Studies

in the History of Statistics and Probability

Vol. 9. Collected Translations

Many papers concern A. A. Markov

Compiled and translated by Oscar Sheynin

Berlin

2017

Contents

Introduction by the compiler
Notation

Notation **S, G,** I means, that either the appropriate rare source, or an English translation of a Russian source is available on, and downloadable from my cite www.sheynin.de which is being copied by Google, see Oscar Sheynin, Home.

General comments on most items

i. Huygens was not really interested in population (or medical) statistics. Indeed, he applied the Graunt table but thought that it was based on exact observations (beginning of Supplement No. 1). Nevertheless, he introduced the probable and expected life and correctly indicated their differing fields of applications. Incidentally, the correspondence of brothers Huygens once more testifies that bets on human life had been quite common. The introduction of an integral distribution function, although in an unusual form (Note 14), was only applied after him in 1829 (Sheynin 2009, § 8.2).

Huygens wrongly solved one of his problems (Note 17), although for those times his mistake seems to be understandable. Another mistake (Note 7) was perhaps just an oversight. Then, when investigating problems of survivorship he applied variable expectations rather than constant probabilities (cf. Korteweg 1920, p. 135) which greatly complicated the calculations. Anyway, in 1709 Niklaus Bernoulli (Todhunter 1865, pp. 195 – 196) managed to solve one of his problems much easier by applying probabilities: Given:

b men who will all die within a years and are equally likely to die at any instant within this time. Required: *the probable duration of the life of the last survivor*.

Answer: $ab/(b + 1)$.

ii. The author scrupulously studied his subject but had not provided information about some scholars which he mentioned and made an unbelievable mistake (see Note 4). He did not recognize translations as scientific work, but I hold that a translation with proper commentaries is no less scientific. To extrapolate: contrary to the established opinion of the scientific community, I am convinced that **honest and knowledgeable** reviewing is a most serious and most important scientific work, see Introduction to my *Black Book* (**S, G,** 80). That both main abstracting journals are barely available I regard as an opprobrium of science.

iii. My Notes show that Wittstein's deliberations were superficial. He restricted his study to Germany (but forgot Chr. Bernoulli). Apart from Graunt and other English scientists whom he neglected (Note 2), he missed De Moivre, Simpson and Wargentin (Nordenmark 1929) and said nothing about Guillard (1855).

The insistence on the introduction and development of mathematical statistics saves his contribution; he also was at least one of those few who coined that term. Nevertheless, we ought to mention Laplace (1814/1995, p. 62): *Let us apply to the political and moral sciences the method based on observations and calculations*. And he himself certainly did apply it!

iv. In his first sections Gnedenko repeatedly stressed disappointment and scepticism allegedly having been felt in the West about the theory of probability, and once he connected it with the rejection of the application of probability to the administration of justice. The unnamed author of the quoted and generally known pertinent statement was Mill (1843/1886, p. 353). It is therefore opportune to quote Gauss (1841, see his *Werke*, Bd. 3, pp. 201 – 204):

The theory of probability can provide the lawgiver a clue for determining the number of witnesses and judges.

This is exactly what Poisson (1837) did. Poincaré and Bernstein, who had surely never read Poisson's contribution, summarily denied such applications. But to return to the main statement (disappointment etc.).

My explanation necessarily begins from afar. Jakob Bernoulli proved the direct LLN (statistical probability tends to its theoretical counterpart), but tacitly applied the inverse form of that law; De Moivre plainly stated that both forms were equally precise, and only Bayes proved that the inverse form of the LLN is less precise. This proposition is indeed qualitatively seen at once: in both forms of the law we have the same trials, but theoretical probability is only known in the direct form.

I conclude that Bayes completed the first stage of the theory of probability during which, as it is additionally seen, the three scholars regarded probability as a discipline belonging to pure science. Laplace, however, transferred probability to applied science, and Poisson (1837, § 84) followed suit:

There exists a very high probability that these unknown chances little differ from the ratio …

The Laplace – Poisson stage was necessary, witness the scientific folklore:

Pure science attains the possible by rigorous means; applied science achieves the necessary by reasonable means.

And now I say: there was no *disappointment* in the West: probability had been applied to most various branches of knowledge. True, the situation changed with time (Kamke 1933, p. 14):

In 1910, a bon mot could be heard in Göttingen: Mathematical probability is a number between zero and unity about which nothing else is known.

The return of probability to the realm of pure science apparently began with Lévy (1925), see Cramér (1976, p. 516):

Lévy made the first attempt to present the theory as a connected whole, using mathematically rigorous methods. It [his book] *contained the first systematic exposition of the theory of random variable, their probability distributions and their characteristic functions*.

Only the axiomatic approach concluded this return. However, statistics, excluding stochastic processes, does not apply it and in this sense remains an applied science.

My next point. Gnedenko repeatedly stressed Chebyshev's rigorous attitude to mathematical reasoning, but here is another opinion.

Bernstein (1945/1964, p. 425):

Especially towards the end of his life Chebyshev deviated from the clearness of formulation and rigour of proofs.

Gnedenko himself (§ 3.2) made a similar remark.

And Kolmogorov's statement about Chebyshev's use of the *full power* of stochastic concepts (quoted by Gnedenko in § 2) was wrong: the *full power* of the notion of random magnitude is possibly not attained even now.

And, for the first time ever, Novikov (2002, p. 330) noted the sad fact: *Chebyshev was a pathological conservative*. And not only he, but his most eminent students, Markov (A. A. Youshkevich 1974, p. 125) and Liapunov (1895/1946, pp. 19 – 20), as well. I quote the last-mentioned scholar:

The partisans of Riemann's extremely abstract ideas delve ever deeper into function-theoretic research and pseudo-geometric investigations […]. *These investigations were recently often connected but have nothing in common with Lobachevsky's geometric research* […].

In 1871, Klein had presented a unified picture of the non-Euclidean geometry!

In § 1 Gnedenko stated that De Moivre had only proved his classical theorem for a particular case and Laplace generalized it. This mistake goes back to Todhunter (1865); actually, De Moivre noted that generalization was easy and, moreover, the long title of his pertinent pamphlet of 1733 included the expression *binomial* $(a + b)^n$. As to Laplace, Gnedenko had thus misunderstood his achievement.

Gnedenko's description of the results of Laplace and Gauss in the theory of errors is all wrong. It is a curious fact, only partly explained by the extreme difficulty of understanding Gauss' main memoir of 1823 on that theory, that scholars, including Chebyshev and Fisher, had hardly studied Gauss. On this subject see Sheynin (2009, §§ 7 and 9; 2010).

That problems of artillery firing led Poisson to the generalization of the Bernoulli theorem is wrong, see Poisson (1837, § 81). Overriding my objection, Gnedenko (Gnedenko & Sheynin 1978/2001, p. 211) stated that those problems had been important for the development of the theory of probability. This conclusion, undoubtedly borrowed from Kolmogorov's similar opinion, had no evidence to support it.

In that section, finally, Gnedenko connected the Laplace – Poisson period with a *philosophical comprehension* of probability theory. Kolmogorov (1947), however, had attributed that achievement to an earlier period, and Gnedenko, who would have never contradicted his mentor, likely misread him.

In § 2 Gnedenko stated that the contributions of Daniel and Niklaus Bernoulli and Euler remained alien to Russian science and culture. Niklaus (the elder brother) died soon after moving to Russia, and the other two great scientists had for a long time *represented* Russian mathematics and natural science.

Lobachevsky attempted to ascertain which geometrical system governed the universe and *developed the theory of errors on a sphere*. I am unable to understand this expression. Lobachevsky applied the theory of errors to astronomical observations, and so did many

astronomers before him. Then, Gnedenko (1949) described Lobachevsky's attempt but had not noticed that the precision of the astronomical instruments should have been much higher. This is all the more strange since Kagan (1944, p. 69) had earlier (indirectly) stated the same.

And in his paper of 1842 Lobachevsky derived the law of distribution of a finite sum of mutually independent random variables, a problem solved by Simpson, Lagrange and Laplace (Sheynin 1973, p. 301).

In Russia, the study of the theory of probability properly began with Buniakovsky and in 1918 P. B. Struve called him a *Russian representative of the French mathematical school* (Sheynin 1991, p. 200). True, he finally left probability and became engaged in statistical inquiries, where his achievements were marred by a serious methodological mistake and lack of sufficient reliable observations.

In § 3.1 Gnedenko describes Chebyshev's first stochastic investigation of 1845. It is, however, doubtful whether this work had been, as intended, a proper textbook. A general survey of the theory of probability and its applications would have been more useful the more so since Chebyshev's reasoning was necessarily burdensome.

In § 3.3, Gnedenko discusses the prehistory of Markov chains. I (2009, §§ 14.2-4) provided a fuller account of this subject.

In § 3.4 Gnedenko repeated Kolmogorov's indirect explanation: Markov had not discussed any applications of his chains to natural science since the Petersburg school was *very remote from statistical physics*. I add another cause, see Markov's letter to Chuprov of 1910 (Ondar 1977/1981, p. 52):

I shall not go a step out of that region where my competence is beyond any doubt.

Gnedenko portrays Markov as an ideal scholar with which I cannot agree, see my paper (2006) whose scope is wider than suggested by its title. And in any case, as a textbook Czuber (1903) was incomparably better than Markov (1900 and later editions).

In § 3.5 Gnedenko mentions Bertrand's treatise (1888) which I have discussed in special paper and shortly described elsewhere (2009, § 11.1). His treatise is impregnated with non-constructive and often unjustified and wrong attitude towards the theory of probability and treatment of observations.

In § 4.2 Gnedenko mentions Borel and Cantelli as the discoverers of the strong LLN. Chetverikov (1959) described Slutsky's debate of 1928 with Cantelli on priority. Slutsky stated that Borel had only considered that law in passing (and only in a particular case).

In § 4.3 Gnedenko mentions, that Kolmogorov's axiomatic justification of probability theory is *generally recognized*. Yes, but not all at once:

Some mathematicians sneered that [...] *perhaps probability needed rigour, but surely not* rigour mortis (Doob 1989).

And the same author once more (Doob 1994, p. 593):

It was some time before Kolmogorov's basis was accepted. [...] *The idea that a* (*mathematical*) *random variable is simply a function with*

no romantic connotation seemed rather humiliating to some probabilists.

In § 4.5 Gnedenko mentions Khinchin's book (1943), so I quote Novikov (2002, p. 334):

Khinchin attempted to begin studying the justification of statistical physics, but physicists met his attempts with deep-rooted contempt. Leintovich [a most eminent physicist] *said* […] *that Khinchin does not understand anything at all*.

In § 4.6 Gnedenko discusses infinitely divisible laws. It was Lévy (1925) who introduced them, but he (p. vii) was entirely mistaken in believing that they were most essential for the treatment of observations (Sheynin 1995, § 5). Again, on p. 79 Lévy mistakenly concluded that the method of least squares is only applicable to observations whose errors obey the normal law.

Gnedenko refers to his own contribution (1939) on the limit laws for sums of random variables. That same year he published another paper on the same subject, but it was printed so badly, that he reasonably passed it over in silence.

Doeblin apparently perished during the war: he was drafted into the French army. Being a Jew, he shot himself rather than find himself in German hands. On his life and work see Bru (1993) and Bru & Yor (2002).

In § 4.8 Gnedenko properly praises the work of Glivenko et al in mathematical statistics, but his statement about the *new direction in statistics* which they originated is not altogether correct. Poisson (1837, § 112, a summary) partly preceded them. *That Soviet scientists have not yet occupied leading positions in mathematical statistics*, can be supplemented by mentioning statistics in general (Sheynin 1998).

Finally, *Romanovsky had for some time worked under a certain influence of the Pearson school*. This is an extremely weak statement. Romanovsky corresponded with Pearson in 1924 – 1925, and with Fisher in 1929 – 1938 (Sheynin 2008). He published reviews of several books of Fisher. In spite of his efforts, Fisher's *Design of Experiments* was not translated, and his *Statistical Methods for Research Workers* of 1934 as only translated in 1958. I (2008, pp. 366 – 368) also noted that Romanovsky had been severely victimized for his connections with western scholars.

I think that a few words about

Shafer G., Vovk V. (2001), *Probability and Finance. It's Only a Game.* New York

are not out of place. *Pt 1*: Probability without measure. *Pt. 2*: Finance without probability. *Preface*. *We show how probability can be based on game theory, free from the distracting and confusing assumption about randomness*. However, without randomness probability seems to be philosophically suspect.

In chapter 2 of pt. 1, pp. 39 – 40 the authors recognize that Kolmogorov's axiomatic approach was important for them.

The book makes a difficult reading and much more thought should have been spent for making it understandable for a wider circle of readers. I have looked through it and noticed that Lindeberg is discussed in detail, but Liapunov is forgotten. The figures in Chapter 1 disgrace the authors. Then, on p. 165 the photo of Legendre is presented as portraying Laplace. That it is a photo of some other Legendre, is a general mistake which lasted for a few decades. No portrait of the *real* Legendre is in existence.

v. We see Koalovich as a worthy correspondent of Markov. He, as

well as Sleshinsky, Yaroshenko and Markov, have not appraised properly the significance of the (yet unnamed) variance and therefore of the second justification of the MLSq by Gauss (1823). At least in 1899 Markov did appreciate this substantiation, but for a wrong reason. A special point is that Markov's early mimeographed lectures are still unstudied. I have omitted many interesting details provided by Ermolaeva in her notes.

ix. The commentary by Youshkevich is the most effective description of Nekrasov's craziness. Note, however, that Nekrasov continued to publish quite *normal* and apparently interesting papers in other branches of mathematics and in mechanics. Youshkevich' reference to Nekrasov's book (1912) only seems to mean that the quote following it is from that very source. When possible, the references correspond to the Bibliography appended to the main text.

xi. This obituary essentially supplements the much later and well-known essay of Markov Junior (1951). I have inserted some critical notes, notably Notes 13 and 9.

I

Christiaan Huygens

Correspondence

Huygens C., *Oeuvres Complètes*, tt. 1 – 22. La Haye, 1888 – 1950

Letter of Huygens to P. de Carcavy
6 July 1656 (OC, t. 1, 1888, pp. 442 – 447)

[…] Fermat solved my problem, and I see that he has a universal method for finding everything that concerns that subject. This was the only point that I wished to find out when proposing my problem. The same ratio 30:31 is contained in the treatise which I have sent to Schoten two months ago (Huygens 1657). The theorem which I apply in each case pertaining to the problem of points is also here. I provide it since otherwise I will be unable to explain everything about the problems which Fermat had proposed me. In some cases the necessary calculations are so long that I do not have the patience for carrying them out completely. This is why you will see below an explanation of that theorem and I will be content to convey my method which allows to approach it. Here is this theorem.

Suppose that the number of chances for obtaining b is p and their number for getting c is q. Then this is the same as having

$$\frac{bp + cq}{p + q}.$$

For example […].

1. The first question of Fermat is this. A and B play with 2 dice. A wins if he throws 6 points, and B wins if he throws 7 points. A begins, then B [just the same] throws the dice 2 times, then A throws them twice and so on until one of them wins. It is required to share the stakes[1].

2. It is supposed that A throws twice, then B throws 3 times and after that A throws 3 times. The method of solution is quite similar and once more I find the same numbers as Fermat did, only he mistakenly transposed them. The share of A is actually the share of B and vice versa, not 87451:72360, but 72360:87451.

3. Gamblers A, B and C play with a complete deck of 52 cards. They extract cards one by one beginning with A and the winner is that gambler who first extracts a heart. There are 13 hearts. If all the other 39 cards are extracted, and it is A's turn to extract the next card, he certainly wins. However, if C is to extract the 39th card, A has 13 chances to win nothing and 1 chance to get the stakes (I denote them by d). According to our theorem this is $1/14 d$. Now, when B is about to extract the 38th card, A has 13 chances to get nothing and 2 chances to obtain $1/14 d$ since then C will extract the 39th card. This is equal to $1/105 d$.

When A extracts the 37th card he has 13 chances to get d and 3 chances to obtain $1/105d$ which is equal to $1368/1680d$. And so, going each time back we at last determine the share of A when he extracts the first card. In the same manner we find B's share whereas the rest goes to C.

4. This problem concerns the game prime with 40 cards. Gambler A undertakes to get the prime whereas B bets that A will not get it after 4 extractions. I am told that a prime means 4 different cards, one of each suit.

I find that the chances of A and B are in the ratio of 1000:8139 so that it is quite possible to bet 8 to 1 against obtaining the prime[2].

5. Two gamblers play piquét. A undertakes to get 3 aces among the first 12 cards and B bets against him. To solve the [pertinent] problem I suppose that A extracts the cards one by one since the manner of extraction is of no consequence. Suppose that A extracted 11 cards and got 2 aces. Then, among the other 25 cards [so there are 36 cards in all] there are 2 more aces. He therefore has 2 chances to win d and 23 chances to obtain nothing. This is equal to $2/25d$.

If he got 2 aces after 10 extractions, he obtains 2 chances to get d and 24 chances to get $2/25d$ which is equal to $49/325d$. However, if among the 10 cards there is only 1 ace, then among the other 26 cards there are 3 aces more, so that he has 3 chances to get $2/25d$ and 23 chances to obtain nothing when there will still be 1 ace among the 11 cards and he will be unable to win. This is equal to $3/325d$. [A detailed calculation follows.]

And so, each time going back we will finally get A's chare at the very beginning and the rest will go to B.

If I were sufficiently acquainted with the problems about the games of chance which Fermat calls the most difficult, I have solved all of them. And I am asking you to do me a favour and inform Mr. Milon[3] about it. Then I will be able to know whether the discovered by Fermat and Pascal conforms to what I have explained. And I very much wish to know whether they had issued from the same theorem as I did. […]

[Similar calculations are contained in a supplement to this letter.]

Letter of Lodewijk Huygens to Christian Huygens
30 Oct. 1669 (OC, t. 6, 1895, pp. 515 – 518)

[…] I provide my not quite proper calculations of the ages but there is so little to say that that is not essential and still less since the English table on which we are basing ourselves is not after all so perfect. That author[4] says that

Those numbers [in his life table, usually called mortality table] *are practically neere enough to the truth, for men do not die in exact proportions* [of one age to another] *nor in fractions.*

So here is the method which I applied. First of all, I calculated the sum of the years which all those 100 men [the initial number in Graunt's table] taken together ought to live, i. e., 1822 years, as you may see on the next page.

36 die before reaching 6 years having lived 3 years in the mean, which makes 108 years

24 die between ages 6 and 16 having lived 11 years in the mean, which makes 264 years […]

 And 1 man who dies between ages 76 and 86, lived 81 year

In all, 1822 years

When dividing these 1822 years equally among the 100, we have 18 years and about 2 months [2.64 months] for each which is the age of each born or conceived (crée ou conceue). Indeed, note in passing that the Englishman is speaking about the conceived which are registered just as those who were born: miscarriages have also been included[5].

Here is how I proceed to calculate and specify how long will each person of a certain age still live. At first, I subtract the 108 (the total age of the 36 children who die before age 6) from the total of 1822 and get 1714. I ought to distribute this number between the 64 remaining people which provides 26 years and about 10 months for each who lived to be 6 years old so that they still have 20 years and 10 months of life.

Then subtract from these 1714 years the sum of the years for the 24 people who die between ages 6 and 16 (i. e., 264 years). The difference is 1450. This number should be distributed among the 40 still living, so that each person

aged 16 has 36 years and 3 months or 20 years and 3 months more
aged 26 has 45 years and 4 months or 19 years and 4 months more
aged 36 has 53 years and 6 months or 17 years and 6 months more […]
aged 86 has none at all

If I intend to determine the remaining life for a person aged between, say, 36 and 46 years, just as you and I, I calculate their future years in proportion to the excess of their age above 36 years, etc.

In accord with the above, I do not understand the grounds of your calculation of [the ratio] 4:3. I think that a bet whether an infant aged 6 years or a young man of 16 years will live 20 years more should be considered as being on almost equal terms. And I am asking for an explanation since I am sending you mine. […]

Letter of Christiaan Huygens to Lodewijk Huygens
21 Nov. 1669 (OC, t. 6, 1895, pp. with supplements 524 – 532)

I have just examined your calculation of the ages and repeated my own since I had lost it. Your calculations are correct although you gave us a somewhat longer life […]. You concluded almost appropriately that if the 100 people taken together have 1822 years of life, but it does not follow that 18 years and 2 months […] is the age of each person born or conceived although you assume that it is certain.

Suppose for example that babies are even sicklier than in reality and that during their first 6 years 90 die out of the 100 but that the rest usually live like Nestors or Methuselahs[6] to 152 years and 2 months. You will have the same 1822 years, but who bets that a baby only (seulement) lives to 6 years will be greatly disadvantaged because only 1 out of 10 lives until that age[7]. Here is another example. Suppose that under the usual premise I bet that each (chacun) of those 100 people lives until age 16. It is clear that usually only 40 people out of the 100 live to age 16 so that to equalize the bet I can only stake 40 against 60 or 2:3.

And so you see that 18 years and 2 months is not at all the longevity of everyone conceived; I find that it only equals about 11 years. Who bets that a 6 years old infant lives until 26 can stake 25 against 39 since out of the 64 such infants only 25 live until 26 whereas 39 die. […]

And so, the calculation is very reliable and quite easy. However, you are asking how I can determine, as you did, the reasonable longevity for a person of a proposed age. So I provide the small English table without bothering to calculate anything, trace a curve and measure with a compass the life of anyone. I see, for example, that at your age of 38 years you can still live 19 years and about 4 months. However, if you will often amuse yourself by asking others to beat you, we will have to subtract something[8].

On another occasion I will send you the line of life[9] with an explanation and even a table of lives for each age, year after year, which will be quite easy for me. […]

Supplement No. 1
21 Nov. 1669

After examining the calculations of my brother Louis.

According to the quite exact observations made in London [the Graunt table is reproduced[10]; it was downright faulty but methodically wonderful]. [It followed that] who bets that a conceived infant lives until 6 years can stake 64 against 36 or 16 against 9. Who bets that it lives until 16 years, can only stake 40 against 60 or 2 against 3 because only 40 live until 16 years.

However, who bets that an infant of 6 years lives until 16, can stake 40 against 24 or 5 against 3 because out of the 64 infants of that age 40 live until 16 and 24 die. [A table based on Graunt's data and reproduced by Louis is adduced. Then 108 is subtracted from 1822 etc.]

A conceived infant therefore has 36 chances to live 3 years; 24 chances to live 11 years; 15, to live 21 year etc. And, according to my rule [to my principle of investigation] for [of] games of chance we should multiply each number of chances by the number of the corresponding years and divide the sum of the products which is here 1822 by the sum of all the chances which is 100. And the quotient, 18 years and about $2_{1/2}$ months will be the value of the chance of a conceived infant.

The method of my brother Louis leads to the same although by a different way. However, although the expectation (espérance) of a

conceived infant is these 18 years and 2$_{1/2}$ months, it does not follow that apparently he will live until that age since it is much more apparent that he will die before that time. Therefore, if someone wishes to bet that the infant will live until that age, he will be at a disadvantage because he can only bet on equal terms that the infant lives until approximately 11 years. Consequently, it is a mistake to say that, the terms are equal when you bet that an infant of 6 years, or a young man of 16 will live 20 years more. We can only bet 25 against 39 in the first case and 2 against 3 in the second although the expectations in both cases are the same (20 years). It is wrong to assume less than 20 years. His (Louis') calculations are proper for annuities.

Find out the time after which two people will die out of 40, each of them 46 years old. Answer: 1 year and 3 months. 4 people between ages 46 and 56 years die out of 10, therefore 2 die in 1 year and 3 months.

A man 56 years old marries a girl of 16 years. Find out how long they both will remain alive and how long will it take for both of them to die. Or, if I am promised 100 francs at the end of each year during which they both remain alive, how much does it justly cost to redeem this obligation. And, given 40 people, each of them 46 years old, how long will it take for all of them to die. And how long for two people each of them 16 years old. Answer: 29 years 2$_{2/3}$ months.

A conceived infant has 18.22 years to live, so he dies at age 18.22 years;
an infant of 6 years, 20.81, dies at age 26.81 years;
a young man of 16 years, 20.25 years, dies at 36.25 years; […]
a man of 76 years, 5.00 years, dies at 81;
a man of 86 years, 0, dies at 86

To find how long lives the last of the two young men, each of them 16 years old, we ought to imagine that each extracts a ticket out of a total of 40.

15 tickets provide 5 years of life; 9 tickets, 15 years; 6, 25 years; 4, 35 years; 3, 45 years; 2, 55 years; and 1, 65 years

Let both of them extract a ticket and suppose that the second extracts the more favourable of them. Now, the first man extracts a ticket and certainly has 15 chances to live 5 years more, 9 chances to live 15 years etc. Next, the second man. If he gets less than 5 years [see below] it will not harm him since the first man had obtained 5 years and any number smaller than 5 counts as 5. However, he, the second man, also has 15 chances, 7$_{1/2}$ of them to live less than 5 years and again 7$_{1/2}$ to live 6, 7, 8, 9 or 10 years[11] which is the same as 7$_{1/2}$ chances to live 8 years. And 25 chances more which, for a young man 16 years old, mean 29.40 years $[(9 \cdot 15 + 6 \cdot 25 + \ldots + 1 \cdot 65):25 = 29.4]$. The chances ought to be only understood in this way since none of them provides less than 5 years.

And so, the first person extracted a ticket with 15 chances for getting 7$\frac{1}{2}$ chances to live 5 years, another 7$\frac{1}{2}$ chances to live 8 years and 25 chances to live 29.40 years. He also has 9 chances to extract a ticket giving him 15 years and then all the less favourable which the second person can extract will give him 15 years. This is the same as 15 chances for 15 years. Out of the 9 chances 4$\frac{1}{2}$ give him less than 15 years but they should be counted as providing 15 years and another 4$\frac{1}{2}$ chances for 16, 17, 18, 19 or 20 years which is the same as 4$\frac{1}{2}$ chances providing 18 years. And 16 chances more to live 37$\frac{1}{2}$ years.

The first person when he extracts a ticket has 19$\frac{1}{2}$ chances for 15 years, 4$\frac{1}{2}$ chances for 18 years and 16 chances for 37$\frac{1}{2}$ years etc. as shown in the marginal note [see below]. So he has [chances multiplied by years]

15 chances for 20.3 = 304.5; 4 chances for 37.6 = 150.4;
9 chances for 24.3 = 218.7; 3 chances for 46.1 = 138.3
6 chances for 30.2 = 181.2; 2 chances for 55.3 = 110.6;
1 chance for 65.0 = 65.0. In all, 1168.7

The second person has 29.22 years [=1168.7:40]. This means that one person out of the two will die until age 45 years and 2$\frac{2}{3}$ months.

Now, to find when one of them dies, we should once more imagine that they extract tickets, one after another, from the total number of them (40). 15 tickets give 5 years; 9 tickets, 15 years etc. just like it was in the previous problem, but now we ought to consider the less favourable tickets.

The first person extracts a ticket and has 15 chances to live 5 years, 9 chances for 15 years etc. And, if he extracts one of those 15, the second person, whichever ticket he extracts, does not give him more than 5 years. It is even possible to shorten that period. Indeed, out of the 15 chances 7$\frac{1}{2}$ give more than 5 years which nevertheless count for 5 years and another 7$\frac{1}{2}$ for 5, 4, 3, 2 years or 1 year. In addition, the second has 25 chances which can only give him 5 years.

And so, the first person has 15 chances or 7$\frac{1}{2}$ chances for 3 years and 32$\frac{1}{2}$ chances for 5 years and 9 chances to obtain 15 years. And if he extracts one of these, the second person cannot obtain more. However, these 15 years can be shortened if he chooses either one of those 15 tickets for 5 years or one of those 4$\frac{1}{2}$ tickets less favourable which count as 13 years. The other 4$\frac{1}{2}$ chances also count as 15 years even if they indicate a longer period. The second person except those 15 + 9 = 24 chances has 16 more which can get him only 15.

And so, the first person has 9 chances more for obtaining 15 chances for 5 years, 4$\frac{1}{2}$ chances for 13 and 20$\frac{1}{2}$ chances for 15 years[12].

The marginal note
15 – 7½ – 5; 7½ – 8; 25 – 29.40 ... 20.3 [20.8][13]
9 – 19½ – 15; 4½ – 18; 16 – 37½ ... 24.3
6 – 27 – 25; 3 – 28; 10 – 45 ... 30.2
4 – 32 – 35; 2 – 38; 6 – 51.67 ... 37.6
3 – 35½ – 45; ½ – 48; 3 – 58.33 ... 46.1

2 – 38 – 55; 1 – 58; 1 – 65 ... 55.3
1 – 39 – 65; 1 – 66½; ... 65.0

Supplement No. 2
21 Nov. 1669

Ages are shown on the horizontal line[14]. For 6 years the perpendicular is 64 parts long [out of the 100] since the English table shows 64 infants remaining alive at that age. For 16 the perpendicular is 40 parts long since […]. And I draw a curve through the ends of these perpendiculars […].

If I wish to know how many persons remain alive after 20 years out of the 100, I choose number 20, let it be point A, on the base and rise a perpendicular which meets the curve at point B. And I say that AB whose length is almost 33 points is the number of persons out of the 100 who live until age 20.

And if I wish to know how much time is reasonably left for a person of, say, 20 years, I choose a half of the perpendicular AB and measure the length of the horizontal segment DE, let AD = DB and E be the point of the curve which is $16_{1/2}$ years. However, out of 33 persons aged 20 years, a half usually dies during the 16 next years. It is therefore possible to bet on equal terms that a man aged 20 lives 16 years more. It is also possible to determine that the life of a conceived baby should be estimated to be 11 years instead of 18 years and 2 months as my brother informs me.

Letter of Christiaan Huygens to Lodewijk Huygens
28 Nov. 1669 (OC, t. 6, 1895, pp. 537 – 539)

Calculations which I have sent you will certainly embarrass you. Since then I have been thinking about them and about yours as well. I find that each of us had a reason to understand the matter in our own way. You give 18 years and $2_{1/2}$ months of life to a conceived baby and it is true that that is his expectation. However, it is not apparent that he will live so long since it is much more apparent that he dies earlier. Therefore, who wishes to bet that he lives so long is at disadvantage. It is only possible to bet on equal terms that he will live approximately 11 years. And I also find in my manner the same expectation, 20 years, as you say, for an infant 6 years old and a young man (garcon)[15] of 16 years. Nevertheless, you cannot conclude that betting on either living 20 years more is on equal terms. You can only stake 25 against 39 in the first case and 2 against 3 in the second case. Or, it is possible to bet on equal terms that the second will live 15 years more.

It is two different things, the expectation or the value of the future age of some person, and the age for which we have an equal possibility of remaining alive or not. The first method is proper for regulating annuities, the second, for bets. I will see whether you make the same distinction. However, your method is very nice and subtly derived. It comes exactly to what I have found by following my rules [my principles of investigation] of chance (1657) […]

For example, a conceived baby has 6 chances to live 3 years, 24 chances to live 11 years etc. because, according to [my] rule the [sought] value is found by multiplying […] and dividing […].

Concerning your captains[16], I think that you have applied the English table. If out of 10 people aged between 46 and 56 years 4 die during 10 years, then out of 40 people of the same age 16 die so that according to the rule of three 2 die in 1 year and 3 months. However, 2 will die out of 40 during 15 months if they are all aged 46 years rather than 50. And we should not suppose that it is exactly 15 months because they do not die uniformly during those 10 years but rather more rapidly at the beginning of that period when their group is more numerous rather than after death takes away some of them[17].

Here is a sufficiently delightful question which seems to me much more difficult than yours about the captains. I have not yet solved it but I see the pertinent method. Two persons, aged 16 years each. How long may they expect to remain both alive. And how long does it take for both of them to die. These are actually two different problems, and each requires some thinking.

If their ages are different, say 16 and 56 years, some changes are necessary but no new essential difficulties since that problem was solved for equal ages.

The curve which I mentioned in my previous letter only serves to solve the problem of bets so that there is no need to send it to you, but it is possible and, moreover, to enlarge it. Then it will supplement your table of the duration of life still left for each age.

Notes

1. It is difficult to follow the provided solution. See Jakob Bernoulli, *Ars Conjectandi*, Proposition 14 from the treatise of Huygens (1657) and, also, the treatment of the Huygens problem.

2. See the solutions of that problem in Part 1 and of Problem 5 from Part 3 of Jakob Bernoulli's *Ars Conjectandi*.

3. Claude Milon (ca. 1618 – ca. 1660) deserves remembrance for his scientific connections with Huygens, Carcavi, Roberval and Fermat.

4. Neither Chistiaan, nor Louis Huygens ever mentioned John Graunt by name. This is strange since both quoted or mentioned his book (1662).

5. Now I myself quote Graunt (1662/1899, § 3.39, p. 360): *The abortives and stillborn are about twentieth part of those that are christened.* This apparently means that they had been entered in the Bills of mortality. See also Note 10.

6. Nestor: a mythological Tsar in ancient Greece who lived until age 200. Methuselah (Genesis 5:27) lived 969 years.

7. It seems that *seulement* should be suppressed. A few lined below Huygens made an obvious mistake possibly occasioned by oversight: *each* should have been *someone*. He himself (see the very beginning of his Supplement No. 1) provided the correct answer: someone has 40 chances out of 100 etc.

8. Hardly anything is known about this episode.

9. See Note 14.

10. The Graunt table registered conceived babies since it was stated in the Bills of mortality that they included *abortives*. C. H.

Exact observations (see a few lines above) is definitely wrong, even Louis knew otherwise, see the very beginning of his letter. It follows that Christiaan had not read Graunt's book, had not been seriously interested in its subject.

11. Fractional chances seem strange but they can easily be excluded by slightly complicating the calculations.

12. Huygens apparently had not completed his calculations. Editor.

13. I have corrected the number 20.3 in the first line. Subsequent corrections are therefore still needed.

14. Christiaan also showed that the probable duration of life could be determined by means of the graph (a continuous curve passing through empirical points given by Graunt's table of mortality; plate between pp. 530 and 531) of the function

$$y = 1 - F(x),$$

where, in modern notation, $F(x)$ was a remaining unknown integral distribution function with admissible values of the argument being 0 x 100.

15. At the time, there were no separate tables for men and women so that *garcon* is strange.

16. This problem seems to be formulated by Louis in his lost letter. It undoubtedly required the calculation of the time during which two captains die out of 40 aged 50 years each. Editor.

17. Huygens examined the expected period of time during which 40 persons aged 46 will die out; and 2 persons aged 16 will both die. The first problem proved too difficult, but Huygens might have remarked that the period sought was 40 years (according to Graunt, 86 years was the highest possible age). True, he solved a similar problem but made a mistake. He assumed there that the law of mortality was uniform and that the number of deaths will decrease with time. However, for a distribution, continuous and uniform in some interval, n order statistics will divide it into $(n + 1)$ approximately equal parts and the annual deaths will remain about constant. In the second problem Huygens applied conditional expectation when assuming that one of the two persons will die first.

Bibliography

Bernoulli Jakob (1713), *Ars conjectandi. Wahrscheinlichkeitsrechnung*. Leipzig, 1899, reprinted in 1999, Frankfurt/Main. Do not trust its English translation by Edith Dudley Sylla (Baltimore, 2006).

Graunt J. (1662), *Natural and Political Observations Made upon the Bills of Mortality*. In vol. 2, pp. 317 – 435 of Petty W. (1899), *Economic Writings*, vols 1 – 2. Cambridge. [London, 1997.]

Huygens C. (1657), De calcul dans les jeux de hasard. In Huygens (1888 – 1950, t. 14, pp. 49 – 91). Reprinted with commentaries by Jakob Bernoulli as part 1 of his *Ars Conjectandi*.

--- (1888 – 1959), *Oeuvres complètes*, tt. 1 – 22. La Haye.

Korteweg D. S. (1920), Apercy de la genèse de [Huygens 1657] et de recherches subséquentes de Huygens sur les questions de probabilité. In Huygens (1888 – 1950, t. 14, pp. 3 – 48).

Sheynin O. (2009), *Theory of Probability. Historical Essay*. Berlin. **S, G,** 10.

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

II

I. Ya. Depman

M. F. Bartels, the teacher of Lobachevsky

Istoriko-Matematicheskie Issledovania, vol. 3, 1950, pp. 475 – 485

1. Biographers of Lobachevsky devoted many pages to his teacher, Johann Martin Christian (Martin Fedorovich, as he was called in Russia) Bartels, and the same is true about the biographers of another of his pupils, or rather of schoolfellows, Gauss. The pedagogic influence of Bartels on Lobachevsky, and his care about his pupil is doubtless. However, this influence is sometimes unfoundedly exaggerated.

In 1893, in his address at the centenary of Lobachevsky's birth, A. V. Vasiliev suggested as a probable hypothesis that Lobachevsky could have got his first incentive for geometric studies from Bartels. The last-mentioned, as it is possible to believe, was acquainted with those Gauss' ideas to which Lobachevsky came later. True, later Vasiliev (1914) indicated that

N*ewly discovered materials prove that that hypothesis is useless. They convince us that Lobachevsky began his studies of the theory of parallel lines quite independently from Gauss*

and that he was possibly led to them by the general interest in that theory which became enlivened at the turn of the 18th century.

This later opinion was published in a reference book and can escape the attention of those who had read Vasiliev's initial statement. There is a somewhat justified fear that those readers will remain under a delusion. Even apart from that argument, it is worthwhile to consider the scientific work of Bartels and find out to what extent he had been acquainted with the new geometric ideas, and how he had regarded them and thus to reveal the entire groundlessness of those assumptions about his role in the discovery of the non-Euclidean geometry.

I was able to get acquainted with Bartels' personal file at the Tartu (Dorpat – Jurjew) University and can add a few new features in Bartels' biography and thus still more strengthen the solidly established in the Soviet literature viewpoint about the total independence of Lobachevsky's great ideas from Gauss.

The study of Bartels' scientific work is all the more necessary since Klein (1928) very confidently stated that the ideas of Gauss prompted Johann Bolyai (through his father, Gauss' fellow student) and Lobachevsky (through Bartels) to develop the non-Euclidean geometry[1]. In the later editions of his book Klein retracted his initial point of view[2], but, as it often happens, such changes of viewpoints do not always become known to the readers of the first editions of books. And here such a case had indeed occurred.

A Riga mathematician, Alfred [Arnold Adolf] Meder (1928), stated that

The possibility of Bartels' talks with Lobachevsky about the Gauss' ideas as well as the assumption that Gauss could have provided, even

if indirectly, an incentive to the work of Lobachevsky, should not be overlooked.

The study of the biography of Bartels shows that such assumptions are groundless.

2. Information about Bartels is rather scanty. He is known to have been born in Braunschweig in 1769 and died in Dorpat (Jurjew, Tartu) on 7 Dec. 1836, old style. For three years he attended the Collegium Carolinum of his home town, then studied the law in Helmstedt. Until entering a secondary school he learned in a primary school and later became there an assistant teacher. As such, he was the teacher of the eight years younger Gauss. They met again at Göttingen University where Bartels found himself owing to the influence of the celebrated mathematician, Johann Friedrich Pfaff[3], after he broke his connection with the initially chosen faculty of law.

Gauss and Bartels struck up a friendship which lasted until the death of the latter. They corresponded, although with long interruptions. The Gauss letters are lost. Vasiliev had asked the readers of two periodicals about them, but got no reply. The Bartels letters are kept in the Gauss archive, and they do not touch on scientific subjects. However, a letter of O. V. Struve, Bartels grandson on the mother's side, makes known that Bartels had been negatively inclined towards Lobachevsky's geometry and did not recognize its scientific importance. This would be impossible had he been acquainted with Gauss' views, see Kagan (1948, p. 395).

In 1801, Gauss who became renowned by calculating the orbit of the small planet Ceres[4], was invited to the Petersburg Academy of Sciences. The pertinent correspondence lasted a few years[5] and ended by Gauss' refusal, although he was initially very much interested in the invitation.

Gauss apparently recommended Bartels instead of himself, and the warden of the just established Kazan University[6], the academician, astronomer S. Ya. Rumovsky, invited Bartels to the chair of mathematics. At first, Bartels only accepted the title of honorary fellow of the University, but in 1807, when Napoleon occupied Prussia and scientific work there became impossible, Bartels accepted that invitation and arrived in Kazan in the beginning of 1808.

The pedagogic work of Bartels in Kazan is sufficiently described in the biographies of Lobachevsky, and it is unnecessary to return to this subject. In the end of 1820 Bartels moved to Dorpat University as ordinary professor of mathematics. Almost nothing is written about his work there or about his scientific activity and I propose to throw some light on this area.

3. The general picture of Bartels' official duties in Dorpat fully coincides with what is known about his work in Kazan. He displayed himself as an honest instructor and a thorough employee. Just as in Kazan, Bartels was elected dean and carried out various administrative duties. His service record shows him as an all-round irreproachable worker.

We find out that in Kazan Bartels read astronomy from 17 May 1816 to 19 April 1818 since the astronomy professor left the university, Then, he was

Never noticed as being weak in the exercise of his service, and as a superior. In spite of his properly demanding attitude, never allowed any discord among his subordinates.

Also significant are the other conclusions:

He never went on holidays […] *was never announced or found guilty of indecent behaviour,* […] *attested in everything.*

Such statements are included in every of Bartels' service record beginning with a certificate of 31 October 1828 signed by the rector, Lobachevsky.

Bartels was elected professor of the Dorpat University as proposed by its physical and mathematical faculty since he *was known as an excellent instructor of mathematics and a person experienced in managing affairs*. On 11 August 1820 the Council of the University voted the proposal; eighteen votes were positive and two, negative. On 29 September the director [not rector?] of Kazan University notified that Bartels was discharged, and on 27 January 1821 Bartels was present for the first time at a sitting of the Council of Dorpat University.

4. Bartels' service record at Dorpat contains no information about his scientific work except a protracted correspondence about the publication of lectures in higher analysis. It almost exclusively describes the attempts of receiving a grant for their publication from the Ministry [see also below]. In 1836, the year of Bartels' death, the warden required the professors to provide lists of their published contributions. Here is his list, accurate and final since written by himself at the end of his life.

The years 1788 – 1791. Translations, including**7**

[William Smellie,] *Philosophy of the Natural History* [1710; 1791; Boston 1846], vols. 1 – 2

[J.-S.] Bailly, *Histoire de l'astronomie*, [1775; 1781;] tt. 1 – 2. Göttingen.

Dissertatio de calculo variationum (a [his] Doctor dissertation)

Disquisitiones quatuor ad theoriam functionum analyticam pertinentes. Dorpat, 1832 (Four Discourses on the Theory of Analytic Functions; *analytic functions* should not be understood in the modern sense)

Vorlesungen über mathematische Analyse, in quarto. Dorpat. 45 – 46 printer's sheets, 1833 [and 1837]. (Bartels intended to publish two more volumes of 60 sheets each.)

I have submitted many papers to the Petersburg Academy:

Apercu abrégé des formules fondamentales de la géométrie à trois dimensions. Lu à l'Académie 1825. Published, as I suppose, in 1830.

Two other papers ought to appear in the next issue [?], but I have not seen them yet:

Sur la parallaxes du Soleil

Sur les trois axes principales d'un corps solide

These two have not ben published (Dinze et al 1936). The *Apercu abrégé* … appeared in 1831 in *Mém. présentés à l'Académie des Sciences par divers savants*, t. 1, pp. 77 – 95.

Bartels' only scientific contributions are therefore his *Disquisitiones* … and the paper read in 1825.

This latter subject, as it seems, interested Bartels in the first place. He is known to have twice suggested themes belonging to this field to his students as a prize problem and both times the answer came from his successor to the chair, Karl Edward Senff (1810 – 1849).

In the published issues of those answers he (1829; 1831) acknowledged that the results were mostly expositions of Bartels' ideas. His first publication coincides with the memoir of Bartels which he submitted to the Petersburg Academy, the second one deserves attention due to an original treatment of the theory of space curves and anticipation of some of the results of Serret (1860) published much later[8].

From 1826 Bartels was corresponding member of the Petersburg Academy, Class of mathematics. We see that his mathematical achievements had hardly justified that status; in Russia, there had undoubtedly been mathematicians who deserved it more justifiably. The election of Bartels was apparently due to the German members of the Academy among whom Bartels' son-in-law, V. Ya. [F. G. V.] Struve enjoyed strong influence.

In 1829, Bartels submitted a plan of the publication of *Mathematische Vorlesungen über hohere Analysis* in 20 parts to the Ministry of Education. Bartels believed that 150 copies taken by the Ministry for its educational institutions will cover the cost of publication. The Ministry, however, only subscribed for 39 copies for military schools. Then, after a prolonged correspondence, agreed on 132 copies, but it only paid for the first volume of the book which appeared in 1833 [see below].

The first Russian mathematical journal, the *Uchebny Matematicheskiy Zhurnal*, enthusiastically apraised that volume. The journal was edited by K. G. Kupfer (1790 – 1838), a gymnasium teacher and mathematics professor since 1835 of the Nezhin Lyceum of prince Bezborodko where he left fondest memoirs of himself[9]. His journal appeared in 1833 and 1834 in Reval [Tallin] Kupfer explained:

The three volumes of the book will comprise a complete course of analysis with applications to geometry, mechanics, calculations of probabilities, – all that which the respectable author had been teaching for 25 years in Kazan and Dorpat.

A reviewer (that journal, 1833, p. 237)[10], the author of the dissertation *On summing of series* (Mitava [Mitau, Jelgava], 1813, in Latin) and of a German book *Essay on the Method of Determining the Number of Imaginary Roots of an Equation of an Arbitrary Power* (Dorpat, 1819), probably exaggerated:

It is impossible not to discern at once a subtle mathematician. Everything, solution of problems, proofs of propositions and methods of computation, is distinguished by simplicity and elegance. This book can be called a collection of refined solutions and methods. Where such had not existed, the author invented them. Add to this advantage his experience, since during his long life the respectable author

diligently and invariably attempted to be useful for his students and all those who looked for his directions.

All the attempts which Bartels undertook to secure the cost of the publication of the next volumes by preliminary orders, his reminders that Liven, the previous Minister of Education, had approved the idea of the compilation of the book, proved futile although the emperor sent him a diamond signet ring.

After Bartels' death Struve, his son-in-law, received an enquiry about the time of the publication of the next volumes of that book. He answered that the second volume was not in print since its publication was not secured.

5. The above includes everything that can be said about Bartels' creative work. We see that it was very modest and this conclusion does not change after studying the fundamental history of the Kazan University (Zagoskin 1903). Its second volume (p. 639 ff) contains a

List of contributions published by the members of the Imperial Kazan University before their work there and their manuscripts. Compiled in the spring of 1819 for the aim of the inspection of the University by M. L. Magnitsky[11].

And there we read:

Ordinary professor of pure mathematics, M. F. Bartels, published various translations such as (see above)*. He has many manuscripts touching on higher mathematics and a complete course of mathematics which he compiled*.

Separately mentioned is his *Memoir on mathematical analysis submitted to the warden, Rumovsky, in 1805*.

I note in passing that on p. 647 this *List* mentions that

Extraordinary professor of pure mathematics, Lobachevsky, compiled <u>The basis of geometry</u> and several discourses on higher mathematics which are not yet published.

This is apparently a reference to a textbook on geometry which Lobachevsky had submitted for publication in 1823 but which was returned to him for revision. The appearance of that manuscript should therefore be attributed to a period preceding 1823.

The same *List* includes information about J. J. Littrow, the astronomy professor:

Published Information about the results of the observation of the comet of 1812 made by Master Lobachevsky and student Simonov, <u>Kazanskie Izvestia</u> 1812, No. 21.

A legend was repeatedly published about Laplace's answer to the question of who was the best mathematician in Germany. He had allegedly answered: *Bartels, since Gauss is the best in the whole world.* It is obvious that Bartels' scientific merits could not have justified such an answer. It is possible that Laplace mentioned Bartels' mentor, Pfaff, rather than Bartels himself, as Shering reported[12]. If true, such an answer testifies to Laplace's reliable flair. The *Pfaff problem*, to which mathematicians of the first rank has been devoting their efforts until and including our time ([the Russian mathematician] N. M. Gyunter) is indeed a sufficient justification for that probable answer.

Returning to the main question about the possibility of Bartels' inspiring Lobachevsky's interest in the non-Euclidean geometry, we ought to answer that nothing that we know about Bartels provides any justification for even an assumption that he understood Lobachevsky's ideas. If they had at least insignificantly interested Bartels, he would have certainly disclosed his interest on some occasion. We do not, however, see this either in his very modest works, or in his course [which course?], or in his letters to Gauss, or, finally, in the reminiscences of his contemporaries about him.

Bartels is known to have been a lover of the history of mathematics who read it both in Kazan and Dorpat. He would have certainly mentioned the ideas of his former student had he only understood or knew them, the more so if in some way he had assisted their origin in Lobachevsky's mind. For Lobachevsky during his early age, Bartels was a kind and reasonable mentor[13], and he returned Bartels his love and gratitude. Bartels was unable to give Lobachevsky anything else, which, for that matter, the latter did not need. Lobachevsky was and remains an original man of genius of Russian science[14].

Notes

1. Kagan (1944/1948, p. 393ff) resolutely opposed this viewpoint. I. D.

2. [Nevertheless,] Klein (1926) expressed in essence the same viewpoint elsewhere. See G. F. Rybkin's paper in this issue, pp. 17 – 19. Editor [A. P. Youshkevich]

3. Johann Friedrich Pfaff (1765 – 1825), professor at Helmstedt. Corresponding member since 1793 and Honorary member from 1798 of the Petersburg Academy of Sciences. In 1815 he posed the *Pfaff problem* belonging to the theory of [partial] differential equations. He should not be confused with his younger brother Johann Wilhelm Andreas (1774 – 1835), mathematician and astronomer, professor at Dorpat (1804 – 1808) before returning to Germany, corresponding member of the Petersburg Academy from 1807. The professorial staff of the Tartu University mistakenly thinks that the professor at Dorpat was the elder Pfaff, who was the author of the celebrated *Methodus generalis aequationis differentiales* […] *complete integrandi*. One more Pfaff (Hans Pfaff) was chair of mathematics at Erlangen, 1869 – 1872, and Klein was his successor. Finally, after Johann Wilhelm Pfaff left Dorpat, it was Gauss who succeeded him in 1809. This episode deserves to be specially described. I. D.

In 1809, Gauss was already chair of mathematics & astronomy and director of the observatory in Göttingen. O. S.

4. Gauss became internationally renowned in 1801, after he published his *Disquisitiones arithmeticae*. The letters of Bartels to Gauss are now published (Biermann 1973). They contain information about Kazan University. O. S.

5. See *Nauchnoe Nasledstvo* (Scientific Heritage), vol. 1, 1948, pp. 784 – 788 and the introductory paper by N. I. Idelson. I. D.

6. The Kazan University was established in 1804, See *Great Sov. Enc.*, third edition, vol. 11, 1973. This edition of the *Encyclopedia* was translated into English, see the same volume 11. Biermann (1975, p. 143) argued that Gauss did not recommend Bartels. O. S.

7. This list is compiled carelessly. The translations (apparently, only the two first books) were not properly separated from the other items. The book of Smellie (I have not found any other author of the *Philosophy* …) was translated by E. A. W. Zimmermann (Berlin, 1791), and probably no other translations were published. Then, Bartels translated two books of Bailly, the histories of ancient and modern astronomy (Leipzig, 1796 and 1797). O. S.

8. Paul Joseph Serret (1827 – 1898), professor at the Catholic University in Paris, should not be confused with Joseph Alfred Serret (1819 – 1885), academician and author of celebrated treatises. I. D.

9. In 1832 – 1840 that Lyceum became a physical and mathematical institute (Youshkevich 1968, p. 308).

10. Depman had not named that reviewer, nor did he provide the German title of his book. I was unable to identify him. O. S.

11. Youshkevich (1968, p. 218) called him an *unworthy offspring* of L. F. Magnitsky, the author of the first published Russian mathematical treatise *Arithmetic* which appeared in 1703, and an arch-reactionary and arch-rascal (p. 232). O. S.

12. Who was Shering, and how to spell correctly his name? Anyway, it was probably Humboldt who asked Laplace that question. Concerning the Pfaff problem, see Note 3. O. S.

13. Youshkevich (1968, p. 232) also highly appraised Bartels as Lobachevsky's mentor. O. S.

14. Biermann (1975, p. 146) maintained, regrettably without substantiation, that Bartels had *many times saved Lobachevsky from repression*. Did religious matters lead to (avoided) repression?

Kagan (1944) briefly described Bartel as a mentor:

At first, Lobachevsky had still been entirely under the influence of his mentors, especially Bartels who was an excellent instructor (p. 52).

As dean of the physical-mathematical faculty from 1813 until his departure, Bartels had been extremely diligently fulfilling his duties, and was an outstanding pedagogue (p. 55)

Bartels was an excellent instructor but absolutely unsuited as an academician (p. 69).

Bibliography

Biermann K.-R. (1973), Die Briefe von Martin Bartels an Carl Friedrich Gauss. *Schriftenreihe f. Geschichte d. Naturwissenschaften, Technik u. Medizin*, Bd. 10, pp. 5 – 22.

--- (1975), Martin Bartels – eine Schlüsselfigur in der Geschichte der nichteuklidischen Geometrie? *Mitt. d. Deutschen Akad. d. Naturforscher Leopoldina*, Bd. 21, pp. 137 – 157.

Dinse O. V., Shafranovsky K. I. (1936), Bibliographic index. In *Matematika v Izdaniyach Akademii Nauk, 1728 – 1935* (Math. in the Publications of the Akademia Nauk).

Kagan V. F. (1944/1948, Russian), *Lobachevsky*. Moscow – Leningrad.

Klein F. (1926 – 1927), *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*. Berlin.

--- (1928), *Vorlesungen über nicht-euklidische Geometrie*. Springer.

Meder A. (1928), article in *Arch. f. Geschichte d. Mathematik, d. Naturwissenschaften u. d. Technik*, Bd. 2, pp. 66 – 67.

Senff K. E. (1829), *Systematische Darstellung der Hauptsätze der analytischen Geometrie im Raume.* Dorpat.

--- (1831), *Theoremata principalia ex theoria curvarum et superficierum*. Dorpat.

Vasiliev A. V. (1914), Lobachevsky. In *Russkiy Biograficheskiy Slovar* (Russ. Biogr. Dict.). Also as a separate edition (Petersburg, 1914).

Youshkevich A. P. (1968), *Istoria Matematiki v Rossii do 1917 Goda* (History of Mathematics in Russia prior to 1917). Moscow.

Zagoskin N. P. (1903), *Istoria Kazanskogo Universiteta* (History of the Kazan Univ.), vols 1 – 2.

III

Theodor Wittstein

Mathematical Statistics and Its Application to National Economy and the Science of Insurance

Mathematische Statistik und deren Anwendung
auf Nationalökonomie und Versicherungs-Wissenschaft.
Hannover, 1867

Translation of Foreword and Chapter 1

Foreword
Statistics, the science which we are still far from understanding

If that remark formulated by Quetelet in 1845 is still valid, an attempt to understand the field of statistics will not be unwelcome. We are making here such an attempt and we suggest to separate a part of our work [from statistics] as an independent science and call it *mathematical statistics*. It should be allocated in its entirety to the domain of mathematicians.

This new science whose definition we offer below does not regrettably yet allow any precise description of its scope. There is really no material on which the methods of mathematical treatment should first develop and we are therefore unable to present it here as a single entity. We must rather direct our utmost urgent intention to making generally known the available data nowadays hidden in offices and archives.

Meantime, we only provide a few isolated studies which should be suitable for drawing a preliminary picture of what we bear in mind by our proposal. Granted the approval of these investigations, we will continue them as possibilities and free time allow[1].

Hannover, August 1867

Contents

Introduction
Based on a report made in 1865

The number of entities comprising natural sciences is known to be incessantly increasing from within. New splittings and divergences continually arise before our eyes so that all the time new sciences isolate themselves from the general multitude and become independent. Already long ago it became impossible anymore for an individual to have control over the total field of all those sciences or of its considerable part.

The need to separate the work (Arbeit) here also had been ever more asserting itself. On the contrary, it is comparatively rare that the number of sciences had increased by enrichment from beyond with inclusion of a never before existed science. And here, the dealing with such a case seems therefore to deserve special attention.

We will prove that a region of science until now not actually belonging to the natural science ever more powerfully requires such a transformation and development so that in the very next future a fully fledged new member, statistics, will be included in the natural science.

Until now, statistics complete with its name existed for a full century. It was created by Achenwall in Göttingen at the mid-18th century[2]. According to its definition, it ought to establish and arrange every remarkable feature, or (Schlözer [1804]) *Staatswürdigkeit* displayed by a state, or, more generally, by a society. However, very soon a supplement had been added which stipulated that almost everything established ought to be expressed by numbers. And we actually see that the contemporary statistical textbooks are mainly accumulations of tables which show everything possible to be somehow represented by numbers about the population, industry, agriculture, trade etc. of a state. In this form statistics belongs to social science and is considered an essential supplementary science of national economy.

However, we should now recall that from the viewpoint of natural science a collection of tables compiled from experimental materials can by no means claim to be a science in the proper sense of that word. The tables only offer the material for constituting a science and in the sense of natural sciences they are simply collected observations. It follows by itself what the next step ought to be. The goal of all the investigations of nature is to elevate the observations to cognizable natural laws. In this case, the mentioned goal also leads to a science, but, having been true to its definition statistics has up to now done almost nothing. It only discovered what little has been directly arresting the attention by looking at the numbers[3], but the realization of the situation leads to something new.

This new something is linked to the point where statistics as known until now had stopped. And since the initial data of this new science are mostly numerical, it is mathematics that aids in solving the posed problem. The new science should therefore be called *mathematical statistics*, or, as a mathematician will perhaps prefer to say, *analytical statistics*, similar to analytical optics and analytical mechanics.

A comparison with known facts will provide a still clearer picture. An astronomer observes and collects his observations in the form of tables. If, however, he thus concluded his work, astronomy would have never had any claims to be called a science and would have remained a collateral branch of statistics[4]. And, approximately, that was the condition of astronomy when Tycho Brahe had carried out his observations of Mars which later became so renown[5]. But then came Kepler and derived his known laws by issuing from those observations. He chiselled, as Kästner wittily remarked, a statue from a marble block provided by Tycho. That transition from Tycho to Kepler elevated astronomy to the rank of real sciences. Later astronomy became theoretically completed by Newton[6], but this is exactly the same transition which statistics still has to achieve.

However, this statement is not to be understood as though statistics only needs a Kepler for being joined to natural science. It still has no Tycho since the data of present statistics are almost without exception very doubtful and unsuited for subsequent mathematical investigation as is shown in detail below. We may say that statistics lacks even a Copernicus who indicated, although only in the most general way, what should be paid attention to while observing.

Today statistics is exactly in the stage of astronomy's childhood when it was only astrology and cast horoscopes. Until now, statistical numbers prove everything needed, it only requires a certain skill for which the French have invented the expression *grouper les nombres*[7]. The situation will improve as soon as mathematics with its relentless testimony seizes statistical materials. This especially concerns the theory of probability which should be applied here and not only in the sense beloved by statisticians according to which it can only provide *inexact* results as opposed to exact calculations.

It is an ingenious discipline which, as testified by all the other branches of mathematics, is enjoying equal rights with them and which Laplace and Gauss so brilliantly developed. Only in this connection statistics will be elevated to a science in the full meaning of this word and only thus completely useful and reliable results will be achieved. Yes, we dare predict that in some future century mathematical statistics will solve problems even about whose formulation we now have not the slightest idea.

What we said about statistics in general does not concern all its parts in the same way. Preliminarily, we may say that only population statistics offers points of contact for mathematical treatment whereas all the other parts of statistics shall in the meantime remain where they are[8]. In population statistics, some steps had been actually taken, and they can be understood as an attempt to originate mathematical statistics. However, they are so insufficient and so little correspond to the aids which analysis in its present state is able to offer, so that we may conclude that those steps had not been made at all and that the attempts ought to be made once more from the very beginning.

The first and the most proper notion with which the mathematical treatment of statistical data on population begins is the notion of mortality, since it is connected with everything concerning population. The solution of the problem of mortality of a given population or any

given society has been attempted by compiling the so-called *mortality tables*. They show the order of extinction in that population/society under the assumption that the mortality at some moment persists in the future. We find many such tables in statistical textbooks but all of them are of a more or less doubtful value. Many attempts have also been made to register those remaining alive as shown in a mortality table by a function with the age as the argument. However, these attempts had to be unsuccessful since the basis for the pertinent investigations did not have the reliability which was justly ascribed to the Tyho observations[9]. Even the function $y = a$ to the power of b^x suggested by Gompertz (1825), which found many followers and led in its application to the logarithm integral and gamma function, will remain a very doubtful hypothesis until sufficiently reliable data verify it[10].

Actually, it is no exaggeration to state that all the existing mortality tables without any exception are far from the desirable degree of reliability. Indeed, the reason is partly that their sources were inadequate and partly that the methods which had been applied for their compilation were also inadequate.

Concerning the first point: until now, censuses of population had ben carried out so unsatisfactorily that only rough approximate calculations were possible. This is true at least about the censuses in the region covered by the German Zollverein[11] whereas in Belgium and France where more money is being spent on them they seem to provide better results. We (1863/1864) have studied the necessary scientific requirements for censuses and do not repeat them. We will only say that the public regrettably has been little attentive. For example, women are known to decrease generally their age. In the latest census in Hannover two women refused to name their age so that the registrar was only able to estimate it. Under such circumstances it will possibly be better to abstain from complete and reliable censuses[12].

Insurance of life, pension and widow funds which are based on longevity provide better materials. However, until now only a very small number of the institutions of insurance had made generally known the statistical material hidden in their books for scientific applications. We therefore have much less mortality tables compiled from such sources than should have been expected. And, just as well, these tables are also unreliable since they are compiled by inadequate methods (see above). Even the two tables which are now thought to be best, and are in general use, the table of Brune based on the experience of the widow fund Allgemeinen Witwenverpflegungs-Anstalt in Berlin, and the table of seventeen English societies, are not exempt from those reproaches. What became known about them does not at all comply with the requirements of a proper theory.

It cannot be denied that the compilation of mortality tables about which statistics had until now taken care of seems completely justified when they are clearly printed in a generally understandable way and allegedly characterize the movement of the population of a given society at a certain moment. In essence, such tables nevertheless

cannot satisfy science since they do not constitute the beginning of a scientific study (Fischer 1860).

For the fundamental notion of the theoretical population statistics with which the treatment of statistical materials is raised [to the proper level] much more important is the notion of the *probability of dying*, or, more precisely, *the probability of a person who belongs to a certain group, to die during a specified period*. Up to now, statistics does not have this notion since in general, as hinted above, it does not correctly understand the notion of the theory of probability with which something essentially new is appearing.

If so desired, we can also discuss *the probability of remaining alive after a given period* which complements the former probability to unity. After one of these probabilities is known for each age only a small amount of work is still needed for compiling a mortality table.

The analytical aids which make the described investigation necessary are so completely prepared by Laplace in his *Théorie analytique des probabilités* that it is surprising that they had not been applied previously[13]. However, we ought to recognize that only quite recently this necessity became acute since various new attempts had been made to solve the pertinent problems. Actually, they managed to make a step forward but did not penetrate the core of the matter. We will attempt to make yet another step by means of the same aids and remark the following which has been overlooked for the time being.

Suppose that from a group of L people L_1 remain alive after one year during which time no one either entered or left it. It has been then stated that the probability sought was

$$W = L/L_1.$$

It was not forgotten to add that, if for another group living under the same circumstances the value of W can change. So then we should logically conclude that the calculation above is inaccurate. Exactly for this reason students assume the strange opinion that the theory of probability provides inaccurate results. Actually, however, the entire reasoning was faulty. The value of W as calculated above does not at all claim to be the *real* value of the unknown probability; W is only its *most probable* value[14]. That value is accompanied by a *probable* or, if so preferred, by *a mean error* which has never been determined.

Then, when W is given, and out of new group of   people living under the same conditions only $_1$ remain alive after a [after the next] year, it has been thought that

$$_1 = W.$$

This calculation des not allow for the possibility that actual observation can provide a number differing from $_1$ so that the same reproach can be made once more. $_1$ is not the real but the most probable value of the number of survivors and once more it is accompanied by a probable or mean error which no one has until now determined.

These cases indicate two gaps. It will be in vain to look for their elimination in the present day statistics. To achieve this aim and go on further is the goal of my investigation whose result is entirely new. The degree of its precision and its possible application become immediately evident so that it seems unnecessary to discuss this subject here.

Notes

1. See Bibliography.

2. Wittstein left out Graunt, Petty and Halley.

3. This is wrong. Beginning with Graunt statistical discoveries had not at all been restricted by *looking at numbers*. The same is true about the application of the statistical method to natural science (Sheynin 2009, § 10.9).

4. Wittstein included astronomy in the field of statistics. Although stellar statistics belongs to astronomy, it is also the application of the statistical method to that science. Cf. Quetelet (1846, p. 275): natural science is alien to statistics!

5. This is wrong. Astronomy became a science in antiquity. A bit below Wittstein mentioned the mathematician Abraham Gottheft Kästner (1719 – 1800).

6. No science is ever completed.

7. Say (1803; later editions) offered an example: in a report of 1818, at the time of disaster, when the commerce was ruined and all kinds of resources were rapidly declining, the French minister of the interior boasted that he had proved by numbers that the country was prospering more than during any other time. T. W.

Lies, damned lies and statistics! It was L. H. Courtney who introduced this saying in 1895. I (2003) thought so and recently noted the same in Wikipedia. O. S.

8. This is wrong, cf. Note 3.

9. This is a strange statement. Did Wittstein really believe that statistics will ever attain such precision?

10. It is now generally thought that no such universal function may exist.

11. This was the customs union of most German states. It existed from 1834 to at least 1871.

12. Once more, a strange statement. The cause of Wittstein's pessimism was rather weak: censuses had always encountered great and numerous difficulties. Nevertheless, they had been carried out even in the 18th century, see for example Nordenmark (1929) and Sheynin (2009, § 10.4) about Wargentin and Buniakovsky respectively.

13. It is astonishing to read Laplace's statement (1814/1995, p. 81): *There is a very simple way of constructing* [mortality tables] *from the registers of births and deaths*.

14. The *most probable value* evidently means that many groups ought to be studied and the mean value of that magnitude, *W*, calculated.

Bibliography

Czuber E. (1903), *Wahrscheinlichkeitsrechnung*, Bd. 2. Leipzig – Berlin, 1921.

Fischer Ph. (1860), *Grundlagen des auf die menschliche Sterblichkeit gegründete Versicherungswesen*. Oppenheim am Rhein.

Guillard A. (1855), *Eléments de statistique humaine ou démographie comparée*.

Laplace P. S. (1814, French), *Philosophical Essay on Probabilities*. New York, 1995. Translated by A. Dale.

Nordenmark N. V. E. (1929), Pehr Wilhelm Wargentin, 1717 – 1783. *Nordic Stat. J.*, vol. 1, pp. 241 – 252.

Quetelet A. (1846), *Lettres sur la théorie des probabilités*. Bruxelles.

Say J.-B. (1803 and later editions), *Traité d'économie politique*.

Schlözer A. L. (1804), *Theorie der Statistik*. Göttingen.

Sheynin O. (2003), Lies, damned lies and statistics. *Intern. Z. f. Geschichte u. Ethik der Naturwissenschaften, Technik u. Med.*, Bd. 11, pp. 191 – 193.

--- (2009), *Theory of Probability. Historical Essay.* Berlin. **S, G,** 10.

Wittstein Th. (1862), *Archiv f. Math. u. Phys.*, p. 267 – . Cited by Czuber (1903/1921, p. 208).

--- (1863/1864), Zur Bevölkerungs-Statistik. *Z. d. Kgl. Preuss. Stat. Bureau*, 3. Jg., No. 1.

--- (1867), The book whose first chapter is here translated. New edition: Univ. of Michigan, 2008. English translation: *J. Inst. Actuaries*, vol. 17, 1872, pp. 178 – 189, 353 – 369, 417 – 435. I had noticed this translation too late, but, anyway, it surely has almost no commentaries.

--- (1881, German), The mathematical law of mortality. Ibidem, vol. 33, 1897, pp. 399 – 411. Abridged translation.

IV

B. V. Gnedenko

The development of the theory of probability in Russia

A typical feature of the development of scientific thought during the latest decades is a rapid advance of statistical concepts in various branches of natural science. By now, it is quite definitely ascertained that the application of stochastic methods to the study of fundamental problems in physics, chemistry, astronomy, biology as well as in technology became imminent. The great progress achieved by the theory of probability during a comparatively short period was indeed due to such applications. And for the same reason the interest in probability has recently considerably heightened.

Not long ago probability occupied an isolated position among mathematical sciences, so neither its problems, nor the methods of solving them found any response from the other branches of mathematics. No deep and far-reaching connections with those branches had been yet established. The theory of probability had been restricted to the solution of separate barely interconnected problems. It seemed that mathematicians were barely interested in essential mathematical problems. Consequently, probability was considered a minor science hardly able to formulate serious general problems.

Even the considerable advances of probability in its applications to various fundamental problems of natural science (theory of errors, kinetic theory etc.) did not change the formed outlook, the less so since the very legitimacy of the application of probability as a method of scientific cognition was questioned. The main notions of probability theory were very imperfect and gave rise to quite justified criticism. The passion for this theory during the first quarter of the 19th century connected with Laplace and Poisson led to the appearance of numerous studies devoted to its application to various problems of natural science and social life. However, many of them had been so poorly justified that later became interpreted as an *opprobrium of mathematics*.

That previous passion had thus been replaced by a deeply felt disappointment in, and utter scepticism about the theory of probability as a method of scientific knowledge. European mathematicians began to consider it mostly as a peculiar mathematical entertainment hardly meriting serious attention. Indeed, Klein (1926 – 1927) had paid no attention to its advancement.

Roughly speaking, only after the WWI it became impossible to continue to ignore probability. The development of modern statistical physics, quantum mechanics and other branches of natural science resulted in the notion of the statistical nature of most of the natural laws caused by the discrete structure of matter.

It was also ascertained how necessary a prompt development of the theory of probability was for the further progress of the entire natural

science. And in many countries, in the first place in the Soviet Union, in France, Sweden, Italy, in the U. S., separate scientists and groups of scholars began in earnest to develop the problems of probability theory. At present, the isolation of probability from the general course of the development of mathematics is to a large extent surmounted. During the latest decades far-reaching connections with other mathematical disciplines, with the set theory and the metric theory of functions of real variable in the first place has been established. The scepticism about the possibility of applying probability theory as a method of scientific study of nature was also overcome. The transformation of probability into a harmonious mathematical discipline with a wide range of its own problems, logically perfect main notions and initial principles assisted all that.

The role of Russian science in the general progress of the theory of probability is very essential. The most eminent scholars devoted a noticeable part of their efforts to its development. Moreover, beginning with Chebyshev's time, that theory became as though a national Russian science. From that time onward, almost all of its main ideas which stirred scientists have been advanced, developed or matured in our country. The Chebyshev tradition of a serious and rigorous mathematical attitude to the problems of probability theory just like to those of any other branch of mathematics, has been attentively maintained. This circumstance allowed the Russian scientists to avoid periods of depression and disillusionment which happened in the West.

The considerable influence of our national scientists on the entire course of the development of probability is to a certain extent testified by the large number of references to recent Soviet contributions. This circumstance is explained first and foremost by the fact that the modern chapters of probability had originated and were created in our country.

My aim is to reveal the role of Russian scientists in the transformation of the theory of probability from its infantile state into a ripe and logically harmonious discipline. Understandably, I have to restrict my account to a very general characteristic of the directions of these studies rather than to delve into detailed descriptions of particular and even fundamental investigations. I left aside a great multitude of factual results, sometimes important for developing one or another point, and many researchers are not mentioned.

1. The theory of probability before Chebyshev

For better to appreciate the contribution of Russian scientists to the development of probability it is necessary to describe, at least in a few words, its state up to the mid-19th century when Chebyshev's first investigations had appeared and determined the direction of studies for several generations of scientists.

The origin of the theory of probability, which occurred in the second half of the 17th century, is connected with Pascal, Fermat, Huygens and especially Jakob Bernoulli. That initial period took place in the second part of the 17th century [?] without any Russian participation.

Having originated during the reign of the deterministic mechanical viewpoint on natural science, the theory of probability had not found serious concrete material for the development of its main initial notions and games of chance constituted its initial field. Nevertheless, probability did not become a simple aid for gamblers but its main notions had been affected. Probability had only been interested in finite sets of possible events.

Apart from the definition of probability and establishment of the main elementary propositions (the addition and multiplication theorems, the formulas of total probability), there appeared a series of limit theorems [?]. At the very beginning of the 18th century the Swiss Jakob Bernoulli published a remarkable theorem […]. Its importance is caused by the possibility of establishing a connection between the results of an experiment and a theoretical coefficient, the probability p and in particular, the possibility to form an opinion about p if it is unknown.

Understandably, his theorem does not yet state that the empirical probability μ/n will necessarily tend to probability p when the number of the trials n becomes infinite. The number μ, the frequency, depends on chance and the deviations of μ/n from p can be very considerable.

The French [the English] mathematician De Moivre found out the probability of the various values of the difference between those probabilities for the case of $p = 1/2$ […]. The De Moivre – Laplace theorem became the second main limit proposition of the theory of probability. […]

At the beginning of the 19th century [in 1836 and 1837] Poisson, the celebrated French mathematician showed that the Bernoulli theorem can be derived from a more general proposition which he called the law of large numbers (LLN). […] The problems of artillery firing which led him to the generalization of the LLN also caused his discovery of a new limit theorem. If, in the Bernoulli theorem, the probability of the occurrence of the studied event is low, then, for large values of n, the probability of $\mu = m$ will approximately be

$$\frac{a^k e^{-a}}{k!}, \; a = np.$$

A large section of the theory of probability was connected with the Bayes theorem which allows the calculation of the probabilities of the circumstances of the observed occurrence of the studied event. A voluminous literature had emerged. It was devoted to the calculation of various social phenomena, for example, of the probability of the correctness of verdicts, but it did not positively influence the development of science.

Cotes, Simpson, Legendre, Laplace, Gauss and other scientists [Lambert!] created a chapter important for applications, the theory of errors. However properly are the observations conducted, an absolutely precise result is known to be impossible. Errors depending on chance are unavoidable. Laplace and Gauss showed that, once the principle of the arithmetic mean is adopted, those errors ought to obey the normal distribution [all wrong]. This principle states that the

arithmetic mean of the observations is the most probable value of the measured magnitude.

The notion of expectation originated in connection with problems in games of chance and the definition of a just game. [Expectation of a discrete random variable with a finite number of possible values is written out.] By the mid-19th century the main properties of expectations became sufficiently well known. At the same time Laplace [who died in 1827] developed the method of generating functions, the prototype of one of the present mightiest analytic tools of the theory of probability, the method of characteristic functions.

Laplace and Poisson completed a long and fruitful initial period in the development of the theory of probability, the period of the philosophical comprehension of its initial notions and of the first attempts to study natural phenomena by its methods. As we stated above, this period led the western scientists to regard the theory of probability less than coldly and to deny its possibility of studying natural phenomena. And thus the period of the stagnation of probability had begun in the West [with no probability to stagnate in Russia].

2. The first studies in probability in Russia

Scientific studies in probability began in Russia after the establishment of the Academy of Sciences and the arrival of Daniel and Niklaus Bernoulli [Daniel was younger] and Euler in Petersburg from Switzerland. For Russia, however, their mathematical contributions remained alien, unconnected with the general level of Russian science and culture. As a reminder of that initial period during which Russia joined stochastic studies history had retained the name of the known paradox, the Petersburg game.

Later all the most eminent representatives of the Russian mathematical life became interested in the theory of probability. Lobachevsky (1835 – 1838), when desiring to ascertain geometrically which geometric system was governing the universe, developed the theory of errors on a sphere. Later he (1842) published a paper devoted to that problem. Ostrogradsky published three ordinary papers on the theory of probability which did not do justice to his mathematical talent. In an introduction to one of them (On generating functions, unpublished) he noted that his problem was *practically important for accepting commodities*. This remark characterizes his entire work: he believed that the progress of the theoretical science was inseparably linked with practical applications of its results.

Buniakovsky published a large number of papers on the theory of probability and especially on its application to statistical problems, insurance and demography. He (1846) also compiled a textbook on probability, excellent for its time. However, none of these studies had contributed any essential new ideas or problems, nor did they create a powerful school of researchers, but they aroused the interest in probability among both the beginners and ripe Russian scholars.

An important step which opened a new page in science was due to Chebyshev. Here is what Kolmogorov (1947, p. 56) stated:

From a methodological aspect, the principal upheaval accomplished by him consisted not only in that he was the first to demand, with categorical insistence, absolute rigour in proving limit theorems (the derivations provided by De Moivre, Laplace and Poisson were not at all irreproachable from the formal logical point of view, although Jakob Bernoulli proved his limit theorem with an exhaustive arithmetical rigour).

The main point is that in each instance Chebyshev strove to determine exact estimates of the deviations from limit regularities taking place even in large but finite numbers of trials in the form of inequalities unconditionally true for any number of these.

Furthermore, Chebyshev was the first to appreciate clearly and use the full power of the concepts of <u>random variable</u> and its <u>expectation</u> (mean value). These notions were known earlier and are derivatives of the main concepts, <u>event</u> and <u>probability</u>. However, they are subordinated to a much more convenient and flexible algorithm. This is true to such an extent that we now invariably replace the examination of event A by consideration of its characteristic random variable $_A$ equal to unity when A occurs and to zero otherwise. The probability P(A) of event A is then nothing but the expectation E_A of $_A$. Only much later the appropriate method of characteristic functions of sets came to be systematically used in the theory of functions of a real variable.

Chebyshev had only devoted four papers to probability, but it is difficult to overestimate their influence on the subsequent development of science. The development of his ideas is still being continued whereas a full solution of his formulated problems has been only obtained during the latest ten – fifteen years.

His ideas had not found an immediate response from scientists abroad, but they became a stimulus for the creation of the Russian school of probability. According to the essence of its investigations, the activity of this school can be separated into two periods. The first one is connected with the names of mostly Petersburg academicians: Chebyshev himself, Markov, Liapunov, and, later, as it seems, Bernstein and Romanovsky.

During that period investigations concentrated around two main subjects: at first, the patterns of independent events, then Markov chains. In this first period I also include the wide initial attempts of academician Bernstein to put into logical order the theory of probability and mathematical statistics.

The second period began after the revolution of 1917. It was connected with the work of a group of Moscow mathematicians, of Khinchin, Kolmogorov, Slutsky and their students. During this period the ideas and methods of the theory of functions of a real variable were included, the scope of probability widened and mathematical analysis became much more applied. These groups were called the Petersburg and the Moscow schools of probability. I keep to this distinction although it is extremely tentative.

3. The Petersburg school of probability

Its origin occurred at the time when Western mathematicians lost confidence in the possibility of widening the scope of probability beyond the narrow confines of games of chance and transforming it into a serious mathematical discipline [nonsense]. The great merit of Chebyshev and his students, Markov and Liapunov, consists in that they were able to reveal useful indications for developing methods of investigating random phenomena in all their multifarious forms from the mistakes of the past and the scepticism of their contemporaries.

Unlike their predecessors Petersburg mathematicians absolutely clearly formulated their statements and admitted the application of the theory of probability only to such phenomena whose occurrence can be experimentally checked. Consequently [?], they brought forward the establishment of inequalities and the estimation of the errors of approximate formulas.

Chebyshev (1846) early remarked in connection with his proof of the LLN:

Regardless of the cleverness of the method used by the celebrated geometer [Poisson] *he does not give the limit of the error* […].

That same principle, as stated above, guided the later investigations of Markov, Liapunov and Bernstein as well.

The main efforts of the representatives of the Petersburg school had been directed towards the investigation of the pattern of sequences of independent random variables. Only later they generalized their results to include Markov chains. The former pattern meant that the studied phenomenon depended on chance and occurred as a result of the action of a very large number of independent causes, each of which only exerted a negligibly small influence $_k$, $k = 1, 2, …, n$. Their joint influence is

$$S_n = {}_1 + {}_2 + … + {}_n. \qquad (3.1)$$

With a large n the probability distribution of the separate terms is, generally speaking, unknown, and the determination of the distribution of S_n is very difficult and sometimes impossible. It is therefore necessary to determine only the limit distribution instead of the real one.

This pattern allowed a satisfactory solution of a great number of various natural scientific problems. Already Laplace assumed that the occurrence of a definite (of the normal) law of distribution which governs the random errors of measurement [?] is due to a deeper cause rather than to the arbitrary principle of the arithmetic mean. He thought that this fact is explained by the composition of the errors under the influence of a large number of independent causes, and he originated the very pattern of the sequence of independent random variables. However, neither he, nor his followers had justified or developed this idea. It was done by the Petersburg school.

In the first decades of its existence the efforts of its representatives had been concentrated on the determination of the most general conditions under which the LLN and the central limit theorem (CLT) became valid; that is, on the main problems of the pattern of sequences of independent random variables.

3.1. The law of large numbers. Chebyshev defended his first study in the theory of probability (1845) as a master dissertation. It was N. D. Brashman, professor of Moscow University, who awakened his interest in this subject. Chebyshev was able by most elementary methods, with only the expansion of $\ln(1 + x)$ into a Maclaurin series taken from mathematical analysis, to prove the Bernoulli theorem and estimate the error of the obtained approximation.

Next year he generalized his result on the LLN in its Poisson form. These initial publications served him as a point of departure for widely generalizing that law and creating a method of proof remarkable for simplicity and power (1867) included since then in each textbook on the theory of probability.

Following Markov, we understand the LLN as the totality of propositions which state that, with probability tending to unity, the arithmetic mean of random variables

$$\xi_1, \xi_2, \ldots, \xi_n \tag{3.1.1}$$

tends as n to the arithmetic mean of their expectations. Apart from the purely mathematical interest of ascertaining the conditions under which the sums of random variables become almost invariable, independent from chance, such theorems are most important for natural science. Here are examples.

If magnitudes (3.1.1) devoid of systematic errors (only corrupted by random errors) are the results of equally precise measurements of some constant a, then, by the LLN, with a sufficiently large n, with probability tending to unity their arithmetic mean will no matter how little differ from a.

Second example: the pressure of a gas on a wall of its vessel. According to the notions of the kinetic theory, it is the result of the hits of separate gas molecules against the wall. The velocities of the molecules are random, occasioned by a great number of their collisions, so the pressure will undergo random fluctuations since the number and the power of those hits are random. Experiment, however, shows the opposite (the Pascal law). The reason? The LLN provides an exhaustive answer.

I formulate now a few theorems concerning the LLN.

The Chebyshev theorem (1867). If the sequence

$$\xi_1, \xi_2, \ldots, \xi_n, \ldots \tag{3.1.2}$$

of mutually independent random variables is such that for every value of n, $n = 1, 2,\ldots$, and some constant c,

$$\mathrm{var}\,\xi_n = E(\xi_n - E\xi_n)^2 \leq c,$$

then, for any $\varepsilon > 0$, the probability of

$$\frac{1}{n}\left|\sum_{k=1}^{n}(\xi_n - E\xi_n)\right| < \varepsilon \tag{3.1.3}$$

is higher than $1 - c/n^2$.

It is easy to see that the Bernoulli and the Poisson theorems are simplest particular cases of this theorem. Indeed, it is sufficient to introduce random variables ξ_k equal to 1 or 0 depending on whether the studied event occurred or not in trial k.

Markov essentially widened the conditions under which the LLN is valid. Here are two of his results (1924) about sequence (3.1.2).

His first condition:

$$\frac{1}{n^2} \sum_{k=1}^{n} \mathrm{var}\, \xi_k \to 0 \text{ as } n \to \infty.$$

His second condition: if there exists such a $\delta > 0$ that for any k and $c > 0$

$$\mathrm{E}|\xi_k|^{1+\delta} \le c,$$

then, as $n \to \infty$, the probability of inequality (3.1.3) tends to unity.

Later investigations by Kolmogorov showed that Markov's first condition was almost necessary. Practical applications of the Chebyshev and Markov theorems required that n ought to be very large, much larger than demanded by the application of the LLN. Bernstein (1937) inserted some additional restrictions not practically burdensome and obtained incomparably more precise results. Denote

$$b_k = \mathrm{var}\, \xi_k,\ B_n = b_1 + b_2 + \ldots + b_n.$$

The Bernstein theorem. Then, if the sequence (3.1.2) of mutually independent random variables is such that for some $H > 0$ and all $k \ge 2$ and $n \ge 1$

$$\mathrm{E}|\xi_n - \mathrm{E}\,\xi_n|^k \le \frac{b_n}{2} H^{k-2} k!,$$

the probability of the inequality (3.1.3) will be higher than

$$1 - \exp(-\frac{\varepsilon^2 n^2}{4B_n}) \text{ for each } \varepsilon \le B_n/nH;\ \exp a \text{ means } e^a.$$

3.2. The central limit theorem. This term covers propositions which assert the conditions under which the distribution of the sum of independent random variables tends to the normal distribution with the increase of the number of terms of such sums. At present, this theorem became one of the main instruments of mathematical statistics and mathematical natural science. Its great importance consists in that an immense number of natural phenomena can be considered as occurring under the impact of a very large number of random causes, each of which only negligibly influencing the course of the phenomenon.

The allowance for the impact of each of these causes or even their simple enumeration is indeed greatly difficult and perhaps even impossible in principle. It is obviously important, therefore, to develop methods of studying their general influence irrespective of the essence of each separate term.

The CLT established by the contributions of Chebyshev, Markov and Liapunov states that that summary impact only negligibly differs from the normal law. Here are examples.

Artillery firing. Deviations of the hit points of a missile from the point of aiming [when laying directly or not] represent the well known phenomenon of dispersion of the missiles. It is the result of the influence of many independently acting causes each of which etc.

Another example. Random deviations from the standard allowed in the manufacturing of machine parts. These deviations are also caused by the summary action of a great many small causes each of which etc.

Last example. The velocity of a gas molecule is the result of the action of other molecules each of which etc. The result is known as the Maxwell law.

I will now formulate the results of the representatives of the Petersburg school concerning the CLT.

The Chebyshev theorem (*1887*). If the expectations of the variables (3.1.2) are zero and the absolute values of the expectations of all of their powers are less than some finite limit, then as, n , the probability that the sum (3.1) divided by the square root of the sum of their variances is located between some magnitudes t_1 and t_2 tends to

$$\frac{1}{\sqrt{2}}\int_{t_1}^{t_2}\exp(-\frac{z^2}{2})dz.$$

Chebyshev had not provided a rigorous proof of his theorem, he did not even introduce necessary restrictions. However, he still is greatly meritorious in that he had emphasised the importance of his problem, devised a method of its proof (the method of moments) and interested his students who concluded the proof of the CLT and extended its conditions almost to their natural boundaries.

In two publications Markov, see Bernstein (1945), was able to formulate quite correctly and to prove a proposition more general than that of his mentor. He applied the Chebyshev method of moments.

The Markov theorem (*1898; 1899*). If the sequence of mutually independent random variables (3.1.2) is such that for all natural values of   3

$$\lim\frac{c_n(\ )}{B_n}=0 \text{ as } n\to\infty,$$

$$B_n^2=\sum_{k=1}^{n}\text{var}\ _k,\ c_n(\ )=\sum_{k=1}^{n}|\ _k-\text{E}\ _k|\ ,$$

Then, as n ,

$$\lim P[\frac{1}{B_n}\sum_{k=1}^{n}(\ _k - E\ _k) < x] = \frac{1}{\sqrt{2}}\int_{-\infty}^{x}\exp(-\frac{z^2}{2})dz. \qquad (3.2.1)$$

In two papers Liapunov (1900; 1901) proved the CLT under much less restrictive conditions, and, first of all, he had not required the existence of the moments of all orders of the random variables $_k$.

The Liapunov theorem. If there exists such a constant > 2 that, as n ,

$$\frac{c_n(\ )}{B_n} \to 0, \ c_n(\ ) = \sum_{k=1}^{n}E|\ _k - E\ _k|\ , \ B_n^2 = \sum_{k=1}^{n}\text{var}\ _k,$$

then

$$P[\frac{1}{B_n}\sum_{k=1}^{n}(\ _k - E\ _k) < x] \to \frac{1}{\sqrt{2}}\int_{-\infty}^{x}\exp(-\frac{z^2}{2})dz.$$

In addition, Liapunov determined the upper boundary of the error caused by the replacement of the exact law of distribution by the limit law. Many works were later devoted to the specification of these estimates [?] and here, first of all, we ought to mention the Cramér studies (1937).

When proving his theorem, Liapunov rejected the method of moments and, instead, developed and applied the method of characteristic functions, which, as it seems, was first used by Lagrange. It has many common features with the Laplacean method of generating functions. They, however, exist by far not for each random variable whereas characteristic functions exist for all of them.

Multiplication of characteristic functions corresponds to the addition of the pertinent random variables and after Liapunov the method of characteristic functions therefore became the main one for solving problems on summation of random variables. It is a very powerful method and it seemed that the method of moments is not capable of such general results as those achieved by Liapunov. Indeed, for one thing, that method of moments requires that the random variable has moments of all natural orders. And Markov (1900/1913, p. 332) stated that he *had thought for a rather long time how to re-establish the shaken importance of the method of moments*.

He coped brilliantly by applying a very witty method, by curtailing the random variables, a procedure which has become standard since then. Instead of the given sequence of random variables (3.1.2) he proposed to consider variables $'_k$:

$$'_k = \ _k \text{ if } |\ _k|\quad N, \text{ and }\quad '_k = 0 \text{ if } |\ _k| > N.$$

The number N is arbitrary, and for its sufficiently large values the equality $'_k = \ _k$ is almost certain. And these new variables have

moments of all orders so that the Markov theorem (above) is applicable to them.

Twenty years had passed after the work of Liapunov and Markov, and only after that period Lévy and Lindeberg exceeded them. However, their contributions actually stressed once again that those Russian scientists had provided all the main results. Later investigations have only to a small extent widened the conditions of the application of the CLT.

Bernstein (1926) published a classical study in which the results of Markov and Liapunov were transferred on a wide class of weakly dependent random variables and on a system of random vectors. His method was, after all, a synthesis of those of his predecessors. It is important to note that Bernstein considered the CLT for independent random variables in sequence (3.1.2) in an essentially new form: he had not anymore required the existence of either expectations or variances. Here is his problem.

Determine the conditions under which there exist such constants A_n and $B_n > 0$ that the laws of distribution of the sums

$$\frac{1}{B_n} \sum_{k=1}^{n} {}_k - A_n \qquad (3.1.4)$$

tended to the normal law as n .

The Bernstein theorem. If, for a given system (3.1.2) of independent random variables and some constant magnitudes c_n,

$$\sum_{k=1}^{n} P[| {}_k | > c_n] \to 0, \quad \frac{1}{c_n^2} \sum_{k=1} \int_{|x| > c_n} x^2 dF_k(x) \to 1,$$

where $F_k(x)$ is the distribution function of ${}_k$ and > 0, then there exist such constants A_n and $B_n > 0$ that the distribution functions of the sums (3.1.4) tend to the normal law as n .

Later Feller (1935 – 1937) showed that the conditions adopted by Bernstein were not only sufficient, but, under a very natural requirement that each separate term exerted a small influence, they are also necessary.

3.3. Markov chains. Markov (1906) began investigations which opened up new paths for the theory of probability as well as new possibilities for mathematically describing most various natural scientific problems. Instead of the Chebyshev pattern of a sequence of independent random variables he considered variables connected in a special way and proposed to call these connections *chains*. The name *Markov chains* has since been established in science.

As it seems, Galton who attempted to formulate mathematically the Darwinian theory of heredity by issuing from immense factual data was the first to consider such connections. The English physicist Rayleigh and, in some particular cases Poincaré had also investigated a pattern of the same kind. However, none of them provided any satisfactory theory, they only considered separate examples. That

several scientists came to the same idea of chains only stresses their importance and vital interest for the entire natural science.

A sequence of random variables (3.1.2) is called a *simple* Markov chain if the probabilities of the values taken by any n only depend on the values previously only taken by that variable in the preceding trial rather on the number of the trial. The chain is called *complex* if those values only depend on the values taken by some *k* preceding values.

Markov had not mentioned any problems of natural science whose study required the application of his pattern, he only created a purely mathematical theory of chains and was not interested in their applications. As an illustration of his theory he only studied the interchange of vowels and consonants in two Russian classical texts.

Later Fokker, Planck, Smoluchowski and Einstein noted the significance of Markov chains for physics. At present, a great number of physical theories depart from the Markov pattern. Suffice it to mention quantum mechanics, the theories of diffusion and Brownian motion.

The Chebyshev pattern is unable to cover, say, the phenomenon of diffusion if the observations of the molecular movement are made very soon to each other: the results of neighbouring observations are then strongly interconnected.

The problems which Chebyshev had formulated for his pattern immediately occurred for the pattern of the chain dependence. The generalization of the LLN had not encountered serious difficulties but the proof of the CLT became incomparably more difficult. Indeed, the method of moments which Markov applied required the calculation of the expectations of all the natural orders of the sums

$$\frac{1}{B_n}\sum_{k=1}^{n}(\,_k - \mathrm{E}\,_k),\ B_n^2 = \sum_{k=1}^{n}\mathrm{var}\,_k$$

and a proof that they tend to the moments of the normal distribution as n . All this demanded cumbersome work with which Markov was only able to cope owing to his exceptional analytic mastery. He based the poofs of these propositions on the decrease of the connection between the terms of the sequence with the increase of the distance between them.

Markov used the same fact for the derivation of a theorem which became the prototype of the now numerous so-called ergodic theorems. He proved that with the increase of *n* the probability distribution of n tends under some conditions to a certain limit distribution which does not depend on the value of the initial variable $_1$ (a distant state of a system hardly depends on its initial state). This theorem is important in natural scientific applications of the chain theory.

Romanovsky (1930; 1932; 1935) who studied Markov chains with a finite and an uncountable set of states essentially forwarded the theory of chains. He connected the former case with the matrix theory and, for obtaining stochastic results, widely developed algebraic tools as well. At present this is one of the main methods in the theory of chains and is widely applied by Soviet and foreign scientists. His student

Sarymsakov (1942) transferred the Romanovsky method on the investigation of Markov chains with a countable set of states. And Romanovsky (Ibidem) had also constructed important patterns of the dependences of random variables which generalized Markov chains.

Kolmogorov (1937) studied chains with a countable set of possible states and inserted new ideas and new problems. Many important work on the chain theory has been done in France (Hadamard, Fréchet, Doeblin, Fortet et al), in [former] Czechoslovakia (Hostinsky), and Rumania (Mihoc and Onicescu). We also ought to mention important studies on the border with the theory of completely determinate processes (N. M. Krylov, N. N. Bogolubov).

3.4. Markov: his *Calculus of Probability*. We ought to dwell on that excellent book which was as though the sum total of many years of work in the Chebyshev direction. It was written as a textbook and ran through four editions from 1900 to 1924. It is easily understood by beginners and at the same time it is a most interesting monograph for readers who master stochastic methods.

Markov began his account by describing the main notions of probability and illustrating them by plenty most various examples with the same strict regard to the simplest and the most complicated problems. He leads the reader to the summit of probability and acquaints him with his own investigations. His treatise remains one of the best in the world in spite of the $40 +$ years which have passed since 1924. The literary and scientific worth of Markov's book has been appreciated at home and abroad and it was soon [?] translated into German [in 1912].

We are now acquainted in general with the state of probability in Russia up to 1917, and now it is necessary to mention that we have only dwelt on few main ideas and results […]. We were unable to provide specimens of that remarkable mastery of analytic calculations peculiar to the representatives of the Petersburg school. I note finally (Kolmogorov 1947, p. 59) that

Only with a considerable delay, in the 1920s or even 1930s, the importance of the work of Chebyshev, Markov and Liapunov was quite appreciated in Western Europe. Nowadays it is everywhere perceived as the point of departure for the entire further development of the theory of probability. In particular, the main Liapunov limit theorem and the theory of Markov chains were exactly what was most of all needed for a reliable substantiation of the developing statistical physics. That the West had slowly adopted the ideas of the Petersburg school may perhaps be partly explained by the fact that that school was very remote from statistical physics. […]

Hopefully, this remark will not lead to an impression that the work of the Petersburg school lacked an animated feeling of connection with the requirements of mathematical natural science. In the second half of the 19th century, Russian physics had ben backward and mathematicians of the Petersburg school were not concentrated on those, possibly most interesting applications of probability theory as already been indicated (the work of Boltzmann covered the period of 1866 – 1898).

A keen sense of reality in formulating mathematical problems was especially characteristic of Chebyshev. Issuing from comparatively special elementary and sometimes somewhat old-fashioned applied problems, he elicited from them with exceptional insight such general mathematical concepts that potentially embraced an immeasurably wider circle of technical and natural scientific problems.

3.5. The role of S. N. Bernstein. A period satiated by grand general ideas and factual results is connected with Bernstein. Above, I noted that up to the mid-19th century the theory of probability did not yet become a mathematical science. Its applications to the study of natural phenomena had been rather weakly justified and led to scepticism especially expressed by Bertrand [1888] in his course on probability. Indefiniteness of the main notions of the science of chance led to a number of paradoxes. True, this fact little troubled naturalists: even a naïve stochastic approach in various branches of science resulted in serious success. Time, however, went on, science made higher demands on the theory of probability. It became necessary to study systematically the main notions of that theory and ascertain the conditions under which it was legitimate to apply its methods and results.

Especially important therefore became a formal logical justification of the theory of probability, and Bernstein (1917) was the first to attempt it. He issued from a qualitative comparison of events according to their probability. The numerical expression of probability only appeared as an arbitrary notion. I am not dwelling on his axiomatics since he described it in detail in his widely known treatise (1927 – 1945), I only note that V. I. Glivenko and Koopman have been later developing Bernstein's concept.

Along with formalizing the theory of probability and the intention to put in order its basis, Bernstein set himself a much wider goal: by issuing from the system of axioms of that theory which he was creating, to construct a logically perfect theory of mathematical statistics and show how to study the most important natural scientific problems absolutely rigorously [!] and strictly. Here are his words (1928/1964, p. 218):

A purely mathematical theory of probability may be uninterested in whether or not the coefficient called mathematical probability possesses any practical meaning, subjective or objective. The only demand that must be observed is the lack of contradictions: when keeping to the admitted axioms, under given conditions various methods of calculating this coefficient should lead to one and the same value.

In addition, if we want the conclusions of the theory to admit of empirical checking rather than to remain a jeu d'esprit, we must only consider such sets of propositions or judgements about which it is possible to establish whether they are true or false. The process of cognition is intrinsically irreversible, and its very nature consists in that some propositions which we consider possible become true (i. e., are realized) whereas their negation thus becomes false or impossible.

It follows that the construction of the theory of probability as a single method of cognition demands that the truth of a proposition be

uniquely, without any exceptions, characterized by a certain maximal value of the mathematical probability, which is assumed to be equal to unity, and that the falsity of an assertion be identical with its minimal probability which is set equal to zero.

In many contributions, both mathematical and natural scientific, devoted to theoretical problems in biology, Bernstein issued from these main ideas. They also inspired him to write a course in probability (1927 – 1945), one of the best in the world literature.

He began there by listing a system of axioms and the main notions of that science and discussing examples of calculating probability by various methods. Then he described classical and his own results concerning the LLN, the Laplace theorems [?], sampling, curves of distribution, the theory of correlation etc. Regular discussions of the practical value of one or another theoretical result and the boundaries of its applicability make the book especially fresh and valuable. It essentially assisted the popularity of the book among mathematicians as well as scientists working in natural science, in economic and technological disciplines.

During the first period of his activity, Bernstein's properly mathematical work represented a brilliant completion of the studies made by Chebyshev, Markov and Liapunov on the limit theorems for sums of random variables. His proof of the CLT for independent variables became so general that the restrictions imposed on that theorem have been later shown to be not only sufficient, but necessary as well. At the same time, wide conditions under which the CLT is still valid were established for sums of dependent terms, see above.

Bernstein was the first to study the conditions for the binary CLT. I illustrate the formulation of the problem by a simple but important example. When firing at a certain objective the missiles, generally speaking, scatter and the probabilities of one or another deviation from the centre of the objective ought to be calculated. Construct a coordinate system with the centre of the objective as its origin, and assume that those deviations result from the summary action of a great many causes depending on chance, each of them negligibly influencing it. Bernstein showed that the deviations obey a special law of distribution, the bivariate normal law. It is often said that in such cases x and y are normally correlated.

Bernstein applied this general mathematical result to biology. Among other findings we ought to mention an important and sudden fact [Bernstein 1922]: under very general natural assumptions the Galton law of the inheritance of quantitative indications does not contradict the Mendelian hypothesis [law] but follows from it. And here is Komogorov (1947, p. 60):

With regard to the scope of Bernstein's work only the works of Richard Mises accomplished at about the same time, of the German mathematician now living in the USA, can be compared with it. They both posed the problems of

1. A rigorous logical substantiation of the theory of probability

2. The completion of research into limit theorems of the type of the Laplace and Liapunov propositions which led to the normal law of distribution

3. *The use of modern methods of investigation which possess full logical and mathematical value for covering to the greatest possible extent the new domains of application of the theory of probability.*

In this last-mentioned direction, the activity of Mises, who headed a well-organized Institute of Applied Mathematics [in pre-Nazi Germany] *was perhaps even wider than Bernstein's research. The latter, however, offered many specimens of using stochastic methods in most various problems of physics, biology and statistics. And in the second, purely mathematical direction, Bernstein accomplished his investigations on a considerably higher methodological and technical level.*

4. The Moscow school of the theory of probability

4.1. Its origin. The ideas of the set theory and the theory of functions which were cultivated by N. N. Luzin and his students determined the essence of the first stochastic investigations made by Moscow mathematicians. An attentive study of the main notions of probability theory, of random event, probability, independence of events, random variable, expectation etc. as well as operations on random events showed that far-reaching analogies can be drawn between them and the main notions of the set theory and metric theory of functions.

These connections between so differing fields of science allowed a new illustration of the logical basis of probability theory, the enrichment of its content by new problems and methods of study and the conclusion of the solution of classical problems.

Khinchin (1923; 1924) initiated the creation of the Moscow school by the study of an absolutely distinctive generalization and strengthening of the LLN. The regularity which he discovered was later called the law of the iterated logarithm, see below.

His contribution became a source for further studies in the direction which he indicated. They were made by Soviet (Kolmogorov, Khinchin himself, I. G. Petrovsky, Gnedenko) and foreign (Cramér, Cantelli, Lévy, Feller, Erdös et al) scientists. At about the same time Slutsky (especially 1925) began to create a new chapter of the theory of probability by the methods of the theory of functions of a real variable, a chapter on the theory of random functions, i. e., random variables depending on a continuously variable parameter. He introduced and studied the notions of stochastic limit, derivative, integral, measurability.

Soon Kolmogorov began his stochastic work. His first joint contribution (Khinchin & Kolmogorov 1925) was devoted to the study of the convergence of series of mutually independent random variables (3.1.2). The authors proved that such series can only tend to some magnitude (in general, to a random variable) with extreme probabilities, 0 or 1. In no case can such a series converge with some intermediate probability, say 1/2.

Later Kolmogorov (1936) provided very wide conditions under which events, depending on an infinite set of random variables, can only occur with probabilities 0 or 1. These Moscow studies also found a considerable response from mathematicians in Western Europe.

The classical problems which had interested Chebyshev and Markov has carried away Moscow scientists as well. Especially significant among their studies became the ascertaining of the conditions under which the LLN is valid as well as the specification of that law. It occurred that the methods and notions of the theory of functions enabled them to determine the definitive (necessary and sufficient) conditions of the classical theorems.

Very soon however, after issuing from the Chebyshev and Markov problems, Moscow mathematicians formulated an absolutely new set of problems which nowadays constitute the most rapidly developing and captivating part of the modern theory of probability, see stochastic processes below.

4.2. The law of large numbers. We have discussed the significance of the LLN for the application of mathematical methods to natural science and practical sciences [?]. This fact was indeed the reason for the ever increasing interest in widening the scope of its applicability. Over several decades most eminent mathematicians busied themselves with that problem which was indeed worth their efforts but finally only Kolmogorov (1928) succeeded.

Such conditions, once established, would have immediately answered whether the LLN or its corollaries were applicable in some definite circumstances. Denote $F_k(x) = P[\ _k - \mathrm{E}\ _k < x]$. Then, this is his proposition.

The Kolmogorov theorem. A sequence of mutually independent random variables (3.1.2) obeys the LLN then and only then when, as n , three conditions are fulfilled:

$$1. \sum_{k=1}^{n} \int_{|x|\geq n} dF_x(x) \to 0,$$

$$2. \frac{1}{n}\sum_{k=1}^{n} \int_{|x|<n} x dF_x(x) \to 0,$$

$$3. \frac{1}{n^2}\sum_{k=1}^{n} \int_{|x|<n} x^2 dF_x(x) \to 0.$$

At about the same time Khinchin discovered that if all the random variables of the sequence (3.1.2) have the same probability distribution, $F_1(x) = F_2(x) = \ldots$, then the existence of their expectations is necessary and sufficient for the LLN.

Investigations on the strong LLN belong to the same set of ideas. There, it is required to determine the conditions under which random variables (3) satisfy the equation

$$P[\lim \frac{1}{n}\sum_{k=1}^{n} (\ _k - \mathrm{E}\ _k) = 0,\ n \to \infty] = 1.$$

In 1909, Borel was the first to formulate this problem and to solve it in the particular case of Bernoulli trials and $p = 1/2$. It attracted the

attention of many scholars. The widest results for independent random variables are here due to Kolmogorov, and to Khinchin for the case of dependent variables. Kolmogorov also discovered that for independent random variables with the same probability distributions the existence of expectation was a necessary and sufficient condition not only for the LLN but for its strong version as well.

The analogies with the theory of functions had been essential in all these investigations. For the LLN, the similarity was with the notion of the convergence in measure, and for its strong version, convergence almost everywhere.

We mentioned the law of the iterated logarithm, and here is a description of Khinchin's work for the simplest particular case of the Bernoulli pattern. The number μ of the occurrences of the studied event in *n* independent trials with a constant probability *p* of its occurrence in each of *n* trials satisfies the relation

$$P[|\frac{\mu}{n} - p| < \quad] \to 1, \quad > 0, \ n \to \infty.$$

In 1909 Borel proved a stronger statement for the case of *p* = 1/2:

In 1917 Cantelli proved that for an arbitrary *p*

$$P[\lim \frac{\mu - np}{n} = 0] = 1.$$

This means that with probability 1

$$\mu - np = o(n)$$

Next year Hausdorff discovered a stronger result for *p* = *q* = 1/2: with probability 1 and an arbitrary > o

$$\mu - np = o(n^{1/2+}).$$

And a year later Hardy & Littlewood showed that with the same probability

$$\mu - np = O(\sqrt{n \ln n}).$$

Finally, in 1923 Khinchin (§ 4.1) showed that

$$\mu - np = O(\sqrt{n \ln \ln n}).$$

In 1924 he additionally discovered that any further improvement was impossible. More precisely, he proved

The Khinchin theorem. If in each of *n* independent trials the probability of the occurrence of event *A* is *p*, 0 < *p* < 1, the number μ of its occurrences in those *n* trials satisfies the equation

$$P[\limsup \frac{|\sim - np|}{\sqrt{2npq\ln\ln n}} = 1] = 1, \ q = 1 - p, \ n \to \infty.$$

Geometrically, this can be represented in the following way. Draw a system of coordinates and mark n on the x-axis, and $y = \mu - np$ on the y-axis. The Borel – Cantelli theorem states that for all sufficiently large values of n the ordinates will almost certainly be contained between straight lines

$$y = \quad n \text{ and } y = - \quad n,$$

which means that almost certainly they will not exceed these boundaries.

However, those straight lines are too widely separated and Khinchin showed that there exist narrower boundaries: for any > 0 and sufficiently large values of n the difference $\mu - np$ will almost certainly be contained between

$$y = (1 + \quad)\sqrt{2npq\ln\ln n}, \ y = -(1 + \quad)\sqrt{2npq\ln\ln n}. \qquad (4.2.1)$$

Moreover, that difference will almost certainly infinitely many times exceed the boundaries between curves

$$y = (1 - \quad)\sqrt{2npq\ln\ln n} \text{ and } y = -(1 - \quad)\sqrt{2npq\ln\ln n}.$$

Later Khinchin (1927) generalized his result on the Poisson pattern, then Kolmogorov (1929) showed that the law of the iterated logarithm for the sum of random variables was valid under very general conditions.

Many scientists at home and abroad have been and still are working on further generalizations and specification of the results of Khinchin and Kolmogorov.

4.3. Axiomatics. During those same 1920s, the period of the supremacy of the ideas of the metric theory of functions, Kolmogorov studied the basis of the theory of probability. From 1926 onward he ordered the ideas of the Moscow school into a harmonious logical system and concluded this work by a monograph (1933a). There, he consistently included the fundamentals of the theory of probability, a science until recently so peculiar, in the sequence of the general concepts of mathematics. Before the creation and wide development of the metric theory of functions any attempt to solve such a problem was almost hopeless. Now, however, when the similarity between the measure of sets and probability of an event; between integral and expectation; orthogonality of functions and independence of random variables, etc. became revealed, the necessity of axiomatizing the theory of probability by issuing from the ideas of the set theory had ripened.

Kolmogorov issued from the set E of *elementary events*. For a logical development of the theory of probability the essence of the

elements of such sets is of no consequence. That theory therefore admits a large number of various interpretations including such which bear no relation to the notion of randomness. Understandably, this circumstance only widens the scope of the application of probability.

Thus, a set F of subsets from E and its elements are called random events. We see that according to Kolmogorov the notion of random event is based on a more elementary notion whereas Bernstein assumes it as an initial notion. Random events and their probabilities obey the following axioms.

1. If random events A and B are included in F, then it also contains events A or B, A and B, not A and not B.

2. F also includes E and all its separate elements.

3. A non-negative real number $P(A)$ called the probability of A is attached to each A.

4. $P[E] = 1$.

5. If A and B do not intersect, and belong to F, then

$$P(A + B) = P(A) + P(B).$$

For infinite sets F the following axiom is also supposed to be satisfied:

6. If the intersection of the sequence of events

$$A_1 > A_2 > \ldots > A_n > \ldots$$
is empty, then

$$\lim P(A_n) = 0, n \qquad .$$

For finite sets this axiom follows from the first five.

Note that from the viewpoint of the Kolmogorov axiomatics the notion of random variable as a function of an elementary event is quite natural.

On the basis of those axioms Kolmogorov constructed the elements of the theory of probability. His investigation seriously helped to definitively establish it as a mathematical discipline. It is widely known and generally recognized and its ideas are guiding modern studies of probability theory and mathematical statistics.

4.4. Theory of stochastic processes. The perfection of physical statistics and of many fields of technology raised many new problems for the theory of probability. They were not confined within the boundaries of classical patterns: a physicist was interested in the study of the random processes, i. e., in magnitudes whose random change depended on one or several continuously changing parameters (time, coordinates etc.), but a mathematician could have only offered him tricks valid for discrete sequences, i. e., for the case in which the parameter changes by leaps and takes only integral values.

A number of physicists Planck, Smoluchowski, Einstein, Fokker et al), biologists (Fisher) and some scientists working in the technological disciplines (Fry) were compelled to construct stochastic patterns all by themselves for answering various particular occasions. Keenly felt was the need for a common mathematical theory which

would allow a general interpretation of all the problems and patterns of the course of random processes.

Around 1900 Bachelier made the first such attempt. His investigations had however been unnoticed and in any case their mathematical level was low. Kolmogorov (1931) provided the first systematic account of the elements of the theory of stochastic processes without aftereffect. At about the same time Khinchin began to develop the theory of another most important class of stochastic processes, the so-called stationary processes.

Many mathematical results, wide possibilities for application in natural science as well as the transformation of the classical problems connected with the theory of stochastic processes, led to a new step in the development of the theory of probability which nowadays became the main field for applying the efforts of all mathematicians.

Processes without aftereffect. I will attempt to describe briefly this new chapter of the theory of probability and I begin by quoting Kolmogorov (1931) who had shown the similarity between the problems of the theory of stochastic processes and those of classical mechanics:

When desiring to treat mathematically phenomena of natural science or social life, it is first necessary to sketch them. Indeed, mathematical analysis can only be applied for studying the process of the change of some system when assuming that each of its possible states is completely ascertained by the known mathematical arsenal, for example, by values taken by a known number of parameters. Such a mathematically determinable system is not reality itself, but only a pattern suited for its description.

Classical mechanics only applies such patterns in which the state of y of a system at moment t is uniquely determined by its state x at any previous moment t_0:

$$y = f(x, t_0, t).$$

If such a simple-valued function exists, as it is always supposed in classical mechanics, we have a pattern of a completely determined process. It will also be possible to ascribe to them such processes in which the state y is not completely determined for the state x by a single moment t_0 but essentially depends in addition on the essence of the change of that x before moment t_0.

However, such a dependence on the previous behaviour of the system is usually avoided by widening the very notion of the state of a system at moment t and, accordingly, by introducing new parameters. In classical mechanics, apart from the coordinates of the position of a system, the components of their velocities are also usually considered.

Beyond the field of classical mechanics, it is usual to consider, along with the patterns of completely determined processes, patterns in which the state x of a system at some moment t_0 only conditions a known probability of a certain state y at some future moment $t > t_0$.

If for any given t_0, $t > t_0$ and x there exists a definite probability function for that state y, we say that this is a pattern of a

stochastically determined process. In general, this function is represented as

$P(t_0, x, t, E)$

where E is some set of states y and P is the probability that one of the states y belonging to that set is realized at moment t.

The reader will easily note that Kolmogorov considered processes which represented a further development of the pattern of Markov chains. It is important that he not only proposed to generalize the Markov idea from a finite on an arbitrary number of states and continuous time; he also established the general laws which govern those processes.

The probabilities *P* are obeying differential equations which he derived and which became named after him. If the set of states $(x_1, x_2, \ldots, x_n)$ in which a system can exist is finite, the states can vary continuously, and the functions

$P_{ij}(s, t) = P(s, x_i, t, x_j)$

are differentiable with respect to *s* and *t*, *s*   *t*, the process, as it occurred, obeys the following differential equations

$$\frac{\partial P_{ik}}{\partial t} = \sum_{j=1}^{n} A_{jk}(t) P_{ij}(s,t), \ i, \ k = 1, \ 2, \ \ldots, \ n,$$

$$\frac{\partial P_{ik}}{\partial s} = -\sum_{j=1}^{n} A_{ij}(s) P_{jk}(s,t), \ i, \ k = 1, \ 2, \ \ldots, \ n.$$

If the value of some real parameter *x* determines the system and

$$P(t_1, x, t_2, \ _y) = \int_{-\infty}^{y} f(t_1, x, t_2, y) dy$$

where $_y$ is the set of those states *z* for which $z < y$, this function satisfies the differential equations

$$\frac{\partial}{\partial s} f(s,x,t,y) = -A(s,x) \frac{\partial}{\partial x} f(s,x,t,y) - B^2(s,x) \frac{\partial^2}{\partial x^2} f(s,x,t,y),$$

$$\frac{\partial}{\partial t} f(s,x,t,y) = -\frac{\partial}{\partial y}[A(t,y)f(s,x,t,y)] + \frac{\partial^2}{\partial x^2}[B^2(t,y)f(s,x,t,y)].$$

The latter equation was derived by Fokker and Planck fifteen years before the appearance of Kolmogorov's study for the process of diffusion with a varying temperature and varying external forces fifteen years before the appearance of Kolmogorov's work. The coefficients *A* represent, so to say, the mean tendency of the process, and *B*, the intensity of the random fluctuations around that mean.

Kolmogorov's work became the source of many studies on the theory of random processes both at home and abroad; we will only mention some of them.

We now see the Fokker – Planck – Kolmogorov equations as the source of theorems of the Laplace – Liapunov type. This point of view explains why the classical normal law of probability is the solution of the equation of heat conductivity

$$\frac{\partial x}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}.$$

In the presence of random perturbations the theory of oscillations is very much in need of studying limiting regularities when the coefficients of the second derivations in the Kolmogorov equations are tending to zero. A number of results in this direction was due to A. A. Andronov, L. S. Pontriagin et al.

Among other applications we ought to mention the works of Kolmogorov and M. A. Leontovich on the Brownian motion, of Leontovich on the theory of bimolecular reactions, Kolmogorov on the theory of congestion in connection with the work of telephone circuits.

Petrovsky (1934) and Khinchin rigorously justified and further developed the mathematical theory of diffusion in the light of stochastic processes. A remarkable event was Khinchin's monograph (1933a). There, he considered a number of problems connected with random walks (diffusion) of a particle along a straight line or in a plane. In the simplest cases these problems are reduced to the well known patterns of series of random variables.

Bernstein (1934a, b) also studied the theory of stochastic processes. Issuing from the equations obeyed by the probabilities of the increments of random variables during a finite interval of time t he proved a number of most important results about their limiting behaviour as $t \to 0$. He deeply analysed the Kolmogorov equations aiming to establish the conditions under which their solutions really satisfy the requirements of the theory of probability.

Feller and V. M. Dubrovsky who both followed the Kolmogorov ideas have been developing the theory of totally discontinuous processes, i. e. such, whose changes occur not continuously but in separate randomly scattered moments of time. I am not dwelling on the numerous examples (radioactive decay, capitals of insurance offices) which were the objects for the application of their theory.

4.5. Stationary processes. Processes without aftereffect do not by any means exhaust all the requirements of natural science on mathematics. Indeed, in many phenomena the previous states of the system very strongly influence the probability of its future states and it is impossible to disregard that influence even in approximate considerations. In many cases the situation can be improved by changing the notion of state, i. e., by introducing new parameters. Thus, when dealing with the changes in the position of a particle in diffusion or Brownian motion as with a process of the Kolmogorov type, it would have meant that we do not allow for its inertia.

Inclusion of the velocity of the particle in addition to its coordinates improves the situation.

And still, there are numerous examples in which that method does not help however many new parameters are included. In the first place this is true for statistical mechanics: the position of the particle only provides a stochastic judgement about its future position. An acquaintance with the previous positions of a point essentially changes our opinion about its future.

Khinchin (1934) selected an important class of processes with aftereffect, the so-called stationary processes homogeneously behaving in time. We say that a process $x(t)$ is stationary if the probability distribution for two finite groups of variables

$$\{x(t_1), x(t_2), \ldots, x(t_n)\} \text{ and } \{x(t_1 + \ ), x(t_2 + \ ), \ldots, x(t_n + \ )\}$$

coincide (and therefore do not depend on). Numbers n and   as well as the moments $t_1, t_2, \ldots, t_n$ can be chosen quite arbitrarily.

Understandably, it is possible to indicate any number of such important stationary processes in various branches of knowledge. We immediately indicate that the deepest cognition of many acoustic (noise, for example) and light phenomena as well as the discovery of latent periodicities so interesting for geophysicists and meteorologists, is only possible in the bounds of stationary processes,

A sharp eye will easily discern phenomena representing stationary processes in any established technological treatment. Consider, for example, spinning. Heterogeneity of the material (of the length of the fibre, of its strength, cross-section), fluctuations in the velocity and homogeneity of its occurrence in the machine at differing moments and of many other parameters, lead to changes in the properties of the yarn. Knowledge of some property of the yarn in one or another part of the skein does not mean its complete knowledge in another part of the skein. However, spinning can be considered established, unchanging in time, so that the stochastic properties of any part of the yarn may be assumed constant and thus represent a stochastic process.

The importance of such investigations was noted even before Khinchin's study. Separate results were due to Slutsky (1927), Romanovsky (1932; 1933) and others, and in particular to some geophysicists. Khinchin, however, provided a general definition of a stationary process and proved its most important properties.

At first, he proved the LLN for the quadratic convergence of means, constructed the theory of correlation of a stationary process and finally proved that the Birkhoff theorem can be generalized on that class of processes. Birkhoff himself (Khinchin 1932) only proved his theorem for dynamic systems and only under additional restrictions.

The Birkhoff – Khinchin theorem. If (t) is a stationary random process with a finite E| (t)| then the limit

$$\lim \frac{1}{T} \int_0^T (t)dt = \text{E} \ (1), \ T \to \infty$$

exists with probability 1.

Its proof and the indication of its importance for statistical mechanics and its connection with the theory of dynamic systems can be found in Khinchin (1943). That the Birkhoff ergodic theorem was included in the bounds of stationary processes as a particular case of their regularities and that in this new aspect Khinchin provided its logically transparent and extremely general proof only confirms the fundamental importance of that [?] theory.

Khinchin (1934) definitively described some points of that theory which only depended on the second moments. Stationarity was here understood in a generalized form: The process $\xi(t)$ is called stationary if for any values of t and τ the magnitudes $E[\xi(t)]$ and $E[\xi(t)\,\xi(t+\tau)]$ do not depend on the parameter t. The central result of that book was the theorem about the spectral representation of the coefficients of correlation[1] for stationary (in the sense just mentioned) processes. Denote the coefficient of correlation between $\xi(t)$ and $\xi(t+\tau)$ by $r(\tau)$. Then

The Khinchin theorem. If the correlation coefficient $r(\tau)$ satisfies the condition $r(+0) = 1$, then there exists a spectral representation

$$r(\tau) = \int_{0}^{\infty} \cos \tau z\, dF(z) \qquad (4.5.1)$$

where $F(z)$ is a non-decreasing function with a unit variation.

The inverse proposition: for any function $r(\tau)$ representable as an integral (4.5.1) there exists a stationary process for which it, $r(\tau)$, is the correlation function.

The spectral representation of the correlation coefficient became the point of departure for all the later studies of the theory of stationary processes. Khinchin himself applied it for deriving both the LLN and a number of results pertaining to statistical physics.

Slutsky, Cramér et al further developed this Khinchin direction of studies which also served as the beginning of the creation of the theory of homogeneous fields of probability (Kolmogorov, A. M. Obukhov, M. G. Krein, Wiener, Schönberg et al). Such fields are functions $A(M)$ of points of an n-dimensional Euclidean space for which the probability distribution of $A(M_1), A(M_2), \dots, A(M_k)$, with arbitrary points $M_1, M_2, \dots, M_k, k = 1, 2, \dots$ remains constant under a parallel shift of those points as a solid system.

Slutsky (1932) deeply analysed the structure of a stationary process as a random function of argument t. Especially harmonious was his particular result: if $F(z)$ in the spectral expansion of the correlation coefficient has a discrete spectrum, then with probability 1 $\xi(t)$ is an almost periodic function in the Bezikovich sense (which means: is almost certainly almost periodic for each realization of $\xi(t)$).

Along with the development of the theory of stationary processes Khinchin (1933b) investigated stationary sequences

$$\{\xi(t)\}, t = 0, \pm 1, \pm 2, \dots \qquad (4.5.2)$$

and received results similar to those for the processes. Wold, Kolmogorov, V. N. Zasukhin continued Khinchin's work.

Kolmogorov (1941) developed the spectral theory of stationary sequences by the arsenal of the spectral theory of operators in the Hilbert space. He was interested in the extrapolation and interpolation of stationary sequences (4.5.2) of random variables. For the sake of simplicity we assume that

$$E\,\xi(t) = 0,\ E\,\xi(t)\,\xi(t+k) = B(k),\ B(0) = 1.$$

Extrapolation means the selection of such constants a_k that, for a given n,

$$L_n = \sum_{k=1}^{n} a_k\,\xi(t-k)$$

as precisely as possible approaches $\xi(t+m)$. The measure of precision is here the mean value of the square of the deviation of L_n from $\xi(t+m)$, i. e. the magnitude

$$\sigma^2(n,\,m) = E[\,\xi(t+m) - L_n]^2.$$

This magnitude obviously does not increase with n and tends to a certain limit. If this limit is 0, an unrestricted extrapolation of the process takes place for a given m.

A spectral representation also takes place for stationary sequences:

$$B(k) = \int_{o}^{\pi} \cos kz\, dF(z)\,.$$

Here $F(z)$ is a non-decreasing function of a unit variation over $(0,\ \pi)$. The derivative $F'(z)$ exists almost everywhere, is non-negative and summable. It occurs that an unrestricted extrapolation takes place then and only then when (for all $m\ \ge 0$ at once) the integral

$$P = \int_{o}^{\pi} \ln F'(z)\, dz$$

diverges. It follows in particular that if $F'(z) = 0$ on the set of a positive measure than an unrestricted extrapolation takes place.

4.6. Influence on classical problems. The theory of random processes essentially widened the science of chance by creating its new chapter and it also illustrated anew the CLT for sums of random variables. It occurred that the main laws of distribution, which had previously been asymptotic, now, in the theory of stochastic processes, became precise solutions of differential equations.

Kolmogorov began the pertinent investigations, and Bernstein, Khinchin et al widely developed them. The CLT of the theory of probability is now perceived as a particular case of a united general theory. Investigations of the classical pattern of the sequences of

random variables experienced an essential stimulus from the theory of stochastic processes.

Kolmogorov's study of homogeneous random processes with independent increments was the first in this direction. He established that all such processes were governed by the so-called infinitely divisible laws[1] and found their analytic representation. Previously, the investigators' interest had been concentrated on the determination of the widest conditions under which the distribution functions of sums of independent terms tended to the normal law, but a number of new problems have since occurred, and the naturalness and importance of their formulation is doubtless. At first, all the probability distributions which can be the limit laws for sums of independent random variables should have been discovered. In other words, given a sequence of random variables, each of them being a sum of independent terms, with the distributions of whose sums tending to a limit distribution law, it was required to investigate that law.

This problem is formulated too generally: any probability distribution can be the limit law in the stated sense. However, always imposed in the theory of probability is a restriction which reflects many problems of statistics and natural science: it required that the separate terms of the sum [of the sums] little influenced it [them]. And now the limit distribution law is not anymore arbitrary. Kolmogorov stated as a hypothesis that the class of the limit laws understood in the above sense coincided with the class of infinitely divisible laws.

His student Bavli (1936), who perished in 1941 in an air raid on Moscow, proved that hypothesis if the variances of the sums, $E(S_n - ES_n)^2$, are restricted by a constant independent from n. Next year Khinchin (1937) provided a complete proof of that hypothesis without any additional restrictions (even without requiring the existence of finite variances of those sums).

At about the same time three authors, Khinchin (1935 – 1936), Lévy and Feller, derived the necessary and sufficient conditions of the CLT, and Khinchin (1938) described the fruitful cycle of investigations of the Moscow school of probability. Further progress in this area was largely based on this monograph.

From those later investigations we ought to mention the work of Khinchin's student Bobrov (1937) on the LLN for positive terms. It allowed Raikov (1938b) to note a curious fact, the connection between the conditions of the LLN and CLT: the sum of independent random variables then and only then tends to the normal law when the sum of their squares obeys the LLN.

After these studies there naturally occurred the problem about the conditions for the existence of a limit law for such sums and for their convergence to each given limit law. Gnedenko (1939) completely solved this problem. And his general method of proving limit theorems for sums of independent random variables allowed him uniformly and without lengthy calculations to describe all the facts collected in this area including those relating to the LLN and CLT, and to derive a number of findings.

Independently and at the same time the Austrian mathematician Doeblin obtained many results in the same direction. After the

German occupation of Austria he emigrated to France but apparently perished during the war.

We see that the theory of limit laws had recently acquired an essentially common character and that the most important problems of the classical theory of probability had been included there as simplest particular cases. That very general viewpoint allowed us to ascertain completely distinctly the fundamental importance of the Gauss law. Exactly this property compelled all the investigators to study it for almost two centuries. It occurred now that the conditions for the convergence to the Gauss law were absolutely general whereas the convergence to other laws required very special conditions.

4.7. Arithmetic of the laws of distribution. Lévy's peculiar findings about the convolution of the distribution functions when random variables are added up found a response in Moscow. The distribution function $F(x)$ of a sum of independent random variables ξ_1 and ξ_2 which obey distributions $F_1(x)$ and $F_2(x)$ is known to be

$$F(x) = \int\limits_{-\infty}^{\infty} F_1(x-z)dF_2(z) = F_1(x) \cdot F_2(x).$$

We will call this operation the multiplication of $F_1(x)$ and $F_2(x)$. Here, such notions as factorization of a law of distribution or a simple law (impossible to factorize) acquire a meaning.

Lévy showed that factorization is, generally speaking, ambiguous and Cramér proved that the Gauss law can always be factorized in multipliers which also obey the normal law. Moscow mathematicians, however, added a number of facts. Gnedenko & Khinchin proved that the division of the distribution functions is not one-valued. Thus, Gnedenko (1937b) showed that there exist such distribution functions $F_1(x)$, $F_2(x)$ and $F_3(x)$ that, in spite of $F_2(x) \ne F_3(x)$,

$$F_1(x)\, F_2(x) = F_1(x) \cdot F_3(x).$$

Then, Raikov (1938a) discovered a class of laws lacking simple multiples. Along with other findings, he proved a theorem similar to the Cramér's proposition about the Poisson law[3].

The Raikov theorem. If the Poisson law is factorized into multipliers $F_1(x)$ and $F_2(x)$, then each of these multipliers is also a Poisson law.

Khinchin (1937) proved that any law of distribution can be represented as the product of two laws, one of them lacking simple multipliers, and the other one, either of the type

$$F(x) = 0 \text{ for } x \le a \text{ and } F(x) = 1 \text{ for } x > a,$$

or a product of finite or countable sets of non-factorable laws.

Arithmetic of the laws of distribution is yet in a rudimentary form. Suffice it to note that there are yet no answers to simplest and natural questions. Thus, the conditions for the law of distribution to be a power of some other law (for example, its square) are unknown.

4.8. Investigations in mathematical statistics. In Russia, they have not acquired the scope deserved by that discipline. Soviet scientists have not yet [!] occupied leading positions and their contribution to the development of mathematical statistics mostly concerned the discovery of separate facts rather than the creation of new concepts. Many of those facts, however, undoubtedly belong to the best achievements of science and will occupy an honourable place in future courses in mathematical statistics.

Many of those stochastic studies are fundamentally important for statistics, as for example those which bear on the LLN or the CLT. However, they do not comprise the main body of mathematical statistics and do not therefore change the opinion about its state in the USSR as formulated above.

Not many are developing the main problems of statistics, although a rather large number of investigators of various concrete applications of statistical methods had obtained valuable results. Here also, as was the case of the achievements of Russian scientists in probability theory, we restrict our account to some main results and, moreover, only to those which were made very recently. The names of many scholars who promoted the dissemination and perfection of the statistical methods are lacking. Neither do we touch on works concerning the development of concrete statistics in agriculture, demography, finance, industry.

We begin by discussing the remarkable cycle of studies initiated by V. I. Glivenko and Kolmogorov and widely developed by N. V. Smirnov. They concerned the solution of the main problem of statistics, the determination of an unknown distribution function given the results of observations, and they also pertained to the study of the approach of the empirical to the theoretical distribution function.

Let some random variable have $F(x) = P(< x)$ as its distribution function and $_1$, $_2$, …, $_n$, be the independent results of observing . The empirical distribution function is then

$$F_n(x) = \frac{k(x)}{n}.$$

Here, $k(n)$ is the number of the observed values of smaller then x. The first general discovery in this direction (Glivenko 1933) was that certainly, for random variable with a continuous distribution function, $F_n(x)$ $F(x)$ as n .

Kolmogorov (1933b) discovered another general fact. If $F(x)$ is continuous, the distribution functions of

$$D_n = \max |F_n(x) - F(x)|/ \ n, |x| < + \quad,$$

converge, as n , to some distribution function (x) which is independent from $F(x)$:

$$\lim P(D_n <) = \quad ( ) = \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2 \ ^2).$$

This theorem can be applied as a test of the concordance between the two distributions. Indeed. Suppose we experimentally established that if the random variable had distribution function $F(x)$, $D_n =$ and that the probability ( ), i. e. the probability of the inequality $D_n <$, was high, then the probability of D_n is low which means that an unlikely event took place. Considering, however, that such events are practically impossible, we ought to believe that the occurred deviation of D_n was not random and that our assumption should be questioned. An essential advantage of this method of estimating concordance consists in the independence of ( ) from the type of the function $F(x)$ (which is unknown). For applying this test tables of the function ( ) was constructed under the guidance of N. V. Smirnov.

Smirnov's later investigations made it clear that the distribution ( ) discovered by Kolmogorov plays the main role in many statistical problems. Here is one among many of those problems which he solved. Let

$x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_m$

be the independent observations on random variables and . For statistics, it is essential to establish rules for judging whether these variables have the same distribution or not. Indeed, suppose that two series of trials had been made to determine the influence of some agricultural measure on the yield, and only one of these series corresponded to the lack of this measure. The deviations between observations can be purely random which means that that measure did not represent any progress. So how to find out whether those deviations were not random? I will not multiply such examples since readers can easily provide them themselves.

Suppose that $F_n(x)$ and $F_m(x)$ are the empirical distribution functions for those series of observations. Smirnov proposed the magnitude

$$D(n, m) = \max |F_n(x) - F_m(x)| \sqrt{\frac{nm}{n+m}}, \ |x| < +$$

as the measure of the deviation. If this magnitude exceeds some boundary the deviation will be considered essential and the identity of the laws of distribution questioned.

The following proposition (Smirnov 1939) completely solved the problem. If n and m increase unboundedly but their ratio $= n/m$ remains constant, then, as n ,

$P[D(n, m) < \ ] \qquad ( ), \ > 0.$

The three theorems considered here sufficiently describe that new direction in statistics which was initiated and is developing in Moscow.

Another important cycle of studies is due to Slutsky and is devoted to the study of cyclic processes. Numerous phenomena of nature, economics and technology are going on as though periodically.

Maximal and minimal values alternate rather regularly, but neither the length of the waves, nor the ordinates are precisely repeated. Many scientists studied such cases by assuming that the irregularities were due to random phenomena which corrupted the regular oscillations. It occurred, however, that this viewpoint was often untenable.

Slutsky, who issued from geophysical and economic problems, essentially promoted the study of such processes. He established the main fact: such pseudo-periodic occurrences can result from the action of random causes rather than being based on periodicity.

Let $_1$, $_2$, …, $_n$, … be a sequence of mutually independent random variables with the same law of distribution, and $a_1, a_2, \ldots, a_n, \ldots$, some constants. Consider a stationary sequence of random variables

$$_1, \ _2, \ldots, \ _n, \ldots,$$

$$_n = a_1 \ _n + a_2 \ _{n+1} + \ldots + a_{k+1} \ _{n+k}.$$

Slutsky proposed to call the formation of sequences of connected random variables a mobile summation. It turned out that such sequences can imitate periodic processes. Moreover, Slutsky (1927a; 1927b) showed that with probability as near to unity as desired and during an arbitrarily long period the terms of such sequences under some conditions do not deviate from the corresponding ordinates of a sine curve more than by .

Later Romanovsky generalized this finding which evidently touches on the problem of determining periodicity. While developing the ideas of his investigation Slutsky studied a number of examples (periodicity of sun spots, the Beveridge wheat price index etc.) and showed that by constructing models of series of the considered type it was possible to question very reliably the hypothesis of periodicity or, in some cases, confirm it.

In 1922, serious work on the theory of probability and mathematical statistics had begun in Tashkent, at first by Romanovsky alone, then by him and his students, among whom we mention T. A. Sarymsakov.

When investigating mathematical statistics, Romanovsky had for some time worked under a certain influence of the Pearson school. However, when selecting methods of work, he followed Chebyshev. Being Markov's student, Romanovsky adopted from him the traditions of the Chebyshev school, and among them a mathematical rigour of considerations and a logical scrupulousness of constructions. This, indeed, was lacking in the work of the English statisticians.

For almost twenty years of work Romanovsky's investigations covered literally all the parts of mathematical statistics (curves of distribution, theory of sampling, distribution of statistical measures, tests of randomness, disclosure of latent periodicities etc.). His studies of the distribution of the coefficients of correlation and regression for samples from normal populations are classical. At the same time Romanovsky actively propagandized statistical methods. He wrote a number of books and thus essentially assisted the upsurge of statistical culture and interest in statistics. Among his books we especially note his elementary course (1921) and the fundamental treatise (1938).

Notes by the author

1. To recall, the coefficient of correlation between magnitudes   and   is

$$r\ \ = \frac{\mathrm{E}\ \ }{\sqrt{\mathrm{var}\ \ \mathrm{var}\ \ }}.$$

2. The law of distribution is called infinitely divisible if a random variable obeying that law can be represented as a sum of *n* independent terms with an arbitrary *n*, all of them distributed according to the same law.

3. A random variable   obeys the Poisson law if its values are of the type $k + $ where   and   are real numbers, *k* takes integral positive values and

$$P(\ \ =\ \ k +\ \ ) = \frac{a^k e^{-a}}{k\,!}$$

with a constant $a > 0$.

Bibliography by the Author

Abbreviation: IMI = *Istoriko-Matematich. Issledovania*
Sel. Works = *Sel. Works*, vol. 2. Dordrecht, 1991

Bavli G. M. (1936), Über einige Verallgemeinerungen der Grenzwertsätze der Wahrscheinlichkeitsrechnung. *Matematich. Sbornik*, vol. 1, pp. 917 – 929.

Bernstein S. N. (1917, Russian), Essay on an axiomatic justification of the theory of probability. *Sobranie Sochineniy*, vol. 4. No place, 1964, pp. 10 – 60. **S, G,** 6.

--- (1926), Sur l'extension du théorème limite de calcul des probabilités aux sommes de quantités dépendantes. *Math. Ann.*, Bd. 97, pp. 1 – 59.

--- (1927), *Teoria Veroiatnostei* (Theory of Probability). Fourth edition, 1945.

--- (1928, Russian), The present state of the theory of probability and its applications. *Sobranie Sochineniy*, vol. 4. No place, 1964, pp. 217 – 232.

--- (1934a), Principes de la théorie des équations différentielles stochastiques. *Trudy V. A. Steklov Fiz.-Mat. Inst.*, vol. 5, pp. 95 – 124; *Sobranie Sochineniy*, vol. 4. No place, 1964, pp. 291 – 315.

--- (1934b), Equations différentielles stochastiques. *Actualités sientifiques*.

--- (1937, Russian), On some modifications of the Chebyshev inequalities. *Doklady Akad. Nauk SSSR*, vol. 17, No. 6, pp 275 – 277. *Sobranie Sochineniy*, vol. 4. No place, 1964, pp. 331 – 333.

--- (1945, Russian), Chebyshev's work in the theory of probability. *Sobranie Sochineniy*, vol. 4. No place, 1964, pp. 409 – 433. **S, G,** 5.

Bobrov A. A. (1938, Russian), On the relative stability of sums of positive random variables. *Uchenye Zapiski Mosk. Gos. Univ.*, No. 15, pp. 191 – 202.

Buniakovsky V. Ya. (1846), *Osnovaniya Matematicheskoi Teorii Veroiatnostei* (Principles of the Math. Theory of Probability). Petersburg.

Chebyshev P. L. (1845, Russian), Essay on an elementary analysis of the theory of probability. *Polnoe Sobranie Sochineniy*, vol. 5, 1951, pp. 26 – 87.

--- (1846), Démonstration élémentaire d'une proposition générale de la théorie des probabilités. *J. reine angew. Math.*, Bd. 33, pp. 259 – 267.

--- (1867), Des valeurs moyennes. *J. math. pures et appl*., sér. 2, t. 12, pp. 177 – 184.

--- (1887), Sur deux théorèmes relatifs aux probabilités. *Acta Math.*, t. 14, 1890 – 1891, pp. 305 – 315.

Cramér H. (1937), *Random Variables and Probability Distributions*. Cambridge.

Feller W. (1935 – 1937), Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung. *Math. Z.*, Bd. 40, pp. 521 – 559; Bd. 42, pp. 301 – 312.

Glivenko V. I. (1933), Sulla determinazione empirica di una legge di distribuzione. *Giorn. dell'Ist. Ital. degli Attuari*, t. 4, pp. 1 – 10.

Gnedenko B. V. (1937a, Russian), On a characteristic property of infinitely divisible laws of distribution. *Byull. Mosk. Gos. Univ*., A, vol. 1, No. 5, pp. 10 – 16.

--- (1937b, Russian), On characteristic functions. Ibidem, pp. 17 – 18.

--- (1939, Russian), On the theory of limit theorems for sums of independent random variables. *Izvestia Akad. Nauk SSSR*, ser. math., No. 2, pp. 181 – 232 + 643 – 647. *Uspekhi Matematich. Nauk*, vol. 10.

Khinchin A. Ya. (1923), Über diadische Brüche. *Math. Z.*, Bd. 18.

--- (1924), Über ein Satz der Wahrscheinlichkeitsrechnung. *Fundamenta Math.*, Bd. 6, pp. 9 – 20.

--- (1932), Zu Birkhoffs Lösung des Ergodenproblems. *Math. Ann.*, Bd. 107, pp. 485 – 488.

--- (1933a), *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*. Berlin.

--- (1933b), Über stationäre Reihen zufälliger Variablen. *Matematich. Sbornik*, vol. 40, No. 2, pp. 124 – 128.

--- (1934), Korrelationstheorie der stationären stochastischen Prozesses. *Math. Ann.*, Bd. 109, pp. 604 – 615.

--- (1935 – 1936), Sul domino di attrazione della legge di Gauss, *Giornale dell'Istituto Ital. degli attuari*, t. 6, pp. 378 – 393; t. 7, pp. 3 – 18.

--- (1937), Contribution à l'arithmétiqe des lois de distribution. *Byull. Mosk. Gos. Univ.*, math.-mech. ser., vol. 1, No. 1, pp. 6 – 17.

--- (1938), *Predelnye Zakony dlya Summy Nezavisimych Sluchainych Velichin* (Limit Laws for Sums of Independent Random Variables). Moscow – Leningrad.

--- (1943, Russian), *Mathematical Foundations of Statistical Mechanics*. New York, 1949.

Khinchin A. Ya., Kolmogorov A. N. (1925, German), On convergence of series whose terms are determined by random events. In Kolmogorov A. N., *Sel. Works*, pp. 2 – 10.

Klein F. (1926 – 1927), *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*. Berlin.

Kolmogorov A. N. (1928), Über die Summen durch den Zufall bestimmter unabhängiger Größen. *Math. Ann.*, Bd. 99, pp. 309 – 319.

--- (1929, German), On the law of the iterated logarithm. *Sel. Works*, pp. 32 – 42.

--- (1931, German), On analytic methods in probability theory. *Sel. Works*, pp. 62 – 108.

--- (1933a, German), *Foundations of the Theory of Probability*. Martino Fine Books, 2013.

--- (1933b, Italian), On the empirical determination of a distribution law. *Sel. Works*, pp. 139 – 146.

--- (1936), *Osnovnye Poniatia Teorii Veroyatnostei* (The Main Concepts of the Theory of Probability). Moscow – Leningrad.

--- (1937, Russian), Markov chains with countable sets of possible states. *Sel. Works,* pp. 193 – 208.

--- (1941, Russian), Interpolation and extrapolation of stationary random sequences. *Sel. Works*, pp. 272 – 280.

--- (1947, Russian), The role of Russian science in the development of the theory of probability. *Uchenye Zapiski Mosk. Gos. Univ.* No. 91, pp. 53 – 64. **S, G,** 7.

Liapunov A. M. (1900), Sur une proposition de la théorie des probabilités. *Izvestiya Imp. Akad. Nauk St. Pétersb.*, sér. 5, t. 13, pp. 359 – 386.

--- (1901), Nouvelle forme de théorème sur la limite des probabilités. *Mém Imp. Akad. Nauk St. Pétersb.*, sér. 8, Cl. phys. math., t. 12, No. 5, separate paging.

Lobachevsky N. I. (1835 – 1838, Russian), *New Principles of Geometry with Complete Theory of Parallels*. Austin, Texas, 1897.

--- (1842), Sur la probabilité des résultats moyens tires des observations répétées. *J. reine angew. Math.*, Bd. 24, pp. 164 – 170.

Markov A. A. (1898), Sur les racines de l'équation […]. *Izvestiya Petersb. Akad. Nauk*, ser. 5, vol. 9, pp. 435 – 446.

--- (1899, Russian), The law of large numbers and the method of least squares. *Izbrannye Trudy.* No place, 1964, pp. 231 – 251. **S, G,** 5.

--- (1900), *Ischislenie Veroyatnostei* (Calculus of Probability). Later editions: 1908, 1913 and, posthumously, Moscow, 1924. German translation: 1912.

--- (1906, Russian), Extension of the law of large numbers on mutually dependent magnitudes. *Izbrannye Trudy.*, pp. 339 – 361. **S, G,** 5.

Ostrogradsky M. V. (1836), Sur le calcul des fonctions génératices. *Bull. Scient. Akad. Imp. Sci. St.-Pétersburg*, t. 1, No. 10, pp. 73 – 75.

--- (1848), Sur une question des probabilités. *Mém. Acad. Imp. St. Pétersburg*, 6me sér., sér. math., phys. et natur., t. 66. No. 21 – 22, pp. 321 – 346.

Petrovsky I. G. (1934), Über das Irrfahrtproblem. *Math. Ann.*, Bd. 109, pp. 425 – 444.

Raikov D. A. (1938a, Russian), On the decomposition of the Gauss and Poisson laws. *Izvestia Akad. Nauk SSSR*, ser. math., No. 1, pp. 91 – 124.

--- (1938b, Russian), On the connection between the law of large numbers and the central limit theorem. *Izvestia Akad. Nauk SSSR*, ser. math., No. 3, pp. 328 – 338.

Romanovsky V. I. (1921), *Elementarnyi Kurs Matematicheskoi Statistiki* . Tashkent. Lithographic edition. Later editions: 1924 and Moscow – Leningrad, 1939.

--- (1930, 1932a), Sur une classe d'équations lineaires. *C. r. Acad. Sci. Paris*, t. 191, pp. 552 – 557. Also *Acta math.*, vol. 59, 1932, pp. 99 – 208.

--- (1932b), Sur la loi sinusoidale limite. *Rendi conti Circ. Math. Palermo*, t. 56, pp. 82 – 111.

--- (1933), Sur une généralisation de la sinusoidale limite. Ibidem, t. 57, pp. 130 – 136.

--- (1935), Recherches sur les chaines de Markoff. *Acta math.*, t. 66, pp. 147 – 251.

--- (1938), *Matematicheskaia Statistika*. Moscow – Leningrad.

Sarymsakov T. A. (1942, Russian), Theory of probability and math. statistics in the works of the Tashkent school. In *25 Let Sovetskoi Nauki v Uzbekistane*. Tashkent, pp. 64 – 72.

Slutsky E. E. (1925), Über stochastische Asymptoten und Grenzwerte. *Metron*, t. 5, No. 3, pp. 3 – 89.

--- (1927a, Russian), Summation of random causes as the source of cyclic processes. *Econometrica*, vol. 5, 1937, pp. 105 – 146.

--- (1927b), Sur un théorème limite relative aux séries des quantités éventuelles. *C. r. Acad. Sci. Paris*, t. 185, pp. 169 – 171.

--- (1932), Sulla successioni stazionari di eventi. *Giorn. dell'Ist. Ital. degli Attuari*, t. 3.

Smirnov N. V. (1939, Russian), On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Byull. Mosk. Gos. Univ.*, math.-mech. ser., vol. 2, pp. 3 – 16.

I am adding

Bernstein S. N. (1922, Russian), Mathematical problems of modern biology. *Nauka na Ukraine*, vol. 1, pp. 13 – 20. **S, G,** 6.

Gnedenko B. V. (1939, Russian), Review of the current state of the theory of limit laws for sums of independent terms. *Uchenye Zapiski Tomsk. Pedagogich. Inst.*, No. 1, pp. 5 – 28. **S, G,** 65. Printed most horribly.

--- (1951, Russian), On Ostrogradsky's work in the theory of probability. IMI, vol. 4, pp. 99 – 123. **S, G,** 5.

Chebyshev's works were almost completely reprinted in his *Oeuvres*, tt. 1 – 2. Pétersbourg, 1899 –1907. Russian – French edition. French edition reprinted: New York, 1962.

Russian works of the appropriate periods are listed in bibliographies included in
Matematika v SSSR za Tridzat Let (Math. in the USSR over 30 Years). Moscow – Leningrad, 1948.
Matematika v SSSR za Sorok Let (Math. in the USSR over 40 Years). Moscow, 1959.

Bibliography by the translator
see also the *Bibliography by the Author*
Abbreviations: AHES = *Arch. Hist. Ex. Sci.*

Bertrand J. (1888), *Calcul des probabilités*. Later editions: Paris, 1907; New York, 1970, 1972.

Bru B. (1993), Doeblin's life and work from his correspondence. *Contemporary Math.*, vol. 149, pp. 1 – 64.

Bru B., Yor M. (2002), Comments on the life and mathematical legacy of W. Doeblin. *Finance Stoch*. Vol. 6, pp. 3 – 47.

Chetverikov N. S. (1959, Russian), The life and scientific work of Slutsky. *Statisticheskie Issledovaniya*. Moscow, 1975, pp. 261 – 281. **S, G,** 40.

Cramér H. (1976), Half a century with probability theory. *Annals of Probability*, vol. 4, pp. 509 – 546.

Czuber E. (1903), *Wahrscheinlichkeitsrechnung und ihre Anwendung*. Several later editions.

Doob J. L. (1989), Commentary on probability. In *Centenary of Math. in America*, pt. 2. Providence, Rhode Island, 1989, pp. 353 – 354. Editor P. Duren et al.

--- (1994), The development of rigour in mathematical probability (1900 – 1950). In J.-P. Pier, Editor, *Development of Mathematics, 1900 – 1950*. Basel, pp. 157 – 170. Reprint: *Amer. Math. Monthly*, vol. 103, 1996, pp. 586 – 595.

Gnedenko B. V. (1949, in Russian), On the work of Lobachevsky in the theory of probability. IMI, vol. 2, pp. 129 – 136.

Gnedenko B. V., Sheynin O. (1978, Russian), Theory of probability. A chapter in *Math. in the 19th Century*. Editors, A. N. Kolmogorov, A. P. Youshkevich. Basel, 1992, 2001, pp. 211 – 282 + 283 – 288.

Kagan V. F. (1944, Russian), *Lobachevsky*. Moscow – Leningrad.

Kamke E. (1933), Über neuere Begründungen der Wahrscheinlichkeitsrechnung. *Jahresber. Deutsche Mathematiker-Vereinigung*, Bd. 42, pp. 14 – 27.

Lévy P. (1925), *Calcul des probabilités*. Paris.

Liapunov A. M. (1895, Russian), P. L. Chebyshev. In Chebyshev P. L. (1946). Moscow – Leningrad, pp. 9 – 21.

Mill J. S. (1843), *System of Logic*. London, 1886.

Novikov S. P. (2002, Russian), The second half of the 20th century and its end result. IMI, vol. 7 (42), pp. 326 – 356.

Ondar Kh. O., Editor (1977, Russian), *Correspondence between A. A. Markov and A. A. Chuprov*. New York, 1981.

Poisson S.-D. (1837, 2003, French), *Researches into the Probabilities of Judgements* etc. Berlin, 2013. **S, G,** 53.

Sheynin O. (1973), Finite random sums. AHES, vol. 9, pp. 275 – 305.

--- (1989), Markov's work on probability. AHES, vol. 39, pp. 337 – 377; vol. 40, p. 387.

--- (1991), The work of Buniakovsky in the theory of probability. AHES, vol. 43, pp. 199 – 223.

--- (1994), Chebyshev's lectures on the theory of probability. AHES, vol. 46, pp. 321 – 340.

--- (1995), Density curves in the theory of errors. AHES, vol. 49, pp. 73 – 104.

--- (1998), Statistics in the Soviet epoch. *Jahrbücher f. Nationalökonomie u. Statistik*, Bd. 217, pp. 529 – 549.

--- (2006), Markov's work on the treatment of observations. *Hist. Scientiarum*, vol. 16, pp. 80 – 95.

--- (2008), Romanovsky's correspondence with K. Pearson and R. A. Fisher. *Archives intern. d'histoire des sciences*, t. 58, No. 160 – 161, pp. 365 – 384.

--- (2009), *History of Probability. Historical Essay*. Berlin. **S, G,** 10.

--- (2010), The inverse law of large numbers. *Math. Scientist*, vol. 35, pp. 132 – 133.

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

Youshkevich A. A. (1974), Markov. *Dict. Scient. Biogr.*, vol. 9, pp. 124 – 130.

Oscar Sheynin

A. A. Markov and B. M. Koialovich

Boris Mikhailovich Koialovich (1867 – 1941) was Markov's student. Mikhelson (1973), whose paper I have not read, and Ermolaeva (2009) described his life and work. The few lines about him below mostly follow her.

As a student, Koialovich was a member of a study group headed by Markov and continued his activity there even after graduating (from the Petersburg University). In 1890, he became a member of the Petersburg Mathematical Society as recommended by Markov and about 1892 – 1893 read there thirteen reports. In 1892 Koialovich became lecturer at the Petersburg Technological Institute (professor since 1893). He read there an elective course in the theory of probability which included the method of least squares (MLSq). In 1928, he became Honoured Scientific Worker.

In 1913, in a letter to Steklov, Markov (Steklov 1991, p. 220) called Koialovich *my talented student*. See also Dobrovolsky (1967, p. 415).

At some time between 1912 – 1915 the empress had invited him to a dinner but he did not come so as to avoid an encounter with Rasputin (Ermolaeva 2009, p. 106, Note 32). He could have been assigned warden of the Petersburg educational district, but Rasputin was unable to form an opinion about him, and Koialovich did not get this high position.

At that time, and since 1903, he was member of the Scientific Council of the Ministry of People's Education and probably rubbed shoulders with Nekrasov, a member of the Council of the Minister himself, but Ermolaeva did not mention Nekrasov. In 1910, Nekrasov asked to be admitted to the Scientific Council as well but was turned down, see Sheynin (2007a, Letter 67b in § 5.4).

I (2004, No. 16) had published, in translation, the extant archival letters from the correspondence between Markov and Koialovich, and Ermolaeva (2009) published a letter from that correspondence previously kept by the descendants of Koialovich. Below, after translating the bulk of her general comments, I translate that letter complete with some of her particular comments, and I had to reprint two of Koialovich' letters to Markov. At least one of Markov's letters written between 25 Sept. and 2 Oct. 1893 is missing.

Explanatory text to Markov's letter of 23 Sept. 1893
Ermolaeva (2009)

Markov's severe criticism of Koialovich was hardly pleasant but their contacts had continued: Grodzensky (1987, p. 126) published a photo of their chess battle and Markov positively referred to Koialovsky's doctoral dissertation[1]. An exchange of opinions also apparently took place.

I think that Koialovsky [partly] took in that criticism. Had he rewritten his course of lectures? We only know that its second part was not published, perhaps not even written, although he possibly had been busy preparing his master dissertation (defended in 1894). Anyway, he continued to lecture until 1896.

Koialovich was able to take some beating and just like Markov he participated in scientific debates. His first minor encounter began in 1895 when he read a report at the Petersburg Mathematical Society and indicated some mistakes in a paper by the academician N. Ya. Sonin. Other incidents in Markov's style also happened.

The discussion of the MLSq was useful for Koialovich: in 1928, he (Mikhelson 1973, pp. 315 – 316) described the treatment of alcoholometric measurements for compiling appropriate tables. He derived a new interpolation formula and calculated its coefficients by the MLSq[2].

Now, Markov. In his [first] letter he partly described his thoughts about the MLSq which he developed later (1899). This contribution consisted of extracts from three of his letters to A. V. Vasiliev (1853 – 1929), the then president of the Kazan Mathematical Society. But why had Markov waited so long before publishing that paper? Chebyshev died in 1894 and it had been inconvenient to criticize him even indirectly[3]. In addition, Markov's remarks as stated in his letter to Koialovich should have been reformulated for publication.

Markov was also connected with Vasiliev by the latter's 27 reviews of his works which appeared in 1884 – 1899 in foreign periodicals. They corresponded from time to time. Then, Vasiliev (1898) published an essay on the work of Chebyshev and touched on his followers, certainly including Markov. Vasiliev said nothing about the incompleteness of the Chebyshev proof of the central limit theorem. Here is his phrase which Markov likely noticed:

The analysis which led Chebyshev to the theorem that <u>can be the basis</u> (my emphasis – N. E.) *of the MLSq will forever remain one of his most glorious achievements*[4].

Markov (1900) had no time to refer to his paper of 1899 and he only managed it in 1908, in the next edition of his treatise.

And so, Markov (1899) proved the Chebyshev theorem (1887) on expectations[5]. He specified its formulation, simplified and proved it more rigorously and indicated that he had achieved this *long ago*.

He [defended the second Gauss' justification of the MLSq] and [wrongfully] criticized Maievsky(1881), a well written and popular book. Maievsky was a general, a professor and corresponding member of the Petersburg Academy of Sciences and a good friend of Chebyshev with whom he collaborated at the Artillery Committee.

Markov (p. 250) also named *some Russian mathematicians*: they *Strove to derive the MLSq by a wrong application of the Chebyshev theorem which states that the probability is larger than some magnitude*.

Those mathematicians, see for example Lysenko (2000), included Yaroshenko and Sleshinsky whom Markov mentioned in his letter

to Koialovich. In the bibliography to [the separate chapters of] his treatise Markov (1900/1908) included ballistics experts Maievsky and N. A. Zabudsky and the astronomer N. Ya. Tzinger. In the edition of 1913, he replaced Zabudsky by Sleshinsky (by *one of the worthiest enemies*) and in 1924 he included Maievsky, Sleshinsky and Tzinger.

Markov's paper (1899) is known to become the guiding star for the justification of the MLSq[6].

Letter of A. A. Markov to B. M. Koialovich, 23 Sept. 1893
Ermolaeva (2009, pp. 92 – 97; her comments, pp. 100 – 108)

Dear Boris Mikhailovich,

Your *Theory of Probability* [lectures of 1892 – 1893; mimeographed edition without date] includes many points which it is quite impossible to agree with. I feel it necessary to make a reservation: these points are not new at all. However, it pains me very much to see them in a work written by my student and I consider it my duty to try somehow to eliminate them. This is the reason why I am writing this letter. I begin and proceed in the order of [your] pages.

On pp. 3, 4, 5 and 6 you attempt to clarify, as you say, the notion of probability, but in my opinion you did not attain this goal[7]. You consider an example in which there are 50 white balls, 25 black and 25 red balls, 100 in all, and you state that common sense indicates that we should expect the extraction of a white ball[8]. Do you really think so? Only a half of the balls are white. And what happens if there are 5 white balls, 2 black, 2 red, 3 blue balls and 1 yellow ball? Will common sense once more indicate that we ought to expect the occurrence of a white ball?

Then, on p. 6 the expression *assumed as certain* should be considered strange. You are thus only obscuring the matter[9]. How is it possible to doubt since we may say that a known ratio is the probability[10]? And was the notion of probability as a measured magnitude established previously? Without clearly, to some extent, establishing such a notion it is strange to attach the sense of some theorem, axiom or hypothesis to a simple definition. As to the significance of this definition, it is made clear subsequently (the Bernoulli theorem).

I note in passing that I am very much surprised by the words on p. 5:

We know how to solve this problem in two cases, although we cannot state that they are the only possible ones.

I certainly do not understand which cases you are speaking about[11].

In the sequel, I regrettably do not see that you pay attention to the change of the probability with the data. Your example of a die with four faces is closely linked with this point. The probability of which event are you discussing? Such an example ought to be preceded by a clearer description of the matter.

When a rather large number of trials is made, it is possible to reason thus, independently from the Bernoulli theorem (which can

hardly be applied here): the future should be a repetition of the past. And, after carrying out the same number of future trials, each event ought to be repeated the same number of times as previously.

We are thus immediately establishing the notion of equally possible cases and of the probabilities corresponding to the results of our trials. Not a quite correct hypothesis is certainly assumed here, but you will be unable to manage without a similar hypothesis. Indeed, when applying the Bernoulli theorem we ought to admit the existence of some constant probability and the mutual independence of the trials. But in the provided example this hypothesis is certainly incorrect. Indeed, after many throws the die should be more or less changed[12].

Among two equally incorrect hypotheses[13] it is natural to choose the simpler. However, in this place you have not introduced any hypothesis and everything remains unclear. Would it not be better to move such unclear points away from the beginning, i. e., to the section on the probability of future events? I do not know how you discuss such probabilities since you have apparently left them for the second part of your work [never published]. So let us go ahead.

On p. 25 you say: we ought to consider that these cases are equally possible since we know nothing about the degree of their possibilities. This is, once more, a very strange expression. You should have just said: these cases are obviously equally possible. Indeed, we have to admit that there exist obvious facts. And what kind of data can there exist when it is said that the number is written on the off chance[14]. If something should be explained here, then it is this expression. You believe that you have explained it on p. 24; actually, the matter should be understood in a simpler way[15].

Turning now to the repetition of events, I do not find any difference between dependent and independent trials in your work. To distinguish between them is however necessary[16]. Then, on p. 43 the Stirling formula is written in an inappropriate way [see Koialovich' letter of 25 Sent. 1893]:

$$1 \cdot 2 \cdot ... \cdot x = x^x e^{-x} \sqrt{2 \ \ x}.$$

The subsequent equations are therefore wrong. Your proof of the Bernoulli theorem is thus non-rigorous, and you have entirely left out its Chebyshev proof[17].

At the end of Chapter 2 you apply a formula whose error is unknown for calculating the probability that the probability is contained between certain boundaries. You do not say that, strictly speaking, such an approach is worthless. Under given conditions, will not the discarded magnitude be rather considerable? In your particular example[18], can you say, e. g., that the probability is 0.9 rather 0.8, 0.7 or 0.99?

When deriving the Laplace formula, I always consider it necessary to say that, although it is in general use, that formula should not carry us away because its error remains indefinite. It is certainly possible to indicate the boundaries of that error but then

everything will become very complicated and the result will barely be good.

I turn now to the MLSq. In the beginning you say that *its practical results almost always had been quite certain*. Many authors certainly say that as well, but do they have any right here? I believe, that none at all. Is it really possible to prove that statement? Would not it be more proper to say the opposite, to say that the results provided by the MLSq had never been quite certain, not in the least more certain than those furnished by other methods.

The main point is that the errors are from the very beginning assumed small. And, if they are very small, the results will be sound even without the MLSq. And, when the errors are large, no method will help. And so, it would have apparently been better not to state that the results are *certain*. Otherwise, it is necessary to prove such a statement.

Then, your reference to the works of Chebyshev, Yaroshenko and Sleshinsky surprised me. As far as I remember, there is no connection between them[19]. Then, if the works of Chebyshev which you mean are his paper *On two theorems*[20] and the previous papers, in which he treats a problem which I had solved, I am all but convinced that you do not know them. More precisely, you certainly know their titles and the author's assurance that he proves some or another propositions and solves such-and-such problems. But such a superficial acquaintance does not give you a right to refer to these works as confidently as you do.

Did the author really prove what he intended and is it possible to apply duly his results to the MLSq? This is the question which can only be answered doubtfully. I am convinced that no one except me has read Chebyshev's paper *On two theorems* since it is based on my results which he stated without proof[21]. Then, I am strongly objecting to Chebyshev's reasoning which make it doubtful. However, anyway, when taking into consideration my works, something can be elicited from his paper, but the MLSq will gain but little.

Then, as far as I remember, the papers of Yaroshenko and Sleshinsky have nothing in common with the works of Chebyshev. At present, I do not regrettably have those papers, but as far as I remember they make little sense. I remember that Sleshinsky had sent me a long paper and an additional page, on which I found an essential mistake which negates his conclusion. And in general it is necessary to regard their work with extreme caution[22].

But the main point is that common sense indicates the nonsense of the urge to prove the MLSq since it is based on arbitrariness. And it is due to common sense that the papers of Yaroshenko and Sleshinsky lack significance. I remind you that, when at the Petersburg Mathematical Society, in the presence of Messrs Shiff[23] you have incidentally expressed regret that mathematicians often forget common sense in their work. Examples of such oblivion are provided by every proof of the MLSq. Bearing in mind common sense I provide no proof and only establish notions necessary for practice.

The MLSq is based on weight[24] which is necessary and established more or less arbitrarily. It is necessary to find out how the weights of some results are determined by the weights of other results. Then the notion of mean square error is necessary for comparing the worth of various results with each other[25]. You certainly mention this error but do not show how to calculate it. In any case, I cannot find it in your work, and, when considering an example, you do not dwell on the determination of the product xh^2. By omitting this essential step[26] you have much decreased the significance of the MLSq. In my course, I do not disregard such essentially necessary elements.

We are really obliged to compare somehow one result with other results. And so, although you have devoted rather much place to the MLSq, you have not expounded all the necessary. However, you included something superfluous, namely, the expression of probability by an integral and the connected notions. It is all the more superfluous since you left a necessary element without a definition.

Then, I consider extremely strange your statement that you have chosen the most elementary method of exposition and your oblivion of the method which I had adopted. I have borrowed it from Gauss as well and it offers everything necessary in the simplest way. It does not require the proof of something which cannot be proved. Note that in my lectures I inserted some explanations which, as it seems, are included in the latest edition of my course[27].

Such an oblivion of the explication chosen by me which I had thoroughly thought out and which constitutes the most independent part of my course is extremely regrettable for me. Can it really be that my students do not understand me at all and are unable to avail themselves of my lectures?

I expected that, with some additions and explanations, you will keep to my exposition. In my lectures, I had always striven to explain that the discovery of the most probable results is not essential at all if the corresponding probabilities are zeros.

In a word, you have chosen the worst way, which disgusts me, to expound the MLSq. Without touching on the arbitrariness of the hypothesis that the arithmetic mean represents the most probable value of the sought number, I cannot pass over in silence your function ( ). How do you choose the sign of d ? You forget that in some instances the notion of the probability of separate values disappears and there appears a notion of the density of probability. Can it really be true that common sense does not tell you that your function ( ) is simply senseless?

You will perhaps say that you follow Gauss but this is wrong since Gauss has

$$(\Delta) = \frac{h}{\sqrt{}} \exp(-h^2 \Delta^2)$$

whereas you[28] write [the same, but with d attached]. The multiplier d makes your expression senseless.

It is seen that Gauss (1823) distinguished his function $\varphi(\ )$ and probability. He called it *la facilité relative* whereas the product was *la probabilité* of the error[29] to be situated between Δ and $\Delta + d\Delta$. Then, it would have been necessary to ascertain to which data does the probability

$$\int \varphi(\Delta)d\Delta$$

correspond. Indeed, for someone who knows the real result (it is indeed possible to imagine such a person)[30], this probability becomes certainty and [or] impossibility. However, our data hardly offer any possibility of establishing the notion of probability.

Finally I am unable to pass over in silence your statement on p. 57: you allege that *all* constant errors can be calculated and eliminated from the observations[31]. You are certainly repeating other authors but common sense tells us that these errors can only be approximately eliminated and, moreover, *only as far as they are known*. It is impossible to guarantee an entire elimination of constant errors, perhaps unknown to us at all.

I am now concluding my letter. A large part of my remarks will certainly remain useless. However, perhaps they will not disappear in vain as my university lectures apparently did. If we may conclude some agreement I am prepared to read your explanations and in general to exchange our thoughts. Perhaps however our viewpoints differ to such an extent that we will be unable to understand each other at all. Then we will have to admit that any explanations are superfluous.

Yours respectfully.

Letter from Koialovich to Markov, 25 Sept. 1893
Russian Academy of Sciences, Archive.
Fond 173, Inventory 1, 10, No. 1

Dear Andrei Andreevich!

I received your letter and am thankful for your attention to my lectures [1893][32]. I myself regarded them as a hasty work only written to provide my students some manual for recollecting my lectures. This explains some of the peculiar features of my book.

I regret very much that my lectures [as published] have thus impressed you, and I am still more sorrowful since I have felt in your letter such an unexpected appraisal of myself with which I can never agree: I consider it undeserved and unjust.

I came to this conclusion when reading for example your description of my application of the Stirling formula. Do you, Andrei Andreevich, really think that I am unaware of its real sense? The point is that I presumed that my readers are familiar with that formula. Therefore, for the sake of brevity and ease of calculation, without fearing any mistake, I replaced the exact formula

$$\lim \frac{1 \cdot 2 \cdot ... \cdot x}{x^x e^{-x} \sqrt{2\ x}} = 1, \text{ when } x = \infty$$

by a conjectural equality

$$1 \cdot 2 \cdot ... \cdot x = x^x e^{-x} \sqrt{2\ x}.$$

I assumed that the terms of the lower order in the right side are ignored.

I have explained all this in detail in my lectures and excluded the explanation from my book only because I considered that the proof of the Stirling formula does not concern the theory of probability. At the end of your letter you mentioned that you were ready to exchange some thoughts with me. Accordingly, I allow myself to offer some explanations and [am following in order your remarks].

It is certainly not for me to judge how properly I have explained the notion of probability[33], and I will only dwell on the example of a container with balls, 50 of them white, 25, black, and another 25, red, to which you object. I think that something is misunderstood here. Two problems are possible. First, if the occurrence of a white ball is set off against the occurrence of a non-white ball; and, second, if it is separately contrasted with the occurrence of a black, and a red ball.

In the first case you are in the right although your objection can be easily removed by only increasing the number of the white balls. However, it seems that I had stated quite clearly that the point really is, which of the three events, of the occurrences of a ball of one of those colours, is apparently the most probable. I believe that in this case you will agree that the occurrence of a white ball is the most probable.

Concerning my phrase on p. 6: *is assumed as certain* etc. It is very easily explained. When issuing from the general view on the theory of probability which I formed for myself, I think that the notion about probability took shape in the human mind long before some other methods of description have occurred.

I was able to choose one of the two methods: either to base myself on the notion of probability of error or to consider the sum of randomly selected magnitudes, just like Chebyshev did in his lectures whose manuscript notes I possessed[34]. I certainly prefer the second method but was unable to apply it owing to lack of time. I was compelled to turn to the first method.

I add the following. It seems that any description of the method of least squares (MLSq) has to be based on some more or less arbitrary assumption[35]. As I understand, it is impossible to avoid an assumption, and moreover it is not necessary. We are not constructing an abstract mathematical theory but have to do with a practical method of treating observations. I am far from having an idea of somehow proving the MLSq, and I have therefore applied the word *justify* (p. 55) rather than *prove*. I think that these words have absolutely different meanings[36].

I also note that among all those hypotheses one of which, anyway, we ought to assume when expounding the MLSq, the assumption about the properties of errors in any case seems the most natural. True, I know very well the difficulties connected with it.

After these general comments I turn to particulars. I called quite certain the results of the MLSq because as far as I know such are the main numerical results of our natural sciences, physics and astronomy[37]. You say that, given sound data, the results will also be good as well even without the application of the MLSq. I agree wholeheartedly, but among these results some can still be better than the others depending on how we combine the observations.

As to the literature, I had not at all thought of connecting the works of Sleshinsky and Yaroshenko with those of Chebyshev. For me, they are only common in that they expound the MLSq not like Gauss did, and it is only in this sense that I have mentioned them. I regret very much that I had not known your works and I apologize for having missed them[38]. They remained unknown to me because I have not seen any references to them and, as far as I can remember, during our talks you have not mentioned them. I had one of Chebyshev's works, which one I do not remember, but for such a short time that I have been unable to get minutely acquainted with it. For me, it was sufficient to see that he expounded the MLSq not like Gauss did.

Concerning the function which I denoted by ( ) I allow myself to turn your attention to the following. Suppose that we have a series of magnitudes

$$-n \;, \ldots, -2 \;, - \;, 0, \;, 2 \;, \ldots, \;, \; + \;, \ldots, \; +n \;.$$

If we assume that   is a finite but very small magnitude and that the error of observation can only take these values, the notion of probability is quite clear whatever is . Now we indefinitely decrease , then our assumption will however near approach a hypothesis of a continuous error and   becomes the differential of . My function   ( ) indeed expresses the probability of [as understood in] the theory of probability. The aim of this theory is not to create the notion of probability anew, but to explain it and make it measurable.

When calling a known expression the probability, we do not explain the difficulty but only move away from such an explanation[39]. Accordingly, I did not regard such a choice permissible.

The two cases mentioned on p. 5 are explained on p. 9. I touched on the problem of the change of probability with the change of data but did not think it necessary to include it in my book[40]. The die will change after many throws, but I may note that the opposite case was assumed. Indeed, otherwise I would have been compelled to mention the tiredness of the gambler, the change of the table or board on which the die was thrown etc.

I understand the notion of equally possible cases in the following way. To say that some cases are equally possible means, that we know nothing about these possibilities. The acknowledgement of one or the other is the same thing[41]. This indeed explains my phrase on p. 25.

As to the known cases of *apparently* equally possible, I would prefer to avoid this expression since it is very easy to consider certain those things which are very uncertain, as it is stated in the problem on pp. 16 – 17.

Concerning the notion of dependent and independent trials, I quite agree with you. I have indeed left a gap, although the subsequent description clarifies what kind of trials is discussed. I vividly remember that I said so in my lectures but omitted it in my book owing to an oversight.

And now I turn to the main point of our disagreement, to the MLSq. You reproach me since I, being your student, digress from expounding it in the same way as you did. Believe me, Andrei Andreevich, among your students there are hardly many of those who respect you more than I do. However, I retain my right to a free choice in scientific matters, and can never renounce that right. You yourself will be hardly pleased if your students restricted their efforts to blindly transmitting your lectures. And the reasons why I have chosen another way of exposition of the MLSq are these

1. I believe that your exposition will be too difficult for my listeners[42].

2. There is one point in your exposition which I was never able to clarify for myself either independently or after conversations with you. Here it is[43]. On p. 159 of your lectures (edition of 1891) you say:

In accord with the above, let us introduce magnitudes

$$x', x'', \ldots, x^{(n)}$$

which represent the possible results of the first, the second, …, the n-th measurement.

It is this place which I was unable to understand. What are these possible results? Under which conditions are they possible, and how do they differ one from another? I was unable to examine this without introducing once more the notion of probability of an error. I have therefore preferred to choose an error equal to   when assuming that $=d$ is an infinitely small magnitude.

The function   ( ) is therefore not only not senseless, it I very useful in that it eliminates the need to introduce the notion of density of probability in the MLSq[44].

I have based my statement about the possibility of eliminating constant [systematic?] errors on my information about the structure and application of astronomical instruments. I think that the very nature of constant errors conditions the possibility of their elimination[45].

I hope that my explanations will put an end to the misunderstandings and disagreements which originated between us.

Yours respectfully

Letter of Koialovich to Markov. 2 Oct. 1893
10, No. 8 of that archival source

Dear Andrei Andreevich,

First of all, I apologize for the delay of my answer caused by lack of free time. I know very well that there are many shortcomings in my course which should be later eliminated. I mentioned this to my friends as soon as my lectures were issued. I am fully aware that no human work is free from shortcomings and will be sincerely thankful for helping me to get rid of them.

You write: *All the same, you cannot write that*

$$1 \cdot 2 \cdot ... \cdot x = x^x e^{-x} \sqrt{2 \ x}$$

since this equality is wrong.

I answer: I wrote this equality when assuming that the terms of the lower orders in the right side are discarded. Under this assumption this equality is correct.

About the example on pp. 4 – 5: you say that when the number of the white balls is increased, we obtain a new example. I answer: yes, but

1. This new example nevertheless confirms the correctness of my words.

2. I clearly stated that we only choose a most probable event. And an event remains most probable however low its probability is if only it is higher than the probabilities of the other events. That low probability (of the most probable event) only indicates that our knowledge is insufficient, but does not influence the order of the events according to these probabilities.

You write that I attempt to obscure both this notion (of equally possible cases) and the concept of probability. I believe that I have the right to protest against such a charge because it rings oddly with respect to a man who nevertheless tries his best to reveal the truth. I can obscure the truth unintentionally but I cannot *attempt* to do so. That would imply such motives which you are unable to assume that I have[46]. To prove that I was justified to reason in a way which, you say, obscures the matter, I am asking you to turn your attention to the following difficulties which, as it seems to me, naturally arise when keeping to your viewpoint.

You write: Why may we not say in the theory that the well-known relation is probability? I answer: We certainly may, and we may say that anything is probability, but for whom will such a theory be compulsory and interesting? Will not everyone have the right to say: Am I concerned with theories operating on notions concocted by you yourself which perhaps have no representation in reality? Will not all this theory become *Übungen für den Verfasser* (as apparently Weierstrass expressed it)[47].

Therefore, Andrei Andreevich, by eliminating my seeming arbitrariness do you not insert instead your own arbitrariness, justified as little as my own is? I have indeed said myself that I am able to explain my arbitrary assumption and I do not explain it only because this problem touches on philosophy, a field in which I am not competent. You will now probably agree that I have grounds for remarking that we do not explain the difficulties but only keep

ourselves away from them when we say that probability is a certain ratio.

You write about the theorem on p. 7: If it really is a theorem, where is its proof? I answer: in my opinion, a proof is a reduction of the studied statement to the main propositions which are accepted as certain whereas the question of their certainty is not considered at all. I believe therefore that I did prove that theorem.

I have solved the problem of finding the number and the comparative significance of the chances in the two cases mentioned on p. 7. In the first case it is solved by applying the proposition mentioned on p. 5, in the second case, by referring to the inverse Bernoulli theorem (Chapter 2). An appropriate indication is on p. 10.

Concerning my example of a die with four faces, you say that my explanations indicate its groundlessness. Is it really so, Andrei Andreevich? Do not my explanations indicate something quite different? When solving this problem, we ought to allow only for those conditions which are stated there[48]. My problem does not mention any changes in the die and we need not therefore consider it.

I shall now dwell on the MLSq. I am naturally acquainted with Laplace's account and completely agree with you that I should have indicated it. I had certain motives for citing Sleshinsky and Yaroshenko; I myself have a low opinion about them, and especially with regard to the former, but these motives are subjective and not obligatory for anyone else.

You write that you perceive that I am not at all acquainted with the works of Chebyshev and others. This surprises me very much. Indeed, I held them in my hands and read them. Or, do not you trust me?

Concerning your account of the MLSq I allow myself to indicate the following. I know of course that your p. 159 carries a reference to previous explanations (probably to p. 157). I have not mentioned it because, to my understanding, nothing is explained either on p. 157 or elsewhere. Even your latest explanation, for which I am of course sincerely thankful, did not explain to me my most serious perplexity, namely:

As far as I understand you, you consider each separate observation as a value of a possible result[49]. Thus, a series of results

$$a_1, a_2, \ldots \qquad\qquad\qquad\qquad\qquad (A)$$

is possible for each measurement, and one of them is realised. I am prepared to understand all this concerning one observation. However, if there are, for example, two observations, then I cannot understand the difference between the series of all the possible results of the first observation (A) and the similar series for the second observation (B)[50]. The problem will certainly be solved at once if you say that the probabilities of the same error in these two series are different, but you will hardly want to introduce the notion of probability of error in your exposition[51].

You are asking how to understand the expression: becomes the differential of . I answer: since an infinitely small increment of the independent variable is indeed its differential.

I never said that *all* the numerical results of physics and astronomy are trustworthy, but only that all the *main* results are such. This is indeed very different. If an astronomer, issuing from numerical data on the motion of the Earth and the Moon forecasts an eclipse with a precision of one second, cannot we say that his results are certain? I think that certainty of the results in mathematics is quite different from certainty in natural science.

Therefore, we should not unnecessarily require the same of the latter as we do of the former. Concerning constant errors it seems to me that we thought about them in the same way and only spoke differently. For example, you write that they [the constant errors] cannot be entirely eliminated, but who doubts it? And who, while discussing their elimination, implied something different from decreasing their influence until the level of random errors? Then, you write: hardly anyone will say that the (personal) error can be entirely eliminated. I agree wholeheartedly, but who had considered that that error was constant? I think that all our misunderstandings occur because, when considering natural science, I applied various notions in the sense in which they are understood there.

Please excuse my illegible handwriting. I wrote hastily to answer you speedily. Respectfully, your obedient servant.

Notes to Ermolaeva's text and Markov's Letter of 23 Sept. 1893
(N. E.) means *according to N. E.*

1. Markov (Zhurnal 1904, p. 40) thus formulated his opinion (in particular):
It can also be regretted that […] *the author only took into consideration a small number of observations.* […] *His estimate of the possible errors in the derived coefficients shows that their precision is low.* (N. E.)

2. This does not at all prove that the *discussion* was useful. O. S.

3. I disagree: Chebyshev was not really interested in the history of the theory of errors (Sheynin 1994). O. S.

4. I certainly agree with the *glorious achievement* but not with its justification by Vasiliev (and Chebyshev himself). The central limit theorem has no place in either justification of the MLSq provided by Gauss. And Chebyshev had understandably not mentioned any other application of that theorem: at the time, the theory of errors was a main field for applying probability. The second and last main field was insurance of life. O. S.

5. Chebyshev first proved that theorem in 1867. O. S.

6. I utterly disagree, see Sheynin (2009, § 14.2-1). O. S.

7. This was impossible, see Rényi (1969, p. 82). The classical definition is simply a vicious circle. Still worse: it is not a definition, but a rule for calculating probabilities and only in the simplest case. Nevertheless, apart from the axiomatic theory, we only have that "definition" and the practically possible Mises approach.

Markov himself (1900/1912, the very first lines) stated that he will not consider in detail the foundations of the theory of probability. Even earlier he (1911/1977; translation: 1981, pp. 149 – 150) made known that he

Will not defend these basic theorems [laws of addition and multiplication] *connected to the basic notions of the calculus of probability* […] *since I know* [since he knows] *that one can argue endlessly on the basic principles even of such a precise science as geometry.*

See also Sheynin (2009, § 14.1-5). O. S.

8. It seems that Markov and Koialovich did not understand this problem in the same way. (N. E.)

9. Koialovich (p. 6) assumed as certain that the probabilities of the extraction of balls (or other objects) from an urn were proportional to the number of those balls. (N. E.)

10. See Note 7. O. S.

11. Koialovich (his second case) considered an irregular die with four faces so that the relevant probabilities were calculated statistically, by applying the Bernoulli theorem. (N. E.)

12. Koialovich properly answered this unreasonable remark. O. S.

13. An unfortunate expression. O. S.

14. Koialovich (p. 25) discussed the Chebyshev problem about the probability of cancelling a random fraction. (N. E.)

See Sheynin (2009, § 13.2-8). O. S.

15. Ermolaeva offered a possible explanation of Koialovich' understanding of that same problem and quoted Markov (1900/1924, p. 242) who had remarked that that problem should have been specified. O. S.

16. The trials in Koialovich' examples were physically independent and perhaps he had not therefore mentioned independence. However, he did consider dependent events. On p. 29 he offered a definition of independent events: their probabilities do not depend on the outcome of the previous events. N. E.

Cf. De Moivre (1718/1738 or 1756, p. 6):

Two events are independent when they have no connections one with another, and that the happening of one neither forwards not obstructs the happening of another. O. S.

17. Koialovich apparently thought that the details were not necessary: he had no time for them. He had not repeated the Chebyshev proof since he did not read his lectures to future mathematicians and, moreover, because that proof required the introduction of auxiliary theorems and of the notion of expectation which he decided to postpone. N. E.

18. That example was the celebrated Buffon problem of throwing a coin 4040 times. Koialovich only derived the boundaries of the probability sought. (N. E.)

19. Koialovich mentioned all the three authors who sidestepped the difficulties encountered in the justification of the MLSq by Gauss. Nevertheless, Koialovich kept to that justification since he thus avoided the necessary introduction of some new concepts and, moreover, first, that he thought that the description according to Gauss was more elementary and more natural, and, second, because the practical results of any approach were identical. (N. B.)

Neither those authors nor Ermolaeva thought about the second justification of the MLSq (Gauss 1823). O. S.

20. I think that Koialovich thought about Chebyshev's paper (1859). (N. E.)

21. See Sheynin (2009, §§ 13.1-4 and 14.2-3). O. S.

22. Markov thought about Sleshinsky (1892). Yaroshenko (1846 – 1917) was professor at the Novorossiysky University in Odessa and his paper (1893a) appeared a bit later. Sleshinsky (1854 – 1931) provided a thorough historical investigation connected with the debate between Cauchy and Bienaymé on the MLSq [cf. Sheynin (2009, § 10.2-6)]. The *additional page* mentioned by Markov was possibly Sleshinsky's short note (1893, perhaps still its manuscript), in which he intended to generalize Chebyshev's theorem about mean values on continuous variables. The first two works had certain merits: Sleshinsky expressed an idea of characteristic functions and Yaroshenko did not consider that the *most probable hypothesis* was serious. See Lysenko (2000) and Gnedenko & Gikhman (1956). N. E.

The introduction of characteristic functions is usually attributed to Poincaré and Liapunov, but Sleshinsky had preceded them. O. S.

23. Markov wrote this sentence carelessly and I adopted its interpretation as understood by Ermolaeva. She also provided information about Messrs Shiff and about the history of the Petersburg Mathematical Society. O. S.

24. Koialovich did introduce weights but elsewhere. (N. E.)

25. As an example, Koialovich considered the treatment of data on the elasticity of alcohol vapours as collected by the French physicist Regnault and calculated the weight of the result. He called his calculations approximate. N. E.

26. See the Gauss formula below. O. S.

27. Markov's lectures (1888) did not include any special explanations and the same is true about their edition of 1891. Note that a reference to Gauss (to the collection of his papers published in 1855) first appeared in his treatise (1900/1908). N. E.

28. Koialovich (p. 67) made a mistake. In two terms of a product he considered in a different way: as an infinitesimal and as a constant. (N. E.)

29. Bertrand called the probability density (a term introduced by Markov in his mimeographed lectures of 1884/1885, 1888 and 1891) *facilité relative*. N. E.

30. Cf. Markov (1924, p. 323; first published in the edition of 1908 and perhaps in 1900 as well):

In the first place, it is necessary to presume the existence of the numbers whose approximate values are provided by observations.

I (2007) studied the mathematical meaning of *the true value of a constant sought*. O. S.

31. Koialovich (p. 57) stated: *all those errors can be calculated and eliminated from the observations.* This is indeed the aim of the study of the instrument. N. E.

Notes to the Sequel

32. This phrase and some letters translated below testify that Markov had apparently attempted to acquaint himself with the entire Russian literature on the theory of probability and statistics. O. S.

33. See Note 7.

34. I can only say that Chebyshev (1879 – 1880/1936, p. 214) mentioned the sum of [random] magnitudes *possessing equal probabilities* but only considered the mathematical treatment of observations from p. 224 onward. O. S.

35. This is what Gauss himself (1823, § 6) stated: *the treatment partly depends on our arbitrariness*, on *free considerations*. O. S.

36. Markov himself (1899/1951, pp. 246 – 247) three times mentioned the *derivation* of the MLSq; in my translation, I mistakenly *corrected* Markov. But why was it necessary to derive or substantiate The MLSq since (Ibidem) it was only a *general method* lacking any optimal properties? Markov thus undermined his defence of the second substantiation of the MLSq by Gauss (1823). See also Koialovich' reasonable objection. O. S.

37. This is a superficial statement, suffice it to recall Newcomb who had encountered great difficulties in determining the astronomical constants (Sheynin 2002). O. S.

38. At the time, only Markov's mimeographed lectures (1882 – 1883/[1884?] – 1891) were published. It is unclear which works of Chebyshev he had in mind (see a bit below). Only his lectures were (partly) devoted to the treatment of observations but his deliberations were unfortunate (Sheynin 2009, § 13.2-7). O. S.

39. Did Koialovich reject altogether the classical definition? Cf. Note 7. O. S.

40. Markov possibly thought about a problem similar to his later, successively complicated problem (1924, pp. 5 – 9). His approach was in line with Laplace's thoughts about revealing the *vérité* by incessant verification and rectification (Sheynin 2009, § 7.2-1). O. S.

41. This approach is based on the *principle of indifference*, as Keynes (1921/1973, p. 44) called it. O. S.

42. Cf. Markov's own words about the *end* of his manual (1900/1908), i. e. about the MLSq (Ondar 1977/1981, Letter 15 of 1910, p. 21):

To my regret, however, I have often heard that my presentation is not clear enough. O. S.

43. The following seems hardly satisfactory: Koialovich apparently distinguished his ( ) and the (not yet studied) probability density whereas Markov introduced only one possible result for each observation. True, at least in 1924 he (pp. 323 and 374) stated the opposite. O. S.

44. But why should we eliminate the density? See also Note 43. When discussing the treatment of observations, Markov (1900/1908) did not say clearly enough that their errors were random, and that therefore they, practically speaking, possessed some density. At the time, the term *random variable* (or, in Russian, regrettably, *random magnitude*) was only emerging. However, random errors have been effectively discussed since Simpson. O. S.

45. This is wrong. O. S.

46. Grave (1993, p. 227) alleged that Markov had hostilely received beginners in science. However, at least here he rather acted in his usual manner. O. S.

47. An unfortunate statement: mathematics commonly deals with abstract notions which have no relation to reality. O. S.

48. If Markov did not carp at Koialovich, he in any case had not bothered to think out his criticisms. O. S.

49. See Note 43. O. S.

50. Koialovich had not inserted series (B). O. S.

51. Why should the observations in those series have different laws of probability? And why Markov would not wish etc.? O. S.

Bibliography

Abbreviation: IMI = *Istoriko-Matematicheskie Issledovania*

Bazhanov V. A. (2002, Russian), Professor A. V. Vasiliev. IMI, vol. 7 (42), pp. 120 – 148.

Chebyshev P. L. (1859), Sur l'interpolation par le méthode des moindres carrés. *Oeuvres*, t. 1. Pétersbourg, 1899. Editors, A. A. Markov, N. Ya. Sonin. Reprint: New York, 1962.

--- (1879 – 1880), *Teoriya Veroiatostei* (Theory of Probability. Lectures). Moscow – Leningrad, 1936. **S, G,** 3.

--- (1887, Russian), Sur deux théorème relatifs aux probabilités. *Acta Math.*, t. 14, 1890 – 1891, pp. 305 – 315.

De Moivre A. (1718), *Doctrine of Chances*. London, 1738, 1758. Reprint of last edition: New York, 1967.

Dobrovolsky V. A. (1967, Russian), Mathematics in the higher technical and special military educational schools. In Stokalo (1967, pp. 408 – 419).

Ermolaeva N. S. (2009, Russian), The method of least squares in a letter from Markov to Koialovich. IMI, vol. 13 (48), pp. 89 – 110.

Gauss C. F. (1809, Latin), *Theorie der Bewegung*, Book 2, section 3. In Gauss (1887, pp. 92 – 117). *Theory of Motion*. Boston, 1857; Mineola, 2004.

--- (1823, Latin), Theorie der den kleinsten Fehlern unterworfenen Combination der Beobachtungen. In Gauss (1887, pp. 1 – 53). English translation: Stewart (1995).

--- (1855), *Méthode des moindres carrés*. Paris. Translator: J Bertrand.

--- (1887), *Abhandlungen zur Methode der kleinsten Quadrate*. Editors A. Börsch, P. Simon. Vaduz, 1998.

Gnedenko B. V., Gikhman I. I. (1956, Russian), The development of the theory of probability in the Ukraine. IMI, vol. 9, pp. 477 – 536.

Grave D. A. (1993, Russian), Autobiographical notes. IMI, vol. 34, pp. 219 – 246.

Grodzensky S. Ya. (1987, Russian), *Andrei Andreevich Markov*. Moscow.

Keynes J. M. (1921), *Treatise on Probability. Coll. Writings*, vol. 8. London, 1973.

Koialovich B. M. (1892/1893, lectures), *Teoria Veroiatnostei* (Theory of Probability). No date, Petersburg.

Lysenko V. I. (2000, Russian), The method of least squares in Russia in the 19th century. IMI, vol. 5 (40), pp. 333 – 361.

Maievsky N. V. (1881), *Izlozhenie Sposoba Naimen'shikh Kvadratov i Primeneniya Ego Preimushchestvenno k Issledovaniyu Resul'tatov Strel'by* (Explication of the MLSq and of Its Application Mostly to the Study of the Results of Artillery Firing). Petersburg.

Markov A. A. (1884, 1885, 1888, 1891), *Teoria Veroiatnostei* (Theory of Probability). Petersburg. Mimeographed editions.

--- (1899, Russian), The law of large numbers and the method of least squares. *Izbrannye Trudy* (Sel. Works). No place, 1951, pp. 231 – 251. **S, G,** 5.

--- (1900), *Ischislenie Veroiatnostei* (Calculus of Probability). Later editions: 1908, 1913, 1924. German translation: Leipzig – Berlin, 1912. The posthumous edition of 1924 was perhaps not completely prepared by the author.

--- (1911, Russian), On the basic principles of the calculus of probability and the law of large numbers. In Ondar (1977/1981, pp. 149 – 153).

Mikhelson N. N. (1973, Russian), Boris Mikhailovich Koialovsky. IMI, vol. 18, pp. 310 – 321.

Ondar Kh. O., Editor (1977, Russian), *Correspondence between A. A. Markov and A. A. Chuprov.* New York, 1981.

Rényi A. (1969, Hungarian), *Briefe über Wahrscheinlichkeitsrechnung*. Berlin, 1969.

Sheynin O. [O. B.] (1989), Markov's work on probability. *Arch. Hist. Ex. Sci.*, vol. 39, pp. 337 – 377; vol. 40, p. 387.

--- (1994), Chebyshev's lectures on the theory of probability. Ibidem, vol. 46, pp. 321 – 340.

--- (2002), Newcomb as a statistician. *Hist. Scientiarum*, vol 12, pp. 142 – 167.

--- (2004), *Probability and Statistics. Russian Papers*. Berlin. **S, G,** 5.

--- (2007a), *Tretiya Khrestomatiya po Istorii Teorii Veroianostei i Statistiki* (Third Reader in History of the Theory of probability and Statistics). Berlin. **S, G,** 16.

--- (2007b), The true value of a measured constant and the theory of errors. *Hist. Scientiarum*, vol. 17, pp. 38 – 48.

--- (2009), *Theory of Probability. Historical Essay*. Berlin. **S, G,** 10.

Sleshinsky I. V. (1892, Russian), On the theory of the method of least squares. *Zapiski Matematich. Otdelenie Novorossiyskoe Obshchestvo Estestvoispytatelei*, vol. 14, pp. 201 – 264.

--- (1893, Russian), On the Chebyshev theorem. *Zapiski Novoross. Univ.*, vol. 59, pp. 503 – 506 of section 1.

Steklov V. A. (1991, Russian), [Correspondence and Reminiscences.] In *Nauchnoe Nasledstvo* (Scientific Heritage), vol. 17. Leningrad.

Stewart G. W. (1995), Translation of Gauss (1823) with Afterword (pp. 207 – 241). Philadelphia.

Stokalo I. Z., Editor (1967), *Istoriya Otechestvennoi Matematiki* (History of National Mathematics), vol. 2. Kiev.

Vasiliev A. V. (1898), P. L. Tchébychef et son oeuvre scientifique. *Bolletin di bibliografia et storia scienze*. A reprint. German version: 1900.

Yaroshenko S. P. (1893a, Russian), On the theory of the method of least squares. *Zapiski Novoross. Univ.*, vol. 58, pp. 193 – 208 of section 2. A reprint was circulated in 1892.

--- (1893b), Sur la méthode des moindres carrés. *Bull. Sci. Math.*, sér. 2, t. 17, pp. 113 – 125.

Zhurnal (1804), *Zhurnal Zasedaniy Soveta Imp. Peterburgsk. Univ. za 1903 God* (Minutes of the Sittings of the Council of the Imp. Petersburg Univ. for 1903). Petersburg.

VI

A. A. Markov

On the solidity of glass

Many physicists will perhaps be surprised that a mathematician who did not experiment decided to discuss the solidity of glass. However, I hope that this surprise vanishes after they find out that I am concerned with appraising the conclusions made about experimental data by the method of mean numbers[1]. The data and conclusions are contained in Galitzin (1902)[2].

I leave aside the question of whether it was necessary, either for scientific or practical aims, to look for a new formula of the braking pressure for a glass tube (Neumann 1885, p. 145)

$$400\frac{n^2-1}{n^2+1}.$$

Neither do I dwell on the question of whether the theoretical section of Galitzin's article should have been devoted to the description of the conclusion of a treatise (Lamé 1852) concerning a problem solved long before it[3]. Nor, finally, do I discuss how proper the conditions of the experiments corresponded to the theoretical assumptions. Instead, I turn to the conclusions of that paper.

First of all I ought to formulate a number of questions about the provided data. Table 1 provides the values of eight magnitudes,

V_m, V, R, R', d, n, P_m and T_m

three of which, d, n and T_m, are not known directly, but[4] determined by their connection with R, R' and P_m. Under these conditions the determination of the errors of n and T_m should have been preceded by the estimation of the possible errors of these three magnitudes but Galitzin had not even hinted at that.

From a mathematical point of view it is impossible not to consider erroneous the calculation of T_m with three significant digits in the cases in which P_m is only given with two digits.

It was very easy to estimate the influence of the errors of P_m on T_m at least for those cases in which, according to Table 1, a series of values of T_m corresponding to some values of P_m are given for the same n, had T_m been quite correctly calculated by n and P_m according to the formulas which the author had accepted. We have to indicate the opposite.

Formulas (25) and (26)[5] show, that for a constant n the increment of T_m is proportional to the increment of P_m. However, in Table 1, for $n = 1.12$, we find

$T_m = 3.22$ for $P_m = 40$ (experiment No. 49),

$T_m = 3.50$ for $P_m = 42$ (experiments NNo. 51 and 54),
$T_m = 3.55$ for $P_m = 44$ (experiments NNo. 48 and 50).

At first the increase of P_m by 2 led to an increase of T_m equal to 0.28, but then only to 0.05. In the same Table for $n = 1.15$ we find

$T_m = 3.44$ for $P_m = 52$ (experiment No. 57),
$T_m = 3.76$ for $P_m = 55$ (experiment No. 58),
$T_m = 3.86$ for $P_m = 58$ (experiment No. 61).

Here, the increase of P_m by 3 led to an increase of T_m at first by 0.32, then only by 0.10.

The influence of the errors of R and R' on n is certainly easy to determine by issuing from the formula

$$n = R/R', \qquad\qquad (1)$$

but this is not worthwhile until the possible errors of R and R' are ascertained. Those errors increase when we turn to T_m as calculated by magnitudes n and P_m. In Table 1 at $P_m = 42$ we have[6]

$T_m = 3.28$ if $n = 1.13$ (experiment No. 36)
$T_m = 3.50$ if $n = 1.12$ (experiment No. 54)
$T_m = 3.90$ if $n = 1.10$ (experiment No. 86)

It is clear that the errors in the third digit of n influence the second digit of T_m. The problem of the errors of R and R' thus becomes very important if, unlike Galitzin, we do not restrict our efforts to the *first approximation*, i. e., to the determination of the mean values of T_m from all observations but attempt to determine some dependence of T_m on n expressed by empirical formulas. Otherwise all the conclusions will possibly be caused by the constant errors, unequal for the internal and external cross sections of the tube since the influence of their errors on n and T_m varies with the variation of R and R'.

Concerning these magnitudes we ought to take into consideration that they are not invariable for different parts of the same tube: $R = 4.37$ for No. 48 and 4.50 for No. 51 although both experiments concern the same tube as indicated by the brackets[7]. This fact can partly occur owing to the errors of measurement and partly because the form of the tube deviated from a right circular cylinder.

But still there are no indications in Galitzin's paper of measurements of R and R' made repeatedly and in various directions. Even their values for the ends of the tube are lacking, and only their arithmetic means are [mean is?] provided.

One of Galitzin's conclusions is that T_m does not depend on the velocity of the increase of the pressure:

There is no essential connection between the values of T_m and the velocity of the increase of the pressure (p. 14);

Under a constant external pressure (1 *atmosphere*) *the solidity of various kinds of glass in the interval of measurement ought to be*

considered independent from the velocity of the increase of the pressure inside the tubes (p. 29).

However, this conclusion cannot be considered well-grounded. The value of T_m with a constant multiplier[8] determined by equality (25) is not only a function of many variables, but of unknown variable circumstances as well which, as we ought to assume, are very important. And so, T_m cannot be considered a function of one variable, either of V_m or n or of any other magnitude. It is very difficult to decide whether T_m depends on some magnitude or not.

In such cases only the method of mean values can offer some, although only a rather shaky indication. Especially shaky if the experiments are randomly and not systematically arranged and all the variables which can influence the studied magnitude are varying at the same time[9].

For proving that T_m does not depend on the velocity of the increase of the pressure Galitzin states that its same values were derived at differing velocities and that sometimes smaller and sometimes larger values of T_m corresponded to larger values of the velocity. However, the author considers it possible to maintain that T_m depends on n, although, taking for example the Thüringen glass, Table 1 shows that for very different values of n there occur values of T_m equal or near to 4.21[10] and that when n increased, T_m increased in some cases, and decreased in other instances (absolutely not according to Table 21 or Table 23[11]). It should have been therefore possible to admit that T_m does not depend on n, and allowing deviations of 50% (as Galitzin did) it would have been possible to obtain

$T_m = 5.4$ or 5.5.

And so, in the first approximation it was apparently possible to conclude, according to the theoretical assumption, that T_m was constant[12]. Had the author restricted his deliberations to this first approximation and disregarded the conclusion of the treatise[13], he would have written a short note possibly suited for some technical edition. I would have taken no notice of such a note.

But the author did not stop at the *first approximation*. He finds it possible to state that T_m depends on n and even attempts to express this dependence by a table or a graph, and I therefore consider it necessary to indicate that he could have just as well at least maintained that T_m depended on V_m.

I have therefore compiled a table [not included here] which allows me to conclude that for the Thüringen glass T_m increases with V_m. I made use of all the numbers of Table 1 concerning that glass except those which pertain to tubes of thick glass (NNo. 12 – 18) since the author himself decided to reject three of those results out of the seven available. I also excluded those for which the values of V_m were not given.

When beginning to compare my table with Galitzin's Table 20 or 21 which served him for conclusions about the dependence of T_m on n, I note first of all that a large part of the numbers in those tables are arithmetic means of three, four or five numbers. We will therefore

estimate the possible errors of the difference between 1) 6.97 [the mean value of T_m in Markov's table for seven experiments and V_m 2.9] and 5.05 [for eight (and V_m 1.5) or nine experiments] and 2) between the two values of T_m in Table 20 for the Jena glass for n = 1.46 and 1.33.

Simple calculations[14] provide the square of the mean error of the number 5.05 (Khvolson 1897, vol. 1, p. 244; Tzinger 1899, Chapters 5, 6 and 7) less than 0.21 for the Thüringen glass at V_m 1.5. For the number 6.97 for the same glass at V_m 2.9 the square of the mean error is less than 0.14. According to a known formula[15] we have for the difference of those numbers, 6.97 and 5.05, the mean error

$$\sqrt{0.21+0.14} < 0.6$$

which is smaller than 1/3 of that difference.

For the Jena glass the difference is between 6.53 and 5.50 (Table 20), and, according to the same formula, it has mean error which is larger than 1/2 of that difference.

Nevertheless, the author (p. 23) decides to say:

Considering now more attentively the numbers of Table 20, we see that in the interval of measurements for the Jena glass T_m steadily increases with n and the dependence is almost linear.

Finally, for the Thüringen glass (Table 21) the mean errors of the difference of T_m between 6.84 and 5.16 at n = 1.36 and 1.18 and, respectively, between 6.84 and 5.59 at n = 1.36 and 1.51, they are

$$\sqrt{0.81+0.13} > 0.95 \text{ and } \sqrt{0.81+0.25} > 1.$$

But still the author decides to say that

For the Thüringen glass T_m at first increases with n until a certain maximal value (*approximately until n* = 1.36) *then gradually decreases.*

These conclusions certainly cannot be recognized as well-grounded for the Jena, and much less for the Thüringen glass not only since the large value of the mean error shows a large disagreement between the mean numbers but also because of the abovementioned causes. For avoiding misunderstanding I repeat that Galitzin provides no data for deciding whether T_m really only depends on (1) but not on R and R' separately.

In concluding, I consider it not superfluous to indicate some facts characterizing Galitzin's paper.

For the Jena glass he provides the formula

$$T_m = A + Bn,$$

and even two formulas for the Thüringen glass:

$$T_m = A + Bn + Cn^2 \text{ and } T_m = A + \frac{B}{n} + \frac{C}{n^2} + \frac{D}{n^3}.$$

One of them allegedly corresponds to certain values of n, the other one, to other values of that magnitude. However, the coefficients of those formulas are not given and without them we certainly cannot say how these formulas correspond to Tables 20 and 21 which served as the basis for their derivation, or to Tables 22 and 23 and the curves on the applied graph. We can only indicate that for the Thüringen glass the number of data in Table 21 is equal to the total number of the coefficients of the empirical formulas and that in spite of such an abundance of those coefficients, some numbers of Table 23 essentially differ from the corresponding numbers of Table 21. At $n = 1.10$ the magnitude T_m is 3.64 and 3.9 respectively; at $n = 1.18$, $T_m = 5.16$, but in Table 23 we have to conclude that T_m is situated between 4.5 and 5.

And, when considering the graph (Fig. 1) appended to the paper, I noticed that the line there for the Thüringen glass does not correspond either to Table 21 or Table 23; at $n = 1.20$ Table 23 gives $T_m = 5.0$ but it is much larger on the graph, approximately equal to 5.5. At $n = 1.26$ $T_m = 5.44$ whereas it is about 6 on the graph.

From the 27 numbers of Table 23 more than a half are in the interval $n = 1.55 – 2.30$ for which there are no observations. Finally, the calculation of the breaking pressure of a tube of thick glass ($n = $) can serve to ascertain the author's attitude towards his empirical formulas and tables. In this calculation (p. 28) T_m is supposed to be 8.03 although it is lacking in both Tables 22 and for $n = 2.3$ Table 23 gives $T_m = 4.4$[16].

We certainly do not know the value of T_m for $n = $ according to the second empirical formula for the Thüringen glass but we cannot forget Galitzin's statement (p. 29):

At first, the solidity T_m increases with n (for both kinds of glass) but later it more or less decreases with the further increase of n (Thüringen glass).

On the same [?] p. 28 the author assumes as an example $n = 10$. I do not know in what measure are the conclusions of the experiments applicable to tubes of such thick glass with n a few [three or four?] times smaller than 10 but I suppose that in any case it is more interesting to choose as an example such a value of n which had occurred in the experiments. At $T_m = 8.03$ and $n = 2.30$ formula (26) provides P_m 469, but according to Table 1 (No. 13) $P_m = 258$ which is much smaller. That experiment was excluded from consideration and Table 25 provides $P_m = 258$ for $n = 2.3$.

I certainly leave for physicists the formulation of a final opinion about Galitzin's paper.

Notes

1. The method of mean values (or, as below, of mean magnitudes) as Markov called it is the method of means. In 1805, Condorcet introduced the term *theory of means* but had not indicated the connection of that theory with the theory of probability. The new term became firmly established not later than in 1830. Later Quetelet applied it and in 1850 Humboldt called it the *method of mean numbers* the only *decisive* method of treating observations. Davidov (1857) published a *Theory of mean values* and even in 1901 Hilbert in his celebrated report on the problems of mathematics stated that it was necessary to develop the theory of mean values in the kinetic theory of gases. At that time the applied term was already dated.

Lambert, in 1765, introduced the term *theory of errors* which neither Laplace, nor Gauss ever used. It only became established in 1845 – 1860 and for a few decades it existed along with the *theory of means*. The latter considered any means (for example, mean stature of men) but the theory of errors only dealt with observations of some constant. Davidov was the first who stressed the unity of treating observations in all cases. See Sheynin (2007, pp. 44 – 46).

2. Markov reviewed the paper Galitzin (1902). This is the author's spelling (as perhaps it is in all his German papers) although the Russian spelling is *Golitzin*. Markov never mentioned him directly and thus complicated his style. I did not follow his strange example. His bibliographic descriptions were incomplete and sometimes inaccurate and I improved them.

3. Markov referred to a *manual* instead of *the manual Evnevich (1868) cited by Galitzin*. We see once more how Markov treated his readers! His remark is far from unquestionable. Galitzin had not touched on the history of his problem and for that matter his paper was mostly devoted to the treatment of observations rather than to theory. Markov himself referred to Khvolson and Tzinger rather than to Gauss. Note that Mendeleev (1874/1939, p. 187; 1875/1939, pp. 481 – 490) investigated the solidity of tubes but had not explained the applied method of treating observations. He determined the burst strength of the tubes when the pressure was applied from within them.

4. In his Table 1, Galitzin provided the results of the investigation of the solidity of glass when the pressure was applied from within. He studied 99 tubes made from different kinds of glass, mostly Thüringen and Jena glass. The meaning of the magnitudes mentioned by Markov is: V_m, mean velocity of the increase of the pressure; V, that velocity just before the tube bursts; R and R', the outer and the inner radii of the tube; $d = R - R'$; $n = R/R'$; P_m, maximal strength of the tube; and T_m, the measure of the solidity of glass.

5. Formula (25) which Galitzin theoretically derived was

$$T_m = \frac{1}{4}\{5P_m + 7[\frac{P_m - 1}{n^2 - 1} - 1]\}.$$

Formula (26) defined P_m through T_m. Galitzin however multiplied T_m by a factor which allowed for the transition from one unit of measurement to another.

6. Experiments NNo. 36 and 54 were made with the Jena glass, and No. 86, with the Thüringen glass.

7. Brackets in Table 1 united those experiments which were made with the same tube cut in parts.

8. See Note 5.

9. Naturalists attempted to avoid a simultaneous change of some arguments (factors) in their experiments. The theory of experimental design originated by Fisher in the 1920s – 1930s became able to allow such changes.

10. For example, experiments NNo. 18, 20, 81 and 89.

11. In all, there were 25 tables. See Note 4 about Table 1. Tables 2 – 19 were derived from Table 1. Each represented experiments with one kind of glass and, with one exception which does not concern us, corresponded to a small interval of the values of n. In addition to n, each indicated the relevant values of d and T_m.

Tables 20 and 21 were summaries for the Jena and the Thüringen glass respectively. Tables 22 and 23 gave the results of the calculation of T_m by Galitzin's empirical formulas (see below). Finally, Tables 24 and 25 were summaries of P_m calculated for the arguments n and T_m taken from Tables 22 and 23.

12. What assumption? This is unclear since one of Galitzin's formulas connects T_m with n, see Note 5.

13. See Note 3.

14. Here and below Markov referred to the formula for the square of the mean square error of the arithmetic mean [of independent terms].

15. Here and below Markov referred to the formula for the mean square error of a sum (a difference) of some number of [independent] terms.

16. Number 8.03 was the maximal value of T_m in Table 1. The tube with the thickest walls in Table 23 had $n = 2.35$ and the corresponding T_m was 4.4.

Bibliography

Davidov A. Yu. (1857), *Teoria Srednikh Velichin* […] (Theory of Mean Values and Its Application to the Construction of Mortality Tables). *Rechi i Otchiet Proiznessennye v Torzhestvennom Sobranii Moskovskogo Universiteta*. Moscow. Separate paging.

Evnevich I. A. (1868), *Rukovodstvo k Izucheniyu Zakonov Soprotivleniy Strolitel'nykh Materialov* (Treatise on Studying the Laws of the Strength of Building Materials). Petersburg.

Galitzin B., Fürst (1902), Über die Festigkeit des Glases. *Izvestia Imp. Akad. Nauk*, ser. 5, vol. 16, No. 1, pp. 1 – 29 of the third paging. His graph is on a separate unnumbered page following p. 52.

Khvol'son O. D. (1897), *Kurs Fiziki* (Course in Physics), vol. 1. Petersburg.

Lamé G. (1852, 1866), *Lecons sur la théorie mathématique de l'élasticité des corps solides*. Paris.

Mendeleev D. I. (1874, Russian), [Report on the study of the solidity of glass.] Extract from the Minutes of the sitting of the Chemical Society. *Sochinenia*, vol. 6, 1939, p. 187. Leningrad – Moscow.

--- (1875, Russian), On the elasticity of gases, pt. 1. Ibidem, pp. 221 – 589.

Neumann F. (1885), *Vorlesungen über die Theorie der Elastizität der festen Körper* etc. Leipzig.

Sheynin O. (2007), The true value of a measured constant and the theory of errors. *Hist. Scientiarum*, vol. 17, pp. 38 – 48.

Tzinger N. Ya (1899), *Kurs Astronomii* (Course in Astronomy), theoretical part. Petersburg.

<div align="center">

VII

Oscar Sheynin

Markov on a paper of Galitzin

</div>

1. Discussions held by academicians (*Prenia* 1903). In 1902, the extraordinary academician of the Petersburg Academy of Sciences, Prince Boris Borisovich Golitzin (1862 – 1916) established that the centres of earthquakes can be determined by data obtained at a single station and thus solved an important problem in seismology. Along with his other scientific merits this subject became the cause of the discussion of his nomination to the title of full academician which was proposed by seven academicians including the astronomer F. A. Bredikhin.

The Discussions (*Prenia* 1903) show that Markov and then Liapunov negatively described some of Golitzin's investigations and he was not nominated.

Liapunov criticized one of Golitzin's contributions belonging to applied mechanics whereas Markov destructively described Golitzin's study (1902) and repeated his criticism in a paper [   ].

Golitzin's study contained a brief theoretical part, but its main aim was the publication and mathematical treatment of his own experiments. This study did not at all belong to the body of Golitzin's scientific interests and someone else could have only admonished him and agreed with those seven academicians. Owing to his character, Markov, however, was unable to restrain himself the less so since the treatment of observations had always been important for him, and has been incessantly essential for any naturalist. On top of that, Golitzin published his paper in a journal of the Academy and just before the described discussion.

Markov (*Prenia* 1903, p. 5) properly stated that Golitzin's paper (1902)

Is scientifically insignificant and hardly considerably useful in the practical sense owing to the large disagreement between the observations discussed there.

He added that there was not time enough for confirming Golitzin's achievements in seismology.

Neither Golitzin, nor Bredikhin who supported him was able to object seriously. They stated that the study of solidity of some substance was only possible with large errors, Golitzin, however, provided allegedly precise quantitative results. One fact additionally characterises his work. Markov noted that Golitzin's graph did not tally with his tabulated data. So what? Golitzin explained that his graph was only a sketch. Markov did not agree since this *sketch* was provided with a coordinate system and, anyway, Golitzin never said anything of sorts.

2. The history of Markov's paper. *Izvestia* of the Academy of Sciences (ser. 5, vol. 18 for 1903, p. XIX) announced that Markov had

submitted his paper [   ] and that it will be published. However, it did not appear either there, or, as I believe after studying some bibliographic sources, anywhere else. Its manuscript had been preserved in Markov's family and a few years ago Markov Junior had given it for a time to Grodsensky who mentioned it in his book (1987, p. 65), returned it to the Archive of the Academy of Sciences and informed me about it.

I say *manuscript.* More properly, it is a text prepared in a printing house in the format adopted in the *Izvestia*; for example, Markov's name is typed in an oblique case. Below the text of the first page there is a stamp with an inscription:

Printing house of the Imp. Academy, 3 proofs. Sent 9.IX.903 (the date is written by hand), *returned* …

There are many non-essential misprints (which, unlike Markov, I had corrected although without stating it) and it ought to be presumed that Markov had not returned this proof. In other words, that after all he refused to publish it. Indeed, the text was not really fit for publication since it was impossible to understand it without reading Golitzin's paper as well. Then, the *Discussions* (*Prenia* 1903) had been, or were about to be published, and, finally, Golitzin could have at least largely privately agreed with Markov.

3. Mathematical treatment of observations. Markov had to spend much time in formulating his criticisms. To repeat, the treatment of observations always interested him; see, however my pertinent remarks [xi, Note 9].

Markov (1900) described the method of least squares together with the study of statistical series, interpolation and the Lexian theory of dispersion. Such an approach is questionable, but at least it reflected his attempt, unsuccessful at the time, to include directly the method in theoretical statistics. He paid due attention to the estimation of the reliability of observations. Here is one of his statements (*Prenia* 1903):

I like very much Bredikhin's rule according to which 'in order to admit the reality of a computed quantity, it should at least twice numerically exceed its probable error'. I do not know, however, who established this rule or whether all experienced calculators recognized it.

Both Newcomb and Mendeleev had also been applying the same rule (Sheynin 2009, §§ 10.10.3 and 10.9.4). However, just like the much more popular rule of *three sigma*, which is still remembered nowadays, it does not depend on the number of the observations.

4. Correlation theory. Markov [vi] noted that the magnitude most interesting for Golitzin

Is not only a function of one variable, but of unknown variable circumstances.

Nowadays this indication means that Golitzin's measurements should have been treated in accord with the requirements of the correlation theory. Up to 1903 it was only beginning to develop but even later Markov (1916/1951, p. 533) did not recognize it:

Its [the *fashionable* correlation theory's] *positive side is not significant enough and consists in a simple usage of the method of*

least squares to discover linear dependences. However, not being satisfied with approximately determining various coefficients, the theory also indicates their probable errors, and enters here the realm of imagination, hypnotism and belief in mathematical formulas that actually have no sound scientific foundation.

Revealing dependences (even if only linear) is nevertheless important. In the posthumous edition of his treatise Markov (1900/1924) did not repeat his criticisms, but neither had he considered correlation theory in any detail.

5. Some events after 1903. Golitzin had naturally continued his scientific work and in 1908 he became ordinary academician. This time Markov offered a positive reference[1]. In 1911 Golitzin was elected president of the International Seismological Association and in 1916 he became a fellow of the Royal Society. He is justly considered a co-founder of modern seismology.

Golitzin's relations with Markov either remained, or became normal once more. Indeed, in a letter of 1913 to Chuprov Markov (Ondar 1977/1981, p. 70) remarked about his report at the forthcoming celebration on the occasion of the bicentenary of the law of large numbers:

I have not dwelt on the question of the importance of the law of large numbers to physics. I will have to talk with prince Golitzin about this.

Nevertheless, his report (Ondar 1977/1981) did not mention physics and nothing is known about his conversation with Golitzin or whether it did occur at all.

In 1916 Markov once more, although indirectly, attacked Golitzin. He formulated his misgivings about correlation theory (§ 4) when criticizing a paper (Tikhomirov 1915) published by the Main Physical (later, Geophysical) Observatory. Its director and editor of that edition was Golitzin. Markov reported his forthcoming paper when Golitzin had still been alive, but it was published after his death.

Note

1. Archive of the Academy of Sciences, Fond 1, Inventory 1a – 1908, Delo 155, pp. 118 – 118 reverse. I am grateful to Natalie Ermolaeva for indicating this source.

Bibliography

Galitzin B. B., Fürst (1902), Über die Festigkeit des Glases. *Izvestia Petersb. Akad. Nauk*, ser. 5, vol. 16, No. 1, pp. 1 – 29.
The *Russian* spelling of his name is Golitzin.

Grodzensky S. Ya.(1987, Russian), *Markov*. Moscow.

Markov A. A. (1900), *Ischislenie veroiatnostei* (Calculus of Probability). Later editions: 1908, 1913, 1924. German translation: 1912.

--- (1914, Russian), Bicentennial of the law of large numbers. In Ondar (1977/1981, pp. 158 – 163.

--- (1916, Russian), On the coefficient of dispersion. In author's *Izbrannye Trudy* (Sel. Works). No place, 1951, pp. 523 – 535.

Ondar Kh. O., Editor (1977, Russian), *Correspondence between Markov and Chuprov on the Theory of Probability and Mathematical Statistics*. New York, 1981.

Prenia (1903), *Prenia mezhdu akademikami* … (Discussions held by Academicians …). Petersburg.

Sheynin O. (2009), *Theory of Probability. Historical Essay*. Berlin. **S, G,** 10.

Tikhomirov E. (1915, Russian), Method of correlation and its applications in meteorology. *Geofizichesky Sbornik*, vol. 2, No. 3, pp. 21 – 48.

VIII

L. I. Emeliakh

The case of the excommunication
of academician A. A. Markov from the Church
(From the history of the struggle of Russian scientists
against the religious obscurantism)

Introductory Note by M. I. Shakhnovich

[…] The cause of Markov's direct break with the Church was the pestering of Tolstoy by clericals[1]. […] In 1912, he requested the Synod of the Russian Orthodox Church to be excommunicated […]. Here is this request whose copy is preserved by the Russian Academy of Sciences (Fond 173, Inventory 1, No. 65, pp. 1 – 2):

To the Most Holy Governing Synod

I have the honour to most obediently ask the Most Holy Synod to excommunicate me from the Church. I hope that a sufficient cause for the excommunication may be a reference to my book *Ischislenie Veroiatostei* (Calculus of Probability). There, I clearly expressed my negative attitude to the legends which underlie the Jewish and Christian religions. Here is an excerpt from this book (1908, pp. 213 – 214):

Independently from the mathematical formulas, which I am leaving aside without attributing them any essential significance, it is clear that we should extremely doubtfully regard stories about incredible events that had allegedly occurred in the long passed time. And we cannot at all agree with academician Buniakovsky (*Osnovania Matematicheskoi Teorii Veropiatnostei* [Principles of the Mathematical Theory of Probability. Petersburg, 1846], *p. 326*) *in that it is necessary to separate a class of such stories, in which, in his opinion, it is reprehensible to doubt.*

So as not to deal with still more severe judges and avoid charges of shaking the foundations I do not dwell on this subject which does not directly belong to mathematics.

For clearing up any remaining doubts about what I am describing, I adduce an excerpt from Buniakovsky's book […].

If this is not sufficient, I am most obediently asking you to take into account that I do not detect any essential difference between ikons and idols which [?] certainly are not gods or their images, and neither do I sympathise with any religion which, like the orthodoxy, are supported by, and in turn lend their support to fire and sword.

        Academician A. Markov. 12 February 1912, Petersburg
[…]

The main article

The [copies of the] quoted documents are being kept in the Central State Historical Archive in Leningrad: 1912, No. 2988. On academician A. A. Markov's request to be excommunicated from the Church.

1. 24 February 1912. According to His Majesty's decree the Most Holy Governing Synod *heard out* the request of academician A. A. Markov […] to excommunicate him from the Church since he negatively regards the legends which underly the Jewish and Christian religions, does not detect any essential difference between ikons and relics on the one hand and idols on the other hand and *does not sympathise* with orthodoxy. […]

Ordered. Send this request together with the decree to the Right Reverend metropolitan of Petersburg with an assignment: to order pastor admonition and conviction for the petitioner. […]

2. Dear Sir, Filosof Nikolaevich,

If the aim of the conversation proposed by you consisted in obtaining from me some useful indications in the line of my speciality, I would have deemed it my duty to converse with you and to assist you to the extent of my possibilities. However, I consider it necessary to avoid conversations which cannot be useful either for me or my interlocutor, since they can only result in loss of time and mutual irritation. […]

<div align="right">Willing to serve A. Markov 17 April 1912</div>

3. To the Right Reverend Antoniy, metropolitan of Petersburg and Ladoga

From archpriest Filosof Ornatsky

Report. […] I consider it my duty to report that […] I have sent Mr. Markov a letter in which I asked him to choose a day and an hour for a conversation with him about his petition. Academician Markov answered me by letter [see No. 2] in which he resolutely refused to converse with me. In his opinion, such a conversation cannot be useful either for him or his interlocutor and can only result […]

<div align="right">20 April 1912</div>

4. To the Most Holy Governing Synod from Antoniy, a member of the Synod, metropolitan of Petersburg and Ladoga

Report. […] Archpriest Ornatsky […] applied to Mr Markov […]. Since Mr. Markov resolutely refused to converse with the appointed archpriest Ornatsky, the Petersburg diocese authorities find that Markov should be considered as seceded from God's Church and that he should be expounded from the list of Orthodox believers. […]

<div align="right">20 May 1912</div>

5. Draft. 1 June 1912 […] The Most Holy Governing Synod […]

Ordered. Perceiving […] that academician Markov, who announced his foolhardy intention to be excommunicated from the Church, refused to hear the admonition and exhortation of the spiritual pastor to abandon his intention, the Most Holy Synod resolved: to send

Markov's petition for a decision to the Petersburg diocese authorities. […]

Note. It is indicated that [excommunication will be] too honourable for Mr. Markov. […]

6. Final text. [It entirely coincides with the draft although the Note is not reproduced.]

7. To the Most Holy Governing Synod from Anatoliy [?], a member of the Synod, metropolitan of Petersburg and Ladoga

Report. […] Academician A. Markov refused […] and remained inexorable in his intention to be excommunicated from the Church. Therefore, the authorities of the diocese, in accord with their decision of 8 – 16 May, resolved, in their new decision of 28 September – 4 October to consider that A. Markov had seceded from the God's Church and that he should be expounded from the lists of Orthodox believers.

This decision should be reported to the Holy Synod and made known to the governor of Petersburg. Independently from this, to order Mr. Markov to report to the consistory the place and year of his birth and baptism and the names and patronymics of his parents. […]

The most obedient servant of your Holiness Antoniy, metropolitan of Petersburg [and Ladoga], 19 October 1912.

8. Draft. 31 October 1912

[…] The Most Holy Governing Synod […] ordered: To consider A. Markov as seceded from the God's Church and to expound him from the lists of Orthodox believers, to order him to report to the consistory the place and year of his birth [baptism is not mentioned] and the names and patronymics of his parents.

To order the arch-public prosecutor to make known the contents of this report of the Right Reverend metropolitan Antoniy to the Minister of People's Education with a request to order a report to the Petersburg spiritual consistory about the abovementioned information concerning academician Markov.

9. Draft. [The text is essentially the same.]

10. The final document. [Essentially repeats the texts of NNo. 8 and 9.]

11. To the Most Holy Governing Synod from Nikandr, the Narva bishop, temporarily managing the Petersburg diocese

Report. In addition to the reports of the deceased Right Reverend metropolitan Antoniy [NNo. 4 and 6] I report […] Markov's answer:

My parents are Nadezhda Petrovna Fedorova and Andrei Grigorievich Markov. I was born 2 June 1856. Do not have my birth certificate at hand. Will not report anything else.

7 December 1912

12. [Information from the minutes of the Synod for 21 December 1912 (7 January 1913). Nothing new.]

Note

In 1901 Tolstoy was excommunicated from the Church. Then, during his last days, the Synod discussed whether he should be *admitted to the bosom of the Church* and decided against it (Anonymous, The Holy Synod and Tolstoy. Newspaper *Rech*', 8 Nov. 1910, p. 3). This fact was probably remembered in 1912, although it does not directly explain Markov's decision. Anyway, Shakhnovich' opinion was unjustified and seems hollow (pestering the deceased Tolstoy twelve years after his excommunication)! I venture to suppose that Markov's request was prompted by the notorious Beilis case (1911 – 1913). Cf. [x, text based on author's pp. 104 – 105]. Markov's bold decision is barely remembered in Russia where the Russian Orthodox Church is now reigning supreme.

Oscar Sheynin

**Markov's letters in the newspaper Den', 1914 – 1915
with Supplement**

1. Introduction

1.1. General explanation. Markov had written many letters to various newspapers. His sharp statements which described burning social and political issues are interesting enough. Grodzensky (1987), see also [x], reprinted apparently most of them (although not those translated below) and thus rendered a serious service to his readers. Regrettably, he did not say which of these letters were indeed printed.

Markov's social activity (undoubtedly including his newspaper letters) earned him the nickname *militant academician* (Nekrasov 1916, p. 9). Neyman (1978; Ondar 1981, p. 4) made known his other nickname *Neistovyi Andrei* which he translated as *Andrew the irrepressible* with the addition of *who does not pull any punches*. In 1981, he somewhat spoiled his account by attributing to Markov Pushkin's satirical verse.

I reprint (now, translate) the three letters devoted to the methodology of, and instruction in mathematics and published in *Den*', a leftist newspaper. After the February 1917 revolution, Mensheviks were included in its editorial staff, but in 1918 it was closed. Minkovsky (1952) briefly described Letter 1 and Nekrasov (1916, p. 8) referred to Letter 3. I discovered Letter 2 by following Markov's vague reference in Letter 2.

Pavel Alekseevich Nekrasov (1853 – 1924), see Sheynin (2003), an alumni of Moscow University, became in 1885 a privat (freelance)-dozent there, then professor and rector. He played a noticeable role in the community of Moscow mathematicians, published interesting investigations in algebra and his doctor dissertation, *The Lagrange series* of 1886, deserves attention. Regrettably, his works were not studied properly; his name is not even mentioned in Russian biographical dictionaries and Youshkevich (1968) provided insufficient information about him.

His former obscurity was certainly caused by his reactionary political views and work as rector of Moscow University (1893 – 1898), as a civil functionary responsible for the Moscow educational region and, since 1905, a member of the Council of the Minister himself of People's Education in Petersburg. Archival documents (*Istoria* 1955, p. 378) show that Nekrasov pursued a tough policy towards revolutionary-minded students of Moscow University as vaguely mentioned by an anonymous author (1898). He hoped that Nekrasov, in his new capacity, will continue to educate young men in the spirit of duty towards God, Tsar and Fatherland.

At the turn of the 19th century Nekrasov began publishing unimaginably verbose and hardly understandable stochastic work,

some of it inseparably linked with religion and politics or simply wrong. We may believe that at least partly that was caused by his administrative work over a long period of time and lack of comprehension of the spirit of the time. Markov and Liapunov did not find any core of good sense in his new contributions and understood that scientific discussions with him became impossible. Moreover, Liapunov (1901, p. 63) concluded about his statements concerning limit that:

All his objections [to the work of most eminent scholars] *are based on various misunderstandings. Some of them are just unsubstantiated declarations*, […] *others either do not at all relate to the subject-matter of the criticized papers or are extremely vague.* […] *If Nekrasov states new objections of the same ilk, I will by right pass them over in silence.*

Markov (1912, p. 215) and Posse, see Nekrasov (1915a; 1915c), made similar statements.

Nekrasov revealed his political views in his unpublished letters of 1916 to P. A. Florensky. Here, for example, is an extract from his letter of 26 November:

The Moscow school advanced the principles of the language of Christian science and repulses the language in the style of Marx, Markov and Ya. A. Linzbach. A comparison of the books of Linzbach, Markov […] *with those of the main representatives of the Moscow philosophical-mathematical school* [Bugaev, Florensky and Nekrasov himself] *clearly shows the crossroads to which the German-Jewish culture and literature are pushing us.*

WWI had been going on, which only to a small extent exonerates Nekrasov. And the book of the completely unknown Linzbach (1916) which I found did not discuss either Marxism or religion.

Thus Nekrasov became a Black Hundreder, as Markov junior (1951, p. 610) stated without beating about the bush. But, to repeat, we should not forget his scientific achievements. Zhukovsky (1890/1949, p. 639) noted that Nekrasov had assisted him when discussing a mathematical problem. Seneta (1984, §§ 6 – 7) noted that Nekrasov had achieved interesting results in the theory of probability and definitely influenced Markov (1914, p. 106) who sometimes had referred to him without criticisms. Furthermore, Markov (1912, p. 215) sometimes considered the refutation of Nekrasov's mistakes as one of the aims of his work, and a similar statement is contained in one of his letters to Chuprov of 1910 (Ondar 1977/1981, p. 5). On the other hand, Nekrasov never heard about variance as Chuprov (letter to Bortkiewicz of 22 Nov. 1896; Bortkevich & Chuprov 2005) definitively found out in 1896, when Nekrasov *admitted* his candidate composition.

1.2. The teaching of probability theory in the school (Letter 1). In 1914, Nekrasov attempted to introduce probability in the curriculum of the secondary school. The discussion of a note by Florov, which Markov mentioned in his letter, was only Nekrasov's first step in this direction. He (1915b) published a very long paper containing the programme of a course in the theory of probability compiled by Florov, responses of many mathematicians and pertinent

materials of the Second All-Russian Conference of 1914 of mathematical teachers. His paper was actually a report about the postal discussion organized by the Ministry of People's Education, whereas Nekrasov only formulated his concluding reasoning in favour of the Florov programme. In particular, Nekrasov included passages from the answers received from Chuprov; in more detail, see Sheynin (2011, p. 32). There, I also note that Chuprov, in his answer of 1904 to the same question from the Free Economic Society remarked that badly prepared instructors will worsen the situation. Youshkevich (1968, p. 311) noted that B. K. Mlodzievsky (1868 – 1923) was the first to touch on this subject, but he had not said when exactly, nor did he know about the initiative of the Free economic Society.

Many participants of that postal discussion went beyond its framework and insisted on the inclusion of the elements of mathematical analysis and analytic geometry in the school curriculum although all of them understood the difficulties of any reorganisations carried out during the war.

As it seems, Markov was not invited to that discussion, but he stated his opinion in a paper (1915). On p. 33 he indicated:

The guiding idea of the Florov & Nekrasov project is […] *the need to acquaint the school students with their works*.

His own opinion (p. 20) was similar to those of Vasiliev and B. M. Koyalovich (Markov's student and an author of textbooks, on the theory of probability in particular, which continued to appear at least up to 1931). In principle, Vasiliev approved the Florov project and his and Koyalovich' objections were mostly directed to Florov's concrete programme. Vasiliev also noted that the theory of probability serves as a good illustration of the theory of combinations and improves logical thinking and that even Kraevich (1864) included a successful stochastic section in his collection of problems.

Markov thus positively regarded the Nekrasov idea although did not say it directly and neither had he offered his own programme. He also indicated that he did not approve the idea of including mathematical analysis and analytic geometry in the curriculum.

Nekrasov applied to the vice-president of the Petersburg Academy of Sciences (Nekrasov 1916b, pp. 55 and 58) with a request to consider the problem of the teaching of probability theory in school, see his letter to K. A. Andreev of 5 Dec. 1915 (Sheynin 1994).

On Markov's initiative (Markov junior 1951, p. 610) the Academy established a special commission for considering Florov's proposal and Markov became its member. The Commission (*Report* 1916, p. 79) agreed with his negative opinion and objected to Nekrasov's attempts

Aimed at a preconceived goal of transforming pure science into a tool bringing religious and political pressure to bear on the rising generation …

The Commission did not enter in essence into the teaching of probability but some of its members (p. 73) objected to its inclusion in principle. In itself, the theory of probability is certainly no ideological weapon at all but we may believe that the conclusions of the

Commission were conditioned by Nekrasov's well known ideological views (§ 1.1).

Nekrasov continued to defend his proposal which had never been realized. In particular, he (1916a, pp. 27 – 29) published the letter of B. V. Stankevich of the same year. He directed the Physical Institute of Moscow University where, as Stankevich wrote, the theory of probability is only *taught during the upper years and only in the mathematical faculty*. He feels himself *essentially restricted* […] *since the listeners are not acquainted with even elements of the theory of probability*. He cannot therefore fail to wish an *introduction of a shortened course of the theory of probability in the school curriculum*.

Stankevich also stated that in the 1880s the astronomer V. Ya. Tsinger (1842 – 1918) *usually* stated that the kinetic theory of gazes was *anarchic* and that in 1890s N.Ya. Sonin (1849 – 1915) *did not approve that theory but regarded it somewhat leniently*[1].

It seems that in 1914/1915 the theory of probability was not taught even in the mathematical faculty (Anonymous ca. 1916) but exactly for this reason Stankevich' opinion about the introduction of probability into the school was hardly convincing.

To the episode described above and to the history of his relations with Markov Nekrasov (1916b) devoted a special booklet. It is not yet studied and I only remark that he (pp. 56 – 62) printed the texts of six of Markov's letters or postcards of 1915 – 1916 to him and maintained that they contained (omitted by him) swear words and stated, not for the first time, that Markov had borrowed some of his results without due references.

I also add that in 1898 – 1889 Nekrasov (Sheynin 1995) applied to A. I. Chuprov (father of A. A. Chuprov, professor at Moscow University and corresponding member of the Petersburg Academy) with a proposal to introduce the theory of probability into the curriculum of the law faculty. His proposal was not realized.

1.3. Seminarians (Letter 2). A considerable fraction of the graduates of theological seminaries (including Nekrasov) entered universities. Thus (Anonymous 1876, p. 90), in 1875 45.7% of the students of the faculty of natural sciences in Petersburg were seminarians, as well as 11.5% in the faculty of mathematical sciences. On the whole, seminarians comprised 29.2% of students of that university. On p. 93 of that source we read:

Both with regard to their training and education, the seminarians were much lower than the graduates of modern gymnasiums.

Nekrasov was an outstanding exception!

In 1875, 53.3% of the students of all the Russian universities had graduated from gymnasiums and the seminarians comprised the *main body* of the other students. I have no similar data about the end of the 19th century, but in the beginning of the 20th century the secular education in theological seminaries undoubtedly worsened. Anonymous (1911, p. 3) reported that the Most Holy Synod, *in discharging the royal will*, developed new regulations for the theological educational institutions *in the ecclesiastic direction*. Before these new regulations were realized (Nikolsky; no date provided, p. 209),

Many scientists left theological academies and the appointment as professors such people, who were absolutely unknown in science but sufficiently known in the sphere of ecclesiastic-political struggles, became possible.

We may think that a similar situation took place in the seminaries as well and that therefore Markov's opinion about the need to deprive the graduates of seminaries of any preferences to the graduates of non-classical schools was quite justified.

1.4. Infinitesimals (Letter 3). An infinitesimal is a variable whose limit is zero. Referring to this definition due to Cauchy, Markov added that he did not ascribe zero to the values of an infinitesimal. True, the context of his letter proved that he had not attached great importance to his restriction[2].

Nekrasov (1912, p. 459) maintained that Markov (1912, pp. 11 – 12) *apparently restores* Euler's terminology. Euler (Youshkevich 1972, p. 267) is known to have considered an infinitesimal equal to zero. Youshkevich traced the development of the notion of infinitesimal in the 18th century, but, since the difference between actual and potential infinitesimals did not influence Markov's scientific work, I leave this point aside. Note, however, that Markov (1912) had not touched on infinitesimals and that this paper lacks pages 11 – 12. But Markov (pp. 223 – 224) quoted Nekrasov (1901, pp. 236 – 237) who had stated that a variable magnitude P should be considered identical to its limit L, i. e., that P/L should tend to unity. Nekrasov generalized this simple statement on the case of a variable L (?) and somehow accused Chebyshev, Liapunov and Markov of their opinion that condition $(P – L)$ 0 is sufficient for such L to be considered the limit of P. Therefore, Nekrasov concluded, this time quite logically, that his opponents had believed that, for example, if $n > 0$ and x 0 x^n has sin$x$ as its limit!

In his Letter Markov indicated that Nekrasov had attempted to *direct the instruction in the secondary school to a wrong track*. Posse (1915, p. 72) concurred:

Nekrasov attempts to discredit among the teachers and school students of the secondary school an entire school of mathematicians whose representatives are all the Petrograd[3] (and not only Petrograd) professors and to direct the instruction of mathematics in the secondary school on a wrong track. I consider it impossible to pass over his attempt in silence since it is published in the official organ of the Ministry of People's Education.

2. The letters

Letter 1, 30 Jan. 1914, p. 4
A question for the Ministry of People's Education
I have recently found out that the Ministry […] is studying the problem of the introduction of the elements of probability theory into the curriculum of secondary schools. I even have in my hands a typed note of P. S. Florov [see § 1.2] about this problem with comments made by P. A. Nekrasov [see § 1.1], member of the Council of the Minister of People's Education.

Concerning the contents of this note I only say now that it is very questionable and cannot be a basis for implementing that proposal. If necessary, I will analyse it in detail [Markov 1915].

Now, however, I ought to indicate that the Ministry had not yet asked a representative of the "theory of probability"[4] in the Petersburg University, i. e., had not asked me, to offer a judgement about this subject. I think that serious problems are not solved without the participation of appropriate specialists if only not deliberately intending to solve them anyhow and badly. Therefore, I ought to ask the Ministry, does it really seriously study the problem of teaching the "theory of probability" in secondary schools or considers it as a pastime for the now idle[5] professor Nekrasov?

Academician A. Markov

Letter 2, 11 Aug. 1915, p. 3

Seminarians and graduates of non-classical gymnasiums

Newspapers have reported that seminarians may enter physical-mathematical faculties of universities without holding special examinations whereas the same problem concerning graduates of non-classical gymnasiums remains open.

It is difficult to agree that this situation is normal. There is no sharp difference between the schools[6]. Latin, which distinguishes the former from the latter, is not necessary for education in physics or mathematics. However, the upbringing of the seminarians trains them for a special kind of reasoning. They must subordinate their minds to indications of the holy fathers[7] and replace them by the texts of the Holy Writ.

The seminarians' wisdom can be very deep [Florensky 1914][8], *but it is remote from the science of reality and can only state religious truth for the believers. Such wisdom easily leads to the desire for subordination of science to* the religious-scientific-political experience under the instructive emblem. I refer readers to [Nekrasov 1915b][9] *where they will find that verbiage as well as specimens of special wisdom. I cannot therefore abstain from expressing serious doubts about the suitability of the seminarians for the physical-mathematical faculties. In any case they should not be preferred to the graduates of non-classical gymnasiums.*

Academician A. Markov

Letter 3, 28 Oct. 1915, p. 3

A letter to the editorial office

Dear Editor, Sir, Allow me to raise two questions for lawyers by means of your newspaper. What measures can be taken, if hoping for success, against the abuse of the press, when perpetrated by an official organ? Does the editor of such a source represent the same responsible figure as the editors of other organs of the press, or is his entire responsibility restricted to an unconditional execution of the will of the authorities?

Then, since the Journal of the Ministry of People's Education has corrupted the facts, I am most humbly asking you to put up the following letter which I had sent to Mr. Radlov but which he refused

to publish. I ought to remark that my request touches on a problem of public interest since Nekrasov, a member of the Council of the Minister of People's Education, in profiting by his influence is attempting to direct the mathematical instruction in the secondary school to a wrong track.

Dear Sir, Ernest L'vovich, You have published two polemic articles of Nekrasov [1915a; 1915c] *which directly concern me, and I am most humbly asking you to find a place in your journal* [Zhurnal Ministerstva Narodnogo Prosveshchenia, Journal of the Ministry of People's Education] *for the following explanation:*

In the September issue, Professor K. A. Posse (1915) *indicated some peculiar traits of Nekrasov's style[10]. Without denying their existence, Nekrasov, in the October issue, mentions that I had infected his style[11]. This statement is untrue since such a style is absolutely alien for me just as I am alien to him in mathematical problems. In particular, I consider it reasonable to pass over in silence everything that does not touch on the business at hand, and I do not permit myself any confusions of mathematics with politics or religion.*

Then, Nekrasov repeatedly refers to scientific societies. In the July issue he[12] says that I had submitted the principles of the theory of limit to the judgement of scientific societies and that our polemic was done away with by them. And in the October issue he maintains that his definition [1912, p. 459] *had passed the crucible of a scientific society so that the reader can be sure of its truth[13].*

These references are inaccurate. I have never submitted those principles to the judgement of scientific societies. And I ought to say that I have not introduced anything new in those principles. In my lectures on differential calculus I defined infinitesimals just as Cauchy had done it: [Markov cites Cauchy (*Analyse algébrique*, 1821) in his original French: an infinitesimal converges to zero] *and just as many other foreign and Russian textbooks* (*for example the mentioned book of Posse*) *define them[14]. I only add: it is important to remark that we do not ascribe zero, the limit of an infinitesimal, to its value[15]. These words do not make me solidary with Nekrasov, but in this case they are decisive.*

I will not repeat or explain what I had earlier said about Nekrasov's discoveries[16], and I am concluding my letter by indicating the resolutions of scientific societies. At the beginning of Nekrasov's paper [1912] *we read*:

Editorial note. This paper is Nekrasov's answer to Markov [1912] *and it is published according to the decision of the Moscow Mathematical Society: to publish in* Matematichesky Sbornik *one paper from each of the two authors on the problem under consideration.*

And the minutes of the sitting of 19 Jan. 1914 of the Kharkov Mathematical Society contains the following:

The letter of Academician Markov about his polemic with Nekrasov was heard out and discussed. The Society, being guarded by the principle that both sides of a polemic ought to be placed, as far as

possible, in an equal situation, resolved to publish Markov's answer if he so desires[17].

Be assured etc.

9 October 1915 A. Markov

For the readers of the newspaper <u>Den'</u> to be able to form some idea about Nekrasov's style, I consider it not superfluous to quote an ending of one of his phrases [1915a, p. 16] *partly provided in my previous Letter:*

Then there only remains the religious-scientific-political experience under the instructive emblem similar to that which is contained in the <u>Arithmetic</u> published by the command of Peter the Great about 1700[18].

Be assured etc.

24 October 1915 A. Markov[19]

Notes

1. For some time, Poincaré (Sheynin 2009, § 11.2) was also dissatisfied with that theory.

2. See Note 15.

3. See [x, Note 15].

4. *Theory of probability* twice appeared in inverted commas. Evidently, it was the Editor who introduced them since (evidently) he thought that that term was insufficiently known.

5. Nekrasov was not idle at all (§ 1.1).

6. Apparently: between the *schools*, seminaries and non-classical gymnasiums.

7. *Holy Father* is the appellation of the Popes, but certainly not used by Orthodox Christians.

8. Florensky (1914) contains a Supplement of a natural-scientific and mathematical essence complete with a list of a few hundred appropriate references. The *Istoriko-Matematich. Issledovania* carried two publications concerning Florensky (Demidov et al 1989; Luzin & Florensky 1989).

9. Markov referred to that source somewhat wrongly.

10. Posse (1915, p. 71) remarked that Nekrasov

Likes to strike his opponents by phrases, which seem very thoughtful, but actually are extremely vague […] *and* […] *while quoting his opponents, sometimes changes their words and ascribes them something that they nowhere and never said.*

11. Nekrasov (1915c, p. 97):

It was the dispute of my opponent [with me] *who challenged me to a dispute and whom I have duly answered becoming infected by his <u>polemic style</u>.*

12. Nekrasov (1915a, p. 12): *Here Markov* [1915, pp. 27 – 28] *had provided as proof the principles of the theory of limit, which he had already submitted to the judgement of scientific societies and got there my proper answer* [Nekrasov 1912, pp. 223 – 224 and 459]. See § 1.4. Here, I add that no judgement of scientific societies is mentioned in those papers.

13. Nekrasov (1915c, p. 101):

My definition [1912, p. 459] *is not to the taste of Posse, but it passed the crucible of the judgement of a scientific society, and the reader can be sure of its truth.*

He never explained what he had meant by *truth of definition*.

14. Markov apparently referred to Posse (1903), but he had not previously mentioned any book by that author.

15. Markov refers to the edition of 1898 of his lectures which I have not seen. Here, however, is the definition from another of its editions (1898/1901, p. 45):

A variable approaching its limit zero is called an infinitely small number [!]. *It is important to note* …

The Report (1916, p. 72), one of whose authors was Markov, explained that the restriction about zero was made for the sake of convenience.

16. Markov (1910):

I never confirmed any discovery of Nekrasov and am unable to confirm them.

17. No such answer was published. I have seen and published the archival correspondence of Markov with Radlov (**S, G,** 16). Radlov and his staff thought that the opinion of Nekrasov did not warrant any detailed discussion and that Markov and Nekrasov had debated under the same conditions.

18. Nekrasov referred to Magnitsky (1703). An emblem (let it be *instructive*) is described by Gnedenko (1946, p. 57). Nekrasov maintained that mathematics, restricted by deduction, which meant, according to the context, lacking probability, was only the *religious-scientific-political experience*. It is hardly possible to understand him, but anyway it is mathematical statistics rather than probability that is connected with induction.

19. P. S. Youshkevich, father of A. P. Y., see Supplement, summarized the polemic of Markov and Nekrasov. Cf. Sheynin (1989).

Bibliography

Abbreviation: IMI = *Istoriko-Matematicheskie Issledovania*
ZMNP = *Zhurnal Ministerstva Narodnogo Prosveshchenia*,
section Sovremennaia Letopis'

Anonymous (1876), On the number of graduates of gymnasiums and of those who entered universities in 1875. ZMNP, vol. 183, No. 2, pp. 88 – 93.

Anonymous (1898, Russian), The new civil functionary responsible for the Moscow educational region [Nekrasov]. Newspaper *Moskovskie Vedomosti*, 13 (25) March, p. 2.

Anonymous (1911), *Zhurnaly Osoboi Kommissii Sinoda dlia Vyrabotki Proekta Ustavov Dukhovnych Sredne-Uchebnykh Zavedeniy* (The Journals of the Special Commission of the Synod for Drawing Up Draft Regulations for Theological Secondary Educational Facilities). Petersburg.

Anonymous (ca. 1916), *Obozrenie Prepodavania na Fiziko-Matematicheskom Fakul'tete Mosk. Universiteta v 1914/1915* (Review of the Teaching at the Physical-Mathematical Faculty of Moscow Univ.). No place, no year. Title page lacking.

Bortkevich V. I., Churpov A. A. (2005), *Perepiska* (Correspondence). Berlin. **S, G,** 9.

Demidov S. S., Parshin A. N., Polovinkin S. M. (1989, Russian), on the correspondence of N. N. Lusin and P. A. Florensky. IMI, vol. 31, pp. 116 – 125.

Florensky P. A. (1914), *Stolp i Utverzhdenie Istiny* (Pillar and the Assertion of Truth). Moscow. Its translation (Princeton, 2004) provides a hardly correct title: *Pillar and the Ground of Truth*.

--- (1986, Russian), Introduction to the discourse *The Idea of Intermittence As an Element of Weltanschauung*. Publication and comments by S. S. Demidov & A. N. Parshin. IMI, vol. 30, pp. 159 – 177.

Gnedenko B. V. (1946), *Ocherki po Istorii Matematiki v Rossii* (Essays on the History of Math. in Russia). Moscow – Leningrad.

Istoria (1955), *Istoria Moskovskogo Universiteta* (The History of Moscow Univ.), vol. 1. Moscow.

Kraevich K. D. (1864, 1867, 1874, 1882), *Sobranie Algebraicheskikh Zadach* (*Coll. of Algebraic Problems*). Petersburg.

Lusin N. N., Florensky P. A. (1989, Russian), Correspondence. IMI, vol. 31, pp. 125 – 191.

Liapunov A. M. (1901, Russian), The answer to P. A. Nekrasov. *Zapiski Kharkovsky Univ.*, vol. 3, pp. 51 – 63.

Linzbach Ya. (1916), *Prinzipy Filosofskogo Yazyka* (The Principles of the Philosophical Language). Petrograd.

Magnitsky L. F. (1703), *Arifmetika*. Moscow.

Markov A. A. junior (1951, Russian), Biography of A. A. Markov senior. In Markov senior (1951, pp. 599 – 613).

Markov A. A., senior (1898), *Differenzial'noe Ischislenie* (Differential Calculus). Petersburg, 1901. Mimeographed editions.

--- (1910, Russian), Correction of an inaccuracy. *Izvestia Akademii Nauk*, ser. 6, vol. 4, No. 5, p. 346.

--- (1914, Russian), On posterior probability. *Soobshchenia Kharkovsk. Matematich. Obshchestva*, ser. 2, vol. 14, No. 3, pp. 105 – 112.

--- (1915, Russian), On the Florov and Nekrasov project. ZMNP, vol. 57, No. 5, pp. 26 – 34.

--- (1951), *Izbrannye Trudy* (Coll. Works). No place.

Minkovsky V. L. (1952, Russian), Pedagogic ideas and the work of Markov. *Matematika v Shkole*, No. 5, pp. 10 – 15.

Nekrasov P. A. (1888), *Teoria Veroiatnostei* (The Theory of Probability). Later editions: 1894, 1896, 1912 (Petersburg). The first two editions were mimeographed.

--- (1901, Russian), Concerning a simplest theorem on the probabilities of sums and means. *Matematich. Sbornik*, vol. 22, pp. 225 – 238.

--- (1912, Russian), The main general method of generating functions. Ibidem, vol. 28, pp. 351 – 460.

--- (1915a, Russian), On Markov's paper (1915). ZMNP, vol. 55, No. 7, pp. 1 – 17.

--- (1915b, Russian), The theory of probability and mathematics in the secondary school. ZMNP, vol. 55, No. 2, pp. 65 – 127; vol. 56, No. 3, pp. 1 – 43; No. 4, pp. 94 – 125.

--- (1915c, Russian), An answer to the objections of K. A. Posse. ZMNP, vol. 55, No. 10, pp. 97 – 104.

--- (1916a), *Prinzip Ekvivalentnosti Velichin v Teorii Predelov i Posledovatel'nom Priblizheonnom Ischislenii* (The Principle of Equivalence of Magnitudes in the Theory of Limit and in the Calculus of Successive Approximations). Petrograd.

--- (1916b), *Srednia Shkola, Matematika i Nauchnaia Podgotovka Uchitelei* (The Secondary School, Mathematics and the Scientific Training of the Teachers). Petersburg.

Neyman J. (1978), Review of Ondar (1977). *Hist. Math.* vol. 5, pp. 485 – 486.

--- (1981), Introduction to Ondar (1977/1981).

Nikolsky N. (no date available, Russian), Secondary theological institutions. *Enc. Dict. Granat*, vol. 19, pp. 202 – 209.

Posse A. K. (1903), *Kurs Differenzial'nogo i Integralnogo Ischislenia* (Course of Differential and Integral Calculus). Petersburg.

--- (1915, Russian), A few words about Nekrasov's paper (1915a). ZMNP, vol. 59, No. 8, pp. 1 – 17.

Report (1916, Russian), Report of the Commission [established] to discuss some issues concerning the teaching of mathematics in secondary school. *Izvestia Akademii Nauk*, vol. 16, No. 2, pp. 66 – 80. **S, G,** 4.

Seneta E. (1984), The central limit theorem and linear least squares in pre-revolutionary Russia. *Math. Scientist*, vol. 9, pp. 37 – 77.

Sheynin O. B. [O.] (1989, Russian), A. M. Liapunov's letters to K. A. Andreev. IMI, vol. 35, pp. 124 – 147.

--- (1990, Russian), *Alexandr A. Chuprov. Life, Work, Correspondence*. V&R, 2011.

--- (1994, Russian), Correspondence of P. A. Nekrasov and K. A. Andreev. IMI, vol. 35, pp. 124 – 147. Co-author, M. V. Chirikov.

--- (1995, Russian), Correspondence of P. A. Nekrasov and A. I. Chuprov. IMI, vol. 36-1, No. 1, pp. 159 – 167.

--- (2003), Nekrasov's work on the central limit theorem. The background. *Arch. Hist. Ex. Sci.*, vol. 57, pp. 337 – 353.

--- (2009), *Theory of Probability. Historical Essay*. Berlin. **S, G,** 10.

Soobshchenia (1915, Russian), *Soobshchenia Kharkovskogo Matematicheskogo Obshchestva* (Communications of the Kharkov Math. Soc.), ser. 2, vol. 14, No. 6, p. 8 of second paging.

Youshkevich A. P. (1968), *Istoria Matematiki v Rossii do 1917-go Goda* (History of Math. in Russia before 1917). Moscow.

--- (1972), Differential and integral calculus. In *Istoria Matematiki s Drevneishikh Vremien do Nachala 19-go Stoletia*, vol. 3. Editor, A. P. Youshkevich. Moscow, pp. 241 – 368.

Zhukovsky N. E. (1890, Russian), On the shape of ships. In author's *Sobranie Sochinenii*, vol. 2, pp. 627 – 639. Moscow – Leningrad, 1949.

P. Youshkevich

On a scientific polemic

For the few latest months, the not very numerous readers of the *Journal of the Ministry of People's Education* had been able to enjoy the spectacle of a protracted scientific polemic struck up between Professor Nekrasov on one side with Academician Markov and Professor Posse on the other side. It was caused by Markov's paper about the project of teaching probability theory in the secondary school which was defended by Nekrasov. However, owing to the latter's efforts, the polemic soon left the confines of a pure pedagogic theme and developed into a polemic on the priority of discovering a certain theorem from the theory of probability, on the understanding the theory of limit and philosophical monism or dualism etc.

Even in this wider framework the scientific polemic could have certainly remained a source for special periodicals if only Professor Nekrasov were not its central figure. He possesses the secret of converting most abstract mathematical problems into very topical issues. Indeed, he is not only a mathematician, but also a great lover of philosophy, of [boring] philosophizing. The philosophizing of this respected scientist is, however, absolutely special. It is some strangest mixture of the thoughtfulness of Gogol's Kifa Mokievich and the annoying nonsense of Shchedrin's Little Judas who became governed by his own ravings. I adduce a typical excerpt from his book (1912) for the readers to appraise his philosophical manner.

Nekrasov is discussing the continuity of historical traditions achieved by various intermediate links, for example, by the link *children* in the chain the *dying out generation of fathers – children – the following generation*. Nekrasov, however, thus describes this simple idea:

In all the series of formulas of historical turnover[1] we need an included intermediate term, an intermediate combined third element. It tears both the clarity of a glance at the future and the definiteness of judgement. In our life, these insertions, combinatorial triangulations[2] and ruptures of the definiteness of judgement call forth the need for the inductive method and a symbolic connection of phenomena; call forth the need for symbolic thinking, and parallel to it, for material symbolic mnemonic (recalling) means of tying and untying the knots of history and of establishing historical legal associations between mental and material possession, associations by means of symbolized (tentative) truths, i. e. of legally active probabilities. A historical genetic and symbolic triangulation is worth the troubles just as a geodetic triangulation is.

Historical genetic and symbolic triangulation; material symbolic mnemonic means; historical legal association! It is easy to imagine

how sweetly Nekrasov himself should have heard all those *combinatorial triangulations* of magnificent showy words. To repeat, this excerpt is not something exceptional in his scribbles. Dozens of his pages are full of such mania-like eruption of speech. We learn from that philosopher-mathematician about *vibrational movement* of *the atmosphere of nominal financial currency* (1888/1912, p. VI), of *formal-analogous meteoric notions* (sweet water and money currency) (p. VII) and about *money water of the debtor and creditor kind which is sometimes charged up with electromagnetic anodic and cathodic means* (*allegorically expressed*), about *remarkable measuring graphic vibration of trust which leads from a nominal to material possession of things and which originates under the influence of facts materializing mental nominal possession of things* (p. XIX) and so on without end.

All this nonsense abounds in charming references to Kant, Hegel, Herbart, Clifford, Hoffding, Vvedensky, Lapshin, Khomyakov et al who are certainly not guilty at all of their duty of being an honorary escort for Nekrasov's outrageous gibberish. However, it has its own method, as in Hamlet's insanity. He is insistently striving for strengthening the religion and social tradition. In Nekrasov's theory of probability one perceives a means for establishing the thought about the existence of the free will (and a Supreme Being) along with determinism. The same role, as he imagines, is played by intermittent magnitudes. Many eminent scholars, for example, Maxwell and Boussinesq, believed in the same prejudice, in that intermittent mathematical magnitudes can be as though prototypes of free action.

It seems that that ideas had a decisive influence on Nekrasov's concept about the two types of differentials. Thus, potential differentials correspond to the change of continuous variables and are able to reach the absolute zero[3], and as an illustration Nekrasov refers to the Zeno of Elea paradox about Achilles and the tortoise. Actual differentials, however, unboundedly tend to zero but never reach it; illustration: the length of all the sides of a regular inscribed polygon, when the number of those sides increases to infinity.

And when his opponents do not agree with this philosophy of differential calculus, he calls their doctrine monistic and declares that it *excludes from science combinatorial moral values of the so-called parallel* (*dual*) *Weltanschauung, see Chelpanov* (1905), *and directly leads to the monism of Haeckel's <u>world mysteries.</u> Will Posse defend the viewpoint of these <u>mysteries</u>*? (Nekrasov 1915a, p. 102).

Without waiting for Posse's answer to that spiteful question, Nekrasov himself, being the author of the book (1912), explains why Haeckel's viewpoint should not be defended:

In the very embryo of his theory of knowledge Haeckel's monism kills the notion about the unities of the highest order which mathematics is teaching, a science which does not desire, in its definitions, to betray the true classical humanistic foundations that directly opposes what is called barbarism, cannibalism, original sin against which the civil science and the Christian civilization[4] are struggling with the sole aim of improving the human nature (*Be perfect as your heavenly father is* [Mathew 5:48]).

It goes without saying that mathematics, the real, *free* mathematics, as Cantor wonderfully called it, rather than a mathematics at beck and call of theology, cannot either betray the true classical humanistic foundations (here also Nekrasov was certainly unable to get rid of the *triangulation* of words) or serve them [serve it]. It, just as any other creative thought, follows its own path, has its own internal criteria of truth and unthruth. And only by following its own path, only remaining an end of itself, it becomes one of the most powerful means of civilization.

The abstract debate about actual and potential differentials thus conceals a rather topical problem of our time, of the struggle for a free mathematics against mathematics which is called upon to become a servant of theology and politics. A number of outstanding mathematicians have repulsed Nekrasov which shows that we should not be afraid of the outcome of that struggle in the scientific milieu. However, in such cases there exists a possibility of acting in an extra-scientific way as well. The hints and nods about Haeckel's monism which apparently conceals in itself *barbarism, cannibalism, original sin*, are perhaps a promising beginning of a campaign for such an extra-scientific propagation of the dualistic theory of infinitesimals.

Notes

1. *Turnover* is an economic term. Nekrasov evidently wished to insert as many such terms as possible but thus obviously spoiled his descriptions.

2. Nekrasov repeatedly applied this term (once even *geodetic triangulation*), no doubt having picked it up in the Moscow Land Surveying Institute where, in 1885 – 1891, he doubled as professor of mathematics and probability theory.

3. Absolute zero is, and likely was in 1915, a definite physical term. It should have not been loosely applied.

4. Nekrasov would have never accepted the (at least nowadays) generally recognized term, *Judeo-Christian civilization*.

Brief Information about Those Mentioned

Kifa Mokievich: a minor character in Gogol's *Dead Souls.*

Little Judas: a central figure of Saltykov-Shchedrin's *The Golovyov Family* (1880, Russian; New York, 2001).

Boussinnesq Joseph Valentin, 1842 – 1929. Mechanician

Haeckel Ernst, 1834 – 1919. Biologist, naturalist, philosopher

Herbart JF, 1776 – 1841. Philosopher, psychologist, father of pedagogy as an academic discipline

Hoffding Harold, 1843 – 1931. Philosopher, theologian

Khomyakov Aleksei Stepanovich, 1804 – 1860. Theologian, philosopher, co-founder of the Slavophile movement

Lapshin Ivan Ivanovich, 1870 – 1952. Philosopher

Vvedensky Aleksandr Ivanovich, 1856 – 1925. Philosopher, psychologist

Bibliography

Nekrasov P. A. (1912), *Vera, Znanie, Opyt* (Belief, Knowledge, Experience). Petersburg.

Chelpanov G. I. (1905), *Vvedenie v Filosofiu* (Introduction to philosophy). Seventh edition, Kiev, 1918.

X

S. Ya. Grodzensky

Andrei Andreevich Markov (excerpts). Moscow, 1987

Introduction by Translator

I translate excerpts from Markov's letters which Grodzensky had discovered in various archives. They are indeed interesting and hardly known abroad. Regrettably, Grodzensky had not examined which of those sent to newspapers were actually published.

I include his short explanations and add my own comments in small print. I (2007) had previously described some of those letters which had dwelt on Markov's unswerving struggle against anti-Semitism and return now to this subject as well.

Pp. 46 – 47. In 1908, because of the students' unrest, the Ministry of public education required professors to pay attention to the political attitude of their mind. On 27 September Markov wrote a letter to the newspaper *Rech*:

I found out that the known circular of the Ministry […] *is being sent to the entire teaching staff of our* [Petersburg] *university.* […] *All of them are considered to be agents of the government.* […] *This circumstance compels me to declare,* […] *that I have always thought to be only a professor.* […] *I can by no means take upon myself the extremely burdensome and absolutely unsuitable role of a government agent.*

On 2 October Markov sent a similar letter to the rector of the university.

P. 48. In December 1910, the Council of Ministers resolved that educational institutes ought to expel students who had participated in disturbances. On 10 December Markov wrote a letter to the physical and mathematical faculty of the university:

[…] *I consider it my duty to declare immediately that under such circumstances I cannot read any lectures.* […]

The same day he wrote to his friend, V. A. Steklov [the future president of the Academy of Sciences]:

I resolved to quit since I have no other means of expressing my sympathy for the students, or, more precisely, my exasperation. […]

His resignation was not accepted.

Pp. 49 – 50. In 1913, the university conferred the statute of honourable members of the University on several professors including Markov. Here is what Steklov read out about Markov in a testimonial signed by him and five other mathematicians:

Academician Markov, an alumni of our university, later a privat-dozent[1] and professor, has served for 25 years for the benefit of science and our university. He only quit after becoming professor emeritus, but is still rendering an important service as an academician by continuing to read a course in the theory of probability and some other courses.

The entire professorial and purely scientific work of Markov, a student and follower of our celebrated Chebyshev, is going on under the eyes of our university, and his high scientific authority is known to the entire educated Russia and abroad. It is therefore unnecessary to describe in detail Markov's scientific deserts, suffice it only to list the most essential of them.

Generally known are his studies of the theory of functions least deviating from zero whose creator was Chebyshev. His application of the theory of continued fractions to various problems of analysis is known to everyone and promptly made him known to the entire scientific world. Markov's numerous and original inquiries into the limiting values of definite integrals, a subject closely linked with his mentioned work on the theory of continued fractions, merit special attention.

I also mention simple and original methods which Markov developed while solving some special problems about maximal and minimal values issuing in the restricted area of variational analysis. There recently appeared a series of his remarkable investigations of the calculus of probability, of the subject with which Markov has been occupied for a long time being the best specialist in this field.

His pertinent studies are remarkably interesting not only for mathematicians, but for all other scientists who are engaged in statistical or economic problems. I indicate Markov's examination of the method of least squares[2], of the law of large numbers (LLN) *whose bicentenary the Academy of Sciences will celebrate on 1 December* [see below].

Finally, Markov had recently developed the theory of the probabilities of events connected into a chain. His results here are the most important among everything achieved after Chebyshev in the area of the celebrated LLN.

Apart from these investigations of outstanding merit with respect to the methods applied, the results obtained and the rigour of the analysis applied, whose general features I described, Markov compiled courses on the <u>Calculus of Finite Differences</u> *and* <u>Calculus of Probability</u> *which ran into two and three editions respectively. Both were translated into German and each represents a remarkable phenomenon in our scientific literature.*

It is even inconvenient to call them courses since they are original tracts mostly compiled from Markov's own studies[3]. They are now reference books for each mathematician as well as for those who work in related fields (for example, in statistics).

P. 50. The Minister did not confirm that Markov's status since he protested against the action of the government.

P. 62. In January 1917, at the general meting of the Academy of Sciences, Markov expressed his opinion about the prizes awarded by the Academy:

The prizes are a real disaster for the Academy. Their number is incessantly increasing and nowadays there are some 50 of them. However, the number of those which undoubtedly deserve to be highly marked by the Academy remains very restricted. Most contributions submitted for awarding can be separated in two groups: obviously

unsatisfactory and mediocre. Those of the first group are annoying but rather harmless since they do not require any detailed consideration. The other ones vainly take much time for their appraisal, and, on the other hand, humiliate the Academy since such works, in the absence of better ones, are sometimes awarded when being regarded too leniently.

It is opportune to recall the opinion of my unforgettable mentor, Chebyshev. He thought that the best way to check the development of science is to gather all the outstanding scientists and charge them with considering the contributions of other scientists[4].

Pp. 62 – 63. For Markov, authoritative was only what he considered true, and even Chebyshev did not escape his criticism. Already in 1891 he wrote to academician Karpinsky:

My colleagues had been certainly unable to regard my works utterly disapprovingly, but they praised me insufficiently. This is proved, incidentally, by their recommendation in connection with Buniakovsky's death. They clearly proposed to fill the vacancy by someone else[5]. *[…] In addition, you have probably never heard from them that in 1884 I had provided a complete review and a suitable proof of Chebyshev's problem*[6], *but that in 1885 he published his solution without proof and without mentioning me and continues in the same vein.*

P. 63. We read in the minutes of the physical and mathematical class of the Academy for 23 March 1921 that

Academician Markov reported that May 26 will be the centenary of the birthday of the unforgettable Chebyshev. Under normal circumstances it would have been necessary to mark this occasion by a special solemn meeting, but under the ugly form of the course of our present life any celebrations are hardly appropriate.

If our class decides that a solemn meeting is necessary, one of the eldest students of Chebyshev, Alexandr Vasilievich Vasiliev[7], *ought to be drawn in.*

Pp. 63 – 64. Markov appraised the work of his colleagues at its true worth. Here, for example, is what he wrote about E. V. Borisov with whom he had competed during his student years:

Borisov is well known to me since he we had at the same time attended lectures at the Petersburg University read by Chebyshev, *[…] and other professors. He undoubtedly belonged to their best students. Not without reason he was awarded a silver prize for a discourse on a subject proposed by the faculty: <u>Integration of differential equations by means of continued fractions</u>. Among those submitted for consideration was my own, but then, as now also, I have had to recognize that Borisov's work was the soundest among all the submitted. It revealed both his diligence and great knowledge.* [Nevertheless, the gold medal was awarded to Markov [xi, § 1].]

Borisov had not published much which does not prevent him from being a very serious scientist. An abundance of publications far from always testifies to a thorough scientific schooling. […]

Pp. 64 – 65. Sometimes Markov changed his opinion about a scientist. Grodzensky cites two of Markov letters to Chuprov dated 6 Nov. 1910 and 1 Dec. 1912, both concerning Karl Pearson, see Ondar (1977/1981, pp. 5 and 60). Grodzensky's second example (p. 65) pertains to Golitsin. At first, Markov

sharply criticized him for an unworthy discussion of some experiments [vi, vii], but later, in 1908, co-signed a favourable testimony about his subsequent work.

P. 80. On 8 Jan. 1905, Markov wrote a letter to the academy. He considered his duty

To remind the General Meeting about the unparalleled case of violating the law: *the honorary academician Peshkov* [Gorky] *even to this day is* […] *deprived of the possibility of enjoying the rights of an honorary academician.*

The cessation of his election to honorary membership was certainly announced in the newspapers as though decided by the Academy, but we know that this announcement is false. Such announcements can only be valid under the reign of unrestricted arbitrary rule.

P. 81. Being exasperated by the vice-president, Markov wrote to [the academician] Lappo-Danilevsky on 12 Oct. 1905:

Today, the liberal permanent secretary together with the vice-president deeply disgusted me by their insistent desire to correct my statement as printed in the proofs of the minutes of the special sitting of the General Meeting of 28 September. They desired that I will not call disgusting a fact which exasperated me not less than their insistent wish. And so, it is desirable to deny even an academician the right to speak out. […]

P. 88. On 22 January (4 February) 1913 the newspaper *Rech* published Markov's appeal:

To the representatives of science and its admirers:

In 1713, there appeared a posthumous work of Jakob Bernoulli, <u>Ars Conjectandi</u>, in which his celebrated theorem, later developed into the LLN, was established for the first time. This theorem directly belongs to mathematics, and more specifically to the theory of probability. However, as Jakob Bernoulli himself had remarked, it has and ought to have numerous applications in each science and to practical problems in which statistical methods have to be used. Believing that this bicentennial jubilee ought to be somehow solemnly marked, I appeal to all those who sympathetically regards this idea to help me to realize it[8].

Pp. 89 – 90. Soon after WWI had broken out the president of the Academy, being frightened by the possibility of *an expression of Royal displeasure*, reminded the vice-president about the need to consider the expulsion from the Academy of the citizens of the states being at war with Russia.

On 30 January 1915 Markov compiled a draft of an answer to the president on behalf of the Academy. In particular, he wrote:

The Academy's connections with its honorary members and corresponding members, citizens of Germany and Austria [Austro-Hungary]*, are interrupted by the war. So the question is, should not the Academy immediately consolidate this situation by making it permanent, i. e., by expelling them or deciding that they left us. For answering that question the Academy may only guide itself by the indications of the past, by the examples of foreign academies and its own duties, but certainly not by the fear of inadmissible accusations of lack of patriotism.*

The Academy has endured many wars, and, if expelling its honorary members, citizens of the countries at war with Russia, it would have

applied this rule to many outstanding scientists: to Laplace (honorary member, 1802 – 1827 [Laplace died in 1827]), *Cuvier* (*1802 – 1832*), [John] *Herschel* (*1826 – 1871*), *Faraday* (*1830 – 1867*) *et al and would have suffered a considerable loss. The examples of foreign academies are also important for the present. The Paris academy had not expelled its German members, nor did the Berlin academy expel its French or Russian members.* […]

Without disturbing its duties, the Academy cannot break the existing ties with other academies or scientific societies. Therefore, the expulsion of individual scientists or even one of them is also impossible, it will not benefit either Russia or itself, but it can involve sudden and undesirable consequences: an interruption of connections with entire societies or academies and failure of the begun international scientific enterprises.

In this particular case, the question is somewhat complicated by the statements of some German scientists[9], but the Academy cannot delve on them. They do not honour those scientists and are explained by military intoxication or psychosis. The intoxication will pass once the war is over. Then, and only then, the Academy will possibly have to admit that our connection with one or another person is definitively interrupted, but such occasions will be rare.

After all, the Academy was nevertheless obliged to expel those honorary members and corresponding members.

P. 94. In 1904, the imminent defeat in the war with Japan became obvious, and the Petersburg newspaper *Novoe Vremia* published an article whose author blamed *all of us*. Next day, on 7 December, Markov wrote a letter to this newspaper:

[…] *Such an accusation is libellous, and an attempt to conceal the real culprits. Happily, this libel is nonsensical and ought to be considered as mocking all of us, and especially the readers whom the newspaper is attempting to take for a ride. For its refutation suffice it to note that here in Russia we have neither a republic, nor a constitution, but only an unrestricted despotism.*

We may safely infer that this letter was never published.

P. 97. On 3 June 1907 the Second Duma was dissolved and a law proclaiming elections to the Third Duma was announced. I do not comment on Markov's letter to the Administration of the Academy of Sciences dated 11 June 1907:

The convocation of a Third Duma is connected with violations of the law, and it will therefore be not an assembly of the people's representatives but some unlawful medley. Therefore I have the honour of asking the Administration not to include me in the list of voters.

And in a letter to the newspaper *Tovarishch* [no date provided] Markov protested against the accusations of the first two Dumas. This, as he stated,

Can only assist in prolonging the state of Russia, which, in the appeal of unknown authors, is described as <u>Scandal and ruin</u>.

Pp. 104 – 105. In 1911 – 1913 a shameful affair of the Jew Beilis had been going on. The preliminary investigation dragged on for two years, but finally, in 1913, the jurors acquitted him of ritualistically killing a Christian boy. However, during these two years a large-scale anti-Semitic campaign had been waged. I

happened to see in the Internet that Florensky, a serious philosopher of mathematics, stated in a private conversation that he himself, had he been a Jew rather than a Russian clergyman, would have killed that boy.

On 1 October 1913 Markov wrote to Steklov:

I do not sympathize with any Jewish sect[10] *[…] but here, however, I see that the essence is not in the Beilis' crime but in the crime of the Russian justice guided by a union of Russian killers*[11]. *Irrespective of the future verdict Russian justice had already condemned itself by compiling a senseless indictment in which Beilis is allotted the least possible part but much irrelevant matters is stated.*

At about the same time Markov sent an open letter to Zamyslovsky, the leader of the extreme right-wingers in the Third Duma:

No court has established that Jews participated in the killing of that boy, but you decided to state, publicly and insistently, that they had tortured him to death. Such an insistence compels me to remark that quite another assumption is also possible, that he was killed not by the Yids[12], *as you express yourself, but by associations which dare to call themselves real Russians or by those who obeyed their indication and assignment.* […]

I mention two points. First, the <u>unionists</u> *attempted to seize immediately the preliminary investigation, to carry it on according to their interests and to remove everything opposing them. Second, the killing of that boy, being shifted onto the Yids, fully corresponded to the aims of the* <u>unionists</u> *as expressed in your pogrom speeches. And it is hardly possible to think that they are scrupulous about the means for achieving their goals.*

And so, I dare to believe that you yourself will admit that my assumption is not groundless although you will hardly associate yourself with my desire to see a most rapid and complete break of the <u>unionists</u>' *activities.*

P. 111. On 5 April 1908 Markov sent a letter to Lappo-Danilevsky:

[…] *The section on the probability of testimonies belongs to the most precarious parts of the theory of probability even if they only deal with confirming or denying a fact. More complicated cases, as far as I know, are not considered at all*[13].

P. 136. On 25 May 1921, at a sitting of the physical and mathematical class of the Academy of Sciences, Markov said:

I have at last received the footwear. However, it is not only badly sewed but not of the right size for me at all. And so, just as previously, I am unable to attend properly the sittings of the Academy. I suggest to place my footwear in an ethnographic museum as a specimen of the present material culture, and I am prepared to sacrifice it[14].

P. 137. On 24 October 1921, at a sitting of the scientific council of the Petrograd[15] university, the following statement of a group of professors was read out:

For successfully studying in the university the student ought only to have a corresponding schooling. Therefore, entrants should be admitted in accordance with their knowledge rather than owing to some class or political considerations. It is all the more necessary for the teaching staff to be properly scientifically qualified which can only be ascertained by the university itself.

The base of the university reform as indicated by the reform of the social faculty [faculty of social science?] *and the statutes of research institutes separate the educational and the scientific functions of universities. This is contrary to the very idea of a university whose main aim is to prepare scientists able to assist the progress of science and at the same time to provide a wide scientific education.*

Markov and Steklov were among that group and there are grounds for assuming that Markov, who was the first to sign that statement, had compiled its text[16].

Notes

1. Privat-Dozent: a freelance instructor.

2. Contrary to general opinion, Markov did *not* examine the method of least squares (Sheynin 2006b). The Gauss – Markov theorem is really due to Gauss alone. It was Neyman who initiated this mistake but later recognized it, see Neyman (1938/1952, p. 228) and Sheynin (2009, § 14.2-1).

3. This is a great exaggeration.

4. I venture to disagree: serious reviews are most important, cf. my commentary on [ii]. Thus, in 1915, the Academy of Sciences awarded Chuprov a gold medal for a review compiled on its behalf (Sheynin 1990/2011, p. 50). I also note that Markov (Ondar 1977/1981, p. 135) had demandingly appraised a certain contribution (which therefore had not been awarded by the Academy).

5. Markov, after all, did begin lecturing after Buniakovsky's death. Other scientists had praised him *insufficiently* possibly because of his *hot temper and unbalanced character* (Grodzensky, p. 63, a correct statement).

6. Instead of Grodzensky's note on p. 63 see Sheynin (2009, § 13.1-4).

7. Vasiliev published a long French paper (1898) which described Chebyshev's life and work. Its German version appeared in 1900. A *solemn meeting* did take place (Sheynin 2006a).

8. Concerning that jubilee see Bernoulli (1986). At about that time at least some Russian newspapers began to indicate the dates of their issues in both styles. The new style was officially adopted in 1918 although the Russian Orthodox Church is still keeping to the old one.

9. German scientists, painters and writers published an *Appeal to the Cultural World*. They justified *Germany's pure attitude* by *the difficult and imposed on it struggle for existence*. Among those who signed it were Ostwald, Planck and Klein. S. G.

Markov explained their behaviour by *intoxication or psychosis*, but the real reason was perhaps the fright at worsening their scientific carrier. O. S.

10. Over the centuries, there were many Jewish sects, some of them of the ultraorthodox direction, but none had practised ritualistic killings.

11. This is a reference to the extremely reactionary and anti-Semitic Union of Russian people. Its members were called *unionists*.

12. In Russia, many centuries ago, it was an official name of Jews, but later it became derogatory. By the end of the 19th century, Yid (*zhid*) disappeared from the educated and progressive circle of Russians.

Here is a remark made about 200 years ago by Schlözer (1823/1827, p. 155) the son of the celebrated August Ludwig von Schlözer and a professor of many years at Moscow University:

The usual Abgeschmacktheit (platitude or indiscretion) *of aliens ascribing a name to a clan or a people*.

13. This is a mistake. Poisson (1837) had considered even too complicated problems which did not really deserve any attention.

14. This episode illustrates the horrible life in Russia during the early 1920s. At the same time it shows the conditions under which Markov spent the last years of his life.

15. After the outbreak of the WWI Petersburg was renamed Petrograd, and, after Lenin's death in 1924 it became Leningrad:

Petersburg has now for some inscrutable reason been given the name of the man who has practically ruined it (Pearson 1978, p. 243).

In 1991, the city once more became Petersburg. Incidentally, contrary to general belief, it was thus named after Apostle Peter rather than Peter the Great.

16. Horrible conditions of the life of universities and of education in general in those years are well known, see for example Rostovtsev (1919).

Bibliography

Bernoulli J. (1986, Russian), *O Zakone Bolshykh Chisel* (On the Law of Large Numbers). Moscow.

Neyman J. (1938), *Lectures and Conferences on Math. Statistics and Probability*. Washington. 1952.

Ondar Kh. O. (1977, Russian), *Correspondence between A. A. Markov and A. A. Chuprov on the Theory of Probability and Math. Statistics*. New York, 1981.

Pearson K. (1978), *History of Statistics in the 17ᵗʰ and 18ᵗʰ Centuries* […]. Lectures 1921 – 1933. Editor E. S. Pearson. London.

Poisson S.-D. (1837, 2003, French), *Researches into the Probabilities of Judgements in Criminal and Civil Cases*. Berlin, 2013. **S, G,** 53.

Rostovtsev M. T. (1919), Proletarian culture. *Russ. Liberation Committee* [*Publ.*] No. 11.

Schlözer Ch. von [junior] (1823, French), *Grundriß der Gegenstände welche in den Theorie der Statistik so wie in der Geschicht*e […] *enthalten sind*. Göttingen, 1827.

Sheynin O. (1990, Russian), *Alexandr A. Chuprov. Life, Work, Correspondence*. V&R Unipress, 2011.

--- (2006a, Russian), On the relations between Chebyshev and Markov. *Istoriko-Matematich. Issledovania*, vol. 11 (46), pp. 148 – 157.

--- (2006b), Markov's work on the treatment of observations. *Hist. Scientiarum*, vol. 16, pp. 80 – 95.

--- (2007), Markov: Integrity is just as important as scientific merits. *Intern. Z. f. Geschichte u. Ethik d. Naturwissenschaften, Techn. u. Medizin*, Bd. 15, pp. 289 – 294.

--- (2009), *Theory of Probability. Historical Essay*. Berlin. **S, G,** 10.

Vasiliev A. V. (1898), P. L. Tchébychef et son oeuvre scientifique. *Boll. bibl. e storia sci. nat*. [t. 1].

XI

V. A. Steklov

Andrei Andreevich Markov (an obituary)

[1] Science has endured a new heavy loss: the death, on July 20, of A. A. Markov, an eldest member of our Academy, a student and follower of Chebyshev. He was of the same age as academician A. M. Liapunov, who had passed away in 1918.

A. A. was the son of a village deacon (Ryasan province), and his mother was a daughter of a civil servant. After graduating from a theological seminary, A. A.'s father had worked as a civil servant, then, when living in Petersburg, became a manager of houses and estates. He was twice married. Apart from daughters and a son, A. A., by the first marriage, he had a son, Vl. A. Markov, by the second marriage. Vl. A. was an outstanding mathematician as well, but he died from tuberculosis at an age of ca. 28 years. Nevertheless, he had time for becoming eminent in Russia and even abroad.

A. A. was born on 14 June 1856. In 1874 he graduated from the Petersburg gymnasium No. 5, and in 1878, from the mathematical department of the physical-mathematical faculty, Petersburg University as a candidate [of science]. At the same time he won a gold medal for a composition on a theme proposed by the Faculty, *Integration of differential equations by means of continued fractions*.

At the end of the 1860s and in the 1870s that faculty of the Petersburg University could have competed with the best universities of Western Europe. Among the professors of the mathematical department were such leading figures as Chebyshev, E. I. Zolotarev, A. N. Korkin and such outstanding instructors and scholars as K. A. Posse and D. K. Bobylev[1].

The three first-mentioned mathematicians exerted an especially strong influence on their students. In addition to reading the prescribed lectures, professors Korkin and Zolotarev devoted special hours, usually at home, to special lessons for, and scientific talks with the most talented listeners, including Markov.

Korkin was a greatly learned person with a rare memory. He knew classical literature, and especially Euler perfectly well and was often able to indicate at once the volume and even the page where Euler had expressed a certain idea, not yet sufficiently appraised or developed. Chebyshev himself frequently formulated new questions and problems and offered valuable advices and indications.

It is easy to understand how great was the influence of such talks with these outstanding thinkers on the development of the talent of their students and how rich was the food which they gave for independent investigations in various branches of mathematics. Markov's mathematical talent rapidly developed in such an exceptionally favourable atmosphere. Two years after graduation, in

1880, he had already passed his master's examination and defended a master dissertation (1951b).

[2] In the autumn of the same year, at the age of 24, Markov was admitted as a privat [freelance]-dozent of the University and began his independent work as an instructor. After four years, in 1884, he defended his doctor dissertation (1884) and in 1886 became extraordinary, and in 1893, ordinary professor of the University.

Also in 1886, on 13 December, A. A. was elected, as proposed by Chebyshev, adjunct (junior scientific assistant) of the Academy of Sciences. In 1890, on 3 March, he became extraordinary, and on 2 March 1896, ordinary academician (mathematics).

In 1905, after 25 years of work and being 49 years old, Markov became an honoured professor of the University and retired from professorship. However, he did not leave pedagogic work: as an academician, he continued to read courses in the theory of probability, continued fractions etc. almost to his last days. As he himself said, he only retired not wishing to stand in the way of younger people. In 1913, as proposed by me, Markov was elected honorary member of the University, but Kasso, the then Minister of People's Education, did not confirm that election.

[3] Markov's scientific merits are very considerable and diverse. His main investigations belong to the following branches of pure mathematics. To the theory of linear differential equations, and especially to the well-known Lamé equation and the equation of the hypergeometric series; to the theory of finite differences and interpolation; the Chebyshev theory of functions least deviating from zero; the theory of algebraic continuous fractions and its applications; the Chebyshev problems on the limiting values of definite integrals; approximate calculation of definite integrals; number theory and especially the theory of quadratic forms; and calculus of probability.

In each of these fields, which cover almost every branch of pure mathematics, A. A. selected essentially important and yet unsolved problems and offered their complete and perfectly rigorous solutions[2]. His investigations of the theory of differential equations (see above) mostly touched on the Lamé equations and the equation of the hypergeometric series which play an important role in analysis and have numerous applications in all the branches of mathematical physics and in astronomy.

These investigations are connected with the works of such eminent mathematicians as Felix Klein, Halphen and Goursat[3] on the so-called reducibility of linear differential equations. For example, Markov completely solved the problem on the reducibility of a differential equation of the third order closely linked with the differential equation of the hypergeometric series with five parameters (1897; 1896b), the problem about the distribution of the roots of the Lamé functions (1896c). To these investigations we may add Markov (1898) which, in turn, is directly linked with Chebyshev's general theory of polynomials as well as with the main theorems of the calculus of probability. In demonstrating them, he had indeed applied the obtained results.

Chebyshev had created the theory of functions (mostly polynomials and a certain type of rational fractions) least deviating from zero. However, we have to discover functions of other types possessing the same property for solving many incessantly appearing applied problems. No general theory covering a somewhat wide class of such functions is known, and, for that matter, such a theory can hardly exist. In each particular case we have to invent special methods for finding such functions, and the success here only depends on the ingenuity of the researcher who has to possess an outstanding talent and be an eminent specialist in analysis. We often have to guess the result, then prove our conjecture.

Markov had that ability in a high measure as shown by many of his works on special types of maximal and minimal magnitudes. Thus, he (1895) directly indicated the final solution of the following problem:

Represent on a plane a part of a surface of rotation between two parallels and two meridians so that 1) the parallels will be shown as concentric circles, and meridians, as their radii; 2) *the ratio of the maximal scale of the image to the minimal scale will be as close to unity as possible*.

After providing the final result, Markov proves that there exist such projections which are determined by his provided formulas and that they indeed satisfy the stated requirement.

His other investigation (1889b) is of the same type. He solves problems about the connection of two given segments by the shortest possible curves under some additional conditions. The main requirement is that at the points in which the curve meets the segments the curvature of the connecting curve ought to be zero and along all the curve it should as little as possible deviate from a given value.

Problems of this kind are essentially important for the construction of railways (when the track changes its direction). Various pertinent rules are described in treatises and manuals for engineers […] and special journals […]. Markov provided rigorous solutions of some of those problems which belong to the theory of functions least deviating from zero; theoretically confirmed some of the earlier rules (for example those, which were provided by Norlund et al in the *Annales de ponts et chaussées* in 1886 on the joining of a straight line and a circumference by a parabola of the third order); indicated necessary changes in some of those rules to avoid the possibility of gaps in the curvature and in some other elements of the curve etc.

Markov's paper (1889a) also deserves attention. There, he very simply proves that only the stereographic projection represents any great circle of a sphere as a circumference or a straight line. A. A. indicated that theorem in one of the propositions in his doctor dissertation (1884). Two years later, M. M. du Chatenet proved the same theorem apparently not knowing the Markov result.

To the field of the functions least deviating from zero also belong Markov's papers (1901; 1890). In the former, he perfected Chebyshev's analysis and final conclusion, in the second paper Markov solved a problem which our celebrated chemist Mendeleev had needed for studying the mixture of spirits and water.

There exists an opinion that A. A. was a theoretician never interested in the applications. The cause of this misunderstanding is that he often objected to wrong practical applications of mathematics, but, in his usual manner, the form of his arguments deceived non-specialists. Actually, Markov only objected to the use of mathematics only for attempting to give substance to groundless figments, to its obviously clumsy application rather than to the essence of the matter[4].

The examples above sufficiently well show that A. A. himself often applied mathematical analysis for the solution of practical problems and considered that work very useful and important. He compiled a treatise on the calculus of finite differences whose second edition appeared in 1911; in 1896 its first edition was translated into German.

That treatise is as worthy as in general are all Markov's works: the proofs are simple and rigorous and it essentially differs from most of the other publications on the same subject. His main attention was here directed to applications of that calculus to interpolation and, second, to compilation and use of mathematical tables. To this subject he devoted the entire first part of the book.

In the second part Markov first of all considered summation and methods of approximate calculations and only then came the equations in finite differences. In a special chapter he described the connection of linear equations of the second order with the theory of continued fractions. Here also the reader will mostly find practical problems which are incessantly encountered in applied sciences, physics, astronomy, statistics[5]. All problems are originally treated, often by Markov's own methods, absolutely rigorously, and illustrated by many numerical examples.

In passing, I note that A. A. was an expert in calculations, and like Chebyshev attached importance to the ability of performing them. Indeed, he remarked that final solutions of most problems lead to numerical calculations. A specimen of his calculations is his table (1888) for $x = 0(0.001)3.790$ with 11 significant digits[6]. Markov checked his table and reprinted it (1900/1913) with six significant digits.

The integral which he considered is of utmost importance in the calculus of probability and its applications, and each researcher seriously working in statistics, insurance, calculations concerning various retirement funds, has to apply values of that integral, and, consequently, its tables.

Markov's table (1909) is of the same kind. It is a continuation of Eisenstein's table whose calculations reached the determinant 20 (but he also included vanishing forms).

[4] In his main works, Markov was a direct follower of Chebyshev, as already seen from the above. Among his other similar contributions are various investigations on the Chebyshev problem about the limiting values of definite integrals. For example on determining the extreme values of the integral

$$\int_a^x p(x)f(x)dx$$

given the moments of the unknown function $f(x)$ until some order μ, i. e., given the integrals

$$\int_a^b x^k f(x)dx, \ k = 1, 2, \ldots, \mu.$$

Here, $p(x)$ is a function which, just like its derivatives up to the $(\mu + 1)^{th}$ order, does not take negative values.

   Markov first offered solutions of such problems in his doctor dissertation (1884). Later he solved more general and complicated problems in various memoirs in the academic *Zapiski*, in *Acta Math.* and *Math. Annalen*. In some cases he thus forestalled the famous Norwegian mathematician Stieltjes. On this point Markov exchanged many letters with him.

   These investigations are important since their results are valuable both theoretically and practically and, in addition, since Markov had applied methods which connected the solutions with the theory of continued fractions. Chebyshev, his mentor, especially emphasized their significance. Markov also connected those studies with interpolation and approximate calculation of definite integrals.

   Note also the absolute rigour of Markov's reasoning which, as stated above, distinguishes his entire work. He himself stated that he did not *recognize non-rigorous proofs if only I do not* [he does not] *see the possibility of making them rigorous*.

   Markov found a [new] application of continued fractions by deriving special formulas of mechanical quadratures similar to Chebyshev's but having alternating signs and, what is especially important, he (1898 Russian, 1904 French) provided exact expressions for the residual terms. He (1884) also was the first to offer the exact expression of the residual term for the known Gauss quadrature formula.

   The theory of continued fractions had been attracting his attention during his entire scientific life. Even his first student composition, as mentioned above, was devoted to their application to the integration of differential equations; almost all of his mentioned contributions were in some way connected with that theory. Not long before his death he read a special course on that same theory and prepared it for publication. As far as I know, one of the now existing publishers had agreed to issue it. I am ignorant of the present situation, but in any case it is desirable that his students carry this business through[7].

   [5] Especially numerous and important are Markov's investigations in the theory of probability. In this field he may be considered one of the best specialists in the world. His *Calculus of Probability* (1900 and later editions) is a treatise of exceptional worth. Especially remarkable are his investigations on the celebrated Jakob Bernoulli theorem (the law of large numbers), on the two main theorems of that calculus first established by Chebyshev, and to the method of least squares [see below].

   In general, the Petersburg school of mathematicians, as whose founder we should recognize Euler himself, introduced very important and original contributions to this field of mathematical knowledge.

Academician Buniakovsky compiled a treatise on the mathematical theory of probability which, for his time, was complete and outstanding.

Chebyshev's investigations introduced a number of new items in this field and turned it in the exact mathematical direction. His students, Markov and Liapunov, introduced new wide generalizations and complete rigour in the methods of investigation. By applying a special method of moments (or expectations, whose idea belongs to Bienaymé), Chebyshev established the two main theorems of the calculus of probability: on the limits of the expectation and of the probability, the first of which was a generalization of the theorems of Jakob Bernoulli and Poisson. However, his formulations gave rise for doubting the rigour of the proof [of the second theorem] and did not sufficiently ascertain the conditions under which his theorems [?] were indeed valid.

By applying a subtle, although often elementary analysis, Markov, in a series of memoirs, and then in his treatise (1900), developed all the pertinent issues with complete rigour. A year later Liapunov used a wholly different method and proved the latter theorem with such generality which apparently was impossible for the method of moments that had been applied by Chebyshev and Markov. In his peculiar frank way, Markov often stated in Liapunov's presence that he *played a really dirty trick on me*. The method of moments assumes the existence of such expectations on which the Liapunov method does not depend at all and at first Markov thought that it was impossible to obtain Liapunov's result by the former method. However, after seven years the *trick* much pleased him. Not ceasing to think about this problem, Markov found a means for generalizing the method of moments. Not only did he thus obtain the Liapunov result for magnitudes independent one from another, which all previous authors had assumed, but he generalized the main propositions of the theory of probability on many cases of magnitudes connected one with another in a definite way.

Markov was thus able to analyse a new wide range of problems which had been previously hardly touched on and which, apart from being interesting in themselves, can have many important practical applications. The last years of Markov's scientific life had been largely devoted to the study of this new chapter of the calculus of probability, to the probability of events connected, as he expressed it, in a chain[8].

He expounded his main ideas in numerous memoirs (mostly in the *Izvestia* of the Russian [of the Imperial] Academy of Sciences) and partly in the third edition of his treatise (1900) which appeared in 1913 on the occasion of the bicentenary of Jakob Bernoulli's discovery of the law of large numbers and in a French book published [in Petersburg] at the same time.

Markov's investigations of the method of least squares also merit special attention […][9].

It goes without saying that the moral expectation had absolutely no place in his *Calculus of Probability*[10].

A. A. did not quit his work almost to the end of his life. Being already sick and bed-ridden, he corrected his manuscript *On continued fractions* for the mentioned publisher and submitted his paper (1924) to the Academy. As usually, he intended to study completely and quite rigorously all the pertinent problems, but the illness prevented it. When passing me this work for a report at the Academy, he asked me to tell [those responsible] that under normal conditions he would have never published not quite a prepared study, but now, as he added, he feels that death is approaching and is afraid that he will be unable to complete the work. He therefore decided to publish it as his last contribution.

[6] My brief essay on the work of the late Markov to some extent clears up, as I think, his remarkable scientific merits. They allow to place him alongside the geometers of the first rank, and his courses on the *Calculus of Probability*, *Calculus of Finite Differences* and *Continued Fractions* (which will been undoubtedly published)[11] are exemplary in every respect both for Russian and foreign sources. For a long time they will serve as reference books for students of higher educational institutions and all those who work in the field of mathematical sciences.

At least during the years of most intensive work, Markov had not participated in social or political activities, but wholly abandoned himself to scientific studies[12]. However, we cannot say that his life proceeded in the atmosphere of calm and tranquillity which distinguished scholars often create for themselves.

His temperament was not at all passive. It was so peculiar that his statements, even in scientific debates or on issues of the Academy, which were sometimes properly appraised, led on other occasions to undesirable misunderstandings. An essential significance in dealing with other people has not only *what* is said, but *how* it is said, and the *how* is often more important than the *what*.

People having the so-called *public vein* possess a special sensitivity, a kind of diplomacy which allows them to *speak the truth to the Tsars with a smile* [G. R. Derzhavin, *Civil Poetry*], to offer their opponents sweetened pills and thus to disarm the short-sighted without irritating them.

Markov absolutely lacked such characteristics, he was organically unable to endure their even tiniest manifestations whoever revealed them. In addition, he was incapable to compromise. He could stand any sharp statement about himself if only they really touched on the essence of the business at hand, not deflected him, not distracted him from the main subject towards personal feelings or a compromise settlement which usually did not satisfy anyone.

Markov always began his objections and statements with a sharp definiteness to which he was accustomed in his scientific studies. This often annoyed touchy people who were not used to such objective and logical forms of *talks*. His opponent often put the essence of the debate aside and began to object diplomatically to its form as shaped by Markov, and that invariably unbalanced him. Such debates led to conflicts and mutual misunderstanding. Markov's proposals,

essentially proper, had been often rejected only because of their discomforting form.

Everyone knows his encounters with academician V. G. Imshenetsky, then with his defenders, Professors K. A. Andreev and P. A. Nekrasov, his special debate with the last-mentioned and with the Moscow Mathematical Society about Nekrasov's frames in the theory of probability, and with academicians B. B. Golizsin and F. A. Bredikhin et al[13].

These and similar cases gave occasion for the dissemination, especially beyond Petersburg, of sharp negative opinions about Markov's disposition and actions which I personally had to hear all the time in Moscow and Kharkov even before my acquaintance with him. Now we may say absolutely impartially that all those negative opinions had been extremely exaggerated and far from just.

I do not at all justify the sharp form in which Markov sometimes formulated his objections, but I ought to say that in most cases he was in the right. He only attached more weight than it was possible to require in the circumstances to *all* the shortcomings in his opponent's considerations, but the notion of weight, just like in the method of least squares, is tentative[14].

Most of Markov's objections, in spite of the often verbose refutations, remained valid and, as I dare believe, would not have led to the appearing aggravations had he been able to formulate them, the objections, in a more usual form and had his opponents attempted to consider deeper the peculiarities of his disposition. But what can be done since the process of his thinking and the properties of his uncommonly frank soul were so peculiar that they did not fit in the usual boundaries? However, we may only respect these, perhaps not always pleasant peculiarities.

[7] Markov's other statements, not purely academic or scientific, had not always been correctly elucidated either and were sometimes ascribed to eccentricity, to the desire to show off. Known, for example, is his statement about the renunciation of ranks and orders which in those times was considered as a great impertinence; his sharp statements about the expulsion of Gorky (real name, A. M. Peshkov) from honorary membership in the Academy after a royal command; his demand to be excommunicated from the Russian Orthodox Church submitted to the Most Holy Synod after the excommunication of Tolstoy[15] etc.

Any protest against actions adverse to the convictions of a given person and especially when stated publicly and not in a form quite usual for the majority of people, can be attributed to eccentricity, but is it really so? We may maintain that his actions were not explained by eccentricity or a show. He had indeed, according to his nature, been *unable to remain silent* and his actions had not been made light-heartedly at all.

Markov had been sincerely indignant at what he considered wrong or unjust and was unable to conceal his feelings, only he expressed them too much in his own way. The explanation of his behaviour by eccentricity was certainly useful for him since he remained unpunished, every cloud has a silver lining, but it was unjust.

[8] In his last years the notion of an imminent death had been extremely oppressing him, but certainly not because he was afraid of dying, but since he had to leave his son who had not yet finished his education, had not yet become firmly independent[16]. But even in such serious moments Markov remained his own self and once expressed his, so to say, protest against the approaching end in the following way. He hopelessly answered my question about his health by reciting, in a quivering voice, Pushkin's *Eugene Onegin*[17]:

My uncle, what a worthy man // […] *But God, what tedium to sample // That sitting by the bed all day // All night, barely a foot away*

And the hypocrisy demeaning // Of cosseting one who's half alive.

Puffing the pillows […] *Thinking with a mournful sigh // Why the devil can't I die.*

[Instead of *can't he die*]. And he began weeping.

For most of us such a form of expressing sorrow at the approaching death will also seem strange and perhaps inappropriate, but may we denounce a prominent man just because he was never able to adapt himself to the general pattern?

Even in his youth A. A. had not been distinguished by a robust health, but he always remained very lively and active[18]. However, during his last two or three years he began to weaken considerably and to be indisposed. Sometimes his temperature rose, but at first he attributed all that to colds or malaria and at times he treated himself. However, the illness dragged on and the weakness increased. A few months before his death Markov became bed-ridden. Sometimes he left his bed, but soon was obliged to lie down again.

In July he felt that for a while he ought to leave without fail the city and have a breath of fresh air. Not long before the trip a small wound appeared on his healthy leg (the other leg had been troubling him from childhood), but he still went with family on a lorry to Novo-Aleksandria. The journey was very jolty, his wound began to bleed and continued to bleed for a long time after arrival.

In spite of some taken steps his condition had not changed and physicians advised him to return to Petersburg and be admitted to a hospital. There, it turned out that the bleeding was caused by an aneurism. Rotting, which extremely developed owing to Markov's general emaciation, was also revealed and soon there appeared indications of blood-poisoning. The physicians though of amputating the leg, but this proved useless[19].

I had returned from Moscow only a day before A. A. died, but one of his students (Besikovich), who had often visited him, reported that almost to the very end A. A. remained fully conscious and attentively listened to the reading with which visitors had been trying to entertain him. As usual, he inserted original remarks and even in details disclosed his intrinsic perfect memory. Only immediately before dying he completely weakened and apparently lost consciousness. He died at about 22 hours, July 22 of this year and was buried in the Mitrofan cemetery.

Notes

1. On Korkin and Zolotarev see Ozhigova et al (1978). A Chebyshev student, Konstantin Aleksandrovich Posse (1847 – 1928) is less known and Dmitry Konstantinovich Bobylev (1842 – 1917) was a physicist and mechanician.

2. Steklov repeatedly mentions perfect or complete rigour, but I rather believe that rigour is a somewhat variable notion. Concerning Chebyshev, I additionally refer to Bernstein, see [iv, my commentary].

3. Georg Henn Halphen (1844 – 1889), Edward Jean-Baptiste Goursat (1858 – 1936).

4. Concerning Markov as an applied mathematician see also [iv, my commentary on § 3.4 of his essay].

5. An unfortunate expression.

6. On this point see Sheynin (2009, § 14.1-3) where I also refer to a reputed reference book. Its authors remarked that two tables of the normal distribution, one of them Markov's and the other one published ten years later, had remained beyond compare up to the 1940s.

7. I can only mention Markov's selected works (1948) and his first relevant contribution (1906).

8. Markov had only investigated his chains up to 1915.

9. This is a mistake and I have omitted Steklov's following description. Instead, see Sheynin (2009, § 14.2-1). Markov defended the second Gauss' justification of the method of least squares, but a few authors had forestalled him. He indirectly denied the first justification because it claimed optimality and preferred the second mostly since it only presumed a tentative optimality. Again indirectly it meant that Markov believed that any method of treating observations was good if it did not spoil robust observations. Then, I indicated that the still alive and kicking Gauss – Markov theorem is due to Gauss alone. See also [v].

10. On moral expectation see Sheynin (2009, § 6.1.1). Suffice it to say here that by 1900 only the economists of the Austrian school were interested in that concept.

11. See Note 7.

12. In § 9 Steklov contradicted himself.

13. Steklov undoubtedly prettified Markov, see Sheynin (2009, § 14.3). Thus, Markov criticized K. A. Andreev for publishing Imshenetsky's incomplete posthumous manuscript, but he himself, being unable to complete a manuscript, allowed its publication, see Steklov's § 8. And here is Andreev' s letter to Nekrasov of 1915 which I had quoted in 2009 in that § 14.3: Markov

Remains to this day an old and hardened sinner in provoking debate. I had understood this long ago, and I believe that the only way to save myself from the trouble of swallowing the provocateur's bait is a refusal to respond to any of his attacks...

14. At the very least this last statement is strange. In both cases the important point is the relative weight of the different circumstances/observations.

15. See [viii].

16. Markov's son, Andrei Andreevich Markov junior (1903 – 1979) was then about nineteen years old. He became a mathematician and corresponding member of the Soviet Academy of Sciences.

17. Translation: A. S. Klein, 2009. Markov junior (1951, p. 601) reported that as a school student his father had apparently been opposed to *Eugene Onegin* since it did not touch on social themes.

18. Markov used to repeat the verdict of a physician addressed to some postman: *You will remain alive as long as you keep walking* (Markov Jr 1951, p. 600).

19. Not amputation but ablation of the aneurism (Markov junior 1951, p. 613).

Bibliography
A. A. Markov

Abbreviations: *Izvestia* = *Izvestia Imp. Akad. Nauk*

Soobshchenia = *Soobshchenia Kharkov Matematich. Obshchestvo*, ser. 2

Zapiski = *Zapiski phys.-matematich. otdelenia Imp. Akad. Nauk*, ser. 8, separate paging

The numbers (3, 10, …) coincide with those provided in Markov (1951, Bibliography of his works, pp. 679 – 714)

3 (**1880,** Russian), On binary quadratic forms of a positive determinant. In Markov (1951, pp. 11 – 85).

10 (**1884**), O *Nekotorykh Prilozheniyakh Algebraicheskikh Drobei* (On Some Applications of Algebraic Continuous Fractions). Petersburg. The doctor dissertation.

14 (**1885**), Sur la méthode de Gauss pour le calcul approché des intégrales. *Math. Annalen*, Bd. 25, pp. 427 – 432.

22 (**1888**), *Table des valeurs de l'intégrale* [of exp ($-x^2$)]. Petersburg.

24 (**1889a,** Russian), On drafting [compiling] maps. *Soobshchenia*, vol. 1, No. 3, pp. 113 – 128.

25 (**1889b,** Russian), A few examples of the solution of special problems on maximal and minimal values. Ibidem, NNo. 5 – 6, pp. 250 – 276.

27 (**1890**, Russian), About one of Mendeleev's questions. *Zapiski*, vol. 62.

45 (**1895,** Russian), On the most advantageous representation of a part of a given surface of rotation on a plane. *Izvestia*, ser. 5, vol. 2, No. 3, pp. 177 – 187.

53 (**1896b,** Russian), On the zeros of an entire function and Lamé functions. *Soobshchenia*, vol. 5, NNo. 1 – 2, pp. 74 – 80.

55 (**1896c**), Nouvelles applications des fractions continues. *Math. Annalen*, Bd. 47, pp. 579 – 597.

57 (**1897,** Russian), On the differential equation of the hypergeometric series with five parameters. *Zapiski*, vol. 5, No. 5.

58 (**1904**), Recherches sur les valeurs extrèmes des integrales et sur interpolation. *Acta math.*, t. 28, pp. 243 – 301.

59 (**1898**), Sur les racines de l'équation exp$x^2 d^m$exp($-x^2$)/$dx^m = 0$. In Markov (1951, pp. 255 – 269).

60 (**1899,** Russian), The law of large numbers and the method of least squares. In Markov (1951, pp. 233 – 251).

70 (**1901,** Russian), On a mechanism of Chebyshev. *Izvestia*, vol. 14, ser. 5, No. 2, pp. 201 – 214.

91 (**1909**), Tables les formes quadratiques terniaires indéfinies ne représentant pas zéro, pour tous les déterminants positivs D   50. *Zapiski*, vol. 23, No. 7.

93 (**1910**), *Ischislenie Konechnykh Raznostei* (Calculus of Finite Differences). Odessa, second edition. First edition, 1889 – 1891.

104 (**1913**), Démonstration du second théorème limite du calcul des probabilités par la méthode les moments. (Suppl. à la 3me édition russe du *Calcul des probabilités*. Petersburg.

125 (**1924**), The difficulty of the method of moments. Two examples of its incomplete overcoming. *Izvestia*, ser. 6, vol. 16, p. 281 – 286.

129 (**1948**), *Izbrannye Trudy po Teorii Nepreryvnych Drobei* etc. (Sel. Works on the Theory of Continued Fractions and Theory of Functions Least Deviating from Zero). Moscow – Leningrad.

196 (**1906**), *O Nepreryvnykh Drobiakh* (On Continued Fractions). Petersburg. Mimeographed edition.

(**1951**), *Izbrannye Trudy* (Sel. Works). No place.

(**1990,** Russian), On the solidity of glass. IMI, vol. 32 – 33, pp. 456 – 467. Written ca. 1903.

Other Authors

Markov A. A. Jr (1951), The biography of A. A. Markov, Sr. In Markov Sr. (1951, pp. 599 – 613). **S, G,** 5.

Ozhigova E. P. assisted by A. P. Youshkevich (1978 Russian), Problems of the number theory. *In Math. of the 19th Century*. Editors A. N. Kolmogorov, A. P. Youshkevich. Basel 1992, 2001, pp. 137 – 209.

Sheynin O. (1989), Markov's work on probability. *Arch. Hist. Ex. Sci.*, vol. 39, pp. 337 – 377; vol. 40, p. 387.

--- (2009), *Theory of Probability. Historical Essay*. Berlin. **S, G,** 10.

Youshkevich A. P. (1968), *Istoria Matematiki v Rossii do 1917-go Goda* (History of Math. in Russian before 1917). Moscow.

XII

Oscar Sheynin

Dostoevsky and his *Jewish Question*

I bear in mind Dostoevsky's *Diary, papers, notebooks,* vol. 3, for 1877. Moscow, 2005, pp. 91 – 116. There, *The Jewish question* was the title of the second chapter of those notebooks. Many authors have described it and I am trying to repeat their commentaries as little as possible.

1. Very often Dostoevsky (e. g., p. 92) calls a Jew an *Yid* but he does not think *it is so offensive*. Terms concerning a multitude of men or, still more, a nation, ought to be only judged by statistical investigations, but he continues: Yid and the derived words *denote a known notion, direction* […]. Indeed, these words are generally encountered in the classical works of Russian literature (Lezhava 2009). However (*Wikipedia*, Yid), in 1787, when Ekaterina II passed through a certain provincial town, some local Jews petitioned her for banishing that humiliating word. The empress commanded to stop using it in official documents and towards the end of the 19th century *Yid* disappeared from the language of progressive educated Russians. Cf. Schlözer [x, note 12].

2. Dostoevsky (p. 102) states that *Jews only prefer one profession, gold trade and perhaps the processing of gold* [the profession of a jeweller], but he concentrates on the first point. And what is their aim? If not to carry the gold off to Jerusalem, then because of their *instinctive irrepressible urge*. He should have known that for a few centuries the Christians had been forbidden to do any business with money and that the Jews had therefore occupied that safety valve. And Jerusalem? Nonsense! Gold is easier to hide from pogrom-makers.

3. Dostoevsky charges Jews of *attacking* emancipated Russian serfs (p. 96) and emancipated American Negroes (p. 97), of *braiding them by their eternal gold trade*. In the first case he added: *the landlords at least tried not to ruin them*. There also (p. 97) Dostoevsky stated that the *Jews almost ruined the Lithuanians near Kowno* [Kaunas] *by vodka*.

Here, however, is Khor from Turgenev's *Khor and Kalynich*, a well-to-do serf on quitrent. *Many times* his master (serf-owner) asked him to pay himself off, but he declined. Was he afraid of Jews? No, of functionaries. So what can we conclude about ordinary serfs?

Much later, after the village community in Russia was broken up (Chuprov 1912), did the Jews *attack* the peasants? Chuprov did not mention Jews at all. And concerning the ruining of peasants (true, Russian peasants) by the Jews (and landowners!) Solzhenitsyn (2001 – 2002/2013, vol. 1, Chapter 1, p. 40) indicated that

They leased various posts belonging to the landowner's privilege – specifically the sale of vodka – and herewith fostered the expansion of drunkenness.

There also he (Chapter 8, p. 308) discusses the wine trade in more detail, and not only in Russia.

Dostoevsky (p. 93) quotes a letter which he received, but does not name its author:

Is the Russian Orthodox kulak, blood-sucker or publican […] *any better than their counterparts from the Yids*?

I note in passing that the term *kulak* has been current very long ago. Dostoevsky agrees: all of them are *bad*, but he only mentions Jews. The anonymous author was Abraham Uria Kovner, whose life was described by Grosman (1924).

4. In connection with the above, I remark: Dostoevsky stresses that Jews disturb the settled way of life, do not care about Russian national traditions, as though burst life itself. Yes, he was in the right or almost so. Solzhenitsyn (Chapter 6, p. 265) explains this phenomenon, but suffice it to repeat his quotation from Sombart, a German economist, sociologist and statistician:

The Jews had been the vanguard who created the capitalist world and mostly in its financial form.

In other words, they had sped up an unavoidable historical process. And I add in accord with Kovner (§3): Russian bourgeoisie had been *worthy* as well. Kuprin, in his *Moloch* of 1896, described one of them.

5. On p. 101 Dostoevsky quotes someone without naming him:

Know that […] *you are God's only one. Destroy everyone else or enslave or exploit them*.

An anonymous author in the Internet thought that that was a concise description of some Talmudic propositions (although *exploit* was hardly contained there) as interpreted by Grinevich (1876). And here is the general opinion about Dostoevsky (the item about him in the *Elektronnaya Evreiskaya Enz.*):

He pathologically hates the God-bearing nation, he is guided by *national religious messiahship*.

Destroying everyone else etc. directed to a tiny minority of the population, if taken from the Talmud, still rings senseless. Perhaps the opinion of the celebrated publicist, *the most subtle,* as Solzhenitsyn vol. 2, chapter 15, p. 103) called him, M. Hershenson (again Solzhenitsyn, vol. 2, p. 17) is more to the point: *The Jewish kingdom is not of this world*.

And here is Dostoevsky's own messiahship (end of Chapter 2, p. 91):

We, Russians, are indeed necessary and inevitable both for the entire Eastern Christendom and for the destiny of the future orthodoxy on our earth for uniting it. This is how our people and our sovereigns always understood it. […] *Constantinople* [Istanbul] *must be ours*.

Dostoevsky (p. 99) also reported that the Jewish inmates of the prison, in which he himself did time as a political prisoner,

Had been in many respects keeping themselves aloof from the Russians, did not wish to eat together with us, all but looked down on us.

Solzhenitsyn (Chapter 4, p. 177) added from the *Evreiskaya Enz.*, vol. 3, p. 334:

Until the mid-19th century even educated Jews with rare exceptions, having mastered the German language, did not know the Russian language and literature.

And on p. 178, citing another author: *The Jew did not want risking separation from his God.*

Disrespect for the basic population and utter nonsense but the Jews did not at all thus attained dominance over the world (or Russia).

The Jewish pseudo-messiahship mentioned by Dostoevsky together with his pronouncements (see below) could have well become the starting point for the author(s) of the later fake *Protocols of the Wise Men of Sion*.

The cited *Enziklopedia* mentions anti-Semitic statements from two of Dostoevsky's books. In *Notes from the House of the Dead* (publ. 1861 – 1862) he derisively and spitefully described a Jew. In *Brothers Karamazov* (publ. 1879 – 1880, just before his death) he included a blood libel. The most renown case of such libel occurred in 1911 – 1913 (the notorious Beilis case), but Dostoevsky forestalled it by about 30 years. The Editors of the German edition of Grosman (1924/1927, p. X) cited Dostoevsky's letter published in 1861 in the Slavophil newspaper *Den*'. He had remarked that a possible negative statement made by Judaism about Christianity can only be answered by quoting Jewish sources, and the Editors concluded that therefore Dostoevsky argued against the equality of rights for the Jews. However, I would say that their argument is rather hollow.

The statements in his books and especially the blood libel make nonsensical Dostoevsky's pronouncement in his Diary for March 1877, Chapter 3 in the edition of 1929 (Solzhenitsyn, vol. 2, p. 17):

The final word about this great tribe is still unspoken.

An American author, Pierce (1979), who described Dostoevsky's attitude towards the Jews, saw fit to gloat over the years after the 1917 coup d'état: the blood-thirsty Jewish commissars who constituted the bulk of the Bolshevist leaders, supervised the slaughter of millions of Russians. He published his paper in a journal of the white Anglo-Saxon protestants, and I would have asked him:

You bastard! Aren't you happy that your forefathers had all but wiped out the indigenous population of the U. S.? And let us recall Solzhenitsyn (Chapter 16, p. 137):

The (Jewish) *chekists were not capable of finding enough self-restraint and self-scrutinizing sobriety.*

But had the pogrom-makers been capable of self-restraint?

A chekist was a member of the Cheka, the Emergency Committee for Combating Counter-Revolution and Sabotage; then, since 1918, … for Combating Counter-Revolution, Profiteering and Corruption.

Nevertheless, I have to continue. In 1919, the leaders of the short-lived revolutionary government of Hungary were mostly Jews. They unleashed such terror that Lenin and Trotsky could have envied them. No wonder that after the revolution was suppressed Hungary became swept over by a wave of anti-Semitism which was barely noticed until then (Solzhenitsyn, Chapter 16, p. 144).

And here is the main point. The slogan *Proletarians of all countries, unite*! meant that neither national peculiarities, nor religion

were important. And who thought so? The Jews, in the first place, and that was what happened in Russia and Hungary. And the Soviet Union disintegrated, in particular, for that same reason. A catastrophe of the century (Putin), only the other way round. However, the first state of workers and peasants had time enough for playing dirty tricks on the whole world. In 1979, that state occupied Afghanistan by a *restricted contingent* of its army. However, the contingent (less its losses) had to return home in disgrace, militant Islam ripened the world over and mankind found itself in a new and extremely dangerous situation.

The anti-Semites have also been using that slogan understood in a generalized form: who is unlike us, is against us, cf. Luke 11:23. They somehow do not see the Islamic peril.

But *revenons à nos moutons*!

6. Dostoevsky (p. 106) recognizes that

Almost nine tenths (of the Jews) *are literally paupers.* […] *In their very work*, […] *in their very exploitation there is something wrong* […] *which leads to its own punishment.*

Its own apparently concerns the work and the exploitation. But who, why and how exploits the Jews? No answer. It is known, however, that the attempts to attract Russian Jews to agriculture proved futile, mostly because they had been carried out formally (Solzhenitsyn Chapter 2, p. 74; Chapter 3, pp. 112 – 114, 118; Chapter 4, p. 161). And, in general, Orthodox Judaism considered many occupations unworthy. Consequently the overcrowding (Solzhenitsyn Chapter 5, p. 216) or rather (!) the uniformity of the remaining trades (Chapter 3, p. 126) led to inter-Jewish economic competition.

And so, in the beginning of the 20th century and perhaps somewhat earlier those young Jews from Eastern Europe and Germany who thought of moving to Palestine, had to learn preliminarily how to saw and chop firewood, dig holes etc.

7. And here is Dostoevsky's general conclusion (p. 106):

The top layer of Jews is ascending over mankind.

Many similar statements are contained in his Notebook for 1880 – 1881 (p. 451):

The Yid community and the Yiddish idea envelopes the whole world

In Europe, there is only the Yid and his bank. He will say *veto* and Bismarck falls down like a mowed blade of grass

The Yid will uproot Christianity [by socialism] and ruin its civilization

All the riches of Europe will perish and only the bank of the Yid will remain

The Antichrist will come and rule unrestricted

The ideal of [moral] beauty is the Russian people

Strictly speaking, there never was any Christian civilization; it was and is, the Judaic-Christian civilization. True, Jews do not wish to read the *New Testament*, and the Christians are only acquainted with its utterly corrupted *churchy* version and are not interested in the *Old Testament*.

Dostoevsky justified his last statement by indirectly arguing that real Christianity only remained in Russia whereas in Europe *the main* (contrary) *idea of the bourgeoisie is reigning* (p. 105). I do not say that Russians are worse than, say, Italians, but where did Dostoevsky find their beauty? Not in Gogol's *Dead Souls*, and neither in the writings of Saltykov-Shchedrin. Did he bear in mind the Russian muzhik? But then, how about fathers living, in the absence of their sons, with their daughters-in-law?

Now, his *Crime and Punishment*. Why was the pawnbroker an old Russian woman rather than an Yid? A possible reason is that in that case Raskolnikov would have not experienced much moral doubts.

Socialism? And even communism? Bad luck: Marx was a baptised Jew. Bank of the Yid? Dostoevsky (§ 2) quite wrongly explained the Jewish inclination to gold.

Much more important: Disraeli became all-powerful in England and very powerful in Europe, but no Jewish kingdom had appeared anywhere (and Dostoevsky certainly knew that). Then, during WWI, Jews fought Jews. And now the year 1933. Contrary to Dostoevsky, no international care about saving the German Jews (for the time being, only them) had occurred. Much worse: in each country, Jews became afraid of an influx of newcomers. What happened was not a Jewish kingdom, but the Holocaust.

So why Raskolnikov had great doubts whereas his creator had none at all, even with respect to a *great tribe*? Then, just as now, when Jews were/are concerned, everything went/goes.

Bibliography

Chuprov A. A. (1912), The break-up of the village community in Russia. *Econ. J.*, vol. 22, pp. 173 – 197.

Grinevich M. I. (1876), *O Tletvornom Vliyanii Evreev* … (On the Pernicious Influence of Jews). Petersburg.

Grosman L. P. (1924), *Ispoved' Odnogo Evreia* (Confession of a Jew). Moscow – Leningrad. The book was translated into German (Munich, 1927) and the author's name became Großmann.

Lezhava Irina (2009), Yid in the works of Gogol, Turgenev, Chekhov … Prosa RU.

Peirce W. (1979), Dostoevsky and the Jews. *National Vanguard*, No. 72.

Solzhenitsyn A. I. (2001 – 2002), *Dvesti Let Vmeste* (Two Hundred Years Together), vols 1 – 2. Moscow, 2013. I have used some quotations from (an incomplete) translation published in the Internet, but I was unable to understand who and when translated Solzhenitsyn.