Studies

in the History of Statistics and Probability

Vol. 16

Compiled by Oscar Sheynin

Berlin
2020

Contents

I am the author of almost all the contributions listed below
(although in one case, a joint author)

oscar.sheynin@gmail.com

Introduction by the compiler

Notation

Notation **S, G,** *n* refers to downloadable file *n* placed on my website www.sheynin.de which is being diligently copied by Google (Google, Oscar Sheynin, Home. I apply this notation in case of sources either rare or translated by me into English.

I

The notion of randomness from Aristotle to Poincaré

Istoriko-Matematicheskie Issledovania, vol. 1 (36), No. 1, 1995, pp. 85 – 105

1. Introduction

Aristotle and even earlier scientists and philosophers attempted to define, or at least to throw light upon randomness, and in jurisprudence, about two thousand years ago, it was indirectly recognized in an ancient Indian book of instructions [1, § 108] which determined the behaviour of man both at home and in social life.

In § 2 I sketch the attempts to direct the concept of randomness into the realm of mathematical science; in §§ 3 – 10 I dwell on various interpretations of randomness that were pronounced in natural sciences and philosophy; my § 11 is devoted to the interrelation between necessity and randomness; and, finally, in § 12, I formulate my conclusions. Since Aristotle, Darwin and Maxwell described (used; indicated) various aspects of randomness, I repeatedly mention each of these great scholars. The history of randomness is especially interesting since the new approach to its understanding that had recently took shape in physics and mechanics and has affected the fundamentals of these sciences [2].

Elsewhere, [3] I examined the work of Poincaré, who paid much attention to randomness. Here, I only mention that he directly linked chance to instability of motion [of the solution of differential equations] and introduced the fruitful method of arbitrary functions [4, pp. 88 – 89]. Then, there is a case for studying the attitude of ancient scientists preceding Aristotle towards randomness. However, my own experience [5, §§ 2.1 and 2.3] is that this topic is extremely difficult since their thoughts may be interpreted in different ways.

Finally, I restrict my paper with the fields of mathematics and natural sciences.

There is no general literature on my subject; one author [6] had discussed randomness from a different point of view, some other [7 – 13] busied themselves with its particular issues; I mention contributions [7 – 9] in the sequel. I myself touched the same topic in many articles published in the *Archive for History of Exact Sciences* and my excuse for doing so, and for returning to the same subject in an ad hoc paper, is that it is patently impossible to compile a contribution such as this one all at once.

2. Mathematics and the concept of randomness

Lambert [14, pp. 238 – 239; 15, p. 246; 5, pp. 136 – 137] made an endeavour to formalize randomness. His interest in this problem may be explained by the fact that he was the first follower of Leibniz in the attempt to create a doctrine of probability belonging to a general science of logic. Lambert's efforts, founded on an intuitive notion of normal numbers, was ahead of its time. True, Cournot [16, pp. 57 – 58] and Chuprov [17, p. 188] had noted Lambert's efforts, but no one became interested in their accounts.

Poisson [18, pp. 140 – 141] hesitatingly offered a definition of a random magnitude as a variable that assumed different values with corresponding probabilities. His definition (independently reintroduced at the end of the 19th century [19, § 15.4]) went unnoticed. Poisson [18, p. 80] also attempted to state the nature of chance. Randomness, he argued, was an ensemble of causes that produced an event without altering its (the event's) chances of happening or failing. His idea seems unsuccessful, but at least Poisson thus maintained that random events possessing stable probabilities of two possible outcomes do occur.

While attempting to construct the theory of probability anew, von Mises [20a, p. 62] reintroduced after Fechner the concept of Kollektiv (of an infinite random sequence) and demanded that the order in which its elements followed each other be random (*mit zufallsartiger Zuordnung* …). Later he [20b, 1939, p. 32] equated chance with *complete* '*lawlessness*', cf. § 6, and (Ibidem, p. 133) noted its *fundamental importance* for the theory of probability. For an evaluation of his efforts see [4]. I only remark here that mathematicians became interested in defining the collective and attempts of such kind are now continued in the modern theory of algorithms. Three approaches are now recognized [22, pp. 199 – 214].

The *frequency* viewpoint had originated with Mises (even with Lambert) and in 1963 Kolmogorov modified it. It demands that the various elements of a random sequence and of its legitimate subsequences appear with stable frequencies.

According to the approach founded on *complexity* (Kolmogorov 1963), the entropy of the initial part of a random sequence should be sufficiently large.

The main idea of the *quantitative* standpoint (Martin-Löf, 1966) is that a random sequence may only have a small number of regularities, and, therefore, that it should pass certain tests. It is easy to see that these approaches are not independent.

In 1963 Kolmogorov additionally outlined the concept of a *finite random sequence*; according to his opinion, a finite sequence is the more random the more complex is the law that describes it. I happened to hear that that concept was not generally recognized.

Quite recently there appeared another Russian paper [23] on the same subject with no reference being provided to the previous one. As in the case of ref. [22], Uspensky was its co-author, but this time, the second, or, rather, the first co-author was Kolmogorov himself.

3. Randomness does not exist

Such was the standpoint of the most eminent thinkers and scholars who believed that a semblance of randomness resulted from ignorance of the relevant causes. Sambursky [9, pp. 40 – 41] described the utterances of ancient Greek authors on that subject, and Kendall [8, p. 11] studied similar ideas due to St. Augustine, Thomas Aquinas, Spinoza and D'Alembert. In turn, I discuss the thoughts of several scientists without dwelling on the writings of Bentley [24, pp. 316 – 318], who somewhat verbosely explicated Newton's point of view, or Lamarck [25, pp. 74 and 97; 26, p. 329]. Here are the statements of

Kepler [27], Laplace [28, p. 145] and Darwin [29, p. 128], in that order.

1. Chance is an idol, an abuse of God Almighty.

2. Chance is only ignorance of the connections between phenomena.

3. That chance occasions variations between individuals is wrong, but this expression *serves to acknowledge* […] *our ignorance* of the relevant causes.

Kepler, however, was unable to deny that the eccentricities of the planetary orbits were random (§ 5). Newton left two pronouncements [30, Query 31; 31, p. 49] which testify that he attached certain importance to chance and to which I return in §§ 7 and 8:

Blind chance could never make all the planets move one and the same way in orb concentrick, some inconsiderable irregularities excepted, which may have risen from the mutual actions of comets and planets upon one another, and which will be apt to increase, till this system wants a [divine] *reformation. Such a wonderful uniformity in the planetary system must be allowed the effect of choice. And so must the uniformity in the bodies of animals*.

Did blind chance know that there was light and what was its refraction, and fit the eyes of all the creatures after the most curious manner to make use of it?

A similar utterance is in [32, p. 544], and another one, formulated in about 1715, in [33, p. 127].

Lamarck [34, p. 450] thought that variations between individuals came into existence because of random causes and the (Darwinian) theory hinged in its entirety on the action of these same causes. It is extremely strange that, in spite of his own statistical explanation of the second law of thermodynamics, Boltzmann failed to recognize either the latter fact or the importance of randomness in nature [35, § 4.3]. Finally, I return to Laplace in § 5.

4. A possibility

Randomness is a possibility. This definition goes back to Aristotle [36, 1064b – 1065a], who, moreover, apparently believed that a chance event had a logical or subjective probability lower than 1/2. Similarly, Thomas Aquinas [37, vol. 19, p. 297] supposed that random events *proceed from their causes in the minority of cases* …

The followers of the Indian teaching of Syadvada, that existed as early as in the 6^{th} century B C, studied the concepts of the possible, the indeterminate, etc. Mahalanobis [38] maintained that this doctrine was interesting for the history of statistics. He had not mentioned randomness, but I believe that the Syadvada indirectly recognized it as a possibility.

Darwin [39, vol. 1, p. 449], drawing on stochastic calculations made at his request by Stokes, decided that a particular deformity in man was passed from parent to child and did not occur by chance [was not merely possible]. William Herschel [40, p. 577] and Struve [41, Note 72] left room for randomness of this kind. In their models of the stellar system; they only restricted the distances of the stars without indicating their actual position. Maxwell [42, p. 274] remarked that neither the form and dimensions of the planetary orbits, nor the size of the Earth were determined by any law of nature [that the relevant magnitudes might have been different]. He had not mentioned randomness and his remark had to do with yet another interpretation of chance (§ 6).

Hegel [43, p. 383], in addition to understanding randomness as a possibility, formulated the converse proposition:

Das Zufällige ist ein Wirkliches, das zugleich nur als möglich bestimmt […] *was möglich ist, ist selbst ein Zufälliges*.

It is easy to illustrate this proposition. If a random variable X takes values x_i with probabilities p_i, $i = 1, 2, \ldots, n$, then any possible x_i is random in a sense that it occurs with probability p_i. Note that Aristotle had not connected any definite probabilities with the possible values of x_i, $i = 1, 2, \ldots$

5. Deviation from laws of nature

Randomness occurs when the purpose of nature is not attained, when hindering causes corrupt the operations of nature. This explanation is due to Aristotle [44, 199b] who thought that nature's accidental mistakes brought about the appearance of monsters and that the birth of female animals was the first departure from the type, and, at the same time, a natural necessity [45, 767b]. His statements were the first to confront necessity and randomness. Indeed, the occurrence of monsters accompanies the necessary acts of regular births whereas the birth of a female is, according to Aristotle, both necessary and random. From a modern point of view the second example is partly wrong, and hardly corresponds to his own belief (§ 4) that the probability of a chance event is lower than 1/2. Referring to the Philosopher, Thomas Aquinas [37, vol. 19, p. 489] pointed out that the birth of a girl was a random event.

Kepler [46, p. 244; 47, p. 932] suggested that only *zufällig* perturbations had forced the planets to deviate from circular motion. True, he also stated that the eccentricities regulated the planets' motions [48, p. 317], but he was naturally unable to say why the eccentricity of a given planet had a particular value rather than any other one. Kant [49, p. 337] repeated Kepler's pronouncement on the elliptic paths of the planets. Lamarck [26, p. 133] maintained that there existed deviations from the divine lay-out of the tree of animal life and explained them by the action of a *cause accidentelle et par conséquent variable*.

The pronouncements described above pertained to determinate laws of nature. However, many natural scientists, while making similar statements, actually thought about mean states. Adanson [50, p. 48] regarded intraspecific variations as digressions from the divine order and believed them necessary *pour l'équilibre des choses*. Lamarck [51, p. 76] argued that *plusieurs causes*, some of them *variables, inconstantes et irrégulières dans leur action*, corrupted [determined!] the [mean] state of the atmosphere. Humboldt [52, p. 68] conditioned the study of all natural phenomena by discovering the appropriate mean values (mean states).

As early as in 1817 he isolated climatology from meteorology [53]. His point of view was not, however, quite consistent in that he had not linked his definition of climate [54, p. 404] with mean states, but at least later scholars improved on him [55, p. 296].

De Moivre [56, p. 253] declared that the value of the parameter of the binomial distribution of male and female births was of divine origin. Quite logically, he regarded as random only the deviations of the number of male (say) births from the corresponding number determined by the binomial law. Random, in modern notation, for De Moivre was not X itself, but rather (X – EX). He (Ibidem, p. 251) also argued that

In process of Time, Irregularities [produced by chance] *will bear no proportion to the recurrency of that Order which naturally results from Original Design*.

And he [57, p. 329] effectively declared that the aim of the theory of probability was to isolate chance from divine design [from purpose], and thus came close to another understanding of randomness (§ 6).

Being greatly influenced by Newton, to whom he devoted the first edition of his book [57], De Moivre had not nevertheless repeated the former's inference on the need for divine reformation (§ 3).

Similarly, for Laplace the theory of probability belonged to natural sciences rather than to mathematics, and its goal was not to study mathematical objects (for example, densities), but the discovery of the laws of nature. He therefore stood in need of analysing observations, of eliminating randomness from them, of separating chance from law.

6. Lack of purpose

Randomness is lack of divine law or goal; it occurs when independent chains of events intersect each other. Again, randomness is lack of purpose, and perhaps, *uniformity* (§ 7) as well. It was in this sense that chance was understood in ancient India, about two thousand years ago, although not in natural sciences, but in civil life [1, § 108]:

If, shortly after giving evidence at a trial, a misfortune befell the witness or his family, it was believed that God punished him [that the evil had not happened without purpose, i.e., not by chance].

6.1. Lack of law or goal. According to Aristotle, an unexpected meeting of two people [36, 1025a] or a discovery of a buried treasure [44, 196b] are chance events. Each of them could have been (but was not) aimed at. Junkersfeld [7, p. 22], who considered numerous examples contained in the great scientist's work, inferred that he would not have thought that coming across a stranger or finding a rusty nail were random.

The ancient Indian Yadrichchha or Chance theory contained a similar interesting illustration of randomness [58, p. 458]:

The crow had no idea that its perch would cause the palm-branch to break, and the palm-branch had no idea that it would be broken by the crow's perch; but it all happened by pure chance.

These examples show that the interpretation of chance as an intersection of chains of events was known even in antiquity. In this connection Cournot [16, p. 56] had quoted Boethius and Bru [59, p. 306] noticed that Cioffari [60, pp. 77 – 84] had discussed or reproduced appropriate passages from several ancient scholars.

Hobbes [61, p. 259] maintained that a traveller *meets with a shower* by chance since *the journey caused not the rain, nor the rain the journey*.

Much the same was the opinion of many modern scientists [5, p. 133] and of course in any reasoning of this kind the interpretation mentioned above simply suggests itself.

Darwin [62, p. 395] argued that he had used the word *chance* only in relation to purpose [to lack of purpose] in the origination of species. He continued: *the mind refuses to look at* [the universe] *as the outcome of chance, – that is, without design or purpose*.

The D'Alembert – Laplace problem merits special attention. The word *Constantinople* is composed of separate letters; is it possible that the choice and arrangement of the letters were random? D'Alembert [63, pp. 245 – 255], who questioned the fundamentals of the theory of probability, maintained that all arrangements of the letters were equally probable only from the mathematical point of view but not in reality. Laplace [28, p. 152; 54, p. XV] came to a different conclusion: Since the word had a certain meaning [answered a particular purpose], the composition was not likely at all to have been accidental [aimless].

This reasoning helps to understand properly a number of earlier pronouncements. Aristotle [65, 289b] believed that it was impossible for the stars to move independently one from another [to move at random] and yet to remain fixed, – they possessed common motion. A similar idea can be traced in the theory of errors. A large deviation of an observation from the appropriate arithmetic mean had rather been assigned to a special reason (though not to a goal, or a law, but to a blunder) than attributed to an unlikely combination of admissible and mutually independent [accidental] errors. Note, however, that observational errors hardly belong to natural sciences.

Kepler [66, p. 337] thought that a possible (a chance, see § 4) appearance of a new star in a definite place and on a particular date was so unlikely that it had to be occasioned on purpose. By implication, he believed that each place (and date) was equally probable. Thus, Kepler understood randomness not only as lack of purpose, but as something [aimlessly] possible (§ 4), and, at the same time, as uniform (§ 7).

6.2. Intersection of chains of events. Randomness is an intersection of such chains. This interpretation is due to La Placette [67, last page of Preface] who devoted his book to proving that games of chance were not contrary to Christian ethics. He contended that

Le Hasard renferme […] *un concours de deux, ou de plusieurs événements contingents*.

Each event had its own cause, the author continued, but we did not know why they coincided. La Placette had not explained randomness; his definition amounted to saying that the cause of any chance event was unknown (cf. § 3).

Cournot [59, § 40; 16, p. 52] took up La Placette's idea and in one instance [16, p. 57] referred to him. He [59] initially mentioned chains of determinate events thus improving on his predecessor:

Les événements amenés par la combinaison ou la rencontre de phénomènes qui appartiennent à des séries indépendantes, dans l'ordre de la causalité, sont ce qu'on nomme des événements fortuits.

In his later work Cournot [16] regrettably omitted the phrase *dans l'ordre de la causalité*. He [59, §§ 41 – 48] apparently thought of using his definition of randomness to present the theory of probability as a science of chance events. He could not have succeeded; what was really needed was a systematic use of the notions of random variable (cf. § 2) and, therefore, of its expectation and variance.

7. Uniformity

Randomness is something uniformly possible, it can occur in one out of several equally possible ways.

7.1. Uniform randomness. In § 6.1 I stated that Kepler had equated chance with uniform randomness. This attitude was characteristic of natural scientists for about two centuries. Arbuthnot [68], in attempting to explain the prevalence of boys among the newborn, contrasted uniform randomness and design without thinking of other possible laws of randomness. The same kind of comparison is implied in both of Newton's pronouncements (§ 3).

Jakob and Niklaus Bernoulli and De Moivre introduced the binomial distribution into the theory of probability; in spite of that, the former understanding of randomness persisted. Boyle [69, p. 43], indicated that a chance composition of a long sensible text was impossible and declared that the world could not have been created randomly. The first part of his statement is also contained in the *Logique de Port-Roya*l [70, Chapt. 16].

Kant [49, p. 230] and Voltaire [71, p. 316] maintained that a uniformly random origin of organic life was even less possible than a similar origin of the system of the world. Daniel Bernoulli [72] and Laplace [73], likely following Newton, calculated the probability that the regularities observed in the Solar system were due to randomness but they only contrasted blind chance and a determinate cause.

Maupertuis [74, pp. 120 – 121] indicated that the seminal liquid *de chaque individu* most often contained parties similar to those of its parents. He also mentioned rare cases when a child resembled one of his remote ancestors (p. 109) as well as mutations (a later term) (p. 121). It could be thought that Maupertuis recognized randomness with a multinomial distribution, but he was not consistent. While discussing the origin of eyes and ears in animals, he [75, p. 146] restricted his attention to comparing *un attraction uniforme & aveugle* and *quelque principe d'intelligence* (and came out in favour of design).

In the 19th century, many scientists, imagining that randomness was only uniform, refused to recognize the evolution of species. While illustrating that idea, both the astronomer John Herschel [76, p. 63] and the biologist Baer [77, p. 6] mentioned the philosopher depicted in the Gulliver's *Travels* [but borrowed by Swift from Raymond Lully, 13th – 14th centuries]. Hoping to get to know all the truths, that good-for-nothing inventor was putting on record each sensible chain of words that happened to appear among their uniformly random arrangements.

Also in the 19th century, Boole [78, p. 256] argued that the distribution of stars was random, if, owing to the ignorance of the relevant law, *it would appear to us as likely that a star should occupy* *one spot of the sky as another* (cf. § 3). And he continued: *Let us term any other principle of distribution an indicative one*. Even in 1904 Newcomb [79, p. 13] called the uniform distribution of stars *purely accidental*. Recalling the definition of a finite random sequence as outlined by Kolmogorov (§ 2), and bearing in mind that the number of stars of the first few magnitudes is finite, I note, however, that Boole's and Newcomb's inferences were quite modern.

The following examples that have to do with finite populations of stars or atoms are similar. Nevertheless, in these instances natural scientists reasonably believed that uniform randomness represented a statistical law of nature. Thus, Forbes [80, p. 49] contended that

An equable spacing of stars […] [is] *far more inconsistent with a total absence of Law or Principle, than the existence of* [regions of condensation and paucity] *of stars*.

He [80, 1850, p. 420] also asked which distributions might be called random [as not representing any law, cf. § 6.1].

In 1906 Kapteyn [81, p. 400] declared that

The peculiar motions of the stars are directed at random, that is, they show no preference for any particular direction.

Struve [82, pp. 132 – 133] pronounced a similar weaker statement even in 1842. Boltzmann [83, p. 237; 84, p. 321] held that gas molecules move with equal probability in whichever direction, but did not mention randomness.

Sometimes chance might have been connected with the state of chaos, i.e., with the absence of any law of distribution. Since this possibility was hardly discussed before the 19th century, I believe that either no one considered it, or, in any case, that it gradually gave way, perhaps unjustly, to uniform randomness. In those times, apparently only De Moivre [56, pp. 251 – 252] mentioned chaos, but even he dismissed it out of hand: *Absurdity follows*, he declared, while considering one or another value of the parameter of the binomial distribution, if a certain event happened not

According to any law but in a manner altogether desultory and uncertain; for then the Events would converge to no fixt Ratio at all.

And, when introducing his definition of probability as the limit of statistical frequency, von Mises [19, p. 60] effectively excluded chaos.

Against the background of the abovementioned examples, it is interesting to name two philosophers of the 18th century who expressly indicated that non-uniform randomness was indeed possible. Hume [85, vol. 1, p. 425], while discussing chance events, illustrated his ideas by considering an imaginary die having *four sides marked with a certain number of spots, and only two with another*. He had not however referred to any law of nature. D'Holbach [86, pt. 2, pp. 138 – 139] maintained that the molecules of various bodies greatly differed one from another and combined with each other in diverse ways. He compared them with *dice pipées* […] *d'une infinite de façons différentes* [with irregular dice].

7.2. Specifying particular problems. During the 19th century, it gradually became clear that the concept of uniform randomness in general was not sufficiently intelligible. The problem of determining the distance between two random points (A and B) on a sphere is highly relevant since Laplace [87, p. 261] and Cournot [59, § 148] understood it in different senses. Laplace believed that B was, with equal probability, any point of the great circle AB whereas Cournot's solution implied that equally probable were all possible situations of B on the sphere. Similarly, Daniel Bernoulli had calculated the probability that the planes of the planetary orbits were close to each other due to uniform randomness (cf. § 6.1), but Todhunter [88, § 396] remarked that it would have been more natural to consider uniform randomness with respect to the closeness of the poles of the orbits.

Darwin [89, pp. 52 – 55] attempted to ascertain whether earth worms carrying small objects into their burrows seize *indifferently by chance* any part of their find. He considered four versions of such randomness with regard to the manner of capturing paper triangles strewn about on the ground. After calculating the appropriate frequencies, Darwin decided that the worms carried the triangles non-randomly, i.e., to a certain extent sensibly. Considering non-randomness on a par with reason, he therefore recognized chance as lack of purpose; in § 6.1 I have mentioned him exactly in this connection.

Bertrand [90, pp. 6 – 7] took up the problem of calculating the distance between random points on a sphere. Without mentioning Laplace or Cournot, he repeated their solutions and concluded that both were correct. In addition, Bertrand maintained that not only small distances but other geometric features as well might be used to characterize an unlikely scatter of the stars over the sky. He hardly knew about Darwin's experiment, but he provided a few more examples including his celebrated problem on the probability of the length of a random chord of a given circle. He thus proved that uniform randomness was not definite enough and justly insisted that in particular instances that concept be specified.

8. Instability of motion

Randomness is instability of motion, or of initial conditions; it involves slight causes leading to considerable consequences. Galen [91, p. 202], without mentioning randomness, asserted that *in old men even slightest causes produce the greatest change*. According to Newton (§ 3), the accumulation of irregularities in the planetary system may be interpreted as an action of slight causes giving rise to considerable effects (true, only gradually). Many examples from § 6.1 can be also considered in this connection.

Maxwell [92, p. 366] prophetically argued that physicists will study *irregularities and instabilities* and thus move away from mere determinacy. Illustrating his idea, he mentioned unstable refraction of rays within biaxial crystals (p. 364). Maxwell thus connected randomness with instability but had not said so directly. He expressed similar thoughts elsewhere [93, pp. 295 – 296]:

There is a very general and very important problem in Dynamics […] *It is this*:

Having found a particular solution of the equations of motion of any material system, to determine whether a slight disturbance of the motion indicated by the solution would cause a small periodic variation, or a total derangement of the motion …

Von Kries [94, p. 58], while discussing the game of roulette, noted that

Eine kleine Variirung der Bewegung hinreichend, um an Stelle des Erfolges Schwarz der Erfolg Weiß herbeizuführen …

His remark was not, however, convincing: the slight variation of the motion could have resulted, first and foremost, in changing the number of rotations travelled by the ball.

Pirogov [95, p. 518] called an event random if its dependence on the relevant causes was complicated and

Mit Hülfe von nur analytischen Funktionen gar nicht ausgedrückt werden kann.

His utterance may be considered as another hint at the connection between chance and instability. As to complicated causes, see § 9.

As stated in § 1, I am not discussing the work of Poincaré, but at least I emphasize that he was the first to say expressly that randomness is instability of motion.

9. Complicated Causes

Randomness occurs when complicated causes are involved. In a heuristic sense Leibniz [96, p. 288] anticipated this explanation by declaring that the *zufällige Dingen* were those

Deren vollkommener Beweis jeden endlichen Verstand überschreitet.

While formulating his celebrated law of the velocities of gas molecules, Maxwell [97] reasonably supposed that the distribution sought sets in

After a great number of collisions among a great number of equal particles.

He had not mentioned randomness. Elsewhere he [98, p. 436] remarked that the motion of heat was *perfectly irregular* and that the velocity of a given molecule could not be predicted. Once more, he did not mention randomness, and he said noting about complicated causes. I have adduced his second pronouncement since it supplements his previous idea. Note that he actually rejected Laplace's famous declaration [64, p. VI] on the possibility of calculating the future states of the universe.

10. Slight causes leading to small effects

Randomness occurs when slight causes lead to small effects. Laplace [99, p. 504] qualitatively explained the existence of trifling irregularities in the system of the world (the different eccentricities of the planetary orbits) by the action of countless [small] differences between temperatures and between densities in the diverse parts of the planets. He had not mentioned randomness. Kepler and Kant (§ 5) referred in similar cases to deviation from purpose.

Laplace was damnably wrong: Newton had explained those differences by the differences of planetary velocities.

Under certain circumstances the same cause leads to chaotic movement not yet studied in the 19^{th} century.

11. Necessity and randomness

In discovering laws and regularities of nature and in studying its mean states, scientists determined necessity. Besides that, they often revealed, or even attempted to isolate, the unavoidable accompanying phenomena of the second order, i.e. randomness. And it was exactly in this manner that many natural scientists imagined the relation between necessity and chance. Recall in this connection Aristotle's opinion (§ 5) on the appearance of monsters, Kepler's reasoning on the eccentricities of the planetary orbits (§ 5), Newton's thoughts (§ 3, also see below) on the planetary system, Lamarck's utterance (§ 5) on the tree of animal life, De Moivre's reasoning (§ 5) on the sex ratio at birth as well as the isolation of climatology from meteorology achieved by Humboldt (§ 5) and William Herschel's and Struve's models of the stellar system (§ 4).

Lamarck's pronouncement [26, p. 169] merits special attention. He apparently believed that necessity and chance were the two main *moyens* of nature. Without proving anything or providing any example, he declared that these *moyens puissans et généraux* were universal attraction and a repulsive molecular action *qui* […] *varie sans cesse* … He also argued that the

Equilibre entre ces deux forces opposées […] *naissent* […] *les causes de tous les faits que nous observons, et particulièrement de ceux qui concernent l'existence des corps vivans*.

Lamarck likely supposed that the molecular action was random since elsewhere (see § 5) he maintained that by definition accidental causes were variable.

Without dwelling on the statistics of marriages, suicides, crime, etc. that reveals laws in apparently free (random] behaviour of man, I note that Kant [100, p. 508] compared the chance birth of a man with the stability of the birth-rate:

Der Zufall im Einzelnen nichts desto weniger einer Regel im Ganzen unterworfen ist …

Only Hegel, after offering his definition of randomness (§ 4), formulated a proposition on the unity [on the interdependence] between necessity and chance. Exactly this unity, he [43, p. 389] declared, *Ist die absolute Wirklichkeit zu nennen*. Engels [101, p. 213] approvingly called this thesis utterly unheard of and urged scientists to study both necessity and chance. It was Poincaré [102, p. 1], however, who provided the most important statement:

Dans chaque domaine, les lois précises ne décidaient pas de tout, elles tracaient seulement les limites entre lesquelles il était permis au hasard de se mouvoir. Dans cette conception, le mot hasard avait un sens présis, objectif …

A few words about the theory of probability. At the end of § 5 I mentioned De Moivre and Laplace in connection with the aim of that scientific discipline. They entrusted the theory with delimiting randomness from necessity. In our days, the same goal is being achieved by mathematical statistics created since then. Pearson [103] remarked that the development of the theory of probability was much indebted to Newton; I shall show that he thought about the great scientist's idea on the relation between necessity and chance. Here are his words:

Newton's idea of an omnipresent activating deity, who maintains mean statistical values, formed the foundation of statistical development through Derham, Süssmilch, Niewentyt, Price to Quetelet *and Florence Nightingale.* […] *A. De Moivre expanded the Newtonian theology and directed statistics into the new channel down which it flowed for nearly a century. The causes which led De Moivre to his* Approximatio [56] [where the normal approximation to the binomial distribution was first discovered] *or Bayes to his theorem were more theological and sociological than purely mathematical, and until one recognizes that the post-Newtonian English mathematicians were more influenced by Newton's theology than by his mathematics, the history of science in the 18th century, – in particular that of the scientists who were members of the Royal Society – must remain obscure*.

Since Newton never mentioned the maintaining of mean values, I believe that Pearson actually thought about divine reformation, necessary, according to Newton (§ 3), for neutralizing the propagation of chance corruptions in the Solar system, for preserving the mean states. Thus, Pearson suggested that Newton's theologically formulated idea concerning the relation between necessity and chance had served as a basis for the development of the theory of probability.

Pearson's general statement about the science in the 18th century may be specified. First, he apparently bore in mind Laplace (end of § 5 and above); second, restricting my attention to the theory of probability, I note that Pearson [104, §§ 10.1 – 10.2] put forward plausible arguments in favour of the thesis on Newton's influence on Bayes (and Price, who communicated and inserted comments in the Bayes memoir).

12. Conclusions

The denial of randomness (§ 3) was only formal and nowadays seems to be deservedly forgotten. Possibility (§ 4) found its way into laws and empirical regularities, but it was Hegel who declared that randomness was a possibility, and, moreover, that the possible was random. Chance as deviation from laws of nature (§ 5) is recognized as a perturbation (a noise) and natural scientists admitted that it indeed was corrupting the laws.

As far as the deviations obey the preconditions of the central limit theorem, this randomness is normal. Randomness as lack of law or purpose (§ 6) may be interpreted as an intersection of independent chains of events.

The definitions of §§ 4 and 6, while reflecting different heuristic features of randomness, essentially coincide.

Randomness is a random variable having a uniform distribution (§ 7), i.e., it is a special case of the possible (§ 4). Therefore, this uniform randomness characterizes lack of determinate law or purpose (§ 6); at the same time, in some instances it signifies the existence of a special statistical law of nature.

Randomness is occasioned by instability (§ 8) and/or complicated causes (§ 9). It can also occur in the context of slight causes leading to slight effects (§ 10). This case partly includes deviations from the laws of nature (§5), as in meteorology and astronomy. The joint action of a large number of such causes can lead to random variables with a normal distribution (above).

It is scarcely possible to comprehend randomness without studying its interconnection with necessity. Hegel stated that these concepts were united. However, even after Hegel scientists had been recognizing randomness only as a phenomenon of the second order accompanying the main event, necessity (§ 11). Until the mid-19th century necessity (divine design) had been contrasted only with blind chance (uniform randomness, § 7).

The explanations and definitions of chance (§§ 4 – 10) are heuristically connected with the modern interpretations of randomness (§ 2). Thus, § 9 is closely linked with the complexity approach and to a lesser degree a similar link seems also to apply to § 8; § 7 illustrates a particular instance of the frequentist approach and the rest of these sections at least do not contradict the quantitative approach. Finally, I note that §§ 4 – 6 and 10 are linked with § 11.

Acknowledgement. B. Bru, B. V. Chirikov and V. V. Nalimov sent me reprints/copies of their contributions or other materials. M. V. Chirikov, a relative of B. V., pointed out a number of mistakes and ambiguities in a preliminary text of this paper. A. Kozhevnikov and E. Knobloch gave me editorial advice, and A. P. Youshkevich counselled me to explain the modern point of view on randomness (§ 2).

References

AHES = *Arch. Hist. Ex. Sci.*
M, Psb = Moscow, Petersburg

1. *Laws of Manu*. Oxford, 1886. (1991.)
2. Chirikov, B. V. The nature of the statistical laws of classical mechanics. In *Metodologicheskie i Filosofskie Problemy Fiziki* (Methodological and Philosophical Issues of Physics). Novosibirsk, 1982, pp. 181 – 196.
3. Sheynin, O. B. Poincaré's work on probability. AHES, vol. 42, 1991, pp. 137 – 171.
4. Khinchin, A. Ya. On the Mises frequentist theory; posth. publ. by B. V. Gnedenko, 1961, in Russian. *Science in Context*, vol. 17, 2004, pp. 391 – 422.
5. Sheynin, O. B. On the prehistory of the theory of probability. AHES, vol. 12, 1974, pp. 97 – 141.
6. Nalimov, V. V., *Yazyk Veroyatnostnykh Prestavleniy* (Language of Stochastic Concepts). M., 1976.
7. Junkersfeld, J. *The Aristotelian – Thomistic Concept of Chance*. Notre Dame, Indiana, 1945.
8. Kendall, M. G. The beginnings of a probability calculus. *Biometrika*, vol. 43, 1956, pp. 1 – 14. Reprint: E. S. Pearson, M. G. Kendall, *Studies in History of Statistics and Probability.* London, pp. 19 – 34.
9. Sambursky, S. On the possible and probable in ancient Greece. *Osiris*, vol. 12, 1956, pp. 35 – 48. Reprint: Ibidem, pp. 1 – 14.
10. David, F. N. *Games, Gods and Gambling*. London, 1962.
11. Byrne, E. F. *Probability and Opinion*. The Hague, 1968.
12. Rabinovitch, N. L. *Probability and Statistical Inference in Ancient and Medieval Jewish Literature*. Toronto, 1973.
13. Gnedenko, B. V. *Iz Istorii Nauki o Sluchainom* (From the History of the Science of the Random). M., 1981.
14. Sheynin, O. B. Newton and the classical theory of probability. AHES, vol. 7, 1971, pp. 217 – 243.
15. Sheynin, O. B. Lambert's work in probability. Ibidem, pp. 244 – 256.
16. Cournot, A. A. *Essai sur les fondements de nos connaissances*, t. 1. Paris, 1851. (Paris, 1975.) **S, G,** 54.
17. Chuprov, A. A. *Ocherki po Teorii Statistiki* (Essays on the Theory of Statistics) (1909). M., 1959.

18. Poisson, S.-D. *Recherches sur la probabilité des jugements.* Paris, 1837 (and 2003. **S, G,** 53.)
19. Sheynin, O. B. *Chuprov*: *Life, Work, Correspondence* (1990, in Russian). Göttingen, 1996 and 2011.
20a. Mises, R. von, Grundlagen der Wahrscheinlichkeitsrechnung (1919). *Sel. Papers*, vol. 2. Providence, RI, 1964, pp. 57 – 105.
20b. Mises, R. von, *Wahrscheinlichkeit, Statistik und Wahrheit* (1928, 1936, 1951). English translation: 1939, 1964.
21. Church, A. On the concept of a random sequence. *Bull. Amer. Math. Soc*., vol. 46, 1940, pp. 130 – 135.
22. Uspensky, V. A., Semenov, A. L. *Teoriya Algoritmov* etc. (Theory of Algorithms: Main Discoveries and Applications). M., 1987.
23. Kolmogorov, A. N., Uspensky, V. A. Algorithms and randomness. *Teoria Veroiatnostei i Ee Primenenia*, vol. 32, 1987, pp. 425 – 455. (R)
24. Bentley, R. *Sermons* (1693). In Newton, I. *Papers and Letters on Natural Philosophy*. Cambridge, 1958.
25. Lamarck, J. B. *Aperçu analytique des connaissances humaines* (MS, 1810 – 1814). In Vachon, M. et al, I*nédits de Lamarck*. Paris, 1972, pp. 69 – 141.
26. Lamarck, J. B. *Histoire naturelle des animaux sans vertèbres*, t. 1. Paris, 1815.
27. Kepler, J. *De stella nova* (1606). German transl.: Würzburg, 2006.
28. Laplace, P. S. Recherches sur l'intégration des équations différentielles (1776). *Oeuvr. Compl*., t. 8. Paris, 1891, pp. 69 – 197.
29. Darwin, C. *Origin of Species* (1859). London – New York, 1958. (Manchester, 1995.)
30. Newton, I. *Optics* (1704). *Opera*, vol. 4. London, 1782, pp. 1 – 264. (London, 1931.)
31. Newton, I. *Theological Manuscripts*. Liverpool, 1950.
32. Newton, I. *Mathematical Principles of Natural Philosophy*. Cambridge, 1934. Revised reissue of the edition of 1729.
33. Manuel, F. E. *A Portrait of Newton.* Cambridge (Mass.), 1968.
34. Lamarck, J. B. Espèce. *Nouv. Dict. Hist. Natur*., t. 10, 1817, pp. 441 – 451.
35. Sheynin, O. B. On the history of the statistical method in physics. AHES, vol. 33, 1985, pp. 351 – 382.
36. Aristotle, *Metaphysica. Works*, vol. 8. Oxford.
37. Aquinas, Thomas, *Summa theologica*. English transl.: *Great Books of the Western World*, vols. 19 – 20. Chicago, 1952.
38. Mahalanobis, P. C. The foundations of statistics. *Sankhya, Ind. J. Stat*., vol. 18, 1957, pp. 183 – 194.
39. Darwin, C. *The Variation of Animals and Plants under Domestication*, vols 1 – 2. London, 1885. (London, 1988.)
40. Herschel, W. Astronomical observations and experiments etc. (1817). *Scient. Papers*, vol. 2. London, 1912, pp. 575 – 591. (London, 2003.)
41. Struve, F. G. W. *Etudes d'astronomie stellaire*. Psb, 1847.
42. Maxwell, J. C. Discourse on molecules (MS, 1873). In Campbell, L., Garnett, W. *Life of Maxwell.* London, 1882. London, 1884, pp. 272 – 274. (New York, 1969.) (Maxwell, *Sci. Papers*. New York, 1965. *Sci. Letters and Papers*, vols 1 – 3. Cambridge, 1990 – 2002.)
43. Hegel, G. W. F. *Wissenschaft der Logik*, Tl. 1 (1812). Hamburg, 1978.
44. Aristotle, *Physica. Works*, vol. 2. Oxford.
45. Aristotle, *De generatione animalium. Works*, vol. 5. Oxford.
46. Kepler, J. *Neue Astronomie* (1609, in Latin). München – Berlin, 1929. (English translation: Cambridge, 1992, 2015.)
47. Kepler, J. *Epitome of Copernican Astronomy*, book 4 (1620, in Latin). In *Great Books of the Western World*, vol. 16. Chicago, 1952, pp. 845 – 1004 (books 4 and 5).
48. Kepler, J. *Welt-Harmonik* (1619, in Latin). München – Berlin, 1939. (English translation: Philadelphia, 1997.)
49. Kant, I. *Allgemeine Naturgeschichte und Theorie des Himmels (*1755). *Werke*, Bd. 1. Berlin, 1910, pp. 215 – 368.

50. Adanson, M. Examen de la question si les espèces changent parmi les plantes. *Hist. Acad. Roy. Paris avec Mém. math. et phys*., 1769 (1772), pp. 31 – 48 of the *Mémoires.*
51. Lamarck, J. B. *Annuaire météorologique* [t. 1]. Paris, pour l'an 8 (1800).
52. Sheynin, O. B. On the history of the statistical method in meteorology. AHES, vol. 31, 1984, pp. 53 – 95.
53. Humboldt, A. Des lignes isothermes. *Mém. Phys. Chim. Soc. d'Arcueil*, t. 3, 1817, pp. 462 – 602.
54. Humboldt, A. *Fragmens de géologie et de climatologie asiatiques*, t. 2. Paris, 1831.
55. Koerber, H. G. Über Humboldts Arbeiten zur Meteorologie und Klimatologie. In Humboldt, *Gedenkschrift*. Berlin, 1959, pp. 289 – 335.
56. De Moivre, A. A method of approximating the sum of the terms of the binomial etc. (1733, in Latin). [57, 1756, pp. 243 – 254].
57. De Moivre, A. *Doctrine of Chances* (1718, 1738, 1756). New York, 1967 (reprint of the edition of 1756).
58. Belvalkar, S. K., Ranade, R. D. *History of Indian Philosophy*, vol. 2. Poona, 1927.
59. Cournot, A. A. *Exposition de la théorie des chances et des probabilités* (1843). Paris, 1984. Editor, B. Bru. **S, G,** 54.
60. Cioffari, V. *Fortune and Fate from Democritus to St. Thomas Aquinas*. New York, 1935.
61. Hobbes, T. *Of liberty and necessity* (1646). *Engl. Works*, vol. 4. London, 1840, pp. 229 – 278.
62. Darwin, C. *More Letters*, vol. 1. London, 1903.
63. D'Alembert, Le Rond J. Doutes et questions sur le calcul des probabilités. In author's *Mélanges de litterature, d'hist. et de philos*., t. 5. Amsterdam, 1768, pp. 239 – 264.
64. Laplace, P. S. *Essai philosophique sur les probabilités* (1814). *Oeuvr. Compl.,* t. 7, No. 2. Paris, 1886. Separate paging. (*Philosophical Essay on Probabilities*. New York, 1995.)
65. Aristotle, *De caelo. Works*, vol. 2. Oxford.
66. Kepler, J. *A Thorough Discussion of an Extraordinary New Star* (1604, in German). *Vistas in Astronomy*, vol. 20, 1977, pp. 333 – 339.
67. La Placette, J. *Traité des jeux de hasard*. La Haye, 1714.
68. Arbuthnot, J. An argument for divine Providence etc. (1712). In *Studies in the History of Statistics and Probability*, vol. 2. Editors, Sir Maurice Kendall, R.L. Plackett. London, 1977, pp. 30 – 34.
69. Boyle, R. Some considerations touching the usefulness of experimental natural philosophy (1663 – 1671). *Works*, vol. 2. London, 1772, pp. 36 – 49. (Bristol, 1999.)
70. Arnauld, A., Nicole, P. *Logique de Port-Royal* (1662). Paris, 1877. (English translation: Edinb. – London, 1850; Cambridge, 1996.)
71. Voltaire, Homélies. Première homélie (1767). *Oeuvr. Compl*., t. 26. Paris, 1879, pp. 315 – 354.
72. Bernoulli, D. Recherches physiques et astronomiques etc., 1735. *Werke*, Bd. 3. Basel, 1987, pp. 303 – 326.
73. Laplace, P. S. Sur l'inclinaison moyenne des comètes (1776). *Oeuvr. Compl*., t. 8. Paris, 1891, pp. 279 – 321.
74. Maupertuis, P. L. M. *Venus physique* (1745). *Oeuvr*., t. 2. Lyon, 1756, pp. 1 – 133.
75. Maupertuis, P. L. M. *Système de la nature* (1751). Ibidem, pp. 135 – 184.
76. Herschel, J. F. Sun (lecture, 1861). In author's *Familiar Lectures on Scient. Subjects*. London – New York, 1866, pp. 47 – 90.
77. Baer, K. *Zum Streit über den Darwinismus*. Dorpat (Tartu), 1873.
78. Boole, G. On the theory of probabilities (1851). In author's *Studies in Logic and Probability*. London, 1952, pp. 247 – 259.
79. Newcomb, S. *On the Position of the Galactic*. *Carnegie Instn of Washington*, No. 10, 1904.
80. Forbes, J. D. On the alleged evidence for a physical connection between stars. *London, Edinb. and Dublin Phil. Mag*., vol. 35, 1849, pp. 132 – 133; vol. 37, 1850, pp. 401 – 427.

81. Kapteyn, J. C. Statistical methods in stellar astronomy. [Repts] *Intern. Congr. Arts and Sci. St. Louis – Boston 1904*. N. p., vol. 4, 1906, pp. 396 – 425.
82. Struve, F. G. W. Review of O. Struve, *Bestimmung der Constante der Präcession*. Psb, 1843. *Bull. Scient. Acad. Imp. Sci. Psb*., t. 10, No. 9 (225), 1842, pp. 129 – 139.
83. Boltzmann, L. Über das Wärmegleichgewicht (1871). *Wiss. Abh.*, Bd. 1. Leipzig, 1909, pp. 237 – 258.
84. Boltzmann, L. Weitere Studien über das Wärmegleichgewicht (1872). Ibidem, pp. 316 – 402.
85. Hume, D. *Treatise on Human Nature*, vols 1 – 2 (1739). London, 1874.
86. D'Holbach, P. H. T. *Système de la nature*, pt. 2. Paris, 1781.
87. Laplace, P. S. *Théorie analytique des probabilités* (1812). *Oeuvr. Compl.*, t. 7. Paris, 1886.
88. Todhunter, I. *History of the Mathematical Theory of Probability* (1865). New York, 1949, 1965.
89. Darwin, C. *Formation of Vegetable Mould* (1881). London, 1945. (London, 1989.)
90. Bertrand, J. *Calcul des probabilités* (1888). New York, 1972.
91. Galen, C. *De sanitata tuenda.* English transl.: *Hygiene*. Springfield, Ill., 1951.
92. Maxwell, J. C. Does the progress of physical science tend to give any advantage to the opinion of necessity etc. (Read 1873). In [42, pp. 357 – 366].
93. Maxwell, J. C. On the stability of the motion of Saturn's rings (1859). *Scient. Papers*, vol. 1 (1890). Paris, 1927, pp. 288 – 376.
94. von Kries, J. *Die Principien der Wahrscheinlichkeitsrechnung (*1886). Tübingen, 1927.
95. Pirogov, N. N. Über das Gesetz Boltzmanns. *Repertorium Phys.*, Bd. 27, 1891, pp. 515 – 546.
96. Leibniz, G. W. Allgemeine Untersuchungen über die Analyse der Begriffe und wahren Sätze (MS 1686). In author's *Fragmente zur Logik*. Berlin, 1960, pp. 241 – 303.
97. Maxwell, J. C. Illustrations of the dynamical theory of gases (1860). *Scient. Papers*, vol. 1 (1890). Paris, 1927, pp. 377 – 410.
98. Maxwell, J. C. On the dynamical evidence of the molecular constitution of bodies (1875). *Scient. Papers*, vol. 2 (1890). Paris, 1927, pp. 418 – 438.
99. Laplace, P. S. *Exposition du système du monde. Oeuvr. Compl.*, t. 6. Paris, 1884. Reprint of edition of 1835.
100. Kant, I. *Kritik der reinen Vernunft* (1781). *Werke*, Bd. 3. Berlin, 1911.
101. Engels, F. *Dialektik der Natur* (written 1873 – 1882, publ. 1925). Berlin, 1971. (English translation: Amazon, 2012.)
102. Poincaré, H. *Calcul des probabilités* (1896). Paris, 1912, reprinted 1923. The Introduction to this edition is a reprint of an article of 1907.
103. Pearson, K. A. De Moivre. *Nature*, vol. 117, 1926, pp. 551 – 552.
104. Pearson, K. *History of Statistics in the 17th and 18th Centuries* etc. Lectures 1921 – 1933. Editor E. S. Pearson. London, 1978.

II

On selection and adjustment of direct observations

Izvestia Vuzov. Geodezia i Aerofotos'emka No. 2, 1966, pp. 107 – 112

Suppose that

$x_1, x_2, \ldots, x_n$

are direct observations of a constant X. It is required to replace the entire variational series by a single number. I am describing the history of the main methods of choosing such a number.

The arithmetic mean. A doctrine of means (and of the arithmetic mean in particular) existed already in the Pythagorean school [23, p. 63]. In antiquity, that mean occurred in most various formulas for calculating areas of figures and volumes of bodies. An Indian commentary of the 16^{th} century stated that the more measurements of the length, width and depth of an excavation were made, the more precise will the determination be of its size and volume [2, p. 97]. In ancient Babylonia, the area of a quadrilateral plot was in two special cases considered to be the product of the half-sums of its opposite sides, viz., when the plot was not quite precisely a rectangle, and when the measurements of the opposite sides were unequal one to another due to the ruggedness of the terrain [32, p. 204].

Thus, the mean should have compensated the inaccuracy of the models, and, possibly, the influence of the systematic errors of measurement. During the epoch of meridian arc measurements the arithmetic mean began to be applied as a universal estimator. Leibniz [21, Book 4, Chapt. 16] testified that it played an important part in developing the [classical] definition of probability and that it has been applied in the sphere of economic relations.

I especially note that even when treating direct observations it was customary to begin by deriving their **binary combinations** and only then taking the mean of these. Thus, Boscovich [22, p. 150], having four values of latitudinal differences between the endpoints of his meridian arc measurement, derived the six paired combinations and calculated their mean. The scattering of these combinations had apparently served as an indicator of error.

Simpson [28; 29] devoted a special memoir to proving stochastically that the arithmetic mean of observations distributed according to the uniform and triangular laws was preferable to a single observation. The immediate cause of his work was to refute the statement of some persons of considerable note who had thought that one single observation taken with due care was as much to be relied on as the mean of a great number of them [28, p. 82].

Such an opinion was possibly occasioned by rapid advances in the technique of observation. Simpson proved (his main result) that for a symmetric triangular distribution the probability of a certain error was essentially less than that of the same error in a single observation. He thus (indirectly) issued from the properties of random errors rather than, as it became fashionable later on, from ignorance of the causes and magnitudes of errors [16, § 4129].

The arithmetic mean has been applied together with **rejection of outlying observations**. Galileo [8, Day Third] recommended such rejection and Lambert applied it systematically.[1] Daniel Bernoulli [3], who did not approve of it, mentioned rejection as something usual, as did Euler. Gauss allowed careful rejection of large deviations; however, if, as he stated, such a deviation was caused by an unfortunate concurrence of circumstances, the pertinent observation ought to be retained [9].Thus, observations might be rejected either if they were indeed corrupted by blunders (Gauss), or if their errors were larger than some magnitude. Struve [31, 1957, § 37] and some German authors [10, p. 68; 11, p. 50] sided with Gauss; or, more precisely, they opposed subjective rejection.

Objective stochastic tests began to be applied to rejection in the second half of the 19th century [25; 5, vol. 2, pp. 558 – 566]. Contrary to Gauss' opinion, they did not take into consideration the causes of deviation.

Some participants in the ensuing discussions stressed that it was reasonable to sacrifice a few possibly sound observations and to avoid the dangerous influence of large mistakes. That attitude was of course in line with modern statistical notions on errors of two kinds. The application of objective tests naturally demanded the knowledge of the appropriate distributions (the normal law was almost always presumed). Robust tests, i.e., such which hardly depended on deviations of the distributions from their assumed type, remained unknown; the only exception was apparently the **criterion of three sigma** [13]. A number of statistical tests (e.g., [30]) were offered in the mid-20th century, but the state of the issue, as Rider [26, pp. 21 – 22] formulated it, did not apparently change.[2]

I conclude by quoting Barnett & Lewis (1978, p. 360):

When all is said and done, the major problem in outlier study remains the one that faced the very earliest workers on the subject – what is an outlier and how should we deal with it.

Posterior weights. Another estimator

$$e = \frac{\sum x_i p(e - x_i)}{\sum p(e - x_i)}. \qquad (1)$$

can be used instead of the arithmetic mean. Here $p(e - x_i)$ are the posterior weights assigned to equally precise observations x_i in accordance with the distances $(e - x_i)$ and e had to be calculated by consecutive approximations. The weights might be discrete or continuous functions of their argument, and, from the 18th century onward, mathematicians and astronomers repeatedly proposed estimators (1). Some authors thought that posterior weights can allow for changing conditions of observation over long periods of time.

For symmetric distributions estimators (1) provide a correction to the arithmetic mean due to the deviation of the observations from pairwise symmetry. In addition, at least in the usual case of posterior weights decreasing to the tails of the distribution, their use enables to do away with rejection, but these weights are assigned subjectively. In some instances (1) is at the same time the maximum likelihood estimator. Suppose indeed [3] that the density law is

$$\varphi(x) = r^2 - (e - x)^2$$

with an unknown parameter *e*. Then, according to the principle of maximum likelihood,

$$\sum \frac{(e - x_i)}{[r^2 - (e - x_i)^2]} = 0,$$

$$e = \frac{\sum p_i x_i}{\sum p_i},\ p_i = \frac{1}{r^2 - (e - x_i)^2}$$

and the weights increased towards the tails. Daniel Bernoulli had not expressly indicated that fact and it might have remained overlooked.

The median. Possibly the most active partisan of the median was Estienne [7].[3] Maintaining that random errors were only characterized by the symmetry of their density, he calculated the probability that, out of n observations, m will be negative, and $(n - m)$ positive. The probability was maximal at $m = n/2$ for an even n and at $m = (n - 1)/2$ for $n = 2m + 1$, hence the median. Estienne then formulated several properties of the median. In particular, he argued that it was closer to the true value of the constant sought than the arithmetic mean [even] if the smaller errors were more probable than the larger ones, and he also stated that the median was the most probable estimator if

$$\varphi(x) = k\exp\,[-\,|f(x) - f(a)|]$$

but did not specify $f(x)$. Estienne did not use the decrease of density to prove his first statement (which thus failed); moreover, it was formulated in a deterministic rather than stochastic sense.

Bervi [4] repeated many of Estienne's assertions and he also proved by a simple reasoning that

$$P(x_1 < X < x_n) = 1 - 2^{-(n-1)}$$

where X was the constant sought. Kornfeld [15] argued that the estimation of precision of observations should be restricted to the use of that formula but this was an anachronism.

According to modern notions [14], for some distributions the median is nevertheless preferable to the arithmetic mean; and, in particular, in the case of unknown densities. It would therefore be sensible to test the use of the median when treating the observations made by modern rangefinders since the densities of their errors are hardly known. In concluding, I note that Mendeleev [24] suggested to separate the variational series into three groups and choose as the estimator of the constant sought the arithmetic mean of the middlemost third. This was a peculiar combination of the median and the mean.

Notes

1. In the theory of errors, Lambert ([17, §§271 – 306], and a large part of [18] and [19]) is the main predecessor of Gauss. He was the first to expound systematically many of its main issues, and to offer the very term, theory of errors. He also was the first to estimate methodically, but unsuccessfully, the precision of observations (by the deviation of the arithmetic mean of all the observations from that of all of them except the most outlying observation) and even before Daniel Bernoulli he put forward the principle of maximum likelihood. At the same time, when deriving the law of distribution of certain observational errors, he issued not from their real properties, but from an alleged lack of causes for any other law.

2. See the passage in [27, p. 113]. The three-sigma test is due to Jordan [12]; the Charlier test, to Czuber [6, p. 206]; and the chi-squared distribution, to Abbe [1].

3. Estienne published two pertinent notes in the *C. r. Acad. Sci. Paris* (t. 130, 1900, pp. 66 – 69 and 393 – 395) and returned to his subject many years later [7a]. The comparison of the median with the arithmetic mean with respect to their precision began with Laplace (1818, the Second Supplement to his *Théorie analytique des probabilités*) and he [20] was also the first to introduce density of the type of (2).

The abstract of Estienne's paper mentioned in [7a] is complemented by a report on the ensuing discussion (Lévy, Hadamard). Lévy stated that, contrary to Estienne's opinion, the precision of the results increased with the number of observations (provided that the errors were not systematic); that the arithmetic mean was best for the normal distribution but the median might be preferable for other cases; that sometimes the mean square error *ne reste pas finie* which is *un argument sérieux* in Estienne's favour; but that it would then be better to reject the extreme observations *dans une proportion détermine* and to take the mean of those retained. Hadamard's remarks were less interesting: Experience proved that precision increased with the number of observations; the increasing precision of astronomical observations revealed that previous results obtained in the classical way by less precise measurements were exact [?]. H. L. Harter (1977, date of preface), *Chronological Annotated Bibliography on Order Statistics*, vol. 1. Wright-Patterson Air Force Base, Ohio, described this material but omitted Hadamard.

References

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

1. Abbe, E. Über die Gesetzmäßigkeit in der Vertheilung der Fehler (1863). *Ges. Abh.* Bd. 2. Olms, 1989, pp. 55 – 81.

2. *Algebra and Mensuration from the Sanscrit of Brahmegupta and Bhascara*. Transl. H. T. Colebrooke. London, 1817. Wiesbaden, 1973.

3. Bernoulli, D. The most probable choice between several discrepant observations etc. (1778, in Latin, with companion commentary by Euler). *Biometrika*, vol. 48, 1961, pp. 1 – 18.

4. Bervi, N. V. Determination of the most probable value of a measured object irrespective of the Gauss postulate. [*Trudy*] *Mosk. Obshchestvo Liubitelei Estestvozn., Antropol. i Etnografii*, section phys. sci., vol. 10, No. 1, 1899, pp. 41 – 45. (R)

5. Chauvenet, W. *Manual of Spherical and Practical Astronomy*, vols 1 – 2 (1863). New York, 1960 (reprint of the edition of 1891).

6. Czuber, E. *Theorie der Beobachtungsfehler*. Leipzig, 1891.

7. Estienne, J. E. Etude sur les erreurs d'observations. *Rev. artill.*, t. 36, 1890, pp. 235 – 259.

7a. Estienne, J .E. Introduction à une théorie rationelle des erreurs d'observation. Ibidem, t. 97, 1926, pp. 421 – 441; t. 98, 1926, pp. 542 – 562 ; t. 100, 1927, pp. 471 – 487. Abstract: *Bull. Soc. Math. France*, sér. 2, t. 55, 1927, pp. 24 – 25.

8. Galilei, G. *Dialogue concerning the Two Chief World Systems (*1632, in Italian). Berkeley – Los Angeles, 1962.

9. Gauss, C. F. Brief nach Olbers 3.5.1827. *Werke*, Bd. 8. Leipzig, 1900, pp. 152 – 153.
10. Gerling, C. L. *Die Ausgleichungsrechnung der practischen Geometrie.* Hamburg – Gotha, 1843.
11. Hagen, G. *Grundzüge der Wahrscheinlichkeitsrechnung* (1837). Berlin, 1867.
12. Jordan, W. Über den Maximalfehler einer Beobachtung. *Z. Vermessungs-wesen*, Bd. 6, 1877, pp. 35 – 40.
13. Kemnitz, Yu. V. Estimating the precision of equally precise geodetic measurements whose errors have non-Gaussian laws of distribution. *Trudy Mosk. Inst. Inzhenerov Zemleustroistva*, No. 2, 1957. (R)
14. Kolmogorov, A. N. The method of the median in the theory of errors (1931, in Russian). *Sel. Works*, vol. 2. Dordrecht, 1992, pp. 115 – 117.
15. Kornfeld, M. On the theory of errors. *Doklady Akad. Nauk SSSR*, vol. 103, 1955, pp. 213 – 214. (R)
16. de Lalande, J. J., *Astronomie*, t. 3 (1771). New York – London, 1966 (reprint of the edition of 1792).
17. Lambert, J. H. *Photometria*. Augsburg, 1760.
18. Lambert, J. H. Anmerkungen und Zusätze zur practischen Geometrie. In author's book *Beyträge zum Gebrauch der Mathematik und deren Anwendung*, Tl. 1. Berlin, 1765, pp. 1 – 313.
19. Lambert, J. H. Theorie der Zuverlässigkeit der Beobachtungen und Versuche. Ibidem, pp. 424 – 488.
20. Laplace, P. S. Sur la probabilité des causes par les événements (1774). *Oeuvr. Compl.*, t. 8. Paris, 1891, pp. 27 – 65.
21. Leibniz, G. W. *Neue Abhandlungen über den menschlichen Verstand*, Bde 1 – 2 (1765). Frankfurt/Main, 1961.
22. Maire, C., Boscovich, R. J. *Voyage astronomique et géographique* etc. Paris, 1770.
23. Makovelsky, A. O. *Dosokratiki*, pt. 1 (Pre-Socratian Philosophers), vol. 1. Kazan, 1914.
24. Mendeleev, D. I. Progress of work on restoring the prototypes of measures of length and weight (1875). *Sochinenia* (Works), vol. 22. L. – M., 1950, pp. 175 – 213. (R)
25. Peirce, B. Criterion for the rejection of doubtful observations. *Astron. J.*, vol. 2, 1852, pp. 161 – 163.
26. Rider, P. R. *Criteria for Rejection of Observations*. *Wash. Univ. Studies*, New Ser., Sci. & Techn. No. 8, 1933.
27. Sheynin, O. B. Mathematical treatment of observations. *Arch. Hist. Ex. Sci.*, vol. 11, 1973, pp. 97 – 126.
28. Simpson, T. On the advantage of taking the mean etc. *Phil. Trans. Roy. Soc.*, vol. 49, 1755 (1756), pp. 82 – 93.
29. Simpson, T. Same title. In author's book *Misc. tracts on Some Curious Subjects* etc. London, 1757, pp. 64 – 75.
30. Smirnov, N. V. On estimating the maximal term in a series of observations. *Doklady Akad. Nauk SSSR*, vol. 33, No. 5, 1941. (R)
31. Struve, V.Ya. (F. G. W.) *Duga meridiana* (An Arc of the Meridian). (1861). M., 1957.
32. Veiman, A. A. *Shumero-vavilonskaya matematika III – I tyshiateletii B C* (Sumerean-Babylonian Mathematics in Third – First Millenia B C). M., 1961.
Barnett V., Lewis T. (1978), *Outliers in Statistical Data*. Chichester, 1984.

III

On the history of the adjustment of indirect observations

Izvestia Vuzov. Geodezia i Aerofotos'emka, No. 3, 1967, pp. 25 – 32

The determination of n unknowns $x, y, z, \ldots$ from the equations

$$a_ix + b_iy + c_iz + \ldots + s_i = v_i,\ i = 1, 2, \ldots, m > n \qquad (1)$$

where v_i are the unavoidable residual free terms, was done from the 18th century onwards when treating meridian arc measurements by imposing, directly or tacitly, various conditions on the values of v_i. One of the first to solve a problem of this type was Euler [7] who determined the figure of the Earth from four arcs. Excluding the parameters of the Earth's ellipsoid, he got two equations between the v_i's (between the corrections to the lengths of one degree of the meridian), and, without applying any definite algorithm, restricted his efforts to rough estimations.

The first classical method of solving systems (1) was the **combination of the equations** (in pairs for the case of two unknowns). All possible combinations of two equations each were formed and the unknowns calculated for each such combination under a tacit assumption that $v_i = 0$. The final values of the unknowns were assumed to be the arithmetic means over all the combinations.

Boscovich used this method to determine the parameters of the Earth's ellipsoid [22] but he made use of another method as well (below). In 1827, even after the introduction of least squares, Muncke [25, p. 872] followed suit.

Moreover, Boscovich [ii] applied the same method for adjusting direct observations. In general, scientists of the 18th century attempted to treat both direct and indirect observations by a single algorithm, and the relation between the two cases was well understood as witnessed by the coincidence of terminology. Lambert [17, § 6] applied the same word *Mittel* to designate both the arithmetic mean and the solution of systems (1); Lalande [15, § 2699] used *milieu* in both these cases.

The method of combinations was also used for a qualitative estimation of precision, which, because of unavoidable systematic errors, should have hardly been based on deviations from the arithmetic mean. Tycho Brahe [5, p. 349] apparently pursued this goal when he, for the first time ever (and certainly before Boscovich, see above) applied the method of pairwise combinations for adjusting direct observations (of the distance between Venus and the Sun when the planet was to the east and to the west from the latter, and, as far as possible, with all other conditions being equal). Recalling the method of measuring angles in all combinations, we may ask whether Gauss arrived at it by issuing from the described method.

The method of means

$$\sum v_i = 0 \qquad (2)$$

was applied by Tobias Mayer [24] who solved a system of 27 equations in three unknowns by forming three preliminary summary equations according to condition (2). Mayer was compelled, as he himself wrote, to introduce this method to avoid the difficult work of deriving and solving all the possible combinations of three equations. More precisely, he thus used a generalization of the method of means (which, in its pure form, allows to determine only one unknown)

Condition (2) might be considered as the limiting case of the method of combinations with a single subset identical with the entire system.

At about the same time Euler [8, § 115] actually applied the same method. Having obtained two equations

$$x = s_i + b_i y + c_i z + \ldots,\ i = 1, 2$$

with pairwise roughly equal coefficients, he assumed that x was equal to the half-sum of their right sides.

Laplace [19, p. 121] mistakenly attributed the method of means to Cotes:

Cotes has prescribed that the equations of conditions be set out in such a way that the coefficient of the unknown element is positive in each of them and that all these equations be then added to form a final equation.

Actually, however, Cotes [4] provided no equations and, in essence, his few lines ran as follows:

The point Z [the centre of gravity] *will be the most probable position of the thing which with the greatest plausibility may be considered its true position*.

My statement does not detract from Cotes, who, incidentally, was well thought of by Newton. In his time, there were no quantitative substantiations of any particular method of treating observations. Even Legendre, more than 80 years later, did not justify least squares by anything other than qualitative considerations.

Like Eisenhart [6], I feel that condition (2) was understood in the 18th century as following from the equal probability of errors of each sign, and, as I shall add, as leading to the arithmetic mean in case of direct observations.

Lambert [17] used a condition of the type of (2) for fitting empirical straight lines and curves to points, – to observations (x_i; y_i). He divided the observations into two (for curves, into several) intervals with lesser and greater abscissas, determined the centre of gravity in each interval and constructed the straight line or curve passing through these. Then, Cauchy [3; 21, Chapt. 14, § 5] also used condition (2).

Boscovich [22, p. 501ff] pointed out the inadequacy of the method of combinations and **proposed a new one** to

Obtain the mean in such a way that it will not be a simple arithmetic mean, but will conform to the rules of random combinations and calculation of probabilities according to a definite law …

Specifically, he proposed to adjust the results of meridian arc measurements under three conditions, the first of which demanding that the connection between the unknowns be of the type of (1). The other two were:

second, that the sum of the positive corrections be equal to the sum of the negative ones;

third, that the sum of all the corrections, positive and negative, be minimal among those possible when the two first conditions are satisfied. ... the second condition is required for an equal degree of probability for the deviations of the pendulum and errors of observation that increase or decrease the length of a degree. The third condition is necessary for a maximal insofar as possible approximation to observations …

Boscovich' requirement of a definite law was legitimate; however, without mastering density functions he was naturally unable to say just how the rules of random combinations corresponded with his conditions.

Later Laplace [20, § 40] used the Boscovich method and Gauss [9, § 186] mistakenly attributed Boscovich' third condition to him.

Gusak [13] considered the history of the minimax principle

$$|v_{\max}| = \min$$

in which the minimum takes place for all possible solutions of (1) and traced it to the Chebyshev problem of the best approximation of an analytical function on a given segment by a polynomial of a certain degree.

Euler [7, §§ 122 – 123], about whose contribution Gusak did not report, was the first to use this principle. Later Laplace [20, Livre 3] and many other scientists applied it. Lambert [16, § 420] knew the minimax principle but admitted that he was unable to devise an appropriate algorithm which was achieved by Laplace. Cauchy [2] busied himself with this problem. A. K. Uspensky had recently recommended the minimax principle for geodetic adjustments.

The minimax method has no optimal properties but it allows to decide whether the observations were good enough and the theory underlying equations (1) was suitable. Indeed, if even the minimax led to an inadmissible maximal v_i, then at least one of the mentioned conditions was not fulfilled.

I shall now dwell on the **connections between least squares and the abovementioned principles.** The solution of (1) by least squares might be obtained [14] as

$$x = \frac{\sum \lambda_i \alpha_i}{\sum \lambda_i^2},\ y = \frac{\sum \lambda_i \beta_i}{\sum \lambda_i^2}, \ldots,$$

where α_i/λ_i, β_i/λ_i, … are the solutions of all the possible subsystems of n equations isolated from (1).

The least-squares solution differs from the one obtained by the method of combinations in that the weights of the partial solutions are there taken into account. In addition, the weights of [the estimators of] the unknowns can be calculated in a similar way by issuing from the appropriate partial weights [12].

Then, both Gauss [9, § 186] and Laplace [18, § 24] noted that the principle

$$\lim(v_1^{2k} + v_2^{2k} + \ldots + v_n^{2k}) = \min,\ k \to \infty,$$

which, in the case of large but finite values of k may be considered as a generalization of least squares, leads to the minimax principle. Indeed, for any sufficiently large k, the term v_i^{2k} with v_i^2 being the maximal term will exert the greatest influence so that the minimax condition will be fulfilled.

It is usually thought that Laplace and Gauss approached the principle of least squares from considerably differing viewpoints. Tsinger [26, p. 1] asserted that Laplace had made

A rigorous [?] *and impartial investigation; it can be seen from his analysis that least squares provide results having more or less significant probability only when the number of observations is large*. […] *Gauss, on the basis of extraneous considerations, attempted to attach to this method an unconditional significance* [a damned lie] ... *it will be easy to see the correctness of Laplace's conclusion; but with a limited number of observations we cannot count on a mutual cancellation of errors and … any combination of observations can … just as well lead to an increase of error as to its diminution*.

Tsinger exaggerated: the arbitrariness of the principle of maximum weight does not yet mean that it is unsuitable; practice had long ago refuted such a conclusion. And Markov [23] unreservedly supported this principle (without ignoring its arbitrariness). It is hardly proper to set off Laplace against Gauss. Their common interest in the treatment of observations enabled these scholars to imagine better the problem that faced them and to approach their goals with clearer understanding of the general situation. And Laplace was no armchair scientist. In particular, he actively participated in the introduction of the metric system of measurements and in the determination of the figure of the Earth which means that he could have hardly restricted his attention to limit theorems. Indeed, he [19, p. 46] pointed out that *it seems natural to use* the method of least squares even when the number of observations was small. And (p. 48)

The optimal procedure is clearly that for which the same error in the results is less probable than it would be under any other procedure.

In the same chapter of [19] Laplace several times returned to his idea about the principle of maximum weight and connected its application with the need for the most rapid decrease of the density function. He understood weight as the positive parameter k of a law of the type $\exp(-kx^2)$ and pointed out that the weight of the mean result *increased like the number of observations divided* [?] *by the number of parameters* (p. 45).

Again in the same source [19, p. 123] Laplace additionally stated that

The slight uncertainty that the observations, when there are not very many of them, leave about the values of the constants […] *causes a slight uncertainty in the probabilities determined by the analysis. But it is almost always enough to know if the probability that the errors in the observed results are contained within narrow limits and approach closely to 1; and when this is not the case, it is enough to know just how many more observations should be taken in order to obtain such a probability that no reasonable doubt remains about the quality of the results*.

Much of the above is also contained in [18, Supplements 1 and 2]. Thus, the optimal result corresponds to the maximal weight, and the weight is inversely proportional to the sum of the squares of the deviations. At the same time, weight is a parameter of the normal law and its maximum corresponds to the minimal probability of errors or the minimal length of the *confidence interval*. The principle of maximum weight is thus formulated, but actually reduced to *confidence probability* with a *confidence interval* of minimal length which makes it impossible to dispense with an assumption of a definite (of the normal) law. Laplace unquestionably issued from the theorem now called after De Moivre and him.

It might be assumed that these thoughts essentially assisted Gauss, but the latter did not mention them. I describe now how he developed the concept of weight. There is no such notion in his *Theoria motus* [9] where we find only *Genauigkeitsgrad* (§ 173). Gauss actually understood it as the root of the weight.

However, he also introduced the *Mass der Genauigkeit h*, a parameter of the distribution

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

but he did not mention the analogy between the *Genauigkeitsgrad* and this *Maass*.

In [10, § 3], issuing from the maximal value of the function

$$h^m\exp[-h^2(\alpha^2 + \beta^2 + \dots],$$

where α, β,… were the errors of m observations, proportional, as it would be said now, to the likelihood, Gauss derived the *most probable* relation

$$h = \sqrt{\frac{m}{2(\alpha^2 + \beta^2 + ...)}}.$$

Finally Gauss [11, § 6] introduced the *mittleren zu befürchtendem Fehler*, the mean error to be feared (*jactura*) m^2, as he called it in § 7, and noted (§ 9) that, for

$$\varphi(x) = \frac{1}{h\sqrt{\pi}} \exp(-\frac{x^2}{h^2}), \; m = h/\sqrt{2}, \qquad (3)$$

and called the magnitude inversely proportional to m^2 the relative weight (*Gewicht*). He could have derived a formula of the type

$$m = \sqrt{\frac{\sum \Delta^2}{n}}, \qquad (4)$$

by issuing from (3) but he did not proceed in such a manner apparently because the result would have depended on the existence of the normal distribution. This fact is extremely important; Gauss had indeed obtained a formula of this type (in § 15) but independently of the density. *The rule of least squares was already concealed* there [27, § 13].

In § 38 Gauss generalized his finding onto the case of several unknowns. According to the context, he was concerned with deriving m through the deviations from the adjusted values and he obtained

$$m = \sqrt{\frac{\lambda_1^2 + \lambda_2^2 + \ldots}{\pi - \rho}} \qquad (5)$$

where the meaning of π and ρ is obvious. His working formula was therefore (5) rather than a generalization of (4). Gauss himself, in his *Anzeige* of [11], noted that the Δ_i in (4) and in its generalization were *always* taken to be the most probable deviations but that it was now possible to apply the more precise formula (5) and thus to observe the *Würde der Wissenschaft*. I doubt that that formula should be called after Bessel. The only writing where he could have preceded Gauss is [1] but (6) is lacking there.

After Gauss' lifetime and Helmert's contributions the *theory of errors* became an engineering discipline with an established sphere of solved problems and its development mostly followed a *technological* direction (its application to the treatment of various geodetic constructions). However, beginning roughly in the 1920s, the theory became a chapter of mathematical statistics although statistical methods (mostly correlation theory and analysis of variance) have been until recently only applied in geodetic literature for special investigations. Without disparaging these at all, it might be said that they did not touch on the essence of the theory of errors.

Quite recently a number of articles on confidence estimation in the theory of errors have appeared; however, neither did this fact essentially change anything since the classical mean square error is also related to such estimation. The basic content of the theory of estimation as applied to the treatment of observations is the attempt to use more fully the information provided by each observation by means of order statistics. According to Gauss, the arithmetic mean of equally precise measurements, independently from the appropriate (but not *bad*) law of distribution, had minimal variance among linear estimators. This point of view is somewhat dated. It is now possible to arrange the observations in ascending (say) order, and to take into account the information furnished by each of them (provided that the pertinent density was known!). For example, observations might receive weights depending on their distribution and number. Extreme observations can obtain weights larger than those of the other ones, some weights might even be negative. This very approach, entirely different from the classical notions, increases the precision of treating observations and actually determines the divide between the classical theory and mathematical statistics.

The arithmetic mean is the best estimator only under normality; decrease of variance is provided by taking account of the dependence between terms of the variational series. Interestingly, no decrease is possible here for the case of the normal law.

To what degree is the increase of precision real? Geodesy, characterized by effective and multiform checking of observations, is the proper sphere for verifying this. A final comment: posterior weights had also been introduced time and time again in the classical error theory and they were assigned to equally precise observations depending on their place in the variational series [26]. They were determined almost without taking into account the appropriate distributions but at least they also aimed at improving the classical arithmetic mean.

Notes

1. These differ from equations (1) and correspond to the second main version of adjustment of observations.

2. Pendulum observations provide the possibility of obtaining the flattening of the Earth's ellipsoid of rotation.

References

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

1. Bessel, F. W. Untersuchungen über die Bahn des Olbersschen Kometen. *Abh. Preuss. Akad.* [Berlin], math. Kl., 1812 – 1813 (1816), pp. 119 – 160.

2. Cauchy, A. L. Sur le système de valeurs etc. (1831). *Oeuvr. Compl*., sér. 2, t. 1. Paris, 1905, pp. 358 – 402.

3. Cauchy, A. L. Sur l'évaluation d'inconnues etc. *C. r. Acad. Sci. Paris*, t. 36, 1853, pp. 1114 – 1122.

4. Cotes, R. Aestimatio errorum etc. (1722). *Opera misc*. London, 1768, pp. 10 – 58.

5. Dreyer, J. L. E. *Tycho Brahe*. Edinburgh, 1890.

6. Eisenhart, C. Boscovich and the combination of observations (1961). *Studies in History of Stat. and Probability*, vol. 2. Editors, Sir Maurice Kendall, R. L. Plackett. London, 1977, pp. 88 – 100.

7. Euler, L. Recherches sur des inégalités du mouvement etc. (1749). *Opera omnia*, ser. 2, t. 25. Zürich, 1960, pp. 45 – 157.

8. Euler, L. Eléments de la trigonométrie sphéroidique etc. (1755). Ibidem, t. 27. Zürich, 1954, pp. 309 – 339.

9. Gauss, C. F. Theoria motus (1809). German transl.: Aus der Theorie der Bewegung etc. In author's book *Abh. zur Methode der kleinsten Quadrate*. Hrsg. A. Börsch, P. Simon. Berlin, 1887, pp. 92 – 117. Latest edition: Vaduz, 1998.

10. Gauss, C. F. Bestimmung der Genauigkeit der Beobachtungen (1816). Ibidem, pp. 129 – 138.

11. Gauss, C. F. Theoria combinationis etc. (1823 – 1828). German transl.: Theorie der den kleinsten Fehlern unterworfenen Combination der Beobachtungen. Ibidem, pp. 1 – 91.

12. Gleinswik, P. The generalization of the theorem of Jacobi. *Bull. Géod*., t. 85, 1967, pp. 269 – 281.

13. Goussac, A. A. La prehistoire et les débuts de la théorie de la représentation approximative des fonctions. *Istoriko-Matematicheskie Issledovania*, vol. 14, 1961, pp. 289 – 348. (R)
14. Jacobi, C. G. J. *Über die Bildung und die Eigenschaften der Determinanten* (1841, in Latin). Ostwald Klassiker No. 77. Leipzig, 1896, pp. 3 – 49.
15. de Lalande, J. J. *Astronomie*, t. 3. Paris, 1771. Reprint of the third edition (1792): New York – London, 1966.
16. Lambert, J. H. Anmerkungen und Zusätze zur practischen Geometrie. In author's *Beyträge zum Gebrauche der Mathematik und deren Anwendung*, Tl. 1. Berlin, 1765, pp. 1 – 313.
17. Lambert, J .H. Theorie der Zuverlässigkeit der Beobachtungen und Versuche. Ibidem, pp. 424 – 488.
18. Laplace, P. S. *Théorie analytique des probabilités* (1812). *Oeuvr. Compl*., t. 7. Paris, 1896.
19. Laplace, P. S. *Philosophical Essay on Probabilities* (1814, in French). Transl. from the edition of 1825: New York, 1995.
20. Laplace, P. S. *Celestial Mechanics*, vol. 2 (1799, in French). Transl. by N. Bowditch. Boston, 1832; reprinted, New York, 1966.
21. Linnik, Yu. V. *Method of Least Squares* etc. (1958, in Russian). Oxford, 1961.
22. Maire, C., Boscovich, R. J. *Voyage astronomique et géographique* etc. Paris, 1770.
23. Markov, A. A. The law of large numbers and the method of least squares (1898, in Russian). In author's *Izbrannye Trudy* (Sel. Works). N. p., 1951, pp. 231 – 251.
24. Mayer, T. Abhandlung über die Umwälzung des Mondes um seine Axe. *Kosm. Nachr. u. Samml*., 1748 (1750), pp. 52 – 183.
25. Muncke, Erde. *Gehlers phys. Wörterbuch*, Bd. 3. Leipzig, 1827, pp. 825 – 1141.
26. Tsinger, V. Ya. (Method of Least Squares). A thesis. M., 1862.

IV

On the history of the De Moivre – Laplace limit theorem[1]

Istoria i Metodologia Estestven. Nauk, vol. 9, 1970, pp. 199 – 211

1. Jakob Bernoulli's *Ars Conjectandi* greatly influenced the development of the theory of probability and was always considered a classic. In particular, Laplace [21, p. 118] highly praised it. I touch on its pt. 4 that includes Bernoulli's law of large numbers (a term due to Poisson) and discuss Karl Pearson's extremely negative opinion about it.

The essence of Bernoulli's law is as follows. At first, he considers the binomial $(r + s)^{nt}$ where $t = r + s$, n is a large number and r and s are natural numbers. He proves that, for a sufficiently large nt, the sum of the $2n$ middle terms, even excluding the middlemost one, will become c times greater ($c > 0$) than the sum of its other terms. Bernoulli then makes use of that algebraic fact in his stochastic reasoning.

Let $p = r/(r + s)$ be the probability of success in each of nt (independent) trials. Then, given a sufficiently large number of these, the probability that the number of successes μ is within the boundaries $n(r \pm 1)$ can be made c times greater than the probability of the contrary event with c being fixed beforehand.

In other words, Bernoulli proved that

$$\lim P[(\mu/n) - p| < \varepsilon] = 1 \text{ as } n \to \infty.$$

Both Markov [22, pp. 44 – 52] and Pearson [26] described the appropriate mathematical steps in detail.

Bernoulli then inverts his problem and maintains, without any special proof, that if some (posterior) probability of success at any trial p is obtained after nt trials, then the probability that the true value of p lies within $[p \pm 1/(r + s)]$ can also be made c times higher than the probability of the contrary event. He also provided a somewhat lesser known estimate: for $r = 30$ and $s = 20$ [and, consequently, for $t = 50$ and $1/(r + s) = 0.02$] it occurred that $c = 1,000$ for $nt = 25,500$; $c = 10,000$ for $nt = 31,258$, etc. Thus, when nt increases by 5,758, c increases tenfold. It was hardly noted that this estimate means that

$$nt = 25,500 + 5,758\lg(c/1,000) = 8,226 + 5,758\lg c, \quad (1)$$

or that

$$c = 10^{(nt - 8,226)/5,758} \quad (1')$$

which is not difficult to write down for base *e*.

Bernoulli did not aim at estimating the change in c with the change in the boundaries of the number of successes. Expressions (1) and (1') are deterministic relations between *nt* and *c* and they show that Bernoulli effectively formulated his law as a local limit theorem.

Then, he actually introduced an exponential function, a prototype of the function

$$\varphi(x) = \frac{m}{2}\exp(-\,m|x|),\ m > 0.$$

In 1913, two hundred years after the publication of the *Ars Conjectandi*, its pt. 4 was translated into Russian under Markov's editorship. The same year Markov put out a third, a jubilee, as he called it, edition of his treatise [22] and supplied it with Bernoulli's portrait. And, again in 1913, the Imperial [Petersburg] Academy of Sciences organized a special sitting devoted to Bernoulli's work in probability with Markov, Vasiliev[2] and Chuprov reading their reports. However, only in 1924 , in the posthumous edition of his treatise, Markov [22, 1924, pp. 44 – 52] improved Bernoulli's numerical estimate (above) obtaining 16,655 instead of 25,500. He ensured the main correction (17,324) by specifying Bernoulli's intermediate inequalities. He did not apply the Stirling formula, apparently because Bernoulli naturally had not known it. [Nevertheless, he did apply that formula in a separate calculation (p. 55ff).] And Markov's residual correction followed from his abandoning the condition that *nt* be divisible by $r + s = t$.

Pearson [26] attained even better results by means of the Stirling formula and secured a practically precise coincidence of his estimate with what would have followed from the normal distribution as the limiting law for the binomial. He (pp. 202 and 210) concluded that

Bernoulli adopted a very crude method of inequalities […]. *He gets most exaggerated values for the needful number of observations, and for this reason his solution must be said to be from the practical standpoint a failure; it would ruin either an insurance society or its clients, if it were adopted. All Bernoulli achieved was to show that by increasing the number of observations the results would undoubtedly fall within certain limits, but he failed entirely to determine what the adequate number of observations were for such limits. That was entirely De Moivre's discovery*.

After all, I think, we must conclude that it is somewhat a perversion of historical facts to call the method […] *by the name of the man who after twenty years of consideration had not got further than the crude values* […] 200 *to* 300 *per cent excesses*.

Bernoulli saw the importance of a certain problem; so did Ptolemy, but it would be rather absurd to call Kepler's or Newton's solution of planetary motion by Ptolemy's name! *Yet an error of like magnitude seems to be made when De Moivre's method is discussed without reference to its author, under the heading of Bernoulli's Theorem. The contributions of the Bernoullis to mathematical science are considerable, but they have been in more than one instance greatly exaggerated. The Pars Quarta of the Ars Conjectandi has not the importance which has often been attributed to it*.

Pearson's opinion is hardly proper since the practical uselessness of the Bernoulli estimate is not that important (to say nothing about his impossibility of applying the Stirling formula). On the contrary, I stress that the very existence of that estimate and of Bernoulli's law of large numbers was extremely essential.

Pearson [25, p. 404] also noted that Bernoulli did not provide a measure of precision determined by $n^{-1/2}$. However, we should not fault Bernoulli for that either. Properly praising De Moivre, whose merits had been attributed to Bernoulli by all French and German authors known to him, Pearson at the same time profaned a great scholar.

2. Niklaus Bernoulli estimated the ratio of the middle part of the binomial series to its other parts and applied his calculations to a stochastic deduction concerning the sex ratio at birth. He communicated his results to Montmort in a letter of 23 January 1713 and the latter included them in his book [24, pp. 388 – 394] published that same year, before or at least independently from the appearance of the *Ars Conjectandi*.

Niklaus issued from Arbuthnot's data[3] and indirectly arrived at the normal distribution. Let the sex ratio be m/f, n, the total yearly number of births, and μ and $(n - \mu)$, the numbers of male and female births in a year. Denote

$$n/(m+f) = r,\ m/(m+f) = p,\ f/(m+f) = q,\ p + q = 1,$$

and let $s = O(\sqrt{n})$. Then Bernoulli's derivation (Montmort 1708/1980, pp. 388 – 394) can be presented as follows:

$$P(|\mu - rm| \le s) \approx (t-1)/t,$$
$$t \approx [1 + s(m+f)/mfr]^{s/2} \approx \exp[s^2(m+f)^2/2mfn],$$
$$P(|\mu - rm| \le s) \approx 1 - \exp(s^2/2pqn),$$

$$P[|\mu - np|/\sqrt{npq} \le s] \approx 1 - \exp(-s^2/2).$$

This result does not however lead to an integral theorem since s is restricted (see above) and neither is it a local theorem; for one thing, it lacks the factor $\sqrt{2/\pi}$

3. A French national, De Moivre (1667 – 1754) [2; 23; 33, pp. 135 – 136; 34; 8][4] was forced to leave France after the revocation of the Edict of Nantes (1685). His mathematical education (his teacher was Ozanam) occurred to be patently insufficient but he managed to fill in the gaps in his knowledge all by himself and was even elected to the Royal Society (1697). Newton favoured and respected him (De Moivre actively participated in editing the Latin version of Newton's *Optics*), and, in his later years, habitually referred to De Moivre those, who asked him questions of a mathematical nature. When the Royal Society appointed a commission for deciding the priority strife between Newton and Leibniz with regard to the analysis of infinitesimals, De Moivre was elected its member (one other member was Arbuthnot).

Todhunter [33, § 233] correctly noticed that

In the long list of men ennobled by genius, virtue and misfortune, who have found an asylum in England, it would be difficult to name one who has conferred more honour on his adopted country than De Moivre.

This, however, is not the whole story. De Moivre's new homeland did not at all secure him a worthy way of life. He was never able to take up a permanent position and had to support himself by private lessons and consultations. In 1735 De Moivre was elected to the Berlin Academy of Sciences, and in 1754, shortly before his death, to the Paris Academy.

Todhunter [33, § 336] concluded that

The theory of probability owes more to him than to any other mathematician with the sole exception of Laplace.[5]

However, when listing De Moivre's concrete achievements, he only mentioned his investigations of the duration of play, his theory of recurring series and his *extension of the value of Bernoulli's theorem by the aid of Stirling's Theorem*. Considering that this *extension* led De Moivre to the normal law, we should estimate his merits much higher.[6] His main pertinent writings are

a) *The Doctrine of Chances* [13] greatly expanded from its initial version [11].

b) *Misc. Anal.* [12] with two supplements apparently bound up to the main text at a later date. Pearson [25; 27] ascertained that not all the copies of the book have the first supplement, and only a few of these have the second one dated 1733 [10], reprinted by Archibald [4]. Owing to its importance, I list it separately:

c) *Approximatio* … [14]. De Moivre included its English translation in the second and the third edition of his *Doctrine* and introduced it [13, 1756, p. 242] in the following way:

I shall here translate a Paper of mine which was printed November 12, 1733, *and communicated to some Friends, but never yet made public …*[7]

Pearson [25] stressed that the *Approximatio* had contained the normal distribution, but he hardly knew that this fact was already noticed by [De Morgan in 1864] Eggenberger [16] and that Czuber [9] and Haussner [6, No. 108, pp. 158 – 159] mentioned the latter's discovery.

I shall first dwell on De Moivre's theological views which he expounded more fully in the second English version of his *Approximatio* [13, p. 253]. There, illustrating his thoughts by a game of dice, he maintained that

The probability of an assigned Chance, that is, of some particular disposition of the Dice, becomes as proper a subject of Investigation as any other quantity or ratio can be. But Chance, in atheistic writings or discourse, is a sound utterly insignificant: It imports no determination to any mode of Existence […] *nor can any Proposition concerning it be either affirmed or denied …*

Arbuthnot clearly formulated the problem of a determinate versus random origin of the observed predominance of male births over those of females, and concluded that that fact was occasioned by Divine design, – but why was it impossible to formulate similar problems *in atheistic writings* with the same clarity?

Derham (1657 – 1735), another Fellow of the Royal Society and a clergyman, pronounced a similar and vigorous statement [15, p. 313] likely known to De Moivre[8]:

Should we be so besotted by the devil, and blinded by our lusts, as to attribute one of the best contrived pieces of workmanship [man] *to blind chance, or unguided matter and motion, or any such sottish, wretched, atheistic stuff*?

And already in 1738 De Moivre [14, p. 251] quite definitively wrote:

Altho' Chance produces Irregularities, still the Odds will be infinitely great, that in the process of Time, those Irregularities will bear no proportion to the recurrency of that Order which naturally results from ORIGINAL DESIGN.

Pearson [27, p. 552] remarked in this connection:

De Moivre expanded the Newtonian theology and directed statistics into the new channel down which it flowed for nearly a century. (See the entire piece in [i]).

Above, I indicated that Newton had respected De Moivre. Here now is a phrase from the *Dedication* of the first edition of the *Doctrine* to Newton, as reprinted in its third edition [13, 1756, p. 329]: He, De Moivre, will think himself *very happy* if he could, by his *Doctrine*,

Excite in others a desire […] *of learning from yours* [Newton's] *philosophy how to collect, by a just Calculation, the Evidences of exquisite Wisdom and Design, which appear in the Phenomena of Nature throughout the Universe …*

In other words, how to choose between Design and Chance. The aim of his theory of probability was thus formulated.

I conclude here by quoting De Moivre's answer to a man

Who, apparently intending to pay him a compliment, remarked that mathematicians had no religion.

He replied: *I will prove that I am a Christian by forgiving you the insult you are offering* (Walker [34, p. 363], repeating an earlier author [2, p. 184]).

Book 5 of the *Misc. Anal*. is called *The binomial a + b ad Potestatem permagnam evecto*. There, De Moivre had provided a long passage from Jakob Bernoulli, described Niklaus Bernoulli's letter to Montmort (above) and solved two problems on expected winnings in games of chance and a few important algebraic problems which he applied later on in his *Approximatio*. While commenting on Niklaus, he [12, p. 98] correctly remarked that Niklaus

Did not investigate the probability that the probability of the number of occurrences or non-occurrences of an event was contained within definite boundaries.

The two abovementioned problems are also in the *Doctrine* (1738 and 1756; NNo. 72 and 73 in the latter)[9].

The second, but not the third edition has a Table of Contents where De Moivre characterized them as tending *to establish the degree of consent that should be attached to experiments* whereas the

Approximatio was modestly described as *the same subject continued further.*[10]

The Corollary to Problem 73 actually states that the statistical probability of an event will be close to its theoretical counterpart, and the closer the more observations are made. Still, De Moivre continued:

Considering the great Power of Chance, Events might at long run fall out in a different proportion from the real Bent …

and he was therefore adducing a translation of the *Approximatio* to solve *the hardest problem that can be proposed on the Subject of Chance …*

Like the *Ars Conjectandi*, the *Approximatio* consists of an algebraic and a stochastic part. In the first supplement to the *Misc. Anal.* De Moivre derived, independently from Stirling and at the same time as the latter, an approximation for $n!$. It was Stirling, however, who informed De Moivre that the constant involved in the formula was $\sqrt{2\pi}$. Commentators [22; 25] indicate that the Stirling formula should be called after both him and De Moivre. This is all the more reasonable since De Moivre, in the same supplement (and also in the *Doctrine* [13, 1756, p. 333]), published a table of lg*n*! with mantissas given to 14 digits for $n = 10(10)900$. When comparing it with a modern table [28, Anhang, Tafel 6, 18-Stellige lg*n*!] I found out that it is correct up to 11 – 12 digits with a single misprint in the fifth digit of the mantissa of lg380!.

De Moivre distinctly recognized the importance of *n* as a measure of precision and called it *the Modulus by which we are to regulate our estimation* [14, p. 248]. True, its first appearance was caused by an *algebraic* fact: the value $l = \sqrt{n}/2$ was the boundary between two methods of integrating the exponential function.

De Moivre (p. 247) also maintained that

The number n should not be immensely great; for supposing it not to reach beyond the 900$^{\text{th}}$ *power, nay not even beyond the* 100$^{\text{th}}$, *the Rule here given will be tolerably accurate, which I had confirmed by Trials*.[11]

He did not elaborate, but the mere fact of checking the precision is remarkable. Walker [34, p. 355] maintained that De Moivre had

Made few practical applications of his discoveries, and he never resorted to physical experimentation or to induction of empirical law from observed phenomena. He did not weigh and measure and count to secure objective verification of his discoveries in the theory of probability. … He does not set up experimental checks … he would doubtless have exhibited extreme astonishment at the suggestion that his Approximatio, which he thought merely an exercise in pure mathematics, contained a law which would …

De Moivre's *Trials* (above) hardly belonged to natural sciences, but they, as well as his Table of lg*n*!, and his calculations of annuities on lives testify that at least in mathematics he carried his work up to practically useful results. Consider also his *Dedication* of the *Doctrine* to Newton (above), and Walker's statement will be dismissed. As to the *Approximatio*, it was written to strengthen statistical deductions (see the description of the Corollary to Problem 73 above).

Bearing in mind that De Moivre, in concluding his *Approximatio*, noted that his deductions might be [readily] extended onto the general case of $(a + b)^n$, and that the title of that contribution included the expression *binomial* $(a + b)^n$, his finding should be interpreted as proving the local and the integral theorems on the convergence of the binomial distribution to the normal law, but of course even Laplace did not know anything about the uniform convergence that takes place there.

Independently from De Moivre, Daniel Bernoulli (1770 – 1771) derived the De Moivre – Laplace limit theorems, and I hope to discuss this topic elsewhere.[12]

Notes

1. The appearance of serious studies [18; 19] as well as of a reprint of Montmort [24] made it possible to leave out some mathematical transformations originally included here. This paper intersects my previous article [29].

2. Aleksandr Vasilievich Vasiliev (1853 – 1929), Professor at Kazan University, a mathematician and historian of mathematics, played an active part in popularizing Lobachevsky's ideas. In 1885 he published a course in probability (Kazan, a mimeographed edition). However, in this branch of mathematics he is primarily remembered as Markov's correspondent. It was in a letter to Vasiliev that Markov expounded his ideas on justifying the method of least squares.

3. John Arbuthnot (1667 – 1735) [1; 5; 32], a physician and mathematician, Fellow of the Royal Society (1704), was well acquainted with Jonathan Swift and Alexander Pope and published a few pamphlets directed against the Whigs. The name of one of his heroes, John Bull (from his *History of John Bull*) is still with us.

Arbuthnot also wrote *An essay on the Usefulness of Mathematical Learning* (1700; reprinted in [1]), and *Tables of the Grecian, Roman and Jewish Measures, Weights and Coins* (1707) and he was the main translator of Huygens' *Of the Law of Chance* (1692).

For my subject, however, the most interesting of his writings is his note [3] where he, for the first time ever [17], tested a statistical hypothesis. At present, such a procedure is understood as a test of the realization of some law of distribution, or of some value of a parameter of some definite law. Arbuthnot, however, attempted to test whether a phenomenon under his study (the prevalence of male births over those of females) was random or determinate, and he decided in favour of the latter, – of Divine design. A number of later scholars (Daniel Bernoulli, Michell) including Laplace tested hypotheses in the same sense as Arbuthnot did.

Newton apparently respected Arbuthbot. Thus, he discussed Flamsteed's observations with him (letter to Flamsteed of 1711 [7, vol. 2, p. 489]).

4. Maty's memoir proved unavailable. However, an article *Sur la vie et sur les écrits de De Moivre* is contained in the *J. Britannique* (La Haye, t. 18, Sept. – Oct. 1755), a periodical edited by him. [That memoir had been translated [5a].

5. This seems too strong.

6. The *Misc. Anal*. is not translated into any modern language [translated into French (Paris, 2009)], and the works of De Moivre are not collected together in any edition. From Lagrange's letter to Laplace of 30.12.1776 [20] it follows that they both thought of translating De Moivre's *Doctrine* into French. De Moivre began his *Approximatio* by stating that only Jakob and Niklaus Bernoulli had preceded him. And he continued:

Tho' they have shewn very great skill, and have the praise which is due to their Industry, yet some things were farther required; for what they have done is not so much an Approximation as the determining very wide limits, within which they demonstrated that the Sum of the terms [of the binomial] *was contained*.

7. In the *Approximatio* itself [13, 1756, p. 243] De Moivre also stated: *It is a dozen years or more since I had found what follows* … These years should be reckoned from 1733 (not 1738) since the Latin version of 1733 mentioned *Duodecim jam sunt anni* … In other words, De Moivre made his outstanding discovery in 1721 or a bit earlier.

8. In a letter of 1714 to Newton Derham [7, vol. 2, p. 520] asked the former to honour his promise of giving *castigations* for the third impression of his *Physico-Theology*.

9. Just as it was in several instances above, I do not describe these problems anymore. However, I refer readers to my later paper [31, p. 236] in connection with the *Spectator* introduced here by De Moivre and with the role of such outsiders.

10. See [30].

11. *Power* likely referred to *binomial to the power of n*.

12. Since then published in 1970, in *Biometrika*, vol. 57.

References

L, M = Leningrad, Moscow

1. Aitken, G. A. *Life and Works of John Arbuthnot*. Oxford, 1892.

2. Anonymous, Eloge de De Moivre. *Hist. Acad. Roy. Sci. Paris*, 1754 (1759), pp. 175 – 184.

3. Arbuthnot, J. An argument for Divine Providence taken from the constant regularity observed in the births of both sexes (1712). In *Studies in Hist. Stat. and Probability*, vol. 2. Editors, Sir Maurice Kendall, R. L. Plackett. London, 1977, pp. 30 – 34.

4. Archibald, R. C. A rare pamphlet of Moivre and some of his discoveries. *Isis*, vol. 8, 1926, pp. 671 – 684.

5. Beattie, L. M. *John Arbuthnot. Harvard Studies in English*, vol. 16. Cambridge (Mass.), 1935.

5a. Bellhouse D. R., Gewert Chr., (Engl. transl. of Maty 1760). *Stat. Sci.*, vol. 22, No. 1, 2007, pp. 109 – 136.

6. Bernoulli, J. *Wahrscheinlichkeitsrechnung* (1713, in Latin). Hrsg. R. Haussner. Ostwalds Klassiker No. 107 – 108 (1899). Frankfurt/Main, 1999.

7. Brewster, D. *Memoirs of the Life of Newton*, vols 1 – 2. Edinburgh, 1855.

8. Clarke, A. M. Moivre. *Dict. Nat. Biogr.*, vol. 38. London, 1894.

9. Czuber, E. *Die Entwicklung der Wahrscheinlichkeitstheorie. Jahresber. Deutsche Mathematiker–Vereinigung*, Bd. 7, No. 2, 1899. Separate paging.

10. Daw, R. H., Pearson, E. S. De Moivre's 1733 derivation of the normal curve: a bibliographic note. *Biometrika*, vol. 59, 1972, pp. 677 – 680.

11. De Moivre, A. De mensura sortis, or the measurement of chance (1711, in Latin). *Intern. Stat. Rev.*, vol. 52, 1984, pp. 229 – 262.

12. De Moivre, A. *Miscellanea analytica de seriebus et quadraturis.* London, 1730.

13. De Moivre, A. *Doctrine of Chances.* London, 1718, 1738, 1756. Reprint of last edition: New York, 1967.

14. De Moivre, A. Approximatio … (1733). Engl. transl.: [13, 1738; 1756, pp. 243 – 254].

15. Derham, W. *Physico-Theology.* Preached in 1711 – 1712. London, 1768 (13^{th} edition).

16. Eggenberger, J. Beiträge zur Darstellung des Bernoullischen Theorems etc. *Mitt. Naturforsch. Ges. Bern* NNo. 1305 – 1334, 1893 (1894), pp. 110 – 182. Also publ. separately: Berlin, 1906.

17. Freudenthal, H. 250 years of mathematical statistics. In *Quantitative Methods in Pharmacology.* Editor H. De Jonge. Amsterdam, 1961, pp. xi – xx.

18. Hald, A. *History of Probability and Statistics and Their Applications before 1750*. New York, 1990.

19. Hald, A. *History of Mathematical Statistics from 1750 to 1930*. New York, 1998.

20. Lagrange, J. L. *Oeuvres*, t. 14. Paris, 1892.

21. Laplace, P. S. *Philosophical Essay on Probabilities* (1814, in French; transl. from edition of 1825). New York, 1995.

22. Markov, A. A. *Ischislenie veroiatnostei* (Calculus of probability). Psb, 1900, 1908, 1913. M., 1924. German transl: Leipzig – Berlin, 1912.

23. Maty, M. *Sur la vie de De Moivre.* La Haye, 1760.

24. Montmort, P. R. *Essay d'analyse sur les jeux de hazard* (1713, second edition). New York, 1980.

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

26. Pearson, K. Bernoulli's theorem. Ibidem, vol. 17, 1925, pp. 201 – 210.

27. Pearson, K. De Moivre. *Nature*, vol. 117, 1926, pp. 551 – 552.

28. Peters, J. *Zehnstellige Logarithmentafeln*, Bd. 1. Berlin, 1922.

29. Sheynin, O. B. On the early history of the law of large numbers. *Biometrika*, vol. 55, 1968, pp. 459 – 467. Translation to appear on my site.

30. Sheynin, O. B. Daniel Bernoulli on the normal law. Ibidem, vol. 57, 1970, pp. 199 – 202. Shortened translation to appear on my site.

31. Sheynin, O. B. Early history of the theory of probability. *Arch. Hist. Ex. Sci.*, vol. 17, 1977, pp. 201 – 259.

32. Stephen, L. Arbuthnot. *Dict. Nat. Biogr.*, vol. 2. London, 1885.

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

34. Walker, H. M. De Moivre (1934). In [13, 1756, pp. 351 – 368].

V

On the work of Adrain in the theory of errors

Istoriko-Matematicheskie Issledovania (IMI), vol. 16, 1965, pp. 325 – 336

In translating my paper I took into account its somewhat revised version appended to my unpublished thesis of 1967 (*Some Issues* …). Its translated part [xiii] does not however include that version. Adrain's articles are now reprinted (see Bibliography) and I have therefore omitted his original and hardly understandable derivations of the normal law. Their latest discussion is due to Hald [9, pp. 368 – 373] and Dutka [8]. Also note that Adrain's paper [2] apparently appeared in 1809 rather than in 1808 [12, p. 170].

Robert Adrain is meritorious for his remarkable findings in the theory of errors. He published two derivations of the normal law of error a year before [or at the same time as] Gauss did and applied it to establish the principles of least squares and arithmetic mean as well as to determine the flattening of the earth's ellipsoid of rotation.

Adrain was born in Ireland (1775) and died in New Brunswick (1843). He learned mathematics mainly by himself and began teaching it at an early age. Then, after participating in the Irish national movement and being wounded in the revolt of 1798, he fled to the United States. Adrain resumed there his teaching activities becoming, in 1809, Professor of mathematics at Queen's College (now, Rutgers College) in New Brunswick. From 1813 to 1826 he was Professor at Columbia University, and, from 1827 to 1836, at Pennsylvania (vice-rector from 1828 to 1836).

Adrain delivered lectures in various disciplines. Thus, in 1829 he taught elementary mathematics, geodesy, cartography, mathematical analysis, mechanics and astronomy. He and Nathaniel Bowditch (1773 – 1838)[1] were among the first American mathematicians. In 1812 Adrain was elected to the American Philosophical Society, and, in 1813, to the Academy of Sciences and Arts. He actively contributed to the first American mathematical periodicals. Coolidge [7] provided a general description of Adrain's work, but his account of the latter's findings in the theory of errors was not comprehensive. In the 19th century several geodesists and astronomers discussed these in more detail (e.g., [1; 9; 26] are from those which I do not mention below) but still not sufficiently. At present [in 1965], however, Adrain is forgotten. Neither Struik [22][2] nor Strasser [21] nor many other authors of general contributions on history of mathematics cite him and Cajori [6, p. 382] only devoted a few lines to Adrain's discovery of the law of error.

I turn now to Adrain's paper [2]. He issued from a prize question: A traverse with measured sides and bearings did not close. It is required to determine the most probable corrections to the computed increments of the coordinates which should have disappeared.

His paper also contained two derivations of the normal law of error; the derivation of the principles of least squares (discovered by Gauss in 1795 or 1794 and offered by Legendre in his publication of 1805) and of the arithmetic mean in the one- and three-dimensional cases; the determination of the most probable position of a ship calculated by dead reckoning given its observed latitude. In concluding, Adrain stated that, owing to lack of space, he had to postpone the publication of his derivation of the most probable flattening of the earth's ellipsoid which he accomplished [in 1818, see below] on the basis of the normal law. Both derivations of the normal law were damned unsatisfactory even in the assumptions made. And, issuing from those derivations, he applied an embryo of the principle of maximal likelihood for deriving the principle of least squares.

1) His first derivation of the normal law concerned linear measurements. He tacitly believed that their errors were independent and obtained a function of the normal type with two essential constants. Neither here, nor in his next derivation Adrain calculated their values; moreover, he had not considered it important.

2) In his second derivation of the normal law Adrain studied the determination of a station B from a given station A by measured distance AB and azimuth of that line. He also obtained the ellipse of concentration (of error) but had not named, or paid any attention to it. He thus missed the opportunity of introducing the bivariate normal law.

Merriman [20] and subsequent authors pointed out that the derivation of the normal law by John Herschel [11] was similar to Adrain's second justification, see below.

3) Adrain derived the principles of least squares and arithmetic mean in a way, similar to that applied by Gauss in 1809.

Adrain's derivation of the principle of least squares (for one unknown) was questioned: Coolidge [7] stated that he had Legendre's book in his library.[4] The proper derivation for the three-dimensional case (also considered by Legendre) was similar.

4) The correction of dead reckoning was similar to the adjustment of the traverse. The observation of only one astronomical magnitude, the latitude, leads to only one (a latitudinal) discrepancy between the reckoned and the observed positions of the ship. Consequently, this case is indeed similar to adjusting the traverse with respect to only one coordinate.

In his next article [3] Adrain determined the flattening of the earth's ellipsoid $\alpha = (a - b)/b$ with semi-axes a and b ($a > b$)[6]. He issued from Laplace's data [17, § 40] on the lengths of the seconds pendulum at various latitudes. The Clairaut formula connecting the latitude λ with the acceleration of gravity r is, in his notation,

$$r = x + y \sin^2\lambda$$

with x and y determining α and Adrain solved a system of such equations under the least-squares condition imposed on the residual free terms (call them v_i) of his (Laplace's) equations getting $\alpha = 1/319$. His trivial application of least squares was interesting in that it was the first one in its field[7] and, in addition, because of the result obtained. Laplace himself got $\alpha = 1/335.78$ which almost coincided with another figure calculated at the same time (in 1799). The difference between Laplace's and Adrain's results was mainly caused by two mistakes made by the former and revealed by the latter. According to Laplace's (Boscovich's) conditions

$$\sum|v_i| = \min,\ \sum v_i = 0,$$

Adrain additionally arrived at $\alpha = 1/316.5$ so that the essential difference between him and Laplace was $1/316.5 - 1/319 = 0.00303$.

I also mention Adrain's article [4]. Believing that for some practical purposes it was convenient to consider the Earth as a sphere, he calculated its radius under seven different conditions: equal volume or equal surface with the appropriate ellipsoid of rotation; equal masses of bodies restricted by these surfaces (sphere and ellipsoid) given a certain law of the decrease of mass with depth, etc. As a first approximation, Adrain arrived at one and the same result, $r = (2a + b)/3$. He then stated that he had determined the most probable values of a and b by means of meridian arc measurements according to the method published in his first paper. The corresponding mean radius, as he added without providing either the initial data or his own calculations, was 3,959.36 English miles. With $\alpha = 1/319$ this was equivalent to a = 3,963.50 English miles, or, assuming that 1 meter equals 39.370113 inches, 6,378.629*km*. Here are some later determinations [13].

1. Delambre, 1800:	a = 6,375.653*km*,	$1/\alpha = 334$
2. Walbeck, 1819:	76.896	302.78
3. Krasovsky, 1940:	78.245	298.3

Herschel [11] derived the normal law by considering a free fall of a ball on a horizontal plane and its deviations from the point above which it was initially situated. He distinctly formulated the symmetry and the decrease of the density sought indicating that those assumptions were a corollary of complete ignorance of the causes of error and the manner of their action.

Thomson & Tait [23, p. 314], without referring to anyone, offered a similar justification. Again, both Tsinger [24] and Krylov [16, Chapt. 8] applied the same pattern for deriving the normal law by considering shooting at a vertical target. Kemnitz [15] noted that Krylov (and therefore his predecessors as well) had not made essential use of the properties of random errors.[8]

Notes

1. A mathematician and astronomer, Fellow of the Royal Society. He is mostly remembered for his work in navigation and the translation of Laplace's *Mécanique Céleste*.

2. As stated in the title of his book, that author had indeed restricted his study to the Yankee, rather than to the American science.

3. This property is characteristic of systematic rather than random errors, cf. Item 5 below.

4. It remains unknown, however, when did Adrain get it [9, p. 371]. More [10, p. 626]: Adrain never used the term *least squares*, nor did he refer to Legendre's treatment of meridian arc measurements, cf. below. Finally, Adrain had not then directly stated that his principle of least squares might be applied to the case of several unknowns.

5. Adrain properly adjusted the directly measured (rather than the calculated) magnitudes.

6. Adrain's definition of the flattening was unusual: the generally adopted formula was and is $\alpha = (a - b)/a$. In some cases, as when comparing his results with those of other authors (below), this is of no consequence. After reading [3], Olbers informed Gauss (24.2.1819; [20, p. 711]) that *ein Amerikaner … schreibt sich … die Erfindung der Methode der kleinsten Quadrate zu*. Gauss made no comment.

7. The first published application of this kind is however due to Biot [5, Additions, pp. 167 – 169].

8. It is generally believed that Maxwell, in his celebrated justification of the normal law of the velocities of gas molecules which assumed the independence of the three components of the velocities, issued from Herschel's derivation. Kac [14] and Linnik [18] had since revised Maxwell's proof. Independence is still needed but in a weaker form; however, it should in addition persist under any choice of the coordinate system.

References

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

1. Abbe, C. Historical note on the method of least squares. *Amer. J. Sci. Arts*, vol. 1, No. 6, 1871, pp. 411 – 415.

2. Adrain, R. Research concerning the probabilities of the errors which happen in making observations (1808). Reprinted in Stigler, S. M., Editor, *American Contributions to Mathematical Statistics in the 19th Century*, vol. 1. New York, 1980. No single paging.

3. Adrain, R. Investigation of the figure of the Earth and of the gravity in different latitudes (1818). Reprinted: Ibidem.

4. Adrain, R. Research concerning the mean diameter of the Earth (1818). Reprinted: Ibidem.

5. Biot, J.-B. *Traité élémentaire d'astronomie physique*, t. 3. Paris, 1811.

6. Cajori, F. C. *History of Mathematics,* 2nd ed. New York, 1929.

7. Coolidge, J. L. Adrain and the beginnings of American mathematics. *Amer. Math. Monthly*, vol. 33, No. 2, 1926, pp. 61 – 76.

8. Dutka, J. Adrain and the method of least squares. *Arch. Hist. Ex. Sci.*, vol. 41, 1990, pp. 171 – 184.

9. Hald, A. *History of Mathematical Statistics from 1750 to 1930.* New York, 1998.

10. Hammer, E. Zur Geschichte der Ausgleichungsrechnung. *Z. Vermessungs-wesen,* Bd. 29, 1900, pp. 613 – 628.

11. Herschel, J. [Review of] Letters on the theory of probabilities … by Quetelet. *Edinb. Rev., or, Critical J.*, vol. 92, No. 185, 1850, pp. 1 – 57. Published anonymously.

12. Hogan, E. R. Adrain: American mathematician. *Hist. Math.*, vol. 4, 1977, pp. 157 – 172.

13. Izotov, A. A. *Forma i Razmery Zemli po Sovremennym Dannym* (Form and Size of Earth according to Modern Data). M., 1950.

14. Kac, M. On a characterization of the normal distribution (1939). [*Sel. Papers.*] *Probability, Number Theory and Statistical Physics*. Cambridge (Mass.), 1979, pp. 77 – 79.

15. Kemnitz, Yu. V. On a derivation of the law of error. *Trudy Mosk. Inst. Inzhenerov Zemleustroistva*, No. 3, 1959. (R)

16. Krylov, A. N. *Lektsii o Priblizhennykh Vychisleniakh* (Lectures on Approximate Calculations). M., 1950.

17. Laplace, P. S. *Traité de Mécanique Céleste*, t. 2 (1799). *Oeuvr. Compl.*, t. 2. Paris, 1878.

18. Linnik, Yu. V. Comments on the classical derivation of the Maxwellian law. *Doklady Akad. Nauk SSSR*, vol. 85, 1952, pp. 1251 – 1254. (R)

19. Merriman, M. List of writings relating to the method of least squares with historical and critical notes. *Trans. Connecticut Acad. Arts Sci.,* vol. 4, pt. 1 – 2.

New Haven, 1877 – 1882, pp. 151 – 232. Reprinted in Stigler (1980, vol. 2), see reference in [2].

20. Schilling, C. W. *Olbers, Sein Leben und seine Werk*, Bd. 2, Abt. 1. Berlin, 1900. Reprint: Gauss, *Werke, Ergänzungsreihe*, Bd. 4. Hildesheim, 1976.

21. Strasser, G. *Ellipsoidische Parameter der Erdfigur, 1801 – 1950*. Deutsche geod. Komm. Bayer. Akad. Wiss., Bd. A19, 1957.

22. Struik, D. J. *Yankee Science in the Making*. New York, 1962.

23. Thomson, W., Tait, P. G. *Treatise on Natural Philosophy*, vol. 1. Oxford, 1867. New York, 2002.

24. Tsinger, N. *Kurs Astronomii. Chast Teoreticheskaya* (Course in Astronomy, Theor. Part). Psb, 1899.

25. Wright, T. W. *Treatise on the Adjustment of Observations.* New York, 1884.

VI

Ladislaus von Bortkiewicz: a scientific biography

Dzieje matematyki Polskiej. Wroclaw, 2012, pp. 193 – 214

1. General Information

1.1. Russia. Vladislav Iosifovich Bortkevich (7 Aug. 1868–15 July 1931) was born in Petersburg into a family of Russified Polish nobility. After moving to Germany he changed his name and became Ladislaus von Bortkiewicz. I abbreviate his name as V. I. and L. B., respectively.

His mother was Helene, née von Rokicka, and father, Iosif Ivanovich, a colonel in the Russian army, later a notary public and teacher of mathematics in a *gymnasium*. V. I. finished a *humanistisches Gymnasium* (UK PA B 347; such references denote codes of the Humboldt University, Berlin, Archive) and graduated from the law faculty of Petersburg University in 1890. In the same source (a questionnaire) Bortkiewicz called himself a Roman Catholic, but he never mentioned religious matters in published works or letters known to me. According to his will (Schumacher 1931, p. 573),

Statt des Vertreters einer Kirche … nur Vertreter der Wissenschaft und der Freundschaft hier [to the cemetery] *heute zum Worte kommen.*

His political standpoint is evident since he attempted to help Gumbel, a noted leftist and later a well–known statistician, to secure a position and to receive a fellowship (Sheynin 2003, pp. 20–21). Without elaborating, Tönnies (1932/1998, p. 319) reported that Borkiewicz *hat der Deutschen Demokratischen Partei angehört,* but political problems were not discussed in his correspondence. Schumacher (1931, p. 576) testified that L. B. had experienced *weitgehenden Uninteressiertheit gegenüber der nationalen Politik*.

His sharp criticism (1903) of P. A. Nekrasov, a talented mathematician who became a double–dyed reactionary, and study of Marx (§ 3.1) show him as a liberal. Bortkiewicz remained a bachelor and devoted all his life to science *als ob sich auf das Bibelwort bezöge: Du sollst keine Götter neben mir haben* [Exodus 20:3] (Schumacher 1931, p. 573).

In Berlin, his unmarried sister Helene was keeping house for him. And, in economics (but actually in everything) he was an *индивидуалист* (Загоров 1929, p. 12). Bortkiewicz fully mastered German which was likely spoken in the family and well taught in the *gymnasium*. Anderson (1932, p. 242/1963, p. 530) was hardly entirely right when stating that

Ist er [V. I.] … *ganz im russischen Kulturkreis aufgewachsen*.

Even before graduating from the University, V. I. began studying statistics and economics and published two papers on population statistics (1890b; 1891) and a study (1890a) criticizing Walras. Altschul (1931, p. 1183) commented:

Walras, der um drei Jahrzehnte ältere Forscher und das anerkannte Haupt der Lausanner Schule, mit ihm, dem Anfänger, in einen langjährigen Briefwechsel eintritt, der die schwierigste Probleme der mathematischen Ökonomik umfasst.

After the University, *er vom Russischen Unterrichtsministerium zur Fortbildung ins Ausland geschickt wurde* (Tönnies 1932/1998, p. 315). But already in 1888, before publishing anything, Bortkevich wrote Knapp a letter making suggestions to that leading scholar *for a reform of the methods used* in estimating mortality (Andersson 1931, p. 9). He *won the appraisal of the master* who asked him *who he was* … Knapp also wrote:

It will please me still more if I should have an opportunity some time of making your personal acquaintance.

Andersson had probably read Bortkevich's posthumous papers before sending them (as I imagine) to Uppsala (§ 2.9).

1.2. Germany. The German (at the time) Straßburg became Bortkevich's *Ausland*. Then, in May 1891 (Andersson 1931, p. 10), Knapp was

The principal of the university. His duties … prevented further scientific teaching so that a special vacation course … was agreed upon. During six weeks, three or four hours a day, … *Knapp demonstrated the results of his mathematical-statistical investigations and found himself richly rewarded by the expert participation of his pupil in this extraordinary undertaking*. *Should I ever receive an inquiring as to your skill*, *Knapp writes in 1894 to V. I.*, *I shall give expression to my great delight*. … *Only think*, *wrote Knapp in 1893*, *that with the exception of Lexis and myself there are no "higher" statisticians and neither are there prospects of any*.

At the beginning of the summer term of 1893 Bortkiewicz moved to Göttingen, to Lexis and (Lorey 1932, p. 199)

Die nationalökonomischen und statistischen Studien auch durch philosophische erweiterte.

In 1893 he obtained there the degree of Doctor of Philosophy (UK PA B347).

Schumacher (1931, p. 575) comments: L. B. was able

Mit überraschend sicherem Instinkt die beiden voneinander grundverschiedenen Männer herauszufinden, die seinem Wesen am meisten adäquat waren: … Knapp und … Lexis.

Once more in Straßburg, he taught insurance of workers and theoretical statistics, and in 1895 became Privat-Dozent (Lorey 1932, p. 200). *Als Habilitätsarbeit diensten wohl die ersten zwei Teilen* of his contribution (1894–1896).

However, he returned to Russia. *С 1 сент. состою на службе в Упр. каз*[*ёнными*] *жел. дор.* (Letter 25 of 1897; such references are to letters included in Борткевич и Чупров 2005). Покотилов (1909) mentioned that he and V. I. had compiled and successfully implemented a plan for the first ever in Russia state insurance of workers. Nevertheless, Chuprov (Letter 59 of 1901) later congratulated him on abandoning *административную деятельность*.

V. I. kept back his work in Russia from Straßburg University (Letter 27 of 1897) because Privat–Dozents were forbidden to pluralize. In autumn of 1899 V. I. in addition began delivering lectures in statistics at the prestigious Aleksandrovsky Lyceum for which Aleksandr Ivanovich Chuprov (father of A. A. Chuprov, an eminent non-mathematical statistician) had recommended him (Letter 25 of 1897).

In 1901 Bortkevich left both his positions and moved to Germany to live there all his life. From 15 Jan. (effectively, as of 1 March) he was appointed *zum außerordentlichen Professor in der Philosophischen Fakultät* [of the Friedrich–Wilhelm, now Humboldt, University in Berlin] (UK PA B347). He was

Verpflichtet die Statistik nebst den verwandten Disziplinen (Versicherungswesen, Bevölkerungswesen etc) … *zu vertreten*, and, if necessary, *auch sonst zur Vervollständigung des Lehrplans auf volkswirtschaftlichen Gebiete beizutragen*.

In particular (Voigt 1994, p. 337),

Besondere wissenschaftlichen Fragen des russisches Staatslebens wünschte Herr Ministerialdirektor Dr Althoff der speziellen Leitung des Prof. von Bortkiewicz zu sehen.

It was resolved and apparently implemented that L. B.

Wird Vorlesungen über die wirtschaftlichen Verhältnisse Rußlands halten and conduct classes *im Zusammenhang mit dem russischen Seminar*.

The invitation came *vermütlich* on Lexis' initiative (Lorey 1932, p. 202). For his part, Andersson (1931, p. 10) reported that

When there was a question of calling Lexis to the University of Berlin at the beginning of the new century, [he] *did not wish to go … himself*, [and] *was able to propose the appointment of von Bortkiewicz in his stead*.

Bortkiewicz (Letter 79 of 1905) dropped a phrase: *если бы даже осуществилось намерение Лексиса передать мне кафедру* [в Гёттингене] … his material condition would not have essentially improved. Bortkiewicz remained in his new position until 1920 after which he (Phil. Fak. 1469, Bl. 67) became *ordentlicher Professor* (*persönlicher Ordinarius*). Schumpeter (1932, pp. 338–339) specified:

This eminent man was never thought of as a candidate for one of the great chairs, either in Berlin or at any other University, and it was not until 1920, when by a measure intended to "democratise" faculties, <u>all</u> extraordinary professors became full professors ad personam, that he obtained that rank, without, however, ceasing to be entirely isolated.

In 1906, L. B. (UK PA B347) was appointed

Dozent im Nebenamt an der Handelshochschule der Korporation der Kaufmannschaft von Berlin and was to deliver two-hours [weekly] lectures in *Versicherungswesen*.

30.7.1931, two weeks after he had died, the *Kuratorium* of the Hochschule sent a letter to his sister Helene (same code) telling her that her late brother

Hat … doch seit Bestehen der [Schule] *bis zum Wintersemester 1922/1923 an ihr in vorbildlicher Weise gelehrt. In höchst* *dankenswerter Weise ist es ihm gelungen, wissenschaftliche Gründlichkeit und allgemein verständliche Lehrweise mit einander zu paaren*. … *Unsere Dankbarkeit gegenüber dem Verstorbenen auch noch dadurch zum Ausdruck zu bringen*, *dass wir Sie bitten, den Betrag von RM 200,– zur Beschaffung eines Grabsteines verwenden zu wollen*.

Bortkiewicz' body was cremated and the remains buried at Wilmersdorf cemetery (since relocated), Abt. B2, Stelle 138. (If erected), his *Grabstein* disappeared.

Finally (WHB 603/1), the *Sekretariat der Wirtschaftshochschule* [the former Handelshochschule, in a document apparently intended for the school's archive] stated that the portrait of L. B., a former *Lehrbeauftragte*, had disappeared from the *Hörsaal*. The Sekretariat suspected that the portrait

Von einem Unbefugten in der irrtümlichen Annahme, Herr von Bortkiewicz sei nicht deutschblütig gewesen, entfernt worden ist.

In essence, the school had exonerated the thief which fully conformed to the situation then existing in Nazi Germany. And Bortkiewicz was not *deutschblütig* at all.

But was Bortkiewicz really a good lecturer? Many authors unanimously declared the opposite. Max Weber, an economist and co–creator of sociology, provided the clearest statement. As quoted by Meerwarth (1936, p. 257), in 1911, L. B. discussed in his report the *Auslese und Anpassung der großindustriellen Arbeiterschaft*.

Diejenige Rede heute, die wenigsten nach der Ansicht der meisten Anwesenden hier die langweiligste gewesen ist, war die, die Herr … *Bortkiewicz gehalten hat, zugleich aber – diejenige, deren Kritik uns sachlich am meisten zu fördern geeignet ist*.

Altschul (1931, p. 1184) agreed:

Die eigentliche pädagogische Wirksamkeit lag ihm aber nicht. Er sah keine Möglichkeit, seine fein differenzierten Gedankengänge in adäquater Weise einem größeren Kreise zugänglich zu machen.

Bortkiewicz himself (Letter 79 of 1905) stated that he did not appreciate himself *особенно высоко как лектора и руководителя*. The last word apparently referred to supervision of postgraduates. Indeed, only in 1913 he was appointed *zum Mitdirektor des Staatswirtschaftlich-statistischen Seminars* of the University (Phil. Fak. 1466, Bl. 186).

From November 1916 to February 1917 Bortkiewicz (Phil. Fak. 1467, Bl. 123 and 195) was a *wissenschaftlich-statistisch Hilfsarbeiter bei der Zivilverwaltung des Generalgouvernements Warschau*. No details are known. He published a paper (1901) in Polish and another one (1930b) in German and Polish. Andersson (1931, p. 24) likely meant the latter:

A memorial written … *at the instance of the Polish Government – covering 44 pages in print – with reference to the life-insurance of mortgagers*.

Two obituaries of Bortkiewicz had appeared in Poland: Neyman (1931) and S. P. (1931). Here is the end of the latter:

Ze śmiercią … Bortkiewicza zeszedł do grobu jeden z najwybitniejszych przedstawicieli nowoczesnej statystyki teoretycznej.

In the Soviet Union, however, Старовский (1933) called him, together with other most eminent statisticians including Chuprov,

теоретиками буржуазной статистики who *при помощи статистических построений доказывают "незыблемость" и "вечность" капиталистического строя и "устойчивость" его законов.*

A damned lie! See also Аноним (1927).

2. The essence and conditions of scientific work

2.1. Lack of mathematical education. Bortkiewicz had no mathematical education which barely told on his published works and is his considerable achievement. Schumpeter (1932, p. 339) even maintained that, as an economist, he *just missed greatness by refusing to put to full use the mathematical tools at his command* …

While still in Russia, Chuprov helped him in his mathematical efforts. Thus (Letter 14 of 1896/1897 and 15 and 17 of 1897), he convinced Bortkiewicz in that he had mistakenly doubted the correctness of one of Gauss' theorems. Some roughness and even mistakes had nevertheless occurred. Lorey (1932, p. 204) noted that Bortkiewicz (1893a) had wrongly thought that continuity of a function led to its differentiability. And, concerning his criticism of Pareto, Chuprov (Letter 25 of 1898) indicated that Bortkiewicz' (mathematical) *аргументация не вполне точная*.

He himself (1917, p. III) remarked that he presupposes

Beim Leser außer der Beherrschung der niederen Algebra lediglich noch die Vertrautheit mit den Anfangsgründen der Wahrscheinlichkeitsrechnung

and added that the application of generating functions is as though *die Gleichung* $2x - 3 = 5$ *mit Hilfe von Determinanten zu lösen*. Let this be a slip of the pen, but, when preparing his booklet (1898b), he (Letter 7 of 1896) also informed Chuprov about his reluctance, contrary to Markov's advice, *прибегать к помощи производящих функций и последовательных дифференцирований*. This was unreasonable, and the more so since the method he meant was applicable to random variables in general rather than only to the binomial distribution which he studied.

Keynes (1921, p. 403n2/1973, p. 440n2) very critically commented on Bortkiewicz' mathematical computations:

The mathematical argument is right enough and often brilliant. But what it is all really about, and what it really amounts to and what the premises are, it becomes increasingly perplexing to decide.

And he did not even mean the law of small numbers, see below.

It is curious that, as Bortkiewicz informed him, Lorey (1932, p. 203n5) was *der einzige eines Mathematiker gewesen sei* who had sent him a *Glückwunsch zur sechzigsten Wiederkehr seines Geburtstages*.

2.2. Critical trend. In 1900, in a letter to his father, Chuprov (Шейнин 1990/2011, с. 55) described his acquaintance with Bortkiewicz:

Борткевич относится к науке как к спорту. Его интересует упражнять и проявлять свои недюжинные силы. … Он очень самолюбив; любит людям импонировать. … На этой почве у него развивается показное отношение к науке. … Силён в Бортк.

и скептицизм. … В существе у него и к науке, и к жизненным вопросам вполне серьёзный и интерес, и отношение, но он предпочитает этого не показывать, ему в большинстве случаев лучше нравится рисоваться скептицизмом. … В душе он сильно полонофил.

No, that scepticism was no pretence; his numerous reviews prove that it was really justified. At first, Bortkiewicz regarded some of his important contributions, for example (1917; 1923–1924) as reviews (and such they formally remained). Furthermore, Андерсон (1929, с. 7) indicated that he

Влага в тех толкова свое, така … осветлява и допълва, че се получава нещо съвсем ново и оригинално.

(He inserts there so much of his own, enlightens and enlarges it so much, that something quite new and original emerges.)

Shumpeter (1932, p. 339) stated that

Even his original contributions assumed the form of criticisms, and that critique became his very breath.

And here are other testimonies.

Woytinsky (1961, pp. 451–452):

In Germany, he was called the Pope of statistics. … The publishers have stopped asking [him] *to review their books* [because of his deep and impartial response]. … [He was] *probably the best statistician in Europe*.

Андерсон (1931): *Борткевич в течение многих лет до самой кончины был своего рода "верховным контролером" научной мысли в области своей специальности, и не один автор, публиковавший работу по теории статистики или политической экономии, с волнением, а иногда с трепетом, ожидал его отзыва, нередко сурового, порой жестокого, но всегда нелицеприятного и обоснованного. Но зато короткое слово одобрения из уст этого аскета науки значило больше, чем самая пламенная похвала со стороны других. Поэтому научное значение Борткевича должно быть измеряемо не только тем, что было написано им самим, но и тем, что было написано другими благодаря ему, под влиянием его критики и в результате его указаний. А если угодно, то в большую заслугу Борткевичу можно поставить и то, наверное, весьма значительное количество посредственных и слабых научных работ, которые не были выпущены в свет из боязни подвергнуться его сокрушительной критике*.

Anderson (1932, p. 245/1963, p. 533):

Bortkiewicz galt allgemein als ein scharfer und galliger Richter, vor dessen Urteil sich auch die prominentesten Gelehrten in Acht zu nehmen hatten. … Wir dürfen aber nie vergessen, dass [his] *Urteil immer sachlich und unparteiisch blieb. Kein Gelehrter stand ihm wohl persönlich näher als Tschuproff, und dennoch hat es zwischen beiden wissenschaftliche Duelle gegeben, bei denen sehr schmerzhafte Hiebe erteilt wurden*.

Anderson apparently described oral battles and letters from one another and to others; in print, they criticized each other (or, rather, Chuprov criticized Bortkiewicz) mildly. In 1897, in a letter to his father (Шейнин 1990/2011, pp. 55), he described his first acquaintance with L. B.:

Очень большое значение имели для меня … сношения с Борткевичем. … Твоё участие, твои советы и замечания были мне всегда … живой поддержкой, но не легко было воспользоваться ими, да и работал я всё в областях, в которых ты не работаешь. … С Борткевичем же нас связывает общность и интересов, и близость направлений, мы работаем над одними и теми же вопросами, примыкая к одним и тем же предшественникам. … Переписка с ним полезна и приятна для меня. … Правда, учителем моим … он быть не может – разница знаний … недостаточно велика.

Their close friendship lasted until Chuprov's death (in 1926); see my Introduction to Борткевич и Чупров (2005).

Schumacher (1931, p. 575):

Jede vage und laxe und schiefe Ausdrucksweise empfand er als eine Versündigung am Geiste der Wissenschaft.

Meerwarth (1936, pp. 256–257):

Es ist nicht zu leugnen, dass seiner Kritik gelegentlich ein pedantischer, langweiliger Zug anhaftete. … Doch hat niemals über den geistigen Gehalt und den Nutzen seiner Kritik ein Zweifel bestanden.

Gumbel (1968, p. 26):

He criticized with equal zeal and profundity important and insignificant mistakes, printing errors and numerical miscalculations.

Other authors barely criticized Bortkiewicz publicly; his contribution (1898b) was an exception, see § 3.3.1. Only Chuprov (1918–1919), in a long note that ended his contribution, stated that Bortkiewicz wrongly opposed the Biometric school and the Continental direction of statistics.

2.3. Lack of summary works. One of the causes was apparently Bortkiewicz' self-criticizing nature:

Hat er zahlreiche Wissenschaftler befruchtet und trotzdem keine eigentliche Schule hinterlassen. Dies lag zum Teil an seinem herben Charakter. Im Grunde unterschätzte er seine eigenen Arbeiten, ja, er zweifelte sogar – zu Unrecht – ihre praktische Bedeutung an. Dies mag dazu beigetragen haben, dass er manchen Begabten nicht genügend herangezogen hat; denn er war zu verantwortungsvoll, um Hoffnungen zu erwecken (Gumbel 1931, p. 233).

And *he had so great an inhibition in giving to the public, that he lost some of his claims to high originality* (Schumpeter 1932, p. 340). Anderson (1932, p. 244/1963, p. 532) described Bortkiewicz' self-criticism: *war ihm auch die geringste Kleinigkeit nicht zu wenig*. He (p. 247/p. 535) also indicated another cause: a German publishing house was prepared to put out *eine Sammlung seiner wichtigeren Untersuchungen*. Bortkiewicz, however, declined:

Setzte er alles daran, um nicht begonenne und halbfertige Werke mit sich ins Grab zu nehmen. Bei seinem unerwarteten Tod wird er wohl doch das meiste mit sich genommen haben!

2.4. A tiny circle of readers. This was occasioned both by the unwillingness of German scholars to read anything mathematical (§ 2.7) and Bortkiewicz' own refusal to help them:

Durch die Gründlichkeit verschwand manchmal die Linearität des Gedankengangs. In jeder Untersuchung waren außer dem zentralen Gedankengang zahlreiche Nebenlinien und umfangreiche Polemiken eingebaut. Daher stellen diese Arbeiten an die Aufmerksamkeit des Lesers große Ansprüche. Dem wirklichen Leser gab er viel (Gumbel 1931, p. 233).

So, how many *wirklichen* readers did he have? Winkler (1931, p. 1030) quoted Bortkiewicz' letter to him (but did not provide its date):

Ich freue mich, in Ihnen einen der erwarteten fünf Leser gefunden zu haben.

Nebenlinien (or even second main subjects) are present in Bortkiewicz (1898b). He begins there his *Vorrede* by stating his aim: *näher zu treten* statistical series consisting of a small number of trials, but he discussed the stability of series in a general setting. Anderson (1932, p. 245/1963, p. 533) remarked that it was difficult to find outlets; many periodicals did not accept mathematically oriented papers which led to an unusual scatter of Bortkiewicz' (and Chuprov's) contributions (and to difficulties of getting hold of them), and, as I myself believe, prompted authors to cram as much as possible into such works.

But then, Bortkiewicz' critical review of Pareto (1898c) was lamely arranged. Chuprov (Letter 35 of 1898) indicated this circumstance, but Bortkiewicz' answer in the next letter was of no consequence. Anderson (1932, p. 245/1963, p. 533) described the situation:

Bortkiewicz schrieb nicht für die weite Öffentlichkeit und war durchaus kein guter Popularisator seiner eigenen Ideen. Er stellte ferner sehr hohe Ansprüche an die Vorbildung und Intelligenz seiner Leser. Mit einer Hartnäckigkeit … weigerte er sich, den Rat … Tschuproff anzunehmen und eine leichtere äußere Form für seine Veröffentlichungen zu wählen. … Hierzu kommt noch, dass Titel und Ort der Veröffentlichungen durchaus nicht immer dem entsprechen, was der Leser billigerweise dort suchen könnte.

The material of this subsection negates Bortkiewicz' caring for his readers (§ 2.1).

2.5. Knowledge of literature. Concerning subjects which interested him, Bortkiewicz was exceptionally knowledgeable about their history and actual situation (Schumacher 1931, Meerwarth 1936). In connection with his own work which, for that reason as well as because of his mentality, was extraordinary, Altschul (1928, p. 1225) concluded: *wie souverän Bortkiewicz weit auseinanderliegende Forschungszweige beherrscht*. Another result was his nickname *Pope of statistics*, see above. The scope of his knowledge was *ungeheuer* (Anderson 1932, p. 244/1963, p. 532), *enzyklopädisch* (Altschul 1931, p. 1183).

Обширностьта на познанията … и кругът на неговите научни интереси су наистина громадни

(The scope of his knowledge … and the domain of his scientific interests are really enormous) (Андерсон 1929, с. 7).

Indeed, Bortkiewicz discussed the work of Aristotle (1906b), Leibniz (1907) and many statisticians and economists; see his study of Marx in § 3.1.

2.6. The negative aspect. It is necessary to add some negative circumstances to Bortkiewicz' *Hartnäckigkeit* (Anderson, see end of § 2.4). The *Gründlichkeit* (Gumbel, § 2.4) was apparently sometimes lacking. He himself (Letter 6 of 1898) agreed with Chuprov: *я и сам сознаю некоторую неоконченность в этих* [прежних своих] *статьях.* And in the same letter, in connection with the not yet published booklet (1898b): *Можно будет впоследствии напечатать Neue Untersuchungen* … [Новое исследование о законе малых чисел] *или Abermals* … [снова закон малых чисел]. Another case concerns his accusation of plagiarism by Gini: in his *great treatise* (1930a), as Andersson (1931, p. 17) called it, on the distribution of incomes, he had not referred to Gini (1912). Andersson had described in detail the whole episode and completely exonerated Bortkiewicz who died soon afterwards. His answer (1931) to Gini appeared posthumously. But still, this is not the whole story. Chuprov received a reprint of Gini's paper, (too) briefly described it to Bortkiewicz (Letter 122 of 1913) and added: *Я могу выслать тебе Джини, буде* [если] *не найдёшь его в библ*. In the next letter Bortkiewicz repeated that Gini's work [or rather the source where it appeared] was not available *в здешней Корол. Библ.* (in the present Staatsbibliothek zu Berlin), so that he has *полное право означенных статей не касаться.* A strange attitude! In spite of their heated discussion of the law of small numbers twenty years ago, he should have mentioned Gini as his possible predecessor.

2.7. Anti-mathematically inclined scientists. At the beginning of the 20th century the curriculum of German student-economists had not included mathematics, and German economists did not, and had no wish to understand that science. Anderson (1932, p. 245/1963, p. 533) mentioned *amathematisch angelegten deutschen Volswirtschaftler*. Earlier he (1929, p. 8) indicated the *неизкоренимата антипатия, която громадното болшинство от немските икономисти храни към математиката*.

(Ineradicable antipathy of a great majority of German economists to mathematics.)

There also, and later (1932, p. 243/1963, p. 531), Anderson stated: *Нашето поколение статистици мучно може да си представи онова блато, в което попада статистическата наука след крушението системата на Кетле и от което тя е била изтеглена само от Лексиса и Борткевича*.

(Our generation of statisticians is hardly able to imagine that mire in which the statistical theory had got into after the collapse of the Queteletian system, or the way out of it which only Lexis and Bortkiewicz have managed to discover.)

Still earlier Чупров (1909/1959, p. 215) indicated roughly the same. Yes, in 1874, after Quetelet had died, German authors declared useless his modest stochastic ideas and (not always) correct conclusions. Statisticians did not wish to delve in the essence of their studied phenomena and restricted their efforts to considering the simplest Bernoulli pattern of trials. Most of them were certainly insufficiently knowledgeable about mathematics, and I mention Knapp and von Mayr.

Knapp abandoned the mathematical direction of statistics, see Letter 30 of 1898 and Шейнин (1990/2011, § 7.2). And now Letter 109 of 1911: as mentioned by Bortkiewicz, his report *заставил Майра выступить с заявлением о ненужности математики в статистике*. And, also there: *самым частным образом* von Mayr told him that *ещё больше, чем математику, он не выносит* [ещё меньше … выносит] *современную теорию познания*.

About 1916 von Mayr, as the Editor of the *Allgemeines statistisches Archiv*, rejected a paper by Bortkiewicz. Anderssen (1931, pp. 14 – 15) quoted the latter's letter to von Mayr in translation:

I have been presented as a mathematician with no understanding for the "State science of statistics".

And still, von Mayr earlier admitted mathematically oriented papers.

This is the first time anything of the kind has happened to me. … I regard my connection with the German Statistical Society, whose organ is the Archives, as severed.

In this atmosphere, Bortkiewicz (Андерсон 1929, p. 8) *се явава … до известна степень "чуждородно тело" и по-скоро требва да се признае за един международен или даже руски, отколкото германски профессор. (От англичаните … Борткевич се различава по високите изисквания – пак в духа на руските математици)*. He (p. 7) also noted that the *научната работа на Борткевич е от своеобразно естество и се доближава донекуде до маниера на Еджворт*.

(To a certain extent an *alien body* and he should be recognized as an international or even Russian rather than German professor. (From the English … Bortkiewicz differs by higher self-requirements, again in the spirit of Russian mathematicians.)

And in 1931 Андерсон remarked: *за границей* [Борткевич] *пользовался несравненно большим признанием, чем в пределах самой Германии (где у него почти не было учеников)*. This has been occurring in spite of his having adapted to German life (Tönnies 1932/1998, p. 319):

Bortkiewicz war ein Mann von nicht alltäglichem wissenschaftlichen Ernst. Die "deutsche" Gründlichkeit hatte in ihm wie in vielen anderen, die nicht durch ihre Geburt zu uns gehören, einen ihrer besten Vertreter. Er war aber durch seine eigene Wahl ein guter Deutscher geworden, auch als bewusster Staatsbürger der deutschen Republik, und hat der Deutschen Demokratischen Partei angehört.

Andersson (1931, p. 11) noted that his work *has been and is of great importance to all the Nordic countries and especially to Sweden*.

2.8. Membership in scientific bodies. In accordance with his scientific activity as well as position in the world, Bortkiewicz had been member of the Swedish Academy of Sciences, the Royal Statistical Society, American Statistical Association and International Statistical Institute. Tönnies (1932/1998, p. 316) also mentioned the

Straßburger wissenschaftlichen Gesellschaft zu Frankfurt a. M. And Bortkiewicz had severed his connection with the German Statistical Society (§ 2.7). His activities in the German Verein f. Versicherungs-wissenschaft are mentioned in § 3.2.

2.9. The archive: lost and found. Andersson (1931, pp. 11, 25 – 26) described Bortkiewicz' rich collection of documents, but did not provide its whereabouts. Apparently, it was still in Berlin (and he quoted some of the documents in translation, see above). Among Bortkiewicz' correspondents Andersson named Scandinavian scientists Frisch, Guldbeg, Meidell, Steffensen, Westergaard and K. and S. Wicksell and called the letters from Walras and Chuprov *most important*. The letters of Bortkiewicz himself (apparently drafts) were also extant and contained *references to fundamental scientific questions*. Finally, there were the texts of lectures on *statistics, social politics, economics and technical insurance … in perfect order*.

G. Rauscher (Vienna) discovered that the archive was in Uppsala, Sweden. Also there, as the University Librarian Doctor Hallberg informed me, are the books which belonged to Bortkiewicz and were sold to them by Helene Bortkiewicz with the assistance of Andersson. The catalogue of the archive is available on request. It lists letters from many other scientists, for example Mahalanobis and Cantelli, from the Russian statisticians A. A. Kaufman, M. V. Ptukha and N. S. Chetverikov, see a brief description in Борткевич и Чупров (2005, p. 10). There also is the correspondence of L. B. with Slutsky, now published (Виттих и др. 2007). Also published are the Walras letters (Jaffé 1965) and the just mentioned book of V. I. and Chuprov is their correspondence kept in Moscow and Uppsala. Letters from L. B. during 1919 – 1926 are lacking; after Chuprov left Prague for a short time (as he thought), he never came back, and those letters, if he had kept them, could have disappeared.

3. The results

3.1. Economics. Here is how Gumbel (1931, p. 232) described Borkiewicz' study of Marx (1906 – 1907):

Wie Lexis, hat er keinen der populären Einwände, welche so häufig gegen Marx erholen wurden, mitgemacht. Als erster hat er das dürre Gerüst der Marxschen Schemata in eine mathematische form gekleidet und das Verfahren zur Umrechnung der Werte in Produktionspreise und zur Bestimmung der Durchschnittsprofitrate nachgeprüft.

Much later he (1968, p. 25) repeated his description in more detail and added:

He made the necessary modifications that rendered the Marxian scheme of surplus values and prices consistent. However, his dry presentation prevented the Marxists (*except for Klimpt* [1936]) *from accepting his method*.

Gumbel concluded his account by stating that L. B. had *made the lonely effort to construct a Marxian econometry* without applying statistical data! Nevertheless, that effort was noteworthy.

Загоров (1929, p. 12), in his French résumé: L. B.

P*rouve que la méthode de Marx de deduction des prix de la valeur est fausse. Les conceptions positives de* [L. B.] *sur la matière sont exposées dans son article (1921), où il s'efforce de réconcilier les* *deux theories contraires bien continues sur la formation des prix: la théorie des frais de production et la théorie de l'utilité*.

In Letter 79 of 1905 L. B. mentioned his future study of Marx, a paper of *теоретического характера, в которой хочу коснуться нового сочинения Туган–Барановского* [1905].

Mises (1932, p. 15) briefly stated that L. B. had made *a Versuch einer positiv gewendeten Zusammenfassung* of the *Werththeorie* and that his criticism (1906a) of Böhm-Bawerk and study of Marx are *von bleibenden Bedeutung*. Mises also maintained that in economics L. B.

Schon zu einer Zeit, in der in Deutschland die historische Schule ganz überwiegend herrschte, zu pflegen wusste.

The very fact that this outstanding scholar published Bortkiewicz' obituary is interesting. On p. 11 Загоров commented on another subject:

В областьта на парите [денег] *Борткевич се явява защитник на умерения метализъм срещу крайния номинализъм.*

(Concerning the theory of money, Bortkevich protected modest metallism against extreme nominalism.)

He also opposed *безогледното приложение на математиката в стопанските* [экономические] *изучвания* and formulated a general opinion on an important issue: *от идеята за пределна полза не може да се изведе наложително и едино целата стопанска теория*.

(reckless application of mathematics to economic studies … it is impossible to derive an urgently needed unified economic theory by issuing from the ideas of marginal utility.)

Many commentators pronounced their opinion about Bortkiewicz' entire work in economics, for example Schumacher (1931, p. 573), Altschul (1928, pp. 1225 and 1226) and Schumpeter (1932, pp. 339–340):

1) L. B. *hat in der Nationalökonomie nicht nur Deutschlands, sondern der Welt eine höchsteigenartige, ja einzigartige Stellung eingenommen; ich wusste niemanden aus der Gegenwart und Vergangenheit zu nennen, der ihm zur Seite gestellt werden könnte, und auch in der Zukunft wird das kaum anders sein.*

But he forgot Chuprov.

2) [Er] *gehört zu den interessantesten und eigenartigsten Gestalten des deutschen Gelehrtentums. Einer der bedeutendsten Nationalökonomen, der zu grundlegenden wirtschaftstheoretischen Fragen in entscheidender Weise Stellung genommen hat. … Alle diese Arbeiten sind auch noch heute … nicht nur lesenwert, sondern bieten zahlreiche Anknüpfungen und Anregungen für die weitere Forschung*.

3) *He upheld the flag of economic theory – professing the Marshallian creed – at an epoch and in a country, in which hardly anyone would hear of it, and he cleared the ground of many battlefields by his powerful sword. By far his most important achievement is his analysis of the theoretical framework of the Marxian system … much the better thing ever written on it and, incidentally, on its other critics. A similar masterpiece is his paper on the theories of rent (1910). … As a writer on monetary theory and policy, he ranks high among German authors. The subjects of the gold*

standard, of banking credit, of velocity of circulation owe much to him. The best he did in this field, however, is his work on index numbers (1923 – 1924).

3.2. The science of insurance. Andersson (1931, p. 12) stated that Bortkiewicz' contributions

Must be regarded as among the most valuable assets of the science of insurance.

Lorey (1932, note 6 and pp. 202 – 204) listed those contributions and described Bortkiewicz' collaboration with the *kurz vorher gegründeten Deutschen Verein f. Versicherungswissenschaft*. He joined it in ca. 1902, from 1903 was a member of its *Ausschuß* and from 1926, the *Vorstand der Abteilung f. Versicherungsmathematik*.

Gumbel (1931, p. 231) stressed the importance of Bortkiewicz' study of mortality tables and singled out his contributions about the connection between *Fehlerausgleichung und Untersterblichkeit* [1910 – 1912], mortality of the *Empfänger von Invalidenrenten* [1899], and *Deckungs-methoden der Sozialversichung* [1909] to which his work in Russia (§ 1.1) should likely be added. Much of the above Gumbel repeated later (1968).

From the very beginning of his scientific work, L. B. had been studying population statistics. He published two papers (1890b; 1891) in a periodical of the Petersburg Academy; a bit earlier (1889, p. 1056) and later (1898a) he severely criticized Buniakovsky's mortality tables. The latter had stressed that his sources were incomplete and inaccurate and that he was dissatisfied with his results; yes, but he made serious methodological mistakes, see Sheynin (1991, pp. 212–213; 1991/1999, p. 71).Also see Bortkevich (1893a; 1893b).

3.3. Theory of probability, statistics and mathematical statistics. Mises (1932) stated without elaborating that

Sehr schöne Leistungen hatte v. Bortkiewicz auf dem Gebiete der angewandten Wahrscheinlichkeitsrechnung und der mathematischen Statistik aufzuweisen and that he *zweifellos zählte* among the *bedeutendsten Vertretern* of mathematical statistics.

Anderson (1931) called him

Одним из крупнейших и в то же время своеобразнейших теоретиков статистики, место которого в одном ряду с Кетле, Лексисом, Чупровым и К. Пирсоном.

A bit later he (1932, p. 242/1963, p. 530) attributed L. B. to

Einen der wenigem wirklich Großen im Bereiche der mathematischen Statistik and maintained that i*n der theoretischen Statistik war er der anerkannte Meister und Führer* [of the Continental direction of statistics].

Apparently Anderson sees here no difference between mathematical and theoretical statistics. I believe that the former, unlike the latter, excludes data analysis and collection of data; L. B. did not study these subjects. Again, in § 2.7 I quoted Anderson's description of Bortrkiewicz' merit in rescuing statistics from the mire of its previous existence. Finally, Anderson (1954/1957, p. 97) mentioned the *alte deutsche mathematisch-statistische Schule von Lexis, Bortkiewicz und Chuprov.*

The theory of probability was Bortkiewicz' *Lieblingsgebiet* (Altschul 1928, p. 1225). He (Ibidem) remarked that L. B. *eine … für die Physik methodologisch richtungsgebende Arbeit vollbracht hat*. And he (1931, p. 1183) maintained that

Als Interpret der Wahrscheinlichkeitstheorie – ein Mathematiker von internationalem Ruf war, mitten aus einer ungewöhnlich vielseitigen, fruchtbaren und bis in die letzte Zeit hinein in neue Gebiete vordringenden Forscherarbeit … der Wissenschaft entrissen worden.

Some of his contributions to economics directly belonged to mathematical statistics as well. I mention the papers on index numbers (1923 – 1924) and distribution of incomes (1930a). One of Bortkiewicz' subjects was the application of the theory of probability to statistics (1898b), see below, and (1904). Among other subjects were the theory of series (1917) and radioactivity (1913), *von bleibendem Werte* (Anderson 1932, p. 244/1963, p. 532). L. B. *подари на физиците една забележителна работа* (Андерсон 1929, p. 7). (He presented physicists a remarkable work). *In Kreise der Physiker* [it had] *großes Ansehen genießt* (Altschul 1931, p. 1183).

Gumbel (1968, p. 24) maintained that L. B. *did classic work in mathematical statistics* and (p. 26) concluded his note by maintaining that

Four of his contributions are decisive: the proof that the Poisson distribution corresponds to a statistical reality (1898b); the introduction of mathematical statistics into the study of radioactivity (1913); the inception of the statistical theory of extreme values [1921]; *and the lonely effort to construct a Marxian econometry* [§ 3.1].

Much of the above belonged to statistics as well as to mathematical statistics, and here Bortkewich' work on mortality (§ 1.1) could be added. One of Anderson's statements seems doubtful: L. B. allegedly

Работил много в областьта на … морална статистика (за което се относят не по-малко от 16 оригинални негови работи).

(He worked much in the region of … moral statistics and published there not less than 16 original contributions.) I only know that he repeatedly studied the statistics of suicides. And Bortkiewicz had published many reviews (some of which undoubtedly remain unknown) on most various statistical subjects.

3.3.1. Stability of series and the law of small numbers. In 1879, Lexis suggested a test, *Q,* for checking whether the probability of the studied event in statistical series remained constant. Suffice it to express *Q* as the ratio of two dependent random variables, ξ/η. He thus originated the mathematical direction of statistics independent from the future Biometric school. L. B. participated in the pertinent studies and, in particular, introduced his law of small numbers, LSN (1898b). For several decades that law continued to be the talk of the town, but Kolmogorov (1954) without elaborating called it *устаревшим названием предельной теоремы Пуассона*, and I (2008) substantiated his opinion. I describe his work on the Lexian theory in general as commented on by Markov, then turn to the LSN, but first I repeat only one of my findings.

L. B. had introduced his own test, Q', not coinciding with the Lexian Q; thus, unlike the latter, Q' could not be less than 1 (1898b, p. 31). Later Bortkiewicz (1904, p. 833) noted that EQ = Q' but mistakenly justified this equality by believing that, for dependent random variables ξ and η, Eξ/η = Eξ/Eη. Then, he (1918, p. 125n) only admitted that the equality was insignificantly approximate. Chuprov (1922) devoted a paper to that subject.

Lexis hardly thought about calculating the mean value and variance of Q (and in any case that was a serious problem). In 1916, Markov, and much better Chuprov derived EQ. L. B. made such calculations much earlier, but his results were only roughly correct.

I turn to Ondar (1977) who published the correspondence between Markov and Chuprov in which Bortkevich was mentioned many times and quoted by Markov. While rendering the texts, Ondar made a great many mistakes and allowed himself numerous fudges. I (1990/2011, chapter 8) corrected all this and added a few letters which he missed. Only my first example is from that latter source.

1) Letter 3a, 6.11.1910, p. 103. *Полагаю, что Бортке*в. <u>*не*</u> *авторитет по теории вероятностей.*

2) However, see Letter 47, 25 (not 26), 11. 1912, p. 64. [Некоторые] *вычисления Борткевича я признал заслуживающими внимания.* And, see Letter 50, 2.12.1912, p. 67: *Я жестоко ошибался относительно значения работ Борткевича.*

3) Letter 11, 18.11.1910, p. 25. *Измышления Лексиса и Борткевича* … However, Chuprov, in Letter 14, 19.11.1910, p. 28, explained that the premises of V. I. and Markov were different.

Markov also criticized the LSN.

4) Letter 66 of 1916, p. 82. He remarked that V. I. основывался на *скудном* материале and that his conclusions were *весьма сомнительны*.

5) Letter 68, 5.3.1916, p. 85. He mentioned the *пресловутый* [notorious] booklet (1898b).

6) His main criticism. Letter 69, postmark of same date, p. 86. *При малых числах* [Q] *не может быть большим* (which was essential for the Lexian theory). Марков (1916) published a pertinent paper.

For his part, Chuprov noted that the LSN admitted four interpretations and that, see Letter 69a, 1916 (Шейнин 1990/2011, pp. 91 – 92), L. B. was reluctant to discuss that law. Even in 1914, in Letter 135 (Борткевич и Чупров 2005), evidently answering Chuprov, he maintained: *о сдаче в архив Q я с Тобой совершенно не с огласен.* Nevertheless, only on 31 Jan. 1921, in Letter 38 to his friend, K. N. Gulkevich, Чупров (2009, c. 88) stated:

Одно из важнейших учений теории статистики, которое я доселе всецело принимал и исповедовал, – лексисова теория устойчивости статистических чисел, – оказывается в значительной мере покоится на математическом недоразумении.

(**One of the most important doctrines of the theory of statistics, which I had been until now adopting and manifesting, the Lexian theory of statistical series, – that theory, as it occurred, essentially rests on a mathematical misunderstanding.**)

Now, Bortkiewicz several times stressed (which is evident in his booklet on the LSN) that his innovation was closely linked with the Lexian theory. I conclude: L. B. became entangled but never admitted his mistakes.

Андерсон (1929, с. 9) politely remarked that *практическо значение* [of that law] *да е много помалко от … основни трудове* [of L. B.] and largely repeated his statement later (1932, p. 243/1963, p. 531):

Dessen praktische Bedeutung nicht als so groß herausgestellt hat, wie es anfänglich erschien.

Mises (1932) noted that L. B. had

Eine lange vernachlässigte Seite der statistischen Betrachtungsweise in der Vordergrund des Interesses brachte.

Yes, the Poisson limit theorem had been all but forgotten and Bortkiewicz proved that important statistical subjects (obey its premises and therefore) can be studied by applying it.

Bibliography

Борткевич В. И.; Bortkiewicz L. von; Władysław Bortkiewicz

JNÖS = *Jahrbücher für Nationalökonomie und Statistik*

1889, О русской смертности. *Врач*, т. 10, № 48, с. 1053–1056.
1890a, Auseinandersetzung mit Walras. *Rev. d'écon. politique*, t. 4.
1890b, Смертность и продолжительность жизни мужского православного населения Европейской России. *Зап. Имп. Акад. Наук*, т. 63, Прил. 8. Отдельная пагинация.
1891, То же название для женского населения. Там же, т. 66, Прил. 3. Отдельная пагинация.
1893a, *Die mittlere Lebensdauer* (*Staatswissenschaftliche Studien*, Bd. 4, No. 6). Jena.
1893b, Russische Sterbetafeln. *Allg. stat. Archiv*, Bd. 3, pp. 23–65.
1898a, Das Problem der Russischen Sterblichkeit. *Allg. stat. Archiv*, Bd. 5, pp. 175–190, 381–382.
1898b, *Das Gesetz der kleinen Zahlen.* Leipzig.
1898c, Die Grenznutzentheorie als Grundlage einer ultraliberalen Wirtschaftspolitik. *Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft im Deutschen Reich*, Jg. 22, pp. 1177–1216.
1899, Über die Sterblichkeit der Empfänger von Invalidenrenten vom statistischen und versicherungstechnischen Standpunkte. *Z. für Versicherungs–Recht und Wissenschaft*, Bd. 5, pp. 563–605.
1901, O stopniu dokładności spółczynnika rozbieżnosci. *Wiadomości Matematyczne*, t. 5, pp. 150–157.
1903, Теория вероятностей и борьба против крамолы. *Освобождение* (Штутгарт), кн. 1, с. 212–219. Статья опубликована в части тиража. Подписано Б.
1904, Anwendung der Wahrscheinlichkeitsrechnung auf Statistik. *Enc. math. Wiss*. Bd. 1/2, pp. 822–851. Submitted 1901.
1906a, Der Kardinalfehler der Böhm-Bawerkschen Zinstheorie. *Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft im Deutschen Reich*, Jg. 30, pp. 943–972.
1906b, War Aristoteles Malthusianer? *Z. für d. ges. Staatswissenschaft*, Bd. 62, pp. 383–406.
1906–1907, Wertrechnung und Preisrechnung im Marxschen System. *Archiv für Sozialwissenschaft und Sozialpolitik*, Bd. 23, pp. 1–50; Bd. 25, pp. 10–51, 445–488. Reprint: Achenbach, 1976.
1907, Wie Leibniz die Diskontierungsformel begründete. *Festgaben für W. Lexis*. Jena, pp. 59–96.

1909, Die Deckungsmethoden der Sozialversicherung. *VI Intern. Kongress f. Versicherungs-Wissenschaft*, Bd. 1, pp. 473 – 505.
1910, Die Rodbertus'sche Grundrententheorie und die Marx'sche Lehre von der absoluten Grundrente. *Archiv für die Geschichte des Sozialismus und der Arbeiterbewegung*, Bd. 1, pp. 391–434.
1910–1912, Über den angeblichen Zusammenhang zwischen Fehlerausgleichung und Untersterblichkeit. *Z. f. die ges. Versicherungs-Wissenschaft*, Bd. 10, pp. 559–564; Bd. 12, pp. 747–752.
1913, *Die radioaktive Strahlung als Gegenstand wahrscheinlichkeitstheoretischer Untersuchungen*. Berlin.
1915, W. Lexis zum Gedächtnis. *Z. für die ges. Versicherungs-Wissenschaft*, Bd. 15, pp. 117–123.
1915, Realismus und Formalismus in der mathematischen Statistik. *Allg. stat. Archiv*, Bd. 9, pp. 225–256.
1917, *Die Iterationen.* Berlin.
1918, Der mittlere Fehler des zum Quadrat erhobenen Divergenzkoeffizienten. *Jahresber. der deutschen Mathematiker–Vereinigung*, Bd. 27, pp. 71–126.
1919, Zu den Grundrententheorie von Rodbertus und Marx. *Archiv für die Geschichte des Sozialismus und der Arbeiterbewegung*, Bd. 8, pp. 248–257.
1921, Variationsbreite und mittlerer Fehler. *Sitz. Ber. Berliner math. Ges.*, Jg. 21, pp. 3–11.
1923 – 1924, Zweck und Struktur einer Preisindexzahl. *Nordisk Statistisk Tidskrift*, Bd. 2, pp. 369–408, Bd. 3, pp. 208–251, 494–516.
1930a, Die Disparitätsmasse der Einkommensstatistik. *Bull. Intern. Stat. Inst.*, t. 25, No. 3, pp. 189–298, 311–316.
1930b, *Anwendung der Versicherung auf das Problem der übermäßigen Grundbesitzzerstückelung*. Warschau. (In German and Polish)
1931, Erwiderung. *Bull. Intern. Stat. Inst.*, t. 25, No. 3, pp. 311–316.
Борткевич В. И., Чупров А. А. (2005), *Переписка* (*1895–1926*). Берлин. **S, G,** 9.

Other authors

Андерсон О. Н., Anderson O. (1929), Професор В. Борткевич. *Тримесячно списание на Главната дирекция на статистиката*, година 1, кн. 1. София, с. 7–9. S, G, 17 (in Russian)
– (1931), Профессор В. И. Борткевич как статистик. *Россия и славянство*, 15 авг. 1931, с. 3. S, G, 17 (in Russian)
– (1932), Ladislaus von Bortkiewicz. *Z. f. Nationalökonomie*, Bd. 3, pp. 242–250. Reprint: *Ausgew. Schriften*, Bd. 2. Editor H. Strecker. Tübingen, 1963, pp. 530–538. **S, G,** 36.
– (1954), *Probleme der statistischen Methodenlehre in den Sozialwissenschaften*. Würzburg, 1957, 3rd edition. Later editions 1962, 1965.
Аноним (1927), Борткевич. БСЭ, изд. 1–е, т. 7, с. 198.
Виттих К., Раушер Г., Шейнин О. Б. (2007), Переписка Е. Е. Слуцкого и В. И. Борткевича. *Финансы и бизнес*, № 4, с. 139–154. (There are errors in the published text.) **S, G,** 40.
Загоров С. (1929). Борткевич като икономист. *Тримесячно списание на Главната дирекция на статистиката*, година 1, кн. 1. София, с. 10–12. **S, G,** 36.
Колмогоров А. Н. (1954), Малых чисел закон. БСЭ, 2–е изд., т. 26, с. 169. Published anonymously.
Марков А. А. (1916), О коэффициенте дисперсии для малых чисел. *Страховое обозрение*, № 2, с. 55–59. **S, G,** 5.
Мордух Я. (1923, Russian**)**, On connected trials complying with the condiitoin of stochastic commutativity. *Trudy Russk. uchenykh zagranitsei*, vol. 2. Berlin, pp. 102 – 125. **S,G,** 6.
Ондар Х. О. (Ondar Kh. O.), редактор, editor (1977, Russian), *Correspondence between Markov and Chuprov*. New York, 1981.
Покотилов А. Д. (1909), *Первый опыт государственного страхования в России. Десять лет пенсионной кассы служащих на казенных железных дорогах по операциям страхования жизни*. СПб. Review: W. Idelson, *Z. f. die gesamte Versicherungs-Wissenschaft*, Bd. 10, 1910, p. 169.

Старовский В. Н. (1933), Экономическая статистика. БСЭ, 1–е изд., т. 63, с. 279–283.
Туган–Барановский М. И., Tugan–Baranovky M. (1905), *Theoretischen Grundlagen der Marxismus*. Leipzig. Четыре сокращенных русских изданий: *Теоретические основы марксизма*. СПб и Москва, 1905–1906.
Четвериков Н. С., составитель и переводчик (1968), *О теории дисперсии*. М.
Чупров А. А., Chuprov A. A. (1909), *Очерки по теории статистики*. М., 1910, 1959.
– (1916), О математическом ожидании коэффициента дисперсии. *Изв. Имп. АН*, т. 10, № 18, с. 1789–1798. **S, G,** 35.
- (1918–1919), Zur Theorie der Stabilität statistischer Reihen. *Skand. Actuarietidskr.*, Bd. 1, pp. 199–256; Bd. 2, pp. 80–133. Русский перевод в книге Четвериков (1968, с. 138–224).
– (1922), О математическом ожидании частного двух взаимно зависимых случайных переменных. *Тр. русск. ученых за границей*, т. 1. Берлин, с. 240–271.
– (2009), *Письма К. Н. Гулькевичу*, 1919–1921. Берлин. Публикация К. Виттиха, Г. Кратца, О. Б. Шейнина.
Шейнин О. Б. (1990), *А. А. Чупров: жизнь, творчество, переписка.* М., 2010. English translation: Göttingen, 1996, 2011.
- (1991), On the work of Buniakovsky in the theory of probability. *Arch. Hist. Ex. Sci.*, vol. 43, pp. 199–223. Russian translation: *Историко–математич. исследования*, вып. 4 (39), 1999, с. 57–81.
– (2003), *Гумбель, Эйнштейн и Россия.* The text in English and in Russian. Only in English: **S, G,** 12.
- (2008а), Bortkiewicz' alleged discovery: the law of small numbers. *Hist. Scientiarum*, vol. 18, pp. 36–48.
Altschul E. (1928), L. v. Bortkiewicz. *Magazin der Wirtschaft*, 7. Jg, pp. 1225–1226.
– (1931), Ladislaus v. Bortkiewicz. Ibidem, pp. 1183–1184.
Andersson T. (1931), Ladislaus von Bortkiewicz, 1868–1931. *Nordic Statistical J.*, vol. 3, pp. 9–26; *Nordisk Statistisk Tidskr.*, Bd. 10, pp. 1–16.
Gini C. (1912), Variabilità e mutabilità. Contributo allo studio delle distribuzioni e relazioni statistiche. *Studio Economico–Giuridici. Univ. Cagliari*, t. 3.
Gumbel E. J. (1931), L. von Bortkiewicz. *Deutsches statistisches Zentralblatt*, No. 8, pp. 231–236.
– (1968), Ladislaus von Bortkiewicz. In: Kruskal W. H., Tanur Judith M., editors. *Intern. Enc. of Statistics*, vol. 1. New York, 1978, pp. 24–27.
Jaffé W., editor (1965), *Correspondence of Leon Walras and related papers*, vols. 2–3. Amsterdam.
Keynes J. M. (1921), *Treatise on Probability*. Reprint: *Coll. Works*, vol. 8. London, 1973.
Klimpt W. (1936), *Mathematische Untersuchungen im Anschluss an L. von Bortkiewicz über Reproduktion und Profitrate.* Berlin. Dissertation 1930/1931
Lexis W. (1879), Über die Theorie der Stabilität statistischer Reihen. JNÖS, Bd. 32. In author's *Abhandlungen zur Theorie der Bevölkerungs– und Moralstatistik*. Jena, 1903, pp. 170 – 212. Translation into Russian: Лексис В., О теории стабильности статистических рядов. В книге Четвериков (1968, с. 5–38).
Lorey W. (1932), Ladislaus von Bortkiewicz. *Versicherungsarchiv*, Bd. 3, pp. 199–206.
Meerwarth R. (1936), Ladislaus von Bortkiewicz, 1868–1931. *Bull. Intern. Stat. Inst.*, t. 26, No. 1, pp. 254–258. Доклад 1931 г.
Mises R. von (1932), Ladislaus von Bortkiewicz. *Chronik der Friedrich–Wilhelm Univ. zu Berlin*, 1931/1932, pp. 14–15. S, G, 17 (in Russian)
Neyman J. (1931), Pamięci profesora dr. Władysława Bortkiewicza. Referat przygotowany na II zjazd matematyków Polskich w Wilnie, we wrześniu 1931. *Kwartalnik statystyczny, Revue trimestrielle de statistique*, t. 8, pp. 1116–1118.
Schumacher H. (1931), Ladislaus von Bortkiewicz. Gedächtnisrede. *Allg. stat. Archiv*, Bd. 21, pp. 573–576.
Schumpeter Jos. A. (1932), Ladislaus von Bortkiewicz (Aug. 7, 1868–July 15, 1931). *Econ. J.*, vol. 42, pp. 338–340.
S. P. (1931), Władysław Bortkiewicz. *Kwartalnik statystyczny, Revue trimestrielle de statistique*, t. 8, pp. 1118–1120.

Tönnies F. (1932), Ladislaus v. Bortkiewicz, 1868–1931. Reprint: *Gesamtausgabe*, Bd. 22. Berlin, 1998, pp. 315–319.
Voigt G. (1994), *Russland in die deutschen Geschichtsschreibung 1843–1945.* Berlin.
Winkler W. (1931), Ladislaus von Bortkiewicz als Statistiker. *Schmollers Jahrbuch f. Gesetzgebung, Verwaltung und Volkswirtschaft im Deutschen Reiche*, 55. Jg, pp. 1025–1033.
Woytinsky W. S. (1961), *Stormy Passage*. New York.

Note. I mentioned Bortkiewicz many times in my later contribution *Theory of Probability. Historical Essay*. Berlin. **S, G,** 11.

VII

Liapunov's Letters to Andreev

Istoriko-Matematicheskie Issledovania, vol. 31, 1989, pp. 306 – 313

The life of Aleksandr Mikhailovich Liapunov is well known [1; 2]. Konstantin Alekseevich Andreev (1848 – 1921) graduated from Moscow University. From 1873 to 1898 he taught in Kharkov, then in Moscow. In 1884 – 1899 he was President of the Kharkov Mathematical Society (KhMSoc) and Editor of its *Soobshchenia* [3]. In 1884 he was elected Corresponding Member of the Imperial (Petersburg) Academy of Sciences.

Gordevsky published passages from Andreev's letters to Liapunov [3][1] but provided no commentaries and neither did he say anything about Liapunov's letters to Andreev which I discovered in the Archive of the Moscow State University.[2]

Andreev's first letter (17.2.1899; all dates are given here in the old style) is not connected with the newly discovered letters of Liapunov, and I leave it alone. Liapunov's first letter (11.6.1897) is also left out since it was only concerned with his stay in Crimea with one of his brothers and described how useful was it for his brother's health. He did not name his brother, but a paper by Academician Boris Mikhailovich Liapunov [1, p. 11], a philologist, mentioned a trip to the Crimea in May – June 1897 by A. M., his wife, mother-in-law and himself (but did not provide any details).

Almost all of Liapunov's subsequent letters are connected with the appearance of his paper [4] as a response to Nekrasov's criticisms [5]. Another subject was the preparation of a new charter for Russia's universities. Liapunov participated in the work of the appropriate commission at Kharkov University [1, p. 11].

On 29 April 1901 the Ministry of Public Education circulated proposals concerning the new charter [6, pp. 1 – 4], but even before that some universities had begun to discuss the causes of the then occurring students' unrest and to suggest "measures for putting university life in order" (Ibidem, p. 5).[2a]

Pavel Alekseevich Nekrasov (1853 – 1924) was an eminent scholar, a professor at, and for several years rector of Moscow University, but at the turn of the 19th century his scientific work (only in the theory of probability and statistics) underwent a radical change. He started connecting mathematics with religion and politics and became unimaginably verbose so that his writings of that period are still unstudied. Here is Pavel Youshkevitch's pertinent opinion from his forgotten newspaper article [7]: Nekrasov is

A great lover of philosophy […] *but the philosophy of this honourable scholar is of an absolutely special nature. It is a strangest medley of senseless profundity with tedious verbiage and dried up words*.

Nekrasov [5] blamed Liapunov for mistakes and shortcomings allegedly committed in the classical memoir [8]. His considerations were however either unfounded or indefinite, or did not bear any relation to the substance of Liapunov's work. A similar conclusion can be made about Nekrasov's accusations of Chebyshev and Markov which he adduced for good measure. Being blinded by his criticism, Nekrasov even mixed up the notions of limit and asymptotic representation of a function.

I reproduce now Liapunov's letters

Liapunov – Andreev, 29.3.1901

Highly respected Konstantin Alekseevich,

I am applying to you with a great request. You have probably already received an offprint of Nekrasov's note [5] where he makes charges against me, Markov and Chebyshev. If you had acquainted yourself with my papers you certainly noticed that Nekrasov completely perverts the truth and does it with impudence beyond any measure. Although I have not the slightest desire to enter into a debate with him, I am compelled to answer him. And I am therefore sending you a manuscript of my Answer and am asking you to assist in its publication in the *Matematichesky Sbornik* , – if possible, in the same issue where Nekrasov's note is to appear. I would not have troubled you with this request if it were possible to publish my Answer in the *Soobshchenia*.[3] But our relations with Silberberg have definitely deteriorated and we are compelled to postpone the printing of the *Soobshchenia* at least for the time being.[4]

This is why I have to apply for help to the *Matematichesky Sbornik*. I think, however, that I have some right to do so since I am a member of the Moscow Mathematical Society [MMS].[5] You will perhaps find the tone of my Answer rather sharp. But what can I do? I attempted to assuage it as much as possible, but everything has its boundary and I cannot go any further in this direction. It is already sufficient that I do not accuse Nekrasov of deliberately perverting the truth (about which I have no doubt) and that I explain his strange attacks against me by his ignorance of the substance of my first paper.

Understanding that my request is putting you to some inconvenience, I am asking you to excuse me after taking into consideration that my position with respect to this case is almost desperate: I am unable to publish my Answer elsewhere. However, your assistance can be restricted only to passing my Answer to the President of the Society[6] and to informing him of my desire. You will do me a great favour by fulfilling my request.

How are things going on at the University? I heard about your Commission, but it is impossible to obtain here any definite information about its purpose. […]

Andreev – Liapunov, 31.3.1901 [3, pp. 40 – 41]

Andreev will pass Liapunov's manuscript to B. K. Mlodzeevsky, the Secretary of the MMS, and will speak to Bugaev. He will not undertake to judge the debate between Liapunov and Nekrasov.

Nekrasov reasons perhaps deeply, but not clearly, and he expresses his thoughts still more obscurely. I am only surprised that he is so self-confident. In his situation, with the administrative burden weighing heavily upon him,[7] it is even impossible, as I imagine, to have enough time for calmly considering deep scientific problems, so that it would have been better not to study them at all.

Andreev is sceptical with regard to the university reform, and he congratulates Liapunov with his being elected Corresponding Member of the Imperial (Petersburg) Academy of Sciences and even with his "future fuller entry into the Academy".

Liapunov – Andreev, 8.4.1901

Highly respected and dear Konstantin Alekseevich,

I thank you for congratulating me with my election to corresponding membership of the Academy, and for your good wishes. As to my fuller entry into the Academy, at which you hint, this is not yet decided, and it is impossible to say how it will be decided. But, since we are discussing this subject, I ought to tell you that I was asked to stand, and gave my consent. But this will only be definitively decided by autumn. At present, it would please me if this business is not spoken about.

V. A. Steklov, who had just arrived from Moscow, visited us today. He told us many interesting things about your university life. It was very pleasant to find out that the report of our faculty committee is finally somewhat on the move and that it is now being used as an initial material by your committee.

I asked Steklov to visit you before leaving Moscow and to take the manuscript of my Answer if its publication was not considered possible. But Steklov informed me that you had already passed it to Mlodzeevsky and that he was unable to ascertain whether it will be published. I would therefore ask you, highly respected Konstantin Alekseevich, to inquire of Mlodseevsky, while meeting him, how was this business decided. If the article will not be published in the *Matematichesky Sbornik*, I would like to receive the manuscript back. In this case I shall publish it in the *Universitetskie Zapiski*[8] (about whose existence, as I ought to add to my shame, I had completely forgotten when sending you my manuscript).

If, however, it is decided to publish the article in the *Matematichesky Sbornik*, I would ask you to inform those responsible that I certainly desire to read the proofs […].

Please excuse me for all the troubles I am inflicting on you. I am very grateful for the assistance rendered me in this disagreeable business. […].

Andreev – Liapunov, 13.4.1901 [3, pp. 41 – 43]

Bugaev and Nekrasov do not want to publish Liapunov's manuscript. Nekrasov, however, agreed to its being put out, but not earlier than in a year, and with his objections added in a separate note. Andreev considers it expedient to extend the manuscript and soften its tone.

Liapunov – Andreev, 21.4.1901

Dear and highly respected Konstantin Alekseevich,

I am grateful to you for sending me the manuscript and for all the troubles encountered when taking it back. From your previous letter I concluded that it was hardly possible to count on its speedy return and therefore began to edit a new version of the Answer. And, in accord with your advice, I have essentially extended my article depicting in detail the entire factual aspect of the business without leaving a single objection of Nekrasov unanswered. And I think that because of this very circumstance my new Answer will cause Nekrasov considerably more annoyance. Perhaps he will even regret (tacitly of course) that he was not quick to publish the Answer in its old version.

Did you conclude work in the commission? If its results were reported to the University's Council, it would be interesting to know the attitude of that body. […]

Liapunov – Andreev, no date

Dear and highly respectable Konstantin Alekseevich,

It was extremely pleasant to hear from you. We came to know that the operation essentially benefited you and that at present your health is largely restored which greatly gladdened us. It would be nice to meet you.

However, […] yesterday I had informed A. N. Krylov over the telephone about your wish to have the book that he published,[9] and he answered me that it will be sent to you in a few days.

This autumn we moved into a kazenny apartment.[10] It is small but cosy and sufficiently spacious for the two of us. And it is very warm, which is indeed valuable in this severe winter. Its only, but really essential defect is that it is somewhat dark: it is on the ground floor and its windows, opening on the street, are directed towards the north-east. […]

I am now busying myself far less diligently than before. The occurring international events hold my attention to such an extent that I do not even wish to think about anything else. In addition, scientific pursuits demand a calm mood whereas the events occurring around us are very often so disgusting that they can only strongly irritate and embitter. In such cases scientific pursuits can only serve for distracting the thoughts and cannot be fruitful.[11] […]

Notes

1. These letters are kept at the Archive of the Academy of Sciences of the Soviet Union in Leningrad. Fond 257, Inventory 1, No. 29.

2. Fond 217, Inventory 1, No. 87.

2a. See Correspondence between P. A. Nekrasov and A. I. Chuprov [ix], Nekrasov's letter of 17 Febr. 1899, concerning the students' unrest at Moscow University and description of similar events in Kiev directly involving Slutsky [9]. See this letter in Russian in **S, G,** 16.

3. Liapunov bears in mind the periodical of the KhMSoc.

4. *The Soobshchenia* were printed in the Kharkov printing office M. Silberberg & Sons. Judging by Liapunov's Imprimatur inscriptions, the appearance of the issues of its vol. 7 had been irregular. The first issue even had two such signs, 30.11.1900 and 10.4.1902.

5. Liapunov was member of that Society from 1892 (*Matematichesky Sbornik*, vol. 16, 1891, p. 845).

6. The President of the MMS was N. V. Bugaev.

7. Nekrasov was then warden of the Moscow educational region and vice-president of the MMS.

8. More precisely, in the *Zapiski Khark. Univ*.

9. The sequel proves that the letter was written during World War I. The only book that he then "published", was his Russian translation of Newton's *Principia*. It appeared in 1915 – 1916.

10. In this particular instance, the proper translation seems to be: apartment, belonging to the Academy.

11. In 1916 Liapunov published two papers, both in the *Izvestia* of the Petrograd Academy of sciences. The same year he submitted one more paper, and it appeared in 1917, in the same periodical [10].

References

M, Psb, R = Moscow, Petersburg, in Russian

1. Liapunov, B. M. A brief essay on the life and work of A. M. Liapunov. *Izv. Akad. Nauk SSSR*, Otdel fiz.-mat. nauk, ser. 7, No. 1, 1930, pp. 1 – 24. (R)

2. Zesevich, V. P. *A. M. Liapunov*. M., 1970. (R)

3. Gordevsky, D. Z. *K. A. Andreev.* Kharkov, 1955. (R)

4. Liapunov, A. M. An answer to P. A. Nekrasov. *Zapiski Khark. Univ*., vol. 3, 1901, pp. 51 – 61. **S, G,** 4.

5. Nekrasov, P. S. Concerning a simplest theorem on probabilities of sums and means. *Matematichesky Sbornik*, vol. 22, No. 2, 1901, pp. 225 – 238. **S, G,** 4.

6. *Roy. Comm. on reforming Acad. Institutions.* Psb, 1903, No. 1. Publ. as a manuscript.

7. Youshkevich, P. [**S.**] On one scientific debate (1915). *Istoriko-Matematich. Issledovania*, vol. 34, 1993, pp. 207 – 209. (R)

8. Liapunov, A. M. Sur une proposition de la théorie des probabilités. *Bull.* [*Izvestia*] *Acad. Imp. Sci. St.-Pétersb*., t. 13, 1901, pp. 359 – 386. Short version : *C.r. Acad. Sci. Paris*, t. 132, 1901, pp. 126 – 128.

9. Chetverikov, N. S. Life and scientific work of Slutsky (1959). In author's book *Statisticheskie Issledovania* (Statistical Investigations). M., 1975, pp. 261 – 281. **S, G,** 10.

10. Lukomskaia, A. M., Liapunov. Bibliography. In Liapunov's *Izbrannye Trudy* (Sel. Works). L., 1948, pp. 495 – 538.

VIII

M. V. Chirikov, O. Sheynin

The Correspondence of Nekrasov and Andreev

Istoriko-Matematicheskie Issledovania, vol. 35, 1994, pp. 124 – 147

1. General information

1.1. Introduction

Pavel Alekseevich Nekrasov (1853 – 1924) [1, § 5; 2], Professor and Rector of Moscow University, then a prominent official at the Ministry of Public Education, was a distinguished mathematician and a religious person. A Platonist according to his philosophical views, he kept to reactionary political convictions. Konstantin Alekseevich Andreev (1848 – 1921) was a Corresponding Member of the Imperial (Petersburg) Academy of Sciences, a geometer and Professor at Kharkov and Moscow. The correspondence of Nekrasov and Andreev[1] was devoted to many subjects: the teaching of probability theory in high school; the encounters of both of them (but mostly of Nekrasov) with Markov;[2] the foundations of mathematical analysis; the central limit theorem (CLT). The main student of the contemporaneous history of that theorem is Seneta [4, §§ 6 and 7].[3]

For us, it suffices to say that in 1898 Nekrasov [5], having applied the methods of the theory of functions of a complex variable, sketched the proofs of the local and integral forms of the CLT for large deviations for sums of lattice random variables. His work made difficult reading and nobody appreciated it; the fate of his later writings proved to be just as dismal (cf. § 1.3).

In his letters to Andreev Nekrasov repeatedly asked him to ascertain the possibilities of discussing some problems, and of reporting at the Moscow Mathematical Society, and we emphasize that the former, although being an *elder* there (see Letter 11), never occupied any official position at the Society. On the other hand, it follows from the concluding salutations in the letters of both Nekrasov and Andreev that there existed ties between their families.

1.2. The teaching of probability theory in schools

Nekrasov's attempts [6] to introduce the theory of probability into high school are well known [2, § 2].[4] Like many other reformers, he had not thought about the difficulties of management which would have arisen had his proposals been implemented. Furthermore, for some reason he based them on P. S. Florov's programme compiled by that mathematician on a low theoretical level. A number of scientists beginning with Markov [8; 9] had therefore come out against Nekrasov's attempt and killed it.

Not feeling himself defeated, Nekrasov continued to explain his theoretically correct viewpoint in private letters. His main step, however, was an appeal to the Vice-President of the Academy of Sciences, the philologist P. V. Nikitin.[5] As a result, on Markov's initiative, the Academy established a commission which sharply denounced the Florov – Nekrasov proposal [10], but, at the same time, missed the opportunity to reform the Russian school programme. Still, Nekrasov stood his ground [3]. In particular, he (pp. 44 – 45) listed the commissions and congresses of the teachers of mathematics, which, over the years, had to do with the school mathematical curriculum and stressed that in 1914 the commercial schools had included elements of probability theory into their programme *in spite of the brakes created by Markov and his colleagues*. For the sake of comprehensiveness we list Nekrasov's writings at least partly devoted to the teaching of probability in schools [11 – 17; 6; 18; 3] and we quote his generalizing declaration [3, p. 51]:

At bottom, my official activities in defining the various types of schools and mathematical programmes […] *are reduced to an ideological struggle that aims at completely upholding the classical values of the mathematical education in all types of the general school*[6].

1.3. Nekrasov's writings. Some conclusions about them

From about 1900 Nekrasov's mathematical writings (only in probability and statistics) became unimaginably verbose, sometimes obscure and confusing, with mathematics being inseparably connected with ethical, political and religious considerations. Markov [19; 20] expressed himself against this manner of exposition and Youshkevich [21] offered a number of Nekrasov's phrase-mongering to illustrate his intolerable style. Much earlier Bortkiewicz, in a forgotten paper [22], accused Nekrasov of *oily words* (p. 215), *reactionary longings* (p. 216) and of attempts to justify *the principles of strong rule and autocracy* by the theory of probability (p. 219).

Nekrasov [13], however, expressly declared that a *consoliditating* [! edinyashchee] *basic education should be of a scientific-religious-national-state nature.* It is still possible to understand this statement but not his own writings compiled according to the same principle. Then, coming out for the introduction of logic into schools of all types, and considering [11, p. v] that school mathematics should be based on logic, he (p. iii) included into it elements of probability theory and the Jakob Bernoulli theorem. Mathematics, as he declared in addition on p. 9, accumulated

Psychological discipline as well as political and social arithmetic or the mathematical law of the political and social development of forces depending on mental and physiological principles!

This monstrous phrase apparently had to do with the works of Quetelet. Elsewhere, in connection with the statistical method, Nekrasov [7, p. 29] mentioned *problems of labour, and public wealth, of credit, life insurance and capacity to work*. Not restricting his efforts by upholding his own views, Nekrasov had been accusing Markov of pan-physicism[7] and of following Nietsche only because his opponent did not lump together mathematics with ethics, philosophy etc. [18]:

The mathematical language [must] […] *embrace supreme ethics*, [be] *together with conscience* (*with theology*) […]. *However, the mathematical language of such pan-physicists as Markov is of another* *kind, it is Nietzschean and does not recognize supreme ethics* (*theology*).

All the above except the support of Florov's unfit program (§ 1.2) is yet compatible with subjective scientific honesty. However, mathematical mistakes and unwarranted statements indicated by Markov [24; 20], Liapunov [25] and Posse [26], also see [2, § 4], impede even this conclusion.[8] The abovementioned peculiar features of Nekrasov's style prevented the recognition of his works, and, to the contrary, favoured his being considered only as a muddleheaded reactionary, We are unable to comment on a statement [27, p. 225] that he suffered from a mental illness, but it is impossible to deny Andreev's opinion formulated by him in a letter to Liapunov in 1901 [vii]: Nekrasov

Reasons perhaps deeply, but not clearly, and he expresses his thoughts still more obscurely. I am only surprised that he is so self-confident. In his situation, with the administrative burden weighing heavily upon him, it is even impossible, as I imagine, to have enough time for calmly considering deep scientific problems, so that it would have been better not to study them at all.[9]

Agreeing with Andreev, we believe that all of Nekrasov's philosophical and mathematical statements should be regarded as doubtful. At the same time, we provide illustrations of his deep thoughts which enable us to consider him as some mathematical Nostradamus. Thus,

1) His ideas about the dominance of logic (above) sounds really modern since it is possible to assume that he also bore in mind mathematical logic.

2) In connection with the mathematical study of indeterminacies Nekrasov [7, p. 23] mentioned almost all the main problems of the then not yet existing theory of catastrophes (and used the term *catastrophe*).

A special point concerns three Nekrasov's mistakes of constructing/spelling and his own coined words.[10] I attempted to preserve the former in translation so that the reader will see *consoliditating* (§ 1.3), *equivalenttion* (Letter 9) and *illiteracism* (Note 38). A mistake of another kind is his calling a pamphlet an article.

1.4. Markov's polemic style

Nekrasov repeatedly complained in his writings about the sharp tone of Markov's polemic statements and even about the rudeness of his private letters [3, pp. 56 – 62]. Andreev believed that scientific debates with Markov were simply impossible. Here is a passage from his letter of 13 April 1901 to Liapunov [28, p. 42][11]:

I have experienced on myself all the annoyance of debating with a man who does not like to restrict his sharp expressions at somebody else's expense. Markov all but scolded me.

Then, Slutsky (letter of 22 Nov. 1912 to Chuprov [68, p. 44]) tactfully remarked that Markov possessed an *unusual* manner of writing private letters, whereas Chuprov, in a letter to an English statistician Isserlis written late in 1925 or early in 1926 (Ibidem, p. 55), indicated that

Markov's temper was no better than Pearson's; he could not tolerate even slightest contradictions either.

However, Markov's very critical letter of 29 April 1913 to N. A. Morozov,[12] a former political prisoner, was polite and ended in a way unusual for him: *Please be assured of my perfect esteem and devotion*.

Acknowledgement. S. S. Demidov acquainted us with Nekrasov's letters to Florensky which are kept by the latter's family.

Notes to § 1

1. Archive of Moscow State University, Fond 217, inventory 1, No. 45 (Andreev's letters) and No. 92 (letters written by Nekrasov).

2. This second topic is also described in [2], which, however, was based on Markov's newspaper letters. Nekrasov himself [3, p. 52] attributed the beginning of his sharp scientific debates with Markov (not yet regarding probability) to the beginning of the 1890s. Seneta [4, p. 70] noted that in those days (in 1892) the relations between the two mathematicians were yet normal: Nekrasov even read out one of Markov's reports to the Moscow Mathematical Society (MMS).

3. Seneta also provided sufficient information about Nekrasov's life. In lesser detail he described Nekrasov's efforts to introduce probability theory into high school.

4. Very interesting are Nekrasov's more general thoughts [7, pp. 30 – 31] on the teaching of mathematics in school. They allow us to perceive his notion of *classical values* of the appropriate course. He recommended to include the theory of probability, elements of analytic geometry and analysis as well as the *consecutive approximate analysis* into the school curriculum. He related the last-mentioned subject to induction (understood in a wider sense) and he mentioned in this connection Laplace, Poincaré and other scholars (p. 19). He attached much importance to the establishment of mathematical classrooms and the educational use of films (pp. 30 – 31).

5. Nekrasov [3, pp. 55 and 58] subsequently published two letters to Nikitin written on 29 Sept. 1915 and 5 April 1916.

6. Cf. Nekrasov's statement from his letter to P. A. Florensky of 2 Nov. 1916:

For the sake of the future of our fatherland, it is necessary to raise the standard of mathematical education in the school but protect it from the Markov & Co's frame of mind by those precepts, emblems and exercises which are included in our native tongue, in Magnitsky's arithmetic, in Bugaev's arithmology, in the theory of probability of Buniakovsky, Chebyshev, Mendeleev and me.

The term *arithmology* introduced by Nekrasov' teacher, Bugaev, meant number theory but later became a synonym of a doctrine of discrete functions and even a Weltanschauung based on discreteness. It is now dated. Mendeleev did not have either any probability theory or even systematized indications on treating observations. True, Nekrasov [17, p. 4] declared that *the maps and the principles of nomography* in the great scholar's book *K poznaniu Rossii* (On Coming To Know Russia) *are adapted to the Chebyshev theorem* but this obscure remark did not explain anything at all. Magnitsky was the author of the first Russian treatise on arithmetic.

7. See Letter 12 and the appropriate commentary. Creation of new words by amateurs was then in vogue.

8. It is hardly known that Nekrasov [23, p. 11; 4, pp. 45 – 46] committed an elementary logical mistake when proving the convergence of an iterative process.

9. This statement was possibly prompted by Liapunov's lost commentary on a letter of 16 March 1901 from Nekrasov to him [29, p. 84]. Nekrasov advised Liapunov not to hurry with publications on probability theory and maintained that the latter's theorems *like those of Chebyshev* were corrupted by mistakes. It is quite possible that Liapunov could have imparted his thoughts about that letter to Andreev. Seneta [4, p. 63] indicated Nekrasov's unjustified declaration [30, § 168] about Liapunov's memoir [31].

10. See Note 7 above and Note 50 about the latter. The last invented word was *juridism* instead of *jurisdiction of sorts*.

11. Andreev apparently had in mind Markov's note [32]. Similar statements are in his Letter 3 (below), and, indirectly, in his Letter 7 (see Note 34). It is sufficient to indicate here that an editorial note attached to [33] explained that Markov had *declared that he considered it* […] *impossible* to replace some phrases in his letter [32] *by insertions without a sharp tone and not containing references to some personal relations*. Consequently, [32] was published with cuts.

12. Archive of the Russian Academy of Sciences, Fond 543, inventory 4, No. 1130. Markov denied Morozov's paper devoted to the application of the statistical method to linguistics and even publicly expressed his opinion [34].

2. The Correspondence between Nekrasov and Andreev

We adduce now these letters (with insignificant abridgements). In a number of cases we had to specify the references.

1. Nekrasov – Andreev, 14 Oct. 1915

Nekrasov had sent Andreev his *pamphlets* including an offprint of [35].

As in the past[13], so now also K. A. Posse [26] appears as Markov's advocate when the latter, in attempting to discredit his opponents, gets entangled in his own nets. This time Posse came out because Markov was painfully flogged in my article [17]. […] This encounter has a double lining. One of these is the natural struggle between schools differing in their principles;[14] but the other one, less visible, represents a tacit aspiration of a group of Petrograd mathematicians for subordinating [other] schools to their practical influence.

Thus, for example, the reviews written by the members of the scientific committee, Posse and Koialovich, and academician Markov, had killed, for all purposes, the talented works of P. S. Florov on analysis and probability theory. You probably know Florov who was a student at Kharkov University; he possesses a gift of explaining issues of higher mathematics in an elementary way.[15]

Leaving aside the second lining since it touches the academic-managerial system, I wish to seek your advice about the first one that concerns the principles of mathematics of main importance for science and education. I am deeply convinced that the comparison of the Brashman – Davidov – Bredikhin – Imshenetsky – Bugaev – Tsinger school[16] with that of Posse – Markov reveals the greater value of the former's principles. At the same time, however, the latter is militant, and, by means of Markov & Co.'s 16-inches' debate [cannonade], it is attempting to overthrow the best principles so as to replace them by their own ones. I alone have to withstand the charge of an entire bloc.

Perhaps the Mathematical Society can objectively (without any personal debate) compile definite decisions on the points of disagreement on principle as they are revealed in a number of my encounters with Markov & Co. Issues to be decided could be formulated about classifying the concepts; on relations of analysis and arithmetic with mechanics and probability theory; and about the preference of some methods to other ones (e.g., about comparing the Bienaymé – Chebyshev – Markov method with the method of Cauchy – Chebyshev – Nekrasov – Pearson[17] […].

2. Nekrasov – Andreev, 15 Oct. 1915

… I have prepared a brief report which I will entitle thus:

The concepts <u>limit</u> and <u>asymptotic equivalent of a function</u> in the calculus of probabilities of sums and means when integer m [apparently: the argument of the function] *increases unboundedly* *applied in the calculus of the convergence of series and the calculus of probabilities of sums and means*[18].

There will be no personal debates in this talk, but, nevertheless, the main concepts will be discussed so exhaustively as to overturn completely the Markov & Co.'s declaration [24; 8; 26] that I, rather than they (Posse and Markov), am abusing the concepts limit and infinitesimal.

The end of the letter is lost.

3. Andreev – Nekrasov, 24 Oct. 1915

Andreev received the manuscript of Nekrasov's report. It had not indeed *contained any polemical sharp words* but it can lead to new discussion. The MMS resolved to transfer the decision about publishing the manuscript in the *Matematichesky Sbornik* to its Bureau.

Naturally, I have absolutely abstained from any personal testimonial about the debate. Assuming that you are interested in my opinion about the differences between you, and Posse & Markov, I venture to formulate it. […] I believe that it is completely out of question to decide who is right in your debate. Its essence is not to determine the correctness or otherwise of some judgements based on rigorously established assumptions, but to establish these very assumptions. Even if this does not belong to metaphysics, it at least lies in the province of intuition in the broadest sense of this word.

Here, along with the mind, […] appear, with a certain degree of being in the right, […] tastes, inclinations, habits, acquired outlooks, sometimes even random points of views, […] about which […] *non est disputandum* […].

My life experience showed me that the champions of either direction sin, each of them in their own way. Some, while attempting to build firmly on a not yet prepared and still quaking soil, are compelled […] to make use of diffuse and tempting explanations, in verbose formulations, etc. That is your sin.

Others, being unable to justify the [mistaken] proposition that progress in science is conditioned by barring the expansion of the mental outlook, resort to sophistry and cannot resist the temptation of stinging their opponents not by scientific, but by journalistic weapons. That is Posse's sin.[19] […] Not being a sinner in either of these senses, [Markov] […] to this day remains 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. […][20]

4. Nekrasov – Andreev, 25 Oct. 1915

Having received no answer to his Letters 1 and 2, and understanding that his requests were difficult to fulfil, Nekrasov is prepared to abandon them, but he asks Andreev to return him the *suggested report* to the MMS entitled *Criticism of the connection and difference between the concepts limit and equivalent of a function of an unboundedly increasing number N.* Nekrasov then supplements the text of his report by the following considerations.

…the concept *equivalent* of a function, that I am widely using in the calculus of discrete functions φ(0) of a discrete and very large number *N*, has also been applied for a long time in another section of mathematics, namely, in the analysis of infinitesimals having to do with continuous magnitudes. Imshenetsky made use of that concept in 1873, in his *Supplements* to [39][21] and thus essentially added to the Lagrange's and Todhunter's concept known as the *theory of analytic functions*, and, at the same time, he paved the way for bringing together the theories of continuous analytic and of discrete (arithmological) functions.

Boussinesq [40] does not apply the term *equivalent*, but, like Imshenetsky, he (pp. 64 – 66) established principles equivalent [tantamount] to this concept and indicates their importance for simplification.[22] These simplifications, as Imshenetsky put it, lead to *imperfect equations*, and, for continuous functions, they lead, in the limiting case, to *perfect* equations between differential coefficients. […]

5. Nekrasov – Andreev, 30 Oct. 1915

Nekrasov thanks Andreev for fulfilling his requests formulated in Letters 1 and 2.[23] He agrees beforehand with any future decision made by the MMS about his debate with his main opponent, Markov, but he asks that his work mentioned in Letter 4 be additionally considered. That letter will *reveal more perfectly the criminal* [?] *sense of what was said by me* [Nekrasov] and Markov [41, p. 459; 30, vol. 22, p. 326; 24, pp. 223 – 224].

The fundamental issues of the calculus of equivalents and limits are closely linked with the main tendency of mathematicians to simplify formulas and to admit, for the sake of simplification and saving of time, […] a *rightful dose of subjectivism* (Boussinesq) and of *active intuition* (of experience, of the ars conjectandi [art of conjecture] in the spirit of Jakob Bernoulli) so as to approach the *objective truth*,[24] – to admit it to an extent which will not lead to overstepping *the right* to make slight errors. Mathematics of approximate and asymptotic calculations has an exact *juridism* with rules, customs, instructions of/in computation which must be categorically observed. The question is, who of us, Markov or I, oversteps this extent in the differential calculus of probability ΔP?[25] […]

My prosecutor in the person of Markov had not, however, calmed down and continues to charge me definitely with introducing fundamental mistakes into the theory of limits, into the doctrine of infinities and of infinitely low probabilities.

Nekrasov listed a number of papers published by Markov, Posse and by himself [24; 9; 26; 17; 35] and continued:

… from 1898 onward, while rendering proper homage to Chebyshev, I am, however, publicly maintaining that his *theorem* on probabilities [37] is rather a *postulate*[26] which demands critical attitude, and that the fundamental faults in the calculus of probability are to be found *not in my work, but in Markov's writings*. Indeed, he, even after the publication of my memoir [5], *persisted in claiming that Chebyshev's postulate is a theorem* [46; 48 – 50; 51, 1900, pp. 88 – 89]. Later, evidently under my influence, Markov [51, 1913, pp. 88 – 97] changed the theorem; however, he introduced a *lacunary* (*a molar*) *reckoning when measuring a varianta*[28]

$$X = \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_m.$$

But he passed over in silence both that lacunarity and the fact that filling it in will demand the recognition of all the power of the fundamental base of my and Pearson's critical attitude towards the known differential calculus of the probability of the value of X.[29] […]

I am quite sharing your opinion that it would have been best to refuse to respond to Markov's provocative attacks so as not to swallow his bait. And I had indeed ignored them until the attacks were perpetrated in some newspaper (the provocation in the newspaper *Den* [20] became Markov's usual business).

However, Nekrasov cannot keep silent when Markov publishes such attacks in the periodical of the Ministry of Public Education [9] or of the Academy of Sciences.[30] The letter in the newspaper [20] impertinently slanders [Nekrasov] as though I attempt to direct the teaching of mathematics in schools on a wrong track[31] […] Actually, my project, and even not my personally, but the collective project compiled by the professors and teachers of the Moscow educational region and the Petersburg Ministerial Commission in 1899 – 1900 [was] later [discussed] at the All-Russian congresses of teachers of mathematics.[32] […]

I would like to ask you to consider personally and carefully the essence of the debate and to take a firm stand with confidence whether I am wrong, or is Markov wrong.[…]

6. Nekrasov – Andreev, 15 Nov. 1915

The debate is going on not about some depths of metaphysics, but only about the calculus of actual infinitesimal probabilities. The abuse of mathematics (petition principii [begging the question]) is not mine, but on Markov's side since he scolds [denies] my right to apply the known simplifying principle of replacing actual infinitesimals by their equivalents […][33]

7. Andreev – Nekrasov, 17 Nov. 1915

The MSS's Bureau indicated that Nekrasov's suggested report concerns the subject of the debate which the Society had earlier (*Matematich. Sb*., vol. 28, 1912, p. 351) resolved to restrict by publishing one paper of each of the debaters. The MMS will apparently approve the Bureau's opinion.

You could have plainly understood [Letter 3] that I consider any investigation of this debate […] as at least a useless business. […] In essence, all that you wish is that at least one such person will be found who explicates your own ideas more clearly […]. I may assure you […] that you are severely mistaken if you think that my firm stand taken with confidence can lead to the resolution of the debate. Not confidence is here needed but clarity. However, an explanation of somebody else's ideas is a thankless and very risky business. […] Only once in my life I had allowed myself to explicate somebody else's ideas [54], the ideas of the late Imshenetsky [55], but I have since felt sorry for myself because I saw that I was occupied with a needless business which did not benefit anyone or anything.[34]

[…] I certainly cannot approve of the academician's [Markov's] appeal to the public opinion formulated as a newspaper feuilleton [20]. However, once he brought himself to come out on that arena, he had the right to use all the means usually applied there, and, in particular, to parade in an ungainly fashion the weak points and the blunders of his opponent. Markov [20] makes use of this weapon very skilfully and deftly […].

Andreev then comments on Nekrasov's unfortunate, to say the least, statement quoted by Markov:

[…] *It is utterly unthinkable that you were seriously convinced that the publication of an article in the transactions of a scientific society may serve as some crucible* […] *or that it indicates the article's approval by the society.* […]

All the previous lines were evoked by a feeling of my sincere liking and goodwill towards you.

8. Nekrasov – Andreev, 5 Dec. 1915[35]

Upon considering Nekrasov's letter to its Vice-President, the Academy of Sciences *resolved to constitute a commission* for looking into the teaching of the theory of probability in the school. The mathematical section of the Pedagogical Museum of the military schools discussed the reports of A. N. Krylov, S. A. Bogomolov, Ya. V. Uspensky and Nekrasov himself.[36]

At the […] Museum I delivered a talk on the more elementary part of my report *Criticism* etc. [see Letters 2 and 4] and ascertained the educational and simplifying significance of the principle of the equivalence of magnitudes.

9. Nekrasov – Andreev, 13 Dec. 1915

[…] my memoir touched on the central issue of the fundamentals of mathematics, namely, on the equivalence of functions. It shows the normal way of induction from the simpler to the more complex.[37]

Nekrasov then mentioned Imshenetsky and Boussinesq (see Letter 4), referred to Barbèra [57] and continued:

My own works on the calculus of functions [of very large numbers], on approximate and asymptotic laws of equivalenttion [!] of functions are completed along the entire line and in all rigor so that I am firm in my conviction against all the insinuations of Markov and Posse […][38].

Nekrasov mentioned the *Zhurnal Ministerstva Narodnogo Prosveshchenia,* cited Markov [24, pp. 223 – 224] and informed Andreev that, in connection with his report at the Pedagogical Museum (see Letter 8) and with the appearance of a Russian translation of Newton's *Principia* [58], he would like to supplement his suggested report at the MMS, and, in addition, to refer there to Barbèra [57].

10. Nekrasov – Andreev, 12 Jan. 1916

Nekrasov once more (see Letter 9) makes known his desire to improve his report indicating the same reasons as before but this time without mentioning Barbèra [57].

11. Nekrasov – Andreev, 5 Febr. 1916

The Commission of the Academy of Sciences took advantage of the school issue only for settling the score with me and for compiling a new pamphlet against me [10] […]

Nekrasov asks permission to answer the Commission via the *Matematich. Sb.*

And I am once more asking you as an elder among the representatives of pure mathematics and pedagogy[39] to support me with all resoluteness […]. The report of the Commission [10, p. 79] includes the main distortion of the basis of my scientific and philosophical concepts.[40] […] I never confuse philosophy […] with pure mathematics[41] […] The booklet [59] contains ideas identical in spirit with mine.[42]

12. Nekrasov – Andreev, 7 March 1916

The MSS had not allowed Nekrasov to publish his answer to the Academic Commission in their periodical but he thanks Andreev for his troubles.

You have correctly indicated that the struggle is going on on three fronts. (The first one: the fundamentals of the analysis of infinitesimals and of their approximate asymptotic calculus; the second one: fundamentals of the theory of probability; and the third front: the bringing of mathematics and probability theory into proper relation with issues in physics, religion and politics.[43]) You think that for me the struggle on the third front is hopeless. However, it is here that the Commission has distorted my works most of all.[44] […] Markov [51, 1908] treats problems in religion and politics (ignorantly, by means rejected not only by Buniakovsky [60, p. 326],[45] but also by Boole, Jevons, Bertrand and Karl Pearson[46]), and this is allowed and even commendable […].

If, however, Nekrasov [43] discusses the same issues, it becomes

An inadmissible abuse of mathematics. […] This is a purely Prussian objectivity of reasoning. [47] […] In spite of the Commission's statements I distinguish two main directions struggling with each other. The motto of the school belonging to one of these is mathematical humanism; the motto of the school of the other direction (trend) is physical-mathematical realism[48]. If these trends can be united into a single one, it might be done only under the first motto, only when [additionally] stating that the ideal of science is the lamp of the truth (V. Ya. Tsinger). Pearson, a mathematician and a humanist, considers the separation of science and philosophy as obscurantism [61, p. 55 of the Russian translation]. […] The founders of the MMS were of the same opinion [62; 63, 64].[49]

Nekrasov then criticizes the Russian translation of Newton's *Principia* [58][50] and declares that the theory of probability is the basis of *a wide mathematical induction in the sphere of disputable but vital issues* (*Poincaré, Pearson, N. A. Umov*) [...][51].

Regrettably, in 1872 the school in Russia took the road to ruthless pseudo-classicism and formalism and reshaped the minds of contemporaries in a different fashion;[52] physical mathematics was substituted for mathematics, humanism became thought of as being opposite of mathematics rather than of the extremes of materialism and heartless formalism.[53]

The mathematical societies in Moscow and Kharkov came into being for struggling against such obscurantism.

Had the Mosc. Mathematical Society wished to defend the humanitarian branch of mathematics with its spiritual culture, it would have very, very strongly supported my just claim to correct the distortions committed by Markov & Co.

Nekrasov begins thinking about leaving the Society. He asks Andreev to inform the MMS about his intention *as about a tentative decision* and to ask them whether they did not become *only physico-mathematical rather than widely mathematical* […]. The formal cause for his leaving will be the Society's refusal to enable him to defend himself from the Academic Commission. The end of the letter is lost.

13. Nekrasov – Andreev, 13 March 1916

After receiving Andreev's (lost) reply to Letter 12, Nekrasov asks him *to do nothing* concerning his intended leaving of the MMS.

Notes to § 2

13. Nekrasov apparently thought about letter [36], see Note 34.

14. The text below makes it clear that Nekrasov contrasts Moscow scientists with those in Petrograd (Petersburg). The lumping together of such mathematicians as Davidov and Bugaev and the astronomer Bredikhin (below) allows us to question whether such a school existed at all, cf. Notes 16 and 17.

15. Posse and B. M. Koialovich were members of the scientific committee of the Ministry of Public Education. Concerning the latter see *Novy Enz. Slovar* (New Enc. Dict.), vol. 23. Petrograd [1915], p. 46. In 1883 – 1888 Florov published not less than eight papers on mathematical analysis (mostly, on differential equations) in the *Soobshchenia Math. Soc. Kharkov Univ*. In 1912 – 1915 he also wrote three superficial articles on the Jakob Bernoulli theorem; on the Buffon needle; and on insurance of life, all of them in *Vestnik Opytn. Fiziki i Elementarn. Mat*., and he reported on annuities at the Second All-Russian Congress of Teachers of Mathematics (1913 – 1914). Nekrasov [17, p. 4] also publicly argued that Florov *possesses a wise gift for explicating great truths in the simplest form* …

We are unaware either of any reviews of Florov's papers on analysis or of Posse's comments on his work in probability. Koialovich [6, No. 3, pp. 18 – 19] regarded Florov's programme as scientifically unsatisfactory.

We also note that Nekrasov called mathematical statistics a shaky and poorly substantiated theory.

In 1910, Nekrasov asked the Minister of Public Education to be appointed unpaid member of the Ministry's Scientific Committee. His request was refused because of the resistance of those mathematicians who already were members of this Committee. They were afraid that Nekrasov's appointment *can lead to very undesirable conflicts* […] *concerning the existing mathematical curricula*, see letter of N. Ya. Sonin to the Minister of 8 May 1910, Ross. Gos. Istorich. Arkhiv, Fond 740, inventory 43, No. 24, p. 2. The entire letter in Russian is in **S, G. 4**. Dr. A. L. Dmitriev (Petersburg), from whom I received a copy of this letter, informed me that it is (mistakenly!) kept among materials concerning politically suspect academics of Tsarist Russia.

16. Nekrasov undoubtedly meant V. Ya. Tsinger, also see his Letter 12. All the scientists whom he named excepting Imshenetsky were founders of the MMS. Also see Note 17.

17. It is difficult to agree with the existence of some single Cauchy – … – Pearson methods. True, Nekrasov applied methods of the theory of functions of a complex variable in probability (§ 1.1) which explains his mentioning Cauchy (but not Chebyshev or Pearson). In essence, Nekrasov repeated his then already published statement [17, pp. 10 – 11] about two directions in mathematics (where he had indeed contrasted *the ideas of Cauchy, and, on the other hand, of Bienaymé*).

The Academic Commission [10, pp. 67 – 73] established to discuss the teaching of mathematics in school, necessarily overstepped its terms of reference and sharply denounced both this statement and Nekrasov's wrong understanding of the principles of mathematics (and his attacks on Chebyshev's memoir [37] based on his mistakes). Also see Letter 5.

18. It can be thought that in Letter 4 Nekrasov described the same suggested report but entitled it differently. Later he published its *more elementary part*, cf. Letter 8. In his works that appeared from 1885 onward Nekrasov understood the term *asymptotic equivalence* of functions $f(x)$ and $g(x)$ with a continuous or discrete argument x in several senses, namely (in modern notation) $f(x) \sim g(x)$; $f(x) = O[g(x)]$; and $f(x) = o[g(x)]$. This ambiguity allowed him to formulate some prophetic statements about the possibility of considering probability as an actual infinitesimal in the spirit of non-standard analysis [38, p. 110]. See Letter 5.

19. It seems that Andreev had correctly noticed some features of the debate between Nekrasov and his opponents but had not wished to consider carefully its mathematical essence. Posse's paper [26] was nevertheless a scientific writing devoted, in particular, to the theory of limits. Finally, Markov, who introduced a new and extremely important object, dependent random variables, into probability, was not interested either in the methods of the theory of functions of a complex variable, or in axiomatizing probability. And for a long time he was regarding the fist steps of mathematical statistics with excessive suspicion.

20. Cf. § 1.4 and Note 34.

21. Imshenetsky supplemented his translation of Todhunter [39] by considering the application of analysis to three-dimensional geometry and by a chapter on infinitesimals (definition; order of magnitude; equivalence) and on differentials. Todhunter's book was an educational treatise and Imshenetsky's chapter was naturally of a methodological rather than scientific nature. Imshenetsky (p. 450) had indeed, see below, introduced imperfect equations of the type $\varphi(\alpha, \beta, 0) = 0$ which replaced equations of the type $\varphi(\alpha, \beta, \gamma) = 0$ when all the variables were infinitesimals with γ being of a higher order of magnitude than α and β whose orders coincided. The attribution of the theory of analytical functions to Lagrange and Todhunter remains on Nekrasov's conscience.

22. On the indicated pages Boussinesq established only one principle by stating that an infinitesimal may be replaced by any other one if their ratio tended to unity. True, he additionally formulated an evident corollary.

23. The extant part of the Letter 2 contains no requests.

24. On the indicated pages Boussinesq [40, pp. 64 – 66] had not mentioned subjectivism. Nekrasov's reference to Jakob Bernoulli is hardly convincing. We are inclined to understand subjectivism in approximate calculations as a replacement of a given function $f(x)$ by a simpler function $\varphi(x)$ that asymptotically estimates $f(x)$; $\varphi(x)$ is chosen subjectively so as to facilitate calculations. We shall also quote Ashby [42]:

The theory of systems should be based on methods of simplification, and, in essence, it represents a science of simplification. […] *I think that the science of simplification was initiated by mathematicians who study homomorphisms.*

25. Nekrasov [43, 1912, p. xiv] also excused, although indirectly, *a relatively infinitesimal error against formal logic*. Markov [9, p. 28] declared that that statement *has nothing to do with common sense*. See Note 29 about the differential calculus of probability.

26. According to the context, Nekrasov meant the central limit theorem (CLT). In any case, in 1901 he [30, vol. 22] groundlessly accused Chebyshev (and, for good measure, Markov and Liapunov as well) in that they had mistakenly understood the foundations of mathematical analysis [2, § 4]. In 1898 Nekrasov [5] had not at all mentioned the Chebyshev memoir [37], and later [44, p. 24] explained that omission by his aspiration for brevity and by the fact that he had applied a more perfect method than the one used by Chebyshev. In 1900 Nekrasov publicly declared that Chebyshev's proof of the CLT was of little value, see my paper on his work in probability (*Arch. Hist. Ex. Sci.*, vol. 57, 2003, pp. 337 – 353 (p. 348).

In 1898 Nekrasov was still able to write concisely, but his arguments were hardly convincing. In 1915 he devoted an article [45] making difficult reading to the memoir [37], also see Note 27, and in 1916 he [7, p. 26] repeated that Chebyshev's statement (the CLT) *is not a theorem in the strict sense but a postulate correct until finite magnitudes of probability are discussed, but having numerous exceptions otherwise*. Nekrasov connected the cases in which the *postulate* failed with an actual infinitesimal probability and referred to his *rigorous proof* and to Pearson's *investigations* but had not mentioned any sources. His reference to Pearson is unconvincing, and it was the limit of the sum of variances of the studied random variables divided by their number that should not have vanished [47, p. 240; 1, § 7.2]. Note that Nekrasov provided his own definition of a postulate [3, p. 54]; quite consistently, he called it a rule spoiled by exceptions.

27. Nekrasov, here [45, p. 318] and elsewhere [17, p. 12], identified a lacuna with a mosaic pattern of landownership and for some reason called the normal distribution lacunary. Earlier he [43, 1912, pp. 141 and 473] understood lacunarity as a defect (lacunarity of a law) or brevity (lacunarity of a table). Also in [45, p. 321] he explained that he had constructed the term molar from the Latin *molares* (a stone block).

28. A variable taking discrete values. The Russian term *varianta* is hardly used anymore although several decades ago Fichtenholz [52, § 22] had applied it and referred to H. C. R. Méray (1835 – 1911) without mentioning any source. The German text of Fichtenholz used the German term *Variante*.

29. It is hardly possible to comment duly on this statement before studying in detail Nekrasov's merits in proving the CLT (which Seneta, see § 1.1, described somewhat cursorily). Note, however, that Nekrasov [17, p. 3] mentioned the *Nekrasov – Pearson differential form* which had something to do with the theory of the Pearson curves. Neither that source, nor [43, 1912, pp. 519 – 520], to which he had there referred, gives any clue to the understanding what exactly had Nekrasov contributed to this form and what was the essence of Pearson's critical attitude towards the *differential calculus of probability*.

30. Nekrasov's reference to the Academy's *Zapiski* is patently wrong. He could have mentioned its *Izvestia* where Markov [53] negatively, although not altogether directly, commented on all of his works in general.

Nekrasov had not (and could not have) answered such a general comment.

31. Markov [20] discussed only the fundamentals of mathematical analysis and his statement about the wrong track remained unjustified.

32. Nekrasov's reference to professors and teachers and to a commission was a fabrication, pure and simple. The discussion of his reports [12; 15; 16] at the congresses took place only partly; and, in addition, it had not at all amounted to their general approval.

33. Nekrasov [17, pp. 13 and 16] also publicly accused Markov of begging the question (again without explanation). Not later than in 1935 the concept of convergence in probability, that had been applied long before that, was fully understood. In probability theory based on classical analysis debates about actual infinitesimals became since then pointless.

34. Andreev had published Imshenetsky's posthumous manuscript [55] whereas Markov [31] sharply criticized it. Then, the former [54], while recognizing the paper's incompleteness, reasonably argued that its appearance was nevertheless useful. Markov [34] however declared that his viewpoint had triumphed since the incompleteness of [55] was not denied.

Markov's opinion was apparently too formal. Here, incidentally, is Bezikovich's testimony [56, p. XIV] about Markov: His last article was the only one

Which lacks a complete solution of the formulated problem […] *He brought himself to publish it only having been afraid that he will be unable to complete it*.

Finally, we disagree with Andreev in that his paper [54] *had not benefited anyone or anything*. Imshenetsky's article [55] continued his previous work of 1887 – 1888 which also provoked debate where Markov, Nekrasov, and Posse et al had participated [36].

35. See § 1.2.

36. The Pedagogical Museum published all these reports in 1916 as separate booklets. Neither of them concerned probability theory, but Nekrasov's contribution [7] was an exception. Also see Note 18.

37. The replacement of an infinitesimal by an equivalent magnitude, as Nekrasov himself indicated (Letter 5), is made for the sake of simplification. His idea becomes clear when recalling his pronouncement about induction, see Note 4.

38. Nekrasov apparently meant the papers by Markov [9] and Posse [26].

39. A list of members of the MMS indicating the year of their entry had been published regularly in the *Matematichesky Sbornik*. Already in 1913 (vol. 29, No. 1 of the periodical) from about 90 members from Russia itself not more than five had joined the Society before Andreev (before 1873) did.

40. Nekrasov perceived the *main distortion* in that he allegedly attempted to prove mathematically the omnipotence of God whereas he stated that without faith mathematics was insufficient for that purpose. Nevertheless, the Commission, in the place indicated, only referred to the bygone (and generally known) attempts to prove God's omnipotence by applying various arbitrary rules for summing divergent number series.

41. In a letter of 13 Dec. 1916 to P. A. Florensky Nekrasov argued that he conciliated mathematics with religion and politics *logically, correctly and rightfully*. Both his statements were wrong, see for example the Introduction to his treatise [43, 1912].

42. Khvolson [59, p. 11] held that *we must, and we can almost only believe* in physical laws. Thus, for small values of *a* the law of universal gravitation cannot be empirically isolated from the family of formulas including (in usual notation) $F = kmM/r^{2+a}$. Then, he stated that hypotheses which indeed included *the veritable essence of physics as a science* (p. 13) were only based on faith (p. 12). Finally (p. 15), Khvolson declared that a number of issues including the problem of free lay beyond the province of knowledge.

43. Nekrasov publicly repeated these definitions of the *fronts* [3, p. 21]. *Mathematics and probability theory* apparently meant *mathematics including* …

44. Cf. Note 40.

45. In 1916 Nekrasov [3, p. 54] mentioned *the main mathematical doctrine* […] *of Buniakovsky and others about the trustworthiness of ancient legends* … Actually, Buniakovsky [60, p. 326] simply stated that *the spiritual world includes such facts which do not obey physical laws*. In 1908 Markov [51, 1924, p. 320] resolutely objected to this, and added later [9, p. 33] that the chapter of probability theory dealing with appraisal of testimonies was its *weakest*. We shall say a few words about Markov's unsuccessful petition, in 1912, for excommunication from the Russian Orthodox Church (the Most Holy Synod resolved that he *fell away* from the Church). A few days before his death in 1910, the Synod resolved that Tolstoy is to remain excommunicated. I think that Markov was prompted by the notorious blood libel case against an ordinary Jew, Beilis.

46. Neither Buniakovsky, nor the other scientists (Boole, Jevons, Bertrand) could have rejected an objection only formulated in 1908. And, anyway, we think that denying atheism is as impossible as denying religious faith. Note also that Nekrasov, in the same letter, mentioned Pearson twice more. Did not he realize that Pearson, together with like-minded associates, had created the Biometric school in order to study mathematically the issues of the atheistic theory (more precisely, hypothesis) of the evolution of species?

47. Recall that Nekrasov wrote this in 1916. A much more ugly statement is in his letter to Florensky of 11 Nov. 1915 (also while the war with Germany was going on):

I quite sympathize with your attempt to teach the mathematical encyclopaedia at the Theological Academy. At your hands, it will differ from an encyclopaedia of Markov & Co. inspired from Berlin.

The letter bears the date 1905, but Nekrasov was obviously mistaken: he also referred to a source published in 1914. In this context, *encyclopaedia* meant a reference book on mathematical analysis.

48. It is impossible to understand why humanism cannot be realistic. Nekrasov, however, apparently thought about Christian principles rather than humanism.

49. Here are the relevant statements. Pearson (translated from Russian): *To distinguish between the fields of philosophy and science means promoting obscurantism*. Tsinger [63, p. 39]: *Mathematical sciences ... are very closely related to philosophy*. The speeches of Davidov [64] and Bugaev [62] lack such pronouncements, but, judging by their contexts, these scientists would have hardly objected to Pearson's opinion.

50. A similar criticism is contained in Nekrasov's letter of 7 Dec. 1916 to Florensky:

Krylov […] *is elected to full membership at the Academy in spite of his scientific illiteracism* [!]. *He translated Newton's book ignorantly* […] *and supplied his biased translation with notes in the spirit of panphysicism, i.e., in the same spirit in which academician Markov distorted the principles of the classical work of academician Buniakovsky* [cf. Note 45] *by his pseudo-interpretation*. […] *Those who collated the authentic Newton's text with its translation also mention many philological mistakes in the latter which had distorted beyond recognition the ideas of the great scholar who believed in God and His prophets*.

Three authors (N. N. Luzin, p. 54; T. P. Kravets, pp. 322 – 323; and T. I. Rainov, p. 343) of the collected articles [65] expressed an opposite opinion. True, we had not found there any detailed discussion of Krylov's work but it is quite possible that Nekrasov had exaggerated. And Krylov [66] expressly stated that he had not kept to Newton's mathematical terminology.

Nekrasov had not explained the terms *panphysicism; physico-mathematical realism; physico-mathematics* (see the same Letter 12). However, since he connected panphysicism not only with Krylov, but also with Markov, who never studied physical problems, it might be assumed that he understood that term as the explanation of nature by mathematical means without turning to God. Indeed, in another letter to Florensky (1 Aug. 1916) Nekrasov argued that *the Moscow school directs the training of teachers in the spirit of panphysicism with an anti-Christian tinge* … One question suggests itself: Was Laplace an panphysicist? Apparently, yes.

51. Not the theory of probability but mathematical statistics is (partly) based on induction. See however Note 4. Umov was a physicist hardly connected with probability.

52. A strict bureaucratic surveillance of schools was implemented in 1866; in 1872 a *Statute concerning city schools* was adopted to weaken the influence of social institutions on education. A *Statute concerning primary public schools* was then introduced in 1874 for guarding the school against *pernicious and ruinous influences* [67, p. 759 – 760]. These facts do not corroborate Nekrasov's statement which apparently reflected the weakening of the influence of faith on natural sciences.

53. We think (cf. Note 48) that, according to Nekrasov, humanism meant Christian or mystic principles. This corroborates our opinion (§ 1.1) that he kept to the Platonic tradition.

References

IMI = *Istoriko-Matematich. Issledovania*
Kazan Izv. = *Izvestia Fiz.-Mat. Obshchestvo Kazan Univ*., 2nd ser.
MSb = *Matematich. Sbornik*
Soobshchenia = *Soobshchenia Kharkov Matematich. Obshchestvo*
ZMNP = *Zhurnal Ministerstva Narodnogo Prosveshchenia*
M, L, R, Psb = Moscow, Leningrad, in Russian, Petersburg

1. Sheynin, O. B. (1989), Markov's work on probability. *Arch. Hist. Ex. Sci.*, vol. 39, pp. 337 – 377.

2. Sheynin, O. B. (1993), Markov's letters in the newspaper *Den*. IMI, vol. 14, pp. 194 – 209. **S, G,** 85.

3. Nekrasov, P. A. (1916), *Srednya Shkola, Matematika i Nauchnaya Podgotovka Uchitelei* (High School, Mathematics and Scientific training of Teachers). Psb.

4. Seneta, E. (1984), The central limit theorem and linear least squares in pre-revolutionary Russia. *Math. scientist*, vol. 9, pp. 37 – 77.

5. Nekrasov, P. A. (1898), The general properties of mass independent phenomena. MSb, vol. 20, pp. 431 – 442. **S, G,** 4.

6. Nekrasov, P. A. (1915), Theory of probability and mathematics in high school. ZMNP, 4th paging, No. 2, pp. 65 – 127; No. 3, pp. 1 – 43; No. 4, pp. 94 – 125. (R)

7. Nekrasov, P. A. (1916), *Prinzip ekvivalentnosti velichin* etc.
(Principle of Equivalence of Magnitudes in the Theory of Limits and in Consecutive Approximate Calculus). Petrograd.

8. Markov, A. A. A question for the Ministry of Public Education. In [2].

9. Markov, A. A. (1915), On the Florov and Nekrasov project. ZMNP, 4th paging, No. 5, pp. 26 – 34. (R) **S, G,** 95.

10. Markov, A. A. et al (1916), Report of the Commission for discussing some issues concerning the teaching of mathematics in high school. *Izvestia Akad. Nauk*, vol. 10, No. 2, pp. 66 – 80. **S, G,** 4.

11. Nekrasov, P. A. (1906), *Osnovy Obshchestvennykh i Eststvennykh Nauk v Sredney Shkole* (Principles of Social and Natural Sciences in High School). Psb.

12. Nekrasov, P. A. (1912), On sections of mathematics necessary for economic sciences. *Matematich. Obrasovanie*, No. 2, pp. 79 – 81. (R)

13. Nekrasov, P. A. (1912), The lyceum system for connecting the education in high school and in universities as a measure for regulating our schools. Newspaper *St. Peterburgsk. Vedomosti*, 17 (30) Oct., p. 2. (R)

14. Nekrasov, P. A. (1913), The intermediate lyceum stage between high school and university. ZMNP, 4th paging, No. 11, pp. 31 – 48. (R)

15. Nekrasov, P. A. (1915), On the educational features of the two directions of the mathematical course in high school. *Doklady, Vtoroy Vseross. S'ezd Prepodavatelei Matematiki* (Reports 2nd All-Russian Congr. Teachers Math.). M., pp. 83 – 93. (R)

16. Nekrasov, P. A. (1915), The second (bachelor) stage in the future high school. Ibidem, pp. 175 – 181. (R)

17. Nekrasov, P. A. (1915), On Markov's paper [9]. ZMNP, 4th paging, No. 7, pp. 1 – 17. (R) **S, G,** 95.

18. Markov, A. A. (1916), Letter to the editorial office. Newspaper *Novoe Vremia*, 7 (20) Dec., pp. 7 – 8. (R)

19. Markov, A. A. (1915), Seminarists and realists. In [2].

20. Markov, A. A. (1915), Letter to the editorial office. In [2].

21. Youshkevich, P. [S.] (1915), On a scientific polemic. IMI, vol. 34, 1991, pp. 207 – 209. (R)

22. Bortkevich, V. I. (von Bortkiewicz, L.) (1903), The theory of probability and the struggle against sedition. *Osvobozhdenie* (Stuttgart), book 1, pp. 212 – 219. Signed "B". Bortkevich claimed his authorship in 1910 (ZMNP, 2nd paging, No. 2, p. 353). Apparently printed only in some copies of the periodical. (R) **S, G,** 4.

23. Sheynin, O. B. (1966), On the history of the iterative methods of solving systems of linear algebraic equations. Manuscript.

24. Markov, A. A. (1912), A rebuke to Nekrasov. MSb, vol. 28, pp. 215 – 227. **S, G,** 4.

25. Liapunov, A. M. (1901), An answer to Nekrasov. *Zapiski Khark. Univ.*, vol. 3, 1901, pp. 51 – 63. **S, G,** 4.

26. Posse, A. K. (1915), A few words about Nekrasov's article. ZMNP, 3rd paging, No. 9, pp. 71 – 76. (R) **S, G,** 95.

27. Mikhailov, G. K., Stepanov S. Ya. (1985), On the history of the problem of the rotation of a solid about a fixed point. IMI, vol. 28, pp. 223 – 246. (R)

28. Gordevsky, D. Z. (1955), *K. A. Andreev etc*. Kharkov. (R)

29. Tsykalo, A. L. (1988), *A. M. Liapunov.* M. (R)

30. Nekrasov, P. A. (1900 – 1902), New fundamentals of the doctrine of probabilities of sums and means. MSb, vol. 21, pp. 579 – 763; vol. 22, pp. 1 – 142, 323 – 498; vol. 23, pp. 41 – 462. (R)

31. Liapunov, A. M. (1901), Nouvelle forme du théorème sur la limite de probabilité. *Mém.* [*Zapiski*] *Akad. Nauk*, vol. 12, pp. 1 – 24.

32. Markov, A. A. (1894), Extract from letter to Andreev. *Soobshchenia*, vol. 4, No. 4, pp. 146 –149. (R)

33. Markov, A. A. (1894), On Andreev's commentary. Ibidem, pp. 175 – 176. (R)

34. Markov, A. A. (1916), On an application of the statistical method. *Izvestia Akad. Nauk,* vol. 10, pp. 239 – 242. (R)

35. Nekrasov, P. A. (1915), Answer to Posse's objections. ZMNP, 4th paging, No. 10, pp. 97 – 104. (R) **S, G,** 95.

36. Korkin, A. N., Bobylev, D. K., Posse, A. K. (1893), Extract from letter to Moscow Math. Soc. MSb, vol. 17, pp. 386 – 391. (R)

37. Chebyshev, P. L. (1887), Sur deux théorèmes relatifs aux probabilités. In Russian. Translation: *Acta Math*., t. 14, 1891, pp. 305 – 315.

38. Uspensky, V. A. (1987), *Chto Takoe Nestandartny Analiz* (What Is Non-Standard Analysis). M.

39. Todhunter, I. (1871), *On the Differential Calculus and the Elements of the Integral Calculus.* 5th edition. London – New York. Translated into Russian by V. G. Imshenetsky. Psb, 1873.

40. Boussinesq, J. (1887), *Cours d'analyse infinitésimale*, t. 1, No. 1. Paris.

41. Nekrasov, P. A. (1912), The general main method of generating functions as applied to the calculus of probability and to the laws of mass phenomena. MSb, vol. 28, pp. 351 – 460. (R)

42. Ashby, W. Ross (1964), [Some remarks]. *In Views on General Systems Theory*. Editor, M. D. Mesarovich. New York. Quotation in text translated back from the Russian translation of this source (M., 1966, p. 177).

43. Nekrasov, P. A. (1888 – 1912). *Teoriya Veroyatnostey* (Theory of Probability). M., 1896; Psb, 1912. Lithogr. editions: 1888 and 1894.

44. Nekrasov, P. A. (1899), On Markov's article [48] and on my report [5]. *Kazan Izv*., vol. 9, No. 1, pp. 18 – 26. **S, G,** 4.

45. Nekrasov, P. A. (1915), Interpretation of Chebyshev's second theorem and its versions. MSb, vol. 29, pp. 315 – 343. (R)

46. Markov, A. A. (1898), Sur les racines de l'équation etc. *Izvestia Akad. Nauk*, vol. 9, pp. 435 – 446 [47, pp. 253 – 269].

47. Markov, A. A. (1951), *Izbrannye Trudy* (Sel. Works). N. p.

48. Markov, A. A. (1899), The law of large numbers and the method of least squares [47, pp. 233 – 251].

49. Markov, A. A. (1899), Application of continuous fractions to calculating probabilities. *Kazan Izv.*, vol. 9, No. 2, pp. 29 – 34. (R)

50. Markov, A. A. (1899), Answer to [44]. Ibidem, vol. 9, No. 3, pp. pp. 41 – 43. **S, G,** 4.

51. Markov, A. A. (1900 – 1924), *Ischislenie Veroiyatnostei* (Calculus of probability). Psb, 1900, 1908, 1913. M., 1924. German transl. of 2nd edition: 1912.
52. Fichtenholz, G. M. (1971), *Differential- und Integral-Rechnung*, Bd. 1. Aufl. 6. (1959, in Russian). Berlin.
53. Markov, A. A. (1910), Correcting an inaccuracy. *Izvestia Akad. Nauk,* vol. 4, p. 346. (R)
54. Andreev, K. A. (1894), A commentary on Imshenetsky's article [55]; an answer to Markov's criticism [32]. *Soobshchenia*, vol. 4, No. 4, pp. 150 – 160. (R)
55. Imshenetsky, V. G. (1894), Comparison of Bugaev's method of discovering rational fractional solutions of differential equations with other methods. Ibidem, No. 1 – 2, pp. 60 – 80. (R)
56. Bezikovich, A. S. (1924). Biographical essay [on Markov]. [51, 1924, pp. iii – xiv]. (R)
57. Barbèra, L. (1876), *Teorica del calcolo delle funzioni*. Bologna. We have not seen it.
58. Newton, I. (1687), [Mathematical Principles of Natural Philosophy]. Russian transl. from the Latin original by A. N. Krylov (1915 – 1916) in his *Sobranie Trudov* (Coll. W orks), vol. 7. M. – L., 1936.
59. Khvolson, O. D. (1916), *Znanie i Vera v Fizike* (Knowledge and Faith in Physics). Petrograd.
60. Buniakovsky, V. Ya. (1846), *Osnovania Matematicheskoy Teorii Veroyatnostei* (Principles Math. Theory of Probability). Psb.
61. Pearson, K. (1892), *Grammar of Science.* London. Russian transl.: Psb, 1911.
62. Bugaev, N. V. (1868), Mathematics as a scientific and pedagogical tool. MSb, vol. 3, 2nd paging, pp. 183 – 216.
63. Tsinger, V. Ya. (1874), Exact sciences and positivism. In *Otchet i Rechi v Torzhestvennom Sobranii Mosk. Univ*. (Report and Speeches, Grand Meeting Mosc. Univ. 1874). M., pp. 38 – 98 of second paging. (R)
64. Davidov, A. Yu. (1857), Theory of mean values with application to compiling mortality tables. Ibidem, first paging. (R)
65. Sbornik (1943), *Sbornik Statei k Trekhsotletiu so Dnya Rozhdenia Nyutona* (Coll. Articles Honouring 300 Years from Newton's Birth). M. – L., 1943.
66. Krylov, A .N. (1916), *Uchenie o Predelakh Kak Ono Izlozheno u Nyutona* (Doctrine of Limits As Explicated by Newton). Petrograd.
67. Falbork, G., Charnolussky, V. (1897), Primary public education. *Enz. Slovar Brokkhaus & Efron* (*Brockhaus & Efron Enc. Dict*.), vol. 40, pp. 728 – 770. (R)
68. Sheynin, O. B. (1996), Chuprov (1990, in Russian). Göttingen, 1996 and 2011.

IX

Correspondence between P. A. Nekrasov and A. I. Chuprov

Istoriko-Matematicheskie Issledovania, vol. 1 (36), No. 1, 1995, pp. 157 – 167

1. Introduction

I have written about Pavel Alekseevich Nekrasov (1853 – 1924) [1, § 1.5; 2; 3, § 1]. Now, I mention other sources [4; 5; 6] throwing light on his biography. The author of [6] explains Nekrasov's Weltanschauung and his later work by his aspiration for permeating social life by arithmology (in its wider sense). Be that as it may, I keep to my previous opinion [3, § 1.3], and, in particular, I am still believing that, from about 1900, Nekrasov's mathematical writings became unimaginably verbose, intrinsically connected with ethics, religion and politics, and therefore obscure. In addition, the term *arithmology*, even in its narrow sense, is no longer in use, and Polovinkin [6] should have explained his reasoning as well as the title of his article. See Note 6 to [viii].

Here, I only repeat that in 1893 Nekrasov became Rector of Moscow University; in 1898, warden of the Moscow educational region; and, in 1905, a prominent official at the Ministry of Public Education. Aleksandr Ivanovich Chuprov (1842 – 1908), Corresponding Member of the Imperial (Petersburg) Academy of Sciences, was a statistician, the father of zemstvo statistics, an economist and writer on current topics [7; 8]. For a long time, until the autumn of 1899, he taught at the Law faculty in Moscow. More widely known is his son Aleksandr.

Two letters from Nekrasov to Chuprov (1898 and 1899) are kept at the Central State Historical Archive (Fond 2244, inventory 1, No. 2124). The first is devoted to the teaching of the theory of probability at the Law faculty of Moscow University (see § 2) whereas the second one characterizes the general situation at the University and I think that it should be also adduced. Here it is.

Nekrasov – Chuprov, 17 Febr. 1899

Dear Sir, Aleksandr Ivanovich, – Today, your lecture, as I heard, had not taken place because of the pressure of a group of students who want to impede the course of studies. Since there exists another group of students seeking after the contrary, I am most zealously [!] asking you not to give in during your forthcoming lecture tomorrow, and, if possible, to carry it out. In this way you will undoubtedly contribute to putting an end to the students' unrest. […][1].

2. The letter of 1898

Nekrasov is known to have been advocating the inclusion of the theory of probability into the school curriculum. It occurred that he also thought of teaching this discipline to student-lawyers. His appeal to Chuprov, who was extremely influential in his field, was hardly official: the latter was not the Dean of the Law faculty.

At the turn of the 19th century, the possibility of using probability in statistics was already proved, – in England, for biological research, and, on the Continent, for the theory of stability of statistical series (Lexis, Bortkiewicz). Furthermore, already Quetelet applied elements of probability for studying moral statistics (the statistics of marriages, suicides and crime), which could have undoubtedly been useful for lawyers. Nevertheless, Nekrasov's programme (below) hardly mentioned that branch of statistics.

In 1896 Nekrasov accepted the *candidate composition* [9] written by Chuprov's son, then graduating from the Physical and Mathematical faculty of the University. There, the future scientist attempted, in particular, to study the interrelations of the statistical method with philosophy and logic, and Nekrasov could have well included the last-mentioned item in § 7 of his program, *The statistical method as one of the methods of cognition*. Finally, § 8 of Nekrasov's programme testifies to his interest in the application of the theory of probability to economics. Later he paid much attention to that issue [10; 11, 1912, Chapt. 5 of pt. 2], and, during 1918 – 1919, he read a special course *On the branches of mathematics necessary for the economic sciences* [12, p. 423] at Moscow University (for a single listener, A. A. Konüs).[2] Here, now, is Nekrasov's letter.

Nekrasov – Chuprov, 27 Jan. 1898

Highly respected Aleksandr Ivanovich, – I am sending you a copy of my memorandum about which I told you during our rendezvous and which I, as a person teaching the theory of probability, intend to submit to the Law faculty.[3] Other mathematicians will also probably sign it. The extent of teaching is determined by the appended programme; for the time being I am raising the issue only in its essence. I am convinced that a proper and skilful teaching of probability will heighten the lawyers' level of education and I hope that you will regard this matter with due sympathy and exert your influence at the Faculty in order to establish this teaching under the most favourable conditions that are especially necessary for an absolutely new undertaking. […]

[*Supplement 1.*]

A rough programme for teaching probability theory with applications to phenomena of public life to students of the law faculty[4]

1. Random phenomena and their probabilities. Examples of direct calculations of probabilities. The case of an infinite number of chances. Moral certitude.

2. The main theorems. The addition theorem. Contrary events. A group of all possible incompatible events.[5] A compound event and the notion of conditional probability. The multiplication theorem. Independent events and the multiplication theorem for them. Hypotheses. The theorem on total probabilities.

3. Probabilities of compound events in numerous trials. The case of constant probability in all trials. The case in which the probabilities of the events change from one trial to another.

4. The law of large numbers (LLN). An elementary derivation of the theory of Jakob Bernoulli and Poisson. Definition of the expectation of a random variable.[6] The Chebyshev form of the LLN. The Poisson and the Bernoulli theorems as particular cases of the

Chebyshev proposition. The boundaries of the action of the LLN. Examples provided by Ettingen and Bertrand.[7]

5. On probabilities a posteriori. The Baye [!] theorem and its corollaries. Examples. The subjective aspect of the notion of probability. The change of the posterior probability depending on the accumulation of data. The application of the Baye theorem to the derivation of the main theorem on testimonies.[8]

6. An elementary theory of the method of least squares. The principle of the arithmetic mean.[9] The measure of precision. Combination of observations having different measures of precision. The weights of the results. The method of least squares in cases of one and many unknowns.

7. Application of probability to statistics. Statistical data and the statistical method. Its field. The need for a special critical appraisal of statistical data. Phenomena in public life and the will as one of its causes.[10] Moral statistics. A classification of mass observations and phenomena. Regularities in phenomena of public life. An empirical determination of probabilities as one of the problems of statistics. Application to determining the probabilities of duration of life. The statistical method as one of the methods of cognition.

8. The influence of chances on estimating monetary undertakings. The importance of the LLN in the Bernoulli and Chebyshev forms for determining the value of sums and undertakings exposed to randomness. On fair money games. On insurance of property and life. On buying annuities.

9. Application of the theory of probability to legal proceedings. A caution regarding the conditions for the application of probability theory to verdicts and testimonies.[11] The difficulties in accomplishing these conditions in full. The change of the probabilities of phenomena after new testimonies and verdicts become known. Criminal statistics.[12]

[Supplement 2.] *To the Law faculty*

During the last half-century, the theory of probability together with its applications made more than a small progress for which it is considerably indebted to Russian scientists, suffice it to mention Buniakovsky, Davidov and especially Chebyshev.[13] The advances in probability were not however reflected upon the level of educating the students of the Law faculty since the teaching of that discipline is not assigned a proper place.[14]

It could hardly be doubted that the subject-matter of a science cannot be isolated from its main methods without causing damage to the teaching. Such an abnormal dissociation always led to stagnation hindering the correct interpretation of the appropriate phenomena, and moreover, precluding the expedient use of methodology. Regrettably, such dissociation, harmful for the success of education, exists between the sciences of jurisprudence, which are in charge of the phenomena in social and political life, and the theory of probability, which provides mathematical methods for their systematic investigation.

Professor [Yu. E.] Yanson, in his *Teoriya Statistiki* (Theory of Statistics). Psb, 1891, p. 490, characterized the abnormality of the situation in the following words:

Regrettably, statisticians are not sufficiently acquainted with the theory of probability whereas the mathematicians, who applied mathematical calculations to analysing numerical data on social phenomena, considered them as abstract magnitudes, did not take into account the special properties of these phenomena and arrived therefore at conclusions bordering on nonsense.[15]

The need to get rid of this dissociation by properly teaching the theory of probability and its applications to phenomena in social life can be justified by many considerations. Thus, the doctrine of probabilities provides a precise formulation and a complete interpretation of the so-called LLN, i.e., of mass phenomena to which social and state phenomena also belong. At the same time, this science offers methods for discovering the most cautious assumptions about future random phenomena, for example those concerning economic and financial life. Lastly, the application of probability theory to testimonies and verdicts cannot remain uninteresting for an educated lawyer.

Since social phenomena cannot at present be investigated without any knowledge of the theory, its conclusions are even now being partly reported at the Law faculty. Regrettably, the information provided is scanty, extremely fragmentary and not always precise. Furthermore, it is offered on trust, without any substantiation which is necessary not only for cogency, but also for ensuring a distinct understanding of the boundaries for applying the reported methods. So as to justify such an abnormal situation, it was reasoned that the mathematical analysis of probabilities demanded the use of higher mathematics, the acquaintance with which could not have been expected from student-lawyers. At present, however, this consideration had lost its meaning owing to the works of Chebyshev and the attempts of later Russian mathematicians.[16]

Chebyshev's outstanding merit consists not only in that he provided a more general expression of the LLN, but also in that he extraordinarily simplified the proof of this most important proposition of the theory.[17] Nowadays, all the essential parts of this doctrine and its applications can be taught in an elementary way when issuing from the mathematical knowledge determined by the gymnasium programme.

This discipline can thus be made intelligible to student-lawyers. It is self-evident that the lectures in probability adapted for them must differ from those read to future mathematicians, but they can retain precise scientific character and be rich in their content.

The reader in probability must naturally take care that its teaching be of an adequate scientific level and sufficiently disseminated in the University. This consideration prompts me to raise before the Faculty the issue in principle about the teaching of the theory of probability with its applications to student-lawyers. This problem interested me for a long time; at present, I have compiled quite a definite plan for its solution and presented it in the subjoined rough programme that can be made use of for teaching the theory of probability with its applications to student-lawyers. Two hours a week lasting for an academic year would be quite sufficient for a conscientious mastering of this course.

When studying the theory according to this programme, student-lawyers will encounter difficulties; these, however, will be caused not by the complexity of mathematical analysis, that will not go beyond the gymnasium curriculum, but by the intricacy of the ideas and notions that form the subject-matter of the science of random phenomena.[18] He who successfully overcomes these difficulties will more distinctly understand the laws of social phenomena. In raising the issue of teaching the theory of probability at the Law faculty, I consider it necessary to state that favourable conditions ensuring adequate success be provided for this. If this issue will be satisfactorily solved, the teaching will not present any difficulties for the personnel at the disposal of the University.[19]

Notes

1. Next year, 9 February 1900, Nekrasov will write to F. E. Kosh (1843 – 1915), a philologist and orientalist:

Up to now, there is absolute order at Moscow University, but the nearest future is full of uncertainty

(Archive, Russian Acad. Sci., Fond 558, inventory 4, No. 235). Also see the appropriate passage in [2, § 1].

2. In November 1989 Konüs told me that Nekrasov's lectures had included an examination of the work of Walras, the founder of the mathematical school in economics. Judging by its title, Nekrasov's course had much in common with his report [13].

3. I emphasize that Nekrasov, still the Rector of the University [14], did not pull rank.

4. The titles of several sections of this programme coincided with those of the appropriate chapters of Nekrasov's treatise [11, 1896 and/or 1912].

5. The term *a group of ... events/phenomena* occurs also in Nekrasov's treatise [11, 1896, p. 13; 1912, p. 221] without any special emphasis on *group*.

6. Nekrasov was one of the first to apply this term (in Russian, *random magnitude* [1, p. 350, Note 17]). However, the absence of the notion of density implies that he restricted his attention here to discrete variables. Incidentally, he was thus unable to mention the normal distribution which considerably worsened his programme.

7. I am unaware of the former and I doubt that either was essential.

8. This term is not in common use. Later Nekrasov [15, p. 14] called the formula of the type

$$P = pA/(pA + qB)$$

the main equation of the probabilities of testimonies; p and A were the probabilities of the truthfulness of the witness of the event in question and of the subsequent narrator, $q = 1 - p$ and $B = 1 - A$. Already Condorcet [16, p. 400] introduced this formula.

9. Otherwise: the Gauss postulate (1809) according to which the arithmetic mean of observations coincided with the mode of the unimodal curve of distribution of their errors.

10. It seems that Nekrasov had grossly exaggerated the importance of (free) will. The regularities in public life, which he obviously had in mind, were caused not by the action of will, but by the specific nature of mass phenomena. My statement excludes public outbursts, or social will. Nekrasov largely devoted his writing [17] to free will.

11. Nekrasov apparently thought about the (non-existent) independence of the judgements passed by the jurors.

12. Vlasov [18], see Note 14 below, had not discussed the subject-matter of Nekrasov's §§ 6, 8, or 9, but he included elements of the theory of stability of statistical series lacking in § 7.

13. Not needed.

14. Here and below Nekrasov directly or implicitly stated that some elements of probability were nevertheless reported at the faculty. According to official documents, the theory of probability was not taught there in 1902 – 1903 or 1912 – 1916 [19], and only in 1907 – 1908 A. K. Vlasov delivered a course of lectures in that theory [18]. At that time, and also in 1912 – 1916, statistics was also taught at the Law faculty. I have no information relating to 1899 – 1907 or 1908 – 1912. Note that in 1908 Vlasov edited a Russian translation of Laplace's *Essai philosophique* and that in 1911, after about twenty years there, he had to leave the University [20] because of the worsening of its social and political atmosphere [21, pp. 375 – 377].

15. Quetelet [22, p. 633] pronounced a similar statement, but he only mentioned *des prétendus savants*. Is it true, however, that mathematicians rather than the statisticians themselves arrived at senseless results? Buniakovsky [23, p. 154], who was both a mathematician and a statistician, remarked, although not in a statistical context, that

Anyone who does not examine the meaning of the numbers with which he performs particular calculations, is not a mathematician.

16. Along with Nekrasov's example below, it can be indicated that Chebyshev, in his Master dissertation [24], had indeed explicated the theory of probability by elementary means but his description was ponderous.

Vlasov (Note 14) also managed without higher mathematics in his textbook.

17. In describing Chebyshev's merits Nekrasov possibly did not want to go beyond the boundaries of his programme. Nevertheless, Nekrasov denied the importance of the Chebyshev's proof of the central limit theorem, see Note 26 in [viii].

18. At the time, this definition of probability theory, although formulated indirectly, was indeed fortunate.

19. I do not know whom Nekrasov had borne in mind. He himself left the University two months afterwards [14]. Moreover, during 1902 – 1904, 1912 – 1913, 1914 – 1915 and 1916 – 1917 no one had been teaching the theory of probability even at the Physical and Mathematical faculty [25]! I have no information about 1904 – 1912, but during 1913 – 1914 and 1915 – 1916 the theory was indeed taught there by L. K. Lakhtin (Ibidem). Incidentally, all this testifies against the Nekrasov – Florov proposal that probability theory be introduced into the school curriculum [26]. Some participants of the then ensuing debate (Ibidem, No. 3) had indeed expressed doubts about the availability of qualified school teachers.

References

L, M, Psb, R = Leningrad, Moscow, Petersburg, in Russian

1. Sheynin, O. B. Markov's work on the theory of probability. *Arch. Hist. Ex. Sci*, vol. 39, 1989, pp. 337 – 377.

2. Sheynin, O. B. (1993), Markov's letters in the newspaper *Den*. IMI, vol. 14, pp. 194 – 209. **S, G,** 85.

3. [viii].

4. Anonymous, Nekrasov. *Novyi Enz. Slovar Brokgaus i Efron* (New Brockhaus & Efron Enc. Dict.), vol. 28, 1916, p. 272.

5. Sluginov, S. P. Nekrasov. T*rudy Matematich. Seminaria Permskogo Gos. Univ.*, No. 1, 1927, pp. 37 – 38. (R)

6. Polovinkin, S. M. The psycho-arithmo-mechanician. Philosophical features of Nekrasov's portrait. *Voprosy Istorii Estestvoznania i Tekhniki* No. 2, 1994, pp. 109 – 113. (R)

7. Kablukov, N. [**A.**] A.I. Chuprov, a biographical essay. In Chuprov, A. I. *Rechi i Statyi* (Speeches and Papers), vol. 1. M., 1909, pp. xiii – xlviii.

8. Koni, A. F. From [my] recollections of Chuprov. Ibidem, vol. 3, pp. ix – xli.

9. Chuprov, A. A. *Matematicheskie Osnovania Teorii Veroiatnostei*

(Math. Principles of Theory of Statistics). Thesis. M., 1896. Unpublished. Gorky Library at Moscow Univ., Section rare books & MSS. The Fond of A. I. & A. A. Chuprov, karton 9, item 1. (R)

10. Nekrasov, P. A. Mathematical statistics, commercial law and financial turnover. *Izvestia Russk. Geografich. Obshchestvo*, vol. 45, 1909, pp. 333 – 398, 565 – 612 and 811 – 896. (R)

11. Nekrasov, P. A. *Teoriya Veroyatnistei* (Theory of Probability). M., 1896; Psb, 1912.

12. Komlev, S. L. On the conjuncture statistics of the 1920s. Conversation with A. A. Konüs. *Ekonomika i Matematich. Metody*, vol. 25, 1989, pp. 423 – 434. (R)

13. Nekrasov, P. A. On the sections of mathematics necessary for the economic sciences. *Matematich. Obrasovanie*, No. 2, 1912, pp. 79 – 81. (R)

14. Anonymous, The new warden of the Moscow educational region. Newspaper *Moskovskie Vedomosti*, 1898, March 13 (25), pp. 2 – 3 and 15 (27), p. 2. (R)

15. Nekrasov, P. A.. *Srednyia Shkola* etc. (High School, Mathematics and Scient. Training of Teachers). Psb, 1916.

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

17. Nekrasov, P. A. *Filosofiya i Logika* etc. (Philosophy and Logic of Science of Mass Manifestation of Human Activities). M., 1902.

18. Vlasov, A. K. *Teoriya Veroyatnistei* (Theory of Probability). M., 1909 and 1916.

19. *Obozrenie Prepodavaniya na Yuridicheskom Fakultete na* … (Review of Teaching at Law Faculty, Imp. Mosc. Univ. for … Academic Year). Appeared yearly. N. p., no dates, no title-pages.

20. Glagolev, N. A. A. K. Vlasov (1868 – 1922). MSb, vol. 32, 1925, pp. 273 – 275. (R)

21. *Istoriya Moskovskogo Universiteta* (History of Moscow Univ.), vol. 1. M., 1955.

22. Quetelet, A. Unité de l'espèce humaine. B*ull. Acad. Roy. Sci., Lettres et Beaux-Arts Belg.*, sér. 2, t. 34, 1872, pp. 623 – 635.

23. Buniakovsky, V. Ya. *Essay on the laws of mortality in Russia and on the distribution of the Orthodox believers by ages*. *Zapiski Imp. Akad. Nauk Psb*, vol. 8, 1866, Suppl. 6. Separate paging. (R)

24. Chebyshev, P. L. *Opyt Elementarnogo Analiza Teorii Veroiyatnostei* (Essay on Elementary Analysis of Theory of Probability). (1845). *Polnoe Sobranie Sochineniy* (Complete Works), vol. 5. M. – L., 1961, pp. 26 – 87.

25. *Obozrenie Prepodavaniya na Fiziko-Matematicheskom Fakultete na …* (Review of teaching at the Phys. and Math. Faculty [of Imp. Moscow Univ.] for … Academic Year). Appeared yearly. N. p., no dates, no title-pages.

26. Nekrasov, P. A. The theory of probability and mathematics in high school. *Zhurnal Ministerstva Narodnogo Prosveshcheniya*, 1915, 4th paging, No. 2, pp. 65 – 127, No. 3, pp. 1 – 43, No. 4, pp. 94 – 125. (R)

Markov and life insurance

Istoriko-Matematicheskie Issledovania, vol. 2 (37), 1997, pp. 22 – 33
Math. Scientist, vol. 30, 2005, pp. 5 – 12

Note by publicator

I am keeping to the original Russian text but certainly take into account its shortened translation of 2005.

1. Introduction

General literature on Markov is [24; 31]. His newspaper letters are not included in the Bibliography of his works [1], neither did Grodzensky [2] cite them although it was he from whom I first came to know about their existence. I comment on these Notes in § 4 and reprint them in § 5, and in § 3 I describe the barely known Markov's activities in insurance. My § 2 barely discusses the history of insurance since its material is largely known. There, I draw, among other sources, on papers [3] and [4]. Hald [5] devoted much attention to this subject.

I define life insurance as any agreement ensuring payments of definite sums either to the heir(s) of the insured should he/she die within a stipulated period of time (a lump sum), or to the insured himself (regular sums, and, especially, an annuity). According to modern ideas, but not in line with the practice of insurance during the 17th and 18th centuries, the price of such agreements must be determined by means of mortality tables depending on the age and the sex of the insured. My definition does not cover all the existing forms of life insurance (§ 4), but it is sufficient for a general understanding of the matter[1]. I also note that various kinds of mutual insurance of several people have also been widely used. Thus, upon paying a necessary sum, a married couple could have enjoyed a fixed annuity until one of them dies, with the surviving spouse continuing to draw it to the end of his/her life.

In England, societies offering mutual insurance had already been in existence in the 17th century. At the turn of the next century that country had several thousands of them. Their members drew insurance in cases of illnesses or death of their wives, and wives received it upon the death of their husbands. It seems [6] that most of such societies existed on voluntary dues, but that in any case there had been no connection between the premiums and the ages of their members.

An operation connected with risk is called fair if the expectation of the winning (ξ) is zero ($E\xi = 0$). For an insured, insurance is never fair: since insurance societies cannot exist without profit, his/her expectation is always negative. Nevertheless, insurance might be advantageous for the insured, if, for example, his family will get a lot of money should he die prematurely. And, indeed, such scientists as Laplace [7, pp. 898 – 890] ardently approved of the institution of life insurance. In 1898, more than 7*mln* people were insured the world over, about 0.1*mln* of them in Russia [8, p. 747] which goes to show the scale of the activities of the main insurance enterprises roughly at the time that directly concerns us.

2. From the history of life insurance

Population statistics had been the most important branch of political arithmetic that emerged in the mid-17th century and at least until the beginning of the 19th century the former remained significant mainly because the developing insurance business were demanding reliable data on mortality and studies of its laws.

These statistical data, insofar as they were being collected by insurance societies, had been kept secret, but the theoretical principles were not concealed. Their development both directly and implicitly heightened the interest in probability and to some extent fostered its advancement.

In 1669, in a letter to his brother Lodewijk devoted to various problems in mortality and published in 1895 [9], Christiaan Huygens calculated the expectations of the order statistics for an empirical distribution, introduced the concepts of mean and probable durations of life and constructed and made methodological use of a graph of a continuous function

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

where $F(x)$ in my notation was an integral distribution function of mortality. It was in this correspondence that the theory of probability went beyond the province of games of chance (as it also did in 1671 at the hands of De Witt).

In 1709 Niklaus Bernoulli [10] considered a number of problems connected with insurance. In one of these he (pp. 296 – 297, also see [11, pp. 195 – 196]) determined the expectation of the maximal element of a sample from a continuous uniform distribution. Issuing from statistical data published by Halley in 1693, De Moivre [12] proposed to describe mortality, beginning with age 12, by a uniform distribution. There also he introduced the expectation of a random variable thus distributed (Problem 20 from pt. 1) and calculated probabilities of the type

$$P(\xi \geq x) = 1 - F(x)$$

for the same distribution (Chapt. 8 of pt. 2, my own notation).

Laplace [13, Chapt. 9] solved several problems on life insurance in the same way as those pertaining to the treatment of observations, but this time he also discussed the so-called Poisson generalization of the Bernoulli trials. Gauss did not shun life insurance either; he had to solve practical problems while managing the pension fund at Göttingen University [14, pp. 61 – 64]. In Russia, Zernov [15] published a treatise in which he paid special attention to life insurance and Buniakovsky, in 1846, devoted a chapter of his celebrated work to the same subject.

3. Markov's work in retirement funds

Retirement funds began appearing in Russia in the second half of the 19th century. After retirement, their members had been drawing lifelong pensions depending on the duration of their work and their final or mean salary. It was indeed possible to estimate the duration of life of the pensioners by applying mortality tables (although, strictly speaking, statistical inferences suitable for the general population will not do for its special groups), but it was extremely difficult to predict the yearly number of the retiring or their salaries whereas the estimation of the number of additional members of a given fund admitted for state reasons (see below) was absolutely impossible. Consider also that the widows and children of dying members were also provided with life annuities or long-term pensions, and it becomes evident that any retirement fund could have went broke, and especially so during its first years of existence when experience was still lacking and unreliable guesswork was necessary.

Ostrogradsky [16] participated in the work of the first Russian retirement fund. A few decades later Markov began to busy himself with similar activities; already in 1884 he [17] published detailed calculations for the retirement fund at the Ministry of Justice. In 1890 he became member of its governing board [18, vol. 2, p. 36]. He actively participated in its sittings, offered his advice about concrete issues and checked book-keeping accounts. Thus, he compiled a note [19, vol. 2] (not mentioned in his Bibliography [1]) on the financial conditions necessary for ensuring the payment of pensions. Vol. 1 of the same source [18] contains many references to Markov, and on pp. 90 and 100 it cites pp. 10 and 6 respectively of a certain note, possibly [17], since [19] is only two pages long.

In 1893, 1894 and 1902 Markov received letters of thanks from the Ministry of Justice [2, p. 59]. Incidentally, its retirement fund was considered the *best established* among the six funds of the *civil departments*, and this fact was naively attributed to Markov's *precise mathematical calculations* [20]. It would have been more correct to mention his prudence and foresight, perhaps his intuition and ability to detect the slightest circumstances.

Markov also occupied himself with similar work at the War Ministry [21]. In 1900 Academician Sonin [22], on behalf of the Physical and Mathematical Class of the Academy of Sciences, acquainted himself with the work of the Ministry's retirement fund and expressed his opinion about it in the following way:

The sole reason for the crisis that it experiences now was the unforeseeable increase in the number of its members which occurs through instructions from above. He recommended to liquidate the fund and to transfer its liabilities to the state.

After hearing this out, Markov (Ibidem) declared that he did not agree with Sonin *on any point*. The Class resolved that, since the problem posed by the War Ministry [before the Academy] was rather of a practical than purely scientific nature, it should only inform the Ministry that *the members of the Academy are always ready to render assistance* to it.

Also in 1900 the same Ministry established a *Special* [standing] *Mathematical Conference* for determining the financial state of its fund [23, p. 10]. Its members included academicians Markov, Sonin and I. I. Yanzhul, other eminent scientists (I. I. Pomerantsev, N. Ya. Tsinger) and actuaries (B. F. Maleshevsky). Regrettably, nothing is known either about the work of this Conference or of Markov's even more active participation in practical life insurance after his retirement in 1906 [24, p. 604].

Again in 1900, Markov devoted to life insurance a short chapter of his textbook [25]. There, not aiming at new results, he acquainted his readers with the main stochastic problems of the contemporaneous insurance business. I indicate, finally, that the Markov Fund (Fond 173, inventory 1) at the Archive of the Russian Academy of Sciences includes three letters directly pertaining to my subject.

1) An undated Markov's letter to Maleshevsky (Delo 60, No. 15). Markov disapprovingly mentioned the *just appeared* book of Savich [26] and noted that it [prompted him] *to turn attention once more* [!] *to the theory of inability to work*. He also discussed one of Maleshevsky's formulas and expressed his opinion about the mortality of the disabled.

2) D. A. Grave's letter to Markov of 21 April 1916 (Delo 5, No. 5). Grave indirectly agreed with Markov in that the granting of some kind of pensions was undesirable.

3) Another letter from Grave to Markov of 13 Dec. 1916 (Delo 5, No. 7). Grave mentioned a *surprising discordance* between the calculations made by Markov and Maleshevsky. These apparently concerned the work of the pension fund in the city of Chernigov.

4. Markov's newspaper publications

In §5 I reprint two polemic newspaper letters published by Markov and devoted to the insurance of children. It may be thought that this kind of insurance was more or less widely practised in Russia from at least the mid-19th century. In any case, Kraevich [27] included an appropriate example in the first three editions of his collection of mathematical exercises for school students. Here it is. Upon the birth of a boy, his father deposits 1000 roubles with an insurance society. In exchange, the son draws x roubles after his 20th birthday, but the money is lost if he dies before that date. Assuming that the insurance is fair, and that the interest rate is 5%, determine x by means of the appended mortality table (whose origin is not explained).

Both this simplest pattern, and the other one criticized by Markov (below), and, as it may be supposed, any other scheme for insuring children, suffers from one and the same essential defect: they necessarily remain unfavourable for the insured (see § 1; fair insurance is only possible in textbooks), and they do not really *insure* him. Here is a relevant passage whose author mentions, among other types of insurance, the insurance of children against death [28, p. 243]:

Cases that, under the guise of insuring life, conceal deals in paying out some moneys upon the occurrence of a stipulated event not inflicting [pecuniary] *loss on the insured, – deals which do not restore actual damage, – should not be attributed to insurance. They abuse the idea of insurance*.

Markov's criticism was justified; regrettably, however, he did not take the occasion to explain to his readers that there exist other forms of life insurance (and of insuring property) advantageous for the insured.

5. Markov's Letters

Letter No. 1. Newspaper *Nasha Zhisn*, 7 April 1906, p. 1

The Benefits of Insurance through the Savings Offices

In order to ascertain the benefits of insuring profits and capitals through the state savings offices […] it is necessary to consider the tariffs. Judging by the number of these (V – X), the insurance of juveniles plays a large part in the new direction of business of the savings offices. Who will benefit from this insurance, except *those engaged in this operation*?[2] To answer this question it is necessary to dwell on the tariffs of insurance from which we extract two lucid examples.

1) According to tariff VI, a down payment of 1200 roubles is necessary for a six-year-old child to draw 2000*r* after reaching the age of twenty; and, should the child die prematurely, only 1200 – 60 = 1140*r* are returned back (5% is retained to cover the expenses). On the other hand, if the same sum, 1200*r*, be kept at a bank with an interest rate of 4% (this is the rate underlying the tariffs) for each full hundred roubles,[3] then, consecutively,

1200 + 4% = 1248 at seven years; […] 1964 + 76 = 2040 at twenty years.[4]

A rough check of this last figure is provided by noting that $1200 \cdot 1.04^{14} = 2078$.

My calculation shows that this insurance is in all cases disadvantageous for the family. If the child survives until age 20, the loss will be expressed by a small sum of 40*r*; otherwise, it can amount to several hundred roubles since the family loses the interest on the down payment.

2) According to tariff VIII, a yearly grant of 600*r* during five consecutive years will be paid out to a six-year-old child after his reaching age 18 for a down payment of 1789*r*; and, should the child die before that age, 1789 – 89 = 1700*r* are returned back.

1789 + 68 = 1857 (age, seven years); […] 2729 + 108 = 2837 (age, eighteen years).

So, when paying out the 600*r* for five years, we obtain consecutively

2837 – 600 = 2237; […] 637 + 24 – 600 = 61.

It is seen that this operation also inflicts a loss for the family. In the favourable case this loss is expressed by a small sum of 61*r*; otherwise, it can amount to a thousand roubles.

As indicated above, I have chosen lucid examples, but similar results are obtained in the other cases as well with the only difference being that, for the most favourable instances, the small loss can become a small profit. However, the possibility of large losses for the family because of a premature death of its child persists.

Letter No. 2. Newspaper *Nasha Zhisn*, 2 May 1906, p. 1

The benefits of insurance through the savings offices

The explanations provided by the Directorate of the savings offices [29] do not explain anything; on the contrary, they obscure the essence of the problem that I have raised.[5] They are composed in such a way as though the whole matter consists in the high cost of insurance through savings offices, which, in my example, was expressed by a small sum of 40*r* out of 2000. Dwelling only on this small loss, the Directorate maintains that it is of no consequence owing to the security of savings through the insurance and is compensated by profit sharing. And, assuming an interest rate of 5% rather than 4%, the Directorate promises the insured a pay out of 2180 – 2280 instead of the 2000.

Thus, the Directorate completely overlooks those cases in which the insured child dies prematurely and the family's loss due to the insurance is expressed already by hundreds rather than tens of roubles. Only by forgetting these instances it is possible to bring oneself to state that the savings are here secured. Meanwhile, in my first Letter I had paid attention to these cases; and, for determining the loss incurred by the insurance to the family, I had adduced, in addition to the figure 2040 on which the Directorate rests its eyes, a number of other ones. The Directorate apparently chose only the most favourable case; and it vainly tries to prove that in this instance the family's loss can be replaced by some profit. Indeed, I had mentioned this possibility in that Letter: suffice it to change the age of the insured and the duration of the contract in such a way that the probability of losing the stipulated insurance heightens.

As to the method by which the Directorate attempts to replace the loss by a profit, it cannot be called proper not only because, instead of providing a detailed calculation, it only indicates an indefinite magnitude between 2180 and 2280, but, mainly, since it admits that its estimation was based on changing the interest rate. The Directorate obviously forgot that after 14 years and assuming a 5% yearly interest rate, a capital of 1200*r,* when saved without any insurance being involved, fetches not 2040, but $1200\ 1{,}05^{14} = 2374r$[6] Nevertheless, I am quite prepared to agree with the Directorate that my calculations were based on a too low rate of interest, witness for example the latest pleasing loan.[7] But an increase in the rate increases the family's loss incurred by the insurance. The Directorate's statement about an exaggeration in my reckoning is therefore absolutely wrong. Thus, my indication that some insurance procedures offered by the savings offices always lead to losses remains unshaken. Neither can it be shaken until the mortality table taken as the basis for computing the tariffs of insurance remains unaltered and the expenses (5%) of carrying out the insurance are not lowered.

Indicating profit sharing, the Directorate says that five years after the insurance operations begins profit will be shared among the insured; but it forgets to mention that a considerable part (up to 25%) of the profit will go to the employees of the savings offices.

Defending its future operations of insuring juveniles, the Directorate refers to [private] insurance societies where such operations are carried out according to higher tariffs, but, regrettably, it does not corroborate this statement by comparative excerpts. My remarks undoubtedly concern these societies as well, but it is also obvious that insurance societies aim at getting rich, and this distinct goal can serve as a warning to those insuring. On the other hand, the fact that some operations are being carried out, is no proof that they should indeed be done. For example, a lot of people gather to play the roulette in Monaco – so should not we therefore arrange that game, or something similar, at the savings offices? For anyone who read my first Letter it should be clear that all the conclusions there contained only concern the insurance of juveniles. The tariffs of insurance as carried out by the savings offices cover, however, not one single form of insurance, as it could be understood from the words of the Directorate, but several forms, so that, according to the number of the tariffs involved, the insurance of juveniles occupies a rather considerable place among the new operations at the savings offices. And I have provided examples concerning two different tariffs.

I have not touched on other kinds of insurance so that the Directorate apparently vainly defends them; and the more so since its arguments reduce to a statement that for 70*r* it is possible, given some conditions, to draw 1000*r*. Is the Directorate really so naïve as to attach serious meaning to this proposition? Having 70*r* and playing the roulette game it is possible to win even more than 1000*r*.

Thus, I have spoken only about some forms of insurance whereas the Directorate itself raised the question about the high, or the low cost of all kinds of its insurance but has not provided any proof of the latter; it did not even adduce comparative passages from its own tariffs and those of insurance societies. For my part, I remark that if, contrary to expectation, insurance through savings offices will prove to be cheap for the insured, it will be expensive for the state, provided of course that the business will be widespread since the expenses will then not be low.[8]

Notes

1. Elsewhere Markov [24], only in the edition of 1908, on p. 97, when referring to his letters and to the Explanation [28], contrasted various forms of insurance. O. S.

There exist also [!] such insurance operations which do not protect against any risks, and in all cases inflict some greater or lesser damage on the insured. Such operations may be justified […] only by a rather doubtful consideration that they compel people to save money. Note, however, that parents (when juvenile insurance is concerned) become directly interested in the pecuniary sense in that their insured children remain alive. A. Markov

2. Markov bears in mind the employees of the offices. O. S.

3. Interest was paid on the sum rounded down to the nearest hundred. A. M.

4. Markov had written down all the intermediate results (here omitted). The same will be true in two other cases below. O. S.

5. [Markov refers to his Letter 1.] The Directorate maintained that the insurance of juveniles ensures but little profit: it *comes close* […] *to simple saving*. Then, private insurance societies offer even worse conditions for the insured; the psychological aspect of being protected from chance by insurance is important; if, after some time, the savings offices show a profit higher than 4%, the surplus will be given over to those insured. A. M.

6. [More correctly, 2376.] It is obvious that the restriction concerning the interest (Note 3) did not apply to the savings offices themselves. A. M.

7. Markov possibly referred to the *Short-Term Treasury Bonds* issued on 9 December 1905 and yielding a 5.5% rate of interest [29, p. 67]. O. S.

8. It seems that the only explanation here is that *cheap* means *almost fair*. O. S.

References

AHES = *Arch. Hist. Ex. Sci.*
M, R, Psb = Moscow, in Russian, Petersburg

1. Alekseeva, V. P. Bibliography of the works of Markov. In Markov, A. A. *Izbrannyie Trudy* (Sel. Works). N. p., 1951, pp. 679 – 714.
2. Grodzensky, S. Ya. *A. A. Markov*. M., 1987. (R)
3. Sheynin, O. B. Early history of the theory of probability. AHES, vol. 17, 1977, pp. 201 – 259.
4. Kohli, K., van der Waerden, B. L. Bewertung von Leibrenten. In Bernoulli, J. *Werke*, Bd. 3. Basel, 1975, pp. 515 – 539.
5. Hald, A. *History of Probability and Statistics* and *Their Applications before 1750.* New York, 1990.
6. Wells, A. F. Friendly society. *Enc. Brit.*, vol. 9, 1965, pp. 935 –938.
7. Laplace, P. S. *Philosophical Essay on Probabilities*. New York, 1995. (French, 1814). Translated by A. I. Dale.
8. Press, A. Insurance. *Brockhaus & Efron Enc. Dict.*, vol. 62, 1901, pp. 736 – 782.
9. Huygens, C. Correspondance (1669). *Oeuvr. Compl*., t. 6. La Haye, 1895, pp. 483 and 526 – 538.
10. Bernoulli, N. *Dissertatio inauguralis mathematico-juridica de usu artis conjectandi in jure* (1709). In Bernoulli, J. [4, pp. 287 – 326].
11. Todhunter, I. *History of the Mathematical Theory of Probability* (1865). New York, 1949, 1965.
12. De Moivre, A. *Treatise on Annuities on Lives* (1725). In author's *Doctrine of Chances,* 1756, pp. 261 – 328.
13. Laplace P. S. *Théorie analytique des probabilités. Oeuvr. Compl.,* t. 7. Paris, 1886.
14. Sheynin, O. B. Gauss and the theory of errors. AHES, vol. 20, 1979, pp. 21 – 72.
15. Zernov, N. *Teoriya veroyatnostei s prilozheniem preimushchestvenno k smertnosti i strakhovaniu* (Theory of Probability with Applications Mainly to Mortality and Insurance). M., 1843.
16. Ostrogradsky, M. V. Note on the retirement fund (1858). *Poln. Sobr. Trudov* (Complete Works), vol. 3. Kiev, 1961, pp. 297 – 300.
17. Markov, A. A. *Zapiska o raschete veroyatnikh oborotov emeritalnoi kassy sudebnogo vedomstva* (Note on Calculating the Probable Turnovers of the Retirement Fund of the Finance Ministry). Psb, 1884.
18. *Trudy po obzoru deistviy emeritalnoi kassy vedomstva Ministerstva justitsii za pervoe piatiletie* (Review of the Activities of the Retirement Fund of the Ministry of Justice for the First Five Years), vols. 1 – 2. Psb, 1890 – 1891.
19. Markov, A. A. Zapiska o vychislenii kapitalov, obespechivayushchikh operatsii kassy (Note on calculating the capitals needed for the operations of the fund) [18, vol. 2, pp. 131 – 132].
20. Lykoshin, A. S. Retirement funds. *Brockhaus & Efron Enc. Dict.*, vol. 40A, 1904, pp. 726 – 729.
21. Markov, A. A. *Zapiska po voprosam, rassmotrennym IV poverochnoi komissiey* [emeritalnoi kassy voenno-sukhoputnogo vedomstva] (Note on Issues Considered by the 4th Control Commission of the Retirement Fund at the Department of Military Land-Forces). [1899].
22. *Protokoly Imp. Akademii Nauk. Fiz.-mat. otdelenie.* Protokol No. 1, (Proc. Imp. Acad. Sci., Phys. and Math. Class., Proc. No. 1), 19 Jan. 1900. Published as a manuscript.
23. *Otchet o denezhnykh oborotakh emeritalnoi kassy voenno-sukhoputnogo vedomstva za 1899 god* (Report on the Turnover of the Retirement Fund, Department of Military Land-Forces for 1899). Psb, 1900.
24. Markov, A. A. Jr, Biography of A. A. Markov, Sr. In [1], pp. 599 – 613. **S, G**, 5.
25. Markov, A. A. *Ischislenie veroyatnostei* (Calculus of Probability). Psb, 1900, 1908, 1913; M., 1924. German transl.: Leipzig – Berlin, 1912.

26. Savich, S. E. *Elementarnaya teoriya strakhovaniya* (Elementary Theory of Insurance). Psb, 1900.
27. Kraevich, K. D. *Sobranie algebraicheskikh zadach* (Collection of Algebraic Problems). Psb, 1864, 1867, 1874, 1882.
28. Nikolsky, P. A. *Osnovnye voprosy strakhovaniya* (Main Issues in Insurance). Kazan, 1895.
29. On insurance through the savings offices. Newspaper *Nasha Zhizn,* 18 April 1906, p. 2.
30. *Ezhegodnik Ministerstva Finansov* (Yearbook of the Ministry of Finance). Issue 36 for 1906/1907. Psb, 1907.
31. Sheynin O. B. Markov's work on probability. AHES, vol. 39, 1989, pp. 337 – 377; vol. 40, p. 387.

Slutsky: Commemorating the 50th anniversary of his death

Istoriko-Matematicheskie Issledovania, vol. 3 (38), 1999, pp. 128 – 137
(lacking § 3.3)

1. Introduction

Many authors [7; 27; 1; 4; 5, 32; 8; 20][1] described the life and work of Evgeni Evgenievich Slutsky (1880 – 1948), an outstanding mathematician, statistician, and economist, and his most important writings are available in a one-volume edition [26]. I am therefore restricting my main goal and publish or describe a few archival letters either written by, or having to do with him (§ 3). In addition, I say a few words about Slutsky's life (below) and throw light on the events which apparently compelled him to abandon economics (§ 2)[2].

In 1920 Slutsky became Professor at Kiev Commercial Institute. However, he had not mastered the Ukrainian language which was then made compulsory for academic institutions, and in 1926 he had to move to Moscow and to start working there at the Central Statistical Directorate [4, p. 268], and, at the same time, at the Conjuncture Institute under the Finance Ministry [5, p. 8].

Already then Slutsky busied himself in real earnest with applying his statistical research to geophysics. Being forced to abandon his activities in economics (§ 2), he [4, p. 270], for a few years,

W*ent over to working in institutes connected with geophysics and meteorology where he* […] *hoped to find application for his discoveries in the field of pseudo-periodic waves*[3].

He had not found suitable conditions for theoretical research (Ibidem), and in 1934 he moved to the Moscow State University, then (in 1939) going over to the Steklov Mathematical Institute. The University conferred on him the degree of Doctor of Physical and Mathematical Sciences honoris causa [4, p. 271].

Slutsky was an original and deep researcher. He is mostly known as a cofounder of the purely mathematical theory of probability and the theory of random processes, and remembered for his application of stochastic ideas and methods in economics and geophysics (especially in studying solar activity) and as a compiler of important mathematical tables which constituted *a masterpiece of the art of calculation* [27, p. 417].

Slutsky's contribution to the theory of consumer's demand is *very valuable* [1, p. 210]. For a very long time before his death he (Ibidem, pp. 213 – 214) remained

Almost inaccessible to economists and statisticians outside Russia […] *His assistance, or at least personal contacts with him would have been invaluable*.

2. Withdrawal from economics

In 1927, N. D. Kondratiev, the Director of the Conjuncture Institute, published a critical article concerning the first Five-Year-Plan. Soon he was elbowed out of science, arrested (1931) and then (1939!) shot [10].

N. S. Chetverikov, Kondratiev's assistant, served four years in prison, and, in 1937 or 1938, was subjected to new *repressive measures* [3]. Slutsky apparently had not suffered[4], but the general situation in statistics became unbearable. Later Chetverikov [4, p. 270] warily remarked that in 1930

The Conjuncture Institute ceased to exist and the Central Statistical Directorate underwent radical change.

I myself add that, also in 1930, the leading statistical journal, *Vestnik Statistiki*, was closed down and only reappeared in 1948[5]; during that period only a meagre number of statistical papers had been published in *Planovoe Khoziastvo*.

Under the changed social conditions, Maria Smit (more correctly, Falkner-Smit), a statistician of the new wave, became especially useful in spite of her crass ignorance (and in 1939 she was even elected Corresponding Member of the Soviet Academy of Sciences!). Pearson, she [30, p. 228] wrote,

Does not want to subdue the real world by a single curve [of distribution] *as ferociously as it was attempted by Gaus* [Gauss] […] *His system* [of curves] *nevertheless only rests on a mathematical foundation, and the real world cannot be studied on this basis at all*.

She [28, p. 168] also declared that Marxist statisticians should help the state security service in exposing the *saboteurs*. Iastremsky (Ibidem, p. 153) effectively agreed and mentioned D. F. Egorov (who died soon afterwards in his exile in Kazan):

I had recently an occasion to hear out […] *the speech of Prof. Egorov, the then not yet exposed saboteur*[6]. *He came out with a programme of sorts saying so ardently, even with a cry in his voice*, *What are you harping here on sabotage*? […] *There are no saboteurs worse than you yourselves, comrades, since you standardize reasoning by popularizing Marxism*.

Also see [22] and [23].

3. Archival sources

Before adducing the promised letters I list similar and already published archival materials concerning Slutsky[7].

1) In three of his letters to Chuprov, Markov, in 1912 [13, pp. 53 – 58] criticized Slutsky's book [24]. In the same source (p. 143) the Editor, in his review of the Markov – Chuprov correspondence, quoted a passage from a letter written by Slutsky to Markov. I translated and published this letter in full [22, pp. 45 – 46].

2) I myself [22] made known a few other archival or hardly known materials:

a) Chuprov's review of Slutsky [24] published in 1912 in a newspaper (pp. 62 – 63).

b) Slutsky's scientific character written by Chuprov in 1916 (pp. 67 – 68).

c) Passages from the correspondence of these scholars with each other (pp. 63 – 67).

3) Seneta [21] published English translations of two of Slutsky's letters to his wife concerning the author's appraisal of the comparative contribution of Borel and Cantelli to the discovery of the strong law of large numbers[8].

3.1. D. A. Grave – A. A. Markov, 4 Nov. 1912, Kiev

Archive of the Soviet Academy of Sciences, Fond 173, Delo 5, No. 1

Highly respected Andrei Andreevich, – I got to know E. E. Slutsky under the following circumstances. I was invited to a sitting of the Society of Economists at K. Comm. [Kiev Commercial] Inst. to attend a report on applying the Pearson theory to statistics. The report was delivered by Slutsky, a young man who had recently graduated from the [Kiev] University with a gold medal awarded for a work on political economy, but, because of some reasons, was not left at the University [to prepare himself for professorship].

I inquired directly of Slutsky's professor of political economy the reasons for this, and his answer surprised my by the justification unusual for a mathematical ear. According to his words, Slutsky is quite a talented and serious scientist, but the professor had not ventured to nominate him for being left at the University because of his distinct sympathy with social-democratic theories. And when I was unable to refrain from stating that at the mathematical faculty the author is not usually asked about his political views, the professor advised me to leave Slutsky at the mathematical faculty. I was naturally obliged to say that I have absolutely no desire to intervene in the business of the law faculty and that I am therefore asking him to leave the mathematical faculty alone. After this encounter Slutsky became my student and protégé. Although I am not at all acquainted with his works and had not understood the mathematical part of his report.

The lawyers, professors at the K. Comm. Inst., who did not understand Slutsky's book [24] but desired to acquaint themselves with the Pearson theory, have asked me to explicate it properly in my course in insurance mathematics [6]. I do not know how to find a way out of this difficult situation: it is simply repulsive to read all this …

[The sequel has no bearing either on Slutsky or probability and/or statistics. As also below, I myself inserted or specified the bibliographic information provided. For Grave, it was *repulsive* to read Pearson; cf. the now published letter of Slutsky to Markov (below).]

3.2. The Extant part of the unsigned and unaddressed letter (obviously, from Slutsky to Markov; no date)

Same Archive, Fond 173, delo 18, No. 5

are not independent in magnitude from the sum of the already accumulated deviations or that the probabilities of equal deviations are not constant, we shall indeed arrive at the formula

$$(1/y)dy/dx = x/F(x).$$

In an infinite number of cases (*naturally, not always*!) F can be expanded into a Taylor series, and the first few (e.g., three) terms will ensure a sufficient approximation. These qualification remarks should have certainly been made.

Only experience can show how often do empirical polygons of distribution, which can with a sufficient approximation be interpolated by a Pearson curve, appear in practice. Much material is already collected for answering this question in the positive. In many cases the

Gauss curve will not do since asymmetric polygons are often encountered in practice. Interpolation by parabolic curves

$$y = a_0 + a_1x + a_2x^2 + \dots$$

is unsuitable since these curves do not give an adequate picture at the edges of the figure: it is impossible to ensure their asymptotic approximation to the X axis; in addition, they lead to many superfluous inflexions.

The Pearson curves constitute the type that occurred to be practically the most suitable. Since the Gauss curve in very many cases is well suited for representing statistical facts, especially in anthropology [anthropometry], it seems desirable also for the asymmetric Pearson curves not only to indicate that they are corroborated by practice, but in addition to provide a theoretical derivation that would put this curve [these curves] in the same line as the Gauss curve on the basis of the theory of probability (hyper-geometric series).

The derivation on pp. 16 – 17 only serves to make the striking practical suitability of these curves less incomprehensible by means of the hypothesis on the action of infinitely many causes combining semi-randomly one with another.

2) The method of moments. Here, I allow myself to remark that neither Pearson, nor Lakhtin [9] say that they proved that the method of moments brings

$$\int (y - Y)^2 dx$$

to its minimal value. They only prove that the method ensures an approximation. It would have been interesting to investigate this problem and to indicate precisely when the method of moments is applicable, and when it is not. Lakhtin does it, but is he not mistaken?

I think that, *quand meme*, approximate formulas should not be objected to. Indeed, you yourself [11, p. iv] admit that such formulas might be used in probability theory even *without estimating their error* since *the aims of applied mathematics* demand this. You also state that approximate formulas should in addition be created for ensuring the calculations [12, p. 77][9]. At the same time, the method of moments is very convenient; and, since it is proved to provide an approximation for a large number of types of functions, its critical investigation is desirable. In many cases it is simply indispensable since the method of least squares sometimes leads to intolerable or even unrealizable calculations. If desired, I shall next time illustrate this proposition.

3) The theory of correlation. Here, I shall allow myself for the time being … [10]

3.3. Slutsky's Letters to Karl Pearson

I [22, pp. 65 – 66] published Slutsky's letter of 31 March 1913 to Chuprov. It occurred that Slutsky sent Pearson two manuscripts for publication in *Biometrika*. Pearson had, however, returned both of them, and Slutsky, considering that he was treated improperly, asked Chuprov's advice. Chuprov recommended that Slutsky submit his work to the Royal Statistical Society, and one of these manuscripts was indeed published there [25]; the other one, on a modification of the difference method, had not appeared anywhere.

Now, I am able to make known three letters from Slutsky to Pearson[11]; Pearson's letters are lost. Slutsky invariably gave his address as the Volodkevich Commercial *Schoole* in Kiev. Volodkevich was the name of his (future?) wife, and I am sure that since 1917 Slutsky never mentioned this private enterprise.

3.3.1. Slutsky – Pearson, 23 April 1912

University College London, Library, Pearson Papers 856/4

Dear Sir, – I am sending for your approval a paper concerning a correction to be made in the theory of contingency. If you find no fallacy in chief results, will not the paper be of some interest to the readers of the *Biometrica*? [!] Should you find any fault making idle the whole of my reasoning, I hope you will not refuse to communicate me your kindly criticism. It is a pleasure to acknowledge beforehand my great debt to you for the slightest of hints on the fallacies possibly made in my work. I am,

Yours faithfully E. Slutsky

P.S. The summary of the results is to be found at the end of the paper.

3.3.2. Slutsky – Pearson, 6 May 1912

Kept at the same place, 856/7

Dear Sir, – I had the pleasure to receive your honored letter on the 3rd May and I must excuse myself for answering so late – the reason is that I wanted much time for translating my letter in English. I thank you very much for your long and very interesting letter and for the proof which I am sorry not to have got yet, probably because it must be censured before I get it. Being you really very thankfull for your suggestiv and very valuable criticism and agreeing with you in many points, I fear nevertheless that I shall not be able to agree with you about their bearing concerning my main thesis. I think I can keep my ancient opinion about the best method of determining the probability we have in view, though after your letter I feel compelled to change its foundation. I take the liberty to begin with some general considerations and then I shall continue with the question in which we disagree.

1. There is not a single method for the determination of the probability that a given system of frequencies has arisen from random sampling.

A) The theoretical frequencies being known à priori, we can determine the probability of the given system of errors:

B) The theoretical frequencies being known à priori, we can determine the probability of the given system of errors:

$e_1 = m_1 - \mu_1$, $e_2 = m_2 - \mu_2$, … $P = Q(\chi_\mu^2; n') -$

in the notation of my paper – where n' is the number of groups,

$$\chi_\mu^2 = \Sigma[{}_\mu R_{ii} e_i^2 / {}_\mu R\, \sigma^2_{\mu_i}] + 2\Sigma[{}_\mu R_{ij} e_i e_j / {}_\mu R\, \sigma_{\mu_i}\sigma_{\mu_j}],$$

$$R = \begin{vmatrix} 1 & r_{\mu_1\mu_2} & r_{\mu_1\mu_3} & \dots \\ r_{\mu_2\mu_1} & 1 & r_{\mu_2\mu_3} & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix}. \qquad (1)$$

Now it is to be remarked that the method, even when applied to the same material, gives us very different results, the value of n' being arbitrary. As you have shown (Pearson 1900, p. 160), by infinitesimal grouping $P = 1$ for any value of χ^2 will appear. There is thus a number of groups n'_m which brings the value of P to the minimum, and I think you will agree that this minimal value of P is that really significant for the probability in question. "Really significant" means but this: we cannot assume a value greater than this P_{minim} to the probability that the given system of frequencies has arisen by random sampling from the supposed theoretical population.

C) Let

$$\theta_1 = f_1(m_1; m_2; \dots; m_n),\ \theta_2 = f_2(m_1; m_2; \dots),\ \theta_q = f_q(m_1; m_2; \dots)$$

be functions of empirical frequencies such that

$$f_1(\mu_1; \mu_2; \dots) = 0,\ f_2(\mu_1; \mu_2; \dots) = 0,\ \dots,\ f_q(\mu_1; \mu_2; \dots) = 0$$

and let $\sigma_{\theta_1}, \sigma_{\theta_2}, \dots, r_{\theta_i\theta_j}, \dots$ be their standard deviations and correlations. Then the probability of our frequency distribution being a random sample of the theoretical population $(\mu_1; \mu_2; \dots; \mu_n)$ can be judged

α) From the probability of the deviation of any θ_i from its zero value. In this case

$$P = \sqrt{\frac{2}{\pi}} \int_{\theta_i}^{\infty} \exp\left(-\frac{\theta_i^2}{2\sigma^2_{\theta_i}}\right) d\theta_i\,.$$

β) From the probability of the set of deviations from their zero values of a correlated system of functions $\theta_1; \theta_2; \dots; \theta_q$

$$P = Q(\chi^2_{\theta_1;\,\theta_2;\,\dots;\,\theta_q}; q+1)$$

where q is the number of independent values $(\theta_1; \theta_2; \dots; \theta_q)$,

$$\chi^2_{\theta_1;\,\theta_2;\,\dots;\,\theta_q} = \sum[{}_\theta R_{ii}\theta_i^2 / {}_\theta R\sigma^2_{\theta_i}] + 2\sum[{}_\theta R_{ij}\theta_i\theta_j / {}_\theta R\sigma_{\theta_i\theta_j}], \qquad (2)$$

and R is the same as (1) but with θ_i replacing μ_i.

The question of the relations between the results obtained by different methods seems to me to be a very difficult one. I think, however, that the following propositions hardly can meet objections.

Proposition 1. From all the values $\chi_1, \chi_2, \ldots, \chi_s$ that is really significant which gives the least value for P. For ex. (Pearson 1902, p. 280 & 283 – 284): In the case (1 – 3) – Motion of bright Line – the probability of the frequency distribution being a random sample from the general population distributed normally equals 1/23 if judged from the value of the criterion χ^2 and it is < 1/1000 if the probable error of the skewness will be taken into account.

Proposition 2. Should we take a great number of random samples from the general population and evaluate all values

$$\chi^2 \text{ with indices } \mu, \theta_1, \theta_2, \ldots, \theta_q, \theta_i\theta_j, \theta_i\theta_j\theta_k, \ldots, \theta_i\theta_j \ldots \theta_s, \ldots$$

for each random sample, the distribution of each χ^2 must be that indicated by the theory within the errors of random sampling.

Proposition 3. Let us have χ_1^2 (for n_1 independent values $\theta_i; \theta_j; \ldots; \theta_k$) and χ_2^2 (for n_2 independent values θ with other indices) and let n_1 not be equal to n_2. Then it is impossible that for all random samples $\chi_1^2 = \chi_2^2 = \chi^2$ say. Indeed, the theoretical distribution of χ_1^2 as given by $Q(\chi^2; n_1 + 1)$ differs from the theoretical distribution of χ_2^2 as given by $Q(\chi^2; n_2 + 1)$ whereas χ_1 being identical with χ_2 their distributions must and will be also identical.

2. I come now to consideration of the point of our divergence and I confess that *"*if I writte

$${}_1e_p = {}_1f_p - N({}_1f_p + {}_2f_p)/(N' + N'')$$

I vary the constitution of the general population for each pair of samples I take, whereas it must really be a constant, as we take all pairs of samples.

For consequence χ^2 proposed by me as the criterion of divergency cannot be regarded as your criterion for goodness of fit as worked out in your paper (Pearson 1900, pp. 160 – 163). In the notation of this letter it is not χ^2_μ. But nevertheless it is significant. Let us have a contingency table [Table 1] and let us look upon the values like

$$m_{ij} - N_i'N_j''/N = \varepsilon_{ij}$$

as on the functions of the group frequencies, varying from sample to sample, and becoming all zeros for the general population. Then my criterion of divergency χ^2_ε [Slutsky wrote out the right side of (2) with ε replacing θ]; the corresponding value of

$$P = Q[\chi^2_\varepsilon; (s - 1)(t - 1) + 1]$$

measures the probability *"*that a given system of deviations from the probable ($\varepsilon_{ij} = 0$) in the case of a correlated system of variables (ε_{ij}) is such that it can be reasonably supposed to have arisen from random sampling*"*. It is quite analogous with my $\chi^2_{\theta_1; \theta_2; \ldots; \theta_q}$ and it is easely to be subsumed under your general theory in Pearson (1900, p. 157 – 160).

Let us suppose there is no correlation in the general population and let a great number of random samples be taken from it. Then the distribution of values of χ^2_ε will be that given by

$Q[\chi^2_\varepsilon; (s-1)(t-1)+1]$.

I have shown in my paper that my criterion of divergency (χ^2_ε) for a fourfold table is identical as to its numerical value with your square continugency χ^2_μ. If so both theories cannot be valid as it is shown in the proposition 3 above.

I am not able now to see any error in my reasoning and it seems me the divergence in our views resolves as follows: We do not know the theoretical frequencies and we use *"*the best available values*",* i. e. $N_i'N_j''/N$ as it occurs in many other cases.

(A) I think that they are not the best, and it seems to me you will agree that we should obtain far better values if we have had a theory of skew surfaces. Then fitting such a surface to the system of values like $N_i'N_j''/N$ and integrating its volume for the base elements of the subgroups we have had indeed the <u>best</u> available values.

(B) Yet supposed the values like $N'N''/N$ be *"*the best available*",* there is still no ground that they are sufficiently <u>good</u>, for we can safely use the theoretical values deduced from the sample itself instead of the unknown quantities relating to the general population <u>only</u> if their probable errors are sufficiently small. That is the case with the standard deviation, when used to determine the probable error of the mean. In determining the goodness of fit we bring into the comparison the empirical frequencies with the theoretical ones deduced from the sample itself. But in using the method of moments for fitting the curves we reduce largely the probable errors of the theoretical group frequencies so that they become small as compared with the empirical frequencies.

For Ex. the frequency in Gaussian distribution, the base element being h, is $\mu_x \approx yh$ whence $\sigma_\mu/\mu = \sigma_y/y$. But in this case

$$\delta y/y(x^2/\sigma^3)\delta\sigma, \text{ so that } \sigma_y/y = (x^2/\sigma^2)\sqrt{2N}.$$

For the empirical frequency m_x we have

$$\sigma_m = \sqrt{m[1-(m/N)]},\ \sigma_\mu/\mu = \sqrt{(1/m)-(1/N)} \approx 1/\sqrt{m}.$$

Let $x = (1/2)\sigma$, $h = (1/8)\sigma$, $N = 450$, $m = \mu$. *Then* $\sigma_\mu/\mu = \sigma_y/y = 0.008$ and $\sigma_m/m = 0.224$ exceeding by 28 times the former value of procentual error of theoretical frequency. Let us take now a fourfold table [Table 2] and suppose the values a, b, c, d be proportional to the values in the general population. Let $a' = (a+b)(a+(1)c)/N$. Then

$$\sigma_a = [a(1-(a/N))]^{1/2},$$

$$\sigma_{a'} = (1/N)[(a+c)^2\sigma^2_{a+b} + (a+b)^2\sigma^2_{a+c} +$$
$$2(a+b)(a+c)\,\sigma_{a+b}\,\sigma_{a+c}\,r_{a+b,\,a+c}]^{1/2}$$

where

$$\sigma^2_{a+b} = (a+b)[1-(a+b)/N],$$
$$\sigma_{a+b}\sigma_{a+c}r_{a+b,a+c} = a-(a+b)(a+c)/N.$$

For $a = b = c = d = 12,\ \sigma_a = 3,\ \sigma^2_{a+b} = \sigma^2_{a+c} = 12,\ r_{a+b,a+c} = 0,\ \sigma_{a'} = 2.45.$

For $a = b = c = d = 12,\ \sigma_a = 30,\ \sigma_{a'} = 24.5.$

Thus, taking for the theoretical frequency $(a + b)(a + c)/N$ as determined by any random sample and dealing with every possible random sample we shall have our errors measured from the point the position of which is subject to errors of random sampling almost so great as the values we are measuring thereof. In consequence we shall obtain the values of χ^2 on the average largely reduced as compared with the case we knew the à priori frequencies in the general population. In my paper are given the values of $\chi^2{}_\varepsilon$ evaluated for random samples obtained by the experiment. The values of e which correspond to the ε in the notation of this letter were measured from the theoretical frequencies deduced from the data. If we measure them from the frequencies known in my case à priori: $a = b = c = d = 12$, we obtain, as a matter of fact, much greater values (given in the table here apart). If we use the same grouping as before we obtain [Table 3].

This seems to me to confirm my views that your theory is to be applied in the cases where we know the à priori frequencies but that in the cases we do not know them your $\chi^2{}_\mu$ must be replaced by my $\chi^2{}_\varepsilon$ which is numerically identical with it, so that the whole difference in the results touches only the value of n' being in the case we use $\chi^2{}_\varepsilon$, $(s - 1)(t - 1) + 1$.

3.3.3. Slutsky – Pearson, 18 May 1912

Kept at the same place, 856/4

Dear Sir, – I take the liberty to write you again, before I have your answer on my previous letter. I am printing now a treatise (or a text-book) on the theory of correlation and I would be very gratefull to you if you let me know whether the probable error of the partial correlation coefficient can be reduced to the same form as the probable error of the total one, as m-r Yule says[12].

I have also brought fast [replace this German word by the proper English *almost* – O. S.] to the end a paper on a General test for Goodness of Fit of the Regression Curves. To keep your valuable time I do not send it to you and I take the liberty only to communicate you an idea of it you will easily appreciate. It is very simple but I am not able to refer to any previous mention of it.

In the notation of your memoir on Skew correlation [16] the criterion will be simply […]

If you will agree with this I can send you a more elaborate – but still a short paper – with the illustrations taken from your memoir on skew correlation [16].

I excuse myself, dear sir, for my very imperfect English and for the trouble I give you and remain very faithfully yours E. Slutsky

3.3.4. Slutsky's Letter to Aleksandr Nikolaevich Shchukarev, a specialist in physical chemistry

Archive of the Moscow State Univ., Fond 276, Inventory 1, No. 114

Slutsky made known his opinion about Shchukarev's unnamed paper, perhaps answering the latter's request. This paper [19], which I located without much difficulty, was written extremely carelessly. In essence, Shchukarev vainly attempted to derive the Maxwellian law without introducing any stochastic ideas and it is therefore sufficient to say only a few words about Slutsky's reply.

Slutsky indicated that Shchukarev had not nevertheless managed without stochastic considerations; admitted (perhaps too modestly) that he *hardly understands* physics but *somewhat grasps* the logical structure of *suchlike theories*; and offered concrete remarks (unnecessarily since the paper was beyond repair).

Notes

1. Short anonymous and barely differing articles on Slutsky are included in the 2nd and 3rd editions of the *Bolshaia Sovetskaia Enziklopedia*; the 3rd edition is available in an English translation (entitled: *Great Sov. Enc.*). My references do not at all exhaust the literature on him. Sarymsakov [18] praised his work in geophysics, and the authors of several sections of [31] described his mathematical achievements. Romanovsky [17] indicated that Slutsky was chairman of a commission on applying statistical methods in industry (as a young man he studied for a few years at the machine-building department of the Munich polytechnic school [4, p. 262]). It seems, however, that because of the negative attitude of the Soviet establishment towards statistics in general (§ 2) that commission was unable to be of essential use.

2. For a background to this section see [23].

3. Slutsky had been applying these discoveries mostly to economics, and his transition to other branches of knowledge was painful: disallowing a report that appeared in 1932 but was delivered by Slutsky in 1928, he had not published anything during 1930 – 1932. I also note that an English translation of his paper of 1927 was published in 1937. It found important application in investigating time series in economics [1, pp. 209 – 210].

4. In 1990 the eminent mathematician Konüs told me that at the time he had also worked at the Conjuncture Institute. He was left alone; as Konüs explained the attitude of those responsible for the decision-making, they had decided: *He is only a mathematician, not responsible for anything*.

5. In 1929 a paper by the mathematician and statistician N. V. Smirnov appeared in the *Vestnik* and Slutsky even before his move to Moscow had published four articles there.

6. Smit [29, p. 4] clumsily declared that *The crowds of arrested saboteurs are full of statisticians*. Anderson, a student of Chuprov, testified [2, p. 294]:

Könnte ich […] *eine ganze Reihe von in Russland früher sehr geschätzten Statistikern und viel versprechenden jüngeren Schülern* […] *Tschuprows aufzählen, deren Namen nach 1930 aus der sowjet-russischen wissenschaftlichen Literatur plötzlich ganz verschwanden*.

7. I also stress that Chetverikov [4] mentions and quotes Slutsky's biography written by his wife, Yu. N. Volodkevich (p. 265), as well as another biography written by Slutsky himself (pp. 267 and 271). In turn, Gnedenko [5, p. 6] quotes Slutsky's autobiography compiled in 1938. They do not provide any information about these sources. Recall (§ 1) that in 1939 Slutsky started working at the Steklov Mathematical Institute.

8. Also see [4, p. 269]. In 1970 Chetverikov had given me (Russian) typed texts of these letters which I turned over to Seneta (their copies are regrettably lost). Seneta acknowledged my help in obtaining *important materials* but had not elaborated. He was concerned that I could have had problems with the Soviet authorities.

9. Slutsky obviously referred not to the paper itself as put out in the *Matematich. Sbornik*, but to its previously published offprint. Indeed, he mentioned the year 1911 and p. 4 neither of which agree with the periodical. The appropriate page numbers in the translation (see References) are 77 and 78.

10. Slutsky discusses the Pearson curves. At the time (and even in 1928, in his letter to Shchukarev, see § 3.3.4, which I only describe but do not quote) he sometimes wrote *theory of probability* instead of the correct Russian … *of probabilities*.

Slutsky derived the equation (see beginning of letter) in his book [26], see p. 25 of its translation.. Also there (pp. 22 – 24 he obtained the normal distribution as the limiting law for the binomial distribution. Assume the unknown law (Y) as, for example, a polynomial of the n^{th} degree, then, in principle, its ($n + 1$) parameters can be determined given the appropriate moments. If the class to which Y belongs is not restricted, its unique determination is impossible even if *all* the moments are given. Slutsky's question apparently touched on this problem of moments.

11. For some reason the pressmarks of two of the letters are identical. Both Pearson and Slutsky (who studied for as few years in Germany) mastered German, but Slutsky needlessly wrote in his very imperfect English.

12. Slutsky's reference is *J. Roy. Stat. Soc.*, 1907, pp. 6 and 47. In both these cases Yule was a participant in discussing the contributions of other authors. The paper that Slutsky mentions just below is apparently [25].

References

MSb = *Matematicheskiy Sbornik*

M, L, Psb, R = Moscow, Leningrad, Petersburg, in Russian

1. Allen, R. G. D. The work of Eugen Slutsky. *Econometrica*, vol. 18, 1950, pp. 209 – 216.

2. Anderson, O. Mathematik für marxistisch- leninistische Volkswirte. *Jahrb. f. Nationalökon. u. Statistik*, Bd. 171, 1959, pp. 293 – 299.

3. Anonymous, Anniversaries and memorable dates. *Voprosy Statistiki,* No. 11, 1995, p. 77. (R)

4. Chetverikov, N. S. Life and scientific work of Slutsky (1959). In author's book *Statisticheskie Issledovaniya* (Stat. Investigations). M., 1975, pp. 261 – 281.
40.

5. Gnedenko, B. V. Slutsky [26, pp. 5 – 11]. (R)

6. Grave, D. A. *Matematika strakhovogo dela* (Insurance mathematics). *Izvestia Kievsk. Kommerch. Inst*., book 16, 1912, pp. i – iv + 1 – 88 of second paging. (R)

7. Kolmogorov, A. N. Slutsky. Obituary (1948, in Russian). *Math. Scientist*, vol. 27, 2002, pp. 67 – 74.

8. Konüs, A. A. Slutsky. *Intern. Stat. Enc*., vol. 2. Editors, W. H. Kruskal, Judith M. Tanur. New York – London, 1978, pp. 1000 – 1001.

9. Lakhtin, L. K. On the Pearson method etc. MSb, vol. 24, 1904, pp. 481 – 500. (R)

10. Makasheva, N. N. D. Kondratiev. Brief biogr. essay. *Mirovaia Ekonomika i Mezhdunarodn. Otnoshenia*, No. 9, 1988, pp. 59 – 61. (R)

11. Markov, A. A. *Ischislenie veroyatnostei* (Calculus of Probability). Psb, 1908.

12. Markov, A. A. A rebuke to P. A. Nekrasov. MSb, vol. 28, 1912, pp. 215 – 227. **S, G,** 4.

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

14. Pearson, K. On a criterion that a given system of deviations from the probable etc. *Phil. Mag*., vol. 50, 1900, pp. 157 – 175.

15. Pearson, K. On the mathematical theory of errors of judgement etc. *Phil. Trans. Roy. Soc.*, vol. A198, 1902, pp. 235 – 299.

16. Pearson, K. *On the general theory of skew correlation* etc. *Drapers' Co. Res. Mem., Biometric Ser*., 2, 1905.

17. Romanovsky, V. I. On the application of math. statistics and the theory of probability in the industries of the Soviet Union. *J. Amer. Stat. Assoc.*, vol. 30, 1935, pp. 709 – 710.

18. Sarymsakov, T. A. Statistical methods and problems in geophysics. *Vtoroe vsesoyuznoe soveshchanie po matematich. statistike* (Second All-Union Conf. on Math. Statistics). Tashkent, 1948, pp. 221 – 239. (R)

19. Schükarev, A. N. Ein Versuch der Ableitung des Maxwellischen Verteilungsgesetzes. *Phys. Z.*, Bd. 29, No. 6, 1928, pp. 181 – 182.

20. Seneta, E. Slutsky. *Enc. Stat. Sci.*, vol. 8. Editors, S. Kotz, N. L. Johnson. New York, 1988, pp. 512 – 515.
21. Seneta, E. On the history of the strong law of large numbers etc. *Hist. Math.*, vol. 19, 1992, pp. 24 – 39.
22. Sheynin, O. *Chuprov. Life, Work, Correspondence.* (1990, in Russian). Göttingen, 1996, 2011. References in text are to the edition of 2011.
23. Sheynin, O. Statistics in the Soviet epoch. *Jahrb. f. Nationalökon. u. Statistik*, Bd. 217, 1998, pp. 529 – 549.
24. Slutsky, E. E. *Teoriya korrelyatsii* (Theory of Correlation). Kiev, 1912. **S, G,** 38.
25. Slutsky, E. E. On the criterion of goodness of fit of the regression lines etc. *J. Roy. Stat. Soc.*, vol. 77, 1914, pp. 78 – 84.
26. Slutsky, E. E. *Izbrannye trudy* (Sel. Works). M., 1960.
27. Smirnov, N. V. Slutsky. *Izvestia Akad. Nauk SSSR*, ser. matematich., vol. 12, 1948, pp. 417 – 420. (R)
28. Smit, M. Planned sabotage and the statistical theory. *Planovoe Khoziastvo*, No. 10, 1930, pp. 139 – 168. (R) Incorporates several reports including that of B. S. Iastremsky.
29. Smit, M. *Teoriya i praktika sovetskoy statistiki* (Theory and Practice of Soviet Statistics), 2nd edition. M. – L., 1931.
30. Smit, M. Against the idealistic and mechanistic theories in the theory of Soviet statistics. *Planovoe Khoziastvo*, No. 7, 1934, pp. 217 – 231. (R)
31. Stokalo, I. Z., Editor, *Istoriya otechestvennoi matematiki* (History of National Mathematics), vol. 4, pt. 2. Kiev, 1970.
32. Youshkevich, A. A. Slutsky. *Dict. Scient. Biogr.*, vol. 12. Editor, C. C. Gillispie. New York, 1975, p. 461.

E. E. Slutsky

On the existence of connection
between the solar constant and the temperature

Zhurnal Geofiziki, vol. 3, 1933, pp. 263 – 281

Summary [in its original English]
Same title

Abbreviation: CC = correlation coefficient
MT = max. temperature
SC = solar constant

1 – 3. The daily Montezuma values of the SC which have been used here, were obtained by the critical examination of the following data:

1) The values found by measuring the ordinates on the enlarged photo-copy of the C. G. Abbot's diagram in Smiths. Misc. Coll., Publ. 3114, p. 2 – 3, covering the period 1924 – 1930;

2) Ten day SC values for the same period (l. c., p. 12);

3) The daily values of the SC published in the Daily Weather Map of the United States Weather Bureau for the period from 24 July 1927 till 31 Oct. 1931.

The errors found by the comparison of our values with the Annals of the Astro-Phys. Obs. of the Smiths. Inst. (vol. 5), which came to us when this study was rather finished, are given in Table 1. Only on one case they are to be imputed to the misreading of the Abbot's diagram, in ten cases to errors in the Daily Weather Map, in the remaining 65 cases to the errors which are to be found in the Abbot's diagram republished now without alteration in the Annals (vol. 5, p. 246). The mean frequency of the errors being less than 1:30 and their influence being found quite negligible in the one of the most doubtfull cases, it is to be hoped that the results of this study cannot be substantially vitiated by the said errors.

4. As we intended to prove the existence of the correlations between the SC and the MT found by H. H. Clayton, the deviations of the ten-daily means from the thirty-daily means of the SC and the MT for Cordoba (Argentina) have been computed. Then we have found the st. d. for every three-months period of the each year and the analogous st. d. based on the data for the whole period 1924 – 1931.

A glance on the Fig. 1 tells us that these st. d. are to be considered as periodic time-functions. Having calculated 3 (resp. 2) harmonics (see the full lines, Fig. 1), we reduced them by the due factors. The momentanious st. d. having thus been found, the original deviations were standardized by dividing them with the values proportional to these standard deviations.

5. From the series of the MT thus obtained we have chosen the partial series corresponding to the 56-th till 155-th and to the 156-th to the 255-th day of each year and we have thus correlated them with the SC values 1) for the same year and 2) for the two preceeding resp. the two following years with the additional lags from 0 to 15 days (see Tables 3a & 3b). The all 16 correlational functions thus obtained for

the corresponding years and a specimen containing 4 functions for the different years are shown on the Fig. 2[1].

After the second partial series had been divided in the two equal parts, the same combinations of years have been considered and for each combination the largest from the CC corresponding to the additional lags from 0 to 15 days was found. They were found thus 8 + 8 + 16 CC between SC and MT values relating to the same year and 26 + 26 + 52 CC between the values relating to the different years, each CC being the largest (as to its absolute value) from the 16 CC corresponding to 16 additional lags from 0 to 15 days which were tried for each combination. These CC are shown on the Fig. 3. Thus, it is fairly evident that there is no signifiants difference between the CC found for the data relating to the same year and the CC found for the data relating to the different years, whence it follows that *the true CC between the values of the SC of radiation and the MT in Cordoba must be quite negligible the correlation c. which can be empiricaly found being nothing else but the errors of the random sampling.*

6. Table 4 gives the values of the momentanious st. d. of the deviations of the ten-daily from the 30-daily means of the SC calculated for the middle points of respective months. In discussing these values the author comes to the conclusion that during six months from the twelve the errors there involved constitute presumably the greater and the true value the lesser part of the values of the said deviations.

7. The distributions of the CC relating to cases when SC preceeds MT and to cases when MT preceeds SC (in both cases with the lags from 1 to 2 years) cannot be considered as significantly different, the value of chi-square being 16.84 and the corresponding probability P = 0.3. Combining both we find $\sigma_r = 0.2672$ instead of 0.1 given by the Pearson's formula (for $r = 0$, $n = 100$), this formula being inapplicable to connected series, i. e., to the series composed by the casual values which are not mutually independent.

Applying further the theory of the R. Fisher's function z for the connected series developed in our paper (J. Geophysics, vol. 2, No. 1(3)), we find $\sigma_z = 0.2873$ which leads to the theoretical distribution of the CC ($\chi^2 = 12.84$, $n' = 15$, $P = 0.5$). The values of z being thus normally distributed, it is possible to find, for instance, the probability of the deviation 0.65, this being the largest CC in the case of the correlation of values of the SC and MT relating to the same year. This probability being 0.007 the mathematical expectation of the number of such cases in the universe of 256 cases will be 1.8 the actual value, as a matter of fact, being only 1. The same theoretical distribution has been compared (see Table 6 and 7) with the distribution of the CC between the values of the SC and of the MT relating to the same year.

The distributions of the Table 6 being at the first sight significantly different, the author analyses the discrepancies and comes to the conclusion that there is probably no significant divergency, the discrepancies being enlarged by the correlation between the CC constituting the set of values under consideration. [See also the paper of the present author in the Journal of Geophysic vol. 2, No 1(3)]. This point of view is confirmed by the distribution of the absolute values of the CC (Table 7), and by the value of the standard deviation for the distribution of the Table 6 (0.250) being not substantially different from the value (0.267) of the st. d. of the CC for the case of different years.

8. There were found further 4·192 CC between the SC values with the lags equal, or nearly equal, to one and to two years and n = 40, 60, 80, 100. The empirical st. d. of these CC are shown in Table 8 where the last column gives the theoretical values according to the formula

$$\sigma_r^2 = \frac{A}{n} + \frac{B}{n^2} + \frac{C}{n^3},$$

A, B, C being found by the method of least squares. As we know (see the paper of the author cited above) the coefficient

$$A = 1 + 2\sum_1^{n-1} r^2(t)$$

whence it follows that

$$\sum_1^{n-1} r^2(t) = 0.5(A-1) = 4.14.$$

As it was found (see Table 9)

$$\sum_1^{31} r^2(t) = 3.45$$

it must be admitted that the values of $r(t)$ for $t > 31$ cannot be regarded as negligible. As it has been necessary to postpone the further study of the serial correlations, the theoretical value of the st. d. between the SC and the MT which (under the supposition of the zero-correlation) is given by

$$\sigma_r^2 = \frac{1}{n}[1 + 2\sum_1^{n-1} r_x(t) r_y(t)]$$

could not be evaluated. Nevertheless it is to be noted that the substitution of the sum

$$\sum_1^{31} r_x(t) r_y(t)$$

in the preceeding formula gives us the value of $\sigma_r = 0.30$ not substantially different from the value 0.27 found above by the direct computation based on 832 CC.

As it follows from the values of the serial correlations for the SC and for the MT given in the Table 9, there is a great similarity between the serial correlations for the periods 1924 – 1927 resp. 1928

– 1931, the relatively small differences being probably of the casual provenience. This fact cannot be underestimated and deserves further studies.

The main text

1. Introductory remarks. The subject of study

Abbreviation: see Summary

The solar constant is the amount of energy received [in 1 minute – not mentioned by Slutsky] from the sun by a surface perpendicular to the sun's rays, $1cm^2$ in area and located outside the earth's atmosphere at the earth's mean distance from the sun.

The remarkable investigations of Abbot and his collaborators (*Annals* 1932) have apparently definitively proved that this magnitude is actually not constant but fluctuates from year to year, from month to month, and perhaps even from day to day. Not so is it with the Clayton – Abbot (Abbot 1931, p. 1) theory of weather. It maintains that exactly those alterations in the intensity of the solar emanation constitute the most essential cause of all meteorological changes which in their totality compose that which is called weather.

The provided justification of that proposition does not seem convincing to us and we aim here to report about the work done for at least partly checking it. Clayton's study that went on year after year led him to conclude that each alteration in the SC produces changes of temperature, of the same sign in the equatorial and polar zones, and of the opposite sign in the temperate zones, and that first of all those alterations are reflected in the equatorial zone and in the high latitudes of the temperate zones. These perturbations move in waves towards the equator and shift eastward travelling at speeds inversely proportional to the length of their periods, and, in the tropical regions, are superimposed on the waves generated in the equatorial zone (Clayton 1923, pp. 215 – 269).

Clayton took into account a large number of stations, ensured a geographical coherence of the entire picture, and, last but not least, his separate, masterly selected illustrations are inspiring. At first, this creates an impression of reliable validity; only after having a closer look you begin to notice that the edifice of Clayton's constructions is not so robust.

First of all, it is necessary to remark that the number of stations indicating a correspondence between the course of meteorological processes and the changes in the SC cannot be especially significant. Since those processes are interconnected, such parallelism observed at one station will almost certainly be revealed in a number of other stations. It is much more important to cover the longest possible period and exactly in this respect Clayton's work leaves too much to be desired.

Clayton, to be truthful, determines CCs many times exceeding their mean square errors. Thus, at Sarmiento in Argentina after two days the CC between the SC and the temperature in winter of 1916 reached 0.82, see Clayton (1923, p. 224); on p. 269 he expressly mentions a small probable error. For 77 days of observation the CC elevenfold exceeded its mean square error.

These data would have provided a reliable guarantee had he issued from series consisting of mutually independent terms. But, when this condition is lacking, as it always does when dealing with wavy series, the usual formula for the mean square error becomes absolutely unsuitable (Slutsky 1929; 1933) and its application can lead to most deplorable blunders.

Indeed, Clayton compares series mostly representing deviations of moving decade averages from similar monthly averages. Suppose that rhythms of about the same length occur in the series of temperatures as well as in that of the SC. That resemblance will be strengthened by averaging and it is not difficult to imagine that when the series are suitably shifted one with respect to the other intervals of 3, 4 and 5 wavelengths will quite often provide sufficiently high CCs.

That, however, is just what Clayton does when he calculates those coefficients after each shift up to 15 days. Shoot the flight of a crow in Moscow and of another one in New York. Measure the ascent of the wing on each film and calculate the CC. If your series are not too long, after a suitable shift [of one film relative to the other one] you will likely find a high coefficient, but does it mean that the flights of those two crows were causally connected?

And so, we decided to restrict our investigation by considering one station, but to take into account the entire period covered by the data on the SC, i. e., the eight years from 1924 to 1931. It was necessary to establish whether Clayton's results pertaining to the country which he especially studied and for which they, the results, occurred most striking were corroborated[2].

2. The Data

When beginning our work, the *Annals* (1932) had not yet appeared whereas (Abbot, no reference provided) it was already known that a large part of the previously published values of the SC should now be considered dated because the methods [of measurement] had been since improved and a number of new corrections introduced. We could therefore only base our study on the following sources.

1. The diagram of the daily values of the SC at mount Montezuma in Chile for 1924 – 1930 (Abbot 1931).

2. The Table of the mean decade and monthly values of the SC (Abbot 1931, p. 12).

3. The *Daily Map* (no date) containing the same data on Montezuma for the period from 24 July 1927 to the end of October 1931[3].

Here is how we proceeded. The ordinates on a photo of the Abbot diagram (22.5•17.5 *cm*) enlarged 2.5 times were measured twice and all the doubtful cases thoroughly considered. A number of values of the SC was thus established. Abbot distinguished satisfactory, almost satisfactory and unsatisfactory data by differing symbols (S, S– and U) and we were therefore able to determine decade and monthly means in which he neglected those of the last-mentioned type. A comparison of our means with his was satisfactory; namely, for all eight years the decade means of the CC were 0.990 with fluctuations in separate years from 0.977 to 0.994 and the monthly means for all that period, 0.9998. Deviation of the former from the latter, 0.963 with fluctuations in separate years from 0.946 to 0.986.

We compiled the series of values of the SC selected for the further work in three parts: from Jan. 1924 to 23 June 1927 (obtained from the Abbot diagram), from 24 June 1927 to 31 Dec. 1930 (the data corrected by critical comparison with the *Daily Map*), and for 1931 (*Daily Map*, the only source here). The *Unsatisfactory* data were neglected.

For comparing the SC with MT, we selected the data pertaining to Cordoba (Carta del Tiempo) in Argentine[4]. They only had a few essential gaps (40 days in succession from 1 Jan. 1929, and 15 days both in Dec. 1928 and Dec. 1931); other gaps were not longer than two days in succession (in the mean, missing was a little less than one day monthly) and we decided that it was permissible to fill those [shorter] gaps by linear interpolation.

3. Comparison of our series of the solar constant with Abbot's final data

Those final data (*Annals* 1932, Table 31, Montezuma 1920 – 1930, pp. 195 – 213) only became available after we concluded our work. We may certainly ignore the deviations concerning the *Unsatisfactory* cases, the rejection of those cases or the change from gaps to *Unsatisfactory* or vice versa as well as the change from *Satisfactory* to *Almost Satisfactory*.

There were 76 deviations left (Table 1), 10 of them (with symbol W attached) based on the *Daily Map*, one of those caused by an unfortunate reading of the Diagram (28 Oct. 1927). The rest 65 cases, as we ought to state regretfully, were mistakes of the Diagram itself, reprinted without change in the *Annals* (1932, p. 246). Concerning their influence on the results, the number of mistakes can be thought unimportant.

The worst case concerns Jan. – March 1925 (16 mistakes). Ten of the other mistakes, each amounting to not more than 1 or 2 units [of the last digit] were absolutely insignificant; 51 that had occurred during 81 month are left, 1 mistake per 48 days, and they certainly cannot discredit our conclusions.

As to the worst case mentioned above, we made the necessary calculations anew. For 100 days of the comparison of MT with the SC (from the 56th to the 155th day of the year) we obtained the highest in absolute value CC of 0.39 for a shift of 10 days instead of 0.40 for a shift of 11 – 12 days according to the previous calculation. Thus, even for the worst case, the error turned out to be absolutely inessential.

4. The treatment of the series

For the sake of convenience we adopted the following artificial calendar (Table 2) considering that each year had 365 days. That assumption would not have been possible to make for a longer period, but for eight years the inaccuracy thus introduced may apparently be neglected.

We bear in mind the study of periods lasting 100 days: from the 56th to the 155th day and from the 156th to the 256th day of the year. The latter approximately corresponds to the period for which Clayton had considered the connection between SC and MT in Argentina, and we indicate the appropriate calendar dates in Table 2. For the calculations below, months were thought to be 30 days long except for December (35 days), and an artificial trick explained below was introduced for ensuring intervals of equal duration.

Following Clayton, we had to study the correlation between the decade and monthly mean deviations, so we began by calculating the appropriate series; the means were taken with respect to the fifth and the fifteenth days of the appropriate moving time intervals. For the MT, because of the filling of the random gaps in the data (§ 2), the number of consecutive terms was always the same (10 and 30); for the SC, we calculated the arithmetic mean for the data at hand in those decade and monthly intervals; following Abbot, we did not exclude cases in which even only one observation was available. The units adopted were 0.001 cal/cm^2 and 1°C and the means and the deviations were calculated to one decimal point.

The numbers in the first series were rounded to integral units; the same was done with those of the second series after multiplying them by 10/3. We then calculated the sums and the squares [of those numbers?] for the moving twelve three-monthly periods of each year (January – March etc.). The lacking data on the SC for Nov. and Dec. 1931 were filled up by the means calculated for the same months of the other years [of all other years?]; and, when calculating the sums for the first three months of 1924 and the last three months of 1931, we replaced Dec. 1923 by Dec. 1931 and Jan. 1932 by Jan. 1924. For each three months we denoted the square of the mean square deviation $\sigma_{3,ij}^2$ where i denoted the month, and j stood for the year.

Then, separately adding up the appropriate numbers of each month for all the years, we called the 12 numbers $\sigma_{3,i}^2$, $i = 1, 2, \ldots, 12$, which described the mean fluctuation of each three months for all the eight years. These numbers are shown on Fig. 1 by small circles, separately for SC and MT. There also, are the $\sigma_{3,ij}^2$ shown by points for each year.

Becoming thus convinced in the presence of a yearly course of fluctuations, we expanded each empirical function $\sigma_{3,i}^2$ in a Fourier series. It occurred that they can be satisfactorily represented by three (SC) or two (MT) first harmonics shown on Fig. 1 by continuous curves. Their parameters were (A_0 – arithmetic mean; A_i and B_i – coefficients of cosines and sines of harmonic i respectively):

SC: $A_0 = 11.191$, $A_1, A_2, A_3 = 2.015,\quad 0.947, 0.777$
$B_1, B_2, B_3 = 0.153,\quad 3.031, 0.239$
MT: $A_0 = 38.958$, $A_1, A_2 = \ -17.255,\ 4.830$
$B_1, B_2 = \ -2.800,\ -0.136$

For three-months periods the arithmetic means of SC and MT are very near to zero, and we will therefore insignificantly violate reality by replacing them below by expectations and by considering those latter equal to zero. And so, let there be m series of random variables

$x_{j1}, x_{j2}, \ldots, x_{jN}, j = 1, 2, \ldots, m,$

$$\mathrm{E}x_{jt} = 0,\ \mathrm{E}x_{jt}^2 = \sigma_t^2 = f(t).$$

Let

$$s_{2n}^2(t+1/2) = \frac{1}{2mn}\sum_{j=1}^{m}\sum_{k=-n+1}^{n} x_{j,t+k}^2$$

be the square of the mean square [literal translation] for the appropriate parts of all the series with centres at [t + (1/2)]. Then, obviously,

$$\mathrm{E}s_{2n}^2(t+1/2) = \frac{1}{2n}\sum_{k=-n+1}^{n} f(t+k) = F(t+1/2).$$

If 2*mn* is sufficiently large, then, according to the law of large numbers, the mean value will be approximately equal to its expectation. But in our case 2*mn* is indeed sufficiently large as can be supposed on the basis of the smooth course of the magnitudes $\sigma_{3,i}^2$ which, owing to their meaning, ought to coincide with $s_{2n}^2(t+1/2)$. Let us call $f(t) = \sigma_t^2$ the *instantaneous*, and $\sigma_{3,i}^2$, the *mean three-month variability*.

As proved above, we will have an approximate equality

$$\sigma_{3,i}^2 = \frac{1}{90}\sum_{k=-44}^{45}\sigma_{t+k}^2$$

where *t* is the fifteenth day of month *i*. Supposing that σ_t^2 is a sum of several sine curves, we recall a well known fact: $\sigma_{3,i}^2$ will then be equal to the sum of the same number of sine curves having the same periods and phases, but altered amplitudes. Knowing the coefficients of the harmonics for $\sigma_{3,i}^2$ and wishing to determine the coefficients of the harmonics comprising the instantaneous variability σ_t^2, it is only necessary to multiply them by

$$Q = \frac{2n\sin(h/2)}{\sin(hn)},$$

where, in our case, 2*n* = 90, *h* = 1, 2, 3 for harmonics 1, 2 and 3 respectively, Note that its 12 ordinates, when expanding $\sigma_{3,i}^2$ into a Fourier series, were treated as being equally spaced in spite of the 35-day long December. This means that December was squeezed into 30 days so that at that stage of our work a year consisted of 360 days. This is exactly why the abovementioned values of *h* were obtained.

Now, after calculating the coefficients of the expansion of σ_t^2, and shifting the origin of the system of coordinates from mid-January 15.5 days back, we multiplied the coefficients of the appropriate harmonics by 4 (for the SC) and divided them by 2.25 (for the MT). Here are their final values.

The solar constant

$a_0 = 44.764$, $a_1 = 8.445$, $a_2 = 4.710$, $a_3 = 4.819$
$b_1 = 3.048$, $b_2 = 19.387$, $b_3 = 9.704$

The maximal temperature

$a_0 = 17.315$, $a_1 = -7.839$, $a_2 = 2.939$
$b_1 = -3.608$, $b_2 = 1.655$

Now, calculating the appropriate sine curves for each day of the 360-day year, then increasing the days of December up to 35 by interpolation, we compiled a table of the values of $10/k\sigma_t$ with $k = 2$ and 2/3 for SC and MT. The deviations of the decade means from the monthly means (see above) were multiplied by those values and the results rounded off to integers. Thus we obtained final series of standardized deviations. The multipliers k were selected so that the absolute values of numbers in the final series will not exceed 21 or 22 which provided sufficient precision and essentially simplified further calculations.

5. Lack of correlated connection between the solar constant and the maximal temperature in Cordoba

That correlational connection was studied according to the following pattern. For MT, two intervals of 100 terms each were selected for each year, – from the 56th to the 155th and from the 156th to the 256th day, and two more of 50 terms each were obtained by dividing that second interval into halves. By comparing the SC with the MT of the day having the same number or a number less by 1, 2, … we were able to obtain CCs with differing "shifts". For the larger intervals CCs with shifts of 0, 1, 2, …, 15 were calculated, and for the shorter intervals, only the CCs maximal in absolute value among the same shifts. When determining these maxima, we were guided by the maximal values of the products, partly by superimposing graphs and we checked our work by calculating a few CCs around the supposed maxima.

As ascertained above, it was impossible to apply in our case the usual formula of the mean square error, but the use of the suitable theory encountered some difficulties (see below), so that we applied the following method. First, we calculated the CC between the values of SC and MT for the same year, i. e., by combining our series in pairs (1924, 1924), …, (1931, 1931). Second, we did the same for differing years, i. e. correlating MT of some year with the SC one or two years apart in either direction (Table 3) [call them combinations A and B].

The course of the CCs for combinations A and both large intervals is shown on Fig. 2. As an illustration, there also we show 4 correlation functions for the second interval and 4 combinations B. Our attention is at once arrested by the lack of any essential difference between combinations A and B. And it is also seen that even for the former combinations it is hardly possible to say that regularities are clearly discerned either in magnitude, sign or the shift corresponding to the maximal in absolute value CCs.

We now take a look at Fig. 3 where all the maximal in absolute value CCs are seen in a decreasing order; horizontal lines separate the larger and the lesser CCs and we clearly see that CCs of the same magnitude appear in both types of combinations and not rarer in the mean in group B. Thus, for the period between the 156th and the 255th day there are 8 [and 26] CCs in groups A and B; a half of those groups is not less than 0.49 and 0.36 respectively. However, we still ought to indicate that almost a quarter among group B reaches 0.49 whereas only 5 CCs from group A are higher than 0.39. It thus occurs that the difference only depends on one CC out of the eight which can well be a random occurrence.

Then, the insignificant superiority of group A in the series 156 – 255 is compensated by a superiority of B over A both in the interval 56 – 155 (the medians almost coincide, but considerably larger CCs are in group B) and in the shorter intervals (superior in both respects).

From all the above it follows that in Cordoba, if judging by the deviations of the decades from the monthly means, correlational connection between SC and MT either does not exist at all, or is quite insignificant and the comparatively high CCs are simply maximal values of random errors.

We will confirm this conclusion by another method (§ 7) whereas § 6 is devoted to a slight digression.

6. On the error of determining the solar constant

When calculating the instantaneous variability σ_t^2 for the middle of each month (see Table 4), we clearly see the magnitude of errors from which the determination of the SC was yet unable to get rid of. Represent the deviation of the mean decade from the mean monthly [values] x as the sum of the real deviation ξ and its error ε and denote the squares of their mean square deviations by σ^2, α^2 and β^2 respectively. For any two months we will have

$$\sigma_1^2 = \alpha^2 + \beta_1^2,\ \sigma_2^2 = \alpha^2 + \beta_2^2.$$

If $\sigma_1^2/\sigma_2^2 = p,$ then

$$\frac{\beta_2^2}{\alpha^2} = p - 1 + p\frac{\beta_1^2}{\alpha^2} \geq p - 1.$$

Comparing now all the months in Table 4 with November we find that for 6 months out of 12 $p \geq 2$. It follows that for these months not less than half of the magnitude of the deviations which we are studying are errors of observation. The deviations of the separate values from the monthly means are certainly corrupted by errors even more. It is hardly necessary to note that these conclusions, being a by-product of our work on which we cannot dwell anymore, should be specified by studying the probable errors of the numbers in Table 4.

7. The mean square error of the coefficient of correlation between the solar constant and maximal temperatures

When shifting the series of SC and MT with regard to each other by 1 or 2 years and some days, from 0 to 15, we obtained, as stated above, 832 CCs, each of them for the two series consisting of 100 terms. Separating them into two groups depending on whether the SC precedes MT (a) or vice versa (b), we obtain two distributions of the CCs (Table 5, columns a and b). For estimating the homogeneity /heterogeneity of those distributions, we can apply Pearson's formula; in our case it will be

$$\chi^2 = \sum \frac{(a_i - b_i)^2}{a_i + b_i}.$$

We obtain $\chi^2 = 16.84$; for $n' = 15$[5], we have $P = 0.3$ which shows a sufficient correspondence between those distributions. This circumstance confirms our assumption that in any case when the shift is 1 year or larger, the CCs between SC and MT vanishes, and the empirical CCs are nothing but "errors". Considering now both groups together (Table 5, column c), we calculate the mean square error of those CCs: $\sigma_r = 0.2672$. Had our series been lacking internal connections, such an error for ($r = 0$) would have taken place if the number of terms $n = 1/(0.2672)^2 = 14$. Or, the presence of such connections influences the square error and the number of terms is lessened from 100 to 14.

Supposing after Fisher that

$$z = \operatorname{arctanh} r = \frac{1}{2} \ln \frac{1+r}{1-r}$$

and, taking into account that in our case we may suppose that the real CC is zero, we find that $\sigma_z = 0.2873$[6]. Assuming that z is normally distributed, we calculate the theoretical numbers corresponding to the group in Table 5 (column m)[7]. If, as it is done after Pearson, the extreme groups having theoretical numbers less than 1 are combined with the neighbouring groups, we will have $n' = 15$

$$\chi^2 = \sum \frac{(m-c)^2}{m} = 12.84$$

and the probability $P = 0.5$ of a random deviation of the empirical distribution from the theoretical.

This fact is not devoid of interest since it again confirms my hypothesis formulated in the abovementioned contribution[8]. In addition, and it is here certainly more important, we become able to estimate the most considerable CCs which occur when comparing SC and MT for the same years. In Table 3 we see that out of 256 CCs of that group not a single one exceeds 0.65. And since

$z = \operatorname{arctanh} 0.65 = 0.7753,$

which exceeds the calculated $\sigma_z = 0.2873$ only by a factor of 2.7, it means that not a single CC out of those 256 deviates from zero by three mean square errors. At the same time, according to the tables of the integral of probability, the theoretical number of deviations $\geq 2.7\sigma$ is $256 \cdot 0.00693 = 1.78 > 1$.

These considerations, as it seems, decidedly confirm the conclusion which we reached by following quite another approach, i. e., that there are no grounds for believing that the CC between the SC and MT in Cordoba appreciably differs from 0.

We will now check this conclusion in yet a different way by comparing the distribution of 256 CCs of group A with the theoretical obtained by studying the 832 CCs for pairs of different years (Table 6). It is not necessary to calculate χ^2 here: we see at once that it ought to be very considerable and the corresponding probability, very low. We ought to recall, however, that, as I had discovered in the quoted above paper, the χ^2 test is suitable, strictly speaking, only for totalities comprised of independent elements. It can be applied to totalities of dependent magnitudes[9], if at all, only tentatively since an entirely adequate criterion is yet lacking.

It seems that dependence has a stronger influence when the number of terms is comparatively small which is well illustrated by Tables 5 and 6. Indeed, a close look at the latter rather sharply brings home that the deviation between the empirical and theoretical distributions occurs owing to the essential accumulation of few cases in which the smoothness of the empirical distribution is grossly corrupted in a way that always takes place exactly in distributions of an insufficient number of elements.

In our case it is easy to explain this. Table 5 consists of 832 CCs, 52 groups of 16 terms each (shifts from 0 to 15 days) whereas only 16 such groups are in Table 6. At the same time the CCs in each separate group between certain series of the SC and MTs provide a series of 16 terms corresponding to shifts of 0 – 15 days closely correlated with each other; this is indeed revealed by the smooth wavy course of the relevant series (Fig. 2).

Therefore, if the maximal range of such a wave is about 0.55, say [?], and the wave forms a smooth stretched peak, a few consecutive CCs will at once be placed in the same cell. Two such waves are sufficient for 6 – 8 superfluous unities to occur, and they very considerably augment the value of the chi-square. Thus, for example, occurred the deviation between empirical and theoretical numbers in Table 6, third cell from above (15 and 7.7). This is easy to become convinced of when having a look at Tables 3a and 3b.

If these considerations are valid, an essential improvement will happen at once when the number of groups is decreased by combining symmetric categories, see Table 7. We get $\chi^2 = 9.61$ and $P = 0.2$. In other words, not more probable deviations occur roughly once in five cases of independent elements. There are therefore no grounds for concluding that that distribution essentially differs from those indicated by the theory when independence is assumed.

Calculation of σ_r by issuing from data of Table 6 provides 0.250 which almost coincides with the case of different years. The conclusion is obvious.

8. Some preliminary results of analysing series of the solar constant and maximal temperatures and derivation of the mean square error of the correlation coefficient

If SC and MT are really not correlated, the mean square error of the empirical CC should be represented by a comparatively simple formula

$$\sigma_r^2 = \frac{1}{n}\sum_{t=-n+1}^{n-1} \rho_x(t)\rho_y(t)$$

in which $\rho_x(t)$ and $\rho_y(t)$ are the true CCs between x_t and x_{i+t} and y_t and y_{i+t}. The difficulty in applying that formula consists in that, instead, we have to make do with the statistical CCs, $r_x(t)$ and $r_y(t)$; for more details, see my paper Slutsky (1932) quoted above. The errors of these CCs can essentially corrupt the results because a large number of terms are being added up. In that previous paper the problem was really solved, at least in principle, for the case of $\rho(t) = 0$, $t > \omega$ and not large values of ω as compared with n. An example of a more difficult case is apparently encountered with the SC. We will assume an obviously highly probable hypothesis that the CCs between the values of SC separated by a year or more are either zero or negligible.

Comparing segments of the series of MT with numbers 156 – 255 taken either entirely ($n = 100$) or by parts with 40, 60 and 80 terms with the corresponding segments of the series of SC differing in time by one or two years in either direction and additionally shifted by 0 – 15 days we have calculated 112 CCs for shifts of about 1 year, and 80 CCs for shifts of about 2 years for each of the cases $n = 40, 60, 80, 100$. Table 8 contains empirical mean square errors of the CCs calculated accordingly and we note that for shifts of about 2 years all the σ's are somewhat smaller which perhaps argues for the presence of some remaining correlation (in any case, quite insignificant) at shifts of about 1 year. This can be checked by a similar study extended to shifts of 3 and 4 years. Anyway, the indicated differences can be neglected in the first approximation, and this is what we do.

Issuing from the known expansion

$$\sigma_r^2 = \frac{A}{n} + \frac{B}{n^2} + \frac{C}{n^3} + \ ...$$

and restricting it to three terms, we determine by least squares that

$A = 9.28$, $B = -164$, $C = 2190$.

The *theoretical* (i. e., the adjusted) values of σ_r^2 are shown in the last column of Table 8.

We consider the satisfactory adjustment as a testimonial that the number of terms allowed for in the formula above was sufficient and that, as I have shown in the paper quoted above, the value of A should therefore satisfy the approximate equality

$$\frac{A-1}{2} = \sum_{t=1}^{\omega} \rho_x^2(t). \qquad (*)$$

Replacing here ρ by empirical CCs r, we can determine an approximate value of ω which is calculated by taking $\rho_x(t) = 0$ for $t > \omega$; if $\omega > (n - 1)$ it should be replaced by that difference.

The next table (Table 9) provides the values of the serial CCs for SC and MT with shifts of 1, 2, …, 31 days and for the 156th – 255th days of each year when correlated for shift t with the segment $(156 - t; 255 - t)$. All these CCs were calculated for the first and the second half of the 8-year period, and for that period as a whole.

The following remark suggests itself first of all: the first and the second 4-year period both for SC and MT provide sufficiently close correlational functions at least when the CCs are still more or less considerable; the discrepancy between them can be certainly explained by random errors[10]. A curious conclusion is that both SC and MT, after eliminating the 30-day level [?], and a suitable standardization of the fluctuations can be considered homogeneous, at least in the first approximation. If the future confirms and extends that inference to other geophysical series, it will be quite an important step in their statistical studies.

We have found the value of the coefficient A, $A = 9.28$. Therefore, the right side of (*) is equal to 4.14. We do not know the true CCs or values of ρ_x, but when replacing them by their approximate values r_x, the sums of the squares of the CCs calculated by means of Table 9 provide

$$\sum_{t=1}^{31} r_x^2(t) = 3.45$$

and it is obvious that, since the further CCs are doubtless small, a large number of them are needed for coming near to 4.14, so that ω should be considerably greater than 31.

However, bearing in mind that the squared sum of all the rest CCs in the series of SC from $t = 32$ to infinity is a magnitude of the order of 0.5 (approximately equal to the difference 4.14 – 3.45), we may hope that the sums of the products of serial CCs for the SC multiplied by the same CCs for the MTs can also be established although somewhat roughly. Multiplying the appropriate values taken from Table 9, we find for $n = 100$ the approximate equality

$$\sigma_{r_{xy}} = \sqrt{\frac{1}{100}[1 + 2\sum_{t=1}^{31} r_x(t) r_y(t)]} = \sqrt{0.0897} = 0.30$$

which is very near to its empirically determined value 0.27.

In all probability, the further CCs (for shifts $t > 31$) are not important and, in addition, the error made by neglecting them was possibly compensated by dropping the term of order $1/n^2$. In any case, it is hardly accidental that the values of the mean square error of the CCs between SC and MT derived by such different methods are so close.

Explanation of tables and figures

Table 1. It lists the values of SC both adopted by Slutsky and either published in the *Annals* (1932) indicating categories *satisfactory* (S), *almost satisfactory* (S–) and *unsatisfactory* (U), or included with symbol W in the *Daily Map*, and the differences between them.

Table 2. Lists the month and day for the 1st, 56th, 155th, 255th and 365th day of an artificial calendar. Example: the 155th day of 1927 = 3 June 1927.

Fig. 1. Cordoba, SC and MT, separately. Shows by points their mean variability σ^2_{3ij} over three months (Jan. – March, Febr. – April, etc) for 1924(1)1931. Their mean variability (the deviations of the decade means from the monthly means) over those eight years σ^2_{3i} shown by small circles. Continuous curves show the sum of three or two harmonics for SC and MT respectively. Translation of legend partly tentative owing to difficult original text.

Table 3a. Lists CCs between SC and MT for period 56th – 155th day, years 1924(1)1931, shifts 0(1)15 days; separately shown are combinations of same year and of different years.

Table 3b. Same for period 156th – 255th day.

Fig. 2. CCs between SC and MT for same year (two upper series) and different years (the lower series), shifts 0(1)15 days. Additional curves shown with inadequate explanation moreover only given in text.

Fig. 3. Maximal in absolute values CCs between SC and MT for same year (A) and different years (B) for series of 100 and 50 days and shifts of 0(1)15 days.

Table 4. Lists magnitude σ^2_t for each of 12 months, year not indicated. Explanation lacking; text (§ 6) only states that SC is meant.

Table 5. Frequency table of CCs between SC and MT for different years, separately for SC preceding MT and vice versa and combined. Theoretical magnitudes additionally provided.

Table 6. Frequency table of CCs between SC and MT for same year, empirical (m') and theoretical (m) values.

Table 7. Same for absolute values of those CCs. Magnitude $[(m' - m)/m]^2$ additionally provided leading to $\chi^2 = 9.61$ and $P = 0.2$, see end of § 7.

Table 8. Lists empirical mean square errors of coefficients of serial correlation for SC, σ^2_r, shifts of about 1 year and about 2 years, and both these shifts combined, periods of 40, 60, 80 and 100 days. Theoretical values of σ^2_r additionally provided.

Table 9. Lists coefficients of serial correlation for SC and MT, shifts of 1(1)31 day, periods 1924 – 1927, 1928 – 1931 and 1924 – 1931, interval 156th – 255th day.

Notes

1. In § 5 of the main (Russian) text, Slutsky wrote: *We show* [on Fig. 2] 4 *correlation functions* etc. Anyway, it is difficult to understand what exactly is shown there. In the context of this paper, *correlation function* means *values of the CCs*. O. S.

2. Abbot (*Annals* 1932, p. 277 and 255ff) has recently put forward a new concept concerning the connection between SC with the weather. He assumes that each periodic component of that constant is reflected in the phenomena of weather with differing shifts moreover variable in time. Separate waves are superimposed upon each other and the connection can be lost in the general picture. The material he adduced for proving this thesis is still too scanty for being convincing but it is extremely interesting, suggests ideas and for the time being compels us to abstain from a final judgement. A check of that new theory was not included in our aims. E. S.

3. Abbot (*Annals* 1932, Table 31, pp. 195 – 213) had since essentially corrected the values of the SC published there before the indicated date. E. S.

4. For Cordoba, Clayton derived one of his best results, CC = – 0.74. True, the CC was even higher for some stations in Argentina, – up to – 0.82 in Sarmiento, – but upon revealing that there were so many missing days we preferred Cordoba. E. S.

5. When being increased by 1, there will be 16 (groups) – 2(connections) + 1 = 15 degrees of freedom, as Fisher called it. E. S.

6. By applying the formula

$$\sigma_z = \sigma_r \sqrt{1 + 2\sigma_r^2 + 2\frac{1}{3}\sigma_r^4 + 4\sigma_r^6 + \frac{1}{5}\sigma_r^8 + 24\frac{2}{3}\sigma_r^{12} + \ldots},$$

see Slutsky (1932, pp. 95 – 96). E. S.

7. I took the values of z corresponding to $r = 0.5$ [0.05?], 0.15, 0.25 etc. from Romanovsky's table (1928, p. 147). E. S.

8. Apparently, Slutsky (1929). O. S.

9. It was Fisher, who, in 1925, showed that the chi-squared test was not suited for studying dependent trials, see Hald (1998, p. 201). O. S.

10. We saw that for sufficiently large values of n and $t > 2\omega$ we may take $\sigma_r = \sqrt{9.28 / n}$ for the CC between SC and MT. According to the above calculations, we have $\sigma_r = 0.267$ at $n = 100$ and we may therefore approximately assume that $\sigma_r = 0.13$ at $n = 400$. Although all the necessary formulas are available, we are not yet able to calculate σ_r for serial CCs at lesser shifts, but the indicated magnitudes probably provide sufficiently correct indications about their order. E. S.

Bibliography

Abbot C. G. (1927), Corrected solar constant values, Montezuma, Chile, from May 27 to August 24, 1927 inclusive. *Monthly Weather Rev.*, September.

--- (1931), Weather dominated by solar changes. *Smithsonian Misc. Coll.*, vol. 85, No. 1 Publ. No. 3114, Washington.

--- (1932), In *Annals Astrophys. Obs. Smithsonian Instn*, vol. 5.

Clayton H. H. (1923), *World Weather*. Washington.

Daily Map (no date), *Daily Weather Map of the United States Weather Bureau.*

Hald A. (1998), *A History of Mathematical Statistics from 1750 to 1930.* New York.

Romanovsky V. I. (1928), *Elementy Teorii Korreliatsii* (Elements of the Theory of Correlation). Tashkent.

Slutsky E. (1929, in Russian), On the [mean] square error of the correlation coefficient for homogeneous connected series. *Trudy Konjunkturn. Inst.*, vol. 2, pp. 64 – 101.

--- (1932, in Russian), On the distribution of the errors of the correlation coefficient for homogeneous connected series. *Zhurnal Geofiziki*, vol. 2, No. 1, pp. 66 – 98. **S, G,** 40. Corrections in No. 2.

Some issues in the history of the theory of errors

Abstract of candidate dissertation. Moscow, 1967.
Published as a manuscript. Inst. Istorii Estestvoznania i Tekhniki

The rise of the theory of errors as a discipline belonging to experimental quantitative science is connected with the spectacular successes of instrumental astronomical observations of the 17th century and the beginning of the epoch of meridian arc measurements (17th – 18th centuries). The advances in, and subsequent new problems of astronomy put onto the agenda various issues concerning the treatment and evaluation of the precision of instrumental observations; and the calculation of arcs of triangulation and the determination of the figure of the Earth demanded, furthermore, the ability to treat redundant systems of linear algebraic equations. The theory of errors thus naturally included a number of problems pertaining to the treatment and estimation of the precision of direct and indirect observations. The treatment of direct observations led to the justification, of a qualitative, and then of a quantitative, stochastic nature, of the long prevalent arithmetic mean. The treatment of indirect measurements demanded the development of a number of algorithms which were independent of probabilistic considerations; later, however, Gauss derived the method of least squares by issuing from the principle of maximum likelihood.

In the process of my work, I became acquainted with a large number of sources in mathematics, astronomy and geodesy. A large part of this literature remained little known or completely ignored. This very situation can explain to a certain degree why I succeeded in getting unexpected important results and among them

1. The description of the work of the unknown in Russia American mathematician, Robert Adrain, who, in particular, published two derivations of the normal distribution in the theory of errors a year before Gauss [or at least not later than he].
2. The establishment of the priority of Ernst Abbe in considering the chi-squared distribution.
3. The first appearance of the principle of maximum likelihood is due to Lambert in 1760.

The literature on the theory of errors and the method of least squares paid great attention to the justification of the normal law. That problem occupies only a secondary place in my work. Indeed, this is perhaps a separate topic in which, first of all, it is necessary to examine the theory of elementary errors and the central limit theorem. And a large number of other derivations turned out to be dead ends and today represent only mathematical exercises.

My work consists of three chapters, an appendix, and an addendum.

Chapt. 1 investigates the early application of the arithmetic mean in approximate calculations, games of chance, in astronomy and theory of probability proper. That it was used in antiquity has been known for a long time. However, the connection of this fact with the stochastic nature of the mean was not noted, – and precisely this fact is stressed throughout almost the entire chapter. Furthermore, a large amount of factual material is collected there.

The science of mean values, including the arithmetic mean, already existed in the Pythagorean school and had been widely applied in antiquity in approximate calculations of the areas of figures and volumes of bodies. The formula for the calculation of the area of a quadrangle as the product of half the sums of opposite sides was used in ancient Babylonia either for not exact rectangles or when the opposite sides were unequal because of the ruggedness of the terrain (A. A. Vaiman). This means that the arithmetic mean was invoked to compensate both for the lack of strict applicability of the formulas (models) and for the systematic (not random, as nowadays) errors of measurement.

Commercial practice aided the spread of the idea of the arithmetic mean, and claims were even made that in this sense the sphere of economics was primary. According to Leibniz, the principle of equal allowance of equally tenable assumptions was the fundamental hypothesis in the contemporaneous theory of probability. I believe that Leibniz thought about the origin of the first stochastic concepts, probability and expectation, and that his idea provided the possibility of the later subjective understanding of probability.

The history of games of chance contains evidence of the widespread knowledge of the idea of the arithmetic mean which served as an intuitive statistical indicator of the totality of possible outcomes and perhaps of the appropriate expectation. During the epoch of meridian arc measurements the arithmetic mean began to be used as an universal estimator in all stages of their treatment. In the same period and even earlier the first qualitative statements about the benefits of applying it had appeared (Copernicus, Kepler, Picard, Condamine).

In 1809 Gauss postulated the principle of the arithmetic mean and essentially used it in his derivation of the normal distribution. This attracted attention to his postulate and a number of authors tried to reduce it to a *more obvious* premise. These attempts were, however, of a purely deterministic nature and I do not dwell on them. In the first half of the 18^{th} century Cotes applied an analogy from mechanics (the centre of gravity) for justifying the arithmetic mean, and Lambert tried to substantiate it on a stochastic basis.

In Chapt. 2, in connection with the history of the treatment of direct observations, I studied the work of Galileo and Lambert, then dwelt on the appropriate memoirs of Simpson and Lagrange. Also there I investigated estimators with posterior weights, the principle of maximum likelihood and the rejection of outlying observations.

Galileo was the first to formulate a number of basic theorems in the theory of errors (Maistrov). Lambert, who laid the foundation for that theory, was the chief predecessor of Gauss in this direction. When substantiating the advantages of the arithmetic mean, Simpson issued from the objective properties of observational errors. His immediate aim was to refute the opinion of “some persons of considerable note” that one careful observation can be relied on as much as on the mean of a great number of them, and in this connection I studied the works of Flamsteed and Bradley. I concluded that the abovementioned opinion, to which these astronomers never subscribed, was the result of the great successes of observational techniques and did not have a lasting influence on experimental science. Lagrange, without mentioning Simpson, reproduced his results and studied several continuous distributions. Just like Simpson, he applied generating functions, even in case of continuous distributions, and thus anticipated the use of characteristic functions.

Beginning with the second half of the 18th century (Short, Euler, De Morgan, Newcomb, Ogorodnikov), estimators with discrete or continuous posterior weights had repeatedly been proposed instead of the arithmetic mean. Some authors thought that posterior weights can allow for the change over time of the parameters of the appropriate law of distribution. In my opinion, with an even law (a natural assumption) such estimators only provide a correction to the ordinary arithmetic mean for the deviation of the observations from pairwise symmetry. However, these estimators may be considered as an historical analogue of some modern statistical estimators.

In 1778 Daniel Bernoulli proposed an estimator with posterior weights increasing towards the tails of the adopted distribution (an arc of a parabola). He sharply criticized the arithmetic mean considering it suitable only for uniform distributions; instead, he proposed the principle of maximum likelihood. The unusual behaviour of the posterior weights would have seemed unacceptable; however, Euler, in a companion commentary, mistakenly concluded that the weights decreased to the tails. Incidentally, in such *unusual* cases posterior weights are no alternative for rejecting the outliers.

I have shown the similarity in the use of the principle of maximum likelihood by Adrain (1808 [or 1809]) and Gauss (1809) for deriving the principle of least squares and the arithmetic mean and investigated the justification of maximum likelihood by inverse probability (Laplace, Gauss).

Rejection of outliers was recommended by Galileo and systematically applied by Lambert. Stochastic criteria for rejection were only devised in the second half of the 19th century. Their appearance was inevitable both because of the desire to abandon arbitrary rejection and of the expansion of the domain of applications of probability. On the other hand, the development of such tests was delayed by the fetish made of the normal law according to which any error was possible and perhaps by the opinion of Gauss who allowed rejection only in cases of gross errors.

The first stochastic criteria for rejecting outliers (Pierce, Chauvenet and others) were based on direct calculation, according to the normal law, of the odds for and against dubious observations and gave rise to a drawn-out polemic where opinions in essence leading to the consideration of errors of the two kinds were expressed: better to sacrifice a few possibly reliable observations but get rid of the dangerous influence of large errors. Thus, Gauss notwithstanding, these criteria rejected large errors regardless of their origin. The errors of these tests resulting from small divergences of the distribution of observational errors from normality (non-robustness of criteria) were not investigated. Furthermore (Barnett & Lewis, 1978), no answer appeared (and will possibly never appear), to the questions: What is an outlier and how to deal with outliers?

At the end of the chapter I mention that a quantitative estimation of precision began to be used relatively late. With the exception of Lambert (who, for that matter, did not norm his measure of precision and therefore could not directly compare several series of observations), no one until Gauss (1823) introduced any such measure. Incidentally, I found a normed estimator of precision in a work of Delambre written sometime during 1818 – 1822.

In Chapt. 3 I study the history of the mathematical treatment of indirect observations (the solution of redundant systems of linear algebraic equations by some supplementary conditions imposed on the residual free terms v_i). Above all, for the frequently occurring case of two unknowns (in particular, the unknown parameters of the terrestrial ellipsoid of revolution), I investigated the practice of using pairwise combinations of measurements (condition: $v_i = 0$). A similar method was traced up to 1827. In addition, I discovered that the method of pairwise combinations was used while treating direct observations (Boscovich) with the subsequent calculation of the arithmetic mean over all the combinations. I assumed that the combinations were applied here for a qualitative evaluation of the precision of observations, and, in addition, to apply a single algorithm for treating both direct and indirect observations.

I also traced the connection of the method of pairwise combinations with the method of means ($\sum v_i = 0$). Tobias Mayer (1750), while solving a system of 27 equations in three unknowns, grouped them in three summary equations. He justified this procedure (a generalized method of means) by the practical impossibility of forming and solving all the possible combinations of the equations in threes. The condition of the method of means was naturally regarded as resulting from the equal probability of errors of each sign and leading to the arithmetic mean in the case of direct observations.

I also describe the treatment of meridian arc measurements by Boscovich. Not being satisfied with the method of combinations, he proposed the conditions

$$\sum v_i = 0, \ |\sum v_i| = \min$$

which were also used later on by Laplace. A. A. Gusak described the history of the minimax method (condition: $\max|v_i| = \min$ with the minimum being taken among all possible solutions of the system). I supplemented his account by several remarks and indicated, in particular, that Euler had applied elements of this method in 1749 (not in 1778).

Gauss (1809) derived the method of least squares on the basis of the normal distribution of random observational errors. However, their usual properties were only connected with this law through the arithmetic mean. In 1823 Gauss published his second derivation of least squares by issuing from the principle of maximal weight. He thus renounced his previous tacit assumption that the normal law was the only possible law of error. The principle of maximal weight for a finite series of measurements was already known to Laplace who thought that the optimal result corresponded to maximum weight with the weights being inversely proportional to $\sum v_i^2$. But at the same time Laplace defined the weight as the positive parameter k of the law of the type $\exp(-kx^2)$ and he thus reduced the condition of its maximum to the least probability of errors, or to the shortest length of a *confidence interval*. This point of view did not allow him to renounce the normal law as the universal law of error. However, one should recognize a greater similarity of Laplace's and Gauss' ideas than it is usually recognized; it is hardly opportune to contrast these great men of science to each other.

Precisely their community of interests in treating observations enabled each of them to formulate better the unsolved problems, and, when attacking them, to rely on the results of each other.

In the **Appendix**, I cite short biographical data on the American mathematician Robert Adrain (1775 – 1843) and investigate in detail his work in the theory of errors. I also trace a number of later derivations of the normal distribution. In 1808 [or 1809] Adrain published an article which contained two derivations of the normal law of random observational errors; a derivation of the principle of least squares (it is barely possible, however, that Adrain was acquainted with the work of Legendre) and of the arithmetic mean; a determination of the most probable position of a ship from dead reckoning and an observation of its latitude; and an adjustment of a closed compass traverse.

In 1818 Adrain published two articles devoted to the derivation of the flattening and the size of the Earth. Using the data collected by Laplace, he applied the principle of least squares and obtained 1/319 for the flattening. In his second article, he arrived at a remarkably good estimate of the Earth's radius. His results were wonderful, but of course his work cannot be compared with the achievements of Gauss either directly or even less in its significance for the later development of the theory of errors since he remained virtually unknown.

In the **Addendum** I attempted to sketch a general outline of the history of the theory of errors, and, in particular, to explain the reason for the existence of two versions of the theory, the mathematical-statistical, and the astronomical-geodetic.

Having worked without a scientific mentor, I consider it my duty to mention with even greater appreciation the participants and the guides of the seminar on the history of mathematics and mechanics at the Moscow Lomonosov State University. The atmosphere at the seminar in general, as well as the advice received, essentially helped me.

Judith A. Behrens and Walter L. Sadowski had translated this piece about thirty years ago; I have revised their work.