Bayes, Markov, Romanovsky

I am presenting three independent papers, the first two of them in translation from their published versions having secured permission to publish them in Russian by the appropriate Editors. Both will eventually be available in *Istoriko-Matematicheskie Issledovania*. The third, concerning Romanovsky, is to appear in a special issue of *Archives Internationales d'histoire des sciences* commemorating Adolph Youshkevich, hopefully in 2007.

К ИСТОРИИ ТЕОРЕМЫ БЕЙЕСА

О.Б.Шейнин

1. Предисловие

В 1764 – 1765 гг. Королевское общество опубликовало мемуар покойного Томаса Бейеса (Байеса; 1702 – 1761) [1] с дополнениями Р. Прайса[1]. Мы рассматриваем недавно возникшее сомнение в авторстве "теоремы Бейеса", но, в соответствии с нашим убеждением и просто для удобства письма, мы придерживаемся этого термина. Начнем с двух предварительных замечаний.

1) Многие авторы, см., например, [5], полагают, что Бейес был учеником Муавра. Это, однако, не доказано и в любом случае обучение могло было быть только частным, поскольку последний никогда не занимал никаких должностей.

2) Почему сам Бейес, видимо, не старался опубликовать свой мемуар? Будучи тонким математиком, он не мог не понимать, что его труд был не совсем совершенен (но тем не менее исключительно важен)[2]. Так, оценка интеграла в формуле (1), см. п. 2, была недостаточно хороша, но тот же упрек можно было бы адресовать и Лапласу. У нас нет ответа на этот вопрос (впрочем, см. конец п. 6) и мы лишь заметим, что и Муавр почему-то лет на 12 задержал свое сообщение 1733 г. о "теореме Муавра – Лапласа" и еще лет на пять (теперь уже возможно вынужденно) – соответствующую полноценную публикацию[3].

Дискуссия о теореме Бейеса началась после того, как С.М. Стиглер [10] со ссылкой на Б. Зингера [11, с. 6] указал на примечательное высказывание Д. Хартли и истолковал его как свидетельство иного авторства теоремы Бейеса. Вот оно:

" ... Муавр показал, что если причины появления события находятся в постоянном соотношении к причинам его непоявления и если количество испытаний достаточно, то [сами] появления должны находиться почти в том же соотношении с непоявлениями; и что последнее соотношение неопределенно [неограниченно] стремится к первому при возрастании количества испытаний ... Изобретательный друг [или: собрат (friend)] ознакомил меня с решением обратной задачи, в которой он определил ожидание того, что первоначальное соотношение причин появления или непоявления события должно уклоняться [уклоняется] на любую заданную степень (degree) от p/q если известно, что это событие произошло p раз и не произошло q раз. И из его решения представляется, что при очень большом числе испытаний это уклонение должно быть незначительным" [12, с. 338 – 339].

В 1986 г. Стиглер [13, с. 98 и 132] сослался и на Хартли, и на свою статью, но свой прежний вывод назвал лишь одним из возможных заключений. Но затем он [14] перепечатал [10] с крохотным примечанием (с. 301), в котором отмахнулся от последовавшей критики этой статьи[4].

2. Теорема Бейеса

1) Впервые, в 1830 г., это выражение применили Дж.У. Луббок и Дж.Э. Дринкуотер-Бетун [6, с. 141], понимая под ним выражение

$$P(b \le r \le c) = \int_b^c u^p(1-u)^q\,du \div \int_0^1 u^p(1-u)^q du. \qquad (1)$$

Вот соответствующее рассуждение Бейеса. Мяч падает на единичный квадрат *ABCD* либо левее, либо правее некоторой прямой *MN*, параллельной *AB* и *CD* и расположенной между ними. Все возможные положения как *MN*, так и точки падения мяча *r* относительно *AB* и *CD* равновероятны. Формула (1) предполагает, что $[b; c]$ – отрезок внутри *AD* и что после $(p + q)$ падений мяча точка *r* оказалась *p* раз правее и *q* раз левее *MN*.

Бейес без труда получил знаменатель правой части формулы (1) и приложил большие усилия для оценки ее числителя. Вся правая часть (1) является разностью значений неполной бета фукции

$$I_c(p + 1; q + 1) - I_b(p + 1; q + 1).$$

Бейес не исследовал случая $n = (p + q) \to \infty$, а Прайс, в своем сопроводительном письме к мемуару, указал, что предельная теорема “недостаточно точна” для конечного *n*. Впрочем, применив остроумный прием, Г.Т. Тимердинг, редактор немецкого издания Бейеса [19], исследовал и предельный случай, хотя и не раскрыл вероятностный смысл входящих в полученную формулу параметров. Сделать это, однако, несложно и оказывается, что по существу предельный результат Бейеса таков:

$$\lim P\left\{\frac{|\overline{p}-a|}{\sqrt{pq/n^{3/2}}} \le z\right\} = \frac{1}{\sqrt{2\pi}}\int_0^z \exp(-w^2/2)\,dw,\ n \to \infty, \qquad (2)$$

$$a = p/n = \mathrm{E}\,\overline{p},\ \sqrt{pq/n^{3/2}} = \mathrm{D}\,\overline{p},$$

$\overline{p}$ – статистическая оценка неизвестной вероятности попадания точки *r* правее *MN*. Подчеркнем, что эта формула недостаточно известна.

2) Под теоремой Бейеса, однако, обычно понимается соотношение (принятые условия и обозначения общеизвестны)

$$P(A_i/B) = \frac{P(B/A_i)\,P(A_i)}{\sum_{j=1}^{n} P(B/A_j)\,P(A_j)}. \qquad (3)$$

Сам Бейес рассматривал лишь частный случай $n = 1$ (известный Муавру [8, 1756, с. 7]), хотя переход к общему случаю вряд ли затруднил бы его. Термин *теорема Бейеса* в смысле (3) восходит к Ог. Курно [20], который, впрочем, вначале (в п. 88) употребил его с оговоркой.

Итак, выражение *теорема Бейеса* исторически недостаточно определенно. Для нас его основным результатом являются формулы (1) и (2). Современное изложение его мемуара, хотя и без материала нашего п. 5, см. [6, п. 8]. По поводу наших пп. 2 и 3 см. наши [21] – [23].

3. Результаты Якоба Бернулли, А. Муавра и Бейеса

3.1. Якоб Бернулли впервые установил соответствие между теоретической (p) и статистической ($\overline{p}$) вероятностями, доказав, что с возрастанием количества испытаний “по схеме Бернулли” $\overline{p} \to p$. Не приводя точного выражения его класического закона больших чисел, заметим, что он определял статистическую вероятность события по его теоретической вероятности и это яснее всего усматривается в формулировке его *Главного предложения* (закона больших чисел [24, с. 56]). Однако, и в самом конце своего сочинения (с. 59), и в гл. 4 Бернулли упомянул обратную задачу и утверждал, что решил и ее также. Вот его характерное, но лишь качественное высказывание:

“... если кто-либо будет рассматривать состояние погоды за очень большое число истекших лет ... или ... наблюдать, сколько раз тот или иной [из двух игроков] оказывается ... победителем, то тем самым откроет соотношение, в котором вероятно находятся числа случаев ...” [24, с. 42].

И. Тодхантер [25, с. 73], и некоторые позднейшие авторы заявили, что Бернулли действительно рассматривал такое обращение, но иные комментаторы [26, с. 351; 27, с. 6 – 7] разумно отрицали это.

3.2. Муавр изучал вероятностное поведение частости (относительного числа) появления события в n независимых испытаниях при заданной вероятности появления события в единичном испытании, – т.е. рассматривал ту же схему Бернулли. Его предельная теорема (теорема Муавра – Лапласа) имеет вид [3, с. 128 – 130][5]

$$\lim P\left(a \le \frac{\mu - np}{\sqrt{npq}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b \exp(-z^2/2)dz,\ n \to \infty,$$

(4)

$q = 1 - p,\ np = \mathrm{E}\mu,\ npq = \mathrm{D}\mu,$

Лаплас же усовершенствовал вывод этой формулы и дополнительно исследовал случай большого конечного n. Ни Муавр, ни Лаплас не владели ни понятием равномерной сходимости, которая имеет здесь место, ни понятем дисперсии[6] и не отличали строгих неравенств от нестрогих.

Как и Якоб Бернулли, Муавр думал и об обращении своей теоремы. В мемуаре [7], который содержит его основной результат, он (только в издании 1756 г.) сформулировал *Замечание 2*[7] с необоснованным указанием на обращенную задачу:

"... обратно, если из бесчисленных наблюдений мы находим, что отношение событий [относительное число появлений события] стремится к детерминированной [определенной] величине, ... то мы заключаем, что это отношение выражает детерминированный закон, в соответствии с которым это событие должно происходить"[8] [8, 1756, с. 251].

Но когда Муавр закончил подготовку этого, уже посмертного издания? Там, на с. XI, помещено следующее анонимное "Объявление":

"В своем преклонном возрасте автор этой книги был вынужден ввиду потери зрения поручить выпуск настоящего издания одному из своих друзей [или: собратьев (friends)], которому он дал экземпляр книги [предыдущего издания] с некоторыми ... поправками и дополнениями ... К ним редактор добавил несколько других ... и расположил все в большем порядке ... Все это [было сделано] в соответствии с планом, согласованным с автором более чем за год до его смерти [в 1754 г.]".

Мы, однако, не знаем, когда Муавр согласовал окончательный вариант своей книги с редактором. В последние годы своей жизни он не только потерял (хотя возможно неполностью) зрение, но и вообще отошел от научной работы. И тем не менее нельзя исключить, что Муавр либо прочел примечательное высказывание Хартли, либо узнал о нем от кого-то (и имел бы тогда субъективный повод упомянуть обращенную задачу).

3.3. Муавр и Бейес: сравнение формул (2) и (4). Дроби в левых частях этих формул можно записать одинаковым образом в виде

$$(\xi - E\xi) \div \sqrt{D\xi}\,,$$

хотя, разумеется, случайная величина ξ имеет в них различный смысл. Примечательно, что Бейес, который также (п. 3.2) не владел понятием дисперсии, видимо понимал, что формула (4) плохо соответствует его обращенной задаче. В этом отношении интересно утверждение Прайса, отрицающее цитированное в п. 3.2 *Замечание 2* Муавра:

"Я не знаю никого, кто показал бы как решить обратную задачу, т.е. как по заданным количествам появления и непоявления неизвестного [неизученного] события определить шанс того, что вероятность его появления должна находиться где-то между любыми двумя назначенными степенями вероятности"[9] [2, с. 135].

4. Николас Сондерсон

На основании различных источников Стиглер [10] заключил, что

1) Хартли в основном закончил свою книгу в 1739 г.

2) Он был хорошо знаком с Сондерсоном (1682 – 1739), слепым математиком штрокого кругозора, автором *Начал алгебры*, см. [28, с. 52].

3) И Р. Котс, и Муавр были высокого мнения о Сондерсоне, который притом ознакомился с мемуаром Муавра [7][10].

4) О возможном знакомстве Бейеса с Муавром или Хартли ничего не известно.

Соответственно, Стиглер решил, что теоремой Бейеса, а точнее формулой (2), владел Сондерсон. Применяя формулу (3), но никак не (2), он даже вычислил соотношение авторства *теоремы*, которое оказалось равным 3:1 в пользу его кандидата. Впрочем, принятое Стиглером равенство априорных вероятностей обеих гипотез, да и назначенные им апостериорные вероятности сомнительны[11].

Стиглер, правда, не отрицает того, что автором *теоремы* мог быть и Бейес, если только тот прочел книгу Хартли и воспользовался высказанной там мыслью. Но в таком случае Бейес мог бы сослаться на Хартли [12], да и утверждение Прайса (п. 3.3) противоречит этому предположению и, наконец, вряд ли у Бейеса был предшественник.

5. Последовавшая дискуссия

1) А.У.Ф. Эдвардс [30] полагает, что Бейес действительно был знаком с высказыванием Хартли и что последний имел в виду Муавра. Второе соображение вряд ли верно, поскольку Хартли упоминает и Муавра, и кого-то иного. К тому же Муавр (п. 3.2) лишь объявлял, что решил обратную задачу, да и сравнение формул (4) и (2) свидетельствует против него.

2) Д.А. Гиллис [31, с. 329] считает существенным, что после публикации мемуара Бейеса никто не заявил о потере своего приоритета. Впрочем, к тому времени ни Сондерсона, ни Хартли не было уже в живых. Гиллис также заметил, что Бейес мог получить окончательный стимул к составлению (или к окончанию?) своего мемуара в 1748 г. от сочинения Д. Юма [32], в котором тот рассуждал о методе индукции. Бейес, допускает Гиллис, сообщил о своих результатах Хартли, тот же успел в последнюю минуту соответственно дополнить свою рукопись.

3) Э. Дейл [26, с. 358] полагает, что сообщение в книге Хартли вполне могло повлиять на Бейеса. Он, кажется, так же как Эдвардс считает, что Хартли имел в виду Муавра.

4) Он же [33] описал рукопись и записную книжку, принадлежавшие, судя по почерку, одному и тому же лицу (с. 313) и привел (там же, с. 322) доводы в пользу того, что этим лицом был Бейес. Наконец, он (с. 322) заметил, что эта записная книжка содержала доказательство *Правила 2* из мемуара Бейеса. Для нас важно только то, что, к сожалению, дата этого доказательства не указана.

5) А. Хальд [34, с. 132; 6, с. 400]. В первом сочинении автор лишь сообщил о дискуссии, но собственного мнения не высказал. Во втором случае он снова сослался на предшествовавших авторов, и, не приводя новых доводов, заявил, что Хартли описал результат Бейеса.

6) Э. Дейл [27, с. 8] четко повторил (см. выше), что Хартли имел в виду Муавра.

6. Заключение

Бейес понимал, что формула (4) не оценивала должным образом изучаемое событие по большому числу бернуллиевых испытаний и он вряд ли думал об окончательном установлении формулы (3), – которую в этом (и только в этом) смысле не следовало бы называть его именем.

Для Бейеса значение формулы (2)[12] могло быть очевидным, потому что, взятая совместно с известными в то время элементами учения о случае (и, разумеется, совместно с формулой (4)) означала завершение его тогдашнего вида. А вот формула (3) и тем менее ее частный случай (п. 2) ничего подобного не обеспечивала.

Быть может существенно, что Бейес фактически покинул свою должность священнослужителя в то же время (в 1749 г.), когда вышла в свет книга Хартли [12]. Будучи чересчур скромным (прим. 2), он возможно решил тогда сообщить свою формулу (2) в частном порядке Хартли, тот же, как предположил Гиллис (п. 5), успел вписать несколько соответствующих строк в свою еще не опубликованную рукопись. В любом случае существование какого-то предшественника Бейеса (мы снова имеем в виду формулу (2)) маловероятно. К тому же ее доказательство содержится (косвенно) только в мемуаре Бейеса.

Признательность. Данная статья появилась на английском языке [35] и мы благодарны редактору журнала Mathematical scientist, Prof. J. Gani, за разрешение на публикацию ее русского варианта (с незначительными изменениями).

Примечания

1. Сжатые биографические сведения об этих ученых и краткое описание мемуара Бейеса см. [3, с. 137 – 139]. Прайс "представил" мемуар к рубликации, но, видимо, лишь в смысле *передал*; он стал членом Королевского общества только в 1765 г., Бейес же был академиком с 1742 г.

2. С его именем связано много современных терминов, например, *бейесовская теорема*; *оценка*; и *бейесовский подход*; *принцип*. И в то же время он, как предположил А.Хальд [6, с. 135], был слишком скромен.

3. Муавр упоминает число 12 уже в сообщении 1733 г. [7].

4. Там же Стиглер [14, с. 52] назвал неких современных авторов учеными первого класса. Опубликовав рецензии на их сочинения [15], [16], мы назовем Т.М. Портера поверхностным автором (а его новую книгу [17] – рыхлой, содержащей грубые ошибки и недостаточно полной), а второго – невеждой, утверждение же Стиглера мы считаем недопустимым. Напомним [18, с. 324], что мы резко осудили его за надругательство над памятью Эйлера и Гаусса.

5. И. Тодхантер [25, с. 192 – 193] поверхностно описал результат Муавра. В частности, он указал, что тот рассматривал лишь случай $p = 1/2$ и эта ошибка до сих пор не изжита. В дополнение к сказанному по этому поводу в [3], заметим, что в заглавии мемуара Муавра [7] упоминается бином $(a + b)^n$, т.е., в современных обозначениях, $(p + q)^n$, а вовсе не $(1/2 + 1/2)^n$.

6. Лаплас вывел формулу (4) в 1812 г., дисперсию же Гаусс определил в 1823 г.

7. Его уже цитировал Э. Дейл [26, с. 352], который, однако, напрасно ссылался и на другие места из мемуара [7].

8. Биномиальное распределение, которое здесь подразумевается, стало у Муавра детерминированным законом; лишь эмпирические уклонения от него оказывались случайными. И далее Муавр философски рассуждал о закономерностях в природе именно исходя из указанного понимания случайного и необходимого.

9. В самом начале своего мемуара Бейес отметил, что не отличает *шанс* от *вероятности*.

10. Добавим из мемуара Муавра [7]:

“... то, что я сделал, было в соответствии с побуждением весьма почтенного джентльмена и хорошего математика [не Сондерсона ли?] ... теперь я добавляю некоторые новые мысли к прежнему” [8, 1756, с. 243] (по контексту, добавляет, очевидно к [29]).

11. Например: поскольку Сондерсон был знаком с Хартли, Бейес же не был связан с последним, соотношение апостериорных вероятностей в пользу Сондерсона оказывается равным 3/4 к 1/4 (?). Учет трех подобных обстоятельств (случайно) привел к тому же окончательному соотношению, а поскольку априорные вероятности совпадают, оно уже не изменяется.

12. Здесь и ниже мы имеем в виду бейесовский вариант этой формулы.

Список литературы

1. *Bayes T.* An essay towards the solving a problem in the doctrine of chances // Philosophical Transactions Royal Society. 1763 (1764). Vol. 53. P. 360 – 418; 1764 (1765). Vol. 54. P. 296 – 325. Первая часть мемуара: Biometrika. 1958. Vol. 45. P. 296 – 315; [2, pp. 134 – 153].
2. Studies in the history of statistics and probability. Edited by E.S. Pearson and M.G. Kendall. London. 1970.
3. Шейнин О.Б., Майстров Л.Е. Теория вероятностей [4, с. 126 – 152].
4. История математики с древнейших времен до начала XIX столетия, под редакцией А.П. Юшкевича. М., 1972. Т. 3.
5. Edwards A.W.F. Bayes. Dictionary of national biography. Missing persons. Oxford. 1993. P. 50 – 51.
6. Hald A. History of mathematical statistics from 1750 to 1930. New York. 1998.
7. De Moivre A. A method of approximating the sum of the terms of the binomial $(a + b)^n$ expanded into a series. Первоначально распространено автором в 1733 г. на латинском языке. Авторский перевод [8, 1738 и 1756], с. 243 – 254 в 1756 г. Латинский текст сохранился в нескольких экземплярах [9].
8. De Moivre A. Doctrine of chances. London. 1718. 1738. 1756. Оба последние издания перепечатаны в 1967 г., в Лондоне и Нью Йорке соответственно.
9. Daw R.H., Pearson E.S. A. De Moivre’s 1733 derivation of the normal curve: a bibliographic note // Biometrika. 1972. Vol. 59. P. 677 – 680.
10. Stigler S.M. Who discovered Bayes’s theorem? // American Statistician. 1983. Vol. 37. P. 290 – 296.
11. Singer B. Distribution-free methods for non-parametric problems: a classified and selected bibliography // British Journal of Mathematical and Statistical Psychology. 1979. Vol. 32. P. 1 – 60.
12. Hartley D. Observations on man, his frame, his duty and his expectations. London. 1749.
13. Stigler S.M. History of statistics. Cambridge, Mass. – London. 1986.
14. Stigler S.M. Statistics on the table: history of statistical concepts and methods. Cambridge, Mass. 1999.
15. Sheynin O.B. Рецензия на книгу Porter T.M. The rise of statistical thinking. Princeton. 1986. // Centaurus. 1988. T. 31. P. 171 – 172.
16. Sheynin O.B. Рецензия на книгу Desrosières A. The politics of large numbers. Cambridge, Mass. – London. 1998. // Isis. 2001. Vol. 92. P. 184 – 185.

17. Porter T.M. Karl Pearson. Princeton – Oxford. 2004.
18. Шейнин О.Б. История теории ошибок // Историко-математические исследования. М., 2000. Вторая серия. Выпуск 5 (40). С. 310 – 332.
19. Bayes T. Versuch zur Lösung eines Problem der Wahrscheinlichkeitsrechnung. Herausgeber H.T. Timerding. Leipzig. 1908.
20. Курно Ог. Основы теории шансов и вероятностей. М., 1970. Первоначально опубликовано на франц. языке (Париж, 1843).
21. Sheynin O.B. On the early history of the law of large numbers // Biometrika. 1968. Vol. 55. P. 459 – 467. Перепечатка [2, pp. 231 – 239].
22. Шейнин О.Б. К истории предельных теорем Муавра – Лапласа // История и методология естественных наук. М. 1970. Т. 9. С. 199 – 211.
23. Sheynin O.B. On the history of some statistical laws of distribution // Biometrika. 1971. Vol. 58. P. 234 – 236.
24. Бернулли Я. О законе больших чисел. Под редакцией Ю.В. Прохорова. М. 1986. Содержит часть 4-ю *Искусства предположений* (1713) автора в переводе Я.В. Успенского (1913), примечания и комментарии и перепечатку статьи А.А. Маркова Двухсотлетие закона больших чисел (1914).
25. Todhunter I. History of the mathematical theory of probability (1865). Перепечатка: New York. 1965.
26. Dale A.I. On Bayes' theorem and the inverse Bernoulli theorem. // Historia Mathematica. 1988. Vol. 15. P. 348 – 360.
27. Dale A.I. History of inverse probability (1991). Второе издание: New York, 1999.
28. Башмакова И.Г., Розенфельд Б.А., Юшкевич А.П. Арифметика и алгебра [4, с. 32 – 100].
29. De Moivre A. Miscellanea analytica de seriebus et quadraturis. London. 1730.
30. Edwards A.W.F. Is the reference in Hartley (1749) to Bayesian inference? // American Statistician. 1986. Vol. 40. P. 109 – 110.
31. Gillies D.A. Was Bayes a Bayesian? // Historia Mathematica. 1987. Vol. 14. P. 325 – 346.
32. Hume D. An enquiry concerning human understanding. Oxford. 2000. Первое издание, 1748. Русский перевод: Юм Д. Исследование о человеческом разумении. М. 1994.
33. Dale A.I. Bayes' work on infinite series // Historia Mathematica. 1991. Vol. 18. P. 312 – 327.
34. Hald A. History of probability and statistics and their applications before 1750. New York. 1990.
35. Sheynin O.B. On the history of Bayes's theorem. Mathematical scientist. 2003. Vol. 28. P. 37 – 42.

Курно Ог. (Cournot A.A.)
Лаббок Дж.У. (Lubbock J.W.)
Лаплас П.С. (Laplace P.S.)
Муавр А. (De Moivre A.)
Портер Т.М. (Porter T.M.)
Прайс Р. (Price R.)
Сондерсон Н. (Saunderson N.)
Стиглер С.М. (Stigler S.M.)
Тимердинг Г.Т. (Timerding H.T.)
Тодхантер И. (Todhunter I.)
Хальд А. (Hald A.)
Хартли Д. (Hartley D.)
Эдвардс А.У.Ф. (Edwards A.W.F.)
Эйлер Л. (Euler L.)
Юм Д. (Hume D.)

Sheynin O.B. On the history of the Bayes theorem
Doubts were recently expressed about the real author of the Bayes theorem. Here, the arguments put forward in favour of another mathematician as well as the ensuing discussion are reviewed. It is argued that the expression *Bayes theorem* is not historically definite and that, anyway, its author was Bayes who understood the essence of his main problem better than De Moivre did.

МАТЕМАТИЧЕСКАЯ ОБРАБОТКА НАБЛЮДЕНИЙ У А. А. МАРКОВА

О.Б.Шейнин

1. Введение

Мы обсуждали исследования Маркова в теории вероятностей, в том числе и математическую обработку наблюдений [1]. Здесь, однако, мы описываем последнюю тему гораздо подробнее, притом с учетом некоторых позднее изученных нами материалов, а именно рукописи Маркова 1903 г. [2] и писем, написанным ему его бывшим учеником, Б. М. Кояловичем [3], впоследствии ставшим видным ученым [4, с. 415]. Наша статья весьма критична, что никак не означает какого-то умаления заслуг Маркова в области теории вероятностей.

Первое (литографированное) издание курса лекций Маркова по теории вероятностей (включая обработку наблюдений) появилось в 1893 г. С тех пор он неизменно обсуждал эту тему не менее чем в четырех последующих литографированных курсах тех же лекций, во всех изданиях своего руководства [6], в мемуаре [7] и в упомянутой рукописи [2]. Мнения о заслугах Маркова в области метода наименьших квадратов (МНКв) были вначале противоречивы, ныне же, как представляется, общепринято считать, что он не достиг многого и мы согласны с этим.

2. Год 1899-й

Название мемуара Маркова [7] показывает, что он исследовал две темы и соотношение между ними.

1) Вот его общие соображения о МНКв: “Мы можем притти [к нему] исходя из следующих соображений”: систематические ошибки отсутствуют; веса “приближенных равенств обратно пропорциональны математическим ожиданиям квадратов погрешностей”; и

“достоинство каждого приближенного равенства мы оцениваем его весом и соответственно этому для каждого неизвестного отыскиваем такое приближенное равенство, вес которого наибольший. Только этот вывод способа наименьших квадратов я считаю рациональным; он указан Гауссом. Рациональным я считаю этот способ [этот вывод] главным образом потому, что он не заменяет [не затемняет?] условность [МНКв]. Мы не приписываем [ему] ... наивероятнейшие или наиблагонадежнейшие результаты, а рассматриваем его только как общий прием, дающий приближенные величины неизвестных с условной оценкой результатов” [7, с. 246].

Впоследствии Марков [6, 1924, с. 323] по существу повторил последние строки. Приближенными равенствами могут быть только n исходных линейных уравнений (возможно имеющих разные веса) с m неизвестными ($m < n$), из которых выводятся m нормальных уравнений, определяющих “приближенные величины неизвестных”. Первые строки приведенного высказывания остаются поэтому непонятными, притом ссылка на Гаусса отпадает. Они могут иметь смысл, пожалуй, лишь для случая одного неизвестного и вот, действительно, соответствующее утверждение Маркова: если получено несколько приближенных равенств $a - X' \approx 0$,

“где X' означает число, вполне определенное [?] результатами наблюдений, то мы будем выбирать из этих равенств, как наилучшее для определения числа a, равенство, вес которого наибольший” [6, 1924, с. 326].

Значения X', как можно полагать, выведены из отдельных рядов наблюдений. Если их веса установлены реально, и они не слишком различны, то зачем тогда отбрасывать почти все наблюдения? Если же веса существенно различны, то зачем вообще были нужны малоточные наблюдения? И, наконец, в дальнейшем изложении [6, 1924] последнее высказывание не используется. Упомянем в этой связи Ньюкома [9]: уравнивая (но не отбрасывая) данные по нескольким астрономическим каталогам, он назначил каждому из них два веса, по-отдельности для систематических и случайных ошибок.

Общую точку зрения Маркова можно дополнительно иллюстрировать письмом 1893 г., полученным им от Кояловича:

“Вы говорите, что если данные хороши, то результаты будут всегда хороши даже без способа наименьших квадратов. Совершенно согласен, но среди этих хороших результатов могут все-таки одни быть лучше других, смотря по тому, как мы будем комбинировать наблюдения” [3].

Возвращаемся к мемуару [7]. Почему оценка точности условна, а вычисленные значения неизвестных не являются ни вероятнейшими, ни благонадежнейшими? Отрицание первого свойства означало отказ от первого обоснования МНКв (1809 г.), от которого, впрочем, отошел и сам Гаусс. Второе свойство Гаусс в 1823 г. посчитал равносильным принципу наибольшего веса (наименьшей дисперсии), который Марков, отрицая оптимальные свойства результатов МНКв, однако, полностью одобрял (см. ниже).

Остановимся на этом заключении. Гаусс не доверял своим собственным формулам оценки точности. Производя геодезические наблюдения, он измерял углы триангуляции до тех пор, пока не убеждался, что дальнейшая работа бесполезна. Так, на одной станции он [10, с. 278 – 281] измерил один из углов 6 раз, а другой – 78 раз! С тех пор стало еще понятнее, что, главным образом ввиду наличия систематических ошибок, надежно оценить точность цепи триангуляции можно только после того, как станут известны "невязки" треугольников и несоответствия между базисами и между азимутами, измеренными на ее концах. Марков, однако, нигде не упоминал о влиянии систематических ошибок, он просто считал, что они исключены.

2) Повторим, что Марков поддержал второе обоснование МНКв [7, с. 247]. Он привел выдержку из известного письма Гаусса Бесселю 1839 г., в котором Гаусс разъяснил, что считает гораздо важнее достигнуть наименьшего значения интегральной меры точности, нежели добиться высшей, но все равно бесконечно низкой вероятности [т. е. вероятнейшего значения]. Марков добавил, что соответствующий вывод

"дает все, что нужно для практики; но он не дает нам вероятной ошибки. Я полагаю [Марков полагает], что эта сомнительная величина не нужна; если же ее необходимо иметь во что бы то ни стало, то следует допустить известное выражение для вероятности в виде [нормального; этот термин Марков так ни разу и не употребил] интеграла ... для каждой ошибки в отдельности, независимо от того, выведена ли она из одного или из нескольких наблюдений".

Снова непонятно: если считать, что соотношение между средней квадратической и вероятной ошибками известно ввиду нормальности распределения, то получается, что не нужна ни та, ни другая мера точности.

Позднее Марков снова без разъяснений заявил, что в теории корреляции и даже в теории ошибок он не придает

"большого значения так называемым вероятным погрешностям и счита[ет] их только средством для условного сравнения достоинства различных наблюдений" [11, с. 535].

Следуя примеру современных ему астрономов, Марков не упомянул средней квадратической ошибки, хотя Гаусс несомненно (и разумно) считал ее основной мерой точности.

3) Марков продолжал: нормальное распределение допускается "согласно с указанием практики", которое, однако, "трудно установить". Другое же обоснование состоит в том, что ошибку наблюдения рассматривают "как совокупность многих ошибок [одного и того же порядка], независящих друг от друга" и ссылаются на "теорему о пределе вероятности".

Марков, однако, отвергнул и возможность указанного представления ошибки, и приложимость центральной предельной теоремы, поскольку она “установлена только со многими ограничениями”, и, разумеется, неограниченность числа наблюдений.

4) В заключение Марков [7, с. 249 – 250] рассмотрел попытки обосновать МНКв, исходящие из закона больших чисел Чебышева.

a) Обозначим истинное значение измеряемой константы через a и среднее арифметическое из n ее наблюдений через $\bar{x}$. Тогда, как заметил Марков, из утверждения Маиевского [12, пп. 31 – 32] о том, что

$$\lim P[|\bar{x} - a| < \varepsilon] = 1,\ n \to \infty \qquad (2.1)$$

[т. е. из того, что $\bar{x}$ – состоятельная оценка a], “ничего не следует”, поскольку Маиевский не сравнил $\bar{x}$ с другими возможными средними, см. также пункт b) ниже. Марков также назвал (2.1) “известным предположением Гаусса”. Гаусс, однако, принципиально рассматривал только конечное число наблюдений (а для вывода принципа наименьших квадратов ему в 1809 г. понадобился и принцип наибольшего правдоподобия).

Первую попытку обоснования МНКв указанного вида сделал Усов [13], которого Марков не назвал. Заметим также, что формула (2.1) явилась следствием более общего предложения Чебышева, которое он непосредственно сформулировал только в своих лекциях [14, с. 165].

b) Наконец, Марков [7, с. 250] заметил, что “некоторые русские математики” пытались вывести МНКв “посредством неправильного применения теорем Чебышева”. Имея в виду неравенство Бьенеме – Чебышева, он продолжил:

“Они требовали от вероятности, чтобы она была больше заданного наперед числа, и при том же неравенстве для вероятности искали теснейшие пределы погрешности. При этом забывалось, что одному и тому же неравенству удовлетворяет бесконечное множество различных чисел и что вероятность может быть больше некоторого данного числа и в то время, когда неравенство Чебышева этого не обнаруживает”.

Неравенство Бьенеме – Чебышева имеет дело с конечным, а не бесконечным, как центральная предельная теорема, числом испытаний или наблюдений, однако и здесь в дополнение к сказанному Марковым требовался принцип наибольшего правдоподобия.

Наиболее интересным в мемуаре [7], поскольку он относился к МНКв, была решительная защита Марковым принципа наибольшего веса. Правда, несколько авторов [15, с. 7; 16, с. 205 и 210; 17] опередили его в этом, но известнее всех стало мнение Маркова (не только в России) ввиду особого авторитета его имени, что даже привело к переоценке его заслуг (п. 5.1). И в то же время его точка зрения непонятна: метод, результаты которого не обладают никакими оптимальными свойствами [7, с. 246; цитировано выше], вряд ли нуждается в каком-либо обосновании.

3. Год 1903-й

Мы имеем в виду развернутую рецензию Маркова [2], опубликованную лишь в 1990 г., см. нашу статью [19]. В 1902 г. Голицын, будущий президент Международной сейсмической ассоциации и член Королевского общества, опубликовал небрежно выполненное исследование [20] прочности стеклянных трубок, Марков же, скрупулезно проверив его работу, опровергнул содержащиеся в ней выводы. Не повторяя нашего описания его справедливой критики, заметим, что он оставался в рамках классической теории ошибок.

Как раз в то время [19] несколько академиков, включая астронома Ф. А. Бредихина, представили Голицына к избранию в действительные члены Петербургской академии наук, а Марков (которого поддержал Ляпунов, раскритиковавший некоторые механические исследования Голицына) успешно противопоставил им свое крайне отрицательное мнение о только что вышедшей статье.

Во время происшедших прений Марков заявил, что ему "очень нравится" правило Бредихина, в соответствии с которым

"Для признания реальности величины вычисленной требуется, чтобы она по крайней мере в два раза превосходила числом свою вероятную погрешность" [19, с. 453 – 454].

Подобное правило применялось в то время в естествознании: Менделеев [21, с. 46] и Ньюком [22, с. 165] независимо друг от друга указывали, что расхождение между двумя эмпирическими величинами существенно, если оно превышало сумму соответствующих вероятных ошибок. Для нормального распределения правило Бредихина, как Марков назвал его, видимо означало, что, поскольку

$$P(|x| < 1.33\sigma) = 0.816$$

(σ – дисперсия "вычисленной величины" x), вероятность противоположного события достаточно низка.

4. Год 1924-й: метод наименьших квадратов

В 1920/21 учебном году Марков

"напряженно работал над четвертым ... изданием своей книги [6, 1924]. Как известно, это издание значительно отличается от предыдущего" [23, с. 612 – 613].

Марков-младший не уточнил, успел ли его отец закончить эту работу. Мы теперь опишем главу 7-ю руководства [6, 1924, с. 323 – 473], названную *Способ наименьших квадратов*. В предшествующем издании 1913 г. соответствующая глава не включала ни исследования статистических рядов (наш п. 4.3), ни краткого обсуждения линеаризации уравнений (с. 469 – 472 в 1924 г.), а материал нашего п. 4.7 был в то время рассмотрен еще менее полно чем в 1924 г.

4.1. Предварительные соображения (с. 323 – 326)

Марков прежде всего указал, что остается при своем прежнем мнении [7] об обосновании МНКв (с. 323 прим.), что подтверждается дальнейшим изложением. На той же странице он "допусти[л] существование чисел, приближенные величины которых доставляются наблюдениями". Скажем: допустил существование истинных значений измеряемых констант. Никто кроме него такого замечания, кажется, не вводил, но оно имеет смысл: траектория визирного луча все-таки переменна, так что же считать величиной угла на местности? Вообще же уже Фурье [24, с. 534] определил "истинный объект исследования" как предел среднего арифметического из его наблюдений при неограниченном возрастании их числа. Это определение не было замечено, но с тех пор многие авторы включая Мизеса вводили его заново, независимо друг от друга [25, с. 100 – 101].

Далее, Марков (с. 323 и 373) рассматривал каждое наблюдение "как частный случай многих". В первом случае он уточнил, хоть и не вполне определенно: каждому наблюдению как бы соответствует одно воображаемое. Наконец, мы здесь напомним об утверждении Маркова по поводу принципа наибольшего веса, которое мы цитировали в п. 2.

4.2. Уравнивание наблюдений одного неизвестного (с. 327 – 344)

a) Возможные наблюдения (с. 327). Предположив, что имеются наблюдения a_i, $i = 1, 2, \ldots, n$, неизвестной константы, Марков указал, что наряду с ними он будет рассматривать возможные наблюдения u_i. Трудно сказать, почему, вопреки Чебышеву [14, с. 227], Марков (еще четче!) повторил свое прежнее соответствующее утверждение (см. п. 4.1). Более того, второе письмо 1893 г. Кояловича Маркову (первое мы цитировали в п. 2) означало, что в 1891 г., в своем литографированном курсе теории вероятностей, он, видимо, придерживался иного взгляда. Вот что писал Коялович:

"Насколько я Вас понял, Вы рассматриваете каждое отдельное наблюдение как одно из значений возможного результата. Таким образом, для каждого измерения возможен ряд результатов ..., один из которых осуществляется на деле. Все это я готов понять для одного измерения, но когда их имеется напр. два, то я не могу понять чем отличается ряд возможных результатов ... первого наблюдения от ряда возможных результатов ... второго измерения. Вопрос, конечно, сейчас же решается, если Вы скажете, что вероятность одной и той же ошибки в этих рядах различна, но ведь Вы вероятно не захотите вводить понятие о вероятности ошибки в Ваше изложение" [26].

Сейчас мы бы сочли наблюдения случайной выборкой из генеральной совокупности. По поводу последней фразы письма (полагая все же, что эта совокупность едина) заметим, что Марков [6, 1924, с. 251] ввел понятие о плотности распределения, но по существу не применил его в рассматриваемой главе. Оно появилось там лишь на с. 420 при обсуждении корреляции; впрочем, Марков обращался к нормальному распределению (см. ниже пункт d).

b) Собственно уравнивание. Марков предположил, что неизвестная величина a является линейной функцией независимых измерений a_i, свободных от постоянных погрешностей

$$a \approx \lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_n a_n, \quad (4.1)$$
$$\lambda_1 + \lambda_2 + \ldots + \lambda_n = 1. \quad (4.2)$$

Здесь и ниже символ приближенного равенства (у Маркова не $\approx$, а $\neq$) применялся им для того, чтобы отличить постоянные величины от их оценок.

Марков (с. 329) заметил, что другие средние (например, среднее геометрическое) не обеспечивает, как бы мы сказали, несмещенность, а на с. 447 указал, что МНКв ограничивается обработкой линейных систем, потому что неизвестными являются малые поправки к приближенно известным величинам, ср. наш п. 4.7. Здесь и ниже он обсуждал наблюдения неравного достоинства, но не пояснил, что от этого ограничения можно отказаться без потери общности.

Марков (с. 327) выбирал веса p_i так, чтобы

$$E(a - u_i)^2 = k/p_i, \tag{4.3}$$

где k – средняя квадратическая ошибка единицы веса (термин классической теории ошибок) и u_i – возможные наблюдения (см. пункт а); математическое ожидание Марков обозначал сокращением м.о.

Следовательно (с. 329 – 330),

$$E(a - [\lambda u])^2 = [\lambda\lambda/p] = P,$$

где мы ввели обозначение Гаусса вида $[abc] = \sum a_i b_i c_i$, а P – вес уравнения (4.1). Ныне проще сказать, что

если $Da_i = \sigma_i^2$, то $D[a\lambda] = [\lambda\lambda\sigma\sigma] = [\lambda\lambda/p]$.

Условие максимума веса P совместно с ограничением (4.2) приводит (с. 332) к

$$a \approx a_0 = [pa]/ \sum p_i = [pa]/P,$$

где a_0, разумеется, удовлетворяет условию наименьших квадратов

$$\sum \sqrt{p_i}\ (a_i - a_0)^2 = \min.$$

с) Определение k. Марков (с. 338 – 339) вычислил математическое ожидание k (тем самым и приближенное значение этой величины):

$$k = \frac{\sum p_i (a_i - a_0)^2}{n-1}. \tag{4.4}$$

Выкладки оказались нетрудными, но их нельзя было бы обобщить на случай нескольких неизвестных. Заметим, что Гаусс счел возможным рассматривать только этот общий случай.

d) Нормальное распределение (с. 341 – 344). Пусть

$$\Delta = a - a_i, \text{ or } \Delta = a - a_0, \tag{4.5a, 4.5b}$$

тогда приближенное (?) значение $E\Delta^2$ (= h) равно (4.4). Далее без пояснений он указал, что h в точности равно второму моменту соответствующей случайной переменной (последнего термина он не употребил). Странно, что Марков оставил без внимания, что (4.5a) и (4.5 b), равно как и соответствующие значения h, отличаются друг от друга.

Далее Марков предположил, что Δ (без уточнения!) распределено, как мы скажем, с дисперсией h, т. е. как $N(0, \sqrt{h})$, и, на с. 342 – 344, упомянул два возможных обоснования. Первое: Δ "рассматривают" как сумму многих независимых ошибок; второе, нормальность соответствует практике (ср. его отрицательное мнение об этих предпосылках в п. 2.-3). Он не уточнил, что частные погрешности должны быть одного и того же порядка, на что, кстати, косвенно указал Лаплас [27, с. 536], который заметил, что нормальное распределение в геодезических работах является следствием применения *повторительных* теодолитов (обеспечивших равенство порядков двух основных ошибок наблюдения).

Марков не предложил никакого количественного правила для проверки указанного соответствия, но на с. 349 он сослался на критерий соответствия хи-квадрат Пирсона, см., однако, наш п. 4.3b.

4.3. Статистические ряды (с. 345 – 373)

Этот раздел Марков назвал *Определение вероятностей по наблюдениям.*

a) Вычисление постоянной вероятности (α) в схеме испытаний Бернулли (с. 345 – 349). Пусть после s испытаний число успехов, чья вероятность постоянна, равна σ; результат каждого испытания, как мы скажем, – индикаторная случайная переменная x_i с биномиальным распределением и

$$E(x_i - \alpha)^2 = k_1 \approx \alpha(1-\alpha) = \frac{\sigma(s-\sigma)}{s(s-1)}. \qquad (4.6)$$

Последнее равенство следует из (4.4). С другой стороны, поскольку $\alpha \approx \sigma/s$,

$$k_2 \approx \frac{\sigma(s-\sigma)}{s^2}, - \qquad (4.7)$$

выражение "не вполне свободное" от постоянной погрешности. Можно полагать (этого в тексте нет), что постоянство α следует из примерного равенства $k_1 \approx k_2$. Одно замечание Маркова (с. 345) представляется, однако, сомнительным (и ненужным). Он указал, что принятые им результаты испытаний означают, что σ раз $\alpha \approx 1$ и $(s - \sigma)$ раз $\alpha \approx 0$. Как же тгда понять постоянство α?

Марков далее обобщил свое описание на несколько рядов испытаний. На этот раз выражениями для k, которые "до известной степени" обеспечивали проверку предпосылок о независимости испытаний и постоянства вероятности α (здесь это уже было указано), были

$$k_1 \approx [1/(n-1)] \sum s_i [(\sigma_i/s_i) - \hat{p}\,]^2 \text{ и } k_2 \approx \hat{p}(1-\hat{p}). \qquad (4.8; 4.9)$$

Здесь n – число рядов и $\hat{p} = \sum\sigma_i / \sum s_i$ – статистическая вероятность успеха.

b) Сравнение теоретической (p) и статистической вероятностей (с. 349 – 353). Марков рассмотрел эксперимент Уэлдона с бросками 12 костей [28]. В 185 бросках ($N_0 = 185$) исходы 5 и 6 не появились; в $N_1 = 1149$ бросках один из этих исходов появился один раз; ...; и в $N_{11} = 4$ бросках – 11 раз. После всех 26 306 бросков статистическая вероятность выпадения пяти или шести очков оказалась равной $\hat{p} = 0.33770$, и $\hat{p} - p$

= 0.00436. Марков применил центральную предельную теорему (а не критерий соответствия Пирсона), чтобы доказать, что вероятность успеха действительно была выше 1/3. Он подтвердил это заключение по теореме Бейеса с переходом к нормальному распределению и указал, что вероятность неравенства $\Delta p \geq 0.00436$ действительно была очень высока.

И Феллер [29, п. 2 из гл. 6], и Хальд [30, с. 201] также рассмотрели указанный эксперимент, равно как и его исследование Пирсоном и Фишером (который заметил, что отклонение опыта от теории могло было быть вызвано и изменением вероятности во времени, и зависимостью между бросками).

Марков дополнительно исследовал тот же эксперимент как объединение 12 рядов, полагая, что $s_i = 12N_i$ и $\sigma_i = iN_i$. Он получил $k_1/k_2 = 1.0049$ и, поскольку вес вероятности $\hat{p}$ был равен $P = 12\cdot 26\,306 = 315\,672$, формула (4.4) привела к математическому ожиданию квадрата его погрешности

$$E\Delta \hat{p}^{\,2} = k/P = 0.71\cdot 10^{-6}.$$

Марков заключил, что вероятная ошибка $\hat{p}$ (здесь и на с. 418 он молчаливо предположил реализацию соответствующего нормального распределения) равна 0.00056, так что с вероятностью 1/2 вероятность выпадения пяти или шести очков в опыте Уэлдона уклонялась от $\hat{p}$ меньше, чем на эту величину.

с) Коэффициент дисперсии (с. 353 – 373). Так Марков (с. 353) назвал величину

$$L^2 = k_1/k_2,$$

заметив, что этим термином обычно называли $\sqrt{k_1/k_2}$ и на с. 355 – 356 сказал несколько слов о его введении и исследовании Лексисом, Борткевичем и Дормуа (который применил иной вариант этого коэффициента). Позже Борткевич [31] сравнил заслуги Лексиса и Дормуа и решительно предпочел первого.

Далее Марков (с. 356 – 360) повторил свое прежнее непростое доказательство [11] того, что $EL^2 = 1$ и лишь выписал свой окончательный результат по поводу $E(L^2 - 1)^2$. В заключение он вычислил EL^2 для двух частных случаев переменной вероятности успеха.

Как мы только что указали, это исследование статистических рядов было непростым; вместе с тем Марков не упомянул примыкающие исследования Чупрова [32; 33], первое из которых (с более приемлемым выводом математического ожидания коэффициента дисперсии) он сам представил к публикации в *Известиях* Петербургской академии наук.

4.4. Уравнивание наблюдений нескольких неизвестных (с. 373 – 397)

Марков рассмотрел определение m неизвестных $a_1, a_2, \ldots, a_m$ по n независимым наблюдениям ($n > m$) $b_1, b_2, \ldots, b_n$, свободным от постоянных ошибок, в соответствии с линейными уравнениями

$$A_{1i}a_1 + A_{2i}a_2 + \ldots + A_{mi}a_m \approx b_i,\ i = 1, 2, \ldots, n. \tag{4.10}$$

Его общие соображения были, конечно же, известны. Как и в п. 4.(2a), необычным было введение возможных результатов наблюдения. Марков исходил из принципа наибольшего веса, получил нормальные уравнения, вычислил веса [оценок] неизвестных, и вывел формулу Гаусса для средней квадратической ошибки единицы веса, т. е. формулу (4.4) с подходяще измененным числителем и со знаменателем ($n - m$).

4.5. Интерполяция (с. 398 – 403 и 427 – 446)

Даны значения y_i, $i = 1, 2, \ldots, n$, неизвестной линейной функции $y(x)$ в точках наблюдения x_i, которые считаются безошибочными. Требуется определить значение функции в произвольной точке x по принципу наибольшего веса (с. 398 – 403).

Введя среднее арифметическое наблюдений $\bar{x}$, и выражая y как

$$y = a_1 + a_2(x - \bar{x})$$

с еще неизвестными коэффициентами, Марков принимает, что $y \approx \sum\lambda_i y_i$ и ограничивает множители λ_i условием

$$\sum\lambda_i y_i = y, \tag{4.11a}$$

которое обеспечивает отсутствие постоянной погрешности. Оказывается поэтому, что

$$\sum \lambda_i x_i = x. \tag{4.11b}$$

Далее, положив

$$\lambda_i = \mu_1 + \mu_2(x_i - \bar{x}), \tag{4.12}$$

Марков вычисляет неизвестные λ_i, μ_1 и μ_2 и указывает, что найденные таким образом λ_i обеспечивают максимальный вес. Да, постоянные ошибки были исключены, но не так как раньше: условие (4.2) было необходимо потому, что каждое наблюденное a_i было примерно равно единственному неизвестному (a), здесь же y_i могут значительно разниться друг от друга. Что касается условий (4.12), остается неясным каким образом они приводят к максимальному весу. Во всяком случае, проблему интерполяции можно непосредственно и гораздо проще решить по МНКв без привлечения множителей λ_i [35, п. 17].

Марков (с. 427 – 446) также исследовал параболическое интерполирование. Он последовал за Чебышевым и привел разработанные примеры (указав несколько недочетов в вычислениях своего учителя). Нам достаточно сказать, что на этот раз Марков избрал прямой путь.

4.6. Корреляция (с. 403 – 427)

Марков рассмотрел линейную корреляцию и применил МНКв для определения параметров линий регрессии. Он также исследовал случайные переменные с плотностью распределения в виде квадратичной формы, – в виде

$$f = a_{11}x_1 + 2a_{12}x_1x_2 + a_{22}x_2$$

для двух переменных, – и даже в виде g (f), где функция g была ограничена только общими аналитическими и стохастическими требованиями. Марков не сослался на одну из своих предыдущих глав, в которой (с. 275 – 287) он рассматривал *связанные* переменные с плотностью e^{-f} и в библиографию к которой он включил книгу Слуцкого [36]. В 1916 г. Марков [11, с. 533] подверг теорию корреляции острой и уже в то время, видимо, незаслуженной (несмотря на противоположное мнение Ю. В. Линника, выраженное им в комментарии к этой статье), [30, п. 27.7] критике, указывая, что она, определяя вероятные погрешности "различных коэффициентов", "вступает в область фантазии, гипноза и веры в математические формулы, которые в действительности не имеют твердого научного основания". В 1924 г. от критики он явно отказался.

4.7. Уравнивание наблюдений нескольких переменных, обобщение (с. 446 – 469)

Пусть, вдобавок к системе (4.10), имеется еще несколько уравнений

$$A_{j1}\, a_1 + A_{j2}\, a_2 + \ldots + A_{jm}\, a_m = c_j,\ j = 1, 2, \ldots, s, \tag{4.13}$$

которые должны быть выполнены точно; простейший пример: после их исправления, сумма измеренных углов плоского треугольника должна равняться 180°.

Марков заметил, что использовать уравнения (4.13) можно либо для исключения s неизвестных из системы (4.10), либо назначив каждому из них бесконечно большой вес и присоединив их все к этой системе. Обе эти возможности он подробно исследовал. Однако, во-первых, рассматриваемое обобщение не получило широкого распространения, хоть Бессель и предложил соответствующий способ его уравнивания. Во-вторых, Марков оставил в стороне важный и обычно применявшийся способ уравнивания триангуляции, – уравнивание прямых наблюдений *с условиями* по *методу условных уравнений.* Именно, начиная с Гаусса [37], уравнения (4.13) вводились в уравнивание совместно с непосредственными измерениями, т. е. не с уравнениями (4.10), а с равенствами типа $a_i \approx b_i$.

В заключение Марков уравнял углы a_i плоского треугольника, полагая, что они были измерены n_i раз с весами каждого измерения, равными p_i. Вряд ли такое обобщение было интересно: можно было бы объединить n_i и p_i в единый параметр.

5. Обсуждение

5.1. Обоснование метода наименьших квадратов

Нейман [38, с. 595] ошибочно приписал Маркову второе (окончательное) обоснование МНКв по принципу наибольшего веса, соединенного с несмещенностью. Затем Дейвид и Нейман [39] повторили эту ошибку и даже ухудшили положение доказательством *обобщенной теоремы Маркова*, фактически принадлежавшей Гауссу. Впрочем, Нейман [40, с. 228] в конце концов признал свою ошибку, но уже после того, как Плакетт [41, с. 460] заключил, что Марков "может быть пояснил предположения [введенные при обосновании МНКв], но не доказал ничего нового". Полезно заметить, что Колмогоров [42] не упомянул Маркова; он, правда, вообще после Гаусса не назвал никого по имени, но, посчитай он это возможным, наверное сделал бы для Маркова исключение.

И после всего этого Леман [43, с. 587], как указал другой автор [44], пустил в научный оборот термин *теорема Гаусса – Маркова*, который Сенета [46, с. 265] справедливо назвал ошибочным и добавил, что в соответствующих работах Маркова *мало оригинальности*. Тем не менее, повторим (см. конец п. 2), что Марков решительно поддержал принцип наибольшего веса. Это тем более важно, что от застарелых ошибок трудно избавиться: даже Фишер [48, с. 260] считал, что МНКв является "специальным приложением метода наибольшего правдоподобия, из которого его можно вывести".

5.2. Примыкающие темы

Обоснование МНКв в указанном только что смысле означает необходимость выбора несмещенных и эффективных оценок, и Линник и др. [49, с. 637] заявили, что Марков "по существу" это и сделал, но то же самое можно было сказать про Гаусса! Ср. Нейман [38, с. 593]: значение соответствующих работ Маркова состоит "в основном в ясной формулировке проблемы". В большой степени он, кстати, тем самым возразил самому себе (см. п. 5.1).

И вот мнение двух авторов: работы Маркова – привели "к полному и завершенному решению основных вопросов ... способа наименьших квадратов" (Безикович [50, с. VII]); – довели метод Гаусса "до высшего логического и математического совершенства" (Идельсон [35, с. 14]). Но это неверно, а Безикович столь же неверно приписал Маркову "совершенно новое развитие" теории корреляции (с. XIV).

5.3. Вопросы методологии

По поводу методологической ценности работ Маркова существуют противоположные высказывания. Бернштейн [51, с. 425] заявил, что руководство Маркова [6] и "его оригинальные мемуары являю[тся] образцами точности и ясности изложения", а Линник и др. [49, с. 615] указали, что для стиля Маркова "характерны ясность и четкость языка, тщательная отделка деталей" и результатов. Мы не согласны с подобными оценками, а по поводу отделки деталей мы напомним об отказе Маркова от рассмотрения важного варианта уравнивания измерений (п. 4.7). Вот подходящее мнение Баушингера:

"Этот случай [уравнивания] встречается на практике очень часто и поэтому должен быть обсужден, хоть он и является частным случаем предыдущего" (который Марков действительно рассматривал) [52, с. 794].

И мы не доверяем Чупрову [53, с. 167], который полагал, что книга Маркова [6, 1924] "может быть признана настольным руководством по теории вероятностей для статистиков" и что для Маркова характерна "прозрачная ясность изложения". Впрочем, он (с. 168) разумно возразил против обсуждения корреляции в главе о МНКв и существенно раскритиковал само ее изложение. По отношению к МНКв Марков сам признался в 1910 г. [54, с. 29], что ему "часто приходилось слышать, что у меня [у него] изложено недостаточно ясно". Вспомним по этому поводу Кояловича (п. 4.2а) и упомянем Идельсона [35, с. 101]: глава о МНКв в руководстве "трудно написан[а]". Да и вообще стиль Маркова был тяжеловесен и иногда [55, с. 341] его было почти невозможно понять.

Марков почти не нумеровал свои формулы, он их переписывал. На с. 328 – 330 равенство (4.2) появилось 5 раз! Он явно избегал указательные местоимения. Действительно, на с. 328 (аналогичные случаи см. с. 379 и 381 – 382) мы читаем: "Выбор коэффициентов [следует выключенная строка] находится в нашем распоряжении. Мы подчиним коэффициенты [эта строка повторена] двум условиям ..." Марков не применял скобок Гаусса (п. 4.2b), а обозначение для среднего арифметического ввел лишь на с. 463. Термины *нормальное распределение* и даже *коэффициент корреляции* у него отсутствуют; он ни разу не сказал ни *случайная ошибка*, ни *случайная величина*. Вот его объяснение 1912 г.:

"Я повсюду, где только возможно, исключу ничего не определяющие слова *случайно* и *наудачу*; где же придется употреблять их, приведу соответствующее ... объяснение" [54, с. 71].

И все-таки *наудачу* он применял, особенно при обсуждении геометрической вероятности [6, 1924, гл. 5], но иногда вместо *случайный* писал *неопределенный*, что было намного хуже. Гораздо правильнее поступил Васильев [56, с. 127 – 131], который одним из первых в России начал применять термин *случайная величина*, а на с. 133 добавил, что "случайные ошибки имеют все свойства случайных величин" (и "свои специальные свойства").

Библиографические ссылки не были у Маркова достаточно определенными. Так, на с. 427 он просто сослался на оба тома (первого) собрания сочинений Чебышева, а на с. 163 привел три библиографические ссылки без дат публикации. Далее, на с. 10 (вне связи с МНКв) он сформулировал аксиому (которая вряд ли пережила его, см. п. 5.5), не выделив ее из контекста и сослался на нее на с. 24, а структура его руководства усложнялась с каждым изданием.

Наконец, глава о МНКв была вряд ли привлекательна. Ни математикам, ни геодезистам не могло понравиться фактическое замалчивание работ Пирсона (см. также п. 5.4), последним же, кроме того, не нужна была ни интерполяция, ни исследование статистических рядов, а вот о корреляции им следовало бы прочесть больше. Далее, отсутствие скобок Гаусса и употребление давно устаревшего термина *практическая геометрия* вместо *геодезия* (с. 462) вряд ли было им по вкусу.

Мы обязаны добавить, что Марков был хорошим педагогом и научным воспитателем. Так утверждает Н. М. Гюнтер [57], один из двух его известных учеников. Другим был Д. А. Граве; впрочем, оба они были также многим обязаны А. Н. Коркину.

5.4. Отношение к работам других авторов

Сюда частично относится наш п. 5.3. Чупров [53, с. 169 – 170] заметил, что Марков "почти не выходит" за рамки своих собственных трудов; добавим, что он не сослался на работы Чупрова (см. п. 4.3c). Марков не упомянул ни многих иностранных ученых (Больмана, Стьюдента, Фишера, Юла), ни даже С. Н. Бернштейна, равно как ни Мизеса, ни Линдеберга. Впрочем, про двух последних Марков может быть знал недостаточно ввиду общего тогдашнего положения в России. Недаром он вообще не сослался ни на одно сочинение, вышедшее в свет после 1914 г.

Кроме того, Марков лишь вскользь упомянул и критерий соответствия Пирсона. Хуже того, его рассуждения в п. 4.3b наводили на мысль, что этот критерий вообще не нужен: Марков применил классические теоремы, пригодные для большого числа измерений, тогда как нововведение Пирсона этого не требовало. Далее, Марков (п. 4.6) еле коснулся корреляции и вообще не упомянул кривых Пирсона. Они, правда, не были достаточно обоснованы, но ведь сам Марков [6, 1913, Введение, перепечатано в 1924 г.] заявил, что в прикладной математике "нельзя совершенно отказаться от приближенных формул, остающихся ... без оценки их погрешности", а в 1915 г. прямо указал, что формулы Пирсона эмпирические, а потому не требуют теоретического доказательства и могут быть обоснованы своим согласием с наблюдениями [58, с. 32].

Возможно, что Марков в этом случае поступал в соответствии со своим правилом, которого вряд ли следовало жестко придерживаться и которое, видимо, поясняет, почему он не сопроводил свои *цепи* указанием на их возможное естественнонаучное приложение. Вот что он написал в 1910 г. [54, с. 59]: "Я ни на шаг не выйду из той области, где компетенция моя не может подлежать сомнению". А ведь Марков мог, например, подкрепить известное рассуждение Пуанкаре 1896 г. о равномерном распределении эклиптических долгот малых планет ссылкой на эргодическое свойство (как бы мы сказали сейчас) однородных цепей с конечным числом состояний. Странно также, что, в противоположность Лапласу, Марков ничего не сказал о принципиальной пользе института страхования жизни в своих газетных статьях 1906 г., резко критиковавших невыгодные для населения условия страхования детей [59].

С другой стороны, по сравнению с 1916 г. (п. 4.6) Марков несколько смягчил свою точку зрения на корреляцию и молчаливо признал (примерное) нормальное распределение ошибок наблюдения (ср. пп. 2.-3 и 4.2d). Он, кроме того, фактически согласился со Слуцким, который, защищая свою книгу [36] от критики, указал, что "недостатки изложения Пирсона – временные" и будут преодолены так же, как это случилось с недостатками анализа XVII – XVIII вв. [60, с. 132].

5.5. Замечание о теории вероятностей

Мы не можем обойти молчанием общие рассуждения Маркова о теории вероятностей. Он (с. 4) привел классическое определение вероятности, но добавил, что "понятия определяются не столько словами ... как нашим отношением к ним [т. е. к понятиям] (с. 2, прим.). Другой пример чрезмерного, на наш взгляд, стремления к строгости (при оставшемся шатком обосновании теории в целом!) виден на с. 1, прим.: "слово *мы* [употребленное им в житейском смысле] не сообщает исчислению вероятностей никакой особой субъективности".

Далее, на с. 10 (см. также п. 5.3) Марков ввел следующую *аксиому*: Если равновозможные события "делятся по отношению к событию *A* на благоприятные и неблагоприятные", то после появления *A* последние "отпадают", первые же "остаются ... равновозможными". На с. 13 – 14 Марков доказал теоремы сложения и умножения вероятностей довольно сложным путем. Так, во втором случае, рассматривая два зависимых события, он сослался на свою аксиому в связи с появлением первого из них. И вот поразительное заключение:

"Теоремы о сложении и умножении вероятностей, в связи с вышеуказанной аксиомой, служат незыблемой основой для исчисления вероятностей как отдела чистой математики" (с. 24).

Добавим, что на с. 241 Марков (не выделяя своего утверждения из контекста) принимает, говоря современным языком, расширенную аксиому сложения (и умножения!).

6. Выводы

Впечатление тягостное ... Ученый, открывший новые горизонты перед математикой и естествознанием, не справился с новой задачей. Глава о МНКв неудачна, а руководство в целом (мы не имеем в виду его приложения, которые, кстати сказать, не следовало включать в подобное сочинение) представило теорию вероятностей на уровне прежних нескольких десятилетий. Попытка преобразовать ее в отдел чистой математики оказалась негодной, и можно заметить [61, с. 15], что первое систематическое изложение теории случайных величин, распределений их вероятностей и их характеристических функций появилось в 1925 г. [62].

Причинами тому были и пошатнувшееся здоровье Маркова, а затем и его неожиданная смерть (которая быть может вообще помешала ему закончить пересмотр руководства), и общее положение в России того времени, чрезвычайно затруднявшее научные связи с зарубежными странами. Но приходится добавить: Марков был слишком своенравен; он плохо воспринимал эвристические понятия; чересчур ограничивал свое поле зрения; недоверчиво относился к чужим нововведениям. Последние два обстоятельства созвучны и с консерватизмом Чебышева [63, с. 330], и с противопоставлением Римана Лобачевскому, на котором странным образом настаивал Ляпунов [64, с. 19 – 20] несмотря на существовавшую единую картину неевклидовой геометрии, представленную Клейном.

Признательность. Этот текст является несколько переработанным вариантом первоначального, который появился на свет на английском языке в журнале Historia Scientiarum, vol. 16, №1, 2006, pp. 80 – 95. Мы благодарны редактору этого журнала за любезное разрешение опубликовать статью в переводе.

Список литературы

1. *Sheynin O. B.* Markov’s work on probability // Archive for History of Exact Sciences. 1989. Vol. 39. P. 337 – 377. Vol. 40. P. 387.
2. *Марков А. А.* К вопросу о прочности стекла (рукопись 1903 г.) // Историко-математические исследования. 1990. Вып. 32 – 33. С. 456 – 467.
3. *Коялович Б. М.* Письмо А. А. Маркову 25.9.1893. Архив РАН. Фонд 173. Опись 1. Дело 10. №1. С. 3 об – 4.
4. *Добровольский В. А.* Математика в высших технических и специальных военных учебных заведениях [5, с. 408 – 419].
5. *Штокало И. З.*, редактор. История отечественной математики. Т. 2. Киев, 1967.
6. *Марков А. А.* Исчисление вероятностей. СПб, 1900. Последующие издания: СПб, 1908; СПб,1912; М., 1924 (посмертное).
7. *Марков А. А.* Закон больших чисел и способ наименьших квадратов (1899) [8, с. 231 – 251].
8. *Марков А. А.* Избранные труды. [М.], 1951.
9. *Newcomb S.* On the right ascensions of the equatorial fundamental stars. Washington, 1872.
10. *Gauss C. F.* Werke. Bd. 9. Göttingen – Leipzig, 1903.

11. *Марков А. А.* О коэффициенте дисперсии (1916) [8, с. 525 – 535].
12. *Маиевский Н.* Изложение способа наименьших квадратов и применения его преимущественно к исследованию результатов стрельбы. СПб, 1881.
13. *Усов Н.А.* Замечание по поводу теоремы П. Л. Чебышева // Математический сборник. 1867. Т. 2. С. 93 – 95.
14. *Чебышев П. Л.* Теория вероятностей. Лекции 1879/1880 г. М., 1936.
15. *Ivory J.* On the method of least squares // London, Edinburgh and Dublin Philosophical Magazine. 1825. Vol. 65. P. 1 – 10, 81 – 88, 161 – 168.
16. *Galloway T.* Treatise on probability. Edinburgh, 1839.
17. *Ellis R. L.* On the method of least squares (1844) [18, p. 12 – 37].
18. *Ellis R. L.* Mathematical and other writings. Cambridge, 1863.
19. *Шейнин О. Б.* Отзыв А. А. Маркова об одной статье Б. Б. Голицина // Историко-математические исследования. 1990. Вып. 32 – 33. С. 451 – 455.
20. *Galitzin B., Fürst*. Über die Festigkeit des Glases // Известия Петербургской академии наук. Сер. 5. 1902. Т. 16. С. 1 – 29.
21. *Менделеев Д. И.* О сцеплении некоторых жидкостей (1860). Сочинения. Т. 5. М. – Л., 1947. С. 40 – 55.
22. *Newcomb S.* A new determination of the precessional motion // Astronomical Journal. 1897. Vol. 17. P. 161 – 167.
23. *Марков А. А. младший*. Биография А. А. Маркова [старшего] [8, с. 599 – 613].
24. *Fourier J. B. J.* Sur les resultats moyennes (1826). Œuvres. T. 2. Paris, 1890. P. 525 – 545.
25. *Шейнин О. Б.* Теория вероятностей. Исторический очерк. Берлин, 2005.
26. *Коялович Б. М*. Письмо А. А. Маркову 2.10.1893. Архив РАН. Фонд 173. Опись 1. Дело 10. №8. С. 11 об – 12.
27. *Laplace P. S.* Théorie analytique des probabilités, Deuxième supplément (1818). Œuvres complètes. T. 7. No. 2. Paris, 1886. P. 531 – 580.
28. *Pearson K.* On a criterion that a given system of deviations can be supposed to have arisen from random sampling // London, Edinburgh and Dublin Philosophical Magazine. 1900. Vol. 50. P. 157 – 175.
29. *Феллер В.* Введение в теорию вероятностей и ее приложения. М., 1964. Перевод с английского издания 1957 г.
30. *Hald A.* History of mathematical statistics from 1750 to 1930. New York, 1998.
31. *Bortkiewicz L. von*. Lexis und Dormoy // Nordic Statistical Journal. 1930. Vol. 2. P. 37 – 54.
32. *Чупров А. А.* О математическом ожидании коэффициента дисперсии // Известия Петербургской академии наук. 1916. Т. 10. С. 1789 – 1798.
33. *Чупров А. А.* К теории стабильности статистических рядов (1918 – 1919, нем.) [34, с. 138 – 224].
34. *Четвериков Н. С.*, редактор. О теории дисперсии. М.
35. *Идельсон Н. И*. Способ наименьших квадратов и теория математической обработки наблюдений. М., 1947.
36. *Слуцкий Е. Е.* Теория корреляции. Киев, 1912.
37. *Гаусс К. Ф.* Теория комбинации наблюдений, Дополнение (1828, латинск.). Избранные геодезические сочинения. Т. 1. М., 1957. С. 59 – 87.

38. *Neyman J.* On two different aspects of the representative method // Journal of Royal Statistical Society. 1934. Vol. 97. P. 558 – 625.
39. *David F.N., Neyman J.* Extension of the Markoff theorem // Statistical Research Memoirs. 1938. Vol. 2. P. 105 – 117.
40. *Neyman J.* Lectures and conferences on mathematical statistics and probability. Washington, 1952.
41. *Plackett R. L.* Historical note on the method of least squares // Biometrika. 1949. Vol. 36. P. 458 – 460.
42. *Колмогоров А. Н.* К обоснованию метода наименьших квадратов // Успехи математических наук. 1946. Т. 1. С. 57 – 71.
43. *Lehmann E. L.* A general concept of unbiasedness // Annals of Mathematical Statistics. 1951. Vol. 22. P. 587 – 592.
44. *David H. A.* First occurrences of common terms in statistics and probability В приложении к [45, p. 209 – 246]: Электронный Препринт 2006г., Iowa State University.
45. *David H. A., Edwards A. W. F.* Annotated readings in the history of statistics. New York, 2001.
46. *Seneta E.* Markov [47, p. 263 – 265].
47. *Johnson N. L., Kotz S.*, Editors. Leading personalities in statistical sciences. New York, 1997.
48. *Fisher R. A.* Statistical methods for research workers (1925). В книге автора *Statistical Methods, Experimental Design and Scientific Inference*. Oxford, 1990.
49. *Линник Ю. В., Сапогов Н. А., Тимофеев В. Н.* Очерк работ А. А. Маркова по теории чисел и теории вероятностей [8, с. 614 – 640].
50. *Безикович А. С.* Биографический очерк [А. А. Маркова] [8, с. III – XIV].
51. *Бернштейн С. Н.* О работах П. Л. Чебышева по теории вероятностей (1945). Собрание сочинений. Т. 4. М., 1964. С. 409 – 433.
52. *Bauschinger J.* Ausgleichungsrechnung. Encyklopädie der mathematischen Wissenschaften. 1900 – 1904. Bd. 1. Tl. 2. P. 769 – 798.
53. *Чупров А. А.* Рецензия на книгу А. А. Маркова [6, 1924] (1925) [54, с. 167 – 170].
54. *Ондар Х. О.*, редактор. О теории вероятностей и математической статистике. Переписка А. А. Маркова и А. А. Чупрова. М., 1977.
55. *Марков А. А.* Распространение закона больших чисел на величины, зависящие друг от друга (1907) [8, с. 339 – 361].
56. *Васильев А. В.* Теория вероятностей. Казань, 1885. Литография.
57. Гюнтер Н. М. О педагогической деятельности А. А. Маркова // Известия Российской Академии наук. 1923. Т. 17. С. 35 – 44.
58. *Марков А. А.* О проекте П. С. Флорова и П. А. Некрасова преподавания теории вероятностей в средней школе // Журнал Министерства народного просвещения. 1915. Май. Отдел современной летописи. С. 26 – 34.
59. *Шейнин О. Б.* А. А. Марков и страхование жизни // Историко-математические исследования. 1997. Вып. 2 (37). С. 22 – 33.
60. *Шейнин О. Б.* Е. Е. Слуцкий: к 50-летию со дня смерти // Историко-математические исследования. 1999. Вып. 3 (38). С. 128 – 137.
61. *Крамер Х.* Полвека с теорией вероятностей: наброски воспоминаний. М., 1979. Перевод с английского издания 1976 г.
62. *Lévy P.* Calcul des probabilités. Paris, 1925.
63. *Новиков С.П.* Вторая половина XX в. и ее итог: кризис физико-

математического сообщества в России и на Западе // Историко-математические исследования. 2002. Вып. 7 (42). С. 326 – 356.
64. *Ляпунов А. М.* П.Л. Чебышев (1895) [65, с. 9 – 21].
65. *Чебышев П. Л.* Избранные математические труды. М. – Л., 1946.

Фишер (Fisher R. A.)
Фурье (Fourier J. B. J.)
Хальд (Hald A.)
Чебышев П. Л.
Чупров А. А.
Юл (Yule G. U.)

O. B. Sheynin. Markov's work on the mathematical treatment of observations.
I discuss Markov's contribution to the treatment of observations. In 1899, he rejected the theory of correlation and refused to recognize any advantages in the application of the method of least squares (and insisted on this opinion to the end of his life) but supported the Gaussian mature justification of the method. In later life, he had been too critical of Pearson but gradually somewhat softened his attitude under Chuprov's influence. However, even the posthumous edition of his treatise (1924) insufficiently reflected the new statistical ideas and methods, partly because of his worsened health and the difficult general conditions in Russia, and partly owing to his personal traits. A special feature of his treatise is an extremely unfortunate attempt to elevate the theory of probability to the level of pure mathematics.

Romanovsky's Correspondence with K. Pearson and R. A. Fisher

Oscar Sheynin

Abstract

Eight letters from Romanovsky to Pearson (1924 – 1925) and 23 letters between him and Fisher (1929 – 1938) are published for the first time. The letters to Pearson were occasioned by Romanovsky's manuscript sent to *Biometrika* (it appeared there in 1925), by his attempts to continue publishing there and the overlapping of their findings. Among the topics discussed in Romanovsky's correspondence with Fisher were the latter's books, his *t*-statistics and field experiments. In a letter from Paris Romanovsky asked Fisher to help a Russian emigrant scientist.

1. Introduction

Vsevolod Ivanovich Romanovsky (1879 – 1954) was an outstanding mathematician and statistician. I refer to his works but in some cases the reader should consult the complete bibliography of his writings in Bogoliubov & Matvievskaia (1997), abbreviated as B&M. These authors described his life and work, and, in particular (p. 85), mentioned his acquaintance with Karl Pearson (which I am unable to confirm) and Ronald Aylmer Fisher. For my part, I make public his correspondence with these most eminent statisticians and specify or add the relevant bibliographic information[1] but omit unimportant everyday details. Two letters from Romanovsky's correspondence, both written in March 1931 on the role of prior distributions in formulas of the Bayesian type, are already published (Bennett 1990, pp. 200 – 202). The reader will notice that Romanovsky's English was very imperfect but understandable and that his terminology is now dated, see Note 8.

As a preliminary, I supplement B&M by a few words. In 1923, Chuprov became interested in Romanovsky's work. Sometimes revealing there "rather large mistakes", he nevertheless at once made him out as a prominent scientist and entered into correspondence with him (Sheynin 1996, pp. 50 – 53), cf. the beginning of Letter 23. I also noted (Ibidem, pp. 40n and 96) that, at the beginning of his scientific career, Romanovsky had certainly overrated the

Bortkiewicz law of small numbers and stressed the natural-scientific essence of the law of large numbers and invariably called it a physical law [21, p. 18; 22, book 1, p. 127].

Romanovsky's ties with Western statisticians were not at all restricted to correspondence. He published many important papers in Europe one of which [11] served as a point of departure for E.S. Pearson and Neyman. Even after several decades, the former did not forget to testify to this (E.S. Pearson 1966). In 1939, on the occasion of his jubilee, the Central Asian University in Tashkent, where Romanovsky was working since its establishment in 1918 until his death, published a collection of papers written by many most eminent Soviet and foreign mathematicians and statisticians in his honour (Zbornik 1939).

Then, Romanovsky published four reviews of Fisher's books. He [14] described in detail Fisher's *Statistical methods* (1934) justly calling it a "remarkable phenomenon" (p. 127) and indicated that it was already translated into Russian and published "as a manuscript in a small number of copies", cf. Letter 28. Nevertheless, an ordinary Russian edition appeared only in 1958 and even then the (state-owned) publisher accompanied it by a critical comment (p. 5). There, he accused Fisher of a "bourgeois narrow-mindedness and formality of views", disregard of the qualitative side of social phenomena etc. For that matter, in Russia, a combined work of mathematicians and sociologists was unheard of at the time.

In [13] Romanovsky indicated that the work of Fisher and his associates was based on experimentation of many years and predicted the importance of their main ideas. Soon Romanovsky [16] described Fisher's new book (1935), *Design of experiments*, and concluded that it deserved "greatest attention" (p. 125). It was not, however, translated. Finally, Romanovsky [19] reviewed the tables Fisher & Yates (1938). He called them valuable but added that they should be published in Russian in a revised and supplemented form. This, however, had not happened either.

From about 1927 the general situation in Russia, and certainly in statistics as well, sharply deteriorated. In particular, it became dangerous to cite Pearson favourably (Sheynin 1998)[2]. And it seems that even Fisher became suspicious. In addition to the above, I indicate an editorial note to Romanovsky's paper [9], see its p. 224:

The editorial staff does not share either the main suppositions of Fisher, who belongs to the Anglo-American empiricists' school, or Romanovsky's attitude to Fisher's constructions …

I stress that, although the Anglo-American statistical school had indeed been empirical to a large extent[3], the only occasion for such an attack could have been the general directive to deny all the *bourgeois*. Thus, Maria Smit (1927/1930, pp. 8 – 9) absurdly accused Romanovsky (and L.K. Lakhtin) of considering random variables with permanent laws of distribution. That, she declared, contradicted both the spirit of Darwinism and the Engels dialectic … I (Sheynin 1998) have already cited that worthy troglodite. Now, I might add that she had no incling of the difficulties connected with studies of such general densities (or, for that matter, of mathematics at all) and that she stated, in equally bad Russian, that "Engels' opinion retains its validness".

Even in 1938 Romanovsly [18, p. 17] nevertheless called Pearson the head of contemporary mathematical statistics; or, more precisely [17, p. 49], its co-creator (together with Galton). In the second case he added that "we also ought to name Fisher, Charlier and Chuprov"[4]. Fisher's subsequent findings advanced him to the very first place in statistics.

After World War II the Soviet authorities launched a new attack against mathematical statistics. In 1948, a Second All-Union Statistical Conference took place in Tashkent (Vtoroe 1948) and Romanovsky naturally became chairman of its organizing committee (B&M, p. 92). In his report, Kolmogorov (1948, p. 220) mentioned the "great" work done by Romanovsky and his school, but another speaker (Sarymsakov 1948, p. 222) blamed his teacher for "following in the Anglo-American direction". Moreover, the conference (Vtoroe 1948, p. 314) carried a resolution that indicated, without naming anyone, that there existed "servility and cringing to all foreign" and approvingly put on record that Romanovsky had admitted his previous ideological mistakes. The resolution also "decisively" condemned Nemchinov, with whom Romanovsky had been in correspondence (B&M, p. 93), for his active opposition to [the notorious humbug] Lyssenko.

The conference had a sudden consequence. Romanovsky published an unfortunate manual of error theory [20]. Like many other mathematicians and statisticians, he was not sufficiently acquainted with that subject[5] and no wonder that Chebotarev (1951) expressed reasonable criticism. However, he also recalled that Lenin had called Pearson a Machian and an enemy of materialism and he attacked Romanovsky (and the historian of astronomy Idelson) from the ideological viewpoint. Finally, Chebotarev formulated a few absurd remarks. Thus (p. 8), Romanovsky, like Mach and Pearson, only attempted to describe phenomena whereas Marx had established that the world should be changed rather than described …

Answering that lackey, Romanovsky [21, pp. 17 – 18] stated that Pearson's mathematics should not be lumped together with his philosophy, and that he, Romanovsky [18], supported the constructions of the Biometric school by a stochastic base and thus amended them[6]. Chebotarev (1953), however, repeated his accusations, although later on, likely being compelled by a somewhat improved general political situation, he (1958, pp. 571 and 586) began to recognize Romanovsky[7].

The correspondence below shows that Fisher held a high opinion of Romanovsky: not only did he describe his own work to his Russian colleague, he also expressed his desire to see Romanovsky's writings, even if in Russian (Letter 29). And Pearson published five of Romanovsky's papers in *Biometrika* and he certainly had to correct the English in each of these. Romanovsky's political inclinations are also felt in his correspondence: the GPU (more correctly, OGPU, the forerunner of the KGB), as he wrote (Letter 11), was "the most dreadfull and mightful organization in the present Russia". And my own experience, on my own scale, tells me that for him the correspondence should have certainly been a vent for fresh air.

2. Romanovsky's letters to Pearson

Letter No.1, 18.12.1924

I send you with this letter a paper on the distribution of the means and standard deviations in samples of an arbitrary number from normal populations with one or two arguments[8]. I think that it can interest the readers of *Biometrika* for it contains the complete and rigid solution of some problems which are partly not completely solved and partly new, as I know. I beg you to pay attention to the following places of my work:

P. 7, formula (21): the generating function of the moments of stand. dev. for one variable, and formula (22) – their general expression.

P. 10, formula (33): the exact value of the mean error of stand. dev. in samples of number s. On the p. 11 I give a short table for $s = 2$ to 30 of the true values of probable error of st. dev.[9] and of approximate values, published in Your *Tables.*

P. 18, formula (44): the generating functions of the moments of means in samples of number s from normal population with two arguments; p. 19, form. (52): their general expression.

P. 28, form. (81): the generating function of the moments of the products of st. deviations in samples from the population with two variables.

P. 31, form. (87): their general expression.

P. 35, form. (101) and (102): the equation of the surface of distribution of st. deviations for the same case.

Many other results I do not mention. I have received them by a method which I discovered some months ago and which I have hold for new till now. When I had finished my work and some other investigations with these method I have received from Prof. Tchuproff some his papers and among them his note (1924) on the book of Mr. Soper (1922) which I do not know, for I could not get it till now. From this note of Prof. Tchuproff I knew that my method in some essential points is contained in the Mr. Soper's book. I do not know the content of this interesting book and can only suppose as far as I know it from the note of Prof. Tchuproff that my method differs from the method of Mr. Soper in some directions (for example, in a symbolical calculus of moments and in its applications to continuous distributions), which are not unimportant. However, I do not pretend much, I state only that I had come independently to the same fundamental ideas as Mr. Soper, and that I developped them in some new directions. I hesitated how to include in a Post-scriptum all these remarks but finally I resolved to omitt them. If you accepte my paper for *Biometrika* (it would be very desirable for me and important), and if You find necessary to accompany it with some remarks on its relation with the Mr. Soper's method, I beg You very much to denote my independence from Mr. Soper.

Now I have finished another research on the product-moments of the form (in samples from normal population) $\bar\mu^h{}_{11}, \bar\mu^k{}_{20}, \bar\mu^l{}_{02}$ where **10**

$$\bar\mu_{11} = (1/s)\sum (x - \bar x)(y - \bar y)^2;\ \bar x = (1/s)\sum x;$$

$$\bar y = (1/s)\sum y;\ \bar\mu_{20} = (1/s)\sum (x - \bar x)^2;\ \bar\mu_{02} = (1/s)\sum (y - \bar y)^2$$

and on the equation of distribution of the quantities $\bar\mu_{11}$, $\bar\mu_{20}$, and $\bar\mu_{02}$. I have found the generating functions of these moments and the equation of distribution of these three quantities. The particular result of the last is the exact value of mean error of coefficient of correlation. It is following:

$$\sigma_{\bar r}{}^2 = \frac{1-r^2}{s-1}\,[F(1;1;(s+1)/2;r^2) + \frac{(s-1)^2 r^2}{(s+1)(1-r^2)}\,F(1;1;(s+3)/2;r^2) -$$

$$\frac{4r^2\Gamma^4(s/2)}{(s-1)(1-r^2)\Gamma^4[(s-1)/2]}\,F^2[1/2;1/2;(s+1)/2;r^2].$$

Here F denotes hypergeometrical function, $\bar r$ is coef. of corr. for samples of number s and r – coef. of corr. of the general population. From this formula it is not difficult to find the ordinary used first approximation $= \sigma_{\bar r}{}^2 (1 - r^2)^2/(s - 1)$, and the approximations of the higher orders.

I prepare now a paper containing the exposition of these results and I shall send it to You if You do not refuse to accept it. I am very glad that my paper on the moments of a hypergeometrical series will be printed in Your journal. If you find that some corrections and alterations in my paper must be made, I beg You to make them. I add to the remark on the $\sigma_{\bar r}$ that

$$\{\Gamma^4(s/2)/\Gamma^4[(s-1)/2]\} = (s^2/4)\,[1 - (3/s) + (5/2s^2) - (1/8s^4) + \dots].$$

Letter No. 2, 9.1.1925

I am returning you the proofs of my paper with many thanks for your corrections. The proofs are excellent and I have found almost nothing to correct in them. I am late with them because they were retained some days in our university befor to be handed to me.

Letter No. 3, 5.5.1925

I am very obliged for your very interesting letter. The problems you write me on are difficult and attracting and I do not see for the moment how to solve them (I mean the distribution of $\sqrt{\beta_1}$ and β_2, ${}_x\mu_3$, ${}_x\mu_4$, etc, $r_{\sigma_x\sigma_y}$ etc being much easier)[11].

The methods of R. A. Fisher or of generating functions seem to be of little use in these problems. I have a method of obtaining very various classes of moments of an arbitrary distribution of two variates, but, applied to your problems, it would involve infinite series of very complicated nature and it would be very difficult to prove the convergence of these series. I am very sorry that for some months I shall not be able to work on statistics, for I am writing now a text-book of analysis[12] for an editor in order to have necessary money for the voyage in England which I hope to do in the beginning of 1926.

For the same reason I could not rework my paper on the distribution of standard deviations very thoroughly[13]. Yet I have much shortened it and added to it my results on the coefficient of correlation. Now I send you this rewritten paper. I hope that it will be more satisfactory and beg you to publish it in *Biometrika.*

Letter No. 4, 2.6.1925

I am very sorry that there is some clashing in our work and that my paper cannot be published in *Biometrika.* I shall try to publish it in *Metron* or in *Nordisk statistisk tidskrift.*

I would be very glad if you added to your paper that many results contained therein are reached by me by a different method and independently from you[14]. Thus I can indicate your equations (v), (viii), (x), (xiv), (xxiii), (xliii), (xliv) and (xlv) which are also in my paper besides some other results which I do not find in your paper (the general formulae of the product-moments $\bar{\mu}_{20}$, $\bar{\mu}_{02}$, $\bar{\mu}_{11}$, of σ_x, σ_y, $\bar{\mu}$ and of correlation coefficient, their generating functions, the equation of distribution of $\bar{\mu}_{20}$, $\bar{\mu}_{02}$ and $\bar{\mu}_{11}$ etc). Perhaps it would be of interest to indicate the mean error of correlation coefficient in my form, which is (in your notations) [here Romanovsky rewrites the same formula from his Letter 1 although in a somewhat different notation. O. S.] and which is different from your equation (xxv).

I thank you very much for sending me your very interesting paper[15].

Letter No. 5, 1.9.1925

I must beg you to excuse me that I answer your letter of the 15-th June only now. I was not in Tashkend.

I am preparing now a paper on the distributions of $\bar{\mu}_{11}$ and $\bar{\rho}_{xy} = \overline{r\sigma}_y / \bar{\sigma}_x$ which I hope soon to send you for *Biometrika.*

This summer I have received some results of purely mathematical and some of statistical character. For example, I have demonstrated that the mean of some variate in samples of number s from a population with any distribution of this variate tends to be normal for $s \to \infty$[16], if some restrictions are laid on the growth of the moments of the variate. Then, I have discovered some interesting relations between the moments of any distribution and the coefficients of its taylorian expansion. I have also constructed an example of non-linear correlation whose correlation coefficient can be made as near to unity as you desire. If these results can be interesting for *Biometrika* I can send you short notes on them.

I have, endly, systematized several general theorems on the distributions [such as "given the distribution $\varphi(x_1; x_2; \ldots; x_n)$, to find the distribution of any functions of $x_1, x_2, \ldots, x_n$"; "given the distribution of x, y, z in a general population, to find the distribution of any functions of the samples of number s from this population" and so on][17]. I consider only continuous distributions.

Letter No. 6, 2.10.1925

I am very sorry that it may seem to you that I have acted incorrectly in regard to you, Fisher and others publishing my notes in the *Comptes rendus*. I have not pretended in them that my results are new and it is not written there. I have only written that "le but de cette note est d'indiquer une méthode nouvelle pour la recherche" etc[18] and I think that my method is indeed new. This does not contest Mr. Fisher in his note published in the *Comptes rendus* (1925) and containing the solution of an integral equation which I have received in one of my notes and could not to solve. Besides my aim in publishing the notes was not the claim of priority or of newness but to indicate a method which can be of use in many similar questions. I am very sorry that, trying to be short as possible, I have omitted all indications of the results already known and of their authors. But I think that this cannot lead anywhom knowing the modern state of statistics in mistake.

I can add to my explanations that my two notes on mathematical statistics were sent together in February befor I have received your letter with your paper (1925) which has reached me at the end of May or the beginning of June (I cannot remember exactly).

It is a very great grief for me that, as you write, I cannot be a contributor to *Biometrika* – the best journal of the theoretical statistics – and still greater one to have lost your good esteem of me. I shall be very obliged to you if you write me how you accept my explanations[19].

Letter No. 7, 5.11.1925

I think that I misunderstood the rules of *Biometrika*: I thought that the quite short abstracts from the contributions to it can be published elsewhere (such is the custum in many mathematical journals). I beg you to excuse me of this misunderstanding and of sending you my paper on the distribution of the regression coefficient: I see now that it could not be printed in *Biometrika*[20].

In order to clear up the matter wholly I beg you very much to write me if you could to accept my other contributions which are and will not be published nowhere or if you refuse in general my contributions to *Biometrika.*

The Statistical Cabinet of the Law faculty of our university through an agent of our Government will purchase an exemplar of *Biometrika* for its library. The Faculty begs you not to refuse to send the journal to the Cabinet or to whom will indicate the agent.

3. Romanovsky's correspondence with Fisher

Letter No. 8. Romanovsky – Fisher 9.10.1929

I would be very glad to see you and to visit the Rothamsted station[21] …

Letter No. 9. Fisher – Romanovsky 10.10.1929

… perhaps you can visit us on Monday, October 14th. And if it suits you stay with us for a while.

Letter No. 10. Romanovsky – Fisher 18.10.1929

… I shall visit you again on Monday if you do not object it. I shall be glad to see you and all your friends … again.

Letter No. 11. Romanovsky – Fisher 28.10.1929

There is [here] in Paris a friend of mine who was some years ago a lecturer of polytical economy at the University of Tashkend and now, as an emigrant, allready two years, lives in Paris. He is an able scientist who has published two books[22]… These books were much praised as original and novel in views and working out of data. They are written, I must add, quite not in an orthodoxal Marxian manner. The name of the author is Alexander [Petrovich] Demidoff. He is now 36 years old and is a bearer of quota immigration visa for U.S.A. In

Paris he was earning his, his wife and his little daughter's life serving at a bank (very little gaining, I must add) and in leisure minutes he was working in the libraries of Paris on a big problem: the present economical state of England, its development and its future. His views, as far as I can judge, are very interesting and in some points original.

Now I come to the aim of this letter. Mr. Demidoff lives in very difficult conditions and has no prospects to finish and publish his work just mentioned. Perhaps, you and Prof. Hotelling[23] could help him to receive the Rock[e]feller's stipend for a year in order he could quietely work on his problem? He will send you a prospect of his work and I ask you and Prof. Hotelling to read it and to write to Mr. Demidoff if he can hope to obtain a help and how he can act further for this aim. You and Prof. Hotelling have many american friends and if you find it to be possible you can help him very much.

There is yet a very important point. If you and Prof. Hotelling resolve to help to Mr. Demidoff, please do not remember at all my name, for it can end with my emprisonment by GPU (Chief Political Administration of Soviet Russia, the most dreadfull and mightfull organisation in the present Russia). It is a crime, and a very heavy one, from its point of view my endeavouring to help an emigrant, although there is no politics in my action but only the desire to help to an able scientist who can do much important work being placed in good conditions. Act if thus as if you knew only Mr. Demidoff's prospect and his book, which, I hope, he will also send you.

To morrow I go to Berlin and from there to Moscow. My best remembrance from my voyage abroad is and will be Rothamsted Experimental Station and the men I had known there. … Please do not write to me in Russia on Mr Demidoff and read this letter to Prof. Hotelling.

Letter No. 12. Romanovsky – Fisher 22.12.1929

Romanovsky begins with season's greetings and continues:

I have received your Christmas card and your last [latest] paper also and bring you my thanks for them. Write me, if you please, the name of the author and the title of the book on statistics of enginiring I have seen at you: I shall purchase it for me. I beg you also to write me a short description of the scheme of the increasing of precision of plot experiments I have seen in your laboratory at the Rothamsted Station. I have forgotten it and some agronomical researchers here are very interested with it.

Letter No. 13. Fisher – Romanovsky 6.1.1930

The title of the book you refer to is [Fry (1928)]. I am not sure what you refer to about Plot experimentation. You have my book[24] and various papers on the subject. There is in my laboratory a diagram illustrating the logical position of the three principles of plot experimentation[25]…

Letter No. 14. Romanovsky – Fisher 22.3.1930

I have received some papers of you and your friends and perused them with great pleasure. Twenty years ago I was much occupied with the theory of the prime numbers and published some papers on them. Then, having no table of the prime numbers under my disposition, I prepared it myself, up to 2000, with the same principle as it is constructed in your note on the sieve of Eratosthenes [Fisher (1929a)]. Thus it procured me much pleasure to see your note and to know that you also have not escaped the fascinating power of the prime numbers – one of the most wonderful things in the world.

I am much occupied in our university and have still no time to study your and Craig's papers on the theory of moments[26]. I shall do it in summer. Some rare hours of leisure I have spent on the investigation of a class of integral equations which I have found in connection with further development of Markoff's chains (you have a note on them)[27]. These equations seem to be novel and I have developped their theory analogous to that of Fredholm. I shall do a communication on this theory and on further generalisations of Markoff's chains this summer at the Congress of the Mathematicians of USSR which will take place in Kharkov

Just in this moment I have received a letter from the Organizing Committee of the Congress (I am one of its members) and there stands that many foreign mathematicians (Borel, Hadamard, Lichtenstein, Levi-Civita, Blaschke, Cartan, Denjoy, Montel, Mandelbroit and others) will take place at the Congress and read communications. Perhaps you should also come and read on your researches in the math. statistics? It would be splendid to meet you at this Congress!

Letter No. 15. Fisher – Romanovsky 11.4.1930

I am afraid I cannot manage the trip to Kharkoff in June next as I seem to have in other ways a very busy year in front of me. Many thanks for the suggestion. I hope you will have a successful meeting.

I am glad to hear of the new class of integral equations; it is a subject that I admire from a distance[28]. The combinatorial procedure for evaluating the higher moments of algebraic statistics may, however, be intimately of interest in this regard. It was a long while before I could see the reason for all the simplifications which the method introduces. Indeed it is still a mystery to me why the algebraic coefficients corresponding to the "patterns" should be so simple.

I worked out the other day the coefficients corresponding to the three symbolic figures … [Fig. 1] which are all that are wanted (in the case of a normal population) for anything like the 4th semi-invariant of the distribution of k_4, such as (with two variates) any 4th order semi-invariant of the simultaneous distribution of k_{40}, k_{31}, k_{22}, k_{13}, k_{o4}[29]. Well, the patterns have eight rows each, and the number of separations of eight parts is very large, so that it was very heavy work before I had the coefficients; but when all is done they are simply

$$n(n + 1)\,(n^4 - 8n^3 + 21n^2 - 14n + 4)/(n - 1)^3\,(n - 2)^3\,(n - 3)^3,$$

$$n^2\,(n + 1)^2\,(n - 2)\,(n - 3)/(n - 1)^3\,(n - 2)^3\,(n - 3)^3,$$

$$n\,(n + 1)\,(n^4 - 9n^3 + 23n^2 - 11n + 4)/(n - 1)^3\,(n - 2)^3\,(n - 3)^3.$$

So that letting N_1, N_2, N_3 stand for these three expressions the 4th semi-invariant of k_4 is simply

$$4 \cdot 12^3\,(9N_1 + 8N_2 + 36N_3)k_2^8$$

and, for example, for the 4th semi-invariant of k_{22} in the bivariate problem, we have only to subdivide the numerical factors by supposing that the four rods which meet at each point are two black and two red, and enumerating the number of ways of linking them up with 0, 2, 4, 6, or 8 black-red junctions (as opposed to black-black or red-red junctions which must be equal in number and supply the factors k_{20}, k_{02}). Thus in every problem the algebraic coefficients are the same, and they are so simple that one feels that one ought to be able to write them down by inspection of the pattern, or of its symbolical diagram.

I am glad you liked the sieve. I feel that Eratosthenes has been too long exposed to the patronising remarks of his critics!

Letter No. 16. Romanovsky – Fisher 28.10.1930

I am very thankful for the copies of your works and of your collaborators and assistents. They are regularly received here and are very interesting and important for me, especially since I am more closely connected with the Cotton Research Institute organised here, in Tashkend. The works of the Rothamsted Experimental Station and your methods for field experiment are of much aid for me and I am propagating them very zealously.

Much time is lost in performing my professional duties and so I am almost unable to write on my personal researches. I am much advanced in the investigation of phenomena connected in chains and depending from random (Markoff's chains as I name them) and the results are very interesting from the point of time series. My intention is to write a memoir on these results but all my time I am spending in new researches: it is not very pleasant to lose it in writing down acquired results. It would be splendid if I could to spend all my time only in the quiet work in libraries like I did it past year in Berlin, Paris and especially in London (The British Museum is the most beautiful and comfortable library).

What are you working on?

Letter No. 17. Fisher – Romanovsky 14.11.1930

I am very glad to hear that my reprints have been safely received, and I shall be much interested to see more of your own researches as they are published. I have long intended to gather together the most important mathematical researches of recent years in a book on Mathematical Statistics, but so far I have not found the time to make any real progress with this task.

I am very glad you found the Library of the British Museum convenient to your work, and hope that you may again have an opportunity to visit us, and carry out the more substantial researches you have in mind.

My family is well. I hope Mrs Romanovsky and your daughter are also in good health[30].

Letter No. 20. Romanovsky – Fisher 22.12.1931

Season's greatings.

Letter No. 21. Fisher – Romanovsky 5.1.1932

"Belated" season's greatings. He continues: I sincerely hope that your country may in time reap the rewards of the great efforts and sacrifices which are being made.

Letter No. 22. Romanovsky – Fisher 19.1.1934

I am very glad to congratulate you with the professorship in the London University. Your field of activity is now widening and I hope it will be to the benefit of the science and yourself.

I would be very content if you send me the prospects or the plans of the researches of the laboratories which are now under your guidance. It interests me very much as also all what concerns the organisation of your laboratories.

Our Physico-Mathematical Research Institute is developping steadily and I hope very soon to send you the proofs thereof: the offprints of papers, mine and of my collaborators, made in the Institute.

Are you now living in London or, as before, in Harpenden?

Letter No. 23. Fisher – Romanovsky 5.2.1934

I am very glad to have your letter, and to see your handwriting again. I am glad to hear of the Physico-Mathematical Research Institute in Tashkent. I have recently been seeing some of the indirect effects of your activity in the improvement of methods of experimentation in Cotton trials. I suppose the new Institute will be concerned with the technology of cotton spinning.

My new department will, I am afraid, only be slowly organised. I want to give Students of Eugenics working here an opportunity to acquaint themselves thoroughly with modern genetical knowledge in animal material. I find I have a nice animal house, and have been engaged since I have been here in getting adequate equipment for the photographic studio, and now for the Laboratory. All the equipment here was very old and bad. I hope later to have a Biological Assistant but he is not yet appointed, and at present I have only two Biological Voluntary workers.

The department of Statistics has been separated from the Galton Laboratory[31], which saves me from having to organise the Statistical teaching, but has the bad effect that Students have not always confidence enough to ask my advice on Statistical points when they need it. I have been lecturing on the Logic of Experimentation, and also on Quantitative Inheritance, and a very good class chiefly of members of the staff have been coming to the lectures, but I am afraid I am not an experienced lecturer and the preparation of the lectures has taken more time than I ought to give.

I shall continue to live in Harpenden, as the new Laboratories are conveniently accessible from there, and I hope some day to welcome you, or perhaps a Student from your University in the Galton Laboratory.

Letter No. 24. Romanovsky – Fisher 4.12.1935

I would be very obliged to you if you indicated me how are established two approximate formulae, p. 221 of your *Statistical methods* …[32] I am also puzzled why you use, in the analysis of variance,

$$z = 1/2 \ln (s_1^2/s_2^2)$$

instead of s_1^2/s_2^2. Many thanks in advance for the answers.

Season's greetings follow.

Letter No. 25. Fisher – Romanovsky 20.12.1935

So far as I remember I obtained the approximation on page[33] for the test of significance of z where both n_1 and n_2 are large by obtaining the moments of the distribution of z, or rather its cumulants, from its characteristic function. I forget the details, but clearly the factor $[(1/n_1) - (1/n_2)]$ is a simple allowance for the third moment, while the first term is derived from the normal distribution.

I had a good many reasons for using z instead of some function of it in the test of significance in the analysis of variance. One important reason was that in order to make a compact table it is necessary that the test value should be well interpolated by what I call asymptotic interpolation using the reciprocals of the numbers of degrees of freedom and this is more true of z than of any other simple function. A second point is that half the tabulation is saved by the fact that reversing the sign of z and interchanging n_1 and n_2 we have the 5% and 1% points at the opposite ends of the distribution. Finally, the close analogy between interclass and intra class correlations is paralleled by that of the values of z obtained from r by the same transformation. The advantages of this transformation I have set out in the book.

Please accept my kind wishes for yourself and family during the coming year. I am sending a copy of a recent book of mine, which may, I hope, interest you.

Letter No. 26. Romanovsky – Fisher 23.1.1936

Many thanks for your excellent and very interesting book [Fisher (1935)]. I shall read it and write a note on it like one I have written on your *Methods for Research Workers* and sent you some time ago. Have you received it[34]?

In some days I shall send you my last [latest] memoir [15].

Letter No. 27. Fisher – Romanovsky 1.2.1937

I am very obliged for the cuttings of the two reviews of the *Design of Experiments* which you were good enough to send me. I am having them translated into English.

Letter No. 28. Romanovsky – Fisher 15.10.1937

One of my pupils, V. Peregoodoff, … has translated in Russian your *Design of Experiments*[35]. The translation will soonly be published and it is intended to accompany it with your portrait in the frontispice. V. Peregoodoff does not dare to beg you to send him it and asked me to write to you. I do it with great pleasure for I appreciate your book very highly. We all shall be very thankfull to you. …

I have read a conference on your book in the Society of Naturalists at our university and now prepare it for publishing.

My time is now very occupied (I am now dean of the physico – mathematical faculty of our university) and I work very little in statistical research and I publish still less. But I hope to publish soonly some of my last [latest] researches in the theory of probabilities and in the math. statistics. At the end of this year will be printed my book [22], a big volume containing much of the recent researches, with demonstrations, more a mathematical work than practical. I shall be glad to send you an exemplar.

Letter No. 29. Fisher – Romanovsky 1.11.1937

I am delighted to learn that one of your pupils has translated my *Design of Experiments*, and, naturally, wish the greatest success to this publication. Nevertheless your request for a photograph does somewhat embarrass me, for the following reason.

I understand that the Soviet Government does not legally recognise the copyright laws of other countries, although, in fact, they make arrangements with the publishers who possess these copyrights. I do not think my publishers, … have been approached, or have given permission for this translation, and in these circumstances I cannot myself cooperate in what they may regard as an infringement of their rights.

I have reason to believe that, if the Department concerned, approached [the publishers] with an offer of no great magnitude, even though payable only in internal currency, they would be satisfied with this formal acknowledgement of their rights, and would at my request not stand in the way of what may be a valuable publication. Would you, or Mr Peregoodoff, take the matter up with the Russian authorities, in which case I should be happy to cooperate.

I am glad to hear that your services in University organisation are now being recognised, even though the additional work may withdraw your time from mathematical statistics. I should very much indeed like to possess a copy of your book when it is published.

Letter No. 30. Romanovsky – Fisher 14.10.1938

I have received from your editors your *Tables*[36]. Many thanks for this valuable presentation. Of course, I shall write and publish a review in some [in one?] of our journals, for I appreciate very much your new statistical tables. I hope to do it as soon as I can (my duties have increased very much: I am now elected as a member of the Supreme Council [Supreme Soviet] (Parliament) of our republic).

Notes

1. Romanovsky's letters to Pearson are kept at Special Collections, Library Services, University College London (Pearson papers 831/3); his correspondence with Fisher is in the Barr Smith Library, University of Adelaide. I myself (Sheynin 1996, pp. 50 – 53) published two letters from Chuprov to Romanovsky (1923 and 1925) as well as Chuprov's later letters concerning him. The Archive of the Russian Academy of Sciences (*Fond* 173, inventory 1, *delo* 17, No. 1) keeps Romanovsky's letter to Markov of 2.11.1916. Taking into account Markov's criticism, Romanovsky revised the proof of one of his theorems, enlarged on his considerations and expressed his desire to publish his manuscript in Petrograd (Petersburg). Markov's answer is not known but the paper in question appeared many years later [12].

2. B&M (pp. 98 – 101) describe Romanovsky's ideas [1] on scientific progress and social phenomena. When mentioning his admiration of Mendel and eugenics, they justly remark that his conclusions were still possible [to publish] in the beginning of the 1920s.

3. Romanovsky (Ibidem, pp. 225 – 226) attributed to Fisher the Mises concept of probability. At the very least, Fisher was indeed an empiricist.

4. Earlier Romanovsky [7, p. 1088] called Chuprov "the greatest Russian statistician".

5. Recall, for example, Fisher (1934, p. 23) who wrongly declared that the method of least squares was a corollary of the principle of maximum likelihood.

6. Kolmogorov (1947, p. 63) favourably cited [18] as well as the Western school of statistics.

7. This book (Chebotarev 1958) was written on the level of the mid-19th century with some elements of linear algebra and mathematical statistics having been added. On p. 579 we find that the Ptolemy system of the world "for 14 centuries held mankind in ideological captivity".

8. In present-day terminology, one-dimensional and bivariate populations. The expression "equation of distribution" (below in this Letter) is also dated. Romanovsky again mentions the same manuscript in Letters 3 and 4. It did not appear in *Biometrika*, but Chuprov (Sheynin 1996, p. 50) later communicated its modified version to *Metron.* Indeed, Romanovsky shortened it and added some new material, see Letter 3. The additions concerned the issue described below in this Letter (see Letter 4).

9. Probable errors calculated for a sample are random variables and do not therefore possess true values.

10. Instead of x and y read x_i and y_i respectively.

11. The Pearson article (1925, p. 181) contains only the last two symbols; they pertained to coefficients of correlation.

12. Such a textbook appeared only in 1939.

13. See beginning of Letter 1.

14. Pearson (1925, p. 199) had indeed indicated:

> *Writing without knowledge of the papers in Biometrika … and naturally without knowledge of my present paper, Professor Romanovsky had reached, dealing only with the algebraic side, many of the published results and certain additional ones. While willing to publish the latter, the present cost of printing prohibited the reproduction of much work already published or about to be published in this Journal. … I sent him a proof of this paper and asked him to cable if he were willing that I should add under the title his name to my own. … He [his Letter 4?] is satisfied with the statement that many results contained in the present and earlier papers have also been obtained by him quite independently and by a different method. I trust for the sake of his additional results that his paper may shortly be published elsewhere.*

And here is Romanovsky's *Remark* [4, p. 208] translated from its Russian version:

After completing this article, I received from Prof. Pearson the proofs of his paper (1925). … It contains some results of my present article derived by means of an utterly different method.

15. The article Pearson (1925), also see Note 14 and Letter 6.

16. The expression "the mean … tends to be normal" is unfortunate. Concerning the indicated findings see [18]. Romanovsky's discovery of the relation between the moments and the Taylorian series is unknown to me; see however Delsarte (1930) where Romanovsky is not cited.

17. See [22, book 2, pp. 47 – 50].

18. Romanovsky quoted these few words from one of his papers [6].

19. Also see Letter 7. In addition to [10], Romanovsky later published two more papers in *Biometrika*, in 1933 and 1936. His first articles there appeared in 1923 and 1924. In the second of these he [3] studied a generalized system of the Pearsonian curves and Pearson added there his remarks.

20. The manuscript was published in Russia [8].

21. The experimental station near Harpenden. Fisher worked there as statistician for 14 years, from 1919 to 1933.

22. The American *National Union Catalog pre-1956 Imprints* mentions three books by Demidov (including those cited by Romanovsky) published in Russia and one more which appeared in Paris in 1931.

23. Harold Hotelling (1895 – 1973), an American statistician and economist. Corresponded with Fisher from about 1927 and worked for a few months in 1929 at Rothamsted.

24. Evidently an earlier edition of Fisher (1934).

25. Fisher wrote out five terms: 1) Replication; 2) Random distribution; 3) Local control; 4) Validity of estimate; of error; 5) Diminution of error; but he numbered only the three first ones. He also indicated by arrows the directions 1 – 2, 1 – 3, 1 – , 1 – 5, and 2 – 4.

26. Apparently Fisher (1929b; 1930) and Craig (1930) if only Fisher (1930) was already published and available.

27. The expression *Markov chains* dates back to Bernstein (1926, §16) who called them *chaines de A. Markoff*. Romanovsky used it in 1929 and 1930. I did not find Fisher's note mentioned by Romanovsky.

28. See however Letter 6.

29. Here is Fisher's marginal note: "There seems to be 34 such bivariate formulae". The so-called Fisherian *t*-statistics k_r (x_1; x_2; …; x_n), $r = 1, 2, \ldots$ are the most general homogeneous polynomials of degree r with mean values Ek_r equal to the r-th cumulants of the appropriate sample distribution. Kendall (1963/1970, p. 442) called Fisher (1929b) "the most remarkable paper he ever wrote" and testified on his next page that Fisher "was never able to explain … to me how he thought of these results". Fisher's letter to Romanovsky is interesting, in particular, in this connection, but it only partly explains his ideas. In his paper (1929b) he investigates, for the same purpose, the partitions of the numbers r (he applies the term *separation* in his letter). Even Wilks (1962, §8.2c) refers readers to special literature on these statistics. Also see [22, book 1, pp. 88 – 99]. Finally, I note that Fisher had not sufficiently explained his figures either. He mentioned points and junctions without distinguishing between them, see below. He adduced other figures in his paper (1929b) with no explanation at all.

30. Romanovsky's daughter died in 1925 (B&M, p. 85). The next two letters are those already published, see §1.

31. The Galton Laboratory of National Eugenics was established in 1907 at London University. In 1933 Fisher replaced Pearson in the new faculty of (mathematical) statistics, but it was E. S. Pearson who taught statistics. Fisher was left with eugenics and biometry and in actual fact (not as he wrote to Romanovsky) that situation disappointed him (Bartlett 1978, p. 353).

32. Read Fisher (1935, p. 221). The formula below is the known Fisher z-transformation which he introduced in 1925 in *Biometrika*; r is the sample coefficient of correlation. In his book of 1935 Fisher introduced the formula on p. 200 and indicated its merits on p. 207. In

Letter 25 he put forward additional pertinent considerations and Romanovsky [14, p. 126] apparently agreed with him. The quotient s_1^2/s_2^2 (in standard notation) is indeed in general usage in analysis of variance but

$$(1 + r)/(1 - r) \neq s_1^2/s_2^2.$$

Later on Romanovsky [22, book 2, p. 21] applied 1/2 $\ln(s_1^2/s_2^2)$ in the same analysis.

33. A blank in the original text.

34. See §1 for the reviews mentioned in Letter 27.

35. See §1.

36. Fisher & Yates (1938).

Acknowledgements

The libraries mentioned in Note 1 kindly permitted me to publish Romanovsky's correspondence. His letters to Pearson have appeared in a preliminary manner in a microfiche collection (mostly of translations) Bernstein et al (1998, pp. 233 – 239), but the publisher has no copyright to it.

I am grateful to Professors Herbert A. David and J. Pfanzagl for indicating some statistical sources and to Dr. A.L. Dmitriev for sending me photostat copies of some of Romanovsky's papers.

Bibliography

Abbreviation

M = Moscow; L = Leningrad; R = in Russian

V.I. Romanovsky

1. Statistical Weltanschauung. *Voennaia Mysl*, No. 1, 1921, pp. 59 – 76. (R)

2. Theory of probability and statistics. On some newest works of Western scientists. *Vestnik Statistiki*, vol. 17, No. 4/6, pp. 1 – 38 and vol. 18, No. 7/9, pp. 5 – 34, 1924. (R)

3. Generalization of some types of the frequency curves of Prof. K. Pearson. *Biometrika*, vol. 16, 1924, pp. 106 – 117.

4. On the moments of standard deviations and of correlation coefficient in samples from normal population. *Metron*, vol. 5, 1925, pp. 3 – 46.

5. Sur certaines espérances mathématiques et sur l'erreur moyenne du coefficient de corrélation. *C.r. Acad. Sci. Paris*, t. 180, 1925, pp. 1897 – 1899.

6. Sur la distribution des écarts quadratiques moyenes dans les observations sur les quantités à distribution normale. Ibidem, pp. 1320 – 1323.

7. On the distribution of the arithmetic mean in series of independent trials. *Izvestia Akad. Nauk SSSR*, vol. 20, 1926, pp. 1087 – 1106. (R)

8. On the distribution of the regression coefficient in samples from normal population. Ibidem, pp. 643 – 648.

9. Theory of statistical constants. On some works of R.A. Fisher. *Vestnik Statistiki*, vol. 25, No. 1, 1927, pp. 224 – 266. (R)

10. Note on orthogonalising series of functions and interpolation. *Biometrika*, vol. 19, 1927, pp. 93 – 99.

11. On the criteria that two given samples belong to the same normal population. *Metron*, vol. 7, No. 3, 1928, pp. 3 – 46.

12. On the moments of means of functions of one and more random variables. Ibidem, vol. 8, No. 1/2, 1929, pp. 251 – 290.

13. On the newest methods of mathematical statistics applied in agricultural experimentation. *Sozialistich. Nauka i Tekhnika*, No. 3/4, 1934, pp. 75 – 86. (R)

14. Review of Fisher (1934). *Sozialistich. Rekonstruktsia i Nauka*, No. 9, 1935, pp. 123 – 127. (R)
15. Recherches sur les chaines de Markoff. *Acta Math.*, t. 66, 1935, pp. 147 – 251.
16. Review of Fisher (1935). *Sozialistich. Nauka i Tekhnika*, No. 7, 1936, pp. 123 – 125. (R)
17. Mathematical statistics. *Bolshaia Sovetskaia Enz.* (Great Sov. Enc.), 1st ed., vol. 38, 1938, pp. 406 – 410. (R)
18. *Matematicheskaia Statistika* (Math. Statistics). M. – L., 1938. (R)
19. Review of Fisher & Yates (1938). *Sozialistich. Nauka i Tekhnika*, No.2/3, 1939, p. 106. (R)
20. *Osnovnye Zadachi Teorii Oshibok* (Main Issues in Theory of Errors). M. – L., 1947. (R)
21. On mathematical treatment of observational results. *Trudy Moskovsk. Inst. Inzhenerov Geodesii, Aerofotosiemki i Kartografii*, vol. 15, 1953, pp. 17 – 20. (R)
22. *Matematicheskaia Statistika* (Math. Statistics), Books 1 – 2. Tashkent, 1961 – 1963. (R)
23. *Izbrannye Trudy* (Sel. Works), vol. 2. Theory of probability, statistics and mathematical analysis. Tashkent, 1964. (R) This is a collection of reprints and translations. In particular, it includes [4; 6; 8; 11; 12].

Other Authors

Bartlett, M.S. (1978), Fisher. *Intern. Enc. Statistics*, vol. 1. Editors, W.H. Kruskal, Judith M. Tanur. New York – London, pp. 352 – 358.

Bennett, J.H., Editor (1990), *Statistical Inference and Analysis. Selected Correspondence of R.A. Fisher.* Oxford.

Bernstein, S.N. (1926), Sur l’extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Math. Annalen*, Bd. 27, pp. 1 – 59. Russian translation 1944 reprinted in vol. 4 of his *Sobranie Sochinenii* (Coll. Works), 1964. No place, pp. 121 – 176.

Bernstein, S.N. et al (1998), *From Markov to Kolmogorov.* Microfiche collection of papers translated from Russian by O. Sheynin. Deutsche Hochschulschriften 2514. Egelsbach.

Bogoliubov, A.N. & Matvievskaia, G.P. (1997), *Romanovsky.* M. (R)

Chebotarev, A.S. (1951), On mathematical treatment of observational results. *Trudy Moskovsk. Inst. Inzhenerov Geodezii, Aerofotosiemki i Kartografii*, vol. 9, pp. 3 – 16. (R)

--- (1953), Same title. Ibidem, vol. 15, pp. 21 – 27. (R)

--- (1958), *Sposob Naimenshikh Kvadratov* (Method of Least Squares). M. (R)

Craig, C.C. (1930), The semi-invariants and moments of incomplete normal and Type III frequency functions. *Ann. Math.*, ser. 2, vol. 31, pp. 251 – 270.

Delsarte, J. (1930), Sur la détermination des coefficients du Taylor d’une fonction de probabilité dont on connaît les moments. *C.r. Acad. Sci. Paris*, t. 191, pp. 917 – 918.

Fisher, R.A. (1925), Sur la solution de l’équation intégrale de Romanovsky. *C.r. Acad. Sci. Paris*, t. 181, pp. 88 – 89.

--- (1929a), The sieve of Eratosthenes. *Math. Gaz.*, vol. 14, pp. 564 – 566.

--- (1929b), Moments and product moments of sampling distributions. *Proc. Lond. Math. Soc.*, ser. 2, vol. 30, pp. 199 – 238.

--- (1930), The moments of the distribution for normal samples of measures of departure from normality. *Proc. Roy. Soc.*, ser. A, vol. 130, pp. 16 – 28.

--- (1934), *Statistical Methods for Research Workers.* Edinburgh – London. First ed., 1925. The 14th ed. appeared in 1970.

--- (1935), *Design of Experiments.* Edinburgh. Not less than seven later editions.

--- (1958), Russian translation of Fisher (1934). Translated by V.N. Peregudov. M.

Fisher, R.A. & Yates, F. (1938), *Statistical Tables for Biological, Agricultural and Medical Research.* London – Edinburgh. Not less than six later editions.
Fry, T. (1928), *Probability and Its Engineering Uses.* Princeton, 1965. Russian tanslation, 1934.
Kendall, M.G. (1963), Fisher. *Biometrika*, vol. 50, pp. 1 – 15. Reprinted (1970) in *Studies in the History of Statistics and Probability*, vol. 1. Editors, E. S. Pearson & M. G. Kendall. London, pp. 439 – 453.
Kolmogorov, A.N. (1947), The role of Russian science in the development of the theory of probability. *Uchenye Zapiski Moskovsk. Gosudarstv. Univ.*, No. 91, pp. 53 – 64. Translated in *Probability and Statistics. Soviet Essays.* Berlin, 2005, pp. 68 – 84. Also at www.sheynin.de.
--- (1948), Main issues in theoretical statistics (abstract). In Vtoroe (1948, pp. 216 – 220). Translated in *Probability and Statistics. Russian Papers of the Soviet Period*. Berlin, 2005. Also at www.sheynin.de.
Pearson, E.S. (1966), The Neyman – Pearson story: 1926 – 1934. In *Festschrift für J. Neyman.* Reprinted together with Kendall (1963), pp. 455 – 477.
Pearson, K., Editor (1914), *Tables for Statisticians and Biometricians*, vols 1 – 2. London.
--- (1925), Further contributions to the theory of small samples. *Biometrika*, vol. 17, pp. 176 – 199.
Sarymsakov, T.A. (1948), Statistical methods and issues in geophysics. In Vtoroe (1948, pp. 221 – 239).
Sheynin, O. (1996), *Chuprov.* Göttingen. Translated from Russian (1990).
--- (1998), Statistics in the Soviet epoch. *Jahrb. f. Nationalökonomie u. Statistik*, Bd. 217, pp. 529 – 549. Russian translation (2001) in *Istoriko-Matematich. Issledovania*, vol. 6 (41), pp. 179 – 198.
Smit, Maria (1930), The dialectics of quantity (1927). In author's collected papers *Teoria i Praktika Sovetskoi Statistiki* (Theory and Practice of Soviet Statistics). M., pp. 7 – 29. (R)
Soper, H.E. (1922), *Frequency Arrays.* Cambridge.
Tschuprov, A.A. (1924), Review of Soper (1922). *Nordisk Statistisk Tidskrift*, Bd. 3, pp. 414 – 417.
Vtoroe (1948), *Vtoroe Vsesoiuznoe Soveshchanie po Matematicheskoi Statistike* (Second All-Union Conf. Math. Statistics). Tashkent. (R)
Wilks, S.S. (1962), *Mathematical Statistics.* New York.
Zbornik (1939), *Zbornik Posviashchennyi 30-Letiu Nauchnoi i Pedagogicheskoi Deiatelnosti V.I. Romanovskogo* (Collection Dedicated to the Thirtieth Anniversary of Romanovsky's Scientific and Pedagogic Work). Central Asian State Univ., ser. math. Issues 19 – 32 (separate pagings) with Introductory note. Tashkent. (R)