в дополнение к их широкой сфере необходимо учитывать его подчас совершенно конкретные ошибки, и, что более существенно, его вызванные темпераментом пристрастия в научных спорах. И всё же научные достижения Фишера настолько разнообразны и проникновенны, что подобные

прегрешения не могут ни затмить их блеска, ни снизить его положения как великого учёного XX в.

[3.] Дополнение

После смерти Фишера возрастающий поток трудов по статистике и популяционной генетике указывают и на использование его идей, и на некоторое освобождение от них. Одобрил бы он или нет новые направления в статистике, как, например, возрождение интереса у некоторых специалистов к байесовским методам или в генетике (теории дрейфа генов, которые на молекулярном уровне, скорее схожи с более ранними взглядами Sewall Wright, чем с его собственными), подобное продолжающееся расхождение взглядов в этих оживлённых областях науки не является ни неожиданным, ни вредным. Оно также не умаляет значимости его собственных трудов, которые теперь доступны в собрании сочинений 1971 – 1974 гг.

Примечания

1. В письме Романовскому 5 февраля 1934 г. Фишер (Sheynin 2008, с. 377 – 378) описал этот переходный период, но никаких особых недостатков не указал.
2. Артур Бальфур (1848 – 1930), английский государственный деятель.
3. Ссылка непонятна, и быть может фактически относится к Фишеру. Книга Эддингтона была переиздана в 2007 г.

Библиография

Р. А. Фишер

- 1912**, On an absolute criterion for fitting frequency curves. *Messenger Math.*, vol. 41, pp. 155 – 160.
- 1915**, Frequency distribution of the values of the correlation coefficient etc. *Biometrika*, vol. 10, pp. 507 – 521.
- 1918**, The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinb.*, vol. 52, pp. 399 – 433.
- 1920 – 1945**, *Contributions to Mathematical Statistics*. New York, 1950.
- 1922**, On the mathematical foundations of theoretical statistics. *Phil. Trans. Roy. Soc.*, vol. A222. В книге автора *Contributions* (1920 – 1945/1950, pp. 10.308a – 10.368).
- 1925a**, *Statistical Methods for Research Workers*. См. книгу автора, указанную ниже (1956/1973, 1990).
- 1925b**, Theory of statistical estimation. *Proc. Cambr. Phil. Soc.*, vol. 22. В книге автора *Contributions* (1920 – 1945/1950, pp. 11.699a – 11.725).
- 1928a**, Moments and product moments of sampling distributions. *Proc. Lond. Math. Soc.*, vol. 30. В книге автора *Contributions* (1920 – 1945/1950, pp. 20.198a – 20.237).
- 1928b**, The general sampling distribution of the multiple correlation coefficient. *Proc. Roy. Soc. London*, vol. 121. В книге автора *Contributions* (1920 – 1945/1950, pp. 14.653a – 14.763).
- 1930a**, *The Genetical Theory of Natural Selection*. New York 1958.
- 1930b**, Inverse probability. *Proc. Cambr. Phil. Soc.*, vol. 26. В книге автора *Contributions* (1920 – 1945/1950, pp. 22.527a – 22.535).
- 1934**, Randomisation and an old enigma of card play. *Math. Gazette*, vol. 18, pp. 294 – 297.
- 1935a**, *Design of Experiments*. См. книгу автора, указанную ниже (1956/1973, 1990).
- 1935b**, The fiducial argument in statistical inference. *Annals of Eugenics*, vol. 6, pt. 4. В книге автора *Contributions* (1920 – 1945/1950, pp. 25.390a – 25.398).
- 1938**, *Statistical Tables for Biological, Agricultural and Medical Research*. Шестое издание: New York. 1963. Соавтор F. Yates.
- 1949**, *The Theory of Inbreeding*. Edinburgh – New York, 1965.

1956, *Statistical Methods and Scientific Inference*. В книге автора *Stat. Methods, Exp. Design and Scient. Inference*. Oxford, 1973, 1990. Три книги автора, каждая с отдельной пагинацией.

1971 – 1974, *Collected papers*, vols. 1 – 5. Adelaide.

Другие авторы

Bartlett M. S. (1936), The information available in small samples. *Proc. Cambr. Phil. Soc.*, vol. 32, pp. 560 – 566.

--- (1965), R. A. Fisher and the last 50 years of statistical methodology. *J. Amer. Stat. Assoc.*, vol. 60, pp. 395 – 409.

Eddington A. S. (1935), *New Pathways in Science*. New York.

Neumann J., Morgenstern O. (1944), *Theory of Games and Economic Behavior*. Princeton, 1953, третье издание. Wiley, 1964.

Neyman J. (1937), Outline of a theory of statistical estimation etc. *Phil. Trans. Roy. Soc.*, vol. A236, pp. 333 – 380.

--- (1967), R. A. Fisher (1890 – 1962): an appreciation. *Science*, vol. 156, pp. 1456 – 1462. Примечание: W. G. Cochran.

Pollack I. (1968), Information theory. *IESS*, vol. 7, pp. 331 – 337.

Shannon C. E. (1948), Mathematical theory of communication. *Bell System Techn. J.*, vol. 27, pp. 379 – 423, 623 – 656.

Savage L. J. (1976), On rereading R. A. Fisher. *Annals of Statistics*, vol. 4, pp. 441 – 500 (with discussion).

Sheynin O. (2008), Romanovsky's correspondence with K. Pearson and R. A. Fisher. *Arch. Intern. d'Hist. des Sciences*, vol. 58, No. 160 – 161, pp. 365 – 384.

Thiele T. N. (1903), *Theory of Observations*. London.

Yates F., Mather K. (1963), R. A. Fisher. *Biogr. Mem. Fellows Roy. Soc.*, vol. 9, pp. 91 – 129.

XI

Томассо Сальвенини

Коррадо Джини

Tomasso Salvenini, Gini, Corrado. IES, pp. 394 – 398

[1. Биография]

Труды Коррадо Джини (1884 – 1965) серьёзно повлияли на многие ветви социальных наук. Сын промышленника¹, он родился в Мота-ди-Ливенца в провинции Тревизо. В университете г. Болонья он изучал юриспруденцию и слушал лекции по математике. Широкий охват его научных работ отражён в разнообразии дисциплин (статистика, политэкономия, демография, биометрия, конституционное право и социология), которые он успешно преподавал в университетах Кальяри, Падуи и Рима.

Джини способствовал развитию социальных наук и помимо преподавания. В 1926 – 1932 гг. он был председателем Центрального института статистики, руководил несколькими научными экспедициями, которые изучали демографические, антропометрические и медицинские характеристики определённых этнических групп в Феццане (историческая область в Африке), Палестине, Мексике, Польше и Литве, Калабрии и Сардинии, равно как и процессы ассимиляции групп иммигрантов вообще.

В 1920 г. Джини основал и редактировал международный статистический журнал *Metron* и журнал *Genus* итальянского Комитета по изучению проблем населения. Его достижения были признаны, и ему были присвоены почётные [honoris causa] докторские степени по экономическим наукам Католическим университетом в Милане (1932), по социологии Женевским университетом (1934), и доктора наук, в связи со своим трехсотлетием, Гарвардским университетом (1936)².

В то время, когда Джини начинал свою научную деятельность, основными течениями мысли в итальянской культуре были позитивизм и идеализм. Лишь в ограниченной степени он примыкал к позитивизму. Его работы эклектичны в примечательной степени, что было вызвано систематическим изучением различных наук, – юриспруденции, экономики, статистики, математики и биологии. Джини неизменно рассматривал изучаемые явления на более общем фоне, привлекая к своей непосредственной теме существенные социальные, экономические и демографические элементы.

Джини прекрасно понимал значимость описательной статистики, что заметно по его особо восхищённому отношению к Людовикджи Бодио, но с то же время он тонко ощущал необходимость изучения методов индуктивной статистики и отыскания подходящих критериев для их оценки.

и. По этой причине он изучал труды [Я.] Бернулли, Лексиса и Чубера³, а также итальянских статистиков Angelo Messedaglia и Родольфо Беннини.

[2.] Вклад в статистический метод. Первые труды Джини были посвящены статистической закономерности редких событий⁴. К этому же периоду (1908 – 1911 гг.) относится его первая работа о понятии и измерении вероятности с особым вниманием и приложением к соотношению мужских и женских рождений. Полученные результаты содержатся в его книге (1908), в которой он исследовал обширный эмпирический материал по различным странам. В ней же он разработал теорию дисперсии, т. е. количественной меры разброса или изменчивости в измеренных величинах одного и того же вида⁵. К этой теме он вернулся позже (1940; 1941)

Джини критиковал принципы статистики, чтобы усилить её логические основы и избавить её от наивного эмпиризма. Он стремился поднять статистику до уровня независимой дисциплины⁶, которая, в частности, должна была бы систематически исследовать преимущества и ограниченность различных индексов или описательной статистики при различных видах погрешностей.

На первом заседании Итальянского общества статистики он прочёл доклад *Об опасностях статистики* (1939), в котором предупредил об ошибочности логического обоснования некоторых статистических методов и подчеркнул важность оценивания значимых параметров. По этой теме он также опубликовал статьи (1943; 1947), во второй из которых указал, что неудовлетворительность статистических обозначений может привести к непониманию и ошибкам.

К примеру, понимая $y = 3 + 7x$ как математическое соотношение между переменными x и y , можно с равным успехом написать $x = -3/7 + y/7$ (что означает обратимость). С другой стороны, в статистическом контексте регрессии то же соотношение, понимаемое нестрого, может означать, что ожидание (или среднее значение) случайной величины Y таким образом линейно соотносится со случайной величиной X . Для ясности следует принимать более полные обозначения, подобные нынешнему обычному

$$E(Y|X = x) = 3 + 7x,$$

но вообще

$$E(X|Y=y) \neq -3/7 + y/7.$$

С другой стороны, приняв безусловное среднее

$$EY = 3 + 7EX,$$

мы осуществили частичное обращение.

Совместная статья (1929) была оригинальна в другой области. Задача в ней состояла в отыскании значимого аналога среднего арифметического для чисто качественного распределения. Определив меру отклонения между двумя любыми классами в качественной классификации и принимая минимизацию, Джини показал, как выбрать один из классов как своего рода среднее⁷.

Тема обобщённых средних постоянно интересовала его. И особенно он заботился о её весьма общем выражении и о существенном разнообразии средних значений (1938). В другой статье с тем же соавтором⁸ он подробно обсуждал параметры сдвига для двумерных распределений, в частности также для проблем географии в совместной статье (1933).

Кроме обобщённых средних большое значение имеют меры изменчивости. Джини подробно обсуждал их и уделял особое внимание описательной статистике, не связанной неразрывно с нормальным распределением (1912а). Он ввёл новую меру изменчивости, *среднюю разность*, т. е. по существу ожидаемое абсолютное значение разности X_1 и X_2 независимых случайных величин с одним и тем же распределением. В основном он имел в виду выборочную среднюю разность. Кроме того, он исследовал меры изменчивости качественных распределений и тех, которые относятся к кривым кучности (*concentration curve*) и считаются существенными в экономическом анализе дохода и богатства (1921).

[3.] Нетипичность типичных областей при переписях.

Джини задался целью отыскать выборку подобных типичных областей Италии. Он показал, что недостаточно просто отыскать выборку так, чтобы выборочные средние различных характеристик совпадали с вычисленными для всей Италии. Даже при исследовании только характеристик, использованных при отыскании выборки, она не окажется репрезентативной в смысле изменчивости и концентрации тех же характеристик и к отношениям между ними. По этой причине Джини включил в критерии репрезентативности указанные две стороны распределений характеристик и исследовал условия, которые каждый из предложенных критериев накладывал на другие критерии⁹.

[4.] Теория индексов цен. Эта теория имеет важнейшее значение для экономической статистики, и Джини опубликовал соответствующую классическую работу (1924), в которой решил некоторые логические и технические задачи, связанные с методами исключения [см. ниже]. Проводя различие между простыми и сложными индексами, он показал, как вообще методы составления сложных индексов могут рассматриваться как частные случаи метода исключения (т. е. отделения различных обстоятельств, влияющих на некоторое явление, с целью его сравнения при прочих равных условиях с аналогичными явлениями в другие периоды или местности).

Задача состояла в том, чтобы отличать явление *изменения цены* от явления *изменения количества*¹⁰. Используя общий метод *типичного населения, подразделённого* на четыре специальные

варианта, Джини вывел различные формулы для индексов в то же время предлагая критерии наилучшим образом приспособленные к частным задачам. Аналогичной была его статья (1937).

[5.] Теория распределений. Джини существенно продвинул изучение соотношений между двумя распределениями вероятностей и между двумя случайными величинами, обладающими совместным распределением. Особо важным было следующее.

[5.1.] Трансвариация. Пусть рассматриваются два распределения вероятностей случайных величин X и Y , которые для удобства рассмотрения приняты независимыми. Пусть, далее, $EX < EY$. Вполне вероятно, однако, что $X > Y$, т. е. что знаки разностей между наблюдениями этих величин и между их ожиданиями противоположны. Джини предложил особую меру для оценивания этого явления, назвав его трансвариацией (1953; 1959a), см. также Kruskal (1957).

[5.2.] Расстояние между двумя распределениями. Джини (1914a) одним из первых, рассматривал задачу значимого измерения этого расстояния в целом. Он предложил меру расстояния, основанную на двух кумулятивных функциях распределения; в математическом смысле она являлась метрикой (Fréchet 1947).

[5.3.] Зависимость между двумя случайными величинами с совместным распределением. Джини отметил существенное отличие между *связью* (т. е. любой статистической зависимостью) и *конкорданцией*, согласованностью (т. е. зависимостью, при которой имеет смысл указывать, что обе величины стремятся изменяться в одном и том же направлении). Корреляционное отношение, к примеру, есть мера связи, а коэффициент корреляции – мера конкорданции.

Джини исследовал предложенные ранее меры связи и предложил иные и для количественных, и для качественных случайных величин, см. статью Goodman & Kruskal (1959), в которой даны ссылки на Джини, Weida и Pietra.

[6.] Демография и биометрия. Джини был пионером в исследованиях, соотносящих демографические явления с социальными и биологическими¹¹. Так, он (1949) изучал причины дифференцированной плодовитости и связал явления миграции с более широкими социальными и демографическими соображениями (1946b).

Он (1915 – 1920) также предложил единую программу исследований евгенических и дисгенических влияний войны, уделяя особое внимание на соответствующую смертность [гражданского населения]¹². К демографическим явлениям примыкают исследования Джини (1934) различных популяций, находящихся в стадиях расширения, попятного движения и вымирания.

[7.] Экономические и социологические исследования. С самого начала своей научной карьеры Джини занимался трудной задачей оценки национального богатства. Он конструктивно критиковал метод интервалов деволуции и впервые

непосредственно вычислил богатство частных лиц Италии (1914b). Кроме того, Джини (1959b) методически расположил данные для сравнительного изучения качественного состава богатства и прояснил условия обогащения наций.

Он впервые предложил включать стоимость человеческого капитала в вычисления богатства (Schultz 1968), что позволило указывать преобладающую причину несоответствия во времени и пространстве богатства и экономического благосостояния различных коллективов (1956a).

Стоимость трудящегося, понимаемая как его собственность, становится фактором при оценке богатства, подобно тому, как доля, определяемая трудом, включается в доход (1914a). Эта идея явственно объединяет экономические соображения с логическими и демографическими рассуждениями.

К понятию человеческого капитала по Джини примыкает его циклическая теория населения. Наблюдая различие скоростей размножения социальных классов, Джини сформулировал теорию социального метаболизма, основанную на аналогии с органическим метаболизмом. Высшие классы с невысокой скоростью размножения будут вымирать, если не воспримут членов из низших классов, обладающих более высокой скоростью размножения (1927). Своей теорией Джини намеревался заменить понятие Парето о круговороте элит.

Он (1923) также разработал теорию стадий роста населения, соотнеся их с социальными и экономическими явлениями. В соответствии с ней, в первой стадии своего развития данное население быстро возрастает, затем скорость снижается, становится примерно постоянной и, наконец, начинается спад и возможно даже полное его исчезновение.

В течение первой стадии капитал очень недостаточен, а социальные классы мало отличаются друг от друга. Затем капитал накапливается, классы всё более различаются друг от друга, а вместе с тем различаются и скорости их размножения. С распространением противозачаточных мер ухудшения качества населения можно избежать только влиянием внешних факторов, как, например, иммиграцией молодых людей из популяций, находящихся в стадии демографического расширения.

Этот подход позволяет отличать нормальные экономические процессы от патологических, при которых общество выведено из состояния равновесия, и приступать к экономическим мерам восстановления. Вообще же Джини верил, что от состояния неуравновешенности можно избавиться регулирующей деятельностью политических институтов и экономических организаций.

В соответствии с его точкой зрения (1956b), существенным элементом развития общества является переход от принудительного труда к свободному, а затем и к непринуждённому труду, т. е. к последней стадии психологического процесса подъёма труда от изначального средства жизни к свободной и независимой деятельности, которая обогащает личность.

[8.] Выдающееся положение Джини в развитии итальянской статистики основано, во-первых, на его работе в научной статистике, включающей важную преподавательскую и редакционную деятельность, и, во-вторых, на его усилиях по развитию официальной статистики, и особенно на объединении в едином учреждении всех агентств по сбору данных, и по расширению охвата изучаемого материала.

Краткие сведения об упомянутых лицах

Benini, Rodolfo; Бенини, Родолфо, 1862 – 1956, экономист

Bodio, Luigi; Бодио Людовикджи, 1840 – 1920, экономист и статистик

Messedaglia, Angelo, 1820 – 1901, социолог и статистик

Примечания

1. По другим сведениям, сыном землевладельца.
2. Также по социологии Кордовским университетом в Аргентине (1963 г.), см. Regazzini, см. Johnson, Kotz (1997, pp. 291 – 296).
3. Вряд ли Джини забыл про Лапласа.
4. Это намёк на дискуссию Джини с Борткевичем по поводу закона малых чисел. Ныне в первую очередь следует сказать, что никакого открытия Борткевич вообще не сделал (Sheynin 2008).
5. По контексту никак не следует, что Джини интересовался теорией Лексиса (его *коэффициентом дисперсии*), но так ли было на самом деле?
6. Это замечание следовало дополнить историческим указанием. Известно, что уже Knies (1850) считал статистику самостоятельной наукой.
7. Это непонятно, притом автор не упоминает здесь соавтора Джини.
8. Автор не назвал эту *другую статью*.
9. Это непонятно.
10. И это также непонятно.
11. Пионером в подобных исследованиях был Граунт.
12. Дисгенетическое, т. е. отрицательное влияние вполне ясно. Евгеническим влиянием автор, видимо, посчитал вымирание больных и старых граждан, но ничто не может сравниться с первым.

Библиография

C. Gini

- 1908, *Il sesso dal punto di vista statistico etc.* Milan.
- 1912a, *Memorie di metodologia statistica*, vol. 1. Roma, 1955.
- 1912b, Variabilità e mutabilità etc. *Studio economico-giuridici Univ. Cagliari*, t. 3.
- 1914a, *Di una misura della dissomiglianza fra due gruppi di quantità etc.* Venice.
- 1914b, *L'ammontare e la composizione della ricchezza della nazioni*. Turin, 1962.
- 1915 – 1920, *Problemi sociologici della guerra*. Bologna, 1921.
- 1921, Measurement of inequality of incomes. *Econ. J.*, vol. 31, pp. 124 – 126.
- 1923, *Patologia economica*. Turin, 1952. Пятое издание.
- 1924, Quelques considérations au sujet de la construction des nombres indices des prix etc. *Metron*, t. 4, pp. 3 – 162.
- 1927, *Il neo-organicismo etc.* Catania.
- 1929, [Extension of theory of mean values to qualitative characteristics]. *Metron*, t. 8, No. 1/2, pp. 3 – 209. Соавтор L. Galvani.
- 1933, Sulla selettività delle cause di morte durante l'infanzia. *Metron*, t. 11, No. 1, pp. 163 – 183. Соавторы L. de Berardinis, L. Galvani.
- 1934, Ricerche sulla popolazione. *Scientia*, t. 55, pp. 357 – 373.
- 1937, Methods of eliminating the influence of several groups of factors. *Econometrica*, vol. 5, pp. 56 – 73.
- 1938, Di una formula comprensiva delle medie. *Metron*, t. 13, pp. 3 – 22.
- 1939, [The dangers of statistics]. *Rivista di politica economica*, t. 29, pp. 901 – 924.

- 1940**, Sur la théorie de la dispersion et sur la vérification et l'utilisation des schémas théoriques. *Metron*, t. 14, pp. 3 – 29.
- 1941**, Alle basi del metodo statistico etc. Там же, pp. 173 – 240.
- 1943**, I testi di significatività. *Atti Soc. Ital. di Statistica*, pp. 241 – 279.
- 1946a**, Actualidades demográficas. *Revista internacional de sociologia*, t. 4, pp. 147 – 169.
- 1946b**, Los efectos demográficos de las migraciones internacionales. Там же, pp. 351 – 388.
- 1947**, Statistical relations and their inversions. *Rev. Inst. Intern. Stat.*, t. 15, pp. 24 – 42.
- 1948**, Evoluzione della psicologia del lavoro e della accumulazione. *Moneta e credito*, No. 2. Весь номер.
- 1949**, Vecchie e nuove osservazioni sulle cause della natalità differenziale etc. *Metron*, t. 15, pp. 207 – 358.
- 1950**, Metodologia statistica etc. *Enciclopedia delle matematiche elementari*, t. 3, pt. 2, pp. 245 – 321.
- 1951**, Caractère des plus récents développements de la méthodologie statistique. *Statistica*, t. 11, pp. 3 – 11.
- 1952**, On some symbols that may be usefully employed in statistics. *Bull. Intern. Stat. Inst.*, t. 33, No. 2, pp. 249 – 282.
- 1953**, The measurement of the differences between two quantity groups etc. *Acta genetica et statistica medica*, t. 4, pp. 175 – 191.
- 1955**, Sur quelques questions fondamentales de statistique. *Annales Inst. Henri Poincaré*, t. 14, pp. 245 – 364.
- 1956a**, Valutazione del lavoro e del capitale nell' Economia lavorista. *Rivista bancaria*, t. 12, pp. 522 – 530.
- 1956b**, *Economia lavorista etc.* Turin.
- 1956c**, *Statistical Methods.* Roma, 1966.
- 1958**, Logic in statistics. *Metron*, t. 19, No. 1/2, pp. 1 – 77.
- 1959a**, *Transvariazione.* Roma.
- 1959b**, *Ricchezza e reddito.* Turin.
- 1959c**, Mathematics in statistics. *Metron*, t. 19, No. 3/4, pp. 1 – 9.
- 1968**, *Questioni fondamentali di probabilità e statistica.* Roma.

Другие авторы

- Castellano V.** (1965), Corrado Gini. С библиографией. *Metron*, t. 24, No. 1 – 4.
- Fréchet M.** (1947), Anciens et nouveaux indices de corrélation etc. *Econometrica*, vol. 15, pp. 1 – 30, 374 – 375.
- (1947 – 1948), Le coefficient de connexion statistique de Gini – Salvemini. *Mathematica*, t. 23, pp. 46 – 51.
- (1957), Sur la distance de deux lois de probabilité. *Publ. Inst. de Statistique*, t. 6, pp. 183 – 198.
- Goodman L. A., Kruskal W. H.** (1959), Measures of association for cross-classifications. II. *J. Amer. Stat. Assoc.*, vol. 54, pp. 123 – 163.
- Heyde C. C., Seneta E., редакторы** (2001), *Statisticians of the Centuries.* New York.
- Johnson N. L., Kotz S., редакторы** (1997), *Leading Personalities in Statistical Sciences.* New York.
- Knies C. G. A.** (1850), *Die Statistik als selbstständige Wissenschaft.* Frankfurt/Main, 1969.
- Kruskal W. H.** (1957), Historical notes on the Willcoxon unpaired two-sample test. Там же, vol. 52, pp. 356 – 360.
- Salvemini T.** (1943), La revisione critica di Gini ai fondamenti della metodologia statistica. *Statistica*, t. 3, pp. 46 – 59.
- Schultz T. W.** (1968), Capital, human. *IESS*, vol. 2, pp. 278 – 287.
- Sheynin O.** (2008), Bortkiewicz' alleged discovery: the law of small numbers. *Hist. Scientiarum*, vol. 18, pp. 36 – 48.

ХП

Деннис В. Линдли

Джон Мейнард Кейнс. Работы по статистике

Dennis V. Lindley, Keynes, John Maynard. Contribution to statistics.
IES, pp. 488 – 489

Кейнс окончил университет, став математиком по специальности, и естественно, что его последовавшая диссертация была посвящена математической теме, а именно теории вероятностей. Свою диссертацию он впоследствии развернул в трактат (1921), который оказался его единственным существенным сочинением по этой теме.

Во второй из его пяти частей Кейнс попытался свести фундаментальные теоремы исчисления вероятностей в традиции математических упражнений к логическим формулам Уайтхеда и Рассела, которые оказали мало влияния на последующих авторов. Другая часть книги, историческая и библиографическая, содержала список 600 источников. Страсть Кейнса к сбору информации проявилась здесь в этом восхитительном сборнике, который довёл до времени издания книги предшествовавшие исторические логические изыскания Тодхантера и Лорана¹.

Третьей чертой книги была тонкая критика некоторых точек зрения, принятых в то время и популярных до сего дня. Среди них была идея о вероятности как о субъективной степени доверия к предложению при его заданных основаниях и понятие о вероятности как о пределе частоты при определённом виде бесконечных последовательностей.

Главное из введенного Кейнсом это мысль о вероятности как о примитивной идее, о логическом отношении между предложением и его основанием, относящемся к верности предложения. Таким образом, вместе с субъективистами он считал, что вероятность – соотношение между предложениями и основаниями, но он поддерживал частотников, полагая, что она объективна.

Он весьма тщательно рассмотрел понятие о вероятности, притом в достойном литературном стиле. Для Кейнса цель математической теории состояла в систематизации процессов умозаключений, и он поэтому пытался сформулировать некоторые правила [действий] с вероятностью и развить её исчисление равно как и логические основания статистических доводов.

И точка зрения, и программа Кейнса были важными и существенно повлияли на специалистов по теории вероятностей и статистиков. К сожалению, они были испорчены серьёзным ограничением, а именно отказом от признания, что все вероятности сравнимы друг с другом. Он был готов признать только их частичное упорядочивание, и соответственно отказывался допустить, что численная мера вероятности всегда уместна².

Ramsey (1923 – 1928) заметил, что этот отказ от введения чисел поразителен, поскольку Кейнс был наверняка знаком с работой Рассела о соответствии между упорядочением и числами. А без чисел прогресс труден, если вообще возможен³. В биографическом очерке о Рамзее Кейнс (1933, с. 300) отказался от своих возражений и признал верность точки зрения Рамзея о вероятности в терминах пари. Он признал и довод Рамзея о том, что правила теории вероятностей являются логическими выводами из надлежащего поведения при заключении пари, а не примитивными аксиомами. Интересное развитие работы Рамзея получили у Savage (1954) в том направлении, которое сейчас часто называется байесовской вероятностью и которое всё более влияет на практическую статистику.

В то время, когда Кейнс работал над своим трактатом, Jeffreys (1939), также в Кембридже, развивал аналогичную объективистскую логическую теорию, но он признавал числа, и поэтому продвинулся намного больше. Джеффрис указал, что и он сам, и Кейнс находились под влиянием У. Э. Джонсона, лектора по философии в Кембридже.

Добавление. Работы Рамзея существенно продолжил де Финетти (1970).

Краткие сведения об упомянутых лицах

Johnson W. E., У. Э. Джонсон, 1858 – 1931, логик, учитель Кейнса

Whitehead A. N., А. Н. Уайтхед, 1861 – 1947, математик, логик, философ, вместе с Расселом основал логистическую школу философии математики

Примечания

1. Эти изыскания нам неизвестны.
2. Хорошо известно, что в 1843 г. Курно ввёл философские (экспертные) вероятности. Фактически чуть раньше их предложил Фриз (Шейнин 2013, § 8.1-5).
3. Древняя наука была качественной, но её никак нельзя сбрасывать со счетов.

Библиография

Шейнин О. Б. (2013), *Теория вероятностей. Исторический очерк.* Берлин. Также Google, Oscar Sheynin, Download area.

von Bortkiewicz L. (1923), *Wahrscheinlichkeit und statistische Forschung nach Keynes.* *Nordisk statistisk tidskrift.* Bd. 2, pp. 1 – 23. Перевод в книге *Пятая хрестоматия по истории теории вероятностей и статистики.* Берлин, 2008. Также Google, Oscar Sheynin, Download area.

de Finetti B. (1970 итал.), *Theory of Probability,* vols 1 – 2. New York – London, 1974 – 1975.

Jeffreys H. (1931), *Scientific Inference.* Cambridge, 1957.

Keynes J. M. (1921), *Treatise on Probability.* *Coll. Writings,* vol. 8. London, 1988.

--- (1933), *Essays in Biography.* *Coll. Writings,* vol. 10. London, 1972.

Ramsey F. P. (1923 – 1928), *Foundations of Mathematics and Other Logical Essays.* New York, 1931.

Savage L. J. (1954), *Foundations of Statistics.* New York, 1972.

ХШ

Мервин Муллер

Случайные числа

Mervin E. Muller, Random Numbers. IES, pp. 839 – 847

[1. Предисловие.] Случайными называются нерегулярные числа, порождённые механизмом, см. уточнение ниже.

Применяются они в четырёх основных целях.

1. Чтобы предотвратить избирательное пристрастие при сборе материала выборочных исследований и опытов. Они же при этом обеспечивают знание вероятностной структуры для статистических вычислений.

2. Чтобы представлять себе поведение сложных механизмов или моделей путём симулирования.

3. Чтобы изучать теоретические свойства статистических методов, как например, эффективность оценивания или мощность статистического критерия.

4. Чтобы приближённо решать другие математические задачи.

В каждом указанном варианте случайные числа симулируют наблюдения, отобранные из произвольного распределения вероятностей. Часто бывает полезно представить себе случайные числа как идеализированные последовательные исходы азартных игр, в которых применяются тщательно изготовленные приборы, например, колёса рулетки.

[2.] Определения. Определением случайных чисел служит уточнение вероятностных свойств процесса, который порождает их, а не свойств самих чисел. Сами числа могут быть только частично испытаны и проверены, поскольку они являются лишь конкретными исходами некоторого процесса. Чтобы прояснить различие между процессом и его исходом, рассмотрим подбрасывание монеты. Утверждая, что вероятность выпадения орла при подбрасывании честной монеты равна $1/2$, подразумевают, что до опыта процесс подбрасывания уже привёл к этой вероятности. Очевидно, что при [новом] броске выпадет либо орёл, либо решетка; положение здесь детерминированное, а не вероятностное¹.

При уточнении процесса порождения случайных чисел участвуют два понятия: независимость и определённый вид недетерминизма. При броске монеты независимость подразумевает, что на этот процесс не влияют предыдущие броски, а недетерминизм означает, что нельзя уверенно предсказать исход броска. В случае честной монеты можно только сказать, что вероятность выпадения орла равна $1/2$.

Говорят, что процесс порождает случайные числа в строгом или теоретическом смысле, если получаемые случайные величины

1. Независимы в смысле вероятности.

2. Подчиняются одному и тому же непрерывному равномерному распределению.

Первое означает, что знание ранее порождённых чисел не помогает предсказанию других чисел. Требование равномерного распределения обеспечивает такой вариант нахождения всех порождённых чисел в некотором интервале, при котором вероятность случайному числу оказаться в какой-то его части равна отношению её длины к длине всего интервала.

Процесс порождает случайные (десятичные) числа, если полученные случайные величины независимы и принимают значения 0, 1, ..., 9 с одной и той же вероятностью $1/10$. Они, стало быть, подчиняются дискретному равномерному распределению. Аналогичные определения имеют место для случайных не-десятичных цифр при любом наборе смежных целых чисел и вообще при любом наборе чисел, подчиняющихся дискретному равномерному распределению.

Никакие реальные процессы никак не могут порождать случайных чисел в строгом смысле, потому что они должны неизменно приводить к рациональным числам. Но термин *случайные числа* всё же часто применяется к процессам, порождающим независимые числа, равновероятные и находящиеся на равных расстояниях друг от друга, если их так много, что дискретный процесс можно считать аппроксимацией непрерывного.

Тем не менее, это отличие теории от практики следует иметь в виду. Аналогично, термины *случайные цифры* и *случайные числа* на практике применяются к реальным процессам, которые предположительно никогда не могут *в точности* приводить к независимости или равномерности. Далее, термин *случайные числа* часто применяется в приближённом смысле в любом из указанных выше возможных вариантов. Наконец, эти термины иногда даже обозначают любой процесс, вероятностный или нет, порождающий числа, которые появляются примерно одинаково часто, а в ряду исходов отсутствует какая-либо видимая схема. Дополнительно об этом см. Hammersley & Handscomb (1964), Шрейдер (1962) и Young (1962). Терминологическая заросль не столь запутана, как может показаться, потому что желаемый смысл обычно ясен из контекста.

[3.] Применение случайных чисел. Случайные числа, цифры и т. д. позволяют выстраивать вероятностные выборки, которые в свою очередь допускают применение современных статистических выводов. Если статистическое исследование в обычном смысле должно быть проведено должным образом, элементы выборок следует вероятностно отбирать из выборочной совокупности. В простейшем случае каждому элементу совокупности присваивается номер (например, от 1 до N). Выбираются случайные целые числа в интервале 1, N и соответствующие элементы образуют выборку. На практике обычно применяется выборка без возвращения, и требуются случайные перестановки, см. ниже. Кроме того, на практике обычно применяются структуры выборок различной степени сложности с расслоением, разбиением на группы и т. д.

При планировании экспериментов аналогично применяются случайные целые числа или перестановки, чтобы решить, какую экспериментальную обработку следует применить к различным физическим элементам (мышам, людям, участкам и т. д.). При обоих указанных применениях случайный механизм защищает от избирательного пристрастия (например, от назначения одной и той же экспериментальной терапии самым здоровым испытуемым). Не менее важно и то, что применение случайных механизмов обеспечивает крепкое основание статистическим выводам.

Случайные числа могут быть также очень полезными при изучении моделей сложных физических механизмов или абстрактных теорий, если эти модели включают стохастические элементы, которые подчиняются произвольным распределениям. К примеру, в моделях усталости или обучения можно симулировать действия многих лиц, порождая наблюдения из соответствующего условного распределения и затем применяя полученные таким образом результаты для оценки действительных действий. Здесь по существу производится выборка из условной совокупности для оценивания ожидаемых действий *реальной* совокупности. Случайные числа применяются для получения наблюдений из произвольного распределения вероятностей.

Если наблюдения $\{x\}$ желательно получить из распределения $F(x)$, а $\{u\}$ обозначает значения случайных чисел в единичном интервале, т. е. наблюдения из единичного равномерного распределения, то x выводится порождением значения u и решением уравнения $x = 1/F(u)$. По поводу вывода и необходимых условий см., например, Mood (1950, с. 107). Часто можно и более экономно переходить от случайных чисел к наблюдаемым значениям $F(x)$. Так, существует несколько методов получения наблюдений из нормального распределения (Box & Muller 1958; Muller 1958; Teichroew 1965; Hull & Dobell 1962).

Рассматривалось и применение случайных чисел для решения физических и математических задач, которые представляются совершенно не-вероятностными. Допустим, к примеру, что задана математическая функция и требуется её проинтегрировать, чтобы определить площадь *под* её кривой между определёнными границами независимой переменной. Пусть для удобства искомая площадь расположена внутри единичного квадрата. Порождая пары случайных чисел внутри него и сравнивая их количество под кривой с их общим количеством, мы оценим искомую площадь. Общее описание этого подхода см., например, у Brown (1956, с. 283).

С появлением современных цифровых компьютеров на этот подход к математическим задачам было обращено существенное внимание. Называется он подходом Монте Карло или выборочным, см. Symposium (1956), Hammersley & Handscomb (1964), Шрейдер (1962). И действительно, подход Монте Карло можно применять для исследования теоретических вопросов статистики, которые трудно непосредственно решать

аналитически (Teichroew 1965). Не входя в подробности, скажем, что случайные числа можно применять и для вывода смешанных стратегий в сложных процессах решений, причём это их приложение становится всё полезнее.

[4.] Источники случайных чисел. Для порождения случайных чисел применяются и физические, и математические процессы. Очень трудно определить, обеспечивают ли они удовлетворительные результаты. Близость реальных процессов и условных моделей измерялась статистическими критериями довольно обычного назначения, которые в лучшем случае обеспечивали только частичный ответ. До последнего времени основным источником случайных чисел были опубликованные таблицы, составленные преимущественно при помощи физических устройств. Некоторые из этих таблиц были исправлены путём математической обработки. Среди использованных физических процессов упомянем

1. Извлечение шарков (билетиков) из урны с возвращением и перемешиванием после каждого тиража.

2. Подача световых сигналов через неопределённые промежутки времени на вращающийся диск, разделённый на пронумерованные секторы.

3. Регистрацию электронных импульсов, предположительно случайно распределённых во времени, и преобразованных в цифровую форму.

4. Подсчёт числа исходящих импульсов, порождённых радиоактивным распадом кобальта 60, в течение постоянных интервалов времени.

Впрочем, в качестве источника случайных чисел физическим процессам присущи недостатки. Порождаемые числа должны быть сохранены, иначе же они оказываются недоступными, и при необходимости проверки или повторения вычислений их использованная конкретная последовательность не может быть получена вновь. Кроме того, ранее применявшиеся типы физических процессов было трудно поддерживать и использовать. Стало интереснее применять математические процессы, которые можно программировать для компьютеров, потому что до сих пор обширные таблицы невозможно экономно вводить и сохранять в их быстрой памяти. Большинство таких математических процессов применяют рекуррентные соотношения, так что некоторый член последовательности определяется одним или более предыдущими.

Эти математические процессы являются детерминированными, т. е. каждый член порождаемой последовательности определяется конкретной функцией предыдущих. По этой причине легко возражать против их применения и предпочитать физические процессы, однако, см. ниже, эти математические методы серьёзно обоснованы практическими соображениями.

При некоторых методах исходный физический процесс считается начальными *случайными* числами и затем исправляется математическими соотношениями, подобными тем,

которые предложили Horton & Smith (1949). Указанная сложная рандомизация представляется весьма перспективной.

[5.] История. Таблицы случайных чисел и цифр появились сравнительно недавно, в 1927г. Это удивительно, потому что для исследования вероятностных проблем механические методы и устройства для тасования карт применяли уже почти столетие раньше. В 1908 г. [ссылка отсутствует] Gosset [Student] сообщил об изучении теоретических свойств различных статистик при помощи выборок. Он исследовал распределение стандартного отклонения и отношения среднего к нему в выборках при помощи тасования и *случайного* извлечения из пачки в 3000 карт. О его статье и ранних выборочных исследованиях с применением игральные кости для оценивания корреляции между событиями, например, о Darbishire (1907) и Weldon (1906), см. Techroew (1965).

Однако, после того как было признано, что случайные числа могут быть надёжным и экономичным средством при отборе выборок, оставалось ещё отыскать удовлетворительный метод их порождения. Пирсон подтолкнул Типпета (Tippet 1927) к составлению таблиц случайных цифр, а в предисловии к этой таблице упомянул об истории их составления. Типпет обнаружил, что извлечение пронумерованных карт из мешка оказалось неудовлетворительным и получил лучшие результаты при составлении таблицы 41 600 цифр путём их отбора из отчёта о переписи 1925 г.

Таблицы Типпета несколько раз исследовали, а затем при помощи иных методов были составлены более обширные таблицы. После появления перфораторных устройств, вычислительных машин и компьютеров были рассмотрены другие методы порождения случайных чисел. Интерес здесь сосредоточился на дополнительных мерах исследования пригодности получаемых при этом чисел.

Понимая, что в длинных последовательностях случайных чисел можно подметить закономерности, против появления которых иногда могли бы быть выдвинуты возражения, Kendall & Smith (1938, с. 153) предложили критерии локальной случайности, чтобы судить о пригодности набора чисел для определённой цели. Против этого нововведения можно возражать, потому что каждая числовая последовательность содержит одну или более систем подпоследовательностей, которые могут появляться или не появляться в некотором желательном порядке. Возникающие практические и философские проблемы не решены, и некоторые усомнились даже в самих основаниях теории вероятностей. Так, Spenser Brown (1957) основательно обсудил очевидный парадокс, возникающий при исключении некоторых расположений чисел.

Первый из рассмотренных ниже математических процессов (middle square method) предложил, видимо, von Neumann (1951), за которым последовал Lehmer (1951) с методом мультипликативной сравнимости. Заметим, кстати, что хоть Фишер до составления (в соавторстве) таблицы (1938) не участвовал в непосредственной разработке методов порождения

случайных чисел, его предложение применять рандомизацию при планировании эксперимента несомненно оказало сильное влияние.

[6.] Таблицы случайных цифр. После таблицы Типпета (1927) появилось несколько обширных таблиц. Упомянутая таблица Fisher & Yates (1938) содержала 15 000 цифр и была основана на таблице логарифмов. В таблице Kendall & Smith (1938; 1939) было уже 100 000 цифр, основанных на отсчётах по вращающемуся диску. В сочетании со своей таблицей эти авторы предложили четыре критерия случайности и понятие о локальной случайности (Brown 1951). Несколько авторов исследовали эти три таблицы. Good (1953) указал на необходимость видоизменения одного из их критериев.

Несколько других таблиц было предложено в период от 1939 г. до середины 1950-х годов. Самую амбициозную и обширную разработали сотрудники корпорации RAND, которые использовали электронный генератор импульсов и метод сложной рандомизации, см. выше. Вначале эти таблицы вышли ограниченным тиражом, но были также размещены на перфокартах. Затем появилась таблица RAND (1955), которая теперь является более или менее стандартной. Включённые в неё отклонения от нормальности были выведены при помощи перевода случайных чисел в наблюдения на основе единичного нормального распределения по указанному выше методу.

Появилось и несколько ранних таблиц таких же отклонений, основанных, к примеру, на таблице Типпета. Обширный список таблиц случайных цифр и случайных наблюдений, основанных на иных распределениях, см. у Greenwood & Hartley (1962, § 13). Каждая таблица случайных цифр содержит наставление по её применению. Существуют, однако, нерешённые проблемы о последствиях повторного использования тех же самых таблиц.

Выше мы указали, что математические методы для получения примерно случайных чисел были введены для удобства и эффективности. При использовании последовательности случайных чисел часто желательно вывести точно ту же последовательность заново, чтобы сравнить друг с другом результаты родственных вычислений. При всех современных математических процессах порождения этих чисел они могут быть воспроизведены. Это означает, что по заданной начальной точке процесса та же самая последовательность может быть получена при его повторении.

Предложенные математические процессы отражают тот факт, что электромеханические калькуляторы и ранние компьютеры имели очень ограниченную память и обычно могли в данный момент производить только одну арифметическую или логическую операцию. Было поэтому нерационально загружать большие объёмы случайных чисел, а для экономии памяти и машинного времени процессы их порождения на каждом шагу обычно довольствовались одной или двумя предыдущими числами процесса и соединяли их при помощи простых операций. Предстоит ещё убедиться в том, будут ли введены более сложные

процессы порождения, полнее использующие успехи в скорости и возможностях компьютеров.

[7.] Математические процессы порождения. Математические процессы детерминированного получения последовательностей чисел не могут порождать последовательностей истинно случайных чисел. И всё-таки эти процессы практически предпочтительнее случайных, потому что последние, хоть это и маловероятно, могут порождать весьма нежелательные последовательности, а потому и сомнительные выборки или планы экспериментирования.

С другой стороны, иногда можно придумать детерминированный процесс, порождающий достаточно беспорядочно расположенные числа, удовлетворяющие критериям примерной, но существенной для применения случайности. Можно спланировать процессы так, чтобы придать некоторым числовым последовательностям особую значимость или вес. Эти методы подобны методам расслоения в выборочных исследованиях, и их рассматривал, например, Brown (1956), Teichroew (1965) и Н. А. Meyer (Symposium 1956). Hammersley & Handscomb (1964, с. 27) применяют термин *квазислучайные числа*, если известно, что их последовательность не случайна, но обладает определёнными желательными статистическими свойствами.

Чтобы предотвратить недоумение при истолковании чисел, порождённых детерминированным математическим процессом, и используемых как случайные, мы будем говорить о процессе, порождающем псевдослучайные числа. Указанные числа таким образом больше связаны с целью, а не действием процесса, пригодность которого следует поэтому оценивать по отношению к каждому определённому применению.

Предположим, к примеру, что при каком-то применении необходимо порождать чётные и нечётные числа с одной и той же вероятностью. Пусть, далее, существует метод порождения статистически независимых чисел с вероятностями r и $q = 1 - r$ того, что они окажутся нечётными и чётными. Желательно, чтобы $r = q$, и этого можно достичь при помощи следующего устройства. Будем порождать и применять числа попарно. Пары, в которых оба числа нечётны (вероятность равна r^2) или чётны (вероятность q^2) не применяются. Если же пара состоит из нечётного и чётного числа, назовём её появлением нечётности, а если из чётного и нечётного, – появлением чётности. Вероятности этих событий совпадают, так что обе условные вероятности равны $1/2$.

Если процесс порождения выдаёт p -значные числа в данной системе счисления (725 – трёхзначное число в десятичной системе), то наибольшее количество отличающихся друг от друга чисел равно основанию системы счисления в степени p (10^p в примере выше). Таким образом, количество различных чисел в трёх разрядах не более 1000. Отсюда следует, что то же ограничение имеет место для псевдослучайных чисел, но оно не обязательно оказывается существенным. Если только

порождается достаточно длинная последовательность p -значных чисел, то некоторые числа будут повторяться.

И если такая последовательность, начиная с некоторого числа, состоит из повторяющихся без разрывов числовых циклов, то длина цикла называется периодом процесса. Ранние методы порождения придумывались таким образом, чтобы при заданном p период последовательности оказывался как можно длиннее. Частично это решение было обосновано убеждением в том, что так легче обеспечить достаточную беспорядочность чисел. Период данного процесса может быть и короче, и длиннее количества различных возможно порождаемых p -разрядных чисел.

Определение периода процесса, если он существует, может оказаться весьма затруднительным и возможно зависящим от того, сколько предыдущих членов последовательности используется для порождения следующего члена. Часто ощущается, что важно иметь период, длиннее количества псевдослучайных чисел, которые требуются при некотором приложении или эксперименте. Кроме того, необходимо быть уверенным в том, что процесс не содержит одного и того же числа слишком часто. Далее, некоторые процессы следует проверять, не вырождаются ли они, т. е. не начинают ли они повторять одну и ту же короткопериодическую числовую схему.

Несмотря на детерминированный характер процесса, иногда очень трудно исследовать его поведение, можно лишь проверять фактические исходы при помощи статистических критериев. Так, трудным оказалось исследование поведения middle square метода Неймана. Впрочем, метод Леймера (Lehmer) гораздо легче исследовать при помощи теории чисел, и у него имеются общие черты со многими позднее предложенными методами, см. ниже. Hull & Dobell (1962) составили весьма тщательный обзор всех математических методов и исследований, проведенных до середины 1962 г.

[7.1.] Мультипликативная сравнимость. Два целых числа, A и B , сравнимы по модулю M , если при делении на него образуется один и тот же остаток; запись: $A \equiv B \pmod{M}$.

Последовательность x_1, x_2, \dots может быть порождена мультипликативной сравнимостью, подобной $x_{n+1} \equiv kx_n \pmod{M}$ следующим образом. Пусть k – заданная постоянная, умножаемая на первый член последовательности, x_1 . Остаток от деления произведения на M считается её следующим членом, x_2 , и т. д. Этот процесс исследовался по поводу выбора k и x_1 относительно M с целью получения длинного периода.

Lehmer упомянул, что соотношение, приведенное выше, обладает неблагоприятной чертой: цифры, находящиеся в правом разряде порождённых чисел, неудовлетворительны в качестве случайных. Для преодоления этого недостатка и в попытке получить период, равный по длине M , Ротенберг [ссылка отсутствует] предложил воспользоваться соотношением $x_{n+1} \equiv kx_n + c \pmod{M}$. Однако, выбор k и c относительно M влияет не

только на период, но и на меру корреляции между членами последовательности.

[7.2.] *Другие рекуррентные соотношения* были также предложены, как, например, использующие два предыдущих члена вместо одного (Hull & Dobell 1962), как и метод, предложенный Ротенбергом, но многое ещё предстоит выяснить прежде, чем поведение этого метода будет полностью понятно. Были, правда, получены псевдослучайные числа, выдержавшие проверку статистическими критериями, но конкретные разряды цифр в этих числах оказались неудовлетворительными.

Можно, однако, разумно вывести псевдослучайные числа, подвергая их сложному процессу рандомизации, подобному тем, которые предложили Horton & Smith (1949) или Walsh (1949). Периоды иррациональных чисел, например, π и e , бесконечны, и поэтому были изучены разряды их цифр. Так, Metropolis и др. (1950) исследовали некоторые свойства первых 2000 цифр этих чисел, а Stoneham (1965) изучил первые 60 000 цифр числа e . Но важно иметь в виду, что, даже применив большое число разумных критериев случайности, нельзя ожидать, что какая-либо данная последовательность чисел или цифр выдержит все подобные критерии при произвольном уровне статистической значимости.

[8.] Случайные перестановки. Они предоставляют возможность особого применения случайных цифр при планировании эксперимента. *Перестановка* указывает, что набору элементов назначен определённый порядок следования. Случайные перестановки применяются, например, для того, чтобы случайно расположить n различных элементов в n мест на прямой.

Говорят, что процесс порождает случайные перестановки объёма n , если каждая из $n!$ различных возможных перестановок порождается независимо с одной и той же вероятностью. Вероятность наблюдения некоторой определённой перестановки окажется равной $1/n!$. Такие перестановки могут быть следующим образом основаны на случайных цифрах.

Пусть n есть число из p разрядов, тогда при соблюдении некоторых ограничений подобные целые числа могут быть отобраны из таблицы случайных цифр. Именно, если отобранное число $\leq n$ и не было отобрано ранее, оно должно быть включено в перестановки, в противном же случае исключено из них. Таким же образом этот процесс продолжается, пока все n мест в перестановках не будут заполнены.

Этот метод осуществим при очень небольших n , но есть и лучшие методы, которые не отклоняют такого большого числа случайных цифр (Cochran & Cox 1950, с. 569; Fisher & Yates 1938, с. 34; Walsh 1957, с. 355).

George W. Snedecor [ссылка отсутствует] предложил составлять таблицы случайных перестановок, и подобные таблицы размером 9 и 16 можно найти у Cochran & Cox (1950, с. 577) и Moses & Oakford (1963), см. также Greenwood & Hartley (1962, с. 460).

[8.1.] Безвозвратные выборки. Они представляют собой вид вероятностных выборок из конечной совокупности, каждый

элемент которой обладает одной и той же и независимой вероятностью быть отобранной, но лишь однократно. Случайные перестановки размером n представляют собой крайний вариант безвозвратной выборки, при котором вся совокупность размером n в определённом порядке включена в выборку. Для безвозвратного отбора выборки размером r из совокупности размером n можно применить случайные перестановки размером n , если использовать только r их первых чисел.

Есть и другие методы применения случайных чисел для отбора безвозвратных выборок. Так, Fan и др. (1962, с. 387) предложили методы, приспособленные для компьютеров, а Jones (1959) изучил непосредственное применение случайных цифр для подсчёта количества необходимых чисел для отбора выборки заданного размера с учётом необходимости отбрасывания повторных и нежелательных чисел.

Библиография

Шрейдер Ю. А., редактор (1962), *Метод статистических исключений (метод Монте Карло)*. М.

Box G. E. P., Muller M. E. (1958), Note on the generation of random normal deviations. *Annals Math. Statistics*, vol. 29, pp. 610 – 611.

Brown G. W. (1951), History of RAND's random digits: summary. В книге Householder A. S., редактор, *Monte Carlo Method*, pp. 31 – 32. U. S. Nat. Bureau Standards, Applied Math. Series. Washington.

--- (1956), Monte Carlo Methods. В книге Beckenbach E. F., редактор, *Modern Math. for the Engineer*. New York, pp. 279 – 303.

Cochran W. G., Cox Gertrude M. (1950), *Experimental Designs*. New York, 1957.

Darbishire A. D. (1907), *Some Tables for Illustrating Statistical Correlation*. *Mem. & Proc. Manchester Literary & Phil. Soc.*, vol. 51, No. 16.

Fan C. T., Muller M. E., Rezucha I. (1962), Development of sampling plans etc. *J. Amer. Stat. Assoc.*, vol. 57, pp. 387 – 402.

Fisher R. A., Yates F. (1938), *Statistical Tables for Biol., Agric. and Med. Research*. New York – Edinburgh, 1963. Шестое издание.

Good I. J. (1953), The serial test for sampling numbers etc. *Proc. Cambr. Phil. Soc.*, vol. 49, pp. 276 – 284.

Greenwood J. A., Hartley H. O. (1962), *Guide to Tables in Math. Statistics*. Princeton.

Hammersley J. M., Handscomb D. C. (1964), *Monte Carlo Methods*. New York.

Horton H. B., Smith R. T. (1949), Direct method for producing random digits etc. *Annals Math. Statistics*, vol. 20, pp. 82 – 90.

Hull T. E., Dobell A. B. (1962), Random number generators. *SIAM Rev.*, vol. 4, pp. 230 – 254.

Jones H. L. (1959), How many of a group of random numbers will be usable etc. *J. Amer. Stat. Assoc.*, vol. 54, pp. 102 – 122.

Kendall M. G., Smith B. B. (1938), Randomness and random sampling numbers. *J. Roy. Stat. Soc.*, vol. A101, pp. 147 – 166.

--- (1939), *Tables of Random Sampling Numbers*. Cambridge.

Lehmer D. H. (1951), Math. methods in large-scale computing units. В книге *Proc. Second Symposium on Large-Scale Digital Calculating Machinery 1949*. Cambridge, Mass., pp. 141 – 146.

Metropolis N. C., Reitwiesner G., von Neuman J. (1950), Statistical treatment of values of first 2000 decimal digits of e and π . *Math. Tables and Other Aids to Computers*, vol. 4, pp. 109 – 111.

Mood A. M. (1950), *Intro. to the Theory of Statistics*. New York.

Moses L. E., Oakford R. V. (1963), *Tables of Random Permutations*. Stanford, Cal.

- Muller M. E.** (1958), Inverse method for the generation of random normal deviates etc. *Math. Tables and Other Aids to Computers*, vol. 12, pp. 167 – 174.
- RAND Corp** (1955), *Million Random Digits with 100,000 Normal Deviates*. Glencoe, Ill.
- Spenser Brown G.** (1957), *Probability and Statistical Inference*. London.
- Stoneham R. G.** (1965), Study of 60,000 digits of transcendental e . *Amer. Math. Monthly*, vol. 72, pp. 483 – 500.
- Symposium** (1956), *Symposium on Monte Carlo Methods 1954*. Редактор Н. А. Meyer. New York.
- Teichroew D.** (1965), History of distribution sampling prior to the era of the computer etc. *J. Amer. Stat. Assoc.*, vol. 60, pp. 27 – 49.
- Tippett L. H. C.** (1927), *Random Sampling Numbers*. Cambridge, 1952.
- von Neumann J.** (1951), Various techniques used in connection with random digits. В книге *Monte Carlo Method*. Редактор А. S. Householder. U. S. Nat. Bureau Standards, Applied Math. ser., pp. 36 – 38. Washington.
- Walsh J. E.** (1949), Concerning compound randomization etc. *Annals Math. Statistics*, vol. 20, pp. 580 – 589.
- (1957), An experimental method for obtaining random digits and permutations. *Sankhya, Indian J. Stat.*, vol. 17, pp. 355 – 360.
- Weldon W. F. R.** (1906), Inheritance in animals and plants. В книге Strong T. B., редактор, *Lectures on the Method of Science*. Oxford, pp. 81 – 109.
- Young F. H.** (1962), *Random Numbers, Mathematical Induction, Geometric Numbers*. Boston.

[9.] Добавление

Теоретические основания и приложения методов порождения случайных чисел продолжали вызывать серьёзный интерес. Теоретические исследования были связаны с сутью случайности и случайных чисел, с понятием о качестве программ проверки генераторов псевдослучайных чисел и программ вычислений. Knuth (1969) и Jansson (1966) представили отличные сводки по указанному вопросу и произвели оригинальные исследования. Halton (1970) составил [соответствующий] обзор с упором на метод Монте Карло, а прекрасную хронологическую и классифицированную библиографию за 1927 – 1971 гг. о порождении случайных чисел и их испытании опубликовал Soweу (1972).

В теоретическом плане заметно весьма интересное объединение дисциплин об основаниях теории вероятностей (Колмогоров, Поппер, Мизес и Вальд), логики (проблемы совместимости, сложности и вычисляемости) и теории информации (Чёрч, Гёдель, Шённон и Тьюринг). Многое, проделанное в этом направлении, описал Knuth (1969).

Некоторые авторы связывают определение случайности со сложностью метода компьютерных вычислений (алгоритма), и этот подход включает предложение о том, что последовательность p чисел случайна, если при заданных первых $(p - 1)$ числах сложность (беспорядочность) метода определения p -го числа максимальна. Особо интересны здесь работы Chaitin (1966), его же популярная сводка (1975) и комментарий Gardner (1968), а также статьи Good (1969), Martin-Löf (1966; 1969) и Stentiford & Lewin (1973).

Разработка и оценка методов порождения псевдослучайных чисел продолжала вызывать интерес. Существенным оказалось введение нового класса критериев, ныне называемых

спектральными и решётчатыми, для оценки распределения n точек

$$X_{i+1}, X_{i+2}, \dots, X_{i+n}, i = 0, 1, \dots$$

Эти работы были вызваны появлением спектрального критерия (Coveyou & MacPherson 1967), основанного на рядах Фурье, для исследования распределения наблюдений в n -мерном пространстве, $2 \leq n \leq 10$. Позднее стало ясно, что непосредственный решётчатый анализ более понятен, чем спектральный критерий, основанный на рядах Фурье, см., в частности, Coveyou (1970), Beyer (1972) и Beyer и др. (1971), а также Good & Gover (1967).

Конгруэнтные рекуррентные отношения, применяемые для порождения псевдослучайных чисел, оставались теоретически интересными. К примеру, Jansson (1964) определил результаты [?] о коррелированности порождённой последовательности, см. также Dieter & Ahrens (1971). Среди других интересных работ назовём Ahrens и др. (1970), Dieter (1971; 1972), Gustavson & Liniger (1970), Lewis и др. (1969), Marsaglia (1970; 1972) и Miller & Prentice (1968).

Спектральные и решётчатые критерии сложны и требуют существенных вычислений. Их применение, однако, было обнадеживающим, потому что вычислительные методы, уже оказавшиеся весьма сомнительными, как, например, применение мультипликативных соотношений с использованием множителя, состоящего из сумм степеней числа 2,

$$\sum_{s=0}^m 2^i s,$$

не выдержали этих критериев.

Далее, конгруэнтные рекуррентные соотношения, которые оказались удовлетворительными, продолжают оставаться приемлемыми, однако сейчас их свойства ощущаются гораздо лучше. Другие критерии указали, что метод middle square не следует считать обоснованным.

Проблемы критериев случайности явно заслуживают большего внимания, и оно им действительно будет уделяться. Если неспециалист начинает всерьёз изучать проблемы случайности или генераторов псевдослучайных чисел, у него в скором времени легко создастся впечатление о сложности и неопределённости этих тем. И на самом деле, неопределённость может легко возникнуть, причём на различных уровнях понимания. Книга Porper (1935) в какой-то степени помогает представить серьёзность этих затруднений, поскольку они относятся к приемлемому основанию теории вероятностей. Вычислительный метод порождения псевдослучайных чисел хорош для конкретного применения (независимо от его значимости при других приложениях), поскольку он соответствует требованиям

примерной статистической независимости наблюдений или их примерной равномерности.

Порождение случайных перестановок и порядковых статистик продолжает быть интересным и привлекает внимание, особо см. алгоритм перестановок в Durstenfeld (1964) и Pike (1965), а Knuth (1969) предлагает другие алгоритмы приложения [чего?]. Rabinowitz & Berenson (1974) обзоревают методы, относящиеся к порядковым статистикам и Ord-Smith (1970) обсуждает порождение наборов перестановок, не обязательно в случайном порядке.

Продолжалось и изучение применений псевдослучайных чисел для непосредственного или косвенного порождения наблюдений при распределениях, отличных от равномерного (например, при нормальном законе), см. соответствующие ссылки у Jansson (1966), Knuth (1969) и Soweу (1972). См. также Bell (1968), Brent (1974), Chay и др. (1975) и Forsythe (1972) по поводу недавних методов порождения нормальных отклонений. Wallis & Roberts (1956, с. 334 – 337 и 631) представили практические советы по применению таблиц случайных чисел.

Интересен комментарий о роли случая и случайности в социальных событиях с библейских времён, см. Fienberg (1971). Он, в частности, исследовал применение рандомизации в отборочной призывной кампании 1970 г. в США¹. Fienberg также указывает, что подобный призыв был, видимо, использован в Австро-Венгрии между 1889 и 1914 гг. и обсуждает, почему в приложениях, требующих вероятностную структуру, необходимы научные определения (независимости и равномерности), аналогичные приведенным здесь, но отличающиеся от обычного повседневного.

Примечания

1. Об отборочной призывной кампании в России упомянул в одном из своих рассказов Шолом Алейхем.

Библиография

- Ahrens J. H., Dieter U., Grube A. (1970), Pseudo-random numbers etc. *Computing*, vol. 6, pp. 121 – 138.
- Bell J. R. (1968), Algorithm 334, normal random deviates. *Communications Assoc. Computing Machinery*, vol. 11, p. 498.
- Beyer W. A. (1972), Lattice structure and reduced bases of random vectors etc. В книге Zaremba S. K., редактор, *Applications of Number Theory to Numerical Analysis*. New York, pp. 361 – 370.
- Beyer W. A., Roof R. B., Williamson D. (1971), The lattice structure of multiplicative congruential pseudo-random vectors. *Math. of Computation*, vol. 25, pp. 345 – 363.
- Brent R. P. (1974), Algorithm 488, a Gaussian pseudo-random generator. *Communications Assoc. Computing Machinery*, vol. 17, pp. 704 – 706.
- Chaitin G. J. (1966), On the length of programs for computing finite binary sequences. *J. Assoc. Computing Machinery*, vol. 13, pp. 547 – 569.
- (1975), Randomness and math. proof. *Scient. American*, vol. 226, May, pp. 47 – 52.
- Chay S. C., Fardo R. D., Mazumdar M. (1975), On using the Box – Muller transformation etc. *Applied Statistics*, vol. 24, pp. 132 – 135.
- Coveyou R. R. (1970), Random number generation is too important etc. *Studies in Applied Math.*, vol. 3, pp. 70 – 111.

- Coveyou R. R., MacPherson R. D.** (1967), Fourier analysis of uniform random number generators. *J. Assoc. Computing Machinery*, vol. 14, pp. 100 – 119.
- Dieter U.** (1971), Pseudo-random numbers. The exact distribution of pairs. *Math. of Computation*, vol. 25, pp. 855 – 883.
- (1972), Statistical interdependence of pseudo-random numbers etc. В книге Zaremba S. K., редактор, *Applications of Number Theory to Numerical Analysis*. New York, pp. 287 – 317.
- Dieter U., Ahrens J. H.** (1971), Exact determination of serial correlations of pseudo-random numbers. *Numerische Math.*, vol. 17, pp. 101 – 123.
- Durstenfeld R.** (1964), Algorithm 235. Random permutations. *Communications Assoc. Computing Machinery*, vol. 7, p. 420.
- Fienberg S. E.** (1971), Randomization and social affairs: the 1970 draft lottery. *Science*, vol. 171, pp. 255 – 261.
- Forsythe G. E.** (1972), Von Neumann's comparison method etc. *Math. of Computation*, vol. 26, pp. 817 – 826.
- Gardner M.** (1968), On the meaning of randomness etc. *Scient. American*, vol. 219 July, pp. 116 – 121.
- Good I. J.** (1969), How random are random numbers? *Amer. Statistician*, vol. 23, No. 4, pp. 42 – 45.
- Good I. J., Gover T. N.** (1967), Generalized serial test and binary expansion of $\sqrt{2}$. *J. Roy. Stat. Soc.*, vol. A130, pp. 102 – 107; vol. A131, 1968, p. 434.
- Gustavson F. G., Liniger W.** (1970), A fast random number generator etc. *Computing*, vol. 6, pp. 221 – 226.
- Halton J. H.** (1970), Retrospective and prospective survey of the Monte Carlo method. *SIAM Rev.*, vol. 12, pp. 1 – 63.
- Jansson B.** (1964), Autocorrelations between pseudo-random numbers. *Nordisk Tidskr. f. Informationsbehandling*, vol. 4, pp. 6 – 27.
- (1966), *Random Number Generators*. Stockholm.
- Knuth D. E.** (1969), *The Art of Computer Programming. Seminumerical Algorithms*, vol. 2. Reading, Mass.
- Lewis P. A. W., Goodman A. S., Miller J. M.** (1969), A pseudo-random number generator for System/360. *IBM Systems J.*, vol. 8, pp. 136 – 146.
- Marsaglia G.** (1970), Regularities in congruential random number generators. *Numerische Math.*, vol. 16, pp. 8 – 10.
- (1972), Structure of linear congruential sequences. В книге Zaremba S. K., редактор, *Applications of Number Theory to Numerical Analysis*. New York, pp. 249 – 285.
- Martin-Löf P.** (1966), Definition of random sequences. *Information and Control*, vol. 9, pp. 602 – 619.
- (1969), Algorithms and randomness. *Rev. Intern. Stat. Inst.*, vol. 37, pp. 265 – 272.
- Miller J. C. P., Prentice M. J.** (1968), Additive congruential pseudo-random number generators. *Computer J.*, vol. 11, pp. 341 – 346.
- Ord-Smith R. J.** (1970), Generation of permutation sequences, pt. 1. *Computer J.*, vol. 13, pp. 152 – 155.
- Pike M. C.** (1965), Remark on algorithm 235. Random permutation. *Communications Assoc. Computing Machinery*, vol. 8, p. 445.
- Popper K. R.** (1935 нем.), *Logic of Scientific Discovery*. New York – London, 1959.
- Rabinowitz M., Berenson M. L.** (1974), Comparison of various methods of obtaining random order statistics etc. *Amer. Statistician*, vol. 28, No. 1, pp. 27 – 29.
- Sowey E. R.** (1972), Chronological and classified bibliography on random number generation and testing. *Intern. Stat. Rev.*, vol. 40, pp. 355 – 371.
- Stentiford F. W. M., Lewin W. W.** (1973), Evolutionary approach to the concept of randomness. *Computer J.*, vol. 16, pp. 148 – 151.
- Wallis W. A., Roberts H. V.** (1956), *Statistics: a New Approach*. Glencoe, Ill.

XIV

Натан Кейфиц

Государственная статистика

Nathan Keyfitz, Government statistics. IES, pp. 413 – 425

Для современного государства государственная¹ статистика является существенной частью более общей информационной системы. Первые составители статистики не отличали достаточно чётко количественных фактов от иных, и при принятии решений частные и общественные объединения всё ещё дополняют количественные данные информацией другого рода. Эти данные незаменимы и в централизованных, и в плюралистических обществах; они важны как для обсуждений в самом правительстве, так и для фирм, государственных учреждений и других производственных объединений. Действенность рыночного механизма в свободном обществе зависит от качества информации, на основе которой принимают решения его отдельные элементы.

Экономическое планирование в Западной Европе в большой степени состоит в предоставлении дополнительной информации частным предприятиям. Жёстко контролируемая централизованная система, если только она не окажется бессознательной жертвой своего собственного сосредоточения власти, в первую очередь нуждается в обратной связи, стержнем которой является собираемая ей статистика.

[1.] Статистика и государство

История, нынешнее состояние и ещё не решённые проблемы государственной статистики отражают обстоятельства жизни современных национальных государств. Правительственная администрация является основным пользователем официальной статистики, существование и действенность которой составляют главное необходимое условие её работы.

В классические времена и в древних империях Среднего Востока время от времени предпринимались попытки подсчёта населения или товаров, но по существу между ними [между их результатами] и современными официальными сборниками нет никакой непрерывной связи.

[1.1.] Современная история. Как очень многое иное в современном государстве, его статистическая система возникла примерно в период Французской революции. Одно из ранних решений Учредительного собрания касалось публикации статистического отчёта о ресурсах страны, который подготовил Лавуазье. В 1800г. в Париже было создано Бюро статистики. 22 июля 1791 г. Учредительное собрание приняло закон о проведении переписи (Faure 1918, с. 277), которая и была проведена в том же году. В циркуляре 15 фруктидора VI года (очевидно, 1 сентября 1799 г.) Франсуа де Нёфшато с характерным революционным преувеличением назвал перепись

мерой силы, источником богатства и политическим барометром мощи государства (там же, с. 284).

Британская нация прорезалась не так неистово, и её статистическая система образовалась постепенно. Ранними вехами были *Domesday Book* [Книга страшного суда, поземельная перепись] Вильгельма I Завоевателя 1086 г.; записи таможенных сборов в лондонском порту в период Эдуарда III [король в 1327 – 1377 гг.], подсчёты мужчин и ресурсов ввиду опасности возможных войн при правлении династии Тюдоров [1485 – 1603 гг.] и регистрация смертей при Генрихе VIII в 1532 г. при широко распространённой боязни чумы.

Со временем зависимость государственности от статистической системы всё более прояснялась. Меркантилисты эпохи абсолютной монархии были озабочены мощью государств в мирное и военное время, а монархи и их советники должны были знать меру числа мужчин, способных носить оружие (men) и других ресурсов (Viner 1968). Британское и французское колониальные администрации в Северной Америке, а позднее британская администрация в Индии проводили переписи. Первой современной была перепись 1666 г. в Канадской провинции Квебек (Linder 1959, с. 330).

С появлением демократии появились новые основания для составления статистических сборников. Для первой из основных государственных конституций, а именно для конституции США, потребовалось определить политический вес входящих в государство объединений, т. е., по-видимому, их относительное население. Первая из регулярных переписей, проводимых каждые десять лет, прошла в 1790 г, хотя их официальная программа была составлена лишь в 1830 г. Доминион Канады в 1867 г. был основан на аналогичном положении о представительстве в соответствии с населением, после чего в 1871 г. произошла перепись. Австралийские колонии объединились в 1900 г., а статистическое бюро было основано там в 1905 г.

После Второй мировой войны династические, колониальные, племенные и иные политические формы управления были по всему миру вытеснены национальными объединениями. Новые страны, наряду с попытками принимать конституцию, проводить выборы и закладывать начала современной промышленности, создали статистические системы и в качестве предварительного условия развития, и как символ нации. Новый статистический стиль оказался гораздо честлюбивее предшествовавшей колониальной модели, в той же мере, в какой национальные цели, взятые в целом, шире колониальных.

Лига наций посредством своего комитета экспертов по статистике серьёзно способствовала расширению статистических работ и повышению их качества. ООН, через свои комиссии по статистике и населению, продолжает эту деятельность. Под её покровительством и при серьёзной поддержке отдельных стран была предпринята попытка проведения всемирной переписи примерно в 1950, а позднее в 1960 г. Если в переписях 1860-х годов было подсчитано 17% оценённого населения планеты, то

примерно в 1960 г. насчитали более чем в 10 раз больше, более 67% всего населения Земли (Yearbook 1962, с. 1). Не везде результаты оказались одинаково удовлетворительными, и анализ качества статистической организации и статистических результатов в различных странах несомненно выявит тесную связь указанного с качеством правительственной администрации в целом. Чили лучше статистически организовано, чем Бирма, а Бирма – лучше, чем Камбоджа.

[1.2.] Классификация государственной статистики. Широкое поле статистических данных стало считаться подходящим для сбора и публикации правительством.

[1.2.1.] Классификация по источникам. Это – одна из возможных классификаций. Посемейные сведения служат источником для переписей населения, исследования семейных бюджетов и подсчётов работающих и безработных. Сведения от фирм для получения данных о производстве и занятости. Сведения от них же, объединённых в корпорации или нет, для сбора данных о прибылях. Сведения от Министерства государственных сборов, для сбора данных о внешней торговле и подоходном налоге. Семьи и коммерческие фирмы часто представляют дополнительные данные такого же (?) рода и поэтому отдельные категории в приведенной классификации не являются единственно возможными.

Если экономическая статистика в основном состоит из сводки сделок, её можно в принципе составить по ответам на вопросы, поставленные любым их участникам. Розничные цены можно устанавливать на основе опросов розничных магазинов и посемейного исследования по отношению к продаже и покупке товаров соответственно.

При более частом обращении к предприятиям появится большее количество зарегистрированных данных и сделок, косвенно относящихся к одному и тому же отчёту. В развивающихся странах, в которых деловая жизнь менее организована, сравнительно более ценятся вопросы, задаваемые потребителям. В отношении данных о занятости и безработице семья в общем видимо является столь же удовлетворительным источником сведений, сколь и предприятия.

В Северной Америке расхождения между соперничающими данными о безработице оказались важным стимулом совершенствования и выборочных исследований, и методов опроса при посемейных исследованиях, и управления самого страхования от безработицы, притом даже в направлениях, далёких от статистики.

[1.2.2.] Классификация по времени публикации. Более грубая форма классификации основана на указании временных тенденций. Некоторые наблюдения публикуются быстро и сообщают последние новости: еженедельная погрузка вагонов, продажи универсальных магазинов, запасы некоторых металлов и зерна на главных элеваторах.

Напротив, полные переписи промышленности и населения²

предпринимаются лишь через длительные промежутки времени, и их результаты публикуются через несколько месяцев после подсчетов, но продолжают появляться в печати в течение нескольких лет. Они ценны, поскольку выявляют соотношения между отдельными видами данных. Ввиду подобного сочетания сроков публикации картина национальной экономики страны на любой момент времени лишь постепенно вырисовывается в течение пяти – десяти лет.

Государственная статистическая деятельность наряду со сбором данных включает их исследование и истолкование. Хотя результаты анализа способствуют сбору данных и наоборот, эти два вида работы различны. Примером разделения труда на указанной основе служит то, которое существует между Бюро переписей и Министерством труда США по статистике занятости и безработице. Бюро собирает данные, а Министерство официально исследует их.

[1.3.] Модернизация. Основной целью начальной стадии построения статистической системы видимо является перепись населения и внешней торговли. Постепенно добавляются другие виды данных, как, например, число начатых и законченных строительством жилых зданий. С развитием страны вместе с разнообразием и трудностями, вызванными усилившимся в экономике дальнейшим разделением труда, возрастает потребность в статистике. Необходимость действительного обоснования своих решений для самих себя и для других видимо оказалась с раннего времени характерной для Северной Америки. В 1893 г. генерал Френсис Уокер сообщил об этом Международному статистическому институту:

Сильная страсть к статистике с раннего времени образовалась в жизни нашего народа, и такие государственные деятели и публицисты как Гамильтон стали по сути статистиками. [...] Ни одно правительство в мире никогда не расточало больше денег, не тратило столько трудов в этом направлении [...] так бодро и терпеливо. Цит. по Cummings (1918, с. 573)³.

Но эта культура фактов уже не ограничивается ни территорией США, ни англоязычными странами, ни европейцами; она становится всемирной.

Утверждение о том, что статистика, собираемая любым правительством, зависит от его интересов и ответственности, можно иллюстрировать примером освобождения колоний. Когда большая часть мира была подчинена государствам Западной Европы, статистическая система сосредоточивалась на внешней торговле. Для британских администраторов было важно знать количество риса, вывезенного из Бирмы и джута из Бенгалии; Нидерланды хотели знать количество кофе, сахара и каучука, вывезенного с Явы в те периоды, в течение которых производство каждого из этих товаров было наивысшим.

Другая стадия колониального развития была связана с земельным налогом, который приводил к необходимости и возможности численно оценивать посевные площади и

продукцию основных культур. Далее появились элементы озабоченности о населении, выраженные в его переписях. Вслед за строительством железных дорог в Индии и на Яве, к примеру, появился сбор транспортной статистики, о которой никто не думал в то время, когда основным средством передвижения товаров были повозки, запряженные волами, парусные суда и человеческие спины.

Эволюция, имевшая место в колониях в Азии, произошла и в Латинской Америке, в которой интересы нескольких государств, хоть и независимых, были почти столь же ограничены, как и у колониальных стран [по отношению к своим колониям]. Переход к независимости в Азии и повсеместное распространение социальных государств расширило охват государственной статистики. Каждая развивающаяся страна стремится расширить свою систему образования, и статистика школ, учителей и учащихся стала почти повсеместной. В Англии и Уэльсе своего рода статистика образования зародилась в 1820-х, 1830-х годах, когда школы впервые начали получать общественные деньги (Baines 1918, с. 377).

Статистика рождений и смертей постепенно совершенствуется и требует внедрения привычки их регистрации не только у руководителей её гражданской системы, но и у медицинских работников и у граждан вообще. В новых странах подобный перелом потребует по меньшей мере одного или двух поколений. Тем временем вводятся выборочные исследования рождаемости и её изменения, равно как и многих других явлений, касающихся населения и экономики.

Исследования семейных бюджетов, начавшиеся уже при колониальном режиме в моменты социальной озабоченности (исследования бюджета кули в 1930-е годы в Голландской Ост-Индии), происходят чаще. Особенно активно в этих и других видах посемейных исследований правительство Индии, действующее посредством Индийского статистического института в Калькутте.

[2.] Организация статистической службы

Каждому правительству приходится рано или поздно обдумывать организационную структуру для сбора статистики. Предоставленные самим себе министерства стремятся сохранять данные любого рода, порождённые их администрацией, и вряд ли они отделяют их сбор от административного контроля своих действий. В 1789 г. (Cumplings 1918, с. 579) в США был принят акт, регулирующий сборы. Он обязал таможенников регистрировать декларации груза судов и другую информацию о торговле. Так в стране начался сбор статистики внешней торговли.

Почтовая служба Англии собирает и публикует сводки о своих доходах, а Министерство внутренних доходов, созданное в 1849 г., о земельном и подоходном налоге, налоге на недвижимость и о гербовом сборе. Однако, с возрастанием объёма работы министерства начали отделять статистическую деятельность от административной. Уже в 1833 г. (Baines 1918, с. 374) статистический отдел был учреждён в Министерстве торговли

Англии, а в дальнейшем в Англии, США, доминионах Англии и странах Западной Европы многие министерства и управления создали свои собственные статистические подразделения. Комиссия по торговле между штатами США с самого начала своей деятельности настаивала на получении данных от железных дорог, работу которых она контролировала. Её публикации составляют важную серию.

Иногда на одно из подобных учреждений возлагали обязанность проводить очередную перепись. В Англии и Уэльсе ответственным за перепись является глава отдела актов гражданского состояния (registrar-general). В Канаде специальное бюро при Министерстве сельского хозяйства, созданное в 1905 г., проводит переписи каждые 10 лет в соответствии с актом британской [колониальной] администрации. Но во всех странах возрастающее внимание к статистике привело к предложениям о её централизации в специальных агентствах.

[2.1.] Степень централизации. Создание даже небольших статистических бюро в министерствах оказалось полезным ввиду некоторого отделения статистики от повседневной административной работы.

[2.1.1.] Централизация в национальных правительствах. Встал вопрос об объединении всей статистической деятельности правительства в едином бюро, чтобы тем самым дополнительно повысить пользу специализации. К доводу действенности [...] добавилось соображение об исключении дублирования, этой неизменной опасности в работе правительства, которое проявилось в статистике с самых ранних лет. Отделение статистической деятельности от иной кроме того ослабляет давление на статистику, требующее искажения результатов для защиты политических или административных интересов. Ниже мы сошлёмся на юридическую сторону работы статистика.

Далее, лучшая координация, достигаемая в централизованной системе, должна позволить выявление пробелов в собранных данных и облегчить принятие единых классификаций. Весьма желательно, к примеру, чтобы производство и экспорт регистрировались на основе одной и той же классификации товаров и оказались поэтому сравнимыми друг с другом. Но каждый, знакомый с тягой к независимости отдельных подразделений правительственного учреждения, поймёт, что единая крыша над несколькими сборниками сама по себе не обеспечивает единообразия. Эти подразделения могут преследовать различные цели.

Централизованной организации присуща особая опасность: она может стать настолько обособленной, что перестанет воспринимать требования пользователей статистических данных. Единая центральная организация обеспечивает должные результаты, если она хорошо внутренне координирована и прислушивается к подобным требованиям.

Административные доводы в пользу централизации уже давно привели к той или иной степени успеха в Нидерландах, Канаде, Австралии, Германии и Италии. В 1848 г. в Нидерландах, в

Министерстве внутренних дел, было создано статистическое бюро, ответственность которого простиралась далеко за пределы табулирования данных, порождённых самим министерством. В Канаде и Австралии централизация оказалась двойной, поскольку статистика была делом центрального правительства не менее, чем правительств штатов или провинций или любых местных органов власти, а среди агентств центрального правительства одно должно было в первую очередь заниматься сбором и табулированием статистики в различных областях.

Канадская система была в принципе полностью централизована с момента рождения Бюро статистики доминиона Канады в 1918 г. Некоторая степень централизации была достигнута в Германии с учреждением Имперского статистического бюро в 1872 г., через год после образования немецкой империи. В Италии централизация началась в 1861 г., т. е. в год образования единого Итальянского королевства, с создания директората статистики, предшественника нынешнего Центрального института статистики.

В США и Соединенном Королевстве практика оказалась менее централизованной. Координирование в Англии осуществляет там бюро в секретариате правительства, которое работает на основе соглашений с главами статистических подразделений Министерств торговли, продовольствия, сельского хозяйства и др. Всеми этими подразделениями руководят статистики профессионального, научного и технического класса, одного из пяти классов гражданской службы. На работу в эти подразделения могут быть приняты члены статистического класса (?) вне зависимости от их места службы. Министерство финансов утверждает расходы статистических подразделений, а по поводу внеплановых расходов их консультирует центральное статистическое бюро.

Американская структура более формальна. Контора статистических стандартов в Бюро бюджета рассматривает все ходатайства о финансировании статистической деятельности для федерального правительства и утверждает все его вопросники, высылаемые им в 10 или более адресов. Это требование появилось в федеральном акте 1942 г. и имело целью облегчить бремя респондентов и исключить дублирование. Там, где оно существует, оно не только напрасно расходует государственные средства, но ожесточает респондентов, обязывая их снова отвечать на те же самые вопросы.

Наилучшая координация не состоит только в отрицательных принудительных предписаниях. Положительным примером служит сотрудничество администрации социального обеспечения США, которое по существу является рабочим агентством, и Бюро переписей в выпуске издания о деловой активности в округах, основанного на данных о социальном обеспечении. Бюро переписей использовало данные о налогах для отмены вопросников, обращённых к миллиону заведений розничной торговли, не нанимающих работников. Исходя из данных социального обеспечения, это Бюро может предпринимать экономические переписи с меньшими расходами.

Координация требует тонких суждений по тысяче вопросов. Должно ли Бюро горных работ совместно с Бюро переписей продолжать составление статистики добычи минералов? Каждые пять лет Бюро переписей собирает данные о продукции по всем типовым (benchmark) заведениям горной промышленности, которых оно охватывает. Бюро горных работ собирает данные чаще, и не только о добыче минералов, но и об инженерных и технических вопросах. Данные каждого из этих двух бюро считаются ценными и достаточно различными, чтобы не считать, что они дублируют друг друга.

[2.1.2.] *Разделение ответственности.* Совсем в другом направлении находится разделение ответственности между национальным правительством и местными органами власти штатов, провинций и крупных городов. В этом направлении Соединённое Королевство сильно централизовано, а в [Западной] Германии возникают проблемы ввиду разделения статистической деятельности по отдельным землям, и даже в Нидерландах существует некоторая передача статистических работ в ведение городских организаций.

Информация для каждой национальной статистической системы должна зависеть от местных источников. На одном конце шкалы находится полный централизованный контроль как при национальных переписях в США и Канаде. В этих странах местные должностные лица назначаются и постепенно инструктируются центром, который указывает им определения и методы, по крайней мере в принципе полностью определяемые центром. Он же оплачивает всех участников статистических работ.

Менее централизованы переписи в таких странах, как Аргентина. Их программы составляет там национальное статистическое бюро, но счётчиками на один или несколько дней назначаются государственные служащие (например, учителя школ), которых оплачивают другие департаменты центра или провинций. Координация таких операций требует совершения дипломатических подвигов.

Ещё менее централизована система регистрации рождений и смертей в США и других странах, в которых местная регистратура докладывает властям штата. Единообразие в масштабе страны, поскольку оно существует, достигается переговорами, в которых национальное бюро может играть главную роль, более в силу убедительности, чем ввиду какой-либо юридической или бюджетной возможности.

[2.2.] *Сравнимость и изменения.* Государственное статистическое агентство должно действовать непрерывно и приспособлять свои сводки к изменяющимся потребностям. Это сочетание достоинств легче предписать, чем выполнять. В некоторых случаях агентства оказываются застывшими. Склонность из года в год собирать те же данные тем же способом оказывается их сильной, но также и слабой стороной. Впрочем, иногда изменения происходят произвольно.

А. Л. Bowley заявил [где, когда?], что перепись 1921 г. в Соединённом Королевстве была намеренно проведена так, чтобы её нельзя было бы сравнивать ни с какой предыдущей, но Greenwood и др. (1932, с. 279) отвергли это обвинение. И здесь тоже требуется суждение. Окончание старой серии и введение новой совместно с видоизменением определений означает, что наблюдения станут полезнее, но их сравнимость с прежними окажется труднее. Этот вопрос должен быть обсуждён более конкретно, но здесь такой возможности нет.

Практические методы, хорошие в одном месте в некотором периоде, могут оказаться неважными в другой экономике и обществе. Попытки перенести отдельные разряды статистических сводок, использованные для характеристики страны с преобладающей денежной [рыночной] экономикой, в страну натурального хозяйства могут оказаться неудовлетворительными. Некоторые наибольшие трудности при сборе статистики в развитых (?) странах происходят в секторе натурального сельского хозяйства.

Взаимодействие статистики и работы промышленности особенно ярко проявилось в наше время в США. Национальное бюро экономических исследований, Бюро переписей и некоторые другие агентства пытаются выяснить, как можно было бы исправить или заменить отдельные серии наблюдений. Товары начали в основном перевозить на грузовых машинах и по воздуху, и погрузка вагонов [как показатель] стала гораздо менее полезной. Продажи универмагов менее значимы, поскольку многие товары продаются и в магазинах сниженных цен, и в аптеках. Нужен дух изобретательности, чтобы дополнить введенные им ранее основные статистические показатели экономики, продажи универмагов и погрузку вагонов.

[2.3.] Взаимоотношения с внешним миром. Государственное статистическое агентство никак нельзя назвать свободным, оно зависит от административной работы других государственных агентств и делового мира. Статистика импорта зависит от оценок, установленных для таможен. К примеру, сравнение экспорта газетной бумаги или металла из Канады в США с импортом номинально тех же товаров в США из Канады покажет, как сильно влияет различие в определении товаров, равно как и их источников и мест назначения. Определения часто используются при таможенном оценивании и поэтому включаются в законы, от которых существенно зависят возникшие финансовые интересы. Расхождения [в определениях и таможенных оценках] между странами вероятно останутся, несмотря на многочисленные соглашения и попытки согласования, имевшие место с брюссельского соглашения 1910 г. до усилий статистической комиссии ООН⁴.

[2.3.1.] Влияние бухгалтерской отчётности. Существенное влияние на работу статистических агентств оказывает бухгалтерская отчётность в промышленности, которая сама частично определяется положениями о подоходном налоге. Искусство составления статистических сводок включает

установление таких способов определений [товаров], которые наиболее соответствуют принятой в деловом мире бухгалтерской отчетности, и предоставляет наибольшую пользу для общественности.

В отличие от фирмы, при сборе промышленной статистики за исходный элемент обычно принимается промышленное предприятие. Практически оно понимается как наименьшая единица, выпускающая продукцию и более или менее полно регистрирующая свою деятельность. Статистическое агентство не всегда играет чисто пассивную роль; часто оно способно оказывать влияние и проявлять инициативу. В Канаде федеральное агентство, занимающееся финансовой статистикой пяти тысяч городов, смогло убедить их принять стандартный метод бухгалтерской отчетности, который оказался полезным и им самим, и национальной статистике.

[2.3.2.] Национальные профессиональные ассоциации. Внешний мир для государственного бюро, составляющего статистику, включает другие правительственные учреждения, производителей товаров, множество профессиональных ассоциаций, а также отдельных лиц, заинтересованных в статистике. В 1844 г. Американская статистическая ассоциация обратилась в Конгресс с ходатайством заново обработать материалы переписи 1840 г. Политика протестов против плохой работы в области федеральной статистики и поддержки хорошей работы продолжается до настоящего времени, хотя необходимость в подобных мерах не так велика, как в то время, когда почти все государственные служащие были любителями (Bowman 1964).

[2.3.3.] Международные профессиональные ассоциации. К национальным ассоциациям теперь присоединились важные международные. Первой был Международный статистический институт со своим неперенным и благотворным влиянием в защиту высоких стандартов и сравнимых классификаций. Экономический и социальный совет ООН⁵ посредством своих комиссий по статистике и по населению побуждал государства проводить переписи с обязательным наименьшим набором вопросов и тем самым серьезно попытаться достичь международной сравнимости этих вопросов и применения современной технологии в целях точности и быстроты публикации собранных данных.

Ощущая желание отдельных стран оставаться вне зависимости от внешних воздействий, совет в основном опирался на довод о международной сравнимости статистики производства и населения. Влияние оказывали также региональные агентства ООН по Азии, Европе, Латинской Америке, а затем и Африке, равно как и Межамериканская статистическая ассоциация, специализированные агентства ООН (ЮНЕСКО, Организация продовольствия и сельского хозяйства (FAO), Международная организация труда и т. д.), каждое в своей области.

Подготовленные ими резолюции и соглашения часто, к сожалению, были сформулированы крайне абстрактно, но они дополняются службой технической помощи развивающимся

странам, в частности посылкой советников и предоставлением возможности обучения. Издания международных организаций были непосредственно полезны учёным и другим читателям и кроме того побудили членов [ООН] собирать данные. Конференции FAO групп соседних государств оказали действенное давление в пользу совершенствования сельскохозяйственной статистики.

[2.3.4.] Другие влияния. Серии публикаций государственной статистики часто осуществлялись под внешним влиянием. В 1926 г. фирма Metropolitan Life Insurance начала собирать статистику текучести работающих в США. Позднее выпуск подобных сводок переняло Министерство труда (Hauser & Leonard 1946, с. 363). В Канаде статистику ставок заработной платы собирала Bell Telephone Company для своих собственных целей, что повлияло на сборники федерального правительства. Статистику имеющихся запасов часто начинали составлять объединения промышленников, однако затем её перенимали общественные статистические агентства.

В Нидерландах централизованная система [статистики промышленности] возникла в некоторой мере под влиянием частного Союза статистики (Stuart 1918, с. 435), который опубликовал много томов статистических данных, хоть и не собирал их. Впоследствии правительство начало финансировать этот союз и в конце концов переняло всю его деятельность. Особенно хорошо известно то давление, которое в 1825 – 1841 гг. Кетле оказывал своими исследованиями в качестве частного лица (после чего он стал должностным лицом). Его ненасытный интерес в закономерности некоторых социальных явлениях, например, в *бюджете эшафотов*, был увековечен в обширной серии уголовной статистики, публикуемой в Бельгии и во многих других странах (Julin 1918, с. 128)⁶. В Англии через три года после выхода первого издания очерка Мальтуса [1798], который привлек серьёзное внимание к населению, а именно в 1801 г., была проведена перепись.

Существует много других примеров влияния личностей и идей, а также и внешних организаций, на государственную статистику. Сборники экономических данных, вышедших при жизни нынешнего поколения, изменились ввиду признания национальных отчётов в качестве общей структуры. Это понятие было разработано учёными в университетах США и Великобритании в соответствии с предложением Кейнса.

Понятие о национальном доходе позволяет расположить существующие и в громадной степени отличающиеся друг от друга серии в различные отделения структуры национального дохода. Результаты, получаемые статистическими системами, уже не представляются произвольными, потому что каждая их часть является мерой определённого вклада в валовой национальный продукт. Пробелы (например, личные услуги) становятся заметными и появляется давление в пользу их заполнения.

Далее, подобное исследование экономической статистики в целом позволяет вычислять общий итог экономической

деятельности страны по меньшей мере тремя способами: по чистой продукции, по полученным доходам, и по расходам. Все эти способы включают существенные элементы, мало зависящие от источников данных. Сопоставление вычислений этого целого, выведенного из более или менее независимых статистических источников, побудило повысить точность всех его компонентов.

Нынешний престиж экономического планирования существенно повлиял на деятельность государственной статистики. Планирование может быть хорошим или плохим, действенным или нет, но в любом случае оно требует исходных данных. Без них нельзя даже приступить к планированию, и государственные планирующие агентства, нуждающиеся в данных, могут особенно настаивать, чтобы для их сбора были выделены необходимые ресурсы.

[2.4.] Статистическая профессия и государство. В основании недавних событий находится быстрая профессионализация статистики. Времена неузнаваемо изменились с того времени, когда Френсис Уокер горевал в своём докладе Американской ассоциации статистики в 1896 г., что он

не знает ни одного человека, являющегося или ранее являвшимся главой статистического бюро или директором статистической службы или статистиком, который имел бы хоть какое-то элементарное образование для своей работы. Цит. по Cummins (1918, с. 574).

Недавнее перечисление штатов Бюро переписей насчитало 576 сотрудников, занимающих должности профессиональных статистиков. Из них только 22 не имело высшего образования, но оказалось 36 докторов философии и 102 магистра искусств (Bowman 1964, с. 14)⁷. Изменение было вызвано расширением сферы деятельности статистического метода, который применяется в промышленности для приёмочного выборочного контроля, для планирования сельскохозяйственных экспериментов, для биотестирования, исследования операций и выборочного обследования населения.

Взаимный обмен [идеями] между этими приложениями и между каждым из них и быстро развивающейся теорией математической статистики имел решающие последствия. Объём книжной и журнальной литературы быстро возрастает, факультеты статистики учреждены в некоторых университетах и курсы по статистике почти во всех из них, а государственные статистические агентства поддерживаются теорией вероятностей и укоренившимися методами её приложения. Они поэтому всё меньше и меньше зависят от талантливых любителей типа Кетле в Бельгии, Книббса в Австралии или Коатса в Канаде.

И всё же нынешнее расширение поля деятельности и продвижение может быть сравнено по качеству с некоторыми работами прежних лет. Флоренс Найтингейл сильно повлияла на Военное министерство Великобритании своей статистикой, которая по отдельности показывала потери в бою и смертность в армии от болезней. Её данные об армейских госпиталях в Индии привели к важным преобразованиям. Современные

профессионалы не то, чтобы были лучше старомодных талантливых любителей, но при необходимости их скорее можно отыскать и они владеют новым инструментарием.

[2.4.1.] Выборочный метод при новом профессионализме.

Среди нынешних крупных выборочных исследований упомянем ежемесячные в США и Канаде, ежеквартальные в Западной Германии и других европейских странах и по существу непрерывные в Индии, а их цели почти столь же разнообразны, как и у самой статистики.

Как и всякое существенное изменение метода, выборочные исследования не только достигают прежних целей с меньшими расходами, они привели к совершенно новому мышлению о всём процессе сбора данных. Статистик теперь не считает, что должен пассивно собирать нужные данные, он ориентируется по цели их сбора и степени требуемой точности. Так, он приходит к мысли учитывать погрешность выборки и той, которая не зависит от неё. При подсчёте погрешностей, контроле качества и оценке результата он ощущает себя профессионалом в высшей возможной степени. И он начинает измерять информацию не по массе данных, которую способен собрать, но скорее по достигаемой точности своего результата. К примеру, подсчитанный процент безработных среди активной части населения может послужить причиной серьёзного экономического решения. А если потребуется повысить точность результата, он должен будет решить, какую из двух упомянутых погрешностей можно снизить в большей степени при одних и тех же усилиях.

Новый подход к погрешностям, а именно понимание их неизбежности и необходимости постоянной борьбы с ними, замечен в Бюро переписей США⁸. Начиная с 1950-го года, помимо результатов переписей населения указываются официальные оценки его недоучёта, которые частично основываются на выборочных исследованиях, выполняемых лучшими счётчиками. С 1951-го года эта практика введена и в Канаде.

Профессиональный статистик не заявляет, что его результат точен, и это его отличительная характеристика. Даже если его исследование произведено наилучшим возможным образом, он удовлетворится тем, что посчитает свой результат произвольно выбранным из тех, которые могли бы быть получены таким же образом в то же самое время.

[2.5.] Стойкие проблемы официальной статистики.

Исследование деятельности статистических агентств даже весьма развитого государства установит, что некоторые устаревшие методы всё ещё в ходу, что публикуются числовые результаты, бесполезные ввиду того, что они были основаны на чисто административных определениях исчисляемых величин, либо потому, что не были указаны их возможно крупные погрешности. Помимо ценных статистических данных, библиотечные полки забиты не относящейся к делу, бесполезной и неточной

статистикой. Это печально, поскольку существуют хорошие методы работы.

[2.5.1.] *Юридическая роль государственной статистики.* Особо опасны [намеренные] неточности, преследующие какие-либо политические цели. Сильнейший довод в пользу статистической централизации относится к тому, что можно назвать юридической ролью официальной статистики. Статистика внешней торговли быть может укажет, что плохо проводится политика государственной торговли; статистика цен быть может установит, что монетарная политика правительства приводит к инфляции. Для защиты населения необходимо, чтобы агентство, ответственное за государственную политику [в данной области], было лишено исключительного контроля своей статистики. В противном случае от него потребуются нечеловеческие усилия, чтобы противостоять искушениям.

В сотне мелких решениях по поводу статистических сводок соответствующее агентство, установив получаемый результат, может уклониться от истины. Обстоятельства публикации также должны быть беспристрастными. Выпустить ли индекс цен в намеченный срок, или на неделю позже, т. е. после всеобщих выборов? От ответа на подобные вопросы зависит достоинство статистика. Если такого достоинства нет, если нет поэтому и доверия общественности к статистике, большая доля затрат на её сбор окажется напрасной. Статистика, в которую не верят, бесполезна.

Для статистика, как и для судьи, невежество менее опасно, чем продажность. Статистик более всего полезен обществу, если его статистическая честь не противоречит какой-либо посторонней ответственности. На более высоком административном уровне лицо, ответственное за политику министерства, имеющего статистическое бюро, должно иметь силу воли, чтобы разрешить статистикам принимать решения, исходя из статистических соображений. Не всегда, правда, эти решения будут верными, но ошибочными они окажутся не из-за корыстных интересов. Сила воли будет не так нужна, если статистическое бюро контролируется комиссией, которая находится полностью вне политики, как это и практикуется в Нидерландах.

Гораздо менее утонченными, чем указанные проблемы, являются искажения, встроенные в политику и заметные в тоталитарных странах. В СССР при Сталине считалось, что статистика относится не к науке, а к революции и мобилизации населения для национального строительства⁹. Если статистика не укажет продвижения, массы разочаруются и их усилия ослабеют. Такой же подход существовал в Китае в 1950-е годы (Li 1962).

Соревнование между районами за первое место в социалистическом производстве было расширено и включило соревнование о выпуске продукции. Усердное объявление высоких показателей достигло своей вершины в 1960 г., когда заявленный общий урожай зерновых в 275 млн тонн был завышен более, чем вдвое.

Считалось демократичным привлекать массы к составлению статистики, и проявившееся социалистическое соперничество видимо действительно ввело и режим, и население в заблуждение, что имело пагубные последствия. Но существуют явные признаки того, что в СССР, как и в более давних индустриальных обществах, заслуживающая доверия статистика больше не считается лишь буржуазной роскошью. Неясно, правда, признаётся ли в Китае зависимость верных решений от добротности статистики.

[2.6.] Научные и производственные достижения. Агентства, пытавшиеся действительно выполнять официально возложенные на них поручения по сбору данных, добились обширных успехов в методах выполнения почти всех разделов статистической деятельности. Можно упомянуть Бюро статистики труда США в связи с методами составления таблиц дожития; введением сезонных поправок во временные экономические ряды; теорией и практикой вычислений индексов, а Бюро статистики Канады по поводу классификации занятий, отраслей промышленности и товаров (Coats 1925). Особо важными для развития демографии были работы XIX в. Уильяма Фарра и других статистиков Бюро генеральной регистрации (General Register) Соединенного Королевства.

Самым поразительным примером прогресса является непосредственное отношение Бюро переписей США к развитию табулирующего и вычислительного оборудования от примитивных перфорирующих устройств 1890-х годов, которые сконструировал в самом Бюро Герман Холлерит, до компьютера UNIVAC, спроектированного частным образом при поощрении, финансировании и технической помощи Бюро стандартов. В этой области пионером было и канадское Бюро переписей со своим крайне оригинальными электрическим табулятором и другим, работающим на сжатом воздухе, – плодами изобретательности Fernard Belisle, неизменного сотрудника канадского Бюро статистики.

Бюро переписей Канады и США в последние годы построили или поощрили конструирование приборов для ввода [информации], чтобы тем самым избежать перфорирования, этой самой утомительной процедуры обработки крупных статистических исследований.

Помимо участия в разработке вычислительных машин, государственные статистические агентства одними из первых весьма активно применили их. Будь то страховое общество, нефтеочистительный завод или подобное агентство, их применение требует существенного переосмысления процессов, которые развивались десятилетиями. Некоторые самые передовые в мире статистические бюро решительно перестроили свою работу, чтобы воспользоваться электронными вычислениями, и, стало быть, шире применять таблицы сопряжённых признаков, лучше контролировать ошибки при обработке данных и экономить средства.

В США по мере того, как пользователи статистики обзаводились собственными компьютерами, увеличивался спрос на результаты, которые считывались машинами. Бюро переписей составило обширный каталог результатов, хранимых на перфокартах или электронных лентах. Эта форма публикации станет безусловно распространённой.

[2.7.] Публикация результатов. Быстрая и полная публикация составляет главнейшее достоинство государственной статистики, если только представленные данные точны и относятся к делу. Это утверждение вызвано ролью статистики в процессе принятия решений. Разумность решения, например, о строительстве фабрики, зависит от обстановки на какой-то будущий срок. Если статистик не сможет оценить количество населения на пять или 25 лет вперёд, он по крайней мере способен описать обстановку, существовавшую на возможно ближайший предшествовавший момент.

О содержании публикаций (какие таблицы сопряжённости признаков распределений и какие проценты указывать, какие схемы и таблицы представить, какие моменты сравнивать) государственный статистик будет лучше судить по мере своих контактов с пользователями его результатов.

При множестве государственных и частных источников возрастает необходимость в хороших указателях и перекрестных ссылках со сравнимыми данными в других публикациях.

При попытке истолкования своих данных государственные статистические бюро вошли в сферу социологии. Конкретно, потребовалось комментировать статистические таблицы, чтобы читатель смог узнать, как были определены сосчитанные величины и указать меру погрешности выборочных исследований и не зависящих от них ошибок; разъяснять вычисления, распределяющие не указываемые респондентами возрасты и вводящие сезонные поправки в ряды экономики. Наконец, нужны были полномасштабные монографии, ставшие одной из обычных форм представления переписей в США и Канаде по меньшей мере в прошлом поколении. В Индии подготавливается серия монографий, основанных на переписи 1961 г., и Пакистан готов последовать этому примеру.

Если представить себе непрерывный ряд от сбора сырых статистических данных до их окончательного использования либо как составной части научных исследований, либо при принятии решений, то место окончания роли государственного статистического бюро и начала роли пользователя окажется неоднозначным. Во всяком случае, государственный статистик должен описать все стороны своего метода сбора данных, которые могут повлиять на истолкование результатов. Следует надеяться, что в будущем ни одно исследование не будет публиковаться без реальной оценки точности числовых результатов, но наступит такой момент, видимо, нескоро.

Несмотря на всё это, некоторая умеренность навязывается официальному статистику. При истолковании данных он не смеет рисковать в том же смысле, в каком свободно может позволить

себе журналист. При явной произвольности мнений, выраженной в его тексте, доверие общественности к точности числовых данных, а потому и их польза, окажется под угрозой. Это соображение часто заставляет его составлять текст, который лишь оставляет благоприятное впечатление о том, что явно выявляется в таблицах, притом возможно, что этот текст не будет так ясен, как сами числа.

С другой стороны, государственный статистик может обоснованно заявить, что истолкование, будь оно живым или тусклым, без глубокого знания о том, на чём основаны числа в тексте, вероятно введёт в заблуждение. Короче, он по крайней мере обязан привести достаточно подробностей о своих методах, чтобы читатель смог правильно истолковать [результаты].

[2.8.] Скрытность. Она имеет двоякое отношение к статистической работе и к ней соответственно применимы прямо противоположные принципы. В отчёте [анкете] отдельного лица государственному агентству, как обычно считается, самое честное сообщение обеспечивается при гарантии доверительности, при уверении, что лица за пределами статистического бюро не смогут установить его авторства и что известными станут лишь обобщённые статистические данные.

Государственный статистик наделён властью потребовать от отдельного лица или объединения ответа на свои вопросы. Обратной стороной этой власти является обязательство сохранять в тайне индивидуальные ответы (Dobrovits 1947). В большом числе государственных исследований, при которых ответы на вопросы не обязательны, доля добровольных ответов может оказываться пропорциональной доверию общественности к мерам по сохранению тайны.

С другой стороны, объединённые результаты каждого исследования должны быть как можно более известны. Одна из трудностей государственных публикаций состоит в том, что правительство не имеет полного доступа к каналам распределения частной книготорговли. Иногда возможна частная публикация, как это произошло в США с монографиями переписи 1950 г. и как это делается в Нидерландах и [Западной] Германии. Принцип всеобщего равного доступа к статистике важен. Он обеспечивается своевременным объявлением срока выпуска, после которого статистика становится доступной одновременно всем заинтересованным лицам.

Во многих странах государственные статистические результаты широко распространяются посторонними промежуточными источниками, которые включают книготорговые ежегодники и научные журналы, поскольку перепечатка обычно допускается без особого разрешения.

Серьёзная проблема возникает ввиду желания правительства сохранять в тайне некоторые статистические результаты. Такова была прискорбная необходимость в США и Канаде во время второй мировой войны по отношению, например, к подробным числовым данным о внешней торговле. Подобную информацию использовали бы державы оси [Берлин – Рим] в своей подводной

войне против союзного мореходства. Но кроме таких исключительных случаев можно справедливо сказать, что общество в целом никогда не окажется в выигрыше от сокрытия статистических отчётов с точными сведениями. Впрочем, правительство и некоторые министерства действительно могут извлечь от этого пользу за счёт общественности. Дальнейшее обсуждение публикаций государственной статистики сливается с более широким вопросом свободного потока информации в обществе в целом.

[3.] Добавление

На более поздней стадии индустриального развития население грамотно, умеет считать и переваривает возрастающий объём статистических данных. Ежегодно спрос на новые серии [статистических исследований] становится настойчивее. Типичной для социальной и экономической статистики были число зарегистрированных рождений и тонн ввезенной стали, а вопросы, задаваемые официальным агентствам, оставались примерно одними и теми же на протяжении десятилетий, теперь, однако, изменения в необходимой статистике сопровождают каждый из многих сдвигов общественного внимания. Государственные агентства чувствительно восприняли новое состояние. Когда преступность стала заметным явлением, были сделаны существенные усилия, чтобы [числа] преступлений, доложенные полиции, оказались более реальными и сравнимыми. Такие исправления имеют предел, ибо большое число преступлений остаётся не известным полиции, и поэтому было испробовано использование нового первичного источника. Национальный центр исследования мнений (по контракту с правительством США), применяя методы выборочных исследований, начал опрашивать пострадавших от преступлений¹⁰. В более крупном масштабе эту работу продолжило Министерство юстиции.

Почти сразу же после того, как началось общественное обсуждение загрязнения окружающей среды, данные о массе загрязнителей в воздухе основных американских городов значительно расширились и стали публиковаться в *Statistical Abstract of the United States*. Как только начали предсказывать нехватку горючего, была представлена статистика расхода топлива на проезд единицы расстояния, мили, автомобилями [различных марок] по американским дорогам, равно как и количества пассажиров в них.

Подобные вызовы, брошенные статистической изобретательности, и, как можно надеяться, столь же действенные отклики будут продолжаться, пока новые экономические и социальные проблемы появляются перед правительством и общественностью, т. е. длительное время. И каждое соответствующее возрастание объёма статистических документов склонно увеличивать число их пользователей.

Помимо новых основных требуемых и предоставляемых данных возрастающее число заинтересованных лиц и их более тесное отношение к линии поведения [правительства] усиливает

спрос на более подробные вычисления. Так произошло в 1940-е годы по отношению к экономической статистике после появления национальных счетов и особенно валового национального продукта [в качестве показателя].

Серия появившихся публикаций заняла своё место и в прежних, и в новых странах наряду с переписями населения как признак нации. Национальные счета объединяются в такой степени, которая далеко превосходит прежние показатели, подобные индексу промышленной продукции. Национальный доход стал приниматься как мера социального прогресса в переходный период от домашнего к заводскому производству.

Совместно с ростом общей продукции он измерял этот процесс, а поскольку обе тенденции считались благоприятными, никто особенно не возражал против их смешения. Официальное предубеждение в пользу производства на рынок и роста экономики отражал мнение общественности, и мало кто замечал какое-либо отклонение от объективности. Но всё же решение о том, что включать в валовой национальный продукт, а что исключать из него, влияет на экономическую перспективу.

Сдвиг к рыночному производству в передовых странах почти завершён (менее 5% работающих заняты там в сельском хозяйстве, и даже фермеры мало что производят для собственного потребления), и появились серьёзные разногласия о направлении дальнейшего прогресса. Тем, кто желает иметь больше товаров, противостоят те, кто стремится сохранять окружающую среду. Упор делается на качество жизни, т. е. на понятие, почти неизвестное до 1970 г. И общественные, и частные усилия направлены теперь на поиск социального показателя.

Для официального статистика различие точек зрения социальных групп, использующих статистику, не ново. Его целью неизменно было предоставление тяжущимся основных материалов. Беспристрастная забота о том, чтобы ни одна группа не снискала исключительной пользы от принятых определений или хотя бы от даты публикаций, была неизменным условием успеха официального статистика в свободном обществе, и уклонения от неё обычно вызывали действенные протесты.

Наибольшая польза от статистических исследований имеет место, когда они пользуются доверием обеих сторон во многих спорах, – между желающими сохранять природу или увеличить выпуск продукцию; предпринимателями и профсоюзами в периоды, когда и тем, и другим ввиду инфляции важны индексы цен; националистами и интернационалистами в периоды давления в пользу защиты от иностранной конкуренции; и неизменно между правительством и оппозицией. Споры должны возникать об истолковании чисел, а не об их непристрастности.

Расширение клиентуры государственного статистика породило новый спрос на истолкование чисел. Массы пользователей не имеют терпения или знания, чтобы воспользоваться традициями строгости и массивности государственных публикаций. Неопределённость, возникшая в 1970-е годы о направлении экономического и социального прогресса, приводит к особым

затруднениям при объективных истолкованиях. Официальная статистика всегда мучила с одной стороны тех, которые хотели узнать, что означают его числа, а с другой, – тех, которые называли его пристрастным, если пояснённый им смысл отличался от представляемого ими.

В какой-то мере смысл определяется в процессе сбора материала, и он лучше всех понимает его. К примеру, ему известны погрешности в исходных данных, и соответствующая часть общего истолкования является его прямой обязанностью. Но что остаётся? Его профессиональное благоразумие должно решать, как поступать при существующих обстоятельствах. Много зависит от того, чтобы официальный статистик продолжал работать эффективно в атмосфере многочисленных перекрёстных давлений, и от того, чтобы правительство вне зависимости от давлений на него всё же видело бы пользу в поддержке его нейтрального профессионализма.

Краткие сведения об упомянутых лицах

Hamilton A., А. Гамильтон, 1755 или 1757 – 1804, отец-основатель США, министр финансов.

Knibbs G. H., Дж. Х. Книббс, 1858 – 1929, занимал руководящие должности в статистической службе Австралии, был активен в международном плане.

Coats R. H., Р. Г. Коатс, 1874 – 1960, главный статистик Министерства труда Канады, главный статистик Канады. Сотрудничал с Лигой Наций.

De Neufchâteau N. L. François, Н.-Л. Франсуа де Нёфшато, 1750 – 1828, политический деятель, писатель.

Hollerith H., Г. Холлерит, 1860 – 1929, статистик и изобретатель

Примечания

1. Government может означать государство, государственный, правительство, правительственный. Автор употребляет это слово во всех его значениях, переведено же оно по смыслу.

2. Переписи могут быть только полными, в противном случае статистическая работа называется исследованием.

3. Автор возвратился к этому докладу в § 2.4.

4. Вот мнение Кетле (1846, с. 364): каждое государство видимо *находит удовольствие* в том, чтобы добиться невозможности *какого-либо вида согласования* данных.

5. Лигу Наций автор упомянул только в § 1.1.

6. Самая полная уголовная статистика начала публиковаться во Франции с 1825 г., до появления соответствующих работ Кетле.

7. Доктор философии примерно соответствует кандидату наук. Магистр искусств, несмотря на это название, должен был изучать философию и социальные науки.

8. Весьма прискорбно, что статистики не знали ничего подходящего из сочинений Птолемея и Кеплера и даже не изучали соответствующих работ Гаусса и Бесселя.

9. Рассуждение автора о Советском Союзе не подкреплено никакими ссылками. Оно крайне поверхностно и частично ошибочно, а потому напоминает *развесистую клюкву*. Даже в 1954 г., т. е. после смерти Сталина, статистику объявляли *общественной наукой, основанной на революционной марксистской теории* (Шейнин 2001, с. 187). В 1952 г. биологический урожай

зерновых составил 130 млн тонн, амбарный же на 1/3 меньше (там же, с. 180 – 181), и вряд ли можно было в 1960 г. собрать 275 млн тонн. Какие явные признаки признания статистики автор нашёл в СССР до 1968 г.? Нам известна лишь статья Орлова (1990). О статистике в России автор, видимо, ничего не слышал.

По поводу социалистического соревнования вспоминаю забытую всеми ахинею. Году в 1935-м, было мне 10 или 11 лет, когда нам в школе разъяснили: пружина капитализма – прибыль, а социализма – социалистическое соревнование ...

10. Если преступления остаются неизвестными, то как можно опрашивать пострадавших?

Библиография

Орлов А. (1990), О перестройке статистической науки и её применения. *Вестник статистики*, № 1, с. 65 – 71.

Шейнин О. Б., Sheynin O. (1986), Quetelet as a statistician. *Arch. Hist. Ex. Sci.*, vol. 36, pp. 281 – 325.

--- (2001), Статистика и идеология в СССР. *Историко-математич. исследования*, вып. 6 (41), с. 179 – 198. Текст произвольно выправлен редакцией. См. подлинный текст в *Российская и европейская экономическая мысль. Опыт Санкт-Петербурга 2005*. СПб, 2006, с. 97 – 119.

Baines A. (1918), History and development of statistics in Great Britain and Ireland. В книге Koren (1918), pp. 365 – 389).

Bowman R. T. (1964), The American Statistical Association and federal statistics. *J. Amer. Stat. Assoc.*, vol. 59, pp. 1 – 17.

Coats R. H. (1925), The classification problem in statistics. *Intern. Labor Rev.*, vol. 11, pp. 509 – 525.

Cummings J. (1918), Statistical work of the federal government of the U. S. В книге Koren (1918), pp. 573 – 689).

Dobrovits A. (1947), Sur le secret en statistique. *Proc. Intern. Stat. Conf. Washington*, vol. 3, pp. 769 – 778. Calcutta.

Eckler A. R. (1972), *The Bureau of the Census*. New York.

Faure F. (1918), History and development of statistics in France. В книге Koren (1918), pp. 217 – 329).

Godfrey E. H. (1918), History and development of statistics in Canada. В книге Koren (1918), pp. 179 – 198).

Greenwood M. и др. (1932), Discussion of the quantity and quality of official statistical publications. *J. Roy. Stat. Soc.*, vol 95, pp. 279 – 302.

Hauser Ph. M., Leonard W. R., редакторы (1946), *Government Statistics for Business Use*. New York, 1956.

Julin A. (1918), History and development of statistics in Belgium. В книге Koren (1918), pp. 125 – 175).

Kahn E. J. Jr (1974), *The American People: the Findings of the 1970 Census*. New York.

Knibbs G. H. (1918), History and development of the statistical system of Australia. В книге Koren (1918), pp. 55 – 81).

Koren J., редактор (1918), *History of Statistics*. New York, 1970.

Kruskal W. H. (1973), The Committee on national statistics. *Science*, vol. 180, pp. 1256 – 1258.

Li Choh-Ming (1962), *The Statistical System of Communist China*. Berkeley.

Linder F. E. (1959), World demographic data. В книге Ph. M. Hauser, O. D. Duncan, редакторы, *The Study of Population: an Inventory and Appraisal*. Chicago, pp. 321 – 360.

Malthus T. R. (1798), *Essay on the Principle of Population. Works*, vol. 1. London, 1986. *Опыт закона о народонаселении*. М., 1895.

Quetelet A. (1846), *Lettres sur la théorie des probabilités*. Bruxelles.

Report (1971), *Federal Statistics: Report*, vols 1 – 2. U. S. President's Commission on Federal Statistics. Washington.

Statistics (1976), *Federal Statistics*. U. S. Office of Management and Budget, Statistical Policy Division. Washington.

Stuart C. A. V. (1918), History and development of statistics in the Netherlands.
В книге Koen (1918, pp. 429 – 444).

Viner J. (1968), Economic thought: Mercantilist thought. IESS, vol. 4, p. 435 – 443.

Yearbook (1948 –), *Demographic Yearbook*. U. N. Ежегодник. Издаётся с 1948 г.

Бернар Лекуе, Энтони Р. Обершаль

Ранняя история социологических исследований

Bernard Lécuyer, Anthony R. Oberschall,
Social research, early history of. IES, pp. 1013 – 1031

[1] Начало исследований в Англии

[1.1] Политическая арифметика. [...]

[1.2] Развитие демографии. [...]

[2] Начало социологических исследований во Франции

[...]

[2.1] **Кольбер.** Поиски качественной и количественной информации о французском обществе начались задолго до Кольбера (который с 1661 г. вплоть до смерти в 1683 г.) был генеральным инспектором финансов), но именно он в основном систематизировал предшествовавшие разрозненные усилия. Он ввёл изучение описаний территориальных единиц, и два опытных исследования (в Эльзасе в 1657 г. и, в 1663г., в Эльзасе, Лотарингии и в Трёх епископствах¹) предшествовали работе 1664 г.

Местных администраторов попросили сообщить о существовавших картах их районов, религиозных делах (особенно о *репутации и влиянии* епископов), военном управлении, знати, отправлении правосудия, состоянии финансов и экономической жизни (Esmonin 1956). Исследование мануфактур в 1665 г. было первой специальной задачей Кольбера, но он также позаботился об изучении состояния населения. Так, в 1667 г. он опубликовал декрет о составлении и поддержании приходских регистров, а через три года под влиянием Граунта ввёл публикацию ежегодных сведений о рожденьях, женитьбах и смертях.

[2.2] **Вобан.** Маркиз [и маршал] Себастьян Ле Претр де Вобан произвёл несколько далеко идущих исследований, привлекая к ним военные власти. В течение 30 лет с 1677 г. он был генеральным комиссаром по военным укреплениям, но затем впал в немилость и был вынужден уйти в отставку².

Вобан описывал территориальные единицы, а в его бумагах содержался *План подготовки переписи населения и описания провинций* возможно 1685 г. и 24 мемуара, описывавшие провинции, финансово-податные округа и города, включая общеизвестное *Описание округа Везле* 1686 г. В том же году он (Vauban 1707) написал *Общий и лёгкий метод переписи населения*, рекомендуя подсчитывать лица, а не хозяйства. И действительно, в 1682 – 1701 гг. было проведено несколько подобных переписей. В дальнейшем, однако, администрация вернулась к прежнему методу. О других исследованиях Вобана см. Dumazedier (1968).

[2.3] **Великое расследование.** Начиная с 1697 г. появилась серия 32 исследований, описывавших каждый административный район страны. Это *Великое расследование*, как его называли,

производили местные администраторы [...] и имело целью показать [...] нежелательные последствия политики Людовик XIV войн и чрезмерного налогообложения (Esmonin 1954; 1956). Каждое исследование основывалось на вопроснике, посланном местному администратору и включало описание территории и *естества населения*, его перепись, число и репутацию духовенства и знати, и 15 вопросов об экономической жизни.

Экземпляры местных расследований были разосланы лицам, занимавшим высокие посты, а администраторы районов использовали их при различных публикациях (Gille 1964). К этому обширному исследованию относится, по оценке, около 900 рукописей, послуживших руководством для администраторов до тех пор, пока в 1762 г. Bertin не начал производить новое исследование. Boulainvilliers (1727) описал содержание Великого расследования и резко критиковал его.

[3] Франция во второй половине XVIII в.

Административные исследования были возобновлены в 1724 г., вскоре после смерти регента и окончания финансового кризиса, в который в 1720 г. страну ввергнул распад французского права. За несколькими исключениями эти исследования были общенациональными, однако общие описания, которые предпочитали Кольбер и его последователи, теперь начали заменяться изучением конкретных проблем, связанных, например, с мануфактурами, общественной администрацией, попрошайничеством и заработной платой.

В 1730 г. Орру, достойный преемник Кольбера, назначил общее экономическое исследование, а в 1745 г., – другое, о *ресурсах народа* и о наборе в добровольную милицию. Оно привело к переписи, то ли действительно проведенной (Dainville 1952), то ли фиктивной (Gille 1964). Кроме того, администраторы должны были распускать слухи об увеличении городских налогов и о формировании милиции, чтобы затем тщательно доложить о реакции граждан; их отчёты сохранились. Существенное внимание было уделено сравнимости ответов и привлечению к работе учёных, находившихся вне правительства. Составление окончательного документа было поручено члену Французской академии³.

[3.1] Учёные и научные общества. Примерно в 1750 г. французское правительство уступило своё ведущее место в социологических исследованиях научным обществам и отдельным лицам, которые и продолжали господствовать над ними вплоть до 1804 г., когда Наполеон, будучи Первым консулом, провозгласил себя императором.

В то самое время, когда политическая арифметика в Англии теряла свою движущую силу⁴, она начала вторгаться во Францию, где её представители ограничивались изучением населения и тем самым избегали критических замечаний Дидро в адрес Петти⁵ (энциклопедическая статья *Политическая арифметика* 1751 г.). Успехи исчисления вероятностей, достигнутые в этом периоде [?], привели Депарсье (1746) к составлению таблицы смертности, и в 1749 г. эти данные использовал Бюффон.

Авторы более оригинальных исследований старались оценить население по числам рождений, указанным в приходских регистрах. В период 1762 – 1770 гг. аббат Expillly опубликовал свой *Географический, исторический и политический словарь Галлии и Франции* (Esmonin 1957). С 1764 г. он перестал основывать свои оценки общего населения на числе печей, а взамен исследовал указанные регистры за 1690 – 1701 и 1752 – 1763 гг., составив списки прихожан каждого прихода и установив искомое население по числу рождений⁶. Результаты по 9000 приходам, опубликованные в 1766 г., показали, что, вопреки распространённому мнению, население Франции возросло и по его оценке составило 22 млн человек.

Тот же метод применил Messance (1766) при исследовании населения некоторых провинций; впрочем, действительным автором мог быть La Michodière, администратор Лиона. Отношения числа жителей к числу рождений выражались переменными коэффициентами (у Expillly подобные коэффициенты доходили до 25). Автор также заключил, что население возросло и по его оценке составило 23 909 400 человек⁷. Эти усилия завершил Moheau (1778) в книге, написанной в 1774 г. Он был секретарём Montyon, бывшего администратора округа Овернь и несомненного автора указанного сочинения (Chevalier 1948; Esmonin 1958).

Montyon сравнил различные методы исследования и заключил, что наилучшее обоснование обеспечивает изучение приходских регистров и чисел рождений. Таким образом он оценил население страны, получив 23 687 409 человек. Его существенная оригинальность состояла в исследовании распределения населения (по возрастным группам и полу, по положению в обществе и гражданскому состоянию), а также и естественных и социальных факторов, влияющих на плодovitость.

[3.2] Исследования, проводимые Академией. Во второй половине XVIII в. Французская академия наук занималась социологическими исследованиями в двух направлениях. Первым было применение исчисления вероятностей к количественным данным. Так, в 1760 – 1761 гг. произошёл спор между Даламбером и Даниилом Бернулли по поводу статистической оценки результатов вариоляции оспы, в 1778 и в 1786 г. Лаплас написал мемуары об отношении мужских и женских рождений в Париже и о рождениях, женитьбах и смертях, а Кондорсе (1785) применил исчисление вероятностей к изучению вердиктов присяжных и результатов выборов (Rosen 1955; Westergaard 1932)⁸.

Вторым направлением было исследование технических проблем, по которым правительство часто советовалось с Академией. В 1762 г. Депарсье подготовил для него отчёт о снабжении Парижа питьевой водой, а в 1764 г. – о наводнениях.

В 1785 г. по инициативе Лавуазье и Кондорсе в новый устав Академии было включено положение об учреждении сельскохозяйственной секции. В том же году генеральный инспектор [?] Calonne образовал сельскохозяйственную

комиссию, включив в неё Лавуазье, а Академия назначила комитет для преобразования Отель-Дьё⁹. Работы Лавуазье в этой комиссии предоставили материал для публикации о территориальном богатстве Франции (Lavoisier 1791). [...]

Лавуазье оценил производство зерновых, сочетая данные о количестве населения (по Messance и Montyon) и о потреблении. Исходя от населения, он перешёл к количеству плугов, затем к *условной переписи*¹⁰ скота и к площади культивируемых земель. Он рекомендовал централизацию официальной статистики и её публикации.

В 1785 – 1789 гг. комитет по преобразованию Отель-Дьё (Bailly, Лавуазье, Лаплас и Tenon) провёл широкие исследования организации больниц во Франции и в остальной Европе. Его официальные выводы, основанные на документах, опросниках и непосредственных наблюдениях, были представлены в трёх отчётах (Bailly 1790). Tenon (1788) кроме того подробно описал организацию парижских больниц, исследовал их недостатки и предложил меры по их более рациональному устройству. Cabanis (1790) предложил более радикальные меры.

[3.3] Идеологи и Институт. Социологические исследования возобновились после падения Робеспьера. В 1795 г. Конвент учредил Национальный институт наук и искусств взамен закрытых в 1793 г. академий. В Институт был включён *второй класс* моральных и политических наук, и его официальной теорией была идеология¹¹, которую сформулировали Дестют де Траси и Кабанис, и появилось новое поколение, поколение идеологов. Их учение с упором на исследовании языка и знаков и понятием об осязаемом соотношении морального и физического повлияло на социологические исследования в этнографии и [общественной] гигиене, равно как и на правительственной администрации.

Выдающимся этнографом был Volney, – врач, а затем востоковед и путешественник, который точно описал и систематически исследовал общество Среднего Востока (1787). Воодушевившись вопросником, который в 1762 г. составил немецкий востоковед Michaelis, а также и инструкциями представителям Министерства внутренних дел, он адресовал свои *Статистические вопросы* (1795) дипломатическим работникам (Gaulmier 1951). [...]

В 1794 г. под влиянием трёх идеологов, Кабаниса, Пинеля и Vichat, Национальный Конвент решил преобразовать медицинское образование, причём особый упор был сделан на [общественную] гигиену и судебную медицину (Rosen 1946; 1958; Foucault 1963). Гигиена рассматривалась в сочетании с улучшением культурно-бытовых условий и филантропией (Cabanis 1803). Министр внутренних дел при Директории, François de Neufchâteau, физиократ и идеолог, приказал перевести сборник английских и немецких работ по *гуманитарным учреждениям*.

В 1802 г. был учреждён Совет по чистому и здоровому округу Сена, который после 1815 г. стал образцом для промышленных

городов. Врач в нём был социальным дознавателем и преобразователем. Гигиеническое движение продолжало развиваться и достигло высшей точки в периоде индустриализации. Полное собрание сочинений Кабаниса появилось в начале 1820-х годов¹². Карьера Gérando служит иллюстрацией непрерывности этого движения. Он был генеральным секретарём Министерства внутренних дел во время Империи и опубликовал две книги (1820; 1839).

[4] Другие исследования в континентальной Европе

В XVIII в. в других континентальных странах Европы информация, подобная той, которая достигается переписями, время от времени собиралась для правительств специально назначенными лицами. Никакого постоянного установления для сбора и табулирования существовавших сведений не было, и не применялись никакие стандартные методы работы. Неудивительно, что получаемые результаты были как правило неполны и ненадёжны.

Главной целью этих исследований было получение информации для обложения населения налогами и для военного планирования, и поэтому неудивительно, что её скрывали. В Австрии, Бельгии и нескольких других странах время от времени проводились подсчёты населения, строений, скота и др. объектов сельского хозяйства, торговли, промышленности и армии, обычно лишь в данном районе или городе. В Дании население было подсчитано в 1760 и 1787 гг.

В Швеции в соответствии с законом 1686 г. приходские регистры должны были указывать количество рождений и смертей и миграцию прихожан, равно как и их списки. Количество населения особенно заботило власти ввиду его заподозренного убывания в начале XVIII в. В 1748 г. было решено регулярно сохранять и исследовать эти записи в центральном архиве. Местное духовенство подсчитывало своих прихожан в стандартных формах и пересылали полученные результаты в центр через посредство церковной иерархии. Сводка за 1749 г., подготовленная Пером Варгентином, была, видимо, самым старым отчётом о национальной переписи, а в 1761 г. Шведская академия наук опубликовала дополнительные данные. Впрочем, большая часть этих сводок оставалась не исследованной, хоть в отличие от Франции результаты шведской переписи были опубликованы¹³.

В многочисленных немецких королевствах, княжествах и свободных городах основной заботой в XVII в. стала перестройка городов после разрушений Тридцатилетней войны. Чувствовалась поэтому необходимость в систематизации сведений о различных территориях. Термин *статистика* происходит от деятельности, направленной на выполнение этой потребности. Вначале *статистика* означала смесь географии, истории, законодательства, политической науки и общественной администрации¹⁴.

Герман Конринг, профессор в Брауншвейге, умерший в 1681 г., разработал набор категорий для характеристики государства. Он

особенно заботился о сравнении различных государств друг с другом, чётко описал свой метод, упорядочил свои источники и указал критерии для оценки их надёжности. В начале XVIII в. немецкие университеты обучали будущих чиновников его системе. Более поздняя гёттингенская школа *статистики*, выдающимися представителями которой были Ахенваль, Шлёцер и Ниман, совершенствовала систему Конринга, а в начале XIX в. *табличные* статистики поставили под сомнение эту описательную статистику. Они перекрестно классифицировали всё возрастающее количество сведений. В то же время в Германии образовались статистические общества, а в нескольких государствах возникли статистические агентства¹⁵. [...]

[5] Исследования в Англии в XIX в.

С 1801 г. под руководством Джона Рикмана начали проводиться переписи с 10-летним интервалом. Первые три сосредоточились на подсчёте жителей, семей и строений. Священники отчитывались за свои приходы, но качество их отчётов не было удовлетворительным.

Перепись 1831 г. впервые исследовала занятия жителей. Она оказалась важной и потому, что для парламентской реформы требовались точные демографические данные. Полученные сведения были включены в несколько сочинения (Colquhoun 1814; Playfair 1801; McCulloch 1837), схожие с нынешними статистическими сводками, ежегодниками, справочниками *Кто есть кто* и информационными бюллетенями.

[5.1] Исследования в сельской местности. Во второй половине XVIII в. в Англии произошла сельскохозяйственная революция¹⁶, и для решения неотложных проблем сельской местности были использованы новые методы социологических исследований. Артур Янг, умерший в 1820 г., проводил их и, в отличие от обширной литературы о путешествиях, которая сосредоточивалась на особых обычаях, он описывал действительный образ жизни в сельской местности и оценивал сельскохозяйственные ресурсы с целью усовершенствования растениеводства и скотоводства. Его обширные отчёты о путешествиях (1771a; 1771b; 1780; 1793) относились к сельской местности в Англии, Франции и Ирландии.

Но количественные методы в исследованиях сельской местности внёс шотландец Джон Синклер. Он был богатым землевладельцем, научным фермером, путешественником, членом парламента и писателем широкого плана. Синклер был знаком с немецкой статистической деятельностью¹⁷, и когда в 1755 г. шотландское духовенство провело частную перепись населения, он задумал аналогичное, но гораздо более амбициозное предприятие.

Оно стало настолько широким монументальным исследованием, что заняло более семи лет и было опубликовано в 21 томах (1791 – 1799). Статистику Синклер (т. 20, с. xiv) определил как

исследование состояния государства с целью установления количества счастья, испытываемого его жителями, и средств для его будущего улучшения.

В нескольких приложениях в последнем томе он ясно пояснил свою методологию. Работу Синклер завершил с помощью шотландского духовенства, получив от него отчёты о каждом из 881 приходов. Сердцевиной исследования был план вопросника, содержавшего более ста вопросов. Позднее он применял и более краткий вопросник, который требовал лишь заполнения чисел в заготовленных таблицах. В этих же приложениях Синклер перепечатал 23 последующих письма, которые он в течение нескольких лет разослал тем, кто не отвечал на его вопросы, и в них он поочерёдно умолял, обхаживал, убеждал и грозил. В конце концов ему пришлось послать нескольких *статистических миссионеров*, чтобы завершить отчёты некоторых приходов. Он составил хронологическую таблицу количеств возвращённых списков и убеждал священников дополнительно посылать ему информацию всякого рода и результаты их собственных изысканий.

Сам вопросник был разделён на несколько частей. Первые 40 вопросов были посвящены географии, геологии и естественной истории приходов. Вопросы 41 – 100 относились к населению (возраст, пол, занятие, религия, сословие и профессия, – знать, нетитулованное дворянство, священники, юристы и пр.), рождения, смерти, самоубийства, убийства, числа безработных, нищих, привычных пьяниц и пр.

Вопросы 101 – 116 имели дело с сельскохозяйственной продукцией, животноводством, минералами, и, наконец, был включён ряд вопросов о заработной плате, ценах, истории приходов, о характере людей, формах землевладения и о сравнении нынешних условий в приходе с предшествовавшими периодами.

Синклер придерживался вполне современного подхода к использованию количественных данных. Исследуя характер прихожан, он спрашивал, нравится ли им военная жизнь и стремился узнать, сколько человек в недавнем прошлом поступили на военную службу. Спрашивая, *склонны ли люди к гуманным и великодушным действиям*, он стремился узнать, *защищают и помогают ли они жертвам кораблекрушений* и т. д.

Ответы сильно различались по качеству и, конечно же, было очень трудно анализировать, сводить воедино и публиковать их. Для составления сводных таблиц по графствам Синклер нанял нескольких помощников, но результаты были окончательно, количественно использованы лишь в 1825 г. [кем?], уже после того, как Синклер покинул общественную жизнь. В конце концов его труд оказался сводкой приходских отчётов с несколькими таблицами по графствам, сведёнными им самим и его помощниками.

Его сочинение послужило полезным примером для переписи и показало, что она возможна. Он старался склонить европейские правительства к установлению переписей с 10-летним интервалом, и действительно способствовал тому, что переписи прошли в первой половине XIX в. во многих странах, а его двухтомное сводное сочинение было переведено на французский. Но его

выбор приходов в качестве отчитывающихся единиц, а духовенства – как счётчиков уже в то время был устаревшим методом, потому что ввиду индустриальной революции и роста городов приход почти перестал быть элементом местного управления. Позднее социологам, занимавшимся индустриальными и городскими проблемами, пришлось придумать иной подход, исследование одного дома вслед за другим. Эту новинку применял ещё Booth и др. (1889 – 1891) [см. § 5.4].

[5.2] Исследование проблем индустриализации. Чтобы понять причину необычного всплеска социологических исследований в Англии в 1780 – 1840 гг., следует вспомнить, что то было время громадных усилий, направленных на преобразование отживших социальных институтов, – законов о бедных, систем образования и местного самоуправления, общественного здравоохранения и самого парламента. Были учреждены независимые власти (authorities), комиссии и союзы по исправлению положения со штатными юристами, деловыми людьми, духовенством, педагогами, врачами и другими специалистами. Многие из этих реформаторов также исследовали социальные проблемы. Социальная политика, социальные исследования и преобразования, законодательные акты оказались частью единого и широкого движения.

Многие врачи были привлечены к реформам и социологическим исследованиям своей профессиональной деятельностью. Их ежедневные контакты с пациентами из рабочего класса и бедняков напоминали им о серьёзности проблем, связанных со здоровьем, питанием, бедностью и антисанитарными условиями жизни. Кроме того, преобладавшая в то время миазматическая теория возникновения и распространения заболеваний оправдывала их реформирующее мировоззрение.

Самая распространённая схема действий в социальной политике и при проведении социологических исследований и социальных реформ и введению законодательных актов была следующей. Во-первых, отдельное лицо или небольшая группа выявляли социальное зло и часто начинали изучать его. В результате этой инициативы более крупные и лучше организованные группы предпринимали новые исследования, ограниченные эксперименты и улучшали что-то. Все эти усилия возбуждали и формировали общественное мнение, привлекали внимание правительства и приводили к его действиям, – к созданию комитетов по расследованию, королевских комиссий, а в успешных случаях, и к законодательным актам, которые нацеливались на искоренение зла и предусматривали систему инспекций и другие подобные меры.

Самыми выдающимися из многочисленных реформаторов были, видимо, Howard, Eden, Kay-Shuttleworth и Chadwick¹⁸. Джон Хоуард, помещик, стал неутомимым исследователем тюремных условий и сторонником их преобразований. По его расчётам, он проехал более 42 тыс. миль [67 тыс. км] по всей

Англии и Европе. В своих книгах он (1777; 1789) подробно описал положение в сотнях тюрем и тюремных больниц: как заключённые проводили своё время; какое питание они получали; какими болезнями они страдали и какие неудобства переносили; каков был образ действий тюремной администрации. Он также указывал численность заключённых с разбивкой по полу и характеру преступлений в каждой английской тюрьме и т. д.

В книгах Хоуарда было полно сравнений обращения с преступниками в различных странах и предложений о его улучшении, и в результате усилий его самого и его сторонников во многих тюрьмах с заключёнными начали обращаться более гуманно.

Сэр Фредерик Мортон Иден был деловым человеком. Инфляция 1794 – 1795 гг. оказалась непосредственной причиной его эмпирического исследования количества бедняков и условий их жизни. Он посетил большое число приходов, вёл обширную переписку с местным духовенством и опубликовал трёхтомное сочинение (1797). В основном оно состояло из местных [приходских?] отчётов, иногда содержащих числовые данные о количестве населения, домов, обложенных налогами [обитатели которых были обложены и т. д.], об основных мануфактурах, о типичной заработной плате рабочих основных специальностей, рентной плате за фермы, ценах на продовольственные товары, об обществах взаимопомощи и их членах, о числе бедняков, и содержало подробное описание условий в приходских работных домах.

Иден кроме того представил 43 подробных семейных бюджетов чернорабочих, ткачей, шахтёров, каменщиков и др. Большую часть этого обстоятельного труда выполнил нанятый исследователь, что было новинкой, которую позднее переняли статистические общества. Свои рекомендации о преобразовании законов о бедных Иден обосновал обнаруженными фактами, а проведенное им изучение бюджетов и его подробный эмпирический метод были социологическими новинками, которые позднее усовершенствовал Ле-Пле и его школа.

Врач Джеймс Филиппс Кей-Шаттлуорт был активным санитарным реформатором и одним из учредителей Манчестерского статистического общества, а позднее помощником уполномоченного по исполнению законов о бедных. Его ранние изыскания (1832) оказались образцом для более подробных исследований, которые позднее начали предпринимать статистические общества¹⁸. После 1840 г. Кей-Шаттлуорт направил свои усилия на введение и совершенствование национальной системы образования.

Эдвин Чедвик был чиновником, который всю жизнь активно участвовал в парламентских и иных комиссиях, и главным составителем *Отчёта (Report 1842)*, приведшего к Постановлению 1848 г. об общественном здравоохранении и к учреждению Центрального отдела здравоохранения¹⁹. Этот же *Отчёт* оказался примером для последующих административных и парламентских исследований социальных проблем.

В 1830-е годы многие граждане, которые участвовали в проведении социальных реформ, создали большое число местных статистических обществ. Старейшие и наиболее действенные общества в Манчестере и Лондоне дожили до наших дней. Кроме того, по настоянию Биббиджа, Мальтуса и Кетле в Британской ассоциации продвижения науки была создана статистическая секция²⁰. Целью этих статистических обществ было улучшение жизни, основанное на количественном изучении проблем общества. Они учреждали комитеты для исследования здоровья, условий проживания, образования, участия в религиозной жизни и условий труда низших классов.

Исследования часто длились по несколько месяцев, и их стоимость исчислялась многими сотнями фунтов. Оплачиваемые счётчики ходили по домам и проводили интервью, основанные на подготовленных вопросниках. Результаты табулировались централизованно и представлялись на годичных собраниях Британской ассоциации продвижения науки или отдельных обществ и часто публиковались, например, в журнале Лондонского статистического общества.

Подобные исследования насчитывались десятками. Иногда в них применялось утонченное многомерное перекрестное табулирование, другие же работы основывались на различных существовавших регистрах.

[5.3] Упадок исследований социальных проблем. Тем не менее к концу 1840-х годов социологические исследования захирели. Основные цели реформаторов всё полнее выполнялись, поскольку парламент принял закон о фабриках и многие другие законы и меры, экономическое положение рабочих заметно улучшилось, а их политическая активность сошла на нет с поражением чартизма.

Было проведено лишь несколько оригинальных исследований, притом по несовершенной методике. Многие местные статистические общества исчезли; в Манчестере большое число членов-учредителей статистического общества всё более вовлекалось в гражданскую и политическую деятельность, а интерес Лондонского статистического общества сместился к общественному здравоохранению, статистике рождений, женитьб и смертей, состоянию здоровья военнослужащих, смертности в колониях, продолжительности жизни лиц различных занятий и т. д. Для этих целей был достаточен анализ опубликованных данных²¹.

Вообще интереса к методологии планирования изысканий, к сбору данных и их анализу по существу не было. Некоторое время Национальная ассоциация по продвижению социологии, основанная в 1857 г., спланировала лиц, занимавшихся эмпирическими исследованиями трудовых отношений, образования и социальными проблемами. В середине столетия отрыв социологических исследований от любых научных соображений частично обусловил их разъединение. Не было постоянного вовлечения заинтересованных исследователей, постоянного совершенства методов работы, не было традиции

заглядывать за пределы краткосрочных целей социальных перемен, не было и регулярной финансовой поддержки. Таковы были некоторые отрицательные стороны разъединения усилий.

Интеллектуальную жизнь Англии всё более формировал социальный дарвинизм. Открытие эволюции видов, успехи биологических наук и возрастающее признание расы и наследственности в качестве фундаментальных понятий социологического анализа изменили интеллектуальный климат. Прежние исследователи связывали преступления с нуждой и малограмотностью, но более новый взгляд разыскивал свидетельства наследственного вырождения или других физических или психологических недостатков. Людей низших классов, бедняков, уголовников и других горемык часто относили к худшему слою человечества. Многие с тревогой рассматривали социальную политику, направленную на облегчение их состояния, поскольку она лишила бы возможности исключительного выживания самых приспособленных, и тем самым обрекла бы нацию на медленное вырождение.

Тот же тип образованного индивидуума среднего или верхней части среднего класса с научным складом ума, который в 1830-е и 1840-е годы мог бы вступить в статистическое общество и проводить приходские исследования рабочего люда от двери к двери, – тот же самый тип позднее в том же столетии становился членом антропологического, этнографического или евгенического общества и проводил время в изучении первобытных культур или современной генеалогии и в измерении сотен черепов и частей тела в неустанных попытках классификации и типологии. Бремя доказательства [?] перешло к защитникам окружающей среды.

[5.4] Бут и исследование бедных. В 1880-е годы вновь началось научное изучение бедноты, высшей точкой которого оказались монументальные социологические исследования Чарлза Бута, который происходил из семьи ливерпульских судовладельцев. Он (1889 – 1891, т. 1, с. 6) хотел показать половине Лондона как жила другая его половина, а точнее пояснить

Количественное отношение бедности, невзгод и развращённости к регулярному заработку и описать общее состояние жизни каждого класса.

Собрав вокруг себя исследователей, подобных Беатрисе Вебб и Октавии Хилл, других социальных работников и экономистов-социологов, Бут начал свою работу в 1886 г., и его многотомный труд (1889 – 1891) выдержал несколько изданий.

В то время в Лондоне было уже 4 млн жителей. Бут вначале собрал существовавшую информацию по результатам переписей и 400 работ инспекторов, следивших за посещаемостью школ и регистрировавших сведения о каждой бедной семье. Всё это он сверил по данным, имевшимся в полиции, у санитарных инспекторов, в обществах взаимопомощи и многочисленных благотворительных организациях и агентствах, изучающих наёмных работников. Позднее Бут начал сверять полученные

данные по личным наблюдениям некоторых улиц и квартир и какое-то время прожил совместно с различными семьями рабочих.

Он подразделил лондонские семьи на восемь классов по величине и регулярности их заработков и, в богатых семьях, по числу домашней прислуги. Каждый городской квартал он охарактеризовал преобладавшим в нём классом семей, показал каждый класс определённым цветом и составил многоцветную схему Лондона.

Кроме того, Бут описал район за районом и улицу за улицей стилем жизни их жителей, их проблемами и видами на будущее. В частности, он указывал участие в религиозной и светской жизни, характеризовал отдых, посещение трактиров и местных добровольных организаций. В нескольких томах Бут описал заработную плату и условия работы рабочих различных профессий.

Окончательным результатом было самое подробное и крупное социологическое описание из числа когда-либо произведенных. Оно взволновало социальное сознание современников и в 1908 г. привело к принятию закона о пенсиях по старости, установлению минимальной заработной платы за тяжёлый труд, государственному обеспечению больных и инвалидов и к зарождению страхования от безработицы.

Труд Бута склонил других исследователей к изучению городской жизни, которые усовершенствовали его методы работы. Самыми примечательными оказались сочинения Rowntree (1901), – обследование одного города, – и Bowley и др. (1915), которые исследовали несколько городов и впервые систематически применили выборочный метод взамен сплошного перечисления²².

Современные критики подметили, что статистический анализ причин бедности у Бута можно было во многом усовершенствовать. Так, Юл (1899), исходя из его данных, впервые применил множественную регрессию и корреляцию. Тем не менее, статистические методы Гальтона, Эджуорта, Пирсона и Юла были включены в социологические исследования лишь в XX в. после того, как они начали применяться в биологических, сельскохозяйственных и психологических исследованиях.

[6] Исследования во Франции в XIX в.

В 1800/1801 гг. появилась *Статистика Франции (Statistique de la République)*²³, см. Gille (1964, с. 121 – 147), а в 1803 г. был упразднён второй класс Института Франции, что снова привело к правительственным расследованиям. До 1806 г. к статистике относились с подлинным воодушевлением, которое поддерживалось публикацией журнала *Annales de statistique* (1802 – 1804) и обществом *Société de Statistique* (1803 – 1806).

Национальное Статистическое бюро публиковало мемуары, составленные отдельными округами и основанные на общем исследовании, которое в 1801 г. затребовал тогдашний Министр внутренних дел Шапталль (Pigeire 1932) и Duquesnoy. Вопросы, заданные тогда префектам, относились к размещению, условиям жизни и передвижениям населения в 1789 и 1801 гг.; к *состоянию граждан* с [их] разбивкой на шесть классов и его изменениям в

период 1789 – 1901 гг.; религиозному и светскому образам жизни; привычкам и обычаям; изменениям в сельском хозяйстве и промышленности, происшедшим с 1789 г.

[В мемуарах] было обнаружено много ошибок, и в 1806 г. их публикация прекратилась, а Статистическое бюро реорганизовано, хотя сочинения, основанные на собранных материалах, продолжали появляться примерно до 1810 г. Бюро также получало регулярные отчёты от префектов и предпринимало специальные исследования, например, так называемую перепись 1801 г., проведенную мэрами (Reinhard 1961), и оценку населения в 1802 г., в которой участвовал Лаплас, описавший её в своей *Аналитической теории вероятностей*. После 1806 г. руководить Бюро стал Coquebert de Monbret, интересовавшийся специальными вопросами.

В 1806 г. было проведено промышленное и сельскохозяйственное исследование, в следующем году – исследование, в соответствии с его описанием, видов занятий и различных религий более чем в 16 категориях. Вторая перепись населения прошла в 1806 г. Идея общего исследования была вновь воспринята в 1810 г. Префектам разослали вопросник из 334 вопросов, а в следующем году было оценено население. Впрочем, провал общего исследования привёл к возврату к специальным обследованиям (например, промышленности и путей сообщения), см. Montavilet (1813). В сентябре 1812 г. Статистическое бюро упразднили²⁴, и с тех пор сбор промышленной и сельскохозяйственной статистики оказался исключительным делом правительства.

[6.1] Исследование проблем индустриализации. Значение, которое различные революционные правительства придавали социологическим исследованиям, поясняет первоначальное недоверие к ним со стороны монархии, реставрированной в 1815 г. Но рост городов и индустриализация сопровождались социальными проблемами, которые выпячивались в парламентских прениях и в прессе, так что для их изучения было создано несколько учреждений.

Филантропическое движение продолжалось при Реставрации, и мы упомянем парижское Королевское общество для улучшения положения в тюрьмах (*Société Royale pour l'Amélioration de Prisons*) и Промышленное общество Мюлуза (*Société Industrielle de Mulhouse*). Появилось и раннее классическое сочинение о пенологии (Villermé 1820). Расширились исследования народного здравоохранения, основанные на наблюдениях и количественных мерах (Ackerknecht 1948) и в 1820 г. мощь этого движения привела к учреждению Королевской академии медицины, Высшего совета здравоохранения в 1822 г. и окружных советов здравоохранения в 1822 – 1830 гг. В 1829 г. появился и журнал *Annales de l'hygiène publique et de médecine légale*, который твёрдо поддерживал роль врача как исследователя и социального реформатора.

В 1817 г. префект Шаброль распорядился о переписи Парижа; результаты были опубликованы в шести томах (*Recherches 1821 –*

1860), в которых помимо данных о населении сообщались сведения о потреблении товаров, уровнях богатства, причинах смертей и самоубийств и т. д. Стала издаваться ежегодная статистика отправления уголовного судопроизводства France (1827 – 1907) с распределением типов преступлений и преступников²⁵ и сборники (France 1819 – 1908), в которых сообщались сведения об уровне образования новобранцев. Эта информация способствовала развитию *моральной статистики*, одним из первых достижений которой было исследование преступлений и образования (Valby и др. 1829).

Июльская революция 1830 г. поспособствовала социологическим исследованиям. В 1832 г. был восстановлен второй класс Института Франции под названием *Академия моральных и политических наук*, что позволило работникам народного здравоохранения и статистикам проводить свои собственные изыскания, поскольку эта академия организовала призовые соревнования. Двумя победителями оказались Frègier (1840) и Buret (1840). В 1832 г. началось создание *Общей статистики Франции (Statistique générale de France)*²⁶. Она же в 1834 и 1836 гг. провела первые надёжные переписи всей Франции за 15 лет и добыла обширные и основательные материалы.

Внимание к рабочему вопросу в 1828 г. привлёк Villeneuve-Bargemont своим отчётом *Rapport sur le département du Nord*, опубликованным в 1837 г., см. Beautot (1939 – 1943). После мятежей 1831 и 1834 гг. эту тему нельзя было больше замалчивать. Рабочим классам посвящалось возрастающее число исследований и книг, см. Rigaudias (1936). В 1834 г. Академия моральных [и политических] наук обязала Benoistin de Chateauneuf и Villermé исследовать как можно точнее физическое и моральное состояние этих классов, см. Villermé (1840).

[6.2] “Моральная статистика” Герри. Герри, адвокат, родившийся в г. Туре, центре округа Эндр и Луара, заинтересовался официальной статистикой преступлений и изучил (1833) соотношение преступности и уровня образования. Ему пришлось столкнуться с двумя методическими проблемами, с отсутствием какой-либо меры статистической корреляции²⁷ и трудностью применения общих мер, как например, преступности или среднего национального уровня образования; исследования проблемы индивидуального поведения и выяснения, к примеру, будут ли те, кто лучше образован, более или менее преступны чем другие.

Установив, что на протяжении шести лет преступность имела склонность к устойчивости (и приписывая это обстоятельство систематическому и постоянному естеству соответствующих причин), он постарался сравнить распределения преступности и образования, расположив 85 округов страны по преступности, а затем по уровню грамотности. Для преступлений против личности крайние округа в первом распределении не совпадали с крайними во втором, и, напротив, совпадали, если преступность рассматривалась только по отношению к собственности.

Герри решил, что между образованием и преступностью не было никакой отрицательной корреляции²⁸, и что значимым возможно являлся уровень индустриализации. О других его исследованиях см. Douglas (1968).

[6.3] Parent-Duchâtelet и проституция. Он был членом группы, основавшей *Annales d'hygiène* ... и автором двух важных сочинений (1834; 1836). Во втором из них он привёл сводку 30 отчётов, составленных с 1825 г. для Совета по чистому и здоровому департаменту Сена (*Conseil de Salubrité de la Seine*), первое же было для своего времени одним из лучших исследований рассматриваемой темы.

Вызвано оно было и филантропическими и административными причинами. Филантроп, знакомый автора, стремился помочь проституткам правдивыми сведениями о них и выяснил, что они живут совсем в ином мире и что прежде всего следовало бы ознакомиться с ними. В то же время муниципалитет, которому удалось ограничить проституцию и которого запрашивали из-за рубежа о применённых им методах, хотел бы увидеть оценку своей работы.

В своём исследовании Parent-Duchâtelet использовал документы (например, полицейские картотеки), личные наблюдения и интервью (необычные для Франции того времени, когда исследователи, в том числе Виллерме, добывали информацию косвенно, по отчётам других лиц) и применяя статистический метод, составив около 150 таблиц. Он попытался определить число проституток и его изменение во времени; их региональное и социальное происхождение; физиологические приметы и социальный характер; их отношение к женитьбе и религии; причины, по которым они стали проститутками; и как они бросают своё занятие.

Parent-Duchâtelet закончил исследование признанием неизбежности проституции и доводами в пользу необходимости моральной и матеральной заботы о проститутках (обсуждаемый в то время вопрос) и создания приютов для проживания раскаявшихся.

Он приложил проект закона о наблюдении за нарушениями общественной благопристойности, вызванными проституцией. Весь его научно-нейтральный подход к этой проблеме, обременённой моральными запретами, равно как и непосредственные наблюдения и интервью, были поразительно современными.

[6.4] Виллерме: изучение жизни рабочих текстильной промышленности. Он родился в Париже в 1782 г. Ему было 52 года, когда он начал это своё грандиозное исследование и прожил шесть лет в наблюдениях в главнейших центрах текстильной промышленности. Его двухтомный отчёт завершил карьеру, которая началась в 1819 г. и была посвящена [общественной] гигиене и статистике. В период между 1819 и 1834 гг. Виллерме опубликовал более 40 статей в семи журналах (Guérard 1864), в том числе простые сводки наблюдений об [общественной] гигиене или других социальных проблемах, которые, однако,

выявили его умение наблюдать. Научился он этому в 1802 – 1814 гг. будучи хирургом в наполеоновских армиях, а затем обучаясь с целью стать врачом (медицинскую степень он получил в 1818 г.)²⁹.

С 1822 г. Виллерме начал в основном интересоваться статистикой смертности в Париже и во Франции, особенно в связи с доходами. В нескольких статьях он показал, что смертность среди бедняков была намного выше, чем среди состоятельных (Vedrenne-Villeneuve 1961).

В своём новом исследовании Виллерме применил статистические данные и свои собственные качественные наблюдения. Некоторые данные он отыскал в ежегодных отчётах округов, хотя в основном собирал их сам. В них сообщалось о числе рабочих (которое трудно было установить, поскольку до 1851 г. сведений о занятости работающих не было); о среднем тарифе заработной платы рабочих различных специальностей (эти сведения предоставляли владельцы предприятий и мастера); продолжительности рабочего дня; о демографических данных (рождениях, женитьбах, количестве детей и внебрачных рождений) и бюджетах рабочих семей.

Собственные наблюдения Виллерме относились к чистоте цехов и квартир рабочих; их одежде и питанию; отдельным сторонам их поведения (к примеру, к потреблению алкоголя и [посещению ?] проституток). При использовании статистики и собственных качественных наблюдений Виллерме вводил показатели. Так, большое число внебрачных рождений он считал надёжным свидетельством нарушения обычаев, а качественными признаками благосостояния служили ежемесячное (а не ежедневное или еженедельное) получение заработной платы, потребление вина за воскресным обедом, наличие оконных занавесок и владение зонтиком.

При истолковании статистических данных Виллерме, как и Герри (§ 6.2), столкнулся с двумя трудностями: отсутствовала мера корреляции, и, что было важнее, по самой своей сути данные описывали районы, преимущественно населённые рабочими, а не самих рабочих. Ему поэтому пришлось прибегнуть к статистическим вычислениям, часто весьма интересным. Так, он выяснил, что в Амьене более 70% призывников из рабочих семей признавались негодными по состоянию здоровья, тогда как отклонялось только 50% призывников из других семей (Villermé 1840, т. 1, с. 311 – 317).

В этом томе Виллерме систематически собрал все свои факты, расположив их почти одинаковым образом в каждом изученном им промышленном районе. Второй том, представленный в 1837 г. Академии моральных [и политических] наук, большинство этих фактов было проанализировано³⁰.

Исследование Виллерме было менее методически разработано, чем у Parent-Duchâtelet. Он интервьюировал только рабочих шёлковой промышленности в Лионе, вообще же пользовался сведениями осведомителей, обсуждение же его темы вызвало гораздо больше споров. В конце своего тщательного описания Виллерме заявил, что участь рабочих слегка улучшилась, однако

социалистические круги, и особенно Бюре, встретили это утверждение в штыки. В то же время его факты о заработной плате и бюджетах выразительно выявили недостаточность средств у рабочих и их жалкий образ жизни, а его осуждение фабричного труда детей очень ранних возрастов сделало его мишенью для атак правоверных либералов и сторонников установленного порядка.

Споры по поводу Виллерме продолжались до недавнего времени. Rigaudias (1936) безоговорочно поддержал Бюре, а Fourastié (1951) одобрил точность его описательных и статистических сведений и основательность его выводов. В мае – июне 1839 г. факты о детском труде, собранные Виллерме, обсуждались и критиковались в обеих палатах парламента Великобритании и 22 марта 1841 г. были приняты во внимание при введении закона, регулирующего детский труд.

[6.5] Исследование сельскохозяйственного и промышленного труда. В мае 1848 г., после падения Людовика XVIII³¹ соответствующее решение приняло Национальное собрание. Исследование занялось теми же вопросами, которые изучали Герри, Виллерме и Бюре. Оно было закончено в 1850 г. и оказалось последним крупным официальным исследованием основных проблем страны.

После 1852 г. правительство назначало либо статистические исследования всей страны, либо подробные изыскания ограниченных проблем. Примером в основном монографического изучения, проводимого отдельными лицами, может служить работа Ле-Пле (Pitts 1968). Решение Национального собрания привело к ожесточённым идеологическим спорам о социальных исследованиях (Rigaudias 1936). В период 1840 – 1848 гг. социалисты неоднократно но тщетно требовали официального обследования условий рабочего класса. Соответствующая петиция была отклонена в 1845 г., но, заимев своим представителем Луи Блана, они потребовали создания министерства труда. Умеренные и консерваторы были против, и вместо этого предложили обследование такого рода, который они раньше отвергали.

Социологические обследования, основанные на наблюдении фактов, стали восприниматься как [проявление] умеренной и консервативной буржуазной идеологии и потому отвергаться различными течениями социалистической мысли. Это открытое противостояние продолжалось до начала периода Дюркгейма и его школы.

Вопросник в указанном выше обследовании [которое, стало быть, состоялось] содержал 29 вопросов об основных проблемах в каждом районе, общем состоянии промышленности, экономическом и социальном положении промышленных рабочих и общем состоянии сельского хозяйства. Ответы должны были представляться местными комиссиями из работодателя и рабочего от каждой индустриальной специальности, заседающими под председательством мировых судей. Но обследование было затруднено неточными формулировками

вопросов (что подчеркнул Gille (1964)) и сомнениями правительства и рабочих в его пользу. По мнению Rigaudias недоверие рабочих оказалось самым серьёзным препятствием.

Ответы были получены из 76% районов, так что их доля была высока, но Национальное собрание после краткого обсуждения распорядилось передать их в Министерство промышленности и торговли без каких-либо указаний об их публикации. Даже к сегодняшнему дню историки опубликовали лишь малую часть результатов указанного обследования 1848 г. [!] (Gossez 1904; Kahan-Rabecq 1934 – 1935; 1939; Vincienne & Courtois 1958). В Национальном архиве документы неполны; отсутствуют материалы 28 округов, почти всех крупных промышленных центров и всех крупных индустриальных городов кроме Марселя. Исследование архивов работы местных комиссий (Vidalenc 1948; Agulhon 1958; Guillaume 1962), о которых сообщил Gille (1964, с. 216), показывает, что ответы на вопросник отражали точку зрения ведущих граждан городов, а иногда просто председательствовавшего мирового судьи.

Провал этого обследования послужил поводом для отказа от национальных социологических исследований, проводимых представителями правительства, и к разрыву между социализмом [!] и подобными эмпирическими исследованиями. Труд Le Play (1855) стал началом нового типа исследований и оказался самым важным событием в истории социологических исследований в течение Второй Империи [1852 – 1870]. Подобно гигиенистам прошлого времени и особенно Виллерме, Ле-Пле использовал непосредственные наблюдения и применил монографический метод.

Основывая схему всемирного преобразования общества на своём *методе наблюдений*, он походил скорее на приверженцев традиции и социалистов, а не на гигиенистов, но отличался от социалистов, поскольку стремился к восстановлению традиционных принципов, а не к новому социальному порядку. Заметим, наконец, что Ле-Пле, как и все предшествовавшие социальные философы и исследователи, находился вне системы университетов. Примирение социалистической мысли и системы университетов с эмпирическими социологическими исследованиями произошло только при жизни поколения Гарда и Дюркгейма и создания Бюро (позже, Министерства) труда, что было одобрено парламентской секцией социалистов.

[7] Исследования в Германии в XIX в.

В Англии социологические исследования в основном проводили отдельные лица и добровольные общества, но в Германии их главным образом осуществляли учёные и профессиональные организации. В то время немецкие университеты представляли собой самую передовую систему высшего образования в мире³². Историческая школа экономики, которая отрицала взгляды английских политических экономистов, внесла наибольший вклад и в социологические исследования, и в постепенное возникновение социологии как отдельной дисциплины.

Экономисты этой школы благоволили людям, находящимися в неблагоприятных условиях и твёрдо верили, что их исследования будут способствовать прогрессивной социальной политике и решению социальных проблем.

В 1860-х и 1870-х годах под влиянием Кетле значительными темами исследований стали моральная статистика и демография. Эрнст Энгель, будущий глава прусского, а затем германского статистического бюро, использовал данные Ле-Пле и Кетле о бюджетах и их методы³³ для доказательства своего закона. В соответствии с ним, вне зависимости от полного размера семейного бюджета, очередность потребностей остаётся без изменений; более того, чем беднее семья, тем значительно больше оказывается доля бюджета, затрачиваемая на её питание. [...]

[7.1] Тённис и его социографический метод. Фердинанд Тённис всю жизнь боролся против узкого понимания эмпирической социологии как простого сбора фактов. В противоположность этому, он ввёл понятие *социографии*, которая включала систематическое наблюдение, изучение характерных случаев и другие качественные методы, равно как и статистическое описание. Целью социографии было выявление эмпирических законов методом индукции, применяемым к систематически собранным данным. Эта дисциплина должна была быть одной из трёх ветвей социологии [?], равной теории по своей значимости. Сам Тённис изучал статистику вместе с Энгелем, который убедил его в значимости достижений Кетле.

С 1895 г. и до конца жизни он публиковал серьёзные статистические монографии о землевладении и социальной структуре сельского хозяйства; демографии; преступности; самоубийствах и голосовании, причём большинство этих изысканий относилось к его родной земле Шлезвиг-Гольштейн. Он ввёл меру ассоциации и разработал метод исследования корреляции между временными рядами. Позднее в жизни Тённис неоднократно призывал к созданию социографических *обсерваторий*, в которых специалисты по многим дисциплинам совместно с лицами либеральных профессий и образованными любителями изучали бы факты социальной жизни, и особенно морально значимые. Тем не менее, многие методы Тённиса были ошибочны, и он не смог воодушевить своих коллег своими планами.

[7.2] Исследования сельскохозяйственных проблем. После объединения Германии в 1871 г. эти проблемы привлекли к себе серьёзное внимание. В сельской местности преобладала нищета и невежество, а немецкие крестьяне постепенно вытеснялись польскими батраками, особенно в Восточной Пруссии. Многие крестьяне эмигрировали; другие постепенно становились деревенскими пролетариями, так как юнкера³⁴ воспользовались покровительственными тарифами и вводили капиталистические методы сельскохозяйственного производства. Прежний *отеческий* тип трудовых отношений отовсюду исчезал.

Первое крупное исследование сельскохозяйственных рабочих, которое занялось этими проблемами под руководством Теодора

фон Гольца, прошло в 1874 – 1875 гг. по заданию Конгресса немецких землевладельцев. Учрежденное в 1872 г. Общество социальной политики провело, в частности, два исследования сельскохозяйственных рабочих, изучило ростовщичество в сельских районах и производство на дому. Общество было и профессиональным союзом, и группой влияния и исследовательской организацией. Следствия из проведенных изысканий обсуждались на конференциях Общества с целью повлиять на социальную политику и законодательство. Его исполнительный комитет запланировал дальнейшие исследования и заручился поддержкой ведущих профессоров, которые в свою очередь привлекли к исследованиям своих студентов. Свою первую эмпирическую работу Макс Вебер (1892) выполнил под покровительством Общества.

Перед указанными исследованиями обычно составлялись вопросники по таким темам как землевладение, сельскохозяйственное производство, заработная плата, условия жизни, состав работающих, степень воровства и пьянства. Ответы предоставляли землевладельцы, духовенство, врачи, нотариусы, учителя, члены сельскохозяйственных обществ и другие знающие лица. Слабость этого метода работы состояла в том, что он предполагал точное знание у осведомителей, что вопросы были сформулированы неточно и сгруппированы случайным образом, что доля полученных ответов была незначительной и притом не использовалась достаточно хорошо.

Короче говоря, качество исследования не заботило членов Общества и только Schnapper-Arndt (1888) резко критиковал его метод. Он (1883) ранее опубликовал весьма подробное полевое изучение сельской жизни, основанное на нескольких месяцах своих усердных и явно обязанных Ле-Пле наблюдениях. К сожалению, воздействие его публикаций на современников оказалось очень небольшим.

Религиозные организации, как, например, Евангелический социальный конгресс, также проводили сельские исследования. В основном они, конечно же, интересовались моралью, религией и грамотностью населения, а сведения добывали от священнослужителей.

[7.3] Исследования промышленных рабочих. После ухода Бисмарка в отставку и отмены в 1890 г. законов, запрещающих социалистическую политическую деятельность, стал важным вопрос о рабочем классе. Повторявшийся успех социалистов на выборах испугал средний и высший классы общества, а возраставшее международное соперничество и требования профсоюзов о сокращении рабочего дня заставили задуматься о продуктивности и возможностях германской промышленности конкурировать с Англией на мировых рынках³⁵.

В 1890 г. Гёре, молодой студент-богослов, а позднее социал-демократический депутат Рейхстага, решил выяснить всю правду, как он надеялся, о рабочем классе. Три месяца он работал на фабрике, притворяясь подмастерьем, и проводил время и на работе, и после неё точно так же, как остальные рабочие. Каждую

ночь он записывал свои переживания, а затем опубликовал их (Göhre 1891). Его книга оказалась примечательным документом социальной структуры рядовых фабричных рабочих, их образа жизни, их устремлений и религиозных убеждений и была широко воспринята научными кругами и общественностью.

Через год он и Макс Вебер объединились под покровительством упомянутого Евангелического социального конгресса в качестве руководителей исследования сельскохозяйственных рабочих, а позднее Гёре отредактировал автобиографии ряда рабочих. В 1907 – 1911 гг. Адольф Левенштейн, рабочий-самоучка, записал результаты видимо первого крупномасштабного исследования поведения и мнений. Он разослал шахтёрам и рабочим сталелитейной и текстильной промышленности 8 тысяч вопросников, использовав для этого лавинообразный метод и начав рассылку с многих своих друзей.

Он получил примечательно много ответов (63%), вопросник же, несмотря на свои недостатки, выявил мнения и отношения рабочих о многих (ко многим и т. д.) существенных тогдашних проблемах: об их материальных и политических надеждах и пожеланиях, устремлениях, религиозных верований, политической деятельности, об их культурных занятиях и развлечениях, удовлетворённости работой или разочарования в ней, привычном потреблении спиртных напитков, – всё это в дополнение к стандартным вопросам о социальном происхождении и заработной плате.

Вначале Левенштейн отказывался публиковать свои труды, но Вебер и другие лица убедили его привести ответы в виде классифицированных таблиц. Именно так он (1912) и составил свою книгу. Во многом Левенштейн последовал советам Вебера (1909) о порядке анализирования подобных исследований, и таким образом предвосхитил некоторые современные методы.

Сам Вебер был в то время основной движущей силой исследования промышленных рабочих Обществом социальной политики. Оно предусматривалось как крупномасштабное изучение карьер, социального происхождения и образа жизни рабочих. Некоторые данные должны были быть получены от них самих. Вебер, однако, имел в виду и испытание некоторых предположений о продуктивности рабочих, и, в частности, выяснение, в какой мере можно было бы приспособить лабораторные методы экспериментальной психофизики к опытам и исследованиям в заводских условиях.

Готовясь к исследованиям, он провёл целое лето, наблюдая рабочих текстильной фабрики и анализируя их производительность. В то же время он составил план действий для исследователей Общества и пояснение теоретического и методического обоснования всей работы. Само исследование закончилось неудачно, потому что значительное большинство рабочих отказалось участвовать в нём. И всё же предварительное теоретическое и статистическое изучение темы провозгласили прорывом. Оно действительно было весьма утонченным, хоть и

отражало тогдашний психофизический подход к продуктивности рабочих³⁶.

Собственное отношение Вебера к нередко утомительным обязанностям, связанными с социологическими исследованиями, ясно заметно в его знаменитой речи (1919/1946, с. 135):

Ни один социолог [...] даже в старости не должен считать себя выше выполнения в уме десятков тысяч простейших вычислений, притом быть может многие месяцы подряд.

Собственный план Вебера поощрения социологических исследований в Германии не был выполнен, но труды, описанные в нашем очерке, ясно показывают, что не было недостатка в предвосхищении его упора на сборе и анализе эмпирических социальных данных. Написать заново историю социологии, чтобы дать полный отчёт об этих прецедентах, – вот давно уже требуемая задача.

[8] Добавление. Исследования в США в конце XIX века

В США, ранние социологи научного направления в основном были авторами учебников, а не исследователями. Вплоть до первой мировой войны основную долю эмпирических социологических исследований независимо от университетов проводили социальные работники, филантропы, работники народного здравоохранения и благотворительных организаций, журналисты и реформаторы, а также некоторые научные работники по социальной патологии.

Эти работы были в слабой степени объединены в широком реформистском движении, которое включало организацию советов и обучение социально слабым слоёв населения в специальных домах и социологические изыскания.

Образцом для них служили английские реформаторы, особенно Бут, а упомянутые дома, впервые появившиеся в Англии, располагались в трущобах. Управляли ими выпускники колледжей, которые, в частности, собирали и публиковали социологические данные об условиях жизни в трущобах, чтобы тем самым побудить городские власти исправить положение. Одним из результатов их деятельности было появление книги Addams (1895), в которую был включён достопримечательный очерк Florence Kelley о работе в условиях усиленной эксплуатации. Woods со своими сотрудниками, работая в Бостоне, полагал, что студенческие общежития в университетах были *лабораториями социологии*. Результатом его исследований были книги (1898; 1902 и др.).

Таковы были первые систематические эмпирические изучения приспособления и ассимиляции иммигрантов [?], т. е. проблемы, которая оказалась в центре внимания социологов. Самого высокого научного уровня достигло в то время исследование городов Du Bois (1899). Нанятый Пенсильванским университетом, он провёл 15 месяцев среди негритянского населения Филадельфии, посещал религиозные, деловые и политические собрания и социальные встречи, а также семьи в домашних условиях. Он, в частности, собрал сведения о расовой дискриминации при найме на работу. Свои результаты Дюбуа

анализировал в составленных таблицах и диаграммах и подкреплял свои выводы систематическими сравнениями условий [на рынке труда?] белых и чёрных рабочих.

Позднее он организовал факультет социологии и *социологическую лабораторию* в университете г. Атланта. Исследования Дюбуа и его студентов об условиях [жизни и работы] различных социальных слоёв негритянского населения в конце XIX в. представляют собой прочный вклад в социальную историю США.

Во втором десятилетии XX в. социологические исследования переживали расцвет. Идея о них зародилась в усилиях Общества организации благотворительности направить общественное мнение к социальным реформам. Всё движение финансировало недавно созданный фонд Russell Sage, а его мотором был Paul Kellogg, редактор журнала *Charities and Commons* (позднейшее название, *Survey*).

Результаты социологических исследований рекламировали на крупных выставках. Самое обширное исследование произошло в Питтсбурге и было опубликовано в шести томах (*Survey* 1909 – 1914), наполненных сведениями и фотографиями. Оно привело к выставке в том же городе, на которой происходили лекции, собрания и конференции городских обществ. В дальнейшем были проведены исследования меньшего масштаба, самым заметным из которых было Спрингфилдское (*Survey* 1918 – 1920) под руководством Шелби Гаррисона. Eaton & Harrison (1930) опубликовали библиографию из 2775 названий и изысканий, в которых так или иначе использовались методы социологических изысканий.

Впрочем, эти методы были усовершенствованы и систематически применены в профессионально руководимых социологических исследованиях лишь в конце 1930-х годов и в годы первой мировой войны. Переход от кабинетной и прикладной социологии, занимавшейся социальной патологией, к социологическим исследованиям в современном смысле произошёл в Колумбийском (в Нью-Йорке) и Чикагском университетах, в этих ведущих учебных центрах социологических исследований тогдашнего времени.

Преподаватели и выпускники Колумбийского университета прокладывали путь к применению [новых] статистических и исследовательских методов. Исключительной особенностью факультета [?] этого университета было наличие статистической лаборатории, оборудованной вычислительной техникой. Студентов обязывали собирать отчёты благотворительных организаций Нью-Йорка и приводить их к *научному виду* при помощи табулирования и статистического анализа. Под руководством Richmond Mayo-Smith и Franklin H. Giddings эта работа ещё не была утонченной, но молодое поколение социологов, особенно Ogburn и Chapin, первыми в американских социологических исследованиях применили статистические методы, разработанные английскими математическими статистиками.

При анализе уровня жизни в Вашингтоне Ogburn (1918 – 1919) применил множественную регрессию, а книга Chapin (1920) была первым современным учебником по этой теме. В нём были описаны случайный выбор, составление планов интервью, кодирование и табулирование данных при помощи перфокарт и специального оборудования. Stuart Rice, другой выпускник Колумбийского университета, был пионером в применении статистических методов в политической науке.

В Чикаго Thomas руководил переходом от [формулирования] социальных проблем к систематической полевой работе и его (в соавторстве) важнейшей работой в этом направлении была книга Thomas и др. (1918 – 1920). В течение восьми периодов полевой работы он собрал для анализа 8 тысяч документов и полностью показал социальные изменения, происшедшие в группе иммигрантов в процессе приспособления к американскому обществу. В этой книге были рассмотрены многие вопросы, которые позднее начали заботить чикагских социологов: иммиграция, ассимиляция, распад семьи в городских условиях, изменения в культуре. Упор был сделан скорее на процессы, а не на структуры. Книга была методическим примером чикагского стиля исследований, более успешного в выявлении, чем в доказательствах.

В конце 1920-х годов ряд организационных мероприятий в Чикагском университете облегчил объединение научных социологических исследований. Так, были назначены профессора по исследованиям, продвижение преподавательского состава стало основываться на результатах исследований, диссертации студентов начали составляться под действенным руководством и посвящались родственным темам, исследования финансировались фондами и общественными организациями и появилось междисциплинарное сотрудничество. Возникла особая сосредоточенная и координированная работа чикагских социологов в области социологии городов. Это привело к преобладанию США [?] в течение 1920-х и 1930-х годов.

Краткие сведения об упомянутых лицах

Blanc, Louis Jean, Бланк, Луи Жан, 1811 – 1882, социалист, историк, журналист

Chaptal, Jean-Antoine, Шапгаль, Жан Антуан, 1756 – 1832, химик, государственный деятель

Coquebert de Monbret, Charles, 1755 – 1831, профессор маркшейдерии

Durkheim, David Emile, Дюркгейм, Давид Эмиль, 1858 – 1917, социолог, психолог, один из основателей социологии

Goltz, Theodore von, Гольц, Теодор, 1836 – 1905, экономист сельского хозяйства

Hill, Octavia, Хилл, Октавия, 1838 – 1912, социолог, без формального образования. В 1905 г. – член королевской комиссии по закону о бедных

Kellogg, Paul, Келлог, Пол, 1879 – 1958, журналист, реформатор

Le Play, Pierre Guillaume Frédéric, Ле-Пле, Пьер Гильом Фредерик, 1806 – 1882, инженер, экономист, социолог.

Исследователь бюджетов рабочих

Montyon, Jean Baptiste Antoine Auge de, Монтион, Жан Баптист Антуан Ож, 1733 – 1820, юрист, филантроп. Оставил значительные средства для поощрения различных наук. Премия Монтиона является высшей французской наградой по статистике

Tarde Gabriel, Тард, Габриель, 1843 – 1904, социолог. В 1901 г. в Петербурге был опубликован перевод его *Социалистической логики*

Webb, Beatrice Potter, Вебб, Беатриса, 1858 – 1943, либеральный реформатор. Основные работы написаны совместно с мужем Сиднеем

Weber, Max, Вебер, Макс, 1864 – 1920, многосторонний учёный, классик социологии

Примечания

1. Историческая область.
2. Вобан предложил ввести единый налог на всех граждан вне зависимости от их сословия, чем разгневал короля.
3. Грубая ошибка, повторенная в § 3.2: Французская академия изучала и изучает французский язык и литературу. Следует читать Парижская академия наук.
4. О политической арифметике в Англии по крайней мере косвенно сообщает FitzPatrick (1960). Особо следует учитывать медицинскую статистику (Greenwood 1941 – 1943).
5. Вне зависимости от мнения Дидро политическая арифметика должна была изучать и изучала экономику.
6. Если списки прихожан составлены, то зачем было нужно дополнительно устанавливать численность населения?
7. Точность и этой, и другой (см. ниже) оценки явно завышена, что, впрочем, соответствовало традиции. Но и авторы (§ 7.3) напрасно, пусть даже и верно, упоминали *8000 вопросников*; мы написали *8 тысяч ...*
8. В известной мере это направление работы Кондорсе использовал Лаплас, однако Годхантер (1865, с. 352) прямо заявил, что *во многих местах почти невозможно установить, что он имеет в виду*. Здесь же авторы упомянули два мемуара Лапласа, но указали только соответствующие годы, да и то неудачно: не было сказано, что мемуар, *написанный* (а точнее, представленный) в 1778 г., был опубликован в 1781 г. Второй же мемуар был фактически посвящен выборочной оценке населения Франции, о чём авторы не сообщили. См. Шейнин (2013, § 8.1, описание гл. 6 *Аналитической теории вероятностей*).
9. Старейшая парижская больница. Её название в переводе означает *Приют божий*.
10. Здесь Лавуазье явно применил элементы выборочного обследования. Условную (у авторов, *гипотетическую*) перепись можно было бы назвать предположительной (Konjectural). Именно так выборочные данные называли позднейшие авторы (например, Чубер). Об истории выборочного метода см. Прим. 22.
11. Кабанис и Дестют де Траси основали учение об идеологии, науке о всеобщем и неизменном образовании идей. Известно, что несколько позже его осмеивал Наполеон.
12. Кабанис был профессором гигиены (видимо, общественной).
13. О Варгентине см. Nordenmark (1929), который сообщает, что в 1766 г. численность населения Швеции считалась государственной тайной.
14. Термин статистика появился в XVI в. и означал описание государства. Авторы правы в том, что она была смесью различных дисциплин, многие из которых (например, забытая ими метеорология) впоследствии обособились. Впрочем, эта смесь стала называться государствоведением, об истории которой

см. Шейнин (2014). Гёттингенская школа (см. ниже) фактически была школой государственного управления.

15. Сомнительно, чтобы статистические общества возникли в (разъединённой) Германии в начале XIX в. О Нимане (Niemann), см. чуть выше, известно явно недостаточно. Табличная статистика, которая описывала государства числовыми таблицами, возникла у Анхерсена в 1741 г. (а не в начале века) но презиралась государственными деятелями (Шейнин 2013, § 7.2.1).

16. Сельскохозяйственной революции в Англии не было, но в течение нескольких столетий произошло немало *огораживаний* земель.

17. Неясно, причём здесь немецкая деятельность.

18. Неплохо бы добавить Флоренс Найтингейл, хоть её деятельность началась только в 1853 г. (во время Крымской войны).

19. Вот впечатляющие данные (Chadwick 1842/1965, с. 228): в Ливерпуле лишь 2/3 детей из среднего класса доживало до пяти лет. Farr (примерно 1857; 1885, с. 148) заявил, что *ежегодная смертность населения, превышающая 17 на 1000, неестественна* и вызывается в основном антисанитарными условиями жизни.

20. На самом деле Лондонское статистическое общество возникло из указанной (преобразованной) секции Ассоциации, см. [vii].

21. Утверждение авторов весьма сомнительно: по крайней мере для некоторых указанных тем опубликованных данных недостаточно.

22. Начало выборочных исследований связано с норвежским статистиком Киэром (1894-й год). О нём и о других выборочных исследованиях, и безусловно о Боули, см. You Poh Seng (1951).

23. Позднее начались работы по *Общей статистике Франции*, см. Прим. 26.

24. Вот утверждение А. И. Чупрова (1910, с. 60):

Первое правильное статистическое бюро было учреждено во Франции в 1801 г. [...] Наполеон ревностно поддерживал его, но, сделавшись императором, стал подозрительно относиться к публикации статистических данных. Бюро ослабевало, замерло до 1830-х годов.

25. Следовало бы упомянуть исследование Пуассона (1837), соответствующая часть которого была главным образом основана на этом источнике.

26. Работы по созданию *Общей статистики Франции* велись с 1833 г. под руководством Моро-ди-Жонеса. До 1851 г. было опубликовано 13 томов её трудов.

27. См. Прим. 28.

28. Вывод непонятен, а вместо отрицательной корреляции следовало бы упомянуть обратную зависимость. До появления работ Карла Пирсона термин *корреляция* вряд ли был уместен, вообще же задача Гёрри решается методами не известной в то время ранговой корреляции.

29. Хирург, стало быть, и не считался врачом!

30. Здесь и ниже авторы указывают, что сочинение Виллерме стало известно до его публикации в 1840 г. Объяснение отсутствует.

31. *Падение* Людовика XVIII непонятно: он не был свергнут.

32. Крайне сомнительно, чтобы немецкие университеты целое столетие оставались бы *самыми передовыми*.

33. Кетле не изучал никаких бюджетов, и совершенно неясно, какие его методы мог использовать Энгель.

34. Юнкерами называли крупных землевладельцев.

35. Во многих источниках утверждается обратное: немецкая промышленность успешно конкурировала с английской.

36. Это замечание непонятно. Психофизический подход применялся и в *тейлоризме*. Если отвлечься от усиленной эксплуатации рабочих, которая предусматривалась Тейлором, то указанный подход окажется весьма положительным.

Библиография

Чупров А. И. (1910), *Учёные труды. Курс статистики*. М.

Шейнин О. Б. (2013), *Теория вероятностей. Исторический очерк*. Берлин. Также Google, Oscar Sheynin, Download area

- (2014), К истории государственного управления. *Финансы и бизнес*, № 1, с. 136 – 158. Также Google, Oscar Sheynin, Download area
- Ackerknecht E. H.** (1948), Hygiene in France, 1845 – 1848. *Bull. Hist. Med.*, vol. 22, pp. 117 – 155.
- Addams Jane** (1895), *Hull House Maps and Papers*. New York.
- Agulhon M.** (1958), L'enquête du Comité du Travail de l'Assemblée Constituante (1848) etc. *Annales du Midi*, t. 70, pp. 73 – 85.
- Bailly J. S.** (1790), *Discours et mémoires*, t. 2. Paris.
- Balbi A., Guerry A. M.** (1829), *Statistique comparée de l'état de l'instruction et du nombre des crimes etc.* Paris.
- Beautot E.** (1939 – 1943), Le département du Nord sous la Restauration. *Rev. du Nord*, t. 25, pp. 243 – 277; t. 26, pp. 21 – 45.
- Booth Ch. и др.** (1889 – 1891), *Life and Labour of the People in London*, vols 1 – 17. London, 1902 – 1903.
- Boulainvilliers H. de** (1727), *L'état de la France*, tt. 1 – 2. London, 1752. Написано в 1711.
- Bowley A. L., Burnett-Hurst A. R.** (1915), *Livelihood and Poverty*. London.
- Buret E.** (1840), *De la misère des classes laborieuses en Angleterre et en France*, tt. 1 – 2. Paris.
- Cabanis P. J.** (1790), *Observations sur les hôpitaux*. Paris.
- (1803), Quelques principes et quelques vues sur le secours publics. В книге автора *Œuvr. Phil.*, t. 2, pp. 1 – 63. Paris, 1956.
- Chadwick E.** (1842), *Report on the Sanitary Condition of the Labouring Population of Gt. Britain*. Edinburgh, 1965.
- Chapin F. S.** (1920), *Field Work and Social Research*. New York.
- Chevalier L.** (1948), Préface à Moheau. *Population*, vol. 3, pp. 211 – 232.
- Colquhoun P.** (1814), *Treatise on the Wealth, Power and Resources of the British Empire*. London.
- Condorcet M. J. A. N.** (1785), *Essai sur l'application de l'analyse à la probabilité des décisions etc.* New York, 1972.
- Dainville F. de** (1952), Un dénombrement inédit au XVIII^e siècle etc. *Population*, vol. 7, pp. 49 – 68.
- Deparcieux A.** (1746), *Essai sur les probabilités de la durée de la vie humaine*. Дополнение: 1760. Paris, 2003.
- Du Bois W. E. B.** (1899), *The Philadelphia Negro*. Millwood, N. Y., 1973.
- Dumazedier J.** (1968). Leisure. IESS, vol. 9, pp. 248 – 254.
- Douglas J. D.** (1968), Suicide, social aspects. IESS, vol. 15, pp. 375 – 385.
- Eaton A. H., Harrison S. M.** (1930), *Bibliography of Social Surveys*. New York.
- Eden F. M.** (1797), *The State of the Poor*, vols 1 – 3. London. 1928.
- Esmonin E.** (1954). Quelques données inédites sur Vauban etc. *Population*, vol. 9, pp. 507 – 512.
- (1956), Les mémoires des intendants pour l'instruction du duc de Bourgogne. *Bull. Soc. Hist. Moderne*, t. 55, pp. 12 – 21.
- (1957), L'abbé Expilly et ses travaux de statistique. *Rev. d'Hist. Moderne et Contemporaine*, t. 4, pp. 241 – 280.
- (1958), Montyon, véritable auteur des *Recherches ...* de Moheau. *Population*, vol. 13, pp. 269 – 283.
- Farr W.** (1885), *Vital Statistics*. London.
- FitzPatrick P. J.** (1960), Leading British statisticians of the 19th century. *J. Amer. Stat. Assoc.*, vol. 55, pp. 38 – 70. Перепечатка: Kendall & Plackett (1977, pp. 180 – 212).
- Foucault M.** (1963), *Naissance de la clinique etc.* Paris.
- 8 **Fourastié J.** (1951, франц.), *The Causes of Wealth*. Glencoe, Ill., 1960.
- France** (1819 – 1908), France, Ministère de la Guerre. *Comptes présentés au roi sur le recrutement de l'armée*. До 1837 г. название было несколько иным.
- France** (1827 – 1907), France, Ministère de la Justice. *Compte général de l'administration de la justice criminelle etc.*
- Frégier H. A.** (1840), *Des classes dangereuses de la population etc.*, tt. 1 – 2. Paris.
- Gaulmier J.** (1951), *L'idéologue Volney (1757 – 1820)*. Beirut.
- Gérando J.-M. de** (1801), *Considérations sur les diverses méthodes à suivre dans l'observation des peuples sauvages*. Paris.

- (1820), *Le visiteur du pauvre*. Paris, 1837. Четвёртое издание.
- (1839), *De la bienfaisances publiques*, tt. 1 – 4. Paris.
- Gille B.** (1964), *Les sources statistiques de l'histoire de France*. Publ. Centre de recherches d'histoire et de philologie. *Hautes études médiévales et modernes*, No. 1. Geneva – Paris.
- Göhre P.** (1891), *Drei Monate Fabrikarbeiter und Handwerksbursche*. Leipzig.
- Gossez A. M.** (1904), *Le département du Nord sous la Diuixième République (1848 – 1852) etc.* Lille.
- Greenwood M.** (1941 – 1943), Medical statistics from Graunt to Farr. *Biometrika*, vol. 32, pp. 101 – 107, 203 – 225; vol. 33, pp. 1 – 24. Перепечатка: Pearson E. S., Kendall M. G., редакторы (1970), *Studies in the History of Statistics and Probability*, vol. 1. London, pp. 47 – 120.
- Guérard A.** (1864), Notice sur M. Villermé. *Annales d'hygiène publique et de médecine légale*, t. 21, pp. 162 – 177.
- Guerry A. M.** (1833), *Essai sur la statistique morale de la France*. Paris.
- Guillaume P.** (1962), Département de la Loire. *Actes 86^e Congrès Soc. Savants de Paris et de Départements*. Paris, pp. 429 – 450.
- Howard J.** (1777), *State of the Prisons*. New York, 1929.
- (1789), *Account of the Principal Lazarettos in Europe*. London, 1791.
- Kahan-Rabecq Marie M.** (1934 – 1935), Les réponses havraises à l'enquête ... en 1848. *La Révolution de 1848*, t. 31, pp. 95 – 113.
- (1939), L'Alsace économique et sociale etc. *Réponses du département du Haut-Rhin etc.*, t. 2. Paris. Приведенное библиографическое описание (без указания с страниц) было неясно, дано нами по смыслу.
- Kay-Shuttleworth J. P.** (1832), *Moral and Physical Condition of the Working Classes etc.* London.
- Kendall M. G., Plackett R. L., редакторы** (1977), *Studies in the History of Statistics and Probability*, vol. 2. London.
- Lavoisier A. L.** (1791), *De la richesse territoriale du royaume de France*. Paris, 1819.
- Levenstein A.** (1912), *Die Arbeiterfrage etc.* München.
- Le Play F.** (1855), *Ouvriers européens*.
- McCulloch J. R.** (1837), *A Descriptive and Statistical Account of the British Empire*, vols 1 – 2. London, 1854. Четвёртое издание.
- Messance** (1766), *Recherches sur la population des généralités d'Auvergne* [и других округов и городов]. Paris.
- Moheau M.** (1778), *Recherches et considérations sur la population de la France*. Paris, 1912.
- Montalivet J.-P.** (1813), *Exposé de la situation de l'Empire etc.* Paris.
- Nordenmark N. W. E.** (1929), Pehr Wilhelm Wargentin, 1717 – 1783. *Nordic Stat. J.*, vol. 1, pp. 241 – 252.
- Ogburn W. F.** (1918 – 1919), Analysis of the standard of living in the District of Columbia in 1916. *J. Amer. Stat. Assoc.*, vol. 16, pp. 374 – 389.
- Parent-Duchâtelet A. J. B.** (1834, франц.), *On Prostitution in the City of Paris*, vols 1 – 2. London, 1857. Третье издание.
- (1836), *Hygiène publique etc.*, tt. 1 – 2. Paris.
- Pigeire J.** (1932), *La vie et l'oeuvre de Chaptal, 1756 – 1832*. Paris.
- Pitts J. R.** (1968), Le Play, Frédéric. *IESS*, vol. 9, pp. 84 – 91.
- Playfair W.** (1801), *The Statistical Breviary*. London.
- Poisson S.-D.** (1837), *Recherches sur la probabilité des jugements etc.* Paris, 2003. Перевод: Google, Oscar Sheynin, Download area.
- Recherches** (1821 – 1860), *Recherches statistiques sur la ville de Paris et le département de la Seine*, tt. 1 – 6. Paris. Редактор первых четырёх томов Ж. Б. Ж. Фурье.
- Reinhard M.** (1961), *Etude de la population pendant la Révolution et l'Empire*. Gap, France.
- Rigaudias Hilde (Weiss)** (1936), *Les enquêtes ouvrières en France entre 1830 et 1848*. Paris.
- Rosen G.** (1946), The philosophy of ideology and the emergence of modern medicine in France. *Bull. Hist. Med.*, vol. 20, pp. 328 – 339.
- (1955), Problems in the application of statistical analysis to questions of health, 1700 – 1800. Там же, vol. 29, pp. 27 – 45.

- (1958), *History of Public Health*. New York.
- Rowntree B. S.** (1901), *Poverty*. London – New York, 1922.
- Schnapper-Arndt G.** (1883), *Fünf Dorfgemeinden etc.* Leipzig.
- (1888), *Zur Methodologie sozialer Enquêtes*. Frankfurt/Main.
- Sinclair J.** (1791 – 1799), *Statistical Account of Scotland*, vols. 1 – 21. Edinburgh.
- Survey** (1909 – 1914), *The Pittsburgh Survey*, vols 1 – 6. Pittsburg.
- Survey** (1918 – 1920), *The Springfield Survey*, vols 1 – 3. New York.
- Tenon J. R.** (1788), *Sur les hôpitaux de Paris*. Paris.
- Thomas W. I., Znaniecki F.** (1918 – 1920), *The Polish Peasant in Europe and America*, vols 1 – 2. New York, 1958, 1971.
- Todhunter I.** (1865), *History of the Mathematical Theory of Probability*. New York, 1949, 1965.
- Vauban S. Le Prestre de** (1707), *Projet d'une dime royale*. Paris, 1943. В книгу вошли Méthode générale ... и Description de l'élection de Vezelay.
- Vedrenne-Villeneuve E.** (1961), L'inégalité sociale devant la mort dans la première moitié du XIX^e siècle. *Population*, vol. 16, pp. 665 – 699.
- Vidalenc J.** (1948), Les résultats de l'enquête sur le travail etc. *Actes Congrès Historique du Centenaire de la Révolution de 1848*. Paris, pp. 325 – 341.
- Villeneuve-Bargemont A.** (1837), Etudes spéciales sur le département du Nord. В книге автора *Economie politique chrétienne etc.* Bruxelles, pp. 216 – 227.
- Villermé L. R.** (1820), *Des prisons*. Paris.
- (1840), *Tableau de l'état physique et moral des ouvriers*, tt. 1 – 2. Paris.
- Vincienne Monique, Courtois Hélène** (1958), Note sur la situation religieuse en France en 1848 etc. *Arch. de sociologie des religions*, t. 6, pp. 104 – 118.
- Volney C.** (1787), *Voyage en Egypte et en Syrie*. Paris, 1959.
- (1795), *Questions de statistique à l'usage des voyageurs*. Paris.
- Weber M.** (1892), *Die Verhältnisse der Landarbeiter in ostelbischen Deutschland*. Schriften Verein f. Sozialpolitik, B d. 55.
- (1909), Zur Methodik sozialpsychologischer Enquêtes etc. *Arch. f. Sozialwiss. u. Sozialpolitik*, Bd. 29, pp. 949 – 958.
- (1919 нем.), Science as a vocation. В книге автора *Essays in Sociology*. New York, 1946, pp. 129 – 156.
- Westergaard H.** (1932), *Contributions to the History of Statistics*. The Hague, 1969.
- Woods R. A., редактор** (1898), *The City Wilderness*. Boston.
- (1902), *Americans in Process. A Settlement Study*. Boston., 1969.
- You Poh Seng** (1951), Historical survey of the development of sampling theories and practice. *J. Roy. Stat. Soc.*, vol. A114, pp. 214 – 231. Перепечатка: Kendall & Plackett (1977, pp. 440 – 457).
- Young A.** (1771a), *The Farmer's Tour through the East of England*, vols 1 – 4. London.
- (1771b), *Six Months Tour through the North of England*, vols 1 – 4. London.
- (1780), *Tour in Ireland*, vols 1 – 2. Dublin.
- (1793), *Travels in France etc.* Cambridge, 1950.
- Yule G. Udny** (1899), Investigation into the causes of changes in pauperism etc., pt. 1. *J. Roy. Stat. Soc.*, vol. 62, pp. 249 – 286.

XVI

Эдгар Боргатта

Социометрия

Edgar F. Borgatta, Sociometry. IES, pp. 1031 – 1036

[1. Общие сведения]

Термин *социометрия* неоднозначен, но исторически он ближе всего связан с работой Морено и особенно с его исследованием (1934) отношения между людьми. По традиции социометрия отождествляется с анализом данных, собранных при *социометрическом эксперименте*, при котором, грубо говоря, каждого члена некоторой группы просят указать, с какими другими членами он хотел бы скорее всего заниматься чем-нибудь. Этот эксперимент разработал Морено с помощниками, которые прекрасно применили его в своих собственных исследованиях.

В начале 1940-х годов появилось несколько высказываний о том, что следует понимать под *социометрией*, и сам Морено (1943) заявил, что смысл этого термина не должен быть ограничен его собственными разработками. И он, и другие полагали, что социометрия должна в общем относиться к измерению социальных явлений, однако на самом деле его арсенал всё ещё является почти всеобщим принятым при обращении к этому термину.

[1.1] Сбор данных. Форма социометрического вопроса и обстановка, в которой собираются данные, в высшей степени разнообразны. Вопрос может указать испытуемому обстановку или возможность выбора. В школьной аудитории подходяще будет сформулировать вопрос соответствующим образом: *С какими учащимися вы хотели бы обсудить эту проблему?* В других случаях испытуемые могут назвать учителя, друзей или родственников, находящихся вне аудитории, или даже экспертов, с которыми они лично не знакомы. Предусматриваемая деятельность (в данном случае, обсуждение в аудитории) должна быть ясно определена, чтобы испытуемые знали цели выборов.

Эту процедуру можно противопоставить *оцениванию*, при котором испытуемые характеризуют других, но не обязаны решать, хотят ли они делать что-то совместно с кем-либо из них.

Если, например, социометрический метод применяет учитель, выбор критериев может определяться практическими целями; он может желать перегруппировать учащихся, предполагая, что они будут лучше работать, если сами выберут себе товарищей. С другой стороны, учитель может преследовать противоположную цель, чтобы добиться большего взаимодействия между теми, кто обычно избегал бы друг друга.

В начальном периоде развития социометрии экспериментаторы, возможно ввиду озабоченности Морено и его помощников, часто чувствовали себя обязанными перегруппировывать испытуемых

вне зависимости от последствий. Впрочем, при более общем использовании этого социометрии подобные чувства исчезли. По существу социометрические вопросы часто формулируются в условной форме, например: *Если ваша группа снова соберётся, с кем из её членов вы больше всего захотели бы участвовать [в чём-то]?* Нет никаких указаний на то, что подобная условная форма вопроса менее полезна, чем та, которая подразумевает реальную перестановку группы, но существует продолжающееся [негласное] предписание делового формулирования вопроса.

Как и в каждом методе исследования, при использовании социометрических вопросов необходимо обращать внимание на общие способности испытуемых. Вот очевидный пример. Пусть требуется опросить детей, которые ещё не умеют ни читать, ни писать. В таком случае социальная обстановка должна обеспечить уединение и доверительность опроса, а опрашивающий обязан быть уверенным в том, что ребёнок не запуган, не боится обстановки, иначе же ответы могут оказаться менее обоснованными.

По существу сбор данных посредством социометрических вопросов несложен и возможен в большинстве случаев, либо когда просят перечислить различные варианты ответов на один-единственный вопрос, либо (вторая крайность) когда желательны ответы на многие вопросы, каждый из которых следует расположить в порядке предпочтений и отказов.

Построение вопросников. Форма сбора данных определяет тип возможных результатов. К примеру, количество запрашиваемых выборов может быть либо ограничено, либо нет. Во втором случае можно сравнивать полное количество выборов, указанных различными группами, составленными из одного и того же числа испытуемых, или несколько раз приведенных одной и той же группой.

В первом случае, однако, подобные сравнения не имеют смысла. Но для данной группы полное число выборов воспринимается сравнительно просто и связано с её *согласованностью и уверенностью в своих действиях*. Goodacre (1951) выяснил, что высокая доля выборов (high rate of choosing) связана с эффективностью действия группы¹. В соответствии с простейшим истолкованием представляется, что если члены группы полагают друг друга подходящими для какого-либо её действия, то она вероятно окажется успешной.

На первый взгляд этот вывод не представляется глубоким, но пока он по существу является единственной надёжной связью с эффективностью группы, если не считать предсказания того, что успешные группы останутся таковыми же и что группы, составленные из лиц с установленными высокими способностями, будут успешными.

Если требуется установить *сети* социальных взаимоотношений, необходимо допустить неограниченное количество выборов. Очевидно, что тесную группу из 10 человек не удастся надёжно выявить в более многочисленной группе, если каждому её члену разрешено лишь три выбора. В начале своей деятельности Морено

занимался подобным анализом структуры групп, и поэтому традиционно рекомендуется допущение неограниченного числа выборов. Тем не менее, были предложены многие иные методы, например допущение определённого числа выборов или выборов в порядке предпочтений; предпочтительное назначение ранга всей группе; парные сравнения внутри группы; оценка периода времени, который член группы хотел бы провести с другими; угадывание особых признаков или репутации других членов; ранжирование каждого члена группы по какому-либо признаку. Допущение отрицательных выборов (отклонений) и применение многих вопросов конечно же увеличивает число возможных выборов.

[2] Социометрическое описание

Схема полученных выборов может быть представлена графически в виде *социограммы*, в которой каждый член группы указывается какой-либо геометрической фигурой (например, кружочком с фамилией внутри него) и соединительными прямыми или стрелками, поясняющими направление выборов. В первых исследованиях применение социограмм носило случайный характер, однако Морено и другие сумели с их помощью вывести некоторое число эмпирически обоснованных и теоретически важных понятий.

Простейшим является понятие *неизбранного*, которого можно считать социально изолированным другими. В прежней традиции *чужеродным* называлось лицо, которое не выбирает никого, и которого никто не выбирает, так что в этом смысле он полностью отделён от группы. Впрочем, при обычном словоупотреблении он не отличается от избранного.

В литературе встречается и термин *недоизбранный*, но он, пожалуй, менее желателен, ибо подразумевает существование какого-то уровня избранности. Разумеется, если вопрос об отклонениях был задан (с кем вы *меньше всего* хотели бы заниматься этим?) наравне с вопросом о выборах, может быть выделен и *отклоняемый*. Отклонение одного члена группы другим подразумевает активную неприязнь, тогда как равнодушие или отказ в выборе может означать просто недостающее общение.

Если при более многочисленных группах применяются и положительные, и отрицательные формы вопросов, то не выбранные лица при положительном вопросе обычно отклоняются при его отрицательной форме и наоборот. Считалось, что лицо, выбранное многократно, находится в желательном положении; в ранней социометрической литературе обычно встречался термин *чрезмерно выбранный*, но теперь он применяется редко. В некоторой степени привлекательным было понятие *социометрической звезды*, – как бы многократно избранного лица, окружённого менее избранными.

Аналогичным является понятие *популярный руководитель*, и многократно избранное лицо наиболее часто связывается с каким-либо руководством. Однако, популярность (многократное избрание) и руководство не совпадают, и отличие между ними

было ясно указано (Criswell & Jennings 1951). Популярное лицо может быть руководителем, но им могут быть и другие лица, обладающие властью, а *властное лицо*, т. е. тот, кого выбрали находящиеся в ведущем положении, может вовсе и не быть популярным. Но самым отчётливым образом отличие между популярностью и руководством выявляется при рассмотрении содержания вопросов. Jennings (1947) обсуждала различие между *социальными* и *психологическими* группами вопросов. Если спрашивают о поставленной задаче, то вопрос считается социальным, и многократно избранное лицо обычно является руководителем. Если же вопрос ориентирован социально (относится к психологической группе), то это лицо обычно популярно или привлекательно².

Упомянутые понятия, хоть и относятся к структуре взаимоотношений, указывают конкретного лица или типа лиц. Но даже в самых ранних социометрических исследованиях серьёзнейшее внимание уделялось сети взаимоотношений. Действительно, отношения между отдельными лицами было легко обозначать; например, очевидны были понятия *взаимная пара* или *взаимное отклонение*. Описание усложняется в так называемом *анализе отношений*, который обычно включает и подсчёт всех возможных выборов и отклонений, и некоторую попытку предсказания будущего распределения тех и других, см., например, Tagiuri (1952).

Изучение ещё более сложных расположений отношений привело к появлению более сложных терминов. Геометрические названия (например, треугольные, четырёхугольные), как оказалось, применимы лишь в ограниченной степени. Теоретически важнее были менее чёткие *цепи* и *кольца*, которые используются при анализе структуры клик³.

[2.1] Социограммы: проблемы и предложения. На ранней стадии развития социометрических методов социограммы были вездесущи. Тем не менее, хоть с ними и были связаны важные понятия, они часто применялись для иллюстрации, а не для анализа. Если лицо, составляющее диаграмму, неявно предполагает что-то, то его чертёж может оказаться выразительным⁴. К примеру, отсутствие выборов между мальчиками и девочками в группах из школьников младших классов будет впечатлять, если символы тех и других расположить на разных концах диаграммы. Если расовые или этнические подгруппы расположены в разных местах диаграммы, то малое число соединительных линий между ними может быть легко заметно в отличие от многих линий *внутри* этих подгрупп. С другой стороны, анализ структуры клик или местоположения сложных сетей выборов при помощи социограмм может оказаться почти невозможным.

Было предложено много других видов социограмм. Так, (Northway 1940) социограмма может подчёркивать *лучшее положение*, определяемое выборами, концентрическими окружностями, показывая наиболее многократно выбранное лицо во внутреннем круге, а отношения между членами группы –

обычным образом, стрелками. Ни этот, ни другие варианты подчёркивания подобного положения не применялись часто. Другой аналогичный метод (Powell 1951) использует символы различных размеров, так что крупные символы соответствуют, например, большому числу выборов. Применяются также расстояния между точками, так что взаимные пары располагаются очень близко, а отклоняющие друг друга – очень далеко (Proctor & Loomis 1951). Существуют и ещё более сложные описательные многомерные диаграммы (Charin 1950), однако повторим, что польза от социограмм для целей анализа, видимо, невелика. Тем не менее, они полезны для описаний. Их стараются упростить, обычно уменьшая число пересекающихся линий (Borgatta 1951).

[3] Социометрический анализ

В отличие от методов описания, аналитические исследования подчеркнули разработку значимых показателей выбора и необходимость систематического анализа матрицы полного набора выборов. Показатели обычно подбираются для применения отобранных понятий, и следует представлять себе, что даже такие произвольные определения как *невывбранный* или *многократно выбранный* уже являются показателями простого понятия социометрического выбора⁵.

Обычно, впрочем, показатель нацелен на достижение сравнимости различных наборов данных. Например, он может учитывать число выбирающих лиц, так что группы, различные по численности, окажутся сравнимыми. Тем не менее, при разработке показателей возникает немало проблем, и литература пестрит предупреждениями о том, что попытка *учесть что-то* в показателе может не только оказаться неудачной, но и вызвать ещё более серьёзную проблему, чем те, решение которых старались облегчить.

[3.1] Матричные методы. Начиная со статьи Moreno & Jennings (1938) большое внимание стало уделяться статистической значимости результатов. Прежние подходы к этому вопросу, см., например, Bronfenbrenner (1943), теперь в общем считаются непрактичными, но их последовавшее обсуждение заставило сосредоточиться на *моделях*, на которых основаны выборы. Чтобы облегчить статистическую обработку социометрических данных, Forsyth & Katz (1946) предложили заменить обременительные социограммы квадратными матрицами размером $N \cdot N$ (N – число лиц в группе). Выбор или отклонение здесь могут быть ясно указаны в соответствующей клетке матрицы знаками плюс и минус. Так, знак *плюс* в 10-м столбце 5-й строки матрицы будет означать, что лицо № 5 выбрало 10-го. Эти авторы посчитали особо важными выборы, располагающиеся вблизи главной диагонали матрицы (эта диагональ, конечно же, означает выбор самого себя) и обратили внимание на скопления взаимных выборов и на примыкавшие скопления с некоторыми совпадающими номерами⁶. Утонченное видоизменение этого метода (Beum & Brundage 1950) было впоследствии обобщено для эффективного компьютерного анализа (Borgatta & Stolz 1963).

Описанные только что методы существенно зависят от перераспределения данных, что уже предусматривалось в матрице размером $N \cdot N$. Однако, умножение матриц (Festinger 1949) подчеркнуло отождествление более формально определяемых структур⁷. Дальнейшие разработки в этом направлении были главным образом направлены на именование и выявление всё более сложных соотношений, см., например, Luce (1950), Katz (1953), Harrary & Ross (1957). Произошёл определённый продвиг от прежних социометрических исследований, которые в основном изучали схемы взаимного выбора.

В другом подходе к исследованию матриц выборов применёна теория графов (Harrary & Norman 1953, Ramanujacharyulu 1964) и факторный анализ (McRae 1960). Он близок указанному методу перераспределения, так что эти два подхода дополняют друг друга. Есть также и метод, основанный на благоприятно встреченном (Ragsdale 1965) кластерном анализе (Bock & Husain 1950).

Интерес к разработке этих аналитических методов достиг высшей точки в начале 1950-х годов. С тех пор количество публикаций уменьшилось, но вероятно удержится на постоянном уровне. Тем не менее, приложений этих методов было немного, возможно потому, что при исследовании взрослых испытуемых они требуют нелегко добываемых видов данных.

[3.2] Надёжность и обоснованность. Некоторые исследования надёжности и обоснованности социометрических методов были проведены, но в последние годы этой стороне вопроса уделялось мало внимания. В одном случае (Mouton и др. 1955) были отмечены ограничения социометрических методов с точки зрения устойчивости измерений. Помимо других обстоятельств, устойчивость измерительного устройства⁸ переплетается с устойчивостью человека и социальных структур. Особенно трудно оценить обоснованность, поскольку социометрические показатели столь часто понимаются как нечто предсказуемое. По существу социометрическая информация представляет собой объективное описание обстановки самыми подходящими судьями, – самими испытуемыми. Поэтому имела место склонность подчёркивания предсказания социометрического состояния на основе других характеристик, а не применения его для предсказания других переменных.

[4] Приложение социометрии

Социометрические методы были включены во многие исследования различного типа. Так, при изучении небольших групп одним из обычных видов информации, добываемой при последующем опросе, является набор ранжировок по поводу действий группы. Здесь следует подчеркнуть, что социометрические методы в классическом понимании обычно сливались с более общими методами ранжирования или оценивания какой-либо деятельности специалистами равной квалификации. Многие исследователи систематически изучали структуру ранжирования и само-ранжирования и произошло

некоторое согласование мнений и сведений об устойчивости измерений (Borgatta 1964). Главной темой исследований оставалась та, которую первоначально установила Jennings (1947) (§ 2), однако в них также повторно появлялись другие понятия.

Социометрические методы были важны для разработки некоторых иных областей исследования, к примеру, для изучения влияния структуры группы на характеристики её членов или на следствия из её действия (её эффективность или моральное состояние). Оказалось необходимой более формальная разработка понятий для сетей связи. В обширном обзоре соответствующих исследований (Glanzer & Glaser 1961) они были рассмотрены в историческом контексте и навели на мысль об ограничениях подобных подходов.

В социологии социометрические методы остаются всеобщее распространёнными, поскольку они родственны исследованиям личности, малых групп, анализу сетей сообщений и структур групп, а также и специальным темам, как например репутациям социального положения в некоторой общине и изучению расовой или иной изолированности.

Примечания

1. Суть вывода указанного автора непонятна, дальнейшее пояснение поэтому неясно.
2. Непонятное разъяснение.
3. Клика (толковый словарь английского языка): небольшая группа общества, члены которой имеют общие интересы и стремятся отгородиться от лиц вне группы.
4. Можно ничего не предполагать заранее, но представить себе по первому взгляду.
5. Почему же, например, термин *невывбранный* произволен?
6. Объяснение отсутствует.
7. Объяснение отсутствует.
8. Какие измерения, какое измерительное устройство?

Библиография

- Beum C. O., Brundage E. G.** (1950), A method for analyzing the sociomatrix. *Sociometry*, vol. 13, pp. 141 – 145.
- Bock R. D., Husain S. Z.** (1950), Adaptation of Holzinger's B-coefficients for the analysis of sociometric data. *Sociometry*, vol. 13, pp. 146 – 153.
- Borgatta E. F.** (1951), A diagnostic note on the construction of sociograms etc. *Group Psychotherapy*, vol. 3, pp. 300 – 308.
- (1964), The structure of personality characteristics. *Behavioral Science*, vol. 9, pp. 8 – 17.
- Borgatta E. F., Stolz W.** (1963), Note on a computer program for rearrangement of matrices. *Sociometry*, vol. 26, pp. 391 – 392.
- Bronfenbrenner U.** (1943), A constant frame of reference for sociometric research. *Sociometry*, vol. 6, pp. 363 – 397.
- Chapin F. P.** (1950), Sociometric stars as isolates. *Amer. J. Sociology*, vol. 56, pp. 263 – 267.
- Criswell Joan H., Jennings Helen H.** (1951), A critique of Chapin (1950). *Amer. J. Sociology*, vol. 57, pp. 260 – 264.
- Festinger L.** (1949), Analysis of sociograms using matrix algebra. *Human Relations*, vol. 2, pp. 153 – 158.
- Forsyth Elaine, Katz L.** (1946), A matrix approach to the analysis of sociometric data. Preliminary report. *Sociometry*, vol. 9, pp. 340 – 347.
- Glanzer M., Glaser R.** (1961), Techniques for the study of group structure and behavior, pt. 2. *Psychological Bull.*, vol. 58, pp. 1 – 27.

- Goodacre D. M.** (1951), The use of a sociometric test as a predictor of combat unit effectiveness. *Sociometry*, vol. 14, pp. 148 – 152.
- Harrary F., Norman R. Z.** (1953), *Graph Theory as a Mathematical Model in Social Science*. Publ. No. 2, Res. Center for Group Dynamics. Ann Arbor.
- Harrary F., Ross I. C.** (1957), A procedure for clique detection using the group matrix. *Sociometry*, vol. 20, pp. 205 – 215.
- Jennings Helen H.** (1947), Sociometric differentiation of the psychegroup and the sociogroup. *Sociometry*, vol. 10, pp. 71 – 79.
- Katz L.** (1953), A new status index etc. *Psychometrika*, vol. 18, pp. 39 – 43.
- Luce R. D.** (1950), Connectivity and generalized cliques etc. *Psychometrika*, vol. 15, pp. 169 – 190.
- McRae D. Jr** (1960), Direct factor analysis of sociometric data. *Sociometry*, vol. 23, pp. 360 – 371.
- Moreno J. L.** (1934), *Who Shall Survive?* etc. Beacon, N. Y.
- (1943), Sociometry and the cultural order. *Sociometry*, vol. 6, pp. 299 – 344.
- Moreno J. L., Jennings Helene H.** (1938), Statistics of social configurations. *Sociometry*, vol. 1, pp. 342 – 374.
- Mouton Jane S., Blake R. R., Fruchter B.** (1955), Reliability of sociometric measures. *Sociometry*, vol. 18, pp. 7 – 48.
- Northway Mary L.** (1940), A method for depicting social relationships etc. *Sociometry*, vol. 3, pp. 144 – 150.
- Powell R. M.** (1951), A comparative social class analysis of [двух населённых пунктов]. *Sociometry*, vol. 14, pp. 182 – 202.
- Proctor Ch. H., Loomis Ch. P.** (1951), Analysis of sociometric data, pt. 2. Jahoda Marie, Deutsch M., Cook S. W., редакторы? *Research Methods in Social Relations etc.* New York, pp. 561 – 585.
- Ragsdale R. G.** (1965), *Evaluation of Sociometric Measures etc.* Univ. of Wisconsin. Диссертация.
- Ramanujacharyulu C.** (1964), Analysis of preferential experiments. *Psychometrika*, vol. 29, pp. 257 – 261.
- Tagiuri R.** (1952), Relational analysis. *Sociometry*, vol. 15, pp. 91 – 104.

XVII

Уильям Краскл

Статистика, её сфера

William H. Kruskal, Statistics, the field. IES, pp. 1071 – 1093

[I. Предисловие]

[Статистика.] Исследователь, собрав эмпирический материал, должен вывести из него какие-то заключения, сделать какой-то вывод, принять какое-то решение. Окончанием этого процесса может быть подтверждение или опровержение какой-либо сложной теории, решение по поводу следующего эксперимента или просто уточнение предположенного интервала значений некоторой константы природы. Возможно даже заключение о бесполезности проведенных наблюдений. В типичном случае исследователь также формулирует утверждение или по крайней мере указывает на ощущение о том, насколько он уверен в своих заключениях.

Ничего подобного, конечно же, не происходит только под влиянием новых наблюдений. Всегда существует какое-то знание фона, есть и интуиция, частично явная, и частично подсудная. Суть науки и состоит в том, что перескок на ошибочное положение либо ввиду неудовлетворительных данных наблюдения, либо вводящего в заблуждение фона, либо, наконец, ввиду неудачного скачка раньше или позже исправляют дальнейшие исследования.

Скачки часто делаются без должного анализа своих действий или самого процесса перехода к заключениям. Так на простом подъёме опытный альпинист может передвигаться от валуна к валуну, но если склон крут, а ухватываться за что-то рукой удаётся лишь изредка, то перед движением следует обдумать его направление, понять, куда можно будет поставить ноги и сообразить, что произойдёт при падении.

Статистика изучает процесс перехода к заключениям, в частности – планирование и анализ экспериментов или исследований, суть ошибок наблюдений и причины изменчивости, которые затемняют лежащие в основе явления, а также действенное сведение воедино наборов данных. Существует расплывчатая граница между статистикой и другими частями философии науки, см. ниже.

Заключения из эмпирических данных приходится делать не только в науке, но и в обыденной жизни и различных областях общественной политики. Так, в 1954 г. составление плана и анализ полиомиелитной вакцины Солка были основаны на статистических понятиях рандомизации и контроля.

Экономические решения частных лиц и общественных институтов иногда зависят от значения и точности суммарных чисел, выведенных из сложных измерений (уровень безработицы, экономический рост, индекс потребительских цен).

Недостаточность статистического обоснования иногда приводит к

неверному истолкованию статистики несчастных случаев или преступности, и подобные ошибки могут происходить и в военной разведке и в дипломатии.

Между статистиками и другими учёными происходит оживлённое двустороннее интеллектуальное общение. Психологи и физические антропологи вызвали появление и сильно повлияли на развитие той ветви статистики, которая называется многомерным анализом; социологи иногда бранят статистиков за недостаточное внимание к проблемам формулирования выводов из исследований населения, а некоторые экономисты являются и потребителями, и производителями статистических методов.

[2.] Теоретическая и прикладная статистика. Теоретическая статистика это формальное изучение процесса, ведущего от наблюдений к выводам, решениям и т. д., если только подобные процессы можно отделить от конкретного эмпирического содержания¹. Указанное изучение не является психологическим исследованием того, как фактически учёные формулируют выводы или принимают решения. Скорее оно имеет дело с последствиями определённого вида выводов или решений и пытается нормативно отыскивать надёжные пути на основе явных критериев.

Теоретическая статистика должна применять более или менее формальный язык, обычно математический, и в каждой конкретной области принимать слабые или существенные предположения, на которых и основывать формальный анализ. Самый архиважнейший математический язык в статистике это язык вероятностей, поскольку в основном статистические рассуждения основываются на случайности, совокупностях, массах и одно-единственное событие считается элементом большого класса событий. Даже с таких позиций, подобных личной [субъективной] вероятности, для которой одиночное событие является основным понятием, существенно используется вероятностный язык.

Понятия и арсенал математики имеют основное значение для большей части статистики, которая, строго говоря, не является ветвью математики. Некоторые важные области теоретической статистики можно обсуждать и совершенствовать без применения малоизвестной математики, и много примечательного в статистике осуществили люди со скромной математической подготовкой.

Прикладная статистика, по крайней мере в принципе, это квалифицированное приложение теоретически исследованных методов, т. е. фактически скачок, совершаемый после изучения его теории. Но на самом деле всё не так просто. Во-первых, теоретическому исследованию статистического метода часто предшествует его интуитивный выбор и использование. Во-вторых, можно почти бесконечно теоретически изучать даже простейшие методы. Практика взаимодействует с теорией и сплетается с ней воедино, так что многие статистики сегодня (или сию минуту) практики, а завтра (или через час) – теоретики.

Искусство прикладной статистики требует ощущения того, как именно теоретические предположения могут нарушаться и возможных последствий этого, равно как и способности быстро видоизменять и обобщать изученные ранее методы. Так, при изучении общественного мнения прикладная статистика имеет дело с планированием и анализом соответствующих исследований; основная ветвь теоретической статистики, которая применяется в таких случаях, относится к выборочным исследованиям, хотя и другие разделы теории тоже могут оказаться полезными (например, теория цепей Маркова в панельных исследованиях, когда требуется устанавливать мнение *одних и тех же* респондентов в последовательные периоды времени).

Опять же, при изучении процесса обучения прикладная статистика включает тщательное планирование и анализ регулируемых лабораторных испытаний с червями, крысами или людьми. Здесь могут применяться статистические теории планирования эксперимента, дисперсионного анализа или конечного числа исходов. И, конечно же, для доброкачественной прикладной статистики важно существенное знание эмпирической стороны темы (общественного мнения, обучения и т. д.).

Статистика молода, и количество её тщательно изученных методов, хоть и растёт с часу на час, всё ещё сравнительно невелико. В её приложениях поэтому обычно приходишь до точки, уравнивающей формальные размышления о конкретной задаче (что редко бывает достаточным, потому что лишь немногие задачи стандартны) и применение методов, понятых в меньшей степени, чем хотелось бы. Стимулирующее подробное обсуждение этого вопроса см. у Tukey (1962); *исследование данных* означает у этого автора что-то подобное прикладной статистике.

Иногда слово *статистика* не означает общего подхода в том смысле, в котором оно схематично пояснено выше, но более узко относится к набору конкретных статистических методов с сопроводительными формулами, таблицами и традициями в нынешнем понимании и применении. Есть и иные обычные значения этого термина, но вряд ли они приведут к путанице. В частности, статистикой часто называют набор чисел, описывающих какое-то эмпирическое явление, например, статистику смертности во Франции в 1966 г. Наконец, статистика часто означает какое-то число, вычисленное по основным наблюдениям.

[3.] Изменчивость и погрешность. Будь жизнь устойчива и проста, состояла бы она из закономерных повторений, от статистического мышления было бы мало пользы. Но тогда вообще некому было бы заниматься им, потому что достаточная устойчивость и простота не допустила бы генетической случайности, этого центрального механизма эволюции. На самом же деле жизнь не устойчива и не проста, хотя некоторые её стороны именно таковы. Можно поэтому сказать, что цель науки

состоит в отыскании и разъяснении этих сторон, и что статистика имеет дело с некоторыми общими методами отыскания схем естественной изменчивости, скрытых в облаке не относящегося к делу и в наблюдениях или измерениях, подверженных ошибкам.

Наибольшая часть статистических рассуждений имеет дело с изменчивостью и погрешностями в наблюдаемых данных, чтобы выводить заключения про скрытые закономерности. Но что означают естественная изменчивость и ошибки измерения? Во-первых, чёткие элементы экспериментов и наблюдений обычно обладают различными характеристиками и ведут себя по-разному. Отдельные лица отличаются друг от друга своими способностями и умением, а некоторые мыши обучаются быстрее других.

Во-вторых, при измерении некоторой величины или качества обычно происходит погрешность, приводящая ко второму виду дисперсии, с которым имеет дело статистика. Не только студенты, обучаемые по новому методу, реагируют на этот процесс по-разному, но и тест, который определяет их успехи, не может быть совершенным прибором. Независимые подсчёты числа кровяных шариков на одном и том же предметном стекле микроскопа двумя наблюдателями обычно не совпадают.

В каждом конкретном эксперименте или изыскании некоторые источники изменчивости целесообразно полагать постоянными². Например, студенты в эксперименте обучения могут быть отобраны из одной и той же географической области. Другие источники изменчивости можно считать случайными; примером может служить колебание оценок студентов из, казалось бы, однородной группы. Часто имеют место более сложные промежуточные формы изменчивости.

Студенты могут быть разбиты на группы и обучаться у разных преподавателей. Поскольку имеет значение принадлежность к группе с одним и тем же преподавателем, появляется простая, но существенная схема зависимости. Идея изменчивости отражается в основном понятии *совокупности*, из которой отбираются выборки. Совокупность может быть группой людей, мышей или машин или воображаемой группой, как совокупность ошибок наблюдений.

Грубо говоря, совокупность численных значений образует распределение, и основной здесь оказывается идея о *случайной переменной*³, значения которой изменяются в соответствии с этим распределением. Так, если случайно отобранного из школьного коллектива студента проверяют на чтение и понимание прочитанного, то его предположительная оценка, не зависящая от его личности и предлагаемого текста, является случайной переменной. Её распределением будет идеально осуществляемое множество всех таких оценок, если этот эксперимент повторен очень много раз без всяких изменений во времени или ввиду общения студентов друг с другом.

Многое в методологии статистики может представляться попыткой понять закономерность в облаке затемняющей изменчивости, но существует много случаев, в которых весьма

интересна сама изменчивость; некоторые из них рассмотрены ниже.

[4.] Планирование. Важной темой в статистике является разумное планирование эмпирических исследований. В примере, приведенном выше, можно выявить более формальные стороны планирования:

Сколько групп следует обучать каждым данным методом?

Сколько студентов из группы проверять?

Учитывать ли другие переменные помимо результатов проверки (например, уровень интеллекта или ранжирование студентов)?

Можно задать и другие вопросы:

Как обучать преподавателей новым методам?

Отбирать ли преподавателей так, чтобы некоторые относились к новым методам с воодушевлением, а другие – скептически?

Каким тестом измерять результат?

Не существует никакой общей теории планирования, которая учла бы все или по крайней мере большинство подобных вопросов, но имеются многие разделы такой теории; что важнее, есть и полезная статистическая точка зрения по поводу планирования экспериментов.

[5.] История. Мы не рассматриваем историю развития статистики, но важно отделить историю самого слова *статистика* и историю статистики в смысле этой статьи. Слово это родственно слову *state* (государство), и первоначально статистической деятельностью назывался вид систематической сравнительной политической науки. Постепенно она сосредоточилась на числовых таблицах экономических, демографических и политических фактов, и статистика стала означать составление и исследование таких таблиц⁴. Нетрудно видеть, как постепенно возникло использованное здесь более философское значение этого слова.

Разумеется, история абстрактного исследования происхождения выводов из наблюдений уходит в далёкое прошлое, а само исследование называлось по-разному (например, теория ошибок⁵ или исчисление вероятностей). Лишь сравнительно недавно слово *статистика* обрело своё нынешнее значение. И даже сейчас по этому поводу полно странных недоразумений. Например, статистику понимают как установившийся порядок сбора неинтересных наборов чисел или в основном как коллекцию математических выражений.

[6.] Функции. Моё описание статистики конечно носит личный характер, но многие статистики в общем согласились бы с ним. Почти любое описание должно было бы включать её следующие общие функции.

– содействие сведению данных (измерений количественных, распределённых по категориям, по порядку и т. д.) воедино и извлечению информации из них

– содействие отысканию и оцениванию закономерностей в данных, затемнённых присущей им случайной изменчивостью

– содействие действительному планированию экспериментов и исследований

– содействие общению учёных друг с другом

Есть и иные роли, которые якобы статистическая деятельность может, к сожалению, играть. Вот две из них

– оправдывать или одобрять. К примеру, известно, что есть руководители диссертаций и редакторы журналов, которые настаивают на некоторых формальных и иногда неподходящих статистических действиях.

– производить впечатление, затуманивать или озадачивать. Некоторые социологические статьи содержат массу непереваренных формул, которые служат только для того, чтобы было видно насколько автор смышлён.

Иногда консультирующие статистики более или менее явно указывают на ответственность или на кодекс их отношений с клиентами, тем самым предохраняя себя, т. е. не оказываясь в роли одобряющего [их действий]. Существует хорошее правило: учёному, имеющему дело с эмпирическими данными, следует применять только те статистические методы, которые он может разумно понять, пусть даже без желания или способности вникать во все подробности математических выкладок. Deming (1965) обсудил общие отношения статистика и клиента и привёл обширную библиографию. В большинстве случаев применения прикладной статистики сам статистик, разумеется, и является клиентом.

[7.] Пример. В качестве иллюстрации указаний вводного комментария рассмотрим следующий условный пример изучения влияния пропаганды. Пусть во время общенациональной политической компании в США 100 студентам колледжей показали фильм, превозносящий кандидата демократов, а 100 других студентов (так называемая контрольная группа) его не видела. Затем всех студентов просят назвать своего предпочтительного кандидата, и пусть 95 из первой группы и только 80 из второй предпочли демократа. Какой вид заключений желателен по поводу действенности пропаганды?

Имеются, конечно, серьёзные вопросы о выборе студентов, о подробностях фильма и применении вопросника, о возможном обмене мнениями студентов, искусственности структуры эксперимента и т. д. Некоторые из них мы затронем ниже.

Будь указанные числа равны 95 и 5, действительное влияние пропаганды было бы установлено без особых забот, хотя методические вопросы появились бы при любой попытке оценить величину влияния. Если же, напротив, оба числа равнялись бы 95, немедленно последовало бы заключение об отсутствии этого влияния, хотя можно было бы усомниться в случайности подобного равенства даже [?] при наличии влияния. Но в конце концов наибольший статистический интерес имеют остальные случаи, как указанные выше числа 95 и 80: достаточно ли они различны, чтобы предположить действие пропаганды?

Простейшая стохастическая модель, пригодная для обсуждения описанного эксперимента, это его аналогия с бросанием каждой

из двух монет неравномерной плотности. Бросок монеты, описывающий действие пропаганды, подобен случайному отбору студента, которому покажут фильм, а бросок второй монеты – отбору студента в контрольную группу и т. д.

Пусть выпадение орла описывает выбор демократического кандидата. Вероятности исходов при броске монет неизвестны и, вообще говоря, не равны друг другу, а интерес представляет разность неизвестных вероятностей выпадения орлов. Пусть студенты отобраны случайно из больших студенческих совокупностей, и для студентов, случайно отобранных из двух указанных групп, вероятности выпадения орла равны p_A и p_B . Допустим, далее, что наблюдаемые предпочтения независимы. Грубо говоря, это означает, что даже при известных p_A и p_B и известной принадлежности этих студентов к той или иной группе предсказание предпочтения одного из них не улучшится от знания предпочтения другого.

Нарушение независимости может произойти, например, если студенты обсудили политические вопросы в промежутке между показом фильма и опросом. В указанных условиях вероятности различных исходов эксперимента при любых значениях p_A и p_B могут быть вычислены при помощи стандартных методов.

Интересующие нас величины (так называемые параметры p_A и p_B) неизвестны, иначе вряд ли был бы нужен эксперимент. Тем не менее, весьма интересно размышление об их возможных значениях и выбор того, что требуется больше всего. К примеру, является ли разность $(p_A - p_B)$ основной величиной? Или быть может p_A/p_B ? Или $(1 - p_B)/(1 - p_A)$? Выбор не безразличен. Если $p_A = 0,99$ и $p_B = 0,95$, применение статистического метода, чувствительного к их разности (к 0,04 в этом примере) может навести [наводит] на мысль о том, что параметры мало отличаются друг от друга, тогда как, если метод чувствителен к отношению $(1 - p_B)/(1 - p_A)$, которое здесь равно 5, он может указать на очень большое действие.

К сожалению, подобными рассуждениями часто пренебрегают, что может привести к неподходящему или искажённому анализу. При недавних обсуждениях возможного отношения между курением сигарет и раком лёгких произошёл спор о том, важны ли в основном отношения или разности смертности. Выбор здесь может привести к весьма различным заключениям.

{1}⁶ При графическом представлении даже как будто малые изменения могут быть исключительно важны. Skinner (1956, с. 225) указал, что в его работе оказался важным переход от графического изображения, которое просто показывало время происходивших событий (движений крысы), к логически равносильному *кумулятивному* изображению, которое показывало число событий, происшедших вплоть до каждого момента времени. В последнем случае доля соответствующих событий часто становится визуально ясной. Соответствующая область называется описательной статистикой, быть может с приставкой *нео*.

Выше было замечено, что по различным причинам предположение о статистической независимости вполне может быть ошибочным. Одна причина такова: студенты могли быть отобраны из групп, в которых обсуждается политика. Другие ошибки в предположениях также вполне возможны. К примеру, отбор студентов из одной и той же совокупности мог не быть случайным; быть может студенты сами выбирали группу для себя, и фильм показали более предприимчивым из них. Другой вид уклонения от первоначального простого предположения (в данном случае, запланированного) может произойти от преварительного уравнивания таких факторов как пол и возраст.

Если предположения сомнительны, их можно упростить (иногда привлекая более сложный, но и более утонченный анализ) или изучая влияние ошибки предположений на анализ, основанный на них. И если это влияние невелико, ошибками можно пренебречь. Эта тема, которую иногда называют *устойчивостью* при ошибочных предположениях о независимости, форме распределения и т. д., трудна и важна.

Другой общий вид вопросов относится к планированию эксперимента. Так, можно заранее выяснять, будут ли группы студентов в 100 человек достаточно (или слишком) велики; полезна ли равночисленность групп; не желателен ли их более сложный состав (быть может с учётом пола и возраста) и т. д. Подобные вопросы могут повлечь за собой формальные статистические рассуждения, но ответы на них в большой степени зависят от глубокого знания [существа дела].

Важно представить себе, что применение лучших методов измерения или изменение структуры опыта могут быть намного более существенными элементами плана, нежели простое увеличение объёма выборок. Вот Skinner (1956, с. 229):

Мы можем уменьшить беспокоящую изменчивость, поставив эксперимент в иных условиях. Выявляя и полностью изучая каждую существенную переменную, мы сможем заранее исключить те различия, которые затемняют анализ.

По поводу описанного выше эксперимента можно иметь в виду несколько подходов. Если все испытуемые – студенты одного пола, возраста, происхождения и жизненного опыта и т. д., то влияние пропаганды может проявиться отчётливее, хотя возможно за счёт общности результата. И вместо прямого опроса о политической предпочтительности это влияние быть может лучше выявится при наблюдении физиологических реакций испытуемых на фамилии или фотографии кандидатов⁷ или при опросе об основных политических проблемах.

Вероятно было бы полезно попытаться использовать общий принцип, при котором каждый испытуемый сам проверяет себя, а именно отмечая предпочтения, имевшие место и до, и после просмотра фильма и сравнивая число их перемен в каждом из обоих возможных направлениях. Но даже в этом случае было бы желательно сохранить контрольную группу (возможно показывая ей предположительно нейтральный фильм), чтобы выявить и

попытаться ввести поправки за искусственное воздействие обстановки эксперимента. Подобные проблемы часто исследуются в очень важных вспомогательных или предварительных изысканиях.

Рассмотрим в предварительном порядке простой конкретный случай двух групп, А и В (контрольная) при простых предположениях. Пусть доли студентов, предпочтвших кандидата демократов равны P_A и P_B . Эти две *случайные переменные* совместно с заданным числом студентов в группах содержат всю полученную и относящуюся к делу информацию и здесь таким образом появляется понятие достаточности. Анализ вполне мог бы быть утончен при учёте пола и других характеристик студентов, но пока что мы оставляем эту возможность в стороне. Прописные буквы мы будем применять для обозначения случайных переменных, а строчные – для соответствующих параметров.

[8.] Оценивание. Случайные переменные $100P_A$ и $100P_B$ распределены биномиально в зависимости от p_A и p_B и объёма выборок или групп. В соответствии с основополагающей предпосылкой большинства статистических методов p_A и p_B следует оценивать по P_A и P_B с учётом их возможных распределений. Одним из простейших видов оценивания является точечный, при котором результат анализа в нашем случае состоит из двух чисел (зависящих от наблюдений), которые считаются разумными оценками p_A и p_B . В нашем случае обычной (не единственной) оценивающей статистикой являются сами P_A и P_B , но уже небольшое изменение этой точки зрения может в некоторой степени затемнить дело. К примеру, пусть требуется точечная оценка отношения p_A/p_B . Совершенно неясно, будет ли в этом случае P_A/P_B хорошей оценивающей статистикой.

{2} Сами по себе точечные оценивающие статистики обычно недостаточны, ибо почти всегда требуется какое-то указание *точности*. Впрочем, есть и такие проблемы, в которых они-то главным образом и интересны, например в справочнике констант природы или, частично, при сделках купли и продажи. По старой традиции точечная оценка сопровождается знаком *плюс-минус* и числом, выведенным из предшествовавшего опыта или данных. Цель здесь состоит в предоставлении понятия о точности этой оценки, о размахе или рассеянии её распределения. В нашем случае один из принятых методов приводит к видоизменённой оценивающей статистике для p_A :

$$P_A \pm \sqrt{P_A(1-P_A)/100}.$$

Здесь добавлена оценивающая статистика стандартного отклонения первого слагаемого, P_A , а 100 – это, разумеется, объём выборки. Но смысл добавленного слагаемого может быть неверно понят, а истолкование меры рассеяния может оказаться косвенным, если распределение точечной оценивающей

статистики недостаточно простое. Обычно предполагают, что это распределение приближенно нормально.

Чтобы обойти эти проблемы часто применяют не точечную оценивающую статистику, а доверительный интервал. Перед экспериментом он случаен и выбирается таким образом, чтобы с заданной вероятностью (обычно близкой к 1) покрывать им неизвестное истинное значение оцениваемого параметра.

Идея доверительного интервала весьма полезна, хотя его утонченность часто приводила к недоразумениям, при которых истолкование неверно производилось в терминах распределения параметра. Впрочем, существуют точки зрения, в соответствии с которыми этот вид истолкования всё-таки верен, т. е. считается, что сами интересующие нас параметры случайны. Самыми важными из этих точек зрения являются байесовская и фидуциальная. Здесь существует много разработанных вариантов, и до сих пор происходят дискуссии по возникающим философским и практическим проблемам.

[9.] Проверка гипотез. Иным и более обычным общим подходом является проверка гипотез (или значимости). Его можно применять, если важно выяснить, совпадают ли p_A и p_B или нет. Проверка гипотез двусторонняя и позволяет выбрать одно из двух возможных действий с известными и контролируемыми шансами ошибок, а в обобщённом виде она приводит к теории проверки решений. Она вызвала к жизни весьма обширную литературу, относящуюся к теоретической статистике, и в ней, в этой теории решающих функций, непосредственно рассматривается цена ошибочных решений, а также и наблюдений, и она близка теории игр и менее родственна эмпирическим исследованиям принятия решений.

Вторая сторона проверки гипотез описательна в большей степени, чем первая и более часто в ходу. С её точки зрения проверка гипотез указывает, насколько необычен некоторый набор данных относительно *начальной гипотезы*. При этом подходе в нашем примере, если начальная гипотеза утверждает, что $p_A = p_B$, следует вычислить, насколько вероятно, что наблюдаемые 95 и 80% должны отличаться друг от друга не меньше, чем на 15%. Лишь недавно было подчеркнута, что имеет смысл дополнительно проверять подобные вероятности относительно иных гипотез.

Иногда с самого начала достаточно ясно, что, как в нашем примере, пропаганда должна оказывать какое-то действие, но в других случаях, например при изучении парапсихологии, могут возникнуть серьёзные вопросы о наличии какого бы то ни было действия⁸.

Существуют и другие виды статистического анализа, например классификация, отбор и скрининг (сплошная проверка), но вероятно будут исследованы намного более разнообразные виды анализа. Они смягчат трудности, которые часто возникают, когда стандартные виды анализа, подобные проверке гипотез, не вполне решают конкретные задачи. Эти виды анализа обычно следует считать приближёнными и требующими осторожного применения.

Одна общераспространённая трудность подобного рода связана с тем, что можно назвать предварительным исследованием данных. Возникает она тогда, когда при исследовании или эксперименте собрано (обычно значительное) количество данных, но либо нет никаких предположений о [скрытых в них] закономерностях, либо таких гипотез слишком много.

Статистик, конечно же, захочет разными способами исследовать свои данные. Он испробует различные виды графического представления, преобразования функций, быть может факторный и регрессионный анализ и другие методы. При этом он несомненно вычислит несколько оценок, проверит гипотезы, вычислит доверительные интервалы и т. д. Но есть здесь существенная трудность: длительное и серьёзное изучение любого конечного множества данных, даже порождённых вполне случайно, выявит в нём некоторую закономерность. Кроме того, следует помнить, что большинство теоретических изучений проверки гипотез, доверительных интервалов и других подобных процедур считает их как бы находящимися в изоляции, предполагает, что они были выбраны до исследования данных.

Пусть, например, проверяется начальная гипотеза о том, что средние уровни интеллекта мужчин и женщин совпадают. Если после наблюдений избран односторонний критерий, то он может вдвое чаще чем ожидалось ошибочно указать статистическую значимость, хоть начальная гипотеза и была верна. С другой стороны, при исследовании данных было бы до нелепости строго отказываться от применения выводных возможностей.

Существуют и два смягчающих подхода.

1. Применение методов (например, многократного сравнения), формулировки которых явно включают элементы исследования.

2. Случайное разбиение данных на две части, далее исследование одной из них любым методом, другой же части – с применением формальных критериев или других выводных методов.

Эта область заслуживает гораздо больше исследований. Selvin & Stuart (1966) перечислили существующие мнения и практические советы.

[10.] Охват выводов. При любом методе анализа важно помнить, что выводы, которые непосредственно связаны с некоторыми статистическими методами, ограничены фактическим исследованием или экспериментом. В примере о пропаганде, если все студенты были отобраны из одного и того же университета, непосредственные выводы будут относиться только к нему. Более общие и обычно желательные выводы предположительно зависят от предшествовавшей информации о существе дела и интуиции.

Охват непосредственных выводов можно, конечно, расширить, например, повторными исследованиями в различные периоды, в разных университетах, в разных районах и т. д. Но за исключением необычных случаев здесь достигается предел, быть может [по крайней мере] только временной, поскольку сегодня нельзя исследовать завтрашних студентов.

Таким образом, в большинстве случаев научный вывод достигается на двух стадиях: непосредственный вывод от выборки к соответствующей совокупности; косвенный вывод от этой совокупности к другой, намного более многочисленной и обычно довольно неопределённой. Вот почему так важно проверять результаты различными способами; к примеру, контролировать обобщения психологических экспериментов, проведенных в одной-единственной культуре, экспериментами совсем в другой культуре.

Формализация и точная теоретическая обработка, проводимая на второй стадии, не изучены нынешней статистикой; исключение быть может имеет место для приверженцев бейесовской методологии. Впрочем, многие утверждают, что эта вторая стадия по существу находится вне статистики. Общая проблема косвенных выводов часто упоминается и часто забывается. См. раннее чёткое рассуждение об этом Bortkiewicz (1909), а современное обсуждение исследований в контексте половой жизни см. Cochran и др. (1954, с. 18 – 19, 21 – 22 и 30 – 31).

{3} Исключительный случай проблемы охвата выводов представляет, например, исследование истории одного-единственного психологически неустойчивого лица. Действительно, некоторые авторы пытаются провести чёткую границу между исследованиями исключительных случаев и тем, что они называют статистическими методами. Я не чувствую, что эта граница столь уж резкая. Во-первых, статистические проблемы измерения надёжности возникают даже при изучении одного-единственного лица. Далее, некоторые особые случаи, например, в физической антропологии, относятся к племени или иной группе лиц, так что вполне могут появиться традиционные вопросы, относящиеся к выборочному методу, при выводах из единого коллективного случая.

Приверженцы исследований особых случаев подчёркивают их гибкость и значение для достижения субъективного интуитивного понимания и пользу как средства для формулирования предположений об интересных теоретических структурах. Если имеется хорошая причина поверить в существование лишь небольшой изменчивости между отдельными случаями, тогда, конечно, один из них многое скажет о более многочисленных совокупностях. Однако, исследователь обязан обосновать предположение об указанной малости.

[II.] Другие темы

[11.] **Линейные гипотезы.** Статистические темы можно классифицировать по виду принятых предположений, т. е. (применительно к приложениям) по структуре предполагаемых экспериментов или исследований, в которых применяются статистические методы. Пример с пропагандой, в котором главными величинами были две пропорции с целочисленными числителем и знаменателем, относится к общей теме исследований величин, принимающих натуральные значения или качественных величин. Эта тема также включает критерии хи-

квадрат, а подобные исследования применимы и к случаю большого числа групп, да и при обобщениях в других направлениях.

Если в указанном примере вместо пропорций, выражающих предпочтения, использовать численный счёт очков и критерии, основанные на многих вопросах, и таким образом количественно указать пристрастие к кандидату или политической партии, то проблему можно будет отнести к общей рубрике линейных гипотез. Пусть, к примеру, групп было не две, а, скажем, четыре. Далее, пусть первая не была подвержена никакой пропаганде, второй показали кинофильм, третьей дали материал для чтения, четвёртая слышала оратора, а затем были сравнены результаты проверки всех групп.

Особо важны здесь будут методы дисперсионного анализа (многие из которых можно считать специальными случаями регрессионных методов). В подобных случаях часто применяют методы многократных сравнений, хотя, строго говоря, они не ограничиваются рамками дисперсионного анализа.

Если эти группы в основном отличались друг от друга в каком-либо количественном смысле (например, по количеству просмотренных фильмов⁹), то можно было бы применить регрессионные методы в более узком смысле. Так, можно было бы предположить, что среднее количество очков примерно является линейной функцией числа просмотров, а затем сосредоточить внимание на константах (угле наклона и пересечения с осями координат) этой функции.

[12.] Многомерная статистика. Слово *регрессия* имеет по меньшей мере два значения. То, которое слегка отличается от только что описанного и исторически предшествовало ему, появляется в статистической теории многомерного анализа, если каждая измеренная единица наблюдалась более, чем одним видом наблюдений. К примеру, при эксперименте о методах преподавания можно руководствоваться не только проверкой правописания, но и результатами экзаменов по грамматике и по чтению и пониманию текста. В области физической антропологии можно измерять несколько частей тела каждого проверяемого.

Простейшая часть многомерного анализа относится к связи только между двумя случайными переменными и, в частности, к важному понятию корреляции. Эти понятия распространяются на большее число переменных, а затем появляются новые возможности. Важной является частичная связь: как соотносятся результаты экзаменов по правописанию и грамматике, если результаты третьего экзамена остаются без изменения?

Это понятие важно в анализе исследований, при которых контролируемый эксперимент часто невозможен. Многомерный анализ также рассматривает статистические методы, относящиеся к объединённой структуре средних из нескольких видов наблюдений, и структуру корреляции в целом.

{4} Факторный анализ является многомерным, но у него своя история и особые отношения с психологией. Методами этого анализа пытаются заменить некоторое число наблюдений

несколькими основными с остатками, обладающими простой вероятностной структурой. Так, можно надеяться, что способности в правописании, грамматике и чтении и понимании текста пропорциональны какой-то единой величине, которая непосредственно не наблюдается и которую видимо можно было бы назвать *лингвистическим мастерством*, изменяющимся от человека к человеку, плюс практически независимые остатки или уклонения.

Стандартная модель факторного анализа принадлежит классу моделей, порождённых процессом, который называется смешиванием распределений вероятностей. Интересна модель этого общего вида со скрытой структурой, хотя и для дискретных, а не непрерывных наблюдений. Другой важной темой многомерного анализа является классификация и различение, т. е. приписывание одного лица двум или более группам на основе нескольких его измерений. Менее хорошо понята родственная задача разбиения на группы или численная систематика, т. е. составление групп лиц [из некоторой совокупности] на той же основе.

[13.] Временные ряды. Ввиду его особого внимания на виды статистической зависимости, многомерному анализу родственна большая область анализа временных рядов, название которой иногда включает легко запоминаемые слова *стохастические процессы*. Наблюдённый временной ряд можно считать реализацией обосновывающего его стохастического процесса. Простейший вид проблем, связанных с временными рядами, может произойти, когда, например, для каждого школьника в эксперименте обучения имеется годичный набор ежемесячных результатов по правописанию.

Более трудные проблемы возникают, когда можно наблюдать только один ряд, когда, например, наблюдаются ежемесячные или ежегодные цены пшеницы. В таких случаях, столь частых в экономике, требуются точные структурные предположения, но даже тогда анализ труден. Эта обработка временных рядов в нынешней *Энциклопедии*¹⁰ начинается с общего взгляда, в основном ориентированного на ряды экономики, затем применяется современная методология, главным образом относящаяся к спектральному анализу.

Временной ряд рассматривается как нечто, похожее на радиосигнал, который может быть разложен на несколько сигналов, расположенных на различных частотах и обладающих различными энергиями. Далее рассматриваются циклы и особо обсуждается, как легко заключить, что они существуют, хотя на самом деле имеют место только случайные вариации. Наконец, изучаются важные технические проблемы, вызванные сезонными вариациями и выравниванием, которое имеет целью исключить или уменьшить их.

Цепи Маркова можно было бы включить в категорию временных рядов, но они всё же отдельны. Понятие об этих цепях является одним из простейших и весьма полезных для смягчения обычного предположения о независимости. Методы, основанные

на идее о цепях Маркова, нашли применение в панельных исследованиях (общественного мнения, бюджета и пр.), при изучении подвижности активного населения, изменения социального класса от одного поколения к следующему и т. д.

[14.] Выборочные исследования и родственные темы.

Выборочные исследования важны и теоретически, и практически. Они возникли для экономической и социальной характеристики населения, но выборочный метод оказался исключительно полезным во многих других областях, например для оценки запасов промышленного оборудования. Изучение выборочных исследований тесно связано с большинством других основных ветвей статистики, особенно с планированием эксперимента, но характеризуется упором на конечные совокупности и сложные планы работы.

Большинство научно ориентированных статистиков, занимающихся выборочными исследованиями, подчёркивают значимость вероятностных выборок, при которых элементы для наблюдения отбираются с явным использованием случайных чисел, так что вероятности возможных выборок становятся известными. С другой стороны, многие выборочные исследования не используют этого метода. Впрочем, случайные числа существенны не только для выборочных исследований, но для планирования экспериментов вообще и для многих вопросов исследования симулирования.

Для выборочных исследований и, если на то пошло, для прикладной статистики вообще, важны ошибки, не зависящие от выборок¹¹. Они появляются, например, при неполучении данных в исследованиях общественного мнения, ввиду склонностей наблюдателя и других систематических сдвигов при измерениях и из-за вычислительных ошибок. Интересное обсуждение этих и многих других проблем, родственных выборочному методу, см. в Cochran и др. (1954).

Социологи издавна интересовались полевыми исследованиями, но особое внимание они обращали не на обстоятельства, которые оказались основными для статистиков. Они намного меньше изучали эффективное планирование и вариации при выборочных обследованиях, и намного больше исследовали сложные методы анализа весьма многомерных данных. Можно надеяться, что эти два научных течения начинают понимать друг друга.

[15.] Непараметрический анализ и родственные темы. Выше я заметил, что важной темой исследований является устойчивость, степень чувствительности статистических методов к погрешностям предположений. Особым видом таких ошибок является неверное предположение о форме распределения, например, о нормальности.

Для устранения этих проблем можно отыскивать другие методы, которые устойчивы по отношению к форме распределения. Их изучение называется непараметрическим анализом; к нему относятся, например, критерий знаков и многие методы ранжирования. Пусть исследуются пары студентов, подобранные по возрасту, полу, интеллекту и пр. Для каждой

пары совершенно случайно решается, которые из них будут подвержены методам обучения А или В. После обучения студентов экзаменуют, и пара считается положительной, если метод А оказался успешнее, и отрицательной в противном случае. Если оба метода равно эффективны, то число положительных пар будет распределено биномиально с равной вероятностью обоих исходов ($p = 1/2$); если А эффективнее, то $p > 1/2$, в противном же случае $p < 1/2$. Число положительных пар является простым критерием гипотезы равносильности методов и основой для оценки возможного превосходства одного из них. Разумеется, подобный план исследования разумен только, если большинство студентов можно объединить в подобные пары.

Порядковые статистики также обсуждаются в рамках непараметрического анализа, хотя они по меньшей мере столь же важны в методах, которые используют чёткие предположения о распределениях.

Никакой резкой границы у непараметрических методов, конечно же, нет. Во-первых, многие методы, основанные на тесных предположениях о распределениях, оказываются устойчивыми, т. е. сохраняют некоторые или все свои характеристики, даже при менее жёстких предположениях. Во-вторых, большинство непараметрических методов являются таковыми лишь частично. К примеру, непараметрический критерий в типичном случае не зависит от формы распределения по отношению к своему уровню значимости, но не мощности (т. е. вероятности отклонения ошибочной начальной гипотезы).

Опять же, большинство непараметрических методов не является устойчивым относительно зависимости наблюдений. Непараметрические методы часто появляются естественно, если материалы наблюдений по существу не выражаются числами, а оказываются, например, ранжированиями. Иногда имеет смысл специально исследовать вид распределения, и для этого применяются методы проверки согласия. К примеру, психологический критерий можно нормировать относительно конкретной совокупности, так, чтобы счёт очков у отдельных лиц оказался почти нормальным $N(0, 1)$. Если теперь подвергнуть испытанию отобранных лиц из иной совокупности, то будет ли и в этом случае то же самое распределение? Это можно проверить критерием согласия, а в более общем смысле можно было бы построить сходный критерий для проверки одной лишь нормальности вообще.

Можно считать, что некоторые методы проверки согласия, а именно методы хи-квадрат, относятся к рубрике данных, принимающих лишь натуральные значения. Других, особенно после их видоизменения для получения доверительной полосы для всего распределения, обычно относят к непараметрическому анализу.

[16.] Рассеяние. Изучение рассеяния или вариации заслуживает большего внимания, чем ему обычно уделяют. К примеру, может оказаться интересным сравнение нескольких методов обучения по неоднородности числа очков, полученных студентами. Какой-

либо определённый метод может привести к желаемому среднему числу очков, если весьма значительно повысит результаты некоторых студентов и оставит их неизменными для других, хотя тем самым усилит неоднородность. Подобный метод явно приводит к последствиям и приложениям, отличающимися от того, который повышает результат каждого студента примерно одинаково. Терминология здесь может привести к путанице. Существенная часть дисперсионного анализа традиционно имеет дело со средними, а не дисперсиями, хоть и исследует дисперсии средних.

[17.] Планирование. Мы уже упоминали планирование эксперимента. Оно имеет дело с такими проблемами, как определение числа наблюдений для достижения заданного уровня точности, выбор способа обработки или факторов для экспериментальных элементов. При исследовании методов обучения такими элементами могут быть школьные классы, расположенные по возрасту учащихся, типы школ и окружающего населения и т. п.

Планирование эксперимента имеет дело с формальными сторонами его структуры. Основное принципиальное значение имеет чёткая рандомизация при назначении процедур (в данном случае, методов обучения) экспериментальным элементам (классам). Впрочем, может оказаться разумным предположить, что рандомизация была присуща этим элементам самой их природой, но чаще оказывается важным применять так называемые случайные числа. Споры сосредоточиваются на случаях, при которых рандомизация непрактична, неэтична или даже невозможна. Иногда, однако, можно отыскать ловкие пути включить её, хоть это и кажется с первого взгляда невозможным.

В противном случае можно применять термин *квазиэксперимент*, чтобы подчеркнуть её отсутствие. Серьёзной проблемой окажется поиск наилучшей защиты от источников смещения, которые можно было бы в основном исключить при рандомизации.

Важной стороной планирования эксперимента является использование должных мер, обеспечивающих *двойной слепой* опыт, при котором испытуемый не знает, какой процедуре он подвергается, а экзаменатор не знает, какой процедуре подверглись те или иные испытуемые.

Многие экспериментальные исследования были испорчены ввиду пренебрежения подобными предосторожностями. Во-первых, испытуемый, знающий, что принимает лекарство, которое, как надеются, улучшает память или уменьшает чувствительность к боли, вполне может изменить своё поведение, поскольку узнал про ожидаемое, и не в меньшей степени чем при физиологической реакции на само лекарство. Поэтому, когда только возможно, члены контрольной группы получают плацебо (нейтральное вещество, не отличимое от настоящего лекарства).

Во-вторых, предвзятое мнение экспериментатора, знающего, какие испытуемые подвергаются той или иной процедуре, может легко и совсем бессознательно сместить его наблюдения.

Проблемы могут возникнуть даже тогда, когда он только знает, какие испытуемые входят в ту же самую группу. Назначение процедур посредством случайных чисел и случайный подбор испытуемых, – вот важные средства для обеспечения беспристрастности.

Число наблюдений по традиции считается установленным до выборочного обследования. Впрочем, в последние годы было много исследований планов последовательного изучения, при котором наблюдения проводятся по сериям (или по сериям их групп), и на каждом шагу [в зависимости от достигнутых результатов] заново принимается решение о необходимости дальнейших наблюдений.

Во многих случаях ответ (или его среднее значение) есть функция нескольких контролируемых переменных. К примеру, среднее время на переучивание правописания списка слов может зависеть от предшествовавших занятий и времени, прошедшего после последнего из них. При изучении поверхности отклика¹² структура зависимости (форма поверхности) исследуется серией экспериментов, обычно со специальным вниманием к окрестности максимума и минимума.

[18.] Философия. В методологическом городе статистика издавна соседствует с философией науки. Она, однако, обычно скромнее по своему охвату и практичнее в своих воззрениях. В строгом смысле слова статистика – отрасль философии науки, хотя на самом деле эти две области обычно изучаются отдельно друг от друга.

Каковы же некоторые проблемы, которые составляют часть философии науки, но обычно не считаются частью статистики? Главная из них – это формирование научных теорий, их тщательное формулирование и подтверждение, полное или частичное. Последнее следует отличать от более узкого, но лучше понимаемого статистического понятия проверки гипотез. Другая проблема, которая по ощущению многих статистиков находится вне статистики, это разрыв между совокупностью, исследуемой выборочно, и целевой. Есть и другие области научной философии, которые обычно не причисляются к статистике. На ум приходят такие понятия как объяснение, причинность, операционализм¹³, свободная воля.

Классическим сочинением, рассматривающим и статистику, и научную философию, является книга Пирсон (1892). Две позднейшие публикации, это Popper (1935) и Braithwaite (1953). В общем, сегодня авторы, называющие себя статистиками, и те, которые считают себя философами науки, часто ссылаются друг на друга, однако связи между ними ограничены и случайны.

Измерения являются важной темой в статистике, и их вполне можно здесь упомянуть, потому что их некоторые стороны явно философские. Грубо говоря, измерение есть процесс приписывания чисел (или категорий) объектам на основе какой-либо операции. Но каково обоснование этого понятия с точки зрения теории познания? Следует ли его расширить и включить другие общие виды данных кроме чисел и категорий?

Важна, в частности, и теоретически, и практически, *шкала* измерений. Естественно сказать, что одно тело вдвое тяжелее другого (при любых единицах измерения), но утверждать, что температура первого вдвое выше было бы глупо, хотя бы потому, что (если не иметь в виду абсолютные температуры) при переходе от одной обычной шкалы температур к другой (например, от шкалы Фаренгейта к шкале Цельсия) отношение температур изменится.

С другой стороны, есть смысл в утверждении, что одно тело на 100° Фаренгейта горячее другого. [...] Некоторые измерения по своему существу являются круговыми (направление ветра или час дня). Имели место оживлённые дискуссии о значимости или обоснованности арифметических действий с различными видами измерений. Есть ли, например, смысл в осреднении субъективных громкостей, если каждое их измерение относится лишь к порядковым соотношениям [к их рангу]?

Упомянем некоторые важные публикации по этой теме: Churchman & Ratoosh (1959); Coombs (1964); Pfanzagl (1959); Adams и др. (1965); Torgerson (1958); Stevens (1946); Suppes & Zinnes (1963).

[19.] Общение и заблуждения. Существует искусство общения статистика и иного специалиста. Первый обязан всегда помнить, что его собеседник в общем непосредственно не заинтересован ни в технических подробностях, ни в профессиональном статистическом жаргоне. А консультирование статистика часто оказывается бесполезным, потому что иной специалист не упоминает те стороны своей работы, которые относятся к статистике. Вообще-то в большинстве случаев учёные являются своими собственными статистиками в том же смысле, в котором мы все за исключением ипохондриков¹⁴ являемся своими собственными врачами.

Статистические заблуждения часто малозаметны, и могут произойти у самых осторожных специалистов. Изучение подобных заблуждений интересно само по себе и помогает смягчить упомянутую проблему общения.

[III.] Критика

Если статистика понимается как общее изучение перехода от наблюдений к заключениям, решениям и т. д.¹⁵, то вряд ли можно будет оспаривать желательность подобного столь широкого исследования. Критика статистики поэтому обычно относится к более узкому виду деятельности, часто к тому виду *статистики*, который критик замечает возле себя.

Допустим, что специалист в какой-либо научной области видит неуклюжий опубликованный анализ, называемый статистическим. Тогда он вполне может составить себе отрицательное мнение о статистике, и так и не узнать, что область статистики шире и что её можно применять разумно, изящно и эффективно.

[20.] Критика приложения вероятностей в статистике. Некоторые критические замечания, сформулированные в философском духе, относятся к применению вероятностных моделей в статистике. Некоторые авторы возражали против

вероятностей ввиду строгого детерминизма своего мировоззрения. Эта точка зрения ныне редка ввиду столь вероятностной сущности квантовых методов в физике и пользы вероятностных моделей для явно детерминированных процессов, например, при изучении влияния ошибок округления в сложных числовых вычислениях.

Но критик, настроенный детерминистски, вероятно скажет, что квантовая механика и вероятностное исследование ошибок округления являются временными приемами, которые будут заменены иными средствами. В 1947 г. Эйнштейн (Born 1949, с. 123) писал:

Статистическое истолкование [как в квантовой механике] [...] во многом верно. Но я не могу всерьёз поверить в него, потому что [подобная] теория противоречит тому принципу, что физика должна представлять реальность. [...] Я совершенно убеждён, что в конце концов мы придём к теории, объектами которой, соединенные законами, будут не вероятностями, а установленными (conceived) фактами. [...] Впрочем, я не могу привести логических доводов в пользу своего убеждения и только вызываю в свидетели свой мизинец, который не может требовать к себе никакого уважения вне моей собственной кожи¹⁶.

Другие критики замечают вредящие противоречия и парадоксы в понятиях вероятности и случайности. Так, Spenser Brown (1957, с. 66) обобщённо заявил, что

Понятие о вероятности, используемое в статистической науке, само по себе не имеет смысла. [...] Как бы оно ни было осмысленно, его смысл тем не менее остаётся бесплодным ввиду невозможности выявить информацию из экспериментальных результатов¹⁷.

Эта довольно-таки нигилистическая точка зрения необычна, и её трудно примирить со многими успешными применениями вероятностных идей. И действительно, этот автор привёл и конструктивные уточнения. Не столь крайнее, но родственное мнение высказал Bridgman (1959, с. 110 – 111). На обоих этих авторов повлияло статистическое применение таблиц случайных чисел, особенно в парапсихологии, в которой пояснение загадочных результатов отыскивали в возможном ошибочном поведении этих чисел.

[21.] Критика ограниченной пользы. Более обычная критика, особенно со стороны некоторых физиков, состоит в том, что статистика для них бесполезна, поскольку случайная изменчивость в их проблемах пренебрегаема, по крайней мере сравнительно с систематическими ошибками. Подобную же точку зрения приняли некоторые экономисты, особенно в связи с индексами. Психолог Скиннер убедительно выразил её вариант. Существует, мол, так много важных проблем, в которых случайная изменчивость пренебрегаема, и он ограничится их изучением. См. Skinner (1956), где представлено это довольно крайнее мнение. И далее он заявляет, что важной проблемой в психологии в целом является отождествление переменных с

пренебрегаемой изменчивостью, которых можно наблюдать непосредственно.

Тем не менее, после подробного исследования часто оказывается, что случайная изменчивость важнее, чем предполагалось, особенно при планировании будущих экспериментов. Далее, тщательное их планирование часто может уменьшить или пояснить систематические ошибки.

Я думаю, что критика указанного выше вида иногда справедлива. В конце концов один-единственный спутник, успешно обращающийся вокруг Земли, уже доказывает возможность этого достижения, но всё же обычно критика отражает нежелательность явно рассматривать статистические методы или же неверное понимание смысла статистики.

Указанным критическим утверждениям родственна точка зрения о том, что статистика хороша для прикладной техники, но не для фундаментальной науки. В своей вступительной лекции в одном из колледжей Лондонского университета Сох (1961) выступил против этой критики:

Ныне существует ощущение, что в некоторых областях фундаментальных исследований статистические идеи могут оказаться не только неуместными, но вредными, будучи признаком излишне эмпирического подхода. Эта точка зрения понятна, но происходит она, видимо, от слишком узкого представления о статистических методах.

Кокс привёл примеры применения статистики в фундаментальных исследованиях в физике, психологии, ботанике и других областях знания.

Иногда заметен другой вариант этой критики (Selvin 1957; Walberg 1966): статистические методы, подобные проверке гипотез, мол, сомнительны, если только эксперимент не проведен классически и не включает рандомизации, контрольной группы, предварительного установления гипотез и других предосторожностей. И без таких мер (которые иногда невозможны или не практичны) могут произойти любые сдвиги.

Эта критика отражает реальную проблему обеспечения разумных заключений при отсутствии подлинного эксперимента, но она не относится только к специальным видам заключений, а равно приложима к любому анализу (формальному или нет, интуитивному). Kish (1959) энергично обсудил эту тему.

{5}[22.] Гуманистическая критика. Некоторые критические выступления против статистики либо вызваны серьёзным непониманием дела, либо относятся к *плохому* статистическому методу, а не к статистике вообще. К примеру, можно иногда слышать, что статистика бесчеловечна, что *нельзя сводить людей к числам*, что гуманисты обязаны бороться со статистикой (и может быть с наукой в целом).

Такова статистическая версия давнишних жалоб, одну из форм которой высказал Walpole (1778) в одном из своих писем:

Эта утонченная эра сводит всё к своему существованию. Все окольные выражения и восклицания настолько вышли из

употребления, что вскоре, как я полагаю, приступить к занятию любовью можно будет только сказавши Ложись!

И вот современный вариант того же мнения (Auden 1946):

Ты не станешь отвечать на вопросники,

Или участвовать в викторинах о положении в мире.

И не будешь, подчиняясь формальным правилам,

Подвергаться испытаниям.

Ты не посмеешь якшаться со статистиками

И не совершишь социологии.

Krutch (1963) сказал:

Я всё ещё думаю, что знакомство с лучшими мыслями и высказываниями писателей полезнее, чем вся статистика социологов.

В подобных утонченных и чарующих критических высказываниях, конечно же, скрыты вполне обоснованные моменты. Легко позабыть, что вещи могут быть сложнее, чем кажутся, что многие важные характеристики исключительно трудно измерить или сосчитать, что учёным (равно как и гуманистам) может не хватать профессиональной скромности и что любой набор измерений исключает иные в принципе возможные измерения. Но гуманитарная критика излишне оборонительна и представляет собой частный случай убеждения в том, что физические и естественные науки по существу отличаются от и противостоят гуманитарным наукам.

{6} [23.] Критика излишних забот о средних. Статистиков иногда дразнят за то, что они интересуются только средними, которые иногда нелепы: 2,35 ребёнка в средней семье, или редкое заболевание, которому подвержены люди в среднем возрасте 40 лет (один из них – двухлетний ребёнок, другой – старик 78 лет).

Skinner (1956, с. 228) заметил по этому поводу, что

никто не идёт в цирк, чтобы увидеть, что средняя собака прыгает через обруч значительно чаще, чем не обученная собака, выросшая в тех же условиях,

а Krutch (1963, с. 15) заявил, что статистика *не принимает во внимание тех, кто предпочитает слушать другого барабанищика.*

В действительности, хоть средние и важны, статистики издавна были серьёзно озабочены рассеянием около них и другими характеристиками распределений, например, крайними значениями. Гальтон (1889, с. 62) поэтически раскритиковал средние:

Трудно понять, почему статистики обычно ограничиваются изучением средних и не наслаждаются более исчерпывающими зрелищами. Их души, видимо, так же безразличны к очарованию изменчивости, как и у уроженца одного нашего плоского графства. Вспоминая о Швейцарии, тот заявил, что, будь возможно сбросить её горы в её озёра, сразу же исчезли бы две пакости.

Даже в то время эта зарисовка была чрезмерной, а сегодня она была бы вовсе неуместной. Другое высказывание Гальтона (с. 62 – 63) относится к упомянутому выше чувству сопротивления статистике:

Некоторые ненавидят само выражение статистические данные, но для меня они интересны и наполнены красотой. Если только не колотить по ним кувалдой, а утонченно обрабатывать возвышенными методами и истолковывать с осторожностью, они будут исключительно эффективно обращаться со сложными явлениями. Они являются единственным средством для прокладки пути сквозь устрашающую заросль трудностей, которые останавливают изучающих науку о человеке.

{7} Один из основных источников непонимания средних состоит в том, что одно-единственное значение может во многом быть средним, но заметно не таковым в других отношениях. Это и оказалось основным затруднением в исторически важном понятии среднего человека у Кетле¹⁸. По этому поводу появился сатирический роман (Aurthur 1953).

Среднее число детей в семье из данной совокупности имеет смысл. Иногда полезно знать его, например, для оценки будущего населения, но средней семьи не существует, хотя бы потому, что семья со средним числом детей (пусть натуральным) не будет средней относительно числа, обратного указанному. Или иначе, нет причины полагать, что у подобной семьи и доход средний, или образование среднее из всех семей, или что она проживает в центре населения своей страны.

[24.] Критика излишней математики. Иногда, притом часто сами статистики, критически заявляют, что статистика чересчур математизирована. Эта критика проявляется в различных формах, например

1. Статистики выбирают для себя исследовательские проблемы ввиду их математического интереса или изящества и потому не занимаются такими, которые действительно заботят статистику. Впрочем, иногда просто имеется в виду ... *которые заботят критика.*

2. Применение математических понятий и языка затемняет статистическое мышление.

3. Ударение на математическую сторону статистики склоняет статистиков к пренебрежению должными целями, значимостью количественных данных и их точностью.

Критику подобного рода высказывали, например, Woytinsky (1954) и Gini (1951; 1959). Аналогичную критику можно найти и у Hogben (1957). Что можно сказать? Происходит ли критика изнутри профессии или вне её? Она имеет почтенную историю, доходящую до ранних дней статистики. Быть может первый подобный спор имел место в те дни, когда слово *статистика* применялось в смысле, отличном от нынешнего, и означало систематическое изучение государства, род политической науки. Спор происходил между теми *статистиками*, которые в свободной манере описывали государства, и теми, которые занимались *табличной статистикой*¹⁹. Последняя применялась в различных вариантах, от удобного расположения словесных сводок до настоящих таблиц рождений, женитьб и смертей.

Описание этих споров см. в Westergaard (1932, с. 12 – 15), Lundberg (1940) и Lazarsfeld (1961, особо с. 293).

Довод к человеку: кто-то главным образом математик и поэтому не способен понять истинно статистические проблемы. Он был и остаётся к сожалению популярным риторическим средством и, видимо, частично является защитной реакцией на почётное положение и престиж математики.

По моему мнению, очень многие споры подобного рода были неуместными, хотя некоторые обвинения, сформулированные той или другой стороной, несомненно верны.

Если какая-нибудь ветвь математики оказывается полезной в статистике, её будут применять. И по мере того, как статистики сталкиваются с математическими проблемами, они станут изучать их, привлекая всё, что смогут из современного математического знания, и быть может поощряя или проводя надлежащие математические исследования. Да, некоторые статистики используют ненужную математическую манеру изложения, и менее математически настроенные коллеги могут посчитать это раздражающим притворством, но кто же способен действительно отличить притворство от естественного способа общения? Tukey (1961) обсудил и осветил отношения между математиками и статистиками, равно как и многие другие темы.

{8}[25.] Критика затемнения. Далее, статистику обвиняют в том, что она является будто бы привлекательным средством затемнения или путаницы: *Ложь, двойная ложь и статистика!* Происхождение этого ошибочного крылатого выражения не вполне ясно (White 1964; Sheynin 2003). Его вариантом является критическое утверждение о том, что статистический анализ, сопровождаемый неразборчивыми схемами и формулами и выраженный на жаргоне, невозможно понять.

Часто это мнение хорошо описывает некоторые статистические или псевдостатистические сочинения, но оно не относится к статистике как к дисциплине. Популярная книга Huff (1954) по существу показывает отвратительные ошибки в статистических описаниях и анализе, но её можно, конечно, использовать и для пагубной софистики. В некоторой степени она была бы аналогична воображаемой книге *Как подделывать деньги*, предназначенной для защиты банковских кассиров или населения вообще.

Lundberg (1940, с. 138) предложил убедительную защиту от одной из форм этой критики:

С учётом глупости, неспособности и алогизмов оказывается, что чем конкретнее применяемые термины и методы, тем грубее ошибки. По этой причине легче выявлять ошибки авторов, имеющих дело с числами. Ошибка такого же рода в жюльнической риторике может показаться менее возмутительной и даже остаться незамеченной или стать почётной избитой фразой.

[26.] Критика выборочного метода самого по себе. Иногда замечаешь утверждение о том, что разумное заключение от выборки к совокупности невозможно, особенно если выборка

является её незначительной частью. Joseph Papp (Kadushin 1966, с. 30) привёл вариант этого мнения:

Эта методология [...] не была научной: они применили выборочный метод, но ведь по выборкам нельзя составить себе полную картину.

Подобная критика обоснована лишь постольку, поскольку она ставит под сомнение *плохие* выборочные методы. Выборки используются всегда, потому что часто не практично или невозможно наблюдать всю совокупность (новое лекарство нельзя испытывать на каждом, и нельзя проверять каждую изготовленную лампочку на продолжительность горения), да и вообще больше информации предоставляет тщательное измерение выборки, чем грубое измерение всей совокупности. Должное выборочное исследование, для которого абсолютная величина выборки намного важнее, чем её доля относительно совокупности, обеспечивает нужную информацию и неизменно применяется с успехом.

[27.] Критика интеллектуального империализма. Иногда заявляют, что статистика не представляет научного метода и практику целиком. Вот Skinner (1956, с. 221):

Ошибочно было бы отождествлять статистическую практику с формализованными структурами [подчёркивание наше – У. К.] статистики и научного метода. У этих дисциплин [?] есть своё место, но оно не совпадает с местом научного исследования. Вопреки тому, что так часто подразумевается, они предлагают некоторый научный метод, но не истинный метод. В качестве формальных дисциплин они появились очень поздно, и большинство научных фактов было обнаружено без их участия.

Я знаю нескольких статистиков, которые настолько высокомерны, что приравнивают свою область с научным методом вообще. Верно, конечно, что большинство научной работы было произведено без помощи статистики, понимаемой в узком смысле слова как определённые формальные и ныне распространяемые методы анализа.

С другой стороны, добрая часть научных сочинений так или иначе связана со статистикой, поскольку в более общем смысле спрашивает, как можно выводить разумные заключения. Скиннер (с. 221, 231) добавил в некоторой степени родственное замечание: ввиду престижа статистики статистические методы (в психологии) заняли исключительное почётное положение.

Статистикам, к сожалению, известно подобное применение статистики в некоторых областях научных изысканий, но профессию в целом нельзя винить за это только на основании некоторых империалистических учебников, многие из которых написаны не настоящими статистиками.

[IV.] Другие области статистики

В конце статьи я кратко обсуждаю остальные статистические статьи из *Энциклопедии*.

[28.] Сгруппированные наблюдения. Проблема сгруппированных наблюдений иногда озабочивает: в

теоретической статистике часто считается, что наблюдения непрерывны, тогда как они всегда дискретны, а группировка поэтому неизбежна. Кроме того, часто желательно группировать наблюдения, чтобы упростить их описание и анализ.

В какой мере дискретность и группирование желательны, и в какой мере опасны?

[29.] Усечение и цензурирование. Наблюдения часто в разумной мере соответствуют какой-либо стандартной модели, но некоторые из них, значения которых оказываются выше или ниже определённых границ, запрещаются (усекаются или отбрасываются). Немного более сложный случай происходит при сравнении оценок студентов в начале и конце некоторого курса лекций. Студенты, показавшие неважные знания, могут быть не допущены к лекциям, и у них вообще не будут никаких заключительных оценок. Впрочем, существуют методы решения подобных проблем.

[30.] Отклоняющиеся наблюдения. Очень часто значения нескольких наблюдений в выборке оказываются необычно большими (или малыми) и эти наблюдения могут считаться отклоняющимися. Как поступать с ними? Если включать их наравне с остальными, они могут исказить анализ; если произвольно исключать их, возможна потеря важной информации; так каких же правил следует придерживаться при их исключении?

[31.] Преобразование данных. Часто оно очень полезно. Так, можно прологарифмировать время реакции, извлечь квадратный корень из экзаменационной оценки и т. д. Цели подобных действий таковы.

1. Упрощение структуры данных, чтобы добиться, к примеру, аддитивности влияний двух видов.

2. Достижение большей согласованности данных с хорошо понятой статистической моделью, добиваясь, к примеру, близкой к нормальности или постоянной дисперсии.

Опасность преобразований состоит в том, что выводы могут оказаться выраженными в шкале, которая не представляет большого интереса.

[32.] Аппроксимация распределений. Она важна и для теории вероятностей, и для статистики. Во-первых, аппроксимирование какого-нибудь теоретического распределения может иметь целью переход к простой аналитической форме или вычисление численных значений. Во-вторых, эмпирическое распределение может быть аппроксимировано для описания и вывода заключений.

[33.] Отождествление: смесь распределений. Проблема отождествления появляется как только точная модель какого-либо явления специализируется и её параметры оцениваются по эмпирическим наблюдениям. Может случиться (и даже оставаться незамеченным), что параметры никак не поддаются оценке по имеющимся данным.

Пусть в модели теории обучения оставшаяся в памяти доля изученного материала после некоторого времени равна отношению двух параметров модели. Если даже эту долю можно

наблюдать без выборочной изменчивости и ошибок измерения, эти параметры нельзя будет определить по-отдельности.

Проблема отождествления возникает, разумеется, в достаточно сложных случаях, так что вряд ли удастся сразу же установить его невозможность. Иногда возникает аналогичная проблема, которую можно назвать отождествляемостью модели.

Классический случай этого происходит при изучении статистики несчастных случаев. Её некоторые виды удовлетворительно описываются отрицательным биномиальным распределением, но само это распределение можно установить по исходу нескольких совсем отличных и более фундаментальных моделей. Некоторые из этих моделей иллюстрируют важное понятие смеси.

Смешивание – важный и полезный способ формирования нового распределения по двум или более статистическим распределениям.

[34.] Приложения. Теперь мы опишем специальные вопросы, связанные с конкретными областями приложений, хотя большинство последних также послужили побудительным источником общей теории.

Контроль качества. Статистический контроль качества был порождён массовым производством, но впоследствии его область применения расширилась. На эту тему в *Энциклопедии* имеется три статьи.

1. Выборочный приёмочный метод. Обычно речь идёт о партиях изделий, и здесь имеется близкое сходство с проверкой гипотез и последовательным анализом.

2. Регулирование процессов (и контрольные карты), которое само по себе иногда называется статистическим контролем качества в более узком смысле. Разработка понятий контрольных карт и методов связано с основными понятиями случайности и устойчивости, ибо важным нормативным понятием является *процесс под контролем*. Он доставляет последовательность чисел, которые ведут себя как независимые и одинаково распределённые случайные переменные.

3. Надёжность и испытание на долговечность. Эта тема также выходит за непосредственные инженерные рамки.

{9} Государственная статистика. Она весьма важна для принятия экономических, социальных и политических решений. Статья о ней [] в *Энциклопедии* рассматривает такие основные вопросы как её применение для политической пропаганды, её значение и точность, проблему доверительности. См. также опубликованные там же статьи Eldridge, Grauman, Liu, Mayer, Moriama, Ruggles, Sauvy, Spengler, Spulber и Whitney.

Индексы. Экономические индексы являются важной частью государственных статистических программ. Три статьи по этой теме [в *Энциклопедии*] обсуждают теорию, практическую сторону индексов и проблемы выборочного исследования.

Статистика в качестве юридического свидетельства. В последние годы применение статистических методов и их результатов в юриспруденции возросло. Споры с профсоюзами были освещены выборочными исследованиями; проблемы

установления отцовства исследовались вероятностными методами; изнашивание и другие учётные величины, возникающие при квазиюридических слушаниях, оценивались статистически. Существуют действительные или кажущиеся противоречия между статистическими методами и юридическими понятиями, подобные тем, которые относятся к свидетельствам, основанным на услышанном от других лиц.

*Статистическая география*²⁰. Приложение статистических и других количественных методов к географии быстро возрастает. В некоторой степени родственным является ранжирование по размеру, при котором эмпирически и теоретически изучаются схемы взаимоотношений между, скажем, населением городов и их ранжированием, начинающимся с наиболее населённых. Другой пример того же рода представляет соотношение между частотой слов и их ранжированием, начинающимся с наиболее часто употребляемых.

Конечное число исходов. Имеется в виду теория и метод, которые могут быть отнесены к данным, принимающим натуральные значения, или к линейным гипотезам с регрессией. Соответствующая задача может быть поставлена для студентов, которые обучались неделю, две недели и т. д. (пусть по сотне различных студентов для каждого заданного периода времени), с регистрацией их доли, сдавшей экзамен. Интересно было бы выяснить при какой длине периода экзамен сдаст в точности половина студентов. С этой же точки зрения могут быть рассмотрены многие традиционные психофизические проблемы.

Очереди. В последние годы их изучение оказалось важным. Иногда оно считается частью исследования операций, но можно его считать и частью изучения случайных процессов. Примером анализа очередей может служить исследование потока транспорта на регулируемом перекрёстке. Подобные исследования имеют эмпирическую, теоретическую и нормативную стороны.

Вычисления. Они всегда переплетены с прикладной статистикой, хотя и отличны от неё. Недавнее вторжение компьютеров привело к ряду качественных изменений в возможных вычислениях. Оно также имело и будет иметь серьёзное влияние на статистику, – не только на обработку данных и анализ, но и на теорию, поскольку многие аналитически нерешаемые задачи теперь могут быть численно исследованы при помощи симулирования.

Кибернетика. Это ныне модное слово применяется к несколько бесформенной области знания и исследований, связанных с обработкой информации и механизмами, включающими живые организмы (Maron 1968; Stagner 1968). Основными понятиями являются здесь регулирование и обратная связь. Влияние компьютеров оказалось существенным. Иногда в эту область включают теории связи и информации, см. статью Pollack (1968), в которой упор сделан на приложения к психологии.

[V.] Добавление

1. Bickel & Lehmann (1975) предложили новый и необычно глубокий подход к описательной статистике. Они подчеркнули

истолкование, пригодное при широких условиях, а также эффективность и устойчивость оценивания.

Новая статья о предварительном исследовании данных, написанная для этой *Энциклопедии*, описывает решительный и несколько рискованный вид описательной статистики, который был в большей части разработан под руководством J. W. Tukey. См. также Tukey (1970).

2. Eisenhart (1963) обсудил неопределённость в символе *плюс минус* и смежные проблемы выражения неопределённости.

3. В широком контексте в связи с изучением особых случаев можно сослаться на речь президента Американской ассоциации продвижения наук (Margaret Mead 1976). Я сослался бы на Dukes (1965), но не знал о его заметке. Не знал, к сожалению, и о предшествовавших сочинениях Redfield (1940), Wallin (1941) и Allport (1942), которые обращаются к нам переступая через годы.

4. Две новые важные статьи, написанные для этой *Энциклопедии*, обсуждают факторный анализ и основные компоненты, и многомерное шкалирование.

5. Гуманистические критические высказывания по-прежнему легко отыскать, как, например, в обзоре Gilman (1975, с. 2):

Это не совсем кульминация книги. Будь это иначе, её невзрачность и поверхностность, её мудрость статистика повредила бы автору сильнее, чем он полагает.

Другим примером служит выступление кардинала Franziskus König (1974):

При помощи статистики можно доказать что угодно и вообще всё, включая противоположное. [...] Вторжение социологии в богословие привело лишь к новому варианту старинной веры в чудеса: к наивной вере в непогрешимость вопросников, представительных исследований и анализов сечений общественного мнения. Это вторжение, никак не скованное сознанием своих ограничений, имело поэтому плачевное влияние.

Есть также трогательный вариант гуманистической критики в книге Elizabeth Barrett Browning *Aurora Leigh* 1856г.:

Рыженький ребёнок в лихорадке. Дотронься до него хоть кончиком пальца – заплачешь. Но если больных – миллион? Тогда плачь хоть над тройным правилом или неправильными дробями²¹.

6. Полно критики и по поводу излишней заботы о средних. Wilkes (1975) описывает год жизни американской семьи. По некоторым характеристикам типа интересующих переписчиков она средняя, но не нужно даже упоминать, что во многом другом она далека от средних. Ошибочное мнение о средних высказал без всяких ограничений Bagley (1968, с. 80):

Он был во всём средний. Ни слишком высок, ни слишком низок, ни слишком толст, ни слишком тощ. Он был одет в средний костюм и выглядел как вполне средний человек. Его мог бы спроектировать статистик. Его мозг был более, чем средний, но этого не было заметно.

7. Этот параграф в расширенном виде см. Kruskal (1973b).

8. Вариант критических замечаний о затемнении сути появился в газетной полосе медицинских советов (Mendelsohn 1976). Он указал (я несколько сократил текст):

Статистически можно доказать почти любую предпосылку и противоположное ей. Если увидите статью, в выводы которой вы не можете поверить, то вспомните следующую историю.

Хорошо известный статистик часто путешествовал по железной дороге, а не на самолётах, потому что статистический анализ выявил, что существует значимый шанс того, что на самолёт пронесли бомбу. Однажды его друг поразился, что тот заходит в самолёт, но статистик сказал, что “вероятность нахождения двух бомб на самолёте не следует принимать во внимание, так что я провожу свою собственную бомбу”.

Этот пример, использованный как критический, особо странен и забавен, потому что его издавна охотно приводят преподаватели статистики, чтобы указать распространённую ошибку, основанную на неверном понимании условной вероятности.

9. Новая статья о статистике и общественной политике, специально написанная для *Энциклопедии*, и обновлённая статья [xiv] совместно представляют подробную картину взаимоотношений статистики и различных проблем и обсуждаемых вопросов, относящихся к [общественной] политике. Следует упомянуть и отчёт 1971 г. Президентской комиссии по федеральной статистике, равно как и три важные книги, две об экспериментальной оценке социальных новшеств, а третья – руководство о статистической стороне общественной политики (Riecken & Boruch 1974; Boruch & Riecken 1976; Fairley & Mosteller 1977).

[35.] **Замечания об источниках.** Дополнительный обзор общих публикаций по теории вероятностей и статистике был бы длинным и неудобным для пользования. Вместо него я перечисляю особо интересные, по моему мнению, статьи, учебники и монографии.

Вот такие статьи или сборники статей: Bartlett (1976), Kendall (1966, 1972), Kruskal (1973a, 1974). Также два тома интересных перепечаток Steger (1971), Lieberman (1971).

Tanur и др. (1972): важное и необычное введение в статистику. Другое подобное сочинение это *Statistics* (1973).

Из учебников я обращаю особое внимание читателей на Cox & Hinkley (1974), Hodges & Lehman [Lehmann?] (1965), Mosteller & Tukey (1977), Tukey (1970). В своей статье мне следовало бы сослаться на

Morgenstern (1950) как на стимулирующее исследование структуры ошибок в экономике. Важно также упомянуть стандартное энциклопедическое сочинение Kendall & Stuart (1958 – 1966).

За последние годы библиографическая работа сильно продвинулась. Частичный список более важных библиографических публикаций см. ниже.

Примечания

1. Автор не включил ни сбора, ни предварительного исследования данных, и тем самым определил математическую, а не теоретическую статистику.

2. ... *целесообразно полагать постоянным?* Этого следует добиваться (если возможно).

3. В русскоязычной литературе принят менее подходящий термин *случайная величина*, хотя впервые применялось выражение *случайная переменная* (Шейнин 2013, гл. 14, Прим. 5). Мы не можем сказать, почему произошло неудачное изменение термина.

4. О возникновении табличной статистики см. Шейнин (2013, § 7.2.1), а о *первоначальной статистической деятельности* см. нашу статью (2014).

5. Теорию ошибок мы считали бы приложением статистического метода к обработке измерений.

6. Числа в фигурных скобках соответствуют тем же числам в *Добавлении* автора.

7. Неискушённые избиратели вполне могут основывать свои предпочтения на внешнем виде кандидатов.

8. Парапсихология продолжает исследоваться в некоторых странах, хоть очень многие считают её псевдонаукой.

9. Утверждение автора неудачно, поскольку оно относится только к одной группе из четырёх.

10. Автор имеет в виду Энциклопедию 1978 г., которую мы обозначаем сокращением IES.

11. Возможные позднейшие наблюдения могут отличаться от произведенных, что и объясняет суть ошибок, зависящих от выборки, а по поводу происхождения прочих ошибок см. сразу же ниже.

12. Поверхность отклика (*response surface*) специально рассматривается в *Энциклопедии*, но недостаточно ясно. Впрочем, для статьи автора этот термин вряд ли нужен.

13. Операционализм: философская концепция перестройки языка науки, оказавшая значительное влияние на методологию физики и психологии. См. БСЭ.

14. Ипохондрик: лицо с преувеличенным вниманием к своему здоровью и необоснованными страхами по его поводу.

15. См. Прим. 1.

16. Вот более известное высказывание Эйнштейна 1944 г. (Born 1969, с. 204): *Вы верите в Бога, играющего в кости, а я – в совершенные законы и порядок в объективно существующем мире ...*

17. Это утверждение трудно понять.

18. Пример: человек (данного пола и возраста) среднего веса не будет иметь среднего роста, но критика автора слаба: Кетле в громадной степени переоценил значимость своего среднего человека (Sheynin 1986).

19. Автор вначале упоминает *систематическое изучение государства*, и поэтому *свободную манеру* можно понимать как словесную форму изложения.

20. Статистическую географию автор упоминает с узкой точки зрения. Её начало положил Spottiswoode (1861), см. Sheynin (1986, с. 314 Прим.). Он рассуждал об установлении *истинного типичного* направления горной цепи по направлениям её хребтов. Вообще же существует в некоторой мере статистическая проблема генерализации изображения (и рельефа) на географических и топографических картах.

21. Сталину приписывают фразу *Умер человек – трагедия, умер миллион – только статистика*. Великий был любитель статистики!

Библиография Общая литература

Boehm G. A. W. (1964), The science of being almost certain. *Fortune*, vol. 69, No. 2, pp.104 – 107,142, 144, 146, 148.

Кас М. (1964), Probability. *Scient. American*, vol. 211, март, pp. 92 – 108.

Kendall M. G. (1950), The statistical approach. *Economica*, vol. 17, pp. 127 – 145.

Kruskal W. H. (1965), Statistics, Molière, and Henry Adams. *Amer. Scientist*, vol. 55, 1967, pp. 416 – 428.

Weaver W. (1952), Statistics. *Scient. American*, vol. 186, янв., pp. 60 – 63.

Сочинения, на которые даны ссылки

Пирсон К. (1892 англ.), *Грамматика науки*. СПб, 1911.

Шейнин О. Б., Sheynin O. (1986), Quetelet as a statistician. *Arch. Hist. Ex. Sci.*, vol. 36, pp. 281 – 325.

--- (2003), Lies, damned lies and statistics. *Intern. Z. f. Geschichte u. Ethik d. Naturwissenschaften, Technik u. Medizin*, Bd. 11, pp. 191 – 193.

--- (2013), *Теория вероятностей. Исторический очерк*. Берлин. Также Google, Oscar Sheynin, Download area.

--- (2014), К истории государственного управления. *Финансы и бизнес*, № 1, с. 136 – 158. Также Google, Oscar Sheynin, Download area.

Adams E. W., Fagot R. F., Robinson R. E. (1965), Theory of appropriate statistics. *Psychometrika*, vol. 30, pp. 99 – 127.

Auden W. H. (1946), Under which Lyre. *Coll. Poems*.

Aurthur R. A. (1953), *Glorification of Al Toolum*. New York.

Born M. (1949), *Natural Philosophy of Cause and Chance*. Oxford, 1951.

---, **Commentator** (1969), *Briefwechsel 1916 – 1955*. München.

Bortkiewicz L. von (1909), Die statistische Generalisationen. *Scientia*, t. 5, pp. 102 – 121.

Braithwaite R. B. (1953), *Scientific Explanation*. Cambridge. В мягкой обложке издано в 1960 г.

Bridgman P. W. (1959), *The Way Things Are*. Cambridge, Mass.

Churchman C. W., Ratoosh P., редакторы (1959), *Measurement: Definitions and Theories*. New York.

Cochran W. G., Mosteller F., Tukey J. W. (1954), *Statistical Problems [...]* *Sexual Behavior in the Human Male*. Washington.

Coombs C. H. (1964), *Theory of Data*. New York.

Cox D. R. (1961), *Role of Statistical Methods in Science and Technology*. London.

Deming W. E. (1965), Principles of professional statistical practice. *Annals Math. Statistics*, vol. 36, pp. 1883 – 1900.

Eldridge H. T. (1968), Population. Population policies. IESS, vol. 12, pp. 381 – 388.

Galton F. (1889), *Natural Inheritance*. London – New York.

Gini C. (1951), Caractère des plus récents développements de la méthodologie statistique. *Statistica*, vol. 11, pp. 3 – 11.

--- (1959), Mathematics in statistics. *Metron*, vol. 19, No. 3/4, pp. 1 – 9.

Grauman J. V. (1968), Population growth. IESS, vol. 12, pp. 376 – 381.

Hogben L. T. (1957), *Statistical Theory*. London.

Huff D. (1954), *How To Lie with Statistics*. New York.

Kadushin Ch. (1966), Shakespeare and sociology. *Columbia Univ. Forum*, vol. 9, No. 2, pp. 25 – 31.

Kish L. (1959), Some statistical problems in research design. *Amer. Sociological Rev.*, vol. 24, pp. 328 – 338.

Krutch J. W. (1963), Through happiness with slide rule and callipers. *Saturday Rev.*, vol. 46, No. 44, pp. 12 – 15.

Lazarsfeld P. F. (1961), Notes on the history of quantification in sociology. *Isis*, vol. 52, pt 2, pp. 277 – 333.

Liu Ta-Chung (1968), Economic data: mainland China. IESS, vol. 4, pp. 373 – 376.

Lundberg G. A. (1940), Statistics in modern social thought. В книге Н. Е. Barnes и др., редакторы. *Contemporary Social Theory*. New York, pp. 110 – 140.

Maron M.E. (1968), Cybernetics. IESS, vol. 4, pp. 3 – 6.

Mayer K. B. (1968), Population composition. IESS, vol. 12, pp. 362 – 370.

Moriyama I. M. (1968), Mortality. IESS, vol. 10, pp. 498 – 504.

Pfanzagl J. (1959 нем.), *Theory of Measurement*. New York, 1971.

Pollack I. (1968), Information theory. IESS, vol. 7, pp. 331 – 337.

Popper K. R. (1935 нем.), *Logic of Scientific Discovery*. New York, 1959. Издание в мягкой обложке 1961.

Ruggles R. (1968), Economic data, general. IESS, vol. 4, pp. 365 – 369.

- Sauvy A.** (1968), Population theories. IESS, vol. 12, pp. 349 – 358.
- Schmeidler Gertrude R.** (1968), Parapsychology. IESS, vol. 11, pp. 386 – 399.
- Selvin H. C.** (1957), Critique of tests of significance in survey research. *Amer. Sociological Rev.*, vol. 22, pp. 519 – 527. Обсуждение: там же, vol. 23, pp. 85 – 86, 199 – 200.
- Selvin H. C., Stuart A.** (1966), Data-dredging procedures in survey analysis. *Amer. Statistician*, vol. 20, No. 3, pp. 20 – 23.
- Skinner B. F.** (1956), A case history in scientific method. *Amer. Psychologist*, vol. 11, pp. 221 – 233.
- Spenser Brown G.** (1957), *Probability and Scientific Inference*. London.
- Spengler J. J.** (1968), Optimum population theory. IESS, vol. 12, pp. 358 – 362.
- Spottiswoode W.** (1861), On typical mountain ranges. *J. Roy. Geogr. Soc.*, vol. 31, pp. 149 – 154.
- Spulber N.** (1968), Economic data. The Soviet Union and Eastern Europe. IESS, vol. 4, pp. 370 – 372.
- Stagner R.** (1968), Homeostasis. IESS, vol. 6, pp. 499 – 503.
- Stevens S. S.** (1946). On the theory of scales of measurement. *Science*, vol. 103, pp. 677 – 680.
- Suppes P., Zinnes J. L.** (1963), Basic measurement theory. В книге R. D. Luce и др., редакторы, *Handbook of Math. Psychology*, vol. 1. New York, pp. 1 – 76.
- Torgerson W. S.** (1958), *Theory and Methods of Scaling*. New York.
- Tukey J. W.** (1961), Statistical and quantitative methodology. В книге D. P. Ray, редактор, *Trends in Social Science*. New York, pp. 84 – 136.
- (1962), The future of data analysis. *Annals Math. Statistics*, vol. 33, pp. 1 – 67, 812.
- Walberg H. J.** (1966), When are statistics appropriate? *Science*, vol. 154, pp. 330 – 332. Дискуссия с ответом автора: там же, vol. 155, p. 953, vol. 156, p. 314.
- Walpole H.** (1778), [Letter] To [...] H. S. Conway. В книге автора *Letters*, vol. 10. Oxford, 1904, pp. 337 – 338.
- Westergaard H. L.** (1932), *Contributions to the History of Statistics*. The Hague, 1969.
- White C.** (1964), Unkind cuts at statisticians. *Amer. Statistician*, vol. 18, No. 5, pp. 15 – 17.
- Whitney V. H.** (1968), Population distribution. IESS, vol. 12, pp. 370 – 376.
- Woytinsky W. S.** (1954), Limits of mathematics in statistics. *Amer. Statistician*, vol. 8, No. 1, pp. 6 – 10, 18.

Дополнительная библиография

- Моргенштерн О.** (1950 англ.), *О точности экономико-статистических наблюдений*. М., 1968.
- Allport G. W.** (1942), *The Use of Personal Documents in Psychological Science*. New York.
- Bagley D.** (1968), *The Viveso Letter*. London – New York. Издание в мягкой обложке 1970.
- Bartlett M. S.** (1976), *Probability, Statistics and Time*. London – New York.
- Bickel P. J., Lehmann E. L.** (1975), Descriptive statistics for nonparametric models. *Annals of Statistics*, vol. 3, pp. 1038 – 1044, 1045 – 1069.
- Boruch R. F., Rieken H. W., редакторы** (1976), *Experimental Testing of Public Policy*. Boulder, Colo.
- Cox D. R., Hinkley D. V.** (1974), *Theoretical Statistics*. London – New York.
- Dukes W. F.** (1965), N = 1. *Psychological Bull.*, vol. 64, pp. 74 – 79.
- Eisenhart C.** (1963), Expression of the uncertainties of final results. *Science*, vol. 160, 1968, pp. 1201 – 1204.
- Fairley W. B., Mosteller F.** (1977), *Statistics and Public Policy*. Reading, Mass.
- Federal Statistics** (1971), *Federal Statistics. Report of the President's Commission*, vols 1 – 2. Washington.
- Gilman R.** (1975), Review of *Humboldt's Gift* by Saul Bellow. *New York Times Book Rev.*, Aug. 17, pp. 1 – 3.
- Hodges J. L. Jr., Lehmann E. L.** (1965), *Elements of Finite Probability*. San Francisco, 1970.

- Kendall M. G.** (1966), Statistical inference in the light of the theory of the electronic computer. *Rev. Intern. Stat. Inst.*, vol. 34, pp. 1 – 12.
- (1972), Measurement in the study of society. В книге W. A. Robson, редактор, *Man and the Social Sciences*. London – Beverly Hills, Calif., pp. 133 – 147.
- Kendall M. G., Stuart A.** (1958 – 1966), *Advanced Theory of Statistics*, vols. 1 – 3. London – New York, 1973 – 1977.
- König F.** (1974), The Future of Religion. *New York Times*, Dec. 21, p. 27, col. 2.
- Kruskal W. H.** (1973a), The Committee on national statistics. *Science*, vol. 180, pp. 1256 – 1258.
- (1973b), Babies and averages. В книге *Statistics* (1973, vol. 1, pp. 49 – 59).
- (1974), The ubiquity of statistics. *Amer. Statistician*, vol. 28, No. 1, pp. 3 – 6.
- Lieberman B.** (1971), *Contemporary Problems in Statistics*. New York.
- Mead Margaret** (1976), Towards a human science. *Science*, vol. 191, pp. 903 – 909.
- Mendelsohn R.** (1976), A numbers game anyone can win. *Chicago Daily News*, May 22 – 23, p. 18, col. 5.
- Mosteller F., Tukey J. W.** (1977), *Data Analysis and Regression*. Reading, Mass.
- Redfield R.** (1940), The folk society and culture. В книге L. Wirth, редактор, *Eleven Twenty-six: a Decade of Social Science Research*. Chicago, pp. 39 – 50.
- Riecken H. W., Boruch R. H., редакторы** (1974), *Social Experimentation*. New York.
- Statistics** (1973), *Statistics by Example*, vols 1 – 4. Редактор F. Mosteller. Reading, Mass.
- Steger J. A., редактор** (1971), *Readings in Statistics for the Behavioral Scientist*. New York.
- Tukey J. W.** (1970), *Exploratory Data Analysis*. Reading, Mass., 1977.
- Tanur Judith M. и др., редакторы** (1972), *Statistics. A Guide to the Unknown*. San Francisco.
- Wallin P.** (1941), Prediction of individual behavior from case studies. В книге P. Horst, редактор, *Prediction of Personal Adjustment*. New York, pp. 181 – 249.
- Wilkes P.** (1975), *Trying Out the Dream. A Year in the Life of an American Family*. Philadelphia.

Библиографические источники

- Dolby J. L., Tukey J. W.** (1973), *The Statistics Cum Index*. Los Altos, Calif.
- Gani J. M.** (1970), On coping with new information etc. *J. Roy. Stat. Soc.*, vol. A133, pp. 442 – 450.
- (1972), Some comments on the bibliography etc. *Intern. Stat. Rev.*, vol. 40, pp. 201 – 207.
- Kendall M. G., Doig Alison G.** (1962 – 1968), *Bibliography of Statistical Literature*, vols 1 – 3. Edinburgh – New York.
- Lancaster H. O.** (1968), *Bibliography of Statistical Bibliographies*. Edinburgh.
- (1970), Problems in the bibliography of statistics. *J. Roy. Stat. Soc.*, vol. A133, pp. 409 – 441.
- Moran P. A. P.** (1974), How to find out in statistical and probability theory. *Intern. Stat. Rev.*, vol. 42, pp. 299 – 303.
- Ross I. C., Tukey J. W.** (1974), *Index to Statistics and Probability*. Los Altos, Calif.
- (1975), *Index to Statistics and Probability. Permuted Titles*, vols 1 – 2. Los Altos, Calif.
- Science** (1961 –), *Science Citation Index*. Ежегодники. Philadelphia.
- Social Science** (1972 –), *Social Science Citation Index*. Ежегодники. Philadelphia.
- Tukey J. W. и др.** (1973), *Index to Statistics and Probability. The Citation Index*. Los Altos, Calif.