

**Studies in the History of Statistics and Probability
Collected Translations**

Compiled and translated by Oscar Sheynin

Berlin, 2009

ISBN 3-938417-94-3

© Oscar Sheynin, 2009

Contents

Introduction by Compiler

- I.** M. V. Ptukha, Sampling investigations of agriculture in Russia in the 17th and 18th centuries, 1961
- II.** A. A. Chuprov, The problems of the theory of statistics, 1905
- III.** N. K. Druzinin, Scientific contents of statistics in the literature of the 20th century, 1978
- IV.** E. E. Slutsky, Theory of Correlation and Elements of the Doctrine of the Curves of Distribution, Introduction, 1912
- V.** E. E. Slutsky, Statistics and mathematics. Review of Kaufman (1916), 1915 – 1916
- VI.** Oscar Sheynin, Karl Pearson 150 years after his birth, 2007
- VII.** H. Kellerer, W. Mahr, Gerda Schneider, H. Strecker, Oskar Anderson, 1887 – 1960, 1963
- VIII.** S. Sagoroff, Oskar Anderson, Obituary, 1960
- IX.** V. I. Riabikin, Oskar Anderson, Chuprov's student, 1976
- X.** O. Anderson, To the memory of Professor A. A. Chuprov Junior, 1926
- XI.** Oskar Anderson, Ladislaus von Bortkiewicz, 1932
- XII.** Oskar Anderson, On the notion of mathematical statistics, 1935
- XIII.** Oskar Anderson, Mathematical statistics, 1959
- XIV.** P. P. Permjakov, From the history of the combinatorial analysis, 1980
- XV.** K.-R. Bierman, Problems of the Genoese lotto in the works of classics of the theory of probability, 1957
- XVIa.** V. Ya. Buniakovsky, Principles of the Mathematical Theory of Probability, Extracts, 1846
- XVIb.** A. Ya. Boiarsky, E. M. Andreev, The Buniakovsky method of constructing mortality tables, 1985
- XVII.** A. M. Liapunov, Pafnuty Lvovich Chebyshev, 1895
- XVIII.** A. A. Konüs, On the definition of mathematical probability, 1978
- XIX.** Oscar Sheynin, Review of Ekeland (2006), unpublished
- XX.** Oscar Sheynin, Antistigler, unpublished

Introduction by Compiler

Following is a collection of papers belonging to statistics or probability theory and translated from Russian and German; in some cases, the dividing line is fuzzy, as it ought to be expected. They are designated by Roman numerals which I am also using just below when providing some general remarks concerning them.

Authors of some rather long papers had subdivided them into sections; otherwise, I myself have done it (and indicated that the responsibility is my own by numbering them [1], [2], etc). References in this Introduction are to the Bibliographies appended to the appropriate papers. Finally, in cases of cross references in the main text, these are such as [V, § 2] and the unsigned Notes are my own.

I. Mikhail Vasilievich Ptukha (1884 – 1961) was a statistician and demographer much interested in the history of statistics, He was a Corresponding Member of the Soviet Academy of Sciences.

Ptukha describes the elements of sampling in agriculture during the period under his study. This could have been dealt with in a much, much shorter contribution, but he also tells us which types of estates had applied these elements and how exactly was each procedure done to ensure more or less reliable data. Of course, much more interesting is the practice of sampling for checking the coining in England which is seen in the very title of Stigler (1977): *Eight centuries of sampling inspection. The trial of the Pyx*. Sampling in agriculture, however, should also be documented.

Ptukha begins by saying a few words about sampling at the turn of the 19th century. On Kiaer (whom Ptukha mentions) and other statisticians of that period, both opposing and approving sampling, see You Poh Seng (1951). It is also opportune to add that Seneta (1985) and Zarkovich (1956; 1962) studied sampling in Russia during the early years of the 20th century.

Meaning of special terms and old Russian measures

Dvortsovaia: dvorets means *palace*, and *dvortsovaia* is the appropriate adjective apparently concerning the Czar's palace.

Sloboda: suburb

Stolnik: high ranking official at court

Tselovalnik: the man who kissed (*tseloval*) the cross when taking the oath of office. In particular, he took part in the judicial and police surveillance of the population

Voevoda: governor of province

Volost: small rural district

Votchina: patrimonial estate owned by the *votchinnik*

Chetverik: 18.2 kg

Dessiatina: 1.09 hectare

Pood: 16.4 kg

Sazhen: 2.13 m

Sotnitsa: some unit; *sotnia* means *a hundred*

II. In 1909, Chuprov published his Russian *Essays on the Theory of Statistics* which became extremely popular and at the time, and even now Russian statisticians consider it a masterpiece although Druzinin [III] critically surveyed his work; my remarks below ought to be supplemented

by his paper. Among mathematicians, Markov (1911/1981), provided a critical comment. Thus (p. 151), the *Essays* “lack clarity and definiteness”.

Chuprov preceded his *Essays* by two long German papers (1905; 1906) the first of which I am translating below from its Russian translation made by his closest student N. S. Chetverikov but I have also checked many places of the Russian version against the original and provide the page numbers (in bold type) of the original to show that the translation was correct. In one case, however, Chetverikov made a mistake and I have additionally inserted there my own translation.

I left out many passages which either seemed not really needed or much more properly belonged to philosophy. Indeed, Chuprov’s exposition was (and to a certain extent remains in the translation) verbose which was possibly occasioned by the author’s wish to satisfy his readers (and perhaps corresponded to his pedagogical activities at Petersburg Polytechnical Institute) and Markov’s remark is appropriate here also.

In 1910 – 1917 Chuprov corresponded with Markov (Ondar 1977) after which (and partly during those years) his most important contributions had been made in the mathematical direction of statistics; I do not say *mathematical statistics* since that discipline had only begun to emerge in those years. Seneta (1982; 1987) described some of Chuprov’s relevant discoveries, but my translation gave me an opportunity to provide some additional related information.

1. Resolutely following Lexis and Bortkiewicz, Chuprov paid much and even most attention to the justification of statistics by the theory of probability. The problem facing statisticians in those times (also much earlier and apparently somewhat later) consisted in that they, perhaps understandably, had been interpreting the Bernoulli theorem in a restricted form, and even Lexis (below) had wavered on this issue. They contended that that theorem was only useful when the theoretical probability (and therefore the equally probable cases) really existed.

Actually, we may imagine its existence (if only not contradicting common sense) given its statistical counterpart. At least towards the end of the 19th century the appearance of the non-Euclidean geometry had greatly influenced mathematicians who became quite justified to treat imaginary objects. Chuprov, although a mathematician by education, had never taken a resolute stand; here, his relevant statement in § 3.2 was at least not definite enough. Lexis wavered over this issue. At first, he (1877, p. 17) stated that equally probable cases might be presumed when a statistical probability tended to its theoretical counterpart, but there also he (p. 14) remarked that, because of those cases the theory of probability was a subjectively based discipline and he (1886, p. 437) later repeated that idea. And in 1903 he (p. 241 – 242) confirmed his earliest pronouncement that the existence of such cases was necessary for “the pattern of the theory of probability”.

Another point pertaining to mathematical statistics if not theory of probability is Chuprov’s failure to state that the ratios of the different measures of precision to each other depended on the appropriate density function. True, he only followed the general (including Lexis’) attitude of the time, but a mathematician should have been more careful.

2. Instead of enthusiastically supporting the Lexian theory of stability of statistical series in 1905 (see below), he became its severe critic and

finally all but entirely refuted it, which his students soon found out; it is strange that his work in that direction is still barely known outside Russia.

Chetverikov inserted a special footnote referring to his (although not the first) relevant paper. Already in 1914, even before publishing his decisive German papers on the stability of series, Chuprov thought about abandoning the Lexian theory. In a letter No 135 to him of that year Bortkiewicz (Bortkevich & Chuprov 2005) disagreed about “shelving” the Lexian coefficient Q .

3. Bortkiewicz’ law of small numbers (1898), which he considered (and which indeed was) inseparably linked with the Lexian theory, was another issue about which Chuprov at least partly changed his opinion. Moreover, since the two scientists had been extremely close to each other with Bortkiewicz stubbornly clinging to his law, it is even possible that Chuprov just did not publicly express his opinion.

In 1905, Chuprov approved Bortkiewicz’ innovation but later he (1909/1959, p. 284n) listed four possible interpretations of that law; there also, on p. 277, he remarked, although without naming Bortkiewicz, that

The coefficient Q cannot be a precise measure of the deviations of stability from its normal level: it does not sufficiently eliminate the influence of the number of observations on the fluctuations.

Then, in 1916 (Sheynin 1990/1996, p. 68), he informed Markov that

It is difficult to say to what extent does it enjoy recognition of statisticians since it is not known what, strictly speaking, should be called the law of small numbers. Bortkiewicz did not answer my questions formulated in the Essays [see above] either in publications or in writing. I did not question him to his face at all since he regards criticisms of that law very painfully.

The title of my paper (2008) explains my understanding of the law of small numbers. See also Heyde & Seneta (1977, § 3.4).

It was Ondar (1977) who had published the correspondence between Markov and Chuprov, but I have since discovered some additional letters, see the quotation above. I take the opportunity to add that he, while moving from manuscript to publication, made a lot of mistakes whose list I adduced in 1990; all or almost all of them were left in the translation of that source.

4. Chuprov (1905) made strange statements about the law of large numbers. He attributed the main merit of its justification (and, as I understand him, even of its Chebyshev generalization) to Cournot but had not referred to him definitely enough. Later Chuprov (1909/1959, p. 166) argued that Cournot had provided the “canonical form” of the proof of that law, see [III, Note 5].

5. Chuprov (1905) discussed the method of comparing statistical series with each other for deciding whether the two relevant phenomena were connected or not. He certainly did not know about rank correlation or the Spearman coefficient ρ , but he could have mentioned some previous contributions to natural sciences. Thus, in 1865 – 1866 the German astronomer, mathematician and optician Seidel (Sheynin 1982, pp. 277 –

279) investigated the possible connection between the number of fever cases, level of subsoil water and rainfall by the same method and even numerically estimated the discovered ties. A bit later, in 1869 – 1870, the celebrated English physician (and inventor of the term *Hospitalism*) Simpson investigated several possible connections concerning surgical diseases (Ibidem, pp. 263 – 264).

6. I am naturally led to a discussion of Chuprov's attitude towards the application of statistics in natural sciences. In 1905, he argued that a creation of two different statistical methods for application in sociology and natural sciences would be unreasonable and he repeatedly (also elsewhere) stated that statistics had *recently* begun to enter those sciences. Later he published two papers on that subject, one in Russian, in 1914, and then its German version (1922a), but he never had time to study sufficiently that issue, see his Letter 124 of 1913 (Bortkevich & Chuprov 2005). In 1905, apart from his statement (above), Chuprov did not say anything on that subject at all.

Actually, statistics had been gradually entering natural sciences from a much, much earlier time, see my five papers in the *Archive for History of Exact Sciences* (vols 22, 26, 29, 31 and 33 for 1980, 1982, 1984 (twice) and 1985).

7. Chuprov (1905) discussed *nomological* and *ontological* relations; the former applied to certain phenomena universally or at least in general, the latter's action was restricted in space and/or time. Such a distinction has been known in natural sciences, for example in gravimetry: pendulum observations clearly showed that there existed local gravimetric anomalies, so was it really necessary to introduce those terms into statistics?

Then, Chuprov (1906, §4; 1960, p. 126) specified: statistics is an ontological science, but is closely connected with the notion of mass phenomena. This seems unnecessarily complicated the more so since he (see here, beginning of § 1.3) defined the aim of the statistical method as discovering, under its own assumptions, the laws of nature.

In 1909, Chuprov continued in the same vein: he discussed the place of statistics among other scientific disciplines by means of two other terms (nomography, the science of the general) and idiography (the science of the particular, the restricted) and pertinent notions borrowed from German philosophers, Rickert and Windelband, whom historians of that science but definitely not statisticians still remember. There, beginning with the second edition of 1910, he inserted an Introduction entitled *The Principal Problems in the Theory of Mass Phenomena* which indirectly defined statistics. Then he somehow separated the statistical method (of nomographical essence, pp. 129 and 130) from statistics (an idiographic study of reality, p. 130) and on p. 301 he largely repeated these ideas and also restricted statistics by studying sociology. Nevertheless, he (p. 75) somehow held that it was obsolete to believe that the sociological phenomena were atypical. See Druzinin [III] for a final verdict, but I have something else to say: Chuprov could have noted that the so-called *numerical method* officially introduced into medicine in 1825 (and unofficially into other branches of natural sciences much earlier), was idiographic (Sheynin 1982, § 4). Thus, compilation of an astronomical yearbook is an idiographic study.

8. Another philosophical subject which Chuprov (1905) discussed at length was induction. He did not however mention incomplete induction whereas Bayes only appeared fleetingly in a footnote and only in respect of his “theorem” (actually lacking in his fundamental contribution).

I ought to add, however, that Chuprov is mentioned in the *Great Sov. Enc.* (Novoselov et al 1973/ 1977, p. 637): “Mill’s theory of inductive inference was subjected to elaboration and criticism, for example in the work of [...] Churprov”.

9. I note finally that randomness was conspicuously absent from his writings in general. Chuprov introduced plurality of causes and effects which was superfluous and misleading because it ought to be explained by the existence of random variables with pluralities of values and probabilities.

III. Nikolai Kapitonovich Druzinin (1897 – 1984) was an eminent Soviet non-mathematical statistician. I met him several times just after he had retired from the chair of statistics at the Plekhanov Institute for National Economy in Moscow, and am sure that he was an honest man and scientist sincerely believing in Marxism. His contribution translated below is the only Soviet or foreign writing attempting to analyse Chuprov’s general viewpoint on statistics more or less critically, and it reveals the weakness of the time, the poor knowledge of the history of the application of statistics in natural sciences.

A special point here concerns the correct statement that statistics studies mass phenomena. Druzinin mentioned it several times (§§ 11, 14 and 16) but forgot Chuprov (1909) and apparently did not know about Poisson et al (1835, p. 174). Chuprov, in the second edition of his book, inserted a special introductory section entitling it *Main Issues of the Theory of Mass Phenomena*, and the French authors wrote:

Statistics in its practical application, which invariably and definitely is the functioning mechanism of the calculus of probability, necessarily appeals to infinite masses, to an unlimited number of facts bearing in mind not only a most possible approach to the truth, but also, by means of known procedures, wishes to destroy, to eliminate as much as possible the numerous sources of error so difficult to evade.

I (1999) have expressed my own opinion about the contents of statistics, the statistical method and the theory of errors and about statistics and mathematics.

Another noticeable point is that Druzinin passed over in silence the pronouncements made at a prestigious Moscow statistical conference of 1954, inconceivable for any statistician worth his salt, see Sheynin (1998, § 5.1).

In accord with its title, the author provided many quotations from early Russian sources hardly available outside Russia. For that matter, he (1963) published a reader on the history of Russian statistics and among his other contributions on the history of statistics, apart from the two papers he himself referred to, I ought to mention his book (1979).

When possible, I have referred to original English editions (and quoted them) rather than to their Russian translations, as did the author.

IV. I (1990/1996, pp. 43 – 49) have described Slutsky’s early work, Chuprov’s opinion about it and Slutsky’s reasonable response to Markov’s criticisms. On Markov’s temporizing attitude see also Ondar (1977/1981, pp. 53 – 58), and, finally, see Kolmogorov (1948) for his high appraisal of Slutsky’s essay, the Introduction to which I am translating. A concise description of the stated above is in Sheynin (2005/2009, § 14.1-4).

A few comments. Bearing in mind Slutsky’s subject, it seems natural that he had nothing to say about the Continental direction of statistics. Without denying the significance of the Biometric school (as it became known) in general, and Pearson’s merits in particular, I note that later events proved that Fisher in the first place rather than Pearson is being remembered alongside the classics of probability of the 19th century. Slutsky had not mentioned either Chebyshev, or Markov, or Liapunov which I do not quite understand.

The first four sections of Slutsky’s pt 1, which he thought most important, were devoted to general notions about curves of distribution; moments; and measures of deviation and error.

V. Slutsky’s essay (rather than a simple review) is noteworthy as being at least until many decades later perhaps the most detailed and convincing examination of the essence of statistics as a science. It is not faultless and I have formulated quite a few relevant remarks in my Notes. Here, I am adding a few points.

First, Slutsky does not explain what does the *third* edition of Kaufman’s treatise (which earbed a prize of the Petersburg Academy of Sciences) mean. In my Bibliography here, I mention the edition of 1912, the German translation of 1913 and the reviewed edition of 1916 (later editions are here of no consequence), but it is rather unusual to include without reservations an edition in a foreign language. True, Kaufman (1909), is considered its the first edition, but Slutsky ought to have been more attentive here the more so since he refers to the second edition.

Second, Slutsky barely documented his account, and I myself supplied many references. Third, after politely but strongly criticizing Kaufman (but only reviewing him insofar as the relations of statistics and mathematics are concerned), Slutsky somehow concludes his work by praising him out of any proportion. Fourth and last, Slutsky is rather careless; some of his phrases are difficult to understand (partly because they are excessively abstract) and he is indiscriminately using two terms, *theory* and *calculus of probability*.

VI. I have not seen the proofs of the original Russian version of this paper; many Russian publishers do not deem necessary to allow authors *to put on airs*. Much worse: someone from the editorial office has unsparingly mutilated my text so that I may well doubt whether copyright exists in Russia (even in Petersburg rather than in a one-horse town).

Porter (2004) had recently put out Pearson’s biography which I (2006b) severely criticised. Indeed, what may we expect from an author who remarks (p. 37) that “even mathematics cannot prove the fourth dimension”, calls Thomson & Tait’s most influential treatise (reprinted in 2002!) “standard Victorian” (p. 199) and forgets that Pearson was elected Fellow, Royal Society?

For the bibliography of Pearson’s writings see Morant et al (1939) and Merrington et al (1983). Many of his early papers are reprinted (Pearson

1948) and his manuscripts are being kept at University College London. His best biography is E. S. Pearson (1936 – 1937) and I abbreviate both that author and his work by ESP.

VII. The authors, being the Editors of Anderson's *Ausgewählte Schriften* (Sel. Works), had also compiled a Vorwort (pp. IX – XI) where 14 publications concerning Anderson were listed, 13 obituaries and one note honouring his 65th birthday; the future co-editors wrote five of these obituaries (Strecker, three, and Kellerer, two). To these, I may add my own short note (*Dict. Scient. Biogr.*, vol. 1, 1970, pp. 154 – 155).

I have met with Professor Strecker who was Editor of my booklet on Chuprov (Sheynin 1990/1996), and I ought to add two points. First, that booklet was only published after an anonymous donor had covered the costs, and it took me rather a long time to figure out that most likely that donor was Strecker himself.

Second, Strecker was Anderson's student and he told me that A. had frequently and reverently mentioned his teacher, Chuprov, and considered himself Chuprov's *son*, and that he, Strecker, was Chuprov's *grandson*.

The Archive of the Ludwig Maximilian University in Munich keeps Anderson's short autobiography written in 1946 (code II-734), and in two places I insert additional information from this source identifying it by his initials, O. A.

My booklet (see above, pp. 58 – 60) contains other archival information about Anderson. In a reference for him Chuprov indicated that in 1910 and perhaps 1911 he participated in the census of the population of Petersburg and “again” proved himself “an excellent worker”. Then, in three letters of 1924 Chuprov made known that Anderson experienced a hard time in Budapest; that he, Chuprov, was “at last able to fix him up” with a position in Bulgaria, and that again he, Chuprov, spent more than two weeks putting in order Anderson's calculations in the manuscript later published in the *Journal of the Royal Statistical Society*. Finally, I note that Anderson's *Ausgewählte Schriften* contain his reminiscences [X] about Chuprov written after the latter's death.

Sheynin O. (1990, in Russian), *Aleksandr Chuprov: Life, Work, Correspondence*. Göttingen, 1996.

VIII. Slavco Sagoroff was an eminent Bulgarian economist and statistician, then Director of the Institute for Statistics at Vienna University, see *Metrika*, t. 14, No. 2 – 3, 1969 (a *Festschrift* commemorating his 70th birthday). He died in 1970, see obituary in the same periodical (t. 16, No. 1, 1970)

IX. I have formulated many critical Notes. Here, I ought to add three points. First, it seems to be generally known that many Soviet scientific contributions had been written carelessly (and published without due reviewing), and this particular paper is a good pertinent example. Second, the author understandably had not mentioned Anderson's attitude (1959, p. 294) towards the Soviet regime. I quote:

I could have cited many statisticians as well as promising younger students [...] of Chuprov much valued earlier in Russia whose names suddenly entirely disappeared after 1930 from the Soviet Russian scientific literature.

In that paper, Anderson severely criticized a Soviet semi-official statistical treatise.

Third and last, it is nevertheless hardly forgivable that the author has not stated that Anderson had become the leading statistician in Bulgaria, then in (Western) Germany and kept most favourably mentioning Chuprov on every suitable occasion.

To the best of my knowledge, Riabikin was the first Soviet author to publish a paper on Anderson, and a favourable at that, in the Soviet Union. The leading Soviet statistical periodical, *Vestnik Statistiki*, did not say a word about Anderson's death. True, I preceded the author, but my short piece appeared abroad (*Dict. Scient. Biogr.*, vol. 1, 1970, pp. 154 – 155).

XIII. In accordance with the source for which this contribution was intended, Anderson paid much attention to social statistics. I myself (1999) studied the definition and essence of statistics but regrettably omitted to mention him. The pertinent point now is that the main difference between theoretical and mathematical statistics, as I see it but as it was not clearly perceived a few decades ago, consists in that the former, unlike the latter, includes preliminary analysis of data (and, in general, pays less attention to mathematical aspects of the necessary work as well).

Another point concerns medical statistics. It does not belong to social sciences, and Anderson did not say anything about it, but, on the other hand, it is closely connected with population statistics, so that a certain gap seems to exist here. The same is true about public hygiene.

Sheynin O. (1999), Statistics, definitions of. *Enc. of Statistical Sciences*, 2nd edition, vol. 12, 2006, pp. 8128 – 8135.

XIV. The title of this essay, as it should, really explains its essence. Its subject does not belong to the history of probability proper, but is related to, and touches it. The author was diligent, but I still had to insert some Notes. Here, I say that he had not always provided the dates of the original publication of his sources and in a few cases of periodicals he preferred to indicate the year *for which* the appropriate paper had appeared. Then, he did not always refer to the latest available edition; thus, Bernoulli (1713) is included without indicating 1975. I have attempted to improve all that (and abandoned the Russian translations). Quotations from Cayley (§ 5) are taken directly from his papers rather than translated back from Russian.

XV. The Genoise lottery still interests us, witness Bellhouse (1991). The author, however, did not study its history in a restricted sense, but went on to describe numerical lotteries (or lottos), although not comprehensively at all. Todhunter (1865), whom he should have consulted, examined the pertinent work of Montmort, De Moivre and Laplace which Biermann (in Russian, Bierman) had not mentioned. Worse: he stated that lotteries met both with approval and opposition, but did not refer either to its harsh criticism by Laplace (1819) or to Poisson (1837, p. 68) who noted in passing that the *loterie de France* was “luckily suppressed by a recent law”.

Concerning the author's Bibliography, I indicate that he referred to page numbers of the initial publications of Jakob Bernoulli and Euler, an attitude certainly discouraging readers. Moreover, the German translation of the *Ars Conjectandi*, for example, additionally carries the page numbers of the 1713 edition and Biermann should have all the more referred to it.

XVI. Before Chebyshev, Buniakovsky (1804 – 1889) was the main Russian student of probability and statistics, and his book, extracts from which follow below, had been widely known. Here is what Markov (1914/1981, p. 162) and Steklov (1924, p. 177) had to say about it: “a beautiful work”; for its time, “complete and outstanding”. I remark, however, that Buniakovsky did not pay attention to Chebyshev’s work: after 1846, he actually abandoned probability for statistics.

I myself (1991) described his work and appended a list of his contributions to both those subjects. Now, I regret that the book itself is here hardly available. Nevertheless, I supplemented the translation of the extracts by many notes, and will only add that, strange as it is, Buniakovsky, while mentioning Buffon, had not indicated that the theory of probability owes him the real introduction of geometric probability.

I am also translating a very short note on the Buniakovsky method of compiling mortality tables and I stress that he was the first to study seriously mortality in Russia.

XVII. Liapunov’s essay written long ago is certainly not suited to appreciate Chebyshev’s heritage; but, apart from being himself a first-class scholar, he was Chebyshev’s student and his recollections (and, for that matter, his opinions) are indeed valuable.

Two circumstances ought to be mentioned. First, in 1895, when Liapunov wrote his essay, he had not yet studied the theory of probability. No wonder that Chebyshev’s relevant achievements are not really described, and in § 5 Liapunov even wrote *law of very large numbers*. Second, Liapunov’s own footnote was accompanied by a note by the Editor of the source, Chebyshev’s *Sel. math. works*, N. I. Achiezer, which I replaced by a much stronger modern comment.

Chebyshev’s almost complete works, although without commentaries, appeared in Russian and French, see my Bibliography. Since then, his *Complete Works* in five volumes were published in Russian, and the bibliography of his works (obviously complete) in the last volume is also mentioned in my Bibliography.

XVIII. Aleksandr Aleksandrovich Konüs (1895 – ca. 1991) was an economist and statistician actively working in economic statistics (Diewert 1987). In 1923 he began work under Kondratiev at the Conjecture Institute in Moscow and had been directly engaged in the mathematical aspect of economics. The leading figures there were arrested, the Institute closed down (and Kondratiev shot in 1938). The general atmosphere in the nation began to worsen essentially in 1927.

Konüs found himself unemployed and was only able to return to successful work in 1945. Not long before his death he explained to me that he had escaped virtually unscathed because, although labelled “apprentice saboteur” (Komlev 1991), those responsible decided that they could have hardly expected anything from a mathematician. While working at the Institute, Konüs attempted to combine the Marxist system with the theory of marginal utility (Komlev).

The author published his paper in a statistical source and, consequently, referred to a statistician (§ 1), elucidated to a certain extent the axiomatic approach to probability (§ 2) and inserted many passages from mathematical works. On the other hand, his most important inference, the alleged justification of the frequentist theory (§ 6), was first and foremost

mathematical. I suspect that Konüs either did not submit his work to any mathematical periodical at all, or had it (its initial version) rejected.

Indeed, no such justification is achieved even nowadays, see Khinchin (1961). Then, Uspensky et al (1990) discussed the most difficult related problem of defining randomness and concluded, in their § 1.3.4, that it seemed “impossible to embody Mises’ intention in a definition of randomness that was satisfactory from any point of view”.

Too late have I found out that the author had also published a certainly related paper (Konüs 1984) and regret that it is hardly available.

I
M. V. Ptukha

**Sampling Investigations of Agriculture in Russia
in the 17th and 18th centuries**

Uchenye Zapiski po Statistike, vol. 6, 1961, pp. 94 – 100

[1] The application of sampling for a deep study of phenomena is spreading ever wider. The theory of probability provides the theoretical principles of sampling, and Laplace was the first to express this idea in connection with a sampling census of the population of France timed to coincide with 22 September 1802, the New-Years Day according to the Republican calendar.

A wide theoretical study of sampling has begun at the end of the 19th century; in 1895, during the fifth session of the International Statistical Institute, Kiaer (1896) initiated its discussion, but a decision was postponed until the next session, then postponed once more. In 1901, the eighth session had recommended statisticians to turn attention to it, but in 1903 debates resumed. Nowadays, a single viewpoint concerning the numerous types of sampling for all cases of applying them and acceptable for all nations does not yet exist.

The logical idea of sampling as a method of judging about the entire totality of objects and phenomena after being acquainted with its part, has been long since inherent in man. When in urgent need, people invariably turned to sampling and have been developing practical methods of its application for establishing approximate values of the parameters of a totality.

[2] The cause for the origin and spread of sampling in agriculture in Russia was the long ago established system of farming large estates belonging to monasteries, nobility, state institutions and of managing royal votchinas. Owing to the shortage of manpower and large storehouses for a great volume of grain, only a comparatively small part of the harvested cereals was being threshed at once with their main volume collected into ricks and threshed as needed during the time free from seasonal work.

For expediently farming a large estate it was necessary to know the entire volumes of the harvest of cereals and of the distribution of the unthreshed remainder stored in great ricks. The landowners distributed the yield into a currently needed part, the seed-fund and the part to be sold. The development of the commodity economy ever more insistently compelled the landowners to establish the volume of the grain as precisely as possible beginning with the harvesting and until its complete exhaustion. Indeed, cereals provided the main inflow of money.

The application of sampling assumes the existence of a system of statistical stock-taking. And we may suppose that in Russia, before Peter the Great, a regulated statistical system had been first of all existing in monasteries. They possessed great estates with a numerous dependent population engaged in agriculture, handicraft and trade. For a long time monasteries had been the centres of enlightenment and culture and the monks were certainly able to take stock. The influence of the Byzantian

culture possibly somewhat fostered the establishment of a system of statistics in the monasteries; Byzantine ideas of management could have been transferred there by the ecclesiastical legislation.

There exists a lot of writings more or less devoted to studying the economic activities of our ecclesiastical institutions, but no summary information on the organization and development of statistical registration in these bodies is in existence. The common deficiency in these writings is that they incompletely and only in passing treat the problem of registration and, specifically, they do not dwell in sufficient detail on statistical methodology.

Much better is the situation with ascertaining the sources for studying the *votchinas* in the 17th and 18th centuries. The sources pertaining to the mid-17th century testify that sampling had been ordinarily applied there as something usual and well known. It may be therefore assumed that that procedure goes back to the 16th century and that it originated in large monasteries. I expressed this idea in 1945, and now it had already been somewhat confirmed.

Instructions of the 18th century essentially differ from the *orders* of the *votchinniks* to their stewards of the 17th century. Even the best orders did not embrace the life of the estate in its entirety, of its management in full. And the contents and even the format of the instructions of the 18th century favourably reflect the spirit of the reforms carried out by Peter the Great. In that century, handicraft and trade became isolated and formed independent branches of the national economy, and commodity-market circulation had been developing. The general need for statistics is far from being as insistent for a natural economy as it becomes after the elements of that circulation permeates it.

The economic well-being of a landowner in the 18th century more strongly depended on the market than it was for a *votchinnik* in the previous century. The landowner had to keep more records of various kinds describing his economy, the economy and personal data of his serfs, he had to calculate and estimate more than his predecessors. The economies of large estates already demanded its own statistics.

[3] Sampling in the *Votchinas* of B. I. Morozov

The Boyar Morozov (1590 – 1662), the tutor of Czar Aleksei Michailovich, applied a properly developed and thought out statistical system (*Akty* 1940). One of the most interesting pertinent documents contains information on the statistics of crop yield and discusses a primitive form of sampling. In July 1648 Morozov sent out instructions to the stewards of [a certain village of] Nizegorodskaia province [...] on harvesting and sowing of cereals in proper time and on keeping books on reaping, threshing and sowing. The instruction says (p. 100):

After storing the winter and spring-sown grain, enter in the reaping books how much is there in a sotnitsa and of what quality is it, and how much quite good, average and bad grain will there be from a sotnitsa after threshing, and send the experimental books [registers of sampling]. And if you will soon complete the threshing of all cereals, send the reaping, threshing and sowing books to Moscow without delay.

Special attention is called here by the expression *experimental books*. The context indicates that such books had been in general use so that the term did not need any explanation. These books indicate that in Morozov's votchinas sampling reaping and threshing had been applied, and, for that matter, for three different grades of grain.

Experimental books were compiled immediately after harvesting and sent to Morozov living in Moscow.

Threshing of all the harvest could have been ended in a short time, and it was in such cases that Morozov demanded that all the three books be sent. I think that such a system of agricultural statistics indicates deep theoretical thought. Because of one or another circumstance much unthreshed cereal could have been left in ricks and barns, and having experimental books and the other ones, the votchinnik was able to estimate both the gross yield to a certain measure of precision and how much grain was still left in the unthreshed cereals.

[4] Sampling in the Economy of A. I. Besobrasov

Novoselsky (1929) had studied the archive of the stolnik Bezobrazov who owned votchinas of average size; the period studied was 1670 – 1680. On p. 104 Novoselsky writes:

Many sowing, reaping and threshing books, although not at all for each year and not for all the villages, are extant in Bezobrazov's correspondence with his stewards. Often the final figures for sowing and threshing are lacking and only provide information for the time when sowing and threshing had not been completed. It is not always possible to translate this information into an intelligible language. Yearly summaries for the estates are not in existence anymore, possibly they did not exist at all.

[5] The Instruction to Managers of Dvortsovaia Volosts of 1731

Many instructions not more in existence or not yet found had been certainly compiled in the first half of the 18th century for the managers of estates. However, their availability does not yet prove that they had been actually applied. The peasants' resistance to the landowners' intent on registering their economies means that some extant instructions testify to statistical thought rather than to actual investigations.

Of special significance are therefore the instructions on sampling carried out by state bodies. They are both specimens of statistical thought and descriptions of practice.

The *Instruction of 1731* is of much interest because it regulated the management of the economy of the dvortsovye peasants over all Russia. The first such instruction dates at 1725, but it was of a general nature and contained no section on cereals mentioning sampling.

The *Instruction of 1731* (Volkov 1951, pp. 176 – 178) however states:

7) Describe exactly after examining and estimating how much threshed grain is left for seeds in barns and how many in ricks as measured by their size. Compile special threshing books and books registering quantity and enter how much quite good, average and bad grain will be threshed out of each sotnitsa separately, also out of the ricks and compare everything with previous reaping and threshing lists...

9) *When the sown cereal of any kind becomes perfectly ripe, the manager will order the peasants to reap and collect it into shocks of one and the same number of sheaves. [...] And to order the peasants to cart the cereal for storage [Stipulated here is registration of quantity by tselovalniks both when collecting and storing the cereals.]*

When storing the cereal a sheaf from each cart should be taken to a special barn. And how much was carted, from how many ricks, and the length and width of each of these in sazhen, and how many shocks in a rick and sheaves in a shock were there should be entered in reaping books.

And from the separated cereal the manager himself with village elders, threshing-ground tselovalniks and elected best peasants and village priests will take a hundred sheaves each of quite good, average and bad cereal, store them for drying on [?] a barn, and seal it up. After the cereal dries off, order the peasants to thresh it, measure the pure grain and the chaff to a quarter of a chetverik and write down in experimental books how much is threshed and measured out of each sotnitsa and order the present village elders, tselovalniks, peasants and priest to sign them. And let those people seal up this experimental grain in a special barn. Measure the grain separated by winnowing, store it in a convenient place and estimate how much grain should there be after threshing. Send the experimental reaping books together with the estimates to the Chief court office in September and October leaving exact copies countersigned by the manager in the department huts.

The new manager [?] measured the threshed grain anew and determined by sampling the probable quantity of grain in the ricks. Elected peasants (village elders, collectors [?], tselovalniks) played an important role here. The Instruction indicated that they were to be elected yearly from the best peasants “in turn from those paying taxes without the help of the village community”.

These people did not pay taxes, nor did they participate in obligatory work which other peasants had been doing instead of them. The Instruction thus solved the complicated problem concerning the bodies of statistical observation and control.

In 1734, about 730 thousand peasants, i. e. more than 6% of the entire tax paying population of Russia, had been living in dvortsovye volosts.

[6] The Stable Regulations Or Statutes of 1733 (*Vtoroe* 1834, No. 6349, pp. 53 – 63)

In all probability, the author of the remarkable statistical documents concerning the Stable Office was Arseny Petrovich Volynsky (1689 – 1740). The government resolved to breed horses for the army at state stud farms. Entire towns (or administrative units), slobodas, large and small villages had been ascribed to that Office. Volynsky, appointed its director in 1732, had chosen them after a detailed statistical and economic investigation.

The *Regulations* certainly drawn up by Volynsky were published on 16 March 1733. The most interesting place there is the description of yearly sampling for studying agriculture in the localities ascribed to the Office. The principles of typological selection are exhibited there clearly and properly and the problem of studying [estimating?] the sown area and the

crop yield is thought out in a broad sense and solved in a surprisingly simple way. Only an exceptionally able man could have formulated so clearly the method of applying sampling in agriculture, and Volynsky indeed was such a person who back in 1724 had compiled a remarkable instruction for the manager of his votchinas. [And here is a passage from the Regulations (p. 54):]

For a better acquaintance with, and establishment of a proper economy it is necessary to send yearly groom and backyard [?] stable men from the Office at the time of ploughing, sowing, reaping and hay making to all the ascribed towns, slobodas, volosts and villages selecting these places in a way suitable for them; order them to test everywhere the arable land by ploughing several dessiatinas of quite good, average and bad land and sowing winter and spring-sowing cereals. Then, during reaping, to reap [both kinds of cereals from all the three grades of land]; to estimate how many dessiatinas of unreaped cereals there is separately [for the three grades of land; then thresh the samples, measure the grain, write it all down, sign the compiled document and have it countersigned by the manager. The same is to be done with hay, only no mention is made here of quite good etc].

In connection with this description of the widely applied sampling I ought to mention three circumstances. 1) It goes back to the 17th century. 2) The probable yields of cereals and hay had been finally estimated at the Stable Office itself. 3) Sampling as described in the *Regulations* had been included in the *Complete Code of Laws*... and could have been widely applied in Russia.

In the 17th and 18th centuries (until 1775) a certain role in the statistical system had been ascribed to the Voevoda (Zaozersky 1937, pp. 53 – 54):

The Voevoda is naturally mostly responsible for an exact execution of the instructions from the capital and therefore for an accurate account of managing and conducting the economy.

The department in charge of the royal votchinas wished to know (Vtoroe 1841, pp. 39, 41)

How much grain is left from previous years, both threshed and unthreshed, how much cereals and of what kind was sown in the year under review, [...] how much was reaped and how many sotnitsas and how much should be expected according to the experimental threshing, and how many haycocks there is and will be from unmown wasteland according to estimation.

Answers to all these questions indeed form the contents of the so-called sowing, reaping and mowing books (Zaozersky 1937, pp. 53 – 54):

Sending partly precise (concerning everything – M. P.) and partly conjectural (experimental) figures pertaining to the entire economic activity to the department, the Voevoda had to narrow extremely the sphere of his unpunishable professional idle profit and at the same time he

was compelled to keep a vigilant watch on the proper execution of everything under his administration.

Such was the statistical system of agriculture in the royal votchinas.

[7] In the 18th century ecclesiastical institutions had been owning great territories with a considerable population attached to them, and the temporal power aimed to seize a part of the clergy's material resources for general state needs. For this reason elements of the statistics of ecclesiastical possessions were created during the reign of Peter the Great; by the mid-18th century that statistics became established in its final form.

On March 6, 1740, an instruction for the members of the Board of economics, one of the then existing governmental organs (*Vtoroe*, see 1841, No. 8029, pp. 39, 41), called *On the Inspection of Bishop and Monastic Votchinas*, with detailed indications on registering all the productive forces and profits of each of them. Its § 2 demanded to compile

A list providing the quantity of cereals of each kind in the granaries of the inspected votchina, the same concerning unthreshed cereals left in ricks and barns and how many shocks or sheaves are there in each and how many grain should there be after threshing out of a shock or from a hundred sheaves.

The instruction (§ 23) precisely indicates how to estimate the probable yield by sampling:

Immediately after harvesting, each kind of cereal should be sampled in the presence of the steward, the village elder and the elected person (if in existence) and to be conducted thus. One sheaf is to be taken from each cart; and, after all the cereal is transported, these sheaves should be threshed and the quantity of grain from a hundred entered into the threshing books. According to that experiment, the entire yield of cereal should be estimated and the reaping books and books of experimental threshing signed by yourself and the steward

and delivered to the Board.

The expression *from a hundred sheaves* ought to be understood as the mean not of all the selected sheaves, but only of the given [?] hundred of them.

[8] In the second half of the 18th century, sampling had been applied in large estates for estimating yields and stocks of cereals and forage. The Free Economic Society established in 1765 aimed at fostering the economic and agricultural development of Russia and worked out a specimen of an order to managers of estates. In 1770 it awarded prizes to and published two such orders compiled by Rychkov (1770) and Bolotov (1770).

Petr Ivanovich Rychkov (1712 – 1777), an educated and knowledgeable person, compiled a specimen list of questions pertaining to separate homesteads and a detailed instruction comprised of 17 sections. Well informed about the practice of managing large estates, he also expressed his opinion about sampling. In his § 13 (1770, pp. 41 – 42) he wrote:

Since they ought to estimate the yield of cereals, or the approximate quantity of grain, they must order those selected and village elders that each peasant, when transporting cereals to the threshing ground, should load one and the same number of sheaves onto his cart. How many sheaves are stocked in each barn should be written down in the votchina's books, and, to know the quantity of each grade of grain yet more precisely, separate sample threshing of the best, the average and the bad [sheaves] should be carried out during harvesting in the manager's presence. Having written down the results, an estimate should be done and, come the new year, the landowner should be notified so that he will know how much grain does he have.

Andrei Timofeevich Bolotov (1738 – 1833) was one of the most eminent agronomists of his time. His specimen of an order is essentially more detailed and concrete, more systematized and complete than Rychkov's. He closely connected sampling with regulating the management of large estates, planning agricultural production and sale of their surplus (1770, pp. 126 – 127):

A good steward must think about the future and about a better use of the yields even when storing the produce. He ought to estimate beforehand how much will he have of everything and how much will be left for selling after expenditure. He needs to know therefore as soon as possible about threshing so that a small amount of each kind of quite good, average and bad quality of cereal should be left for experiments. For more convenient calculations these three grades should be collected into different ricks rather than mixed up and what grade of cereals came from which dessiatina and in which rick were they collected should be registered in the cereal book. Then, after threshing, it will be possible to estimate how much grain will there be of each grade.

[...] Bolotov's *Order* is a remarkable attempt to embrace the investigation of the productive forces of a large estate by a unified statistical system. The empirical types of sampling for estimating the entire yield of cereals and forage, methods of periodical censuses of homesteads and agricultural statistics pertaining to the current situation worked out during many years, were later critically applied by the Russian zemstvo statistics.

Bibliography

Akty (1940), *Akty Khoziastva Boiarina B. I. Morozova* (Acts of the Economy of the Boiar B. I. Morozov), pt. 1. Moscow.

Bolotov A. T. (1770, in Russian), An Order to manager or steward on the manner of ruling villages in the owner's absence. *Trudy Volnoe Ekonomich. Obshchestvo*, pt. 16, pp. 69 – 230.

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

Kiaer A. N. (1896), Observations et expériences concernant des dénombrements représentatives. *Bull. Intern. Stat. Inst.*, t. 9, pp. 176 – 186.

Novoselsky A. A. (1929), *Votchinnik i Ego Khoziastvo v XVII Veke* (Votchinnik and His Economy in the 17th Century). Moscow – Leningrad.

Rychkov P. I. (1770, in Russian), Order to manager or steward concerning proper maintenance and management of villages in the owner's absence. *Trudy Volnoe Ekonomich. Obshchestvo*, pt. 16, pp. 13 – 68.

- Seneta E.** (1985), A sketch of the history of survey sampling in Russia. *J. Roy. Stat. Soc.*, vol. A148, pp. 118 – 125.
- Stigler S. M.** (1977), Eight centuries of sampling inspection: the trial of the pyx. *J. Amer. Stat. Assoc.*, vol. 72, pp. 493 – 500.
- Volkov S. I.** (1951, in Russian), Instruction to managers of dvortsovy volosts, 1731. *Istorich. Arkhiv*, vol. 6.
- Vtoroe** (1834), *Vtoroe Polnoe Sobranie Zakonov [Rossiiskoi Imperii]* (Second Complete Code of Laws [of the Russian Empire]). Petersburg.
- Zaozersky A. I.** (1937), *Czarskaia Votchina XVII Veka* (The Czar's Votchina in the 17th Century). Moscow. Second edition.
- Zarkovich S. S.** (1956), Note on the history of sampling methods in Russia. *J. Roy. Stat. Soc.*, vol. A119, pp. 336 – 338. Reprinted: Kendall & Plackett (1977, pp. 482 – 484).
- (1962), A supplement to Zarkovich (1956). *Ibidem*, vol. A125, pp. 580 – 582. Reprinted *Ibidem*, pp. 486 – 488.
- You Poh Seng** (1951), Historical survey of the development of sampling theories and practice. *J. Roy. Stat. Soc.*, vol. A114, pp. 214 – 231. Reprinted: Kendall & Plackett (1977, pp. 440 – 457).

II A. A. Chuprov

The Problems of the Theory of Statistics

Die Aufgaben der Theorie der Statistik. *Jahrbuch f. Gesetzgebung,
Verwaltung und Volkswirtschaft im Deutschen Reiche*,
Bd. 29, 1905, pp. 421 – 480

[0. Introduction]

[0.1] After having a look at those sections of statistical manuals that have to do with the theory of statistics, it is impossible to deny that statisticians have a very peculiar notion about it. Instead of a system of propositions we find a medley of pronouncements of most diverse contents lacking in any connections between them. Advice based on experience about rules which ought to be obeyed when studying large totalities; instructions on summarizing and publishing the collected numerical data; formulas for fostering the determination of statistical regularities given such data but justified by personal tastes rather than by generally recognized principles.

And, along with all this, and devoid of any theoretical foundation, deliberations on the concept of mass phenomena; on the essence of statistical regularity; on contrasting the statistically knowable and the typical and its close relations with the probable; isolated technical directions obeying no theoretical principles and even entirely independent from them – this, in short, is the picture that can compel any theoretically minded researcher to abandon statistics.

The theory of statistics had only looked so hopelessly during the life of the previous generation, from the time when the attacks of German scientists annoyed by the exaggerations made by Quetelet's followers demolished his elegant building of "social statistics"¹. Still, however Quetelet's constructions seem now intrinsically contradictory, his social physics had nevertheless been a powerful scientific theoretical system. In its philosophical viewpoints, his system rested upon the Comte positivism² and methodologically applied the newest mathematical achievements. Together with Quetelet's doctrine of mean values and *Average man* and his *inclinations*, with the dependence of the characteristics of that being on the conditions of the *social system* and the reduction of statistical regularities as well as the methods of their study to the theory of probability, – that system had been unique in its peculiarity over the entire history of statistics.

[0.2] After having defeated *Queteletism*, the investigative ardour of the new generation had been directed towards particular concrete research rather than to erecting a building of the theory of statistics from the ruins of the demolished system. The immediate achievements were very substantial: an immense volume of numerical material of most various essence has been gradually amassed; and, at the same time, statisticians have become able to treat their data so precisely as was hardly possible earlier.

However, more remote results of that turn of events proved not really comforting: the more observations had been collected, the more refined had the statistical art becoming, the keener felt was the absence of a theoretical base. Nowadays, statistics is almost lacking in generally accepted principles on whose foundation the correctness of conclusions made and the expediency of the methods applied could have been established. Debators differing in opinion find themselves in a helpless situation. It is impossible to compel the opponent to acknowledge defeat and as a rule the controversy remains undecided although points important in various aspects are [often] connected with its outcome.

The bias of contemporary statisticians towards pure empiricism ought to be done away with if only to restrict the researcher's discretion and the results of statistical studies to become obligatory and thus acquire that most important property of science. Until recently, most statisticians forthwith rejected the demand for a theoretical base as a dream of no practical importance.

[0.3] However, during the recent years the theory of statistics gained many partisans, especially in England³ and it seems that we may hope that the next turn of events is not far off. This is all the more true because a researcher, the only one among his contemporaries not satisfied with the easy victory over Queteletism, has been labouring all his life to erect the theory of statistics anew, has become head of the school of theoreticians, and has recently published a contribution (Lexis 1903) serving as a code of sorts for statisticians of his direction. There, "essential principles of population statistics"⁴ stated by a master are provided in an easily accessible form for all advocates of theoretically based statistics.

Although Lexis entitled his latest writing (1903) *Contributions to the Theory of Population and Moral Statistics* rather than *Theory of Statistics*, it contains all the essential elements of the latter; his choice of title was apparently conditioned by the nature of exposition. The book includes papers which Lexis had written over 30 years; he had reworked the material and took into account the published investigations made by other scientists under his influence. However, he only included those so far as they did not go out of the subject matter of his own previous papers, otherwise he only considered them more or less thoroughly in footnotes.

Thus, the book primarily reflects his own scientific work, but since his research is the starting point for all the sources of the development of the modern theory of statistics, his contribution contains the elements from which that theory is composed. Only some rearrangement of separate parts is required for the entire theoretical system to become clear.

[0.4] And it is indeed my aim to trace those problems of the theory of statistics and the means at the disposal of the statistician for constructing it as called forth by the work of Lexis and the researchers who align themselves with him. It is necessary [however] to indicate first of all a digression [my digression] from the Lexian statistical viewpoint based on the principle perhaps explained by his responding to the excessively naturalistic way of thought of the adherents of Queteletism, – to the idea that "man's behaviour is in essence individual and falls out entirely beyond the bounds of natural regularity". This Lexian thesis does not touch on the inferences from those investigations made by him which seem important for statisticians, cf. Bortkiewicz (1904, p. 241n1), and

even his closest students hardly share it: statistical methods are being ever widely applied in natural sciences when studying those phenomena the existence of whose causes is unquestionable and these Lexian notions are absolutely insufficient for the creation of a theory of statistics.

When wishing to exclude the creation of two theories of statistics, one for the social sciences, the other one for natural sciences, the results of the Lexian investigations ought to be as though transposed to another philosophical key. In reducing statistical regularities to constructions of the theory of probability Lexis perceives a replacement of the impossible detailed explanation of the causal connections of phenomena. On the contrary, we ought to recognize that the theory of probability itself should become a means for establishing these connections. New contributions in the field of logic of the probable, and I especially mention the works of von Kries⁵, so closely linked that theory with causality that it does not seem anymore difficult to throw a bridge across the gulf that separates them.

1. The Inductive and the Statistical Methods

[1.1] The behaviour of man is based on recognizing that everything in the world is going on regularly and that like causes bring about like effects⁶. We have only to imagine what our life would have been if this assumption did not hold, if one and the same set of phenomena led now to one corollary, now to another one; if water sometimes extinguishes fire, and sometimes catches it, if heavy bodies either fall down or fly up. We would be able to live even in such a world of randomness but only [...]

However, nature's general regularity is not sufficient for serving as a basis of our conduct; that abstract pattern ought to be filled up with definite subject matter. We must know precisely which effects follow from a given set of phenomena and what kind of circumstances lead to a concrete action. Only after availing ourselves with such knowledge we will be able to influence actively the processes going on in the universe and to fulfil our plans⁷.

The realm of man's knowledge coincides with the sphere of his power because ignorance of causes deprives him of action. Nature can only be subdued by resigning to it and causality revealed by contemplation becomes the guiding principle of our activities.

Knowledge is power. We partly acquire such kind of knowledge instinctively by building up the so-called experience of life. A child learns to walk without thinking about the laws of equilibrium of solids, and for learning to swim it is not necessary to acquaint oneself with the results of the Archimedean investigations.

This path is however very prolonged and difficult. A lot can be thoroughly learned in the school of life, but too much will depend on chance, and too much and too often we will dearly pay for our knowledge. [...]

Those sciences which, as contrasted with descriptive sciences, can be called *nomological*, take upon themselves the aim of systematically studying the definite connections between phenomena. Logic is called upon to show the way which they ought to follow and therefore by issuing

from the assumption of nature's regularity to treat scientifically the material collected by experience.

[1.2] Experience establishes that in a given locality phenomena A_1, B_1, C_1, D_1 follow after phenomena A, B, C, D . Granted the assumption mentioned above, it is possible to conclude (under some premises which I leave aside for the time being) that the same will happen always and everywhere. However, for isolating those elements of the initial set which are [really] necessary for the appearance of definite phenomena belonging to the subsequent set, it is necessary first of all to modify somewhat and to develop our ideas of regularity. We must assume that not only does the general state of the universe determine its state in the next moment (**p. 426**)⁸, but that phenomenon A is connected with phenomenon A_1 in such a manner that the occurrence of A is always followed by A_1 and that A_1 never happens without being preceded by A . [...]

It can be shown that, if a number of initial sets has only one common element A , and the corresponding series of subsequent sets has only one common element A_1 then phenomena A and A_1 are connected as cause and effect. [...] Such patterns [...] are called inductive methods. [...]

On the face of it, it seems that they are quite sufficient for discovering the laws of nature; [...] actually, however, only for constructing abstract schemes but not for studying concrete phenomena. The correctness of our inferences rests on our tacit assumptions which very rarely come true. The first assumption is that the set of observed phenomena ought to be exhaustive. [...] [which can be wrong.]

Other difficulties also exist [...] and especially important is Mill's forcefully stressed plurality of causes. [...] I will try to prove, however paradoxically it seems, that plurality of causes does not really exist in Mill's sense, but that nevertheless a researcher ought to invariably allow for plurality of both causes and effects. [...]

[1.3] The set of methods for investigating causal connections under plurality of causes and effects comprises exactly what is usually collectively called *statistical method*⁹. In formal logic, the statistical method thus understood is earning a position of equal rights with inductive methods. Both have the same goal, viz., to discover the laws of nature by scientifically treating observational material, but each is only being applied when their special assumptions are holding. If we may state that the investigated phenomena are inseparably causally linked, then the various kinds of induction can be made use of. If however we are unable to prove that plurality of causes and effects will not be encountered, we ought to apply one or another form of the statistical method. These two methods thus supplement each other; when one is found unfit, the other one replaces it.

2. Notion of Probability. The Law of Large Numbers

[2.1] The logical purpose of the statistical method elucidated above directly conditions the goals of the theory of statistics. Similar to the theory of induction which develops its methods of likeness and distinction for studying inseparable causal ties, the theory of statistics has to create special methods for revealing incomplete causal connections, for indicating how to treat observational material when there exists no inseparable link between cause and effect.

In spite of its practical importance, the theory of statistical method regrettably lags far behind the theory of induction. On one hand, this backwardness is caused by the greater variety and complexity of those forms of causal relations which are studied by the statistical method; on the other hand, until recently that method had only been applied to study mass phenomena of human life.

Logicians had been therefore inclined to deny any general significance of the statistical method in the same way as for example the methods of microscopic investigations. I have indicated however that until recently statisticians [themselves] little cared for thoroughly justifying their methods. Only after Lexis had investigated the stability of statistical series and the substantiation of statistics by the calculus of probability, thus showing the possibility of elevating oneself above naked empiricism, did they begin to become more interested in statistical theory.

The erection of the theory is not yet completed in all its details, but after Bortkiewicz, Edgeworth, Kries, Pearson, Westergaard, to mention but a few most outstanding names, had joined Lexis¹⁰, the essential features of the theory of statistics showed through sufficiently clearly and an attempt to trace the general outline of the arising building's façade can be made.

The field of the statistical method's application, as I have sketched it, consists of investigating incomplete and not inseparable links so that the causes and effects can bring forth other effects and be conditioned by other causes respectively. This essentially negative definition is however not in general sufficient for constructing methods of statistical investigations. It is necessary to distinguish and define more exactly the kinds of causal links to be studied. Most important is the kind having plurality of effects which therefore can be characterized by the pattern $\alpha, \beta \rightarrow \alpha_1, \beta_1, \mu_1, \nu_1$.

[2.2] We will use it for considering in more detail how a statistical investigation is carried out¹¹. First of all, we ought to define precisely the connection between cause and effect in a manner fit for application as the base of statistical methods. [...]

The number of various possible effects following a given cause immediately suggests itself as the foundation of this definition. When throwing a coin, two effects are possible; when throwing a die [...]; when extracting a ball from an urn containing a hundred various [?] balls, a hundred different effects is possible. In many cases, the integers 2, 6 and 100, or, better, the fractions $1/2$, $1/6$ and $1/100$ can clearly indicate the closeness of the link between cause and effect.

This principle really serves as the foundation of the statistical methods, but it ought to be specified. The number of possible effects conditioned by one and the same cause is not definite in itself because it depends on the choice of the unity. Suppose that the urn contains 99 black balls and one white ball. Then the connection of the cause and the extraction of the white ball is characterized by the fraction $1/100$. If however we study the appearance of a ball of one of the two colours, the connection of that effect with the cause will be defined by the fraction $1/2$. For our goal of defining causal relations such multivalued and unstable indications are useless.

The notion of equally possible effects following some cause comes to the rescue¹². [...] After establishing the set of equally possible effects

corresponding to some cause, their relation to the latter can be defined uniquely [...] An effect not included in that set can be subdivided into a number of equally possible effects. [...]

For example, the connection between extracting a black ball from the urn mentioned above [with its cause] is represented by the fraction 99/100. Such fractions which are sometimes expressed by irrational numbers (p. 441!) are called objective probabilities and they uniquely determine the connection between cause and effect. That probability does not immediately supplement our knowledge, it only expresses it by a single number. [...]

[2.3] The notion of objective probability is methodologically justified by the law of large numbers which establishes the relation between it and the empirically discovered frequencies of events. [...] All statistical inferences rest on that law. [...] Even political arithmeticians had divined it; thus, Halley indicated that the errors from which his tables of mortality were not free, were partly random (p. 442)¹³ and could have been eliminated had the number of years of observation been much larger (20, say, instead of 5). The Dutchman 'sGravesande presented the same idea in a more general form¹⁴:

The regularity that often escapes us when considering a small number of conclusions reveals itself if a large number of them is taken into account.

Süssmilch illustrated the influence of large numbers by the sex scatter of population. For separate families and in small settlements the ratio of the sexes fluctuated absolutely irregularly, but in large cities and entire countries it was constant.

The law of large numbers acquired its precise mathematical definition in the works of the great theoreticians of the science of probability, J. Bernoulli, Laplace [?] and Poisson¹⁵. However, its real substantiation was provided by the talented French mathematician, philosopher and economist Cournot, who for a long time had remained in oblivion.

His proof consists of three parts. First of all, referring to his notion of mutually independent elementary causes and effects [§ 39], he inferred [§ 43] that phenomena whose objective probability was very low, cannot occur often. The second, mathematical part of the proof adjoining the first one is in the forms of either the well-known Bernoulli theorem or in its generalized forms of Poisson or Chebyshev. In essence, these theorems reduce to proving that the objective probability of obtaining frequencies of a separate kind of effect following some cause essentially deviating from their corresponding probabilities is very low given a large number of observations and can be made lower than any small magnitude by increasing the number of observations.

In itself, this proposition is a combinatorial theorem which does not lead us beyond the sphere of probability and nothing testifies about the behaviour of phenomena in reality. But, coupled with the earlier established relation between low objective probability and essential rarity, this mathematical theorem leads to the third part of Cournot's deliberations, to the conclusion named above the law of large numbers which tells us that for large statistical totalities essential deviations of the

frequencies of the various kinds of effects occasioned by a common cause from proportionality to corresponding objective probabilities occur very rarely¹⁶.

Which numbers of observations ought to be considered large, and what should we understand as an essential deviation, and under which kinds of premises does the main proposition hold, i. e., what are the demands which the totality of phenomena should obey, – we find all this in the mathematical formulation of the second part of Cournot's proof. They establish the connection between the number of observations and the degree of precision with which frequencies reflect their underlying probability.

[2.4] Denote the probability of any event by p , the number of observations constituting the totality by M , and the number of occurrences of the event by m . The frequency of the event will then be m/M and the probability $f(u)$ that the difference between p and m/M will not exceed

$$u\sqrt{\frac{2p(1-p)}{M}}, \quad (1)$$

or, if expressed in a more convenient form,

$$u\sqrt{\frac{2m(M-m)}{M^3}}. \quad (2)$$

The values of $f(u)$ corresponding to different values of u can be found in tables included in many statistical contributions, for example in Lexis (1903, pp. 252 – 253)¹⁷. These tables show that as u increases, $f(u)$ rapidly approaches unity. Thus, at $u = 3$, $f(u)$ is already 0.99998. We may therefore be very sure that the difference between p and m/M will not exceed

$$3\sqrt{\frac{2m(M-m)}{M^3}}.$$

[...] The magnitude [the square root in (1) or (2)] is called *modulus*¹⁸ and the result obtained can also be thus stated: The difference between the objective probability and frequency does not thrice exceed the modulus. And since the modulus becomes smaller as m increases, we conclude that [...]. This indeed is why statisticians invariably aim at large numbers.

The law of large numbers makes it possible to pass from frequencies derived from trials on to the objective probability which constitutes their foundation and therefore to make inferences about the kind of connection of the [studied] phenomenon with the general causes defining the totality. [...] If, for example, a proper die the occurrence of whose faces has probability 1/6 is thrown 100 times, no face will turn up more than

$$100 \left[\frac{1}{6} + \sqrt{\frac{2 \cdot \frac{1}{6} \cdot \frac{5}{6}}{100}} \right] = 32$$

times. And if any face occurs oftener, we may state that the die is faked.

Had the roulette wheel in Monte Carlo been quite properly constructed, a change of colour would have occurred not oftener than 2236 times in 4274 rounds; on the contrary, the same colour would have remained twice and thrice in succession not less than 984 and 470 times respectively. However, Professor Pearson (1894, p. 189), drawing on observations of a series of rounds, showed that repeated colours only happened 945 and 333 times so the construction of the wheel had something leading to an oftener change of colour and its lesser repetition than it should have been¹⁹.

This case is practically important because probability $p = 1/2$ [how does it appear here?] plays an essential methodological part: it means that the considered phenomena are independent since each was observed equally often whether the other one was present or not. In a great majority of cases all that the statistician has to show is that the studied phenomena are not mutually independent.

For proving it, he can compare the series (see below) and the method of applying the law of large numbers just considered is the base of this, practically speaking, most important scientific method at the disposal of statisticians. The French statisticians Guerry and Dufau had empirically developed it in a most possible systematic way.

Contrary to this method which refutes independence there is another method of affirmatively applying the law of large numbers. Reliable conclusions from statistical observations, without their having been scientifically treated and therefore [still] based on empirical frequencies, cannot be immediately reached. For understanding this, scientific training is not needed. If in some locality 12,345 people had died in a year, and 12,344 in the next year, it will be clear to each sensible person that it does not mean that sanitary conditions had improved. The difference is so insignificant as to having possibly [?] been caused by chance. But what is chance in this context? Where does it end? When do the differences become sufficiently large for us to be able to say that they are not accidental?

If the answers to these questions are left to the statisticians' personal feelings, a sceptic will overrate the bounds of the accidental, and an optimist, underrate it. When aiming, however, to discover a precise and obligatory method, the answer can only be found in the law of large numbers. [...] If we can show that for two statistical totalities the objective probabilities of some event, of death for example, are [indeed] different, we will have to conclude that the general causes [of death] in these two cases are [also] different.

Although objective probabilities are reflected in empirical frequencies, this reflection is not sharply outlined; depending on the number of observations comprising the totality, it always remains more or less indefinite. Different frequencies can correspond to the same probability, and vice versa, which explains why empirical frequencies cannot directly tell us whether the compared statistical totalities were conditioned by the same general causes. Consideration of probabilities must always be a link in the chain of our deductions. Only when able to prove that the objective probabilities are different in these two cases, we may infer that the general causes differ as well.

But how to prove that the probabilities are different? There are two possibilities. First, drawing on the relation between probabilities and frequencies in each case, which is determined by the law of large numbers, we can establish the boundaries within which these probabilities are included. If these boundaries do not intersect, they, the probabilities, obviously cannot be equal.

Thus, in Vienna in 1897 33,181 inhabitants out of 1,558,129 died. [...] In Prague, the same year, 6,392 and 193,097. [...] We may state that [...] in respect of mortality the conditions of life in Prague were less favourable than in Vienna. But what in essence does this depend on [...] we cannot say at once. [...]

The other possibility is preferable because it restricts randomness more tightly. Mathematical theory indicates that, when comparing two frequencies with moduli m_1 and m_2 , their difference has modulus $\sqrt{m_1^2 + m_2^2}$. If therefore the actual difference exceeds thrice that modulus, we cannot anymore say that the frequencies randomly deviate from the same objective probability; on the contrary, we are then justified in stating that the general causes of the two cases were different.

3. Theory and Experience: the Study of Stability (Lexis) and the Law of Small Numbers (Bortkiewicz)

[3.1] Raw empirical frequencies established by statistical observation can be scientifically treated and the essence of phenomena cleaned up from the dross of the accidental. However, before taking the next step in constructing statistical methods and ascertaining how to apply the conclusions about the main causes, it is necessary to show that the deductively justified notion of the connection between empirical frequencies and objective probabilities is a construction not made up of thin air, but, on the contrary, substantiated by experience.

Do conditions exist under which it is really possible to show that connection? And, if so, are the circumstances of given trials of that kind [...]? Only after answering both questions affirmatively further construction of the theory of statistics makes sense. The first question can be answered by referring to experiments on games of chance. Carried out thoroughly, [...] they utterly correspond to preliminary theoretical calculations. [...]

In accord with a proposal made by De Morgan, a coin had been thrown until 2048 heads had appeared; tails had come out 2044 times. Jevons threw a coin 20,480 times and the frequency of heads was 0.505. Quetelet experimented with an urn containing 20 white balls and the same number of black ones. After 4096 extractions there appeared 49.3% and 50.7% of them respectively. Westergaard repeated the same experiment with one ball of each colour (white and red) and got 5011 and 4989 of them respectively.

Consider now the Jevons experiment as 2048 series of 10 throws each; the frequency 1/2 for heads appeared in 24% of them, with theoretically expected 25%. Cases in which the number of heads deviated from 5 by 2 (say), and thus equalled 7 or 3, occurred 482 times (according to theory, 480 times). Then, there were 5 series only consisting of heads or tails (theoretically, 4)²⁰. It follows, that experiments do not refute our theoretical construction; on the contrary, they are considerably important for understanding the occurring phenomena because they show how to

explain the unaccountable, on the face of it, negligible fluctuations of the empirical frequencies given invariable general causes.

[3.2] Nevertheless, before delving into an explanation of that so surprising constancy of statistical figures, it is necessary to prove that the probabilistic pattern, only compared until now with experiments on games of chance, is also applicable for statistically studying mass phenomena. Obviously, however, it is impossible to consider each relative figure based on observations as an empirical expression of a certain objective probability; certain assumptions are here needed. But are these assumptions fulfilled in statistical work? We cannot deductively establish it because that work encompasses phenomena as well as their causes which are not known with sufficient precision. It is necessary to study the statistical material itself, and Lexis is greatly meritorious for accurately justifying by theoretical considerations the need to apply the calculus of probability when studying statistical data, and inventing appropriate precise methods.

When deciding whether the frequencies correspond to probabilities, separate statistical figures obviously cannot help at all. [...] There will be no way to solve this problem. If, however, we have a number of frequencies belonging to different totalities, we can ascertain, as Lexis had shown, whether all of them ought to be considered as empirical expressions of one and the same probability.

Lexis offered two methods. Assuming that separate frequencies were empirical expressions of one and the same objective probability, we can calculate the modulus

$$\sqrt{\frac{2p(1-p)}{M}} \quad (3)$$

with p replaced by the mean frequency of the phenomenon over the totality or over all the totalities together and to take the mean number of events in these totalities instead of M ; the separate numbers should not considerably differ from each other.

Then that modulus is made use of to calculate how would have the separate frequencies been grouped around their mean supposing that all of them had one and the same objective probability as their base. Indeed, we know that no deviations exceeding three moduli should have been found at all, and the table of the values of $f(u)$ will show that beyond the bounds of two moduli there should be about 0.005 of the frequencies; that approximately a half of them will be beyond those that correspond to 0.4769 of the modulus [to the probable error], etc.

If the actual distribution of the frequencies does not correspond to those theoretically calculated, we ought to consider as refuted our assumption of all of them being based on one and the same probability. In the opposite case we may assume that experience had justified our hypothesis.

[3.3] The other method created by Lexis applies an obvious criterion and is therefore preferable. If the separate frequencies v_1, v_2, \dots, v_n are grouped around their mean in accord with theory, we can mathematically prove that

$$\sqrt{2 \sum_{k=1}^n \frac{(v_k - v)^2}{n}} \quad (4)$$

should be equal to the modulus (3)²¹. It follows that if the actual fluctuations of the frequencies around their mean are exactly such as prescribed theoretically, the ratio of these two magnitudes, Q , should approximately equal unity²². This ratio plays an essential part in the Lexian theory. It indeed provides us with a reasonable measure of the constancy of statistical numbers so surprising everyone previously.

The stability of statistical figures, their insignificant fluctuations from year to year explained the entirely justifiable interest in statistics expressed both by specialists and laymen. [...] As a rule, figures only significant for a certain moment are not determined at all. For example, in Germany, owing to the irregular weather in summertime, the proposed census of people employed in navigation was abandoned: it was impossible to reach an idea corresponding with usual conditions.

Actually, the stability of statistical returns serves as the foundation of the very notion of society as an organic whole of outwardly uncoordinated elements which, in spite of freedom of movement, constitute a general totality existing in stable equilibrium²³. [I am omitting a piece which included relevant short comments on Graunt, Süssmilch and Quetelet.]

[3.4] The overrating of stability had led to statements that [...] order rested on compulsion from the outside which made Queteletism vulnerable and stirred up the storm that demolished it. [...] The cause of the mistake made is obvious. The materialistically minded followers of Quetelet made conclusions from statistical regularities which intrinsically offended the opponents of Quetelet. Resentment suggested to the opponents that not everything in such inferences was proper, but they lacked theoretical foundation which would have allowed them not only to reject, but to refute those followers.

Is the constancy of statistical figures indeed stable? Rehnisch (1876) resolutely denied it. Quetelet had been amazed by the stability of the figures of the French criminal statistics, but Rehnisch was able to show that Quetelet had treated them wrongly; he failed to note that in the middle of the period under his consideration the manner of publishing the data had changed (**p. 454**)²⁴. When taking the actually comparable figures for the entire period, it would have been necessary to double the data pertaining to half of them. And, in spite of such a change, the new series of figures displayed the same stability as previously, so how can we say that the figures were stable?

But what exactly did Rehnisch prove? At best, if any weight at all can be attached to his conclusion, his was an *argumentum ad hominem* which compromised the leader of his opponents. Indeed, if an inferior series accidentally happened to be weakly fluctuating, it does not really follow that the fluctuations of other series are large. An objective measure for stability of statistical figures ought to be discovered and only then will it be possible to make inferences about their actual behaviour. Many examples of diversity of opinion in regard to stability of one and the same statistical indicator show how acute is the need for such a measure.

[Chuprov provides a long quotation from Lexis (1903, pp. 171 – 172). Below, I only partly adduce it.]

It seemed natural to assume as the test and measure of fluctuation the mean deviation of the separate terms of a series from their mean level. [...] But that is a purely empirical method and only a theoretical investigation of the issue will indicate the boundaries within which it can be considered proper.

[3.5] And so, Lexis himself outlined the starting point of his investigations of the stability of statistical series. In connection with his study of the applicability of the theory of probability to statistics he was really able to discover a general solution to that problem. The magnitude Q which is the ratio of (4) or the calculated physical modulus to its combinatorial value (3) as described above is indeed the theoretically justified and therefore reasonable measure of stability of statistical figures. If the actual fluctuations of those figures exactly correspond to theoretical calculations resting on the assumption that the base of all the statistical totalities is one and the same probability, Q is approximately equal to unity²⁵.

Lexis called this case normal stability or normal dispersion. The case of $Q > 1$ signifying subnormal stability or supernormal dispersion is only possible if the actual fluctuations are greater than theoretically expected. On the contrary, the case of $Q < 1$ in which the fluctuations are smaller than those expected on the strength of the calculus of probability, the stability is supernormal, or, in other words, the fluctuations are subnormal.

Lexis and his students have applied these concepts and methods for studying the fluctuations of various statistical series. The investigation of the sex ratio at birth showed complete coincidence of the empirical data and probabilistic calculations. [...] That ratio also evinces normal stability when separately considering born in and out of wedlock and live newly born and stillborn.

Almost complete coincidence was also typical for the sex ratio of those dying during the first five years of life and in extreme old age (those above 80). For other age groups it clearly occurred subnormal; thus, for ages 50 – 75 years Q did not fall lower than 4. On the contrary, mortality of old age males was normally stable and obviously subnormal for children.

Complete failure²⁶ also befell the investigators when they attempted to apply the same test to the data of moral statistics: all the indicators revealed subnormal stability. Thus, when considering the part of various civil [social?] states of those entering into marriage in Sweden, we find $Q = 2.3$; the relative number of suicides in Germany, $Q = 5$. Greater values were also encountered: not rarely $Q = 10, 20$ and even larger (Bortkiewicz 1898, p. 28).

There are thus cases in which empirical values are so close to theoretical, that the empirical frequencies may be rightfully considered as reflections of one and the same objective probability. However, the majority of the studied frequencies revealed subnormal stability: for cases in which all of them had been based on one and the same probability, their actual fluctuations were essentially greater than theoretically expected.

[3.6] Two important points were here noted. On the one hand, the coincidence with theory proved the better the less were the totalities for which the frequencies were calculated. For example, the relative frequency of suicides in Germany had $Q = 5$ but for the eight smallest German states and female suicides $Q = 1.15$. Bortkiewicz also proved that for small totalities the stability became invariably normal. On the other hand, it was established that stability is the closer to normal the greater part of the properties of relative probability does the frequency possess. Thus, the sex ratio of the dead infants of age less than a year is more stable than of boys and girls [of the same age group] considered separately.

How to explain this? And what should we conclude if the dispersion is normal or supernormal? Why is stability of small totalities higher? Does it testify to the applicability of the calculus of probability to statistics or refute that possibility?

To answer these questions we ought to turn once more to the tested method, to studying games of chance. We know already when is normal dispersion possible. When throwing a coin or a die, or extracting a ball from an urn, the frequencies of heads and tails, of the number of points and of the white and black balls in several series of trials are grouped around their means in a way we call normal. Poisson proved that the scatter also remains normal when, for example, the balls are extracted from several urns having differing ratios of balls of those colours provided that before each extraction the proper urn is randomly chosen²⁷.

Thus, if one urn only contains black balls, and another one is only filled with white balls, and their choice is random [...], the frequency of the white ball after M extractions will be ca. $1/2$ in the mean and the approximation will be the same as if there were only one urn with an equal number of balls of each colour. When repeating that experiment n times with an invariable M , the frequencies of the white balls will be grouped around their mean just like when extracting the balls from a single urn with an equal number of balls of each colour.

This generalization of the pattern of normal dispersion made by Poisson by means of mathematical calculations but obvious without them as well (p. 458)²⁸, is the cause of the appearance of the normal dispersion and therefore of very high theoretical significance. It proves the utter groundlessness of the notion which assumes that the totality from which the frequencies are calculated ought to be homogeneous. [...]

[3.7] It follows that the normal dispersion does not at all mean that all the elements comprising the statistical totality be essentially identical as the Queteletists had been inclined to believe. On the contrary, they can be worlds apart. [...] In their time, while opposing Queteletism, Rümelin and then Schmoller rightfully denied, and called “psychologically monstrous” the statement that

It is possible to say to the face of each member of a nation, and even of the noblest of them, that his inclination to crime is expressed with mathematical precision by the fraction 21/10,000.

However, these scientists acknowledged that statistics was in the right when proclaiming to everyone that “he must die during the next year with probability $1/49$ and to bemoan heavy losses to the circle of those dear to

him with even a higher probability” and that he must “humbly submit to the sorrowful truth”. This only shows that their judgement, although correct, had been based on indignation rather than on sober and rational thinking. Indeed (Bortkiewicz 1894 – 1896, Bd. 10, p. 359n), for his part, “The strongest and healthiest representative of the nation” could have quite as justifiably “denied as psychologically monstrous the forecast made by statistics that he will die with probability equal to its general mean value”.

To describe clearly the conditions under which appears subnormal stability, I make use of the same example with which I had explained the Poisson generalization of the pattern of normal dispersion. An urn only contains black balls and another one is only filled by white balls. Choose the urn for the first extraction by lot and extract all the M balls of a series from it. Then repeat all this to obtain the second equally numerous series. If n series are thus composed, the mean frequency of white balls over all the series, as in the previous case, will approximately equal $1/2$.

However, the particular frequencies over each series will deviate from the mean essentially greater than previously, since, depending on the outcome of the random choice of the urn, separate series will wholly consist either of black or white balls. In this case, the predominance of small deviations from the mean which took place previously is impossible; moreover, the only possible deviations are now $1/2$ or $-1/2$. This is why the stability of the entire series will be essentially lower than normal with supernormal dispersion.

Imagine now a third experiment. If the urn for the first extraction is chosen by lot, but the first series of M balls is this time made from both urns in turn, the mean frequency of the white balls in n [such] series will also be approximately $1/2$. However, the grouping of the serial frequencies around that mean will reveal something special because for an even M there will be an equal number of balls of each colour, and otherwise these numbers will differ by unity. The possibility of large deviations from the mean is thus excluded and all frequencies will be situated very near to their mean. Stability will thus be supernormal with subnormal dispersion.

[3.8] When comparing the three examples we get a clear notion about the conditions for the appearance of the normal, subnormal and supernormal dispersion. If the separate elements are included in a statistical totality independently from those previously included (as in the first example when before extracting each ball the urn was chosen randomly) normal dispersion will appear. If, however, previous results influence the following ones (as in the second example when the urn was chosen once for the whole series of extractions [...]) the dispersion will be supernormal. And it will be subnormal if a deviation in one case leads to deviation in the opposite direction so that there arises a certain mutual compensation (as in the third example [...]).

Both super- and subnormal dispersion testify that the separate trials united into a statistical totality are not independent one from another. In one case the results of the trials are more similar to each other than to the results in some other totality; in the other case the totality was composed in such a way that the deviations of each element from the mean was smoothed over by an opposite deviation of another element. Both these

cases can be got by a certain choice of the separate trials constituting these totalities.

Supernormal dispersion can also be obtained in another way without the result of one trial influencing those of the other ones. The results in each series being based on one and the same probability (differing from that in the other series) are interconnected. It is not therefore necessary for the trials constituting a series to influence each other because the extractions from the chosen urn are made in the usual manner. The dispersion however became supernormal because the extractions of one and the same series are interconnected by being made from the same urn.

As to the subnormal dispersion, it also nevertheless testifies to the presence of certain rules regulating the totality's composition from subsets being in exact numerical ratios to each other. Thus, a subnormal dispersion of accidents in the army is occasioned because their probabilities in infantry, in the cavalry, in artillery and supper units are very different whereas the ratios of the strengths of the arms of the service to the total strength of the army are stipulated by laws and can be considered constant, at least over a long period (Lexis 1903, pp. 228 – 229).

[3.9] It follows that the interpretation of statistical data as applied by the Queteletists against the proponents of free will would only have been proper for supernormal stability of the indications of moral statistics. Indeed, in that case it would have been difficult to reject the assumption that “the seemingly arbitrary actions of man” obey the laws establishing certain norms of numerical ratios in the studied statistical series. But even then such norms could have been interpreted in many different ways. Be that as it may, until we are able to establish cases of supernormal stability, regularity will accept all the demands which the proponents of free will could have ever put forward.

Indeed, suppose that the significance of external circumstances of, and intrinsic reasons for human behaviour only consisted in determining the sole possible way of acting out of an infinite set of those thinkable, and that free will had the choice of those acts without being guided by any considerations (similar to the ball in the roulette game that stops in one or another sector independently from its colour). Then, ultimately, exactly the normal stability should appear in the mean²⁹.

However, it actually occurs rather rarely, usually stability is subnormal, so the Queteletists' attempt to solve the metaphysical problem of free will by simple empirical *experimenta crucis* (deciding experiments) ought to be rejected as obviously unsuccessful. The actual stability of statistical figures, even if it most approaches normality, only allows to infer that its foundation is the existence of common causes of the statistical totalities under comparison.

Here is how Lexis (1903, pp. 226 – 227) discusses the real significance of the assumptions that the main causes were invariable:

There exist infinitely many combinations of conditions and circumstances leading to the possibility of a 30 years old person to die during the next year. Partly they depend on his physical condition and partly on external factors of life. It is as though during the next year all people of that age contained in a given population are subjected to a test

on whether there exists for them a probability to die or not (p. 463: whether these possibilities of death will indeed come true for them).

This can give us some idea about the probabilities of death during the next year for that age and exactly in the same way, as, for example, when obtaining an approximate idea about the relief map of the bottom of the sea by measuring its depth in many definite places. The bottom remains invariable, so that a second series of measurements will give us approximately the same picture. However, the physical condition of cripples of that age also repeats itself to a certain extent. Although not constant, it is similar in respect of the probability of dying during the next year.

The combinations of conditions and circumstances leading to death can change in infinitely many ways, but we will still consider them invariable. Just the same, a die on an infinitely differing parts of the surface of a table has infinitely many arrangements of that edge around which it rotates before falling down and showing the same number of points³⁰.

As I mentioned, 30 years old individuals of consecutive observed generations are interchangeable in regard to their condition, both physiological and pathological, and external circumstances of life. The same idea can be extended on human actions interesting for moral statistics. They are determined by a combination of conditions caused partly by social, economic and other external circumstances of life and partly by the psychological and moral peculiarities of the individual. However, these types of individuals in spite of their possible infinite variety are also sufficiently homogeneous in respect of conditioning their behaviour. The totality of external and subjective circumstances in regard to a studied event changes comparatively weakly even for various individuals replacing each other from year to year and determined to a certain extent abstractly which indeed leads to the observed statistical regularity.

If, given a normal dispersion, both the general circumstances of life and the structure of the groups of population, homogeneous in respect to the studied event, change from year to year not more than it happens under the pattern of independent trials in a game of chance, it becomes immediately clear why does the majority of statistical series reveal stability lower than normal. Events, being united over a calendar year (say), are usually independent not to the degree necessary for normal dispersion. Some causes among those, influencing the process of events all year long or for its considerable part, often act in the same direction, then making way for other causes acting in the opposite sense.

[3.10] Again, take for instance mortality. In a given year climatic peculiarities [?], a random outbreak of an epidemic disease or, on the contrary, its cessation lead now to increase, now to decrease in mortality in regard to its mean level. It is therefore wrong to consider that the objective probability of mortality remains invariable over the whole period of many years. The correct conclusion is that each year has its own probability and in that regard can essentially differ from the previous year and future years.

The fluctuations in the empirically established indications of mortality are therefore caused by two sets of causes. The first one leads to the

distinction between the objective probabilities for separate years and is reflected in the fluctuations of the probabilities around their mean level. The second set of causes involves the reproduction of the objective probability of mortality in each year to occur not precisely, but with random deviations. It remains the only acting set when all the objective probabilities are the same, and when, consequently, we should have expected a normal dispersion.

The first set however entails an excess of fluctuations of the empirical frequencies over the level of normal dispersion. It is thus easy to understand why is the stability of the probability of children's deaths lower than that of others. It is obvious that the climatic peculiarities of the year in question stronger influence the delicate health of a child than the health of an adult. It is only sufficient to recall here epidemic diseases and gastric diseases occasioned by intense heat in the summer. [...]

It is also clear that the dispersion is the closer to normality the greater part of the properties of relative probability does the frequency possess, cf. the "second point" in § 3.6. This circumstance can only be explained by the more or less equal action of many factors which distinguish one year from another one on both groups of phenomena contained in the relative probability. These factors therefore only weakly reflect on their ratio. Thus (Lexis 1903, pp. 208 – 209),

The ratios of boys and girls dying in the group of 0 – 1 years old to the live newly born of the same sex represent approximate values of absolute probabilities. When studying these empirical data over a number of years, we discover their supernormal dispersion. This result could have been foreseen because we are able to name many external causes (cholera, strained circumstances etc) leading to essential changes in the normal conditions of mortality. On the contrary, there are no grounds for the external circumstances differently influencing the health of boys and girls because the way of life of both sexes during infancy is identical. This is why the relative probability of death of children of both sexes remains constant in spite of the change in the absolute probabilities, which is testified by the discovered normality of dispersion of the empirical values of that probability.

The correctness of this explanation is also corroborated by the supernormal dispersion revealed by the relative probability of death of people of both sexes in certain age groups in which the conditions of life of people of different sexes are not anymore identical.

[3.11] It remains to explain the following fact unacceptable to the majority of statisticians blindly trusting the law of large numbers: the stability of statistical figures is as a rule the lower than normal the more numerous is the totality made use of when calculating frequencies. There are two causes of this circumstance. The first one is an optical illusion of sorts. The measure of dispersion in our investigation is the coefficient Q , the ratio of the modulus calculated by issuing from the actual frequencies to the theoretical modulus corresponding to the normal dispersion. To recall, in case of supernormal dispersion (beginning of § 3.10) the numerator of the fraction consists of two components, of the fluctuations of the objective probabilities from year to year and the random deviations

of the empirical frequencies from their corresponding probabilities whereas only the second component serves as the denominator of that fraction.

Therefore, if the first component is greater than the second one, Q will essentially deviate from unity; if it is considerably smaller, Q will be close to unity. However, the random deviations are known to be the greater, the less is the number of observations used for calculating the frequencies. If the measure of fluctuations of the objective probability remains invariable from year to year, then Q ought to be smaller for a small number of observations than for larger numbers. It follows that although the stability of a series usually approaches normality given a lesser number of observations, no real significance ought to be attached to this result since it depends on the method of measurement (p. 486).

This explanation is not however exhaustive. Bortkiewicz proved that, after eliminating the fictitious influence of small numbers, the stability of the repetition of events in that case more closely approaches normality and that, on the contrary, when that number becomes ever greater, stability becomes ever lower. He also proved that that fact remains partly unexplained.

Drawing on the Bienaymé and Cournot theory of causes acting in concert, Bortkiewicz called this fact *law of small numbers*³¹ and rigorously justified it. The essence of the matter is this. We have satisfied ourselves in that the supernormal dispersion occurs because the elements united in a statistical totality are not mutually independent, that there exists some kind of affinity between them. We may show mathematically that the excess of the dispersion as compared with normality depends on the number of the thus interconnected elements. If only two consecutive elements are connected, the dispersion deviates from normality not really significantly; if, however, there are three or more such elements, the deviations become ever greater. The interconnection of the parts comprising the totality can occur owing to two causes. One of these is occasioned by peculiarities of the studied phenomenon. Consider as an example explosions of steam boilers. Fatalities are then not independent one from another because in such accidents several persons usually perish at once. The number of those killed in such instances of *acute solidarity of single cases*, as Bortkiewicz called them, depends on how many people in the mean are working in the boiler rooms.

Interconnection can also occur because there exist factors uniformly acting on the process of some events during a certain period of time (for example, such climatic conditions as heat acting on mortality). Bortkiewicz called this kind of interconnection *chronic solidarity of single cases*. When following the wrong path because of a mistakenly understood significance of the law of large numbers, and amassing observations, collecting single cases over large regions, the totalities become composed of large groups of interconnected elements. If, however, you restrict that territory to small districts with a small number of [the studied] events occurring in each time unit, the volume of the totalities of interconnected elements will be the less the smaller are the corresponding district and the relevant numbers.

Issuing from the Bortkiewicz formulas, I will study sufficiently small numbers. As a conclusion it occurs that the dispersion is always close to

normality so that the statistical figures will better correspond to probabilistic calculations. And this will justify the fact that “mathematical probabilities or their functions are the foundation of all figures derived in population and moral statistics” (Bortkiewicz 1898, p. 39)³².

4. Establishing Causal Connections by Probabilistically Treated Observations. The Corrupting Influence of Ontological Connections. Comparing Series

[4.1] And so, it is hardly possible to question the right to consider statistical figures as empirical expressions of objective probabilities (or their functions). But how to apply this principle for justifying the rules of the statistical method; how, by issuing from it, to study causal connections between the studied phenomena?

Two cases ought to be distinguished. Statistical totalities at our disposal are either directly obtained from observations (for example, by comparing indications of mortality in a number of localities), or arbitrarily compiled for the investigation with the single cases being brought together into totalities in accord with the indication selected in advance. In the first case, having established that the distinctions between empirical frequencies are real, we immediately turn to induction. Suppose that the probability of dying in city A is higher than in city B. This is only explained by the difference between the general causes influencing mortality. [...] To establish such a distinction is however almost as difficult as when applying inductive methods in the usual way. True, it is not necessary to consider the infinitely many individual particulars in each case, but nevertheless the majority of other general causes remains practically boundless so that disjunction, the foundation of the inductive method, is not exhaustive and the indefiniteness of the retained remainder of causes makes conclusions impossible.

For example, applying the method of standard population, we have established that the differences could not have been occasioned either by the age or the sex structure of the populations or by climatic conditions or numerous other factors. But the conclusion that the sole cause [differently] influencing mortality is the lack of water supply [in city A] will nevertheless remain in thin air because it is always doubtful that there are no other unnoticed causes.

True, there are cases in which numerous consecutive considerations of differences and the researcher’s know-how make it possible to eliminate at once so many causes that the conclusion seems almost certain. For example, Professor Tugan-Baranovsky had studied industrial crises in England wishing to establish how did they influence the industrial population in the demographical sense. He compared the number of births, marriages and deaths in English agrarian and industrial counties both during the years of industrial prosperity and crises. That double comparison in space and time indeed eliminated all other factors and it seems that nothing was left except the influence of industrial crises on the only part of the population which they could have affected.

[4.2] Such cases are however rather rare. [...] On the contrary, when artificially composing statistical totalities in accord with the aims of investigation, the circumstances are better, or at least so they seem to be on the face of it. Here, indeed, phenomena are collected in accord with definite indications so that the ensuing totalities can only differ in these or

in indications connected with them. Other indications can only randomly influence the properties of a totality as is explained by the very notion of independence.

Therefore, if we are able to establish that the distinctions between the totalities are not accidental, we are seemingly justified to conclude without any reservations that the sought causal connection between these distinctions and the chosen indications [really] exists. However, more closely studying this conclusion we establish its wrongfulness (Kries 1886, pp. 85 – 87). Apart from those casual connections in which we are primarily interested, and whose significance is not restricted either in time or space, – apart from the so-called *nomological* connections, – there exists an interconnection only occurring during a certain time period and over a certain region, resulting, as Mill termed it, from the original distribution of causes. Such interconnections are called *ontological*.

The state of the universe in each given moment, even in the opinion of those who admit complete determination of every occurrence, is conditioned not only by causal connections, but also by the essence of the distribution of separate “causal elements” at the assumed initial moment. The concept of independence on which the conclusion above rests, only considers nomological ties, but the possible ontological connections are able to corrupt considerably our conclusions either by strengthening or opposing the nomological dependences.

[There follows an example concerning the lease of land in Russia after the abolition of serfdom.]

We always have to do with various causes and deep investigations are needed for examining them. And so we fall again into the trap which we expected to avoid. There is no other means except applying induction, but in such cases, as in other instances, its methods are of little use (p. 472)³³. The situation is extremely difficult but the knot can be still cut, namely by the mentioned method of comparing series. It is based on the following principle: If two phenomena are not connected with each other either nomologically or ontologically, the presence or absence of one of them will not influence the appearance of the other one. If both phenomena admit quantitative description and, bar random deviations whose measure can be calculated in advance, any value of one of them can with the same probability (and therefore with the same facility) coincide with any value of the other one.

[4.3] This principle can be formulated more precisely in accord with various methods of investigation. Practically the most important method consists in grouping separate cases in accord with the value of an indicator investigated with a view to detect possible connections between two phenomena. Then the values of that magnitude (the *characteristic*) that best quantitatively describes the other phenomenon are calculated.

Thus, all the villages of a province are separated into 10 groups in accord with the mean allotment per capita, and the per cent of those who lease additional land is calculated for each group. If the prevalence of leases and the size of the owned plot are connected nomologically, and there are no opposing ontological connections, we may expect that, as the indication applied for grouping the villages increases, the characteristic will regularly increase or decrease³⁴.

If, however, the two phenomena are perfectly independent, then no regularity (if only it does not occur accidentally) in the variability of the characteristic can be detected which shows that the phenomena are independent both nomologically and ontologically. In contrary cases it is first of all necessary to calculate the probability that the regularity could have occurred randomly. [...] Our method is certainly unable to reveal the essence of the connection between the phenomena [if it exists], whether it is nomological or ontological, but that is impossible to achieve not only analytically, but even experimentally and under the most favourable conditions³⁵. Only when invariably detecting the same connection under most different conditions of space and time we may admit its real existence and be sure that its essence is nomological. But [even] this conclusion only becomes obligatory if we are able to prove that ontological connections are lacking or are of the kind which, as is known from other sources, are able to compensate the nomological dependence and all by themselves must cause an opposite change in the values of the characteristic.

5. The Collection of Separate Observations into Totalities As the Precondition for Comparing Series. The Principles of Constituting Totalities

[5.1] Minute probabilistic calculations do not lead to immediately applicable results and the meagre conclusions ultimately made are as a rule arrived at by comparing series. A question therefore arises: Does it make sense at all to compose groups, calculate frequencies and determine the bounds of objective probabilities since the same method can be applied directly to separate observations? The concept underlying it does not at all presuppose dealing with totalities and probabilities. Is it not possible to spare time demanded by statistical calculations?

After a closer scrutiny doubts about the expediency of the statistical methods become unfounded. In principle, certainly nothing at all hinders the comparison of series given isolated and not statistically treated observations, but in practice such an attempt will in most cases prove unsuccessful because only very rarely would have been the connection between phenomena detected.

The point is that a conclusion made after comparing series is based on their intimate parallelism which cannot be explained away by random coincidences. However, to reveal it (if there exists at all a connection between the phenomena) it is necessary for that cause which served to arrange the first series in accord with its numerical characteristic, to be more significant than the other ones. And, if we consider isolated cases, the action of each cause is too often corrupted by other causes which either strengthen or hinder it.

[There follows an example concerning lease of land.]

When grouping observations only allowing for the indication which interests us we neutralize to a certain extent the infinitely many factors independent from it which corrupt its action since within a group they more or less compensate one another. On the contrary, the action of the selected indication strengthens similar to what happens when a photographer, while attempting to decipher the text of a faded manuscript, superimposes several transparent negatives one on top of another.

Grouping aims at detecting connections in those cases in which an immediate inference is impossible and it is necessary to detect more clearly a regularity to serve as a foundation for conclusions in a series of characteristics.

The replacement of empirical frequencies by objective probabilities underlying them pursues the same aim: we thus exclude the corrupting influence of chance. And so, for comparing series successfully a preliminary treatment of isolated observations, their gathering into totalities and the calculation of probabilities, is necessary.

[5.2] Here arises a theoretical problem: how to group observations and to select numerical characteristics of both studied phenomena so as to achieve best results when comparing series? Innumerable mistakes which are made by statisticians day after day testify to the great importance of this problem. For example, often a characteristic of the phenomenon is chosen on account of a formal arithmetical connection of its values with the magnitude of the indication and it cannot therefore serve for proving a real dependence between phenomena. [There follows an explanation by an example of leasing of land.]

Is some real relation between the two phenomena concealed behind the formal dependence of numbers established by us ourselves and therefore lacking any interest? This cannot be ascertained by such treatment of the numerical data [?]. Very much useless labour is being spent on these or similar methods of work. Statistics (Knapp 1868, p. 71)

Busies itself with empirical calculations of relative numbers so much that specialists forget to ask themselves, What, indeed, is the meaning of those magnitudes from which the relative numbers are calculated? Often only the process of calculation is important for statistics, it hardly thinks about the results, and, least of all, about the aims and subject matter of that work. It had broken the habit of preliminary ascertaining what is desirable to investigate and therefore it is often lacking in the need to ask [itself] what is the meaning of one or another magnitude for the study.

An effective means against such squandering of force and time can only be created either if each method of investigation used in separate branches of statistics be deeply analyzed in respect to its methodological value, or a general theory of grouping and of selecting numerical characteristics of phenomena to be studied be developed. We do not regrettably have such a theory; moreover, even the very problem of creating it was hardly ever formulated in a general way³⁶. Still, there exists a branch of statistical investigations which is satisfactorily developed in respect of all abovementioned demands.

[5.3] Because of the successful work of German mathematicians cum statisticians during the 1860s – 1870s the statistics of mortality is now so well methodologically regulated that only those who had not studied the special literature can methodologically blunder.

On the contrary, in all the other branches we even nowadays are groping our way ahead just as it had been in mortality statistics before the works of Knapp, Becker, Zeuner, Lexis and others had appeared³⁷. At first, representatives of the mathematical direction in statistics had only been engaged in methodological problems connected with investigating

mortality. They learned to delimit clearly the various problems; they established such sets of those dead and living whose comparison allowed to solve theoretically and precisely each of these problems; they showed that the passage from the groups usually applied in statistical practice to those needed by theory was possible; they discovered a simple method of improving statistical practice by separating the dying in a given year both by year of birth and age.

In general, they created a system of scientific principles which elevated mortality statistics to such a high theoretical level which can be hardly exceeded. However, broadening the sphere of problems caused by studying mortality had not even been attempted. True, migration has begun to be allowed for, but it was not considered as a phenomenon interesting in itself and only studied insofar as it interfered with the natural movement of the population, of *Bevölkerungswechsel*, as Knapp called this new scientific discipline. Only Lexis, who joined the other scientists when but little could have been added to it, discovered new paths for investigations in this sphere. By generalizing the conclusions from the studies made by those of the mathematical direction he showed that they, the studies, can be easily extended from phenomena of mortality to “changes of status” such as death or birth, marriage or widowhood, emigration or immigration.

Believing that the main aim of demography consists in studying the course of life of the “abstract man”, he ascertained how to apply the methods developed by students of mortality to a rational construction of demography.

On our younger generation falls the problem of continuing the work of our teachers and rationalizing the statistical method in other spheres of its application. Then perhaps we will be able to fill the gap yet existing to this day between our general ideas about the aims of the method and the ways to fulfil our work.

[5.4] I have attempted to describe in general terms the theory of the statistical method as it is apparently being gradually reconstructed out of the wreckage of the Quetelet system. I believe that this method serves as a substitute of the method of induction, which, having been created for investigating continuous causal ties, becomes helpless when treating practically more important free connections.

I have therefore attempted to show how does the statistician, when collecting separate phenomena in statistical totalities, exclude corrupting chance from the empirical frequencies, reveal their underlying objective probabilities and establish by comparing the obtained series whether there exists an interdependence between the studied phenomena. I may hardly expect that my attempt to collect various streams of theoretical statistics on the basis of the principle of causality into a single whole by broadening the sphere of scientifically studied causal connections will find sympathy of the creators of that theory. The points of departure of individual researchers are too diverse and different.

Mathematicians and philosophers are being engaged in creative work not less than statisticians whereas natural scientists and especially biologists are recently carrying it out with remarkable success. It is obviously impossible to say whether all the streams will come together exactly at the point which I have outlined above, but come together

somewhere nearby they certainly will. Indeed, in spite of all the distinctions in the details of building up a methodology and in its philosophical underpinning, the common goal of all the attempts is the search for the roots of the statistical method in the calculus of probability. This general agreement ought to be attributed to the influence of that scientist, whose latest work prompted me to compile this paper and whose name runs all through its pages. I cannot therefore conclude without saying a few words about the significance of Lexis as the creator of a new theory of statistics.

Bortkiewicz (1904) has skilfully analyzed how Lexis had treated his predecessors and it is meaningless to dwell on this point anymore. As to his influence on the current development of the theory of statistics, it is really sufficient to note that in each country where some activity in the field of statistical theory now begins to be felt, its sources issue from the Lexian scientific principles. The powerful development of that theory in England, the country a short way from becoming the leader [in that direction], owes its being not only to Galton, but to Edgeworth as well whose work bears obvious traces of his influence. The Dutchman Westergaard, the Russian Bortkiewicz, not to mention the German students of Lexis, owe him the principles of their scientific approach to statistics and the first impulse to their theoretical statistical work.

The studies of stability which Lexis published in a separate book (1877) and especially those which appeared in the *Jahrbücher*, had long ago become the ABC of sorts for researchers interested in justifying the theory of statistics.

Now that his papers published in periodicals are included in a single contribution (1903) and became available, it will hardly be a mistake to expect that the interest, at least to young specialists and even outside England in theoretical research will flourish anew and that the circle of those joining Lexis will be gradually able to throw critical light on the empirical methods of work in statistics and to justify them in a rational manner.

And if statistics will ever enjoy equal rights with, and become full member of the fraternity of other theoretical disciplines, it will testify to the essential merit of the courageous man who elevated himself above the everyday practical troubles and paved the way to a rational theory of statistics through a desert of compiled fruitless prescriptions replacing general principles, and who achieved all this at the time when the work of statisticians had been completely and exclusively reduced to counting and calculations.

Notes

1. Chuprov discusses the exaggeration in sociological conclusions based on stability of statistical series. N. C.

Chetverikov provided no references. I would say: sociological in a narrow sense only concerning individuals.

2. It is opportune to mention that Comte (1830 – 1842/1877, t. 2, p. 255; 1854, p. 120) decidedly opposed the application of the theory of probability to social sciences. He (1830 – 1842, 1839/1908, p. 4) also alleged that Quetelet had plagiarized him. See quotation in Sheynin (1986, p. 295). I do not see any special merit in Quetelet's treatment of the theory of means (*Ibidem*, pp. 311 – 312).

3. See Note 10 to § 2.1.

4. Both the original German text (p. 423) and its Russian translation state: “principles of population and demological (demologische) statistics” which I do not understand.

5. See Kries (1886; 1888).

6. In the sequel, I am borrowing the ideas of naïve realism since everyone invariably uses its language whichever personal point of view he is keeping to and since their translation from it into the language of one or another metaphysical or theoretical system is usually the easiest. A. C.

7. Chuprov inserted here three lines in their original Latin. I translated them from their Russian translation provided by Chetverikov who also stated that their author was Francis Bacon.

8. A strange statement. Where is randomness (and chaos)? Cf. similar statement in § 4.2.

9. The generally accepted indications underlying innumerable definitions of the statistical method can be easily deductively derived from my formula. It is, in particular, applied for studying numerically described mass phenomena, see below. A. C.

10. Later Chuprov himself (1918 – 1919, Bd. 2, pp. 132 – 133) noted the difference (already existing in 1905) between the Continental direction of statistics and the Biometric school and stated:

Not “Lexis against Pearson” but “Pearson cleansed by Lexis and Lexis enriched by Pearson” should be the slogan of those, who are not satisfied by the soulless empiricism of the post-Queteletian statistics and strive for constructing its rational theory.

Chuprov did not prove that Edgeworth had joined Lexis. In a few years he (1909/1959, pp. 27 – 28) described Edgeworth’s work in general terms but had not said anything of the sort. In 1896, Edgeworth had criticized Bortkiewicz (1894 – 1896) who answered him at the end of his contribution. Chuprov, in Letter No. 5 of 1896 (Bortkevich & Chuprov 2005) sided with his colleague.

11. The same methods can be however applied in other cases as well. Thus, the notion of probability, worked out for being immediately used under plurality of actions and comprising the base for statistical methods, when generalized on the case of plurality of causes serves as the logical subject matter of those propositions of the calculus of probability which adjoin the Bayes theorem. [...] A. C.

12. It was Poincaré who initiated discussions about equiprobable cases, see Khinchin (1961/2004, pp. 421 – 422) and von Plato (1983).

13. In 1896 Chuprov (Sheynin 1990/1996, p. 95) discussed the possibility of proving the law of large numbers in a non-mathematical way with Bortkiewicz and referred to Cournot in the same vein as in 1905 (see below). Bortkiewicz, however, called this opinion a pipe-dream. A bit earlier Chuprov (1896, p. IV) even mentioned an attempt at systematising the theory of probability on the basis of the logical analysis of probability.

In 1923, in a letter to Chetverikov, Chuprov (Sheynin 1990/1996, p. 97) expressed a different idea:

Like you, I even now, just as when writing the [Essays (1909)], [...] see no possibility of throwing a formal bridge across the crack separating frequency from probability.

However, Chuprov (1909/1959, p. 168, also see p. 301) stated there that the law of large numbers “establishes the connection between probabilities of phenomena and their frequencies”.

Chuprov never mentioned the definitely known to him (Ibidem) strong law of large numbers. A few years after his death Khinchin (1929, pp. 124 – 125) explained that it, rather than the Bernoulli theorem, was the “veritable foundation” of statistics.

Halley mentioned irregularities (which could have been occasioned by systematic influences) rather than chance.

14. Translation of the original Latin from the Russian translation provided by Chetverikov.

15. Chuprov only mentioned Chebyshev somewhat below, which was hardly proper.

16. Instead of choosing such roundabout ways, it was possible to prove directly this proposition. However, such attempts would have been unconvincing and, in addition, defective in that they do not provide a precise measure of the expected deviations from the underlying probability. A. C. This is not clear.

17. Lexis provided a table of the normal law so that Chuprov assumed it without any restriction. His formula (2), which originally correctly appeared in a German translation by Timerding of the celebrated Bayes memoir (Sheynin 2003), was wrong; the denominator should have been $M^{3/2}$.

18. In addition to the modulus especially popular with English statisticians other magnitudes connected with it by simple relations can also be applied for describing the precision with which frequency reproduces its underlying probability. Such are for example fluctuations equal to the square of the modulus, the mean error = 0.707 of the modulus, the probable error = 0.47696 of it or the reciprocal of the modulus. Lexis prefers the probable error. A. C.

In standard notation, the modulus is $\sigma\sqrt{2}$, and the mean error = σ . Chuprov tacitly assumed that the frequencies were distributed according to the normal law. Pearson was also in the habit of retaining an excessive number of digits, and Fisher may be mentioned here as well. See discussion of that practice in *Science* (vol. 84, 1936, pp. 289 – 290, 437, 483 – 484 and 574 – 575).

19. Pearson's study certainly attracted attention. Newcomb (Sheynin 2002, p. 160), in a letter to him of 1907, asked K. P. whether he had collected any new pertinent information.

20. See Czuber (1903) for further examples. A. C.

21. Chuprov had not explained the meaning of v (in my notation); it obviously is the arithmetic mean of v_i .

22. Approximately, because for not very large M and n very essential random deviations can be expected both in the numerator and denominator of that fraction. For the range of random fluctuations of the coefficient Q see Bortkiewicz (1901). A. C.

Dormoy (1874) had provided a similar method two years before Lexis but his paper had been unnoticed until the latter's publications. A. C.

23. Physicists, working on the kinetic theory of gases, have developed the notion of totality consisting of irregularly moving components unrestrained by any interrelations directed to re-establish equilibrium but remaining essentially invariable. Sociologists however are applying that concept without bothering to deliberate about the logical structure of the concept of [thermo]dynamical equilibrium. A. C.

But how to explain Chuprov's own words (above) about the society's stable equilibrium? In the Soviet Union, Starovsky (1933, p. 279) called Süssmilch, Quetelet, Bowley as well as Bortkiewicz and Chuprov theoreticians of bourgeois statistics who had attempted to prove the firmness of the capitalist system and the stability of its laws.

24. Quetelet did not account for a law which, in particular, had reduced the number of capital offences. Furthermore, Rehnisch quoted official sources concerning the increase in suicides. See Sheynin (1986, pp. 299 – 300).

25. Chuprov (1922) essentially restricted that proposition. N. C.

26. *Failure* seems to be an unfortunate term here.

27. Poisson had nothing in common with the Lexian theory. Chuprov mentioned him several times more in the same vein.

28. Does this really mean that the Poisson law of large numbers did not demand "calculations"?

29. In 1858, C. Lewis, the celebrated English writer, applied this simple but deciding argument in his review of Buckle in the *Edinburgh Review*. Nevertheless, surprising as it is, the opponents of Queteletism never attempted to rest on the theory of probability if only the hardly successful reasoning of Drobisch be discounted. A. C.

Chuprov had only quoted Drobisch (without offering a proper reference) as saying "The study of moral statistics leaves a sting of doubt in each scientifically and philosophically minded person". I left out this quotation together with the relevant fragment.

30. This is not sufficiently definite.

31. Bortkiewicz only referred to them in his later contributions.

32. For simplifying the exposition as much as possible, I have only considered empirical frequencies and their underlying objective probabilities because the statistical method in its simplest form is [only] applied for investigating such magnitudes which in essence are probabilities. In his practical work, the statistician has often to deal with magnitudes which could be reduced to probabilities but nevertheless are only their functions (in the mathematical sense). These functions can be known; for example, the number of newly born boys corresponding to 100 newly born girls considered as a function of the probability of a male birth. In such cases this circumstance does not lead

to essential changes in the methods of statistical work. The decisive significance here is the mathematically proved proposition: If $y = f(x)$, the modulus of y equals the modulus of x times $f'(x)$. On the contrary, if such functional dependences are unknown, which as a rule happens in anthropology [anthropometry], deeper changes in the methods of work are needed, but these are out of place here. A. C.

33. As formulated, induction becomes entirely useless!

34. It should not be thought however that the ratio between these two magnitudes will always be exactly the same, as it appears owing to the very notion of causal dependence. [...] As a rule, we will have to do with magnitudes only somehow depending on the numerical values of the considered causes and actions. [...] A. C.

35. Previous theoreticians of statistics had been inclined to contrast the statistically knowable (man and society) and the typical (nature). Insofar as this understanding contains a methodologically valuable kernel, it can be reduced to the idea that a natural scientist can immediately perceive nomological connections by treating observations. "A physicist, after observing that a drop of mercury freezes at -40°C [-39°C], does not doubt that other drops of that metal will always and everywhere freeze at the same temperature". [...] This conclusion is only valid under the condition *ceteris paribus* (all other circumstances being the same). The circumstances supposed invariable ought to include the various ontological relations. [...] A. C.

[...] When considering atypical phenomena, we ought to allow for the discrepancies between observations [...]. Later Chuprov (1909/1959, pp. 69 – 70) stated his views on the same subject: "Nowadays the doctrine of typicalness of natural phenomena and atypical essence of human communal life can be thought irrevocably obsolete". N. C.

Again in his *Essays*, Chuprov (1909/1959, p. 75) provided the missing reference to the physicist: Gabaglio (1888, t. 2, p. 54).

36. I (1904) have attempted to collect some of the related material and allow myself to refer readers wishing a more thorough development of the abovementioned reasoning to this paper. A. C.

37. Among the new contributions to this sphere I ought to indicate especially the works of Professor Bortkiewicz. They differ from earlier studies mostly devoted to the methods of composing the groups by treating in the first place the choice of indicators. A. C.

Bibliography

Abbreviation: JNÖS = *Jahrbücher f. Nationalökonomie u. Statistik*

Bortkevich, V. I., Chuprov A. A. (2005), *Perepiska, 1895 – 1926* (Correspondence). Berlin. Also at www.sheynin.de

Bortkiewicz, L. von (1894 – 1896), *Kritische Betrachtungen zur theoretischen Statistik*. JNÖS, 3. Folge, Bd. 8, pp. 641 – 680; Bd. 10, pp. 321 – 360; Bd. 11, pp. 671 – 705.

--- (1898), *Das Gesetz der kleinen Zahlen*. Leipzig.

--- (1901), Über den Präzisionsgrad des Divergenzcoefficienten. *Mitt. Verbandes öster. u. ungar. Versicherungs-Techniker*, No. 5, pp. 1 – 3.

--- (1904), Die Theorie der Bevölkerungs- und Moralstatistik nach Lexis. JNÖS, Bd. 27 (82), pp. 230 – 254.

Chuprov, Tschuprow, A. A. (written 1896), *Matematicheskije Osnovania Teorii Statistiki* (Math. Principles of the Theory of Statistics). Unpubl. Gorky Library, Moscow State Univ. The Chuprov Fond, 7/1.

--- (1906), Statistik als Wissenschaft. *Arch. f. soz. Wiss. u. soz. Politik*, Bd. 5 (23), No. 3, p. 647 – 711.

--- (1909), *Ocherki po Teorii Statistiki* (Essays on the Theory of Statistics). Moscow. Later editions: 1910, 1959.

--- (1918 – 1919), Zur Theorie der Stabilität statistischer Reihen. *Skand. Aktuarietidskr.*, Bd. 1, pp. 199 – 256; Bd. 2, pp. 80 – 133.

--- (1922a), Das Gesetz der großen Zahlen und der stochastisch-statistische Standpunkt in der modernen Wissenschaft. *Nord. Stat. Tidskr.*, Bd. 1, pp. 39 – 67.

--- (1922b), Ist die normale Stabilität empirisch nachweisbar? *Ibidem*, pp. 369 – 393.

--- (1960), *Voprosy Statistiki* (Issues in Statistics). Moscow. Coll. articles, reprinted or translated by N. S. Chetverikov.

- Comte, A.** (1830 – 1842), *Cours de philosophie positive*, t. 2. Paris, 1877; t. 4. Paris, 1839. Many later editions. I also refer to the 1854/1908 edition of t. 4.
- Cournot, A. A.** (1843), *Exposition de la théorie des chances et des probabilités*. Paris, 1984.
- Czuber, E.** (1903), *Wahrscheinlichkeitsrechnung und ihre Anwendung etc.* Leipzig. Later editions: 1908, 1914, 1924. Reprint of edition of 1908: New York – London, 1968.
- Dormoy, E.** (1874), *Théorie mathématique des assurances sur la vie. J. des actuaires franç.*, t. 3. pp. 283 – 299, 432 – 461. Incorporated in book (1878) of same title.
- Gabaglio, A.** (1888), *Teoria generale della statistica*, vol. 2, parte filos. e technica. Milano. Second edition.
- Heyde C. C., Seneta E.** (1977), *I. J. Bienaymé*. New York.
- Khinchin, A. Ya.** (1929, in Russian), The strong law of large numbers and its significance for mathematical statistics. *Vestnik Statistiki*, No. 1, pp. 123 – 128.
- (1961, in Russian), R. Mises' frequentist theory and the modern concepts of the theory of probability. *Sci. in Context*, vol. 17, 2004, pp. 391 – 422.
- Knapp, G. F.** (1868), *Über die Ermittlung der Sterblichkeit aus den Aufzeichnungen der Bevölkerungs-Statistik*. Leipzig.
- Kries, J. von** (1886), *Die Principien der Wahrscheinlichkeitsrechnung*. Freiburg i. B. Tübingen, 1927.
- (1888), Über den Begriff der objektiven Möglichkeit und einige Anwendungen desselben. *Vierteljahrschr. wiss. Philosophie*, Bd. 12, pp. 179 – 240, 287 – 323, 393 – 428.
- Lexis, W.** (1877), *Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft*. Freiburg i. B.
- (1886), Über die Wahrscheinlichkeitsrechnung und deren Anwendung auf die Statistik. *JNÖS*, Bd. 13 (47), pp. 433 – 450.
- (1903), *Abhandlungen zur Theorie der Bevölkerungs- und Moralstatistik*. Jena.
- Markov, A. A.** (1911, in Russian), On the basic principles of the calculus of probability and on the law of large numbers. In *Ondar* (1977/1981), pp. 149 – 153).
- Mill J. S.** (1843), *System of Logic*. London, 1856 (4th edition).
- Novoselov M. M., Kuzicheva Z. A., Biriukov B. V.** (1973, in Russian), Logic, History of Logic. *Great Sov. Enc.*, vol. 14, 1977.
- Ondar, Kh. O., Editor** (1977, in Russian), *The Correspondence between A. A. Markov and A. A. Chuprov etc.* New York, 1981.
- Pearson, K.** (1894), Science and Monte Carlo. *Fortnightly Rev.*, new ser., vol. 55, pp. 183 – 193.
- Plato, J. von** (1983), The method of arbitrary functions. *Brit. J. Phil. Sci.*, vol. 34, pp. 37 – 47.
- Rehnsch, E.** (1876), Zur Orientirung über die Untersuchungen und Ergebnisse der Moralstatistik, second part. *Z. Phil. u. phil. Kritik*, Bd. 69, pp. 43 – 115.
- Seneta, E.** (1982), *Chuprov. Enc. Stat. Sciences*, 2nd edition, vol. 2, 2006, pp. 944 – 947.
- (1987), Chuprov on finite exchangeability, expectation of ratios and measures of association. *Hist. Math.*, vol. 14, pp. 243 – 257.
- Sheynin, O.** (1982), On the history of medical statistics. *Arch. Hist. Ex. Sci.*, vol. 26, pp. 241 – 286.
- (1986), Quetelet as a statistician. *Ibidem*, vol. 26, pp. 281 – 325.
- (1990, in Russian), *Aleksandr A. Chuprov: Life, Work, Correspondence*. Göttingen, 1996.
- (2002), Newcomb as a statistician. *Historia Scientiarum*, vol. 12, pp. 142 – 167.
- (2003), On the history of the Bayes theorem. *Math. Scientist*, vol. 28, pp. 37 – 42.
- (2008), Bortkiewicz' alleged discovery: the law of small numbers. *Hist. Scientiarum*, vol. 18, pp. 36 – 48.
- Starovsky, V. N.** (1933, in Russian), Economical statistics. *Bolshaia Sov. Enz.*, first edition, vol. 63, pp. 279 – 283.
- Venn J.** (1866), *Logic of Chance*. London, 1876 (2nd edition).

III N. K. Druzinin

Scientific Contents of Statistics in the Literature of the 20th Century

Problemy Teorii Statistiki (Issues in the Theory of Statistics).
Moscow, 1978, pp. 5 – 32

[1] In the first half of the 19th century statistical science had been marked by such an important event as the publication of Quetelet's contributions in which he had put forward fundamental concepts of that science. The end of that, and the beginning of the next century had also left a noticeable trace in the history of theoretical statistics. New ideas had then been formed on Russian soil and were connected with the Petersburg Professor Chuprov, who may be rightfully called one of the most prominent theoreticians not only in the Russian, but also in the world statistical science.

His merit consisted in an attempt to throw new light on the theory of statistics in regard to its main logical issues and to allow for its part in the common family of science. He thought it necessary because, in his opinion (1909/1959, p. 17), that theory was in a disastrous state:

There are periods in the life of each science when the specialists scornfully reject the attempts to analyse theoretically its principles. [...] Such feelings prevailed in statistics until recently. After the downfall of the Quetelet theoretical system under the blows dealt by German criticisms, years of extreme empiricism ensued in statistics.

And he (p. 18) therefore believed that a “new building of statistical theory out of the ruins of the Quetelet doctrine” should be erected¹. When appraising the actual situation, Chuprov undoubtedly exaggerated. The Quetelet doctrine was not at all reduced to entirely useless “ruins”, only its German critics thought that it did not contain something resembling a stone to be laid in the foundation of the new building. In reality, its general methodological concept advancing the statistical study of mass phenomena as the object of statistical studies, firmly entered the theory of statistical science. While Quetelet had been building up his doctrine, statistics did not yet gain sufficient practical experience resting on which it would have been able to form a more perfect system of the statistical theory. And the general state of both sociological and natural sciences had not allowed it either whereas the whole history of the concepts of the theory of statistics did not foster the creation of such a system.

And still, Chuprov could have certainly discovered material suitable for using it for compiling his own system in the writings of Quetelet and of later theoreticians as well. For the theory of statistical science, Chuprov's ideas nevertheless acted as fresh air also because in his time statistics had essentially advanced both in theory and practical applications and because of the general development of scientific knowledge in particular in natural sciences into which statistics had been extensively entering².

[2] When building his statistical doctrine, Chuprov issued from the Neo-Kantian philosophy which the representatives of the Baden school [a direction of that philosophy], Rickert and Windelband, had been developing. In accord with its agnostic concept Chuprov numbered statistics among the so-called *idiographic* sciences, or sciences of *individuality* whose subject was the knowledge of the concrete rather than the general, and contained in definite boundaries of time and space.

There exists the necessity of such knowledge because the “struggle against the immense universe”, if carried on by the *nomographic* sciences by simplifying reality and constructing general notions cannot completely satisfy the cognizing mind. The idiographic nature of statistics constituted one of the main propositions of Chuprov’s statistical doctrine and apparently corresponded to recognizing it as an empirical discipline. This, indeed, was the firm idea of statisticians of the 19th century working under the criticisms of the Quetelet system and the influence of Mill who had put forward the concept of empirical laws. That Chuprov turned to other philosophical notions for justifying his point of view did not essentially change anything.

When explaining why was statistics interested in idiographic knowledge, Chuprov (1909/1959, pp. 70 – 71) did not contradict the main principle of the theory of statistical science established from the time of Quetelet and stating that the subject of statistical studies is mass phenomena:

Leaving aside all the inherent peculiar features of the objects on which it is concentrated except for a few and without looking for precise data about each of them individually, statistics is interested not in an individual object, but in “totalities”. An individual horse does not interest the statistician even if it is a stallion costing a lot of money. [...] He is only interested in the horses of a given uezd (district), province and state. A Dreadnought has no bearing on statistics; its only subject is the totality of ships, a fleet.

In his time, Quetelet stated the same, only in another wording. He wrote that, when studying phenomena of social life, “we ought to leave aside the man taken be himself, and only to consider him as part of mankind”, to be interested not in the individual, but in the totality of men.

Chuprov indicated that the statistical approach to the phenomena in the real world also created peculiar logical forms of knowledge connected not with the essence of the object, as it was generally thought, but with the viewpoint from which we consider it³. This is one of the initial principles of Chuprov’s doctrine. However, before describing its main propositions, it is necessary to indicate that his subject was the statistical method in itself rather than statistics as an independent science.

[3] Concerning the existence of such a science, Chuprov held to that dualistic notion which had been developing in the German literature as well as by some Russian theoreticians in the second half of the 19th century according to which the statistics as a science should be distinguished from the statistical method. He (1906/1960, p. 127) therefore believed that statistics as an independent science ought to be a discipline studying social phenomena because the generally recognized cultural and

practical values primarily had to do with the society: “This only allows the birth of a statistical science restricted in its subject”. He also expressed himself in the sense that also in future the preservation of statistics as an independent science can only be imagined in the form of a descriptive science similar to the *Staatswissenschaft* (or *University statistics*) of old.

In a review of a book by Zizek Chuprov (1922, p. 339) wrote:

Zizek convinces me in the correctness of the opinion that the future of material statistics lies in the sharp advancement of idiography, in a systematic return to the old descriptive Staatswissenschaft. [...] Statistics as an independent science will assume the form of a systematic description of remarkable facts, of various social formations precisely restricted in time and space. [...] Otherwise it will be entirely unable to enjoy the significance of an independent science.

Here, Chuprov mixed up two different aims of statistics, the collection and systematization of statistical data, of great practical importance for the work of contemporaneous statistical organizations, but not properly speaking of a scientific nature, and the analytical problem posed by the researcher applying statistical methods⁴. Chuprov’s views testify once again how strong in science is sometimes the influence of previous delusions and we ought to agree with Chetverikov (Chuprov 1960, pp. 3 – 4) who remarked on this point that “Nowadays the defence of such reactionary directions in statistics is not anymore possible”.

While developing the theory of the statistical method, Chuprov declared that the statistical forming of notions ought to be connected not with the essence of the studied object, not with the real distinctions between phenomena, but with our viewpoint concerning them [see previous note]. He formulated this thesis wishing to refute the traditional concept yet existing in the relevant literature that the statistical method was allegedly mostly applicable for studying phenomena of the society and only plays a restricted and subordinated part in natural sciences.

[4] This concept is known to be based on metaphysically contrasting the individuality of social phenomena and typicality of natural phenomena. The first to state it in Russian writings on the theory of statistics was Poroshin (1838): in the “moral” world there is development and “life endlessly renews its forms” but the movement in nature is only circular along “forever invariable” paths and the general might be therefore directly cognized in the particular.

Chuprov mostly attempted to refute this metaphysical viewpoint by pointing out a number of branches of natural sciences into which statistical methods of research had already entered. He (1909/1959, pp. 75 – 76) also corroborated this negative attitude by general considerations indicating that natural sciences had abandoned the trust in that under any conditions it was possible to form a general conclusion from a single observation considered as “typical” and that any doctrine stating that natural phenomena are typical and the sociological phenomena are atypical “can be thought irrevocably obsolete”.

Accordingly, Chuprov insisted that the statistical method was universal and that the “statistical forming of concepts” should not be restricted to the phenomena in social life. Believing that the way of thinking about an

object determined the statistical approach to phenomena, he (1909/1959, pp. 77 – 78) wrote:

To keep nowadays to the opinion that the essence of the statistical method is in any measure connected with peculiarities of man and society as an object of study is only possible for a statistician who does not follow the development of scientific thought beyond the sphere of his speciality. The aim of the investigation, the statistician's viewpoints on approaching the studied phenomena rather than the features of the material define the unification of separate phenomena into totalities.

According to him (1909/1959, p. 75) the traditional doctrine of the typicalness of natural phenomena and atypical phenomena of society was not only theoretically wrong, it did not foster the successful building of the theory of the statistical method either:

To subordinate the peculiar distinctions of the statistical methods of scientific work to the features of society considered as an object of study had led, on the one hand, to the logicians' scornful attitude towards the problems of statistical methodology. Their peculiarity, as it seemed, had a technical rather than general significance caused by the essence of the material and deserved a place in the general theory of science as little as for example the theory of microscopic technique. On the other hand, it tore statistics away from natural sciences and hindered the introduction of methods developed in sociology into the sphere of similar problems in those sciences. As a result, logic, the general theory of statistics, and the separate disciplines applying statistical methods had been suffering⁵.

Even before Chuprov the writings on the theory of statistics had indicated that the statistical methods were entering various branches of natural sciences. Some theoreticians noted that in many cases the typicalness of natural phenomena was only apparent, that what seemed to be a dynamical law (in the philosophical language) was actually a regularity of a mass phenomenon hidden behind randomness. Thus, A. Wagner (Drobisch 1867, p. 141), an eminent German statistician of the 19th century, wrote that “even the law of the fall of a body actually lacks a quite exact expression since it presumes vacuum and is modified by air resistance, friction etc”. Drobisch (Kaufman 1928, p. 7), another celebrated German statistician, also indicated that “in a rigorous sense” even the Keplerian laws, strictly speaking, are included among those deduced from a large number of cases “because they only determine the mean planetary paths from which the planets continually deviate now in one direction, now in another one”⁶.

However, before Chuprov no one stated that the principle of the universality of the statistical method alone which presupposes a “statistical formation of notions” only in connection with the viewpoint on the studied object allows to reveal the logic of that method, that the theory of statistics lacking in the mentioned principle is impossible.

[5] To appreciate Chuprov's standpoint it is necessary to bear in mind that he published his contributions at the time when the general principles of science had changed. The reasoning about the typicalness of the natural

phenomena which restricted the application of statistical methods in that field could have seemed convincing for his predecessors, but at the turn of the 19th century, because of the general development of natural sciences, that position became groundless.

It was Darwin, the creator of the evolutionary theory, in the first place, who refuted the former notion about the invariability of species (Engels 1873 – 1882/1985):

Darwin, in his epochal contribution, issues from the most general factual basis which rests on randomness. Exactly these infinitely many accidental distinctions between individuals of the same species can intensify until they find themselves beyond the bounds of the species' indication. Their immediate causes can only be established in rarest cases and exactly they compel him to doubt the previous foundation of any regularity in biology, the idea of a species in its previous metaphysical ossified and invariable form.

Engels also noted that in the 19th century natural sciences had changed their essence, becoming a science not of invariable objects, but rather of their origin and development, a science of processes.

Apart from the Darwinian evolutionary doctrine other great discoveries were made in the second third of that century which have indeed been changing natural sciences into a science of processes, viz., the cell theory and the discovery of the law of conservation and conversion of energy. And by the very end of the century, in the mid-1890s, there occurred, as Lenin expressed it, “the newest revolution in natural sciences”. In particular, it was marked by the discovery of the first laws of the quantum theory. Given this situation, natural sciences really had to abandon the “trust in the possibility of formulating a general conclusion valid under any conditions on the basis of one single observation considering it typical”⁷.

This meant that in their experimental work natural sciences should have inevitably turned to statistical methods in spite of sometimes meeting opposition to that innovation⁸. Chuprov (1909/1959, pp. 21 – 22) stated that

After a fierce battle, statistics invariably held the captured place. Meteorology and anthropometry (together with anthropometry) were the first to submit. Then came psychology (physiological psychology) and the biological sciences. In botany and zoology, the development of a number of main problems gradually passes on to the statistical sphere. Nowadays, in its empirical foundation, the theory of evolution rests mostly on the material of mass observations. [...] And during the latest years statisticians belonging to the mathematical direction, encouraged by the success in the field of biological sciences, raise the hand against astronomy as well⁹. At the same time statistics is celebrating victory in some branches of applied natural sciences, especially in agronomy where ever more weight is being assigned to mass observation, ever clearer becomes the insufficiency of a single experiment for firmly justifying conclusions and the complicated methods of calculations not even

avoiding the mathematical theory of probability are spreading ever broader.

Chuprov could have therefore rightfully claim that the statistical method was universal and that only allowing for that fact was it possible to ascertain its logical nature.

[6] Turning to the logic of the statistical method, Chuprov indicates that when studying mass phenomena it should replace the usual rules of inductive logic whose application presupposes unique connections between phenomena. In reality, conditions of scientific work, necessary for inferences based on these rules, are lacking, and Chuprov (1903, p. 34) states:

The researcher always ought to suppose that he is treating a connection defined by plurality of causes and actions, a free connection as we will call it, so that one and the same cause can lead to several different consequences, and several different causes can precede one and the same consequence. He is only justified in admitting that the connection which he aims at establishing is continuous, is such that neither a cause without its consequence, nor a consequence without its cause can occur, when obtaining a direct proof of that fact. But inductive methods provided by logic are unsuitable for establishing free connections.

It followed (1903/1960, p. 34; [II, § 1.3]) that such new methods of investigation were needed which did not rest on the peculiar features of continuous connections:

The set of such methods of studying causal connections and suitable for the case of free connections constitutes exactly what is usually called the statistical method.

Its defining distinction, as Chuprov believed, was its being based on the theory of probability [II, § 5.4]:

I have therefore attempted to show how does the statistician, when collecting separate phenomena in statistical totalities, exclude corrupting chance from the empirical frequencies, reveal their underlying objective probabilities and establish by comparing the obtained series whether there exists an interdependence between the studied phenomena.

The statistician, while treating statistical data and grouping them in accord with the values of one or another indicator,

Excludes the corrupting chance from the empirical frequencies, reveals their underlying objective probabilities and establishes by comparing the obtained series whether there exists an interdependence between the studied phenomena.

The entire investigation is here based on the law of large numbers, which, as Chuprov expressed it, serves as a bridge of sorts thrown between empirical frequencies and the objective probabilities which can only be

determined a priori¹⁰. Such, as he imagined, were the logical functions of the statistical method which replaces the usual inductive methods when dealing not with continuous, but with free connections, more important for scientific research.

Even before Chuprov some statisticians indicated that the logical aspect of the statistical method was peculiar and distinct from the *canonical* methods of inductive logic as formulated long ago by Bacon and Mill. Kaufman (1928, p. 274)¹¹ remarked that unique connections cannot be reflected by contradictory statistical indications varying under the influence of different chance factors:

It was noted a long time ago [...] that in the sphere of those phenomena studied by the statistical method induction presents some essential features so that the pertinent inferences are not as firm and indisputable as those rendered by induction.

And even in the 19th century G. Rümelin (Yanson 1879, p. 171), an eminent German statistician, while ascertaining the peculiarities of the statistical method, stressed that it “asserted itself exactly where induction was unable to derive conclusions by considering a typical concrete case suitable for other ones”.

Also before Chuprov, Reichesberg (1893/1898, p. 117), a Swiss statistician, considered the distinction of the statistical method from the usual rules of inductive logic:

In our opinion, the statistical method ought to be considered as enabling to understand the complicated mechanism of mass phenomena of every possible kind when it is impossible to infer either inductively from the particular to the general, or deductively from the general to the particular.

He (Ibidem) also referred to A. Onken who even previously, in 1870, stated that, when applying the methods of usual logic, mass phenomena remained incomprehensible. All these remarks, however, were only made in passing, whereas, for the first time in the history of theoretical statistics, Chuprov vividly elucidated the logical aspect of the statistical method.

[7] Chuprov’s idea about plurality of causes and effects is the pivot of his entire doctrine and it is close to the propositions of materialistic dialectics on the interconditionality and interconnection of phenomena. However, he indicated that the problem of plurality of causes occurs in the researcher’s mind since he is unable to establish completely and thoroughly the complicated previous phenomena which he aims to connect with the considered action. This renders his entire concept metaphysical and subjective and leads away from correct ideas about real objective connections.

It also leads to a wrong explanation of the interconnection of phenomena as a simple sum of *pairs* of elementary causes and effects provided that the researcher had been able to derive them. An imaginary simplification of reality which leads the researcher in each separate case to decide about the existence of a causal connection occurs not by a simple separation of the elements of causation from their sum, as Chuprov

believes, but by tearing the studied phenomenon from its natural or historical ties.

Be that as it may, the idea of plurality of causes and effects (actions), the base for Chuprov's logical conception of the statistical method, enabled him to present clearer its features. Indeed, the *canonical* rules of the so-called [!] complete induction can misfire when applied to a statistical study¹². The general inference is known only to be made there after having collected such isolated data that exhaust all possible cases of the given kind, the researcher unites them in some logical group. [...] But in a statistical investigation, when, as Kaufman expressed it, "a full list of causes and effects is lacking", complete induction is useless.

[8] However, since complete induction is impossible for studying free connections (as Chuprov calls them), incomplete induction is still at our disposal and conclusions are made when only having a repetition of a part of the needed and are therefore stochastic. [...] When discussing the notion of probability Chuprov nevertheless slides off into the sphere of ideas corresponding to the agnostic Neo-Kantian *Weltanschauung* which he follows in his philosophical premises and statistical doctrine. He (1924/1960, pp. 188 – 189) believes that probability is objective, but not because it expresses some relations between real events:

It is not more real than a centre of gravity of a body, the Greenwich meridian, than the equator. But, like that centre, it is objective in the sense of being the outcome of theoretical thought free from an admixture of the feature of subjectivism peculiar to an individual cognizing mind.

Thus, for him, probability remains a method of study, of arranging the chaos of the complicated phenomena of the Universe by the human mind rather than a quantitative expression of an objectively existing possibility of the occurrence of an event. But he correctly understands the action of the law of large numbers, which, as he (1909/1959, p. 143) wrote, was a bridge connecting objective probabilities with the empirical frequencies: "It expresses those complicated half-nomographic and half-idiographic interlacing of events in the form of a free causal connection".

[9] I cannot pass over in silence Chuprov's attitude to the problem of stability as put forward by the German statistician W. Lexis and, almost at the same time [a bit earlier] by the French mathematician E. Dormoy. Even the early researchers among the political arithmeticians are known to have been surprised by those regularities which they were able to establish in statistical figures¹³. Graunt, Petty, Süssmilch, the founders of that science, indicated some regularities revealed in studied events when sufficient statistical data were at hand. This is certainly an important fact for comprehending the history of the origin and development of the statistical science.

Such discoveries made by means of numerical data had been able to prompt an idea of whether the researchers were dealing with something possibly forming a new science distinct from the previous knowledge. Had not a thought already then appeared of replacing the merely descriptive *Staatswissenschaft* by a new direction enabling to reveal and explain some regularities peculiar to society?¹⁴

Quetelet is known to have put forward such an idea in the first half of the 19th century. He clearly and minutely stated what the political arithmeticians had been vaguely feeling about establishing the regularities of social phenomena; it was necessary to study them in the mass rather than in isolation. And, drawing on concrete statistical material, he also attempted to show what kind of regularity can exist.

It follows from his deliberations and investigations that the discovered regularities were expressed by stability of statistical figures weakly changing in time. Exactly that problem of stability became the issue of the dispute that came about between German statisticians, opponents of the *statistical determinism* allegedly denying free will, and vulgar materialists following Quetelet.

[10] That dispute over the problem of such importance for comprehending the logic of the statistical method had only been speculative and therefore fruitless until Lexis and Dormoy attempted to indicate a concrete way of solving it. Lexis is known to have introduced a special indicator, Q , the coefficient of dispersion (of divergence) for measuring the stability. It allowed to compare the empirical dispersion with its theoretical counterpart by tentatively equating them over the whole series of observations. [...]

Chuprov certainly had to consider that problem of stability since it directly bore on the fundamental principles of the statistical method. He (1909/1959, p. 295) stressed first of all that the stability of statistical figures was a phenomenon changing together with the conditions of the environment of life. He thought that normally stable series did not actually exist, only relatively stable phenomena changing comparatively slowly can be found, and he also pointed out that the circumstances leading to one or another degree of stability of the studied phenomena were intricate:

We are able to indicate definitely such circumstances in the conditions of the existence of the society that are inclined to raise over the norm the stability of mass phenomena as well as such that raise their fluctuations. As a rule, these conflicting influences do not however act separately but are continually intermixing so that each separate mass phenomenon bears the stamp of their mixed set.

When interpreting the problem of stability, he thus attempted to issue from reality with which the most important problems of the theory of the statistical science were connected rather than from the philosophical disputes about *statistical determinism* and free will. For explaining the very mechanism of the origin of one or another degree of stability, he reasoned on the presence or absence of connections between separate trials¹⁵. The pertinent propositions threw additional light on the action of the law of large numbers on dependent trials and at the same time led to a deeper development of the main principles of the theory of sampling, which the British statisticians later achieved by means of their concept of within-class correlation.

[11] Chuprov's contributions certainly clarified the logical aspect of the statistical method. When abandoning the false and in essence unneeded references to the Baden philosophical school, his doctrine connecting the

statistical method of research with stochastic logic constituted an important period in the development of the statistical theory.

But did it end those debates over the scientific contents of statistics so widely spread during the second half of the previous century? This did not happen because statistics contained controversial points as well. Of course, much should have been cleared up here in the 20th century since it became possible to take into account all the previous development of the theory of statistics and the essential experience in applying statistical methods gained both in scientific investigations and practical work.

As I will show below, the idea that statistics is a doctrine of the methods of quantitatively investigating mass phenomena has been winning an ever greater place in the writings of Russian and foreign theoreticians of statistics. But the search for other answers to the question of *What is indeed statistics as an independent science* was continuing. Sometimes it led to a restoration of principles which were put forward almost a century ago and apparently could have been irrevocably abandoned as obviously unfounded.

At the beginning of the 20th century, some theoreticians also attempted to restore the idea, prompted by Quetelet's notions and developed by his followers, that statistics is a *science of sciences* of sorts among other disciplines studying society, is a science studying phenomena of social life in the widest sociological aspect. The eminent statistician A. Fortunatov was the most fervent representative of this direction in the Russian literature. In an inaugural lecture read in 1897 in the Novoaleksandrovsky Institute [Novoaleksandrovsk, Stavropol krai (territory)] but only published in 1905, he (1909, pp. 9 – 15) bluntly called statistics the “elder sister of sociology”. True, he also added a qualifying remark: “In future, when sociology shows its ability to forecast scientifically, statistics must become its younger sister”.

Elsewhere, while reasoning on various social sciences, or, as he called them, parts of social sciences, Fortunatov (1900, p. 11) wrote:

Some of these parts are restricted by a single aspect of social life. Thus, political economy aims at studying economics; jurisprudence busies itself with laws and ethics studies morality. Other parts of social sciences strive for embracing the entire sphere of social life and in any case from various sides. Such are history and its own sister, statistics.

Other Russian statisticians of the beginning of this century, for example A. N. Kotelnikov (1908, pp. 30 – 31) expressed similar thoughts. He also contrasted statistics as a most general science embracing various manifestations of social life, and particular sciences studying some of its separate aspects:

The analysis of general distinctive properties of social phenomena and the formulation of general laws of social life ought to belong to the statistician. It follows that statistics is a science of social elements, of general principles and general laws concerning all that is occurring in society, that it is the main science about society.

He then stressed that exactly statistics is such a main science rather than sociology falsely thought of having precedence:

Although the names of really outstanding authorities are connected with such an idea, we cannot admit it because not a sociologist but the statistician is applying the statistical method, is formulating some problems for statistically studying society. To him therefore belongs the problem of studying society as a whole.

Such ideas about statistics [...] had been justified in its time by the very novelty of statistical investigations allowing to establish some regularities in social life, by enthusiastic feelings about *the new science* apparently opening up extremely wide possibilities of such investigations. Even at the time when Yu. Yanson [1835 – 1893], for example, had been active, it was possible to state that statistics was a universal social science studying “everything happening in society”. However, it is difficult to explain why were those old and obsolete ideas restored even in the beginning of the 20th century after the scientific contents of statistics had been debated for almost a hundred years and the boundaries of the applications, and the very nature of the statistical method were sufficiently clearly determined. I can only say that such a setback was one of the causes for the continued delay in ascertaining the true situation during the entire course of the development of theoretical statistics.

[12] The dualistic concept to which Chuprov had been keeping (see above) was another brake. In this part of his doctrine he followed a number of his predecessors, those among German statisticians in the first place. I have remarked that the inherited ideas of forerunners, even if obsolete and abandoned by the development of science, can strongly influence the viewpoints of a scientist. Chuprov, who dared at the same time demolish another traditional concept, not less harmful for the progress of the statistical science, about the typicalness of natural phenomena, was deluded and his mistake certainly ought to be attributed to a peculiar *inertia of thought*. On this point he obviously contradicted himself, opposed the very spirit of his generally speaking progressive ideas.

However, in spite of his mistake, it would seem that later theoreticians of statistics should have surmounted it; actually, it occurred extremely persistent and along with other points had been preserved in a number of their writings and continued to swerve statistics from the proper path. As to the Russian theoreticians at the beginning of the 20th century, their such an eminent representative as N. A. Kablukov (1922, p. 33 ff) had continued to support the *concept of dualism*. He was inclined to consider the results provided by the statistical method in studying social phenomena, or, as he expressed it, the material part of statistics, as a special branch of knowledge. He also advanced the idea of statistics as a method, as a doctrine of that method.

He indicated that mass phenomena were the contents of statistics as an independent material science and believed that, being independent, it belonged to the group of social sciences with the statistical method as the method mostly studying phenomena of social life.

Not less clearly was the concept of dualism represented by the Ukrainian statistician K. G. Vobly (1918, p. 8). Supposing that statistics was “a science numerically studying mass phenomena of social life” he at the same time mainly restricted the sphere of application of the statistical method by these phenomena:

When investigating physical phenomena the quantitative (the statistical – N. D.) method is sometimes applied, but as compared with the observation and study of isolated objects, its significance is only secondary.

It is difficult to distinguish his pronouncements from the relevant considerations found in the writings of German and some Russian statisticians of the 19th century.

[13] Finally, addressing the past was also that treatment of the scientific contents of statistics that can be found even in modern textbooks on the theory of statistics (*Statistika* 1969, pp. 14 – 35). It consists in that statistics is an independent social science which, unlike political economy engaged in general laws of social phenomena under definite conditions of the given social and economical structure, quantitatively studies these phenomena under definite conditions of time and space.

When recalling that even Quetelet and his contemporary, Moreau de Jonnés (1847) expressed similar ideas, it becomes clear that the authors of the above definition had not advanced the comprehension of the scientific contents of statistics as compared with their remote predecessors.

Almost a century and a half ago, Moreau de Jonnés (1847/1859, pp. 2 – 3), while wishing to distinguish political economy connected with statistics “by tightest ties” and statistics proper, wrote that “political economy is a transcendental science bravely turning to the elevated field of speculative systems” whereas statistics is “only a science of facts”.

This interpretation of the contents of statistics as an independent science is mistaken since it is impossible to classify sciences as integral systems of knowledge belonging to definite fields and at the same time to separate them into theoretical and empirical since this is philosophically unfounded. To insist that political economy studies the essence of social phenomena and the general laws of their development whereas statistics, existing alongside as an independent science, provides their concrete numerical illustrations, is a denial of one of the main principles of dialectical materialism on the indissoluble unity of the qualitative and quantitative definiteness of objects and phenomena¹⁶. Such an uncritical repetition of the mistakes made by remote predecessors is obviously anachronistic (Druzinin 1952; 1961).

[14] In this century, did the theoreticians of statistics attempt to formulate anew the problem of the contents of statistics as an independent science? Yes, some Russian and foreign statisticians had indeed made such attempts. One of them, Ptukha (1916, p. 28ff), a noted Russian and Soviet theoretician and historian of statistics, excludes economic phenomena from the sphere of statistics and therefore only acknowledges population and moral statistics as the independent science of statistics. He justifies his opinion in the following way.

In the economic activities of man “separate individual cases [...] are so homogeneous that they are explained by one or another single cause or by their set”, so mass observations are not needed for their study. In addition, when invading economical phenomena, statistics forfeits its significance of an independent science; indeed, for such a science to exist, “it is also necessary that the same phenomena are not scientifically discussed by any other science” whereas the results of economical statistics do not satisfy that demand and only represent a “supplementary branch” of social politics, or management of the economy of the nation.

Actually, only concrete mass phenomena pertaining to the person or to his actions constitute the only sphere for the results of investigations made by statistical methods to build up an independent material science. As to the statistical method proper, Ptukha indicates that it acquires special significance exactly in the sphere of mass phenomena of social life, but he also admits that it can be applied to natural sciences as well in spite of their additionally possessing another method, “the natural scientific method precisely establishing causes and measuring the power of their actions”.

[15] The attempt to represent statistics as an independent science on population and behaviour of man was obviously unsuccessful. In each case in which statistical methods are applied, the studied phenomena belong to the field of the pertinent science. In demography, statistical methods are only studying processes constituting the subject of that science, the science on the laws of population in various social systems. Just the same, moral statistics is a means for concretely investigating problems posed by sociology. The proposal made by Ptukha only to recognize population and moral statistics as an independent science apparently ought to be also regarded as an attempt to return to the past. Indeed, in its time, it was exactly the statistical investigation of man’s behaviour that served as a basis for declaring the origin of a new science, of *social physics*, as Quetelet named it. In addition, Ptukha’s viewpoint undoubtedly includes traces of a dualistic concept.

Among eminent Russian statisticians attempting to define the subject of statistics as an independent science I also ought to name Slutsky [V, § 5] as well. He believed that that problem can be solved by analyzing those theoretical principles which are the base of the statistical methodology:

Isolating that which relates to the properties of, first, judgements and concepts, i. e., to logic, and then of the properties of quantitative images upon which it (statistics – N. D.) is operating¹⁷, i. e., of mathematics, we nevertheless obtain some remainder for which no acknowledged sanctuary is in existence, which remains uncoordinated and homeless until we perceive its special theoretical essence and provide it with the missing unity in the system of judgements fully deserving the name of theoretical statistics.

Slutsky names the doctrine on the properties of totalities, on curves and surfaces of distribution, on mean values etc. as playing the role of such a remainder which, as he supposes, can form an independent statistical science (Ibidem):

All this is not a logical doctrine of the world of judgement and concepts, but statistical doctrines of the world of phenomena in [the entirety of] their forms and mutual conditionality.

Slutsky's viewpoint is very close to the idea about the scientific contents of statistics held by Fisher (1925/1990, pp. 1 – 2), a famous British statistician and mathematician whose writings strongly influenced the development of mathematical statistics in this century¹⁸:

The science of statistics is essentially a branch of Applied Mathematics and may be regarded as mathematics applied to observational data.

And, when defining the contents of statistics as a science, he held that

Statistics may be regarded as (i) the study of populations, (ii) as the study of variation, (iii) as the study of methods of the reduction of data.

Fisher added vague explanations, but it is difficult to understand, does he consider statistics an independent science or a universal method. Thus, he writes:

In a real sense, statistics is the study of populations, or aggregates of individuals, rather than of individuals. Scientific theories which involve the properties of large aggregates of individuals [...] such as the Kinetic Theory of Gases, the Theory of Natural selection, or the chemical Theory of Mass Action, are essentially statistical arguments. [...] Statistical methods are essential to social studies, and it is principally by the aid of such methods that these studies may be raised to the rank of sciences. This [...] has led to the painful misapprehension that statistics is to be regarded as a branch of economics [...].

It can hardly been thought that Fisher was able to approach satisfactorily the solution of the issue of the scientific contents of statistics. His and Slutsky's viewpoints do not represent quite successful results of studying the pertinent sphere. All that, which they attribute to the contents of statistics actually belongs to it not as to an independent material science but rather as to a doctrine of methods.

The doctrine of totalities and their properties certainly cannot consist in a description of the concrete totalities encountered in statistical practice in all their infinite variety. The contents of that doctrine is the establishment of the general principles guiding the separation of such totalities for being statistically analysed and the determination of the methods of that analysis. And, as Fisher himself justly remarked, the idea of a totality is here applicable not only to a collection of [human] beings or simply to material objects, it can also refer to the mass results of any measurements. And here the researcher enters the field of the theory of errors which belongs not to statistics, but to mathematics (to the theory of probability)¹⁹. Again, the doctrine of totalities is connected with the distribution of values along with [the probability of] variations.

These problems are also methodological and again theoretically justified by the theory of probability. And, finally, the doctrine of means

which, as Slutsky believes, is also a component of statistics as an independent science, bears on the concrete methods of calculating them, and, in the aspect of the theory of knowledge, is an issue connected with the action of the law of large numbers whose interpretation is the most fundamental element of the entire statistical methodology.

[16] Thus, it is quite obvious that we infer that the scientific contents of statistics is the doctrine of methods applied for quantitatively studying mass phenomena. In truth, this is a fully logical conclusion from all the prolonged discussion which has been going on since the formation of that theory had begun.

It is therefore not surprising that in the 19th century some theoreticians of statistics (A. Guerry, A. Wagner, and, among the Russians, D. Zuravsky) were keeping to this view and that in the next century it had won a number of partisans (the Russians A. A. Kaufman and R. Orzentsky, the British statisticians G. U. Yule and M. G. Kendall, the noted American economist and statistician F. Mills, the most eminent Polish statistician O. Lange, the French scientist R. Dumas, the eminent Italian statistician R. Boldrini).

In the Russian statistical literature, this methodological direction had begun to form at the very end of the 19th century. Relevant pronouncements can be found in the writings of L. Hodsky, L. Fedorovich, G. Ster and I. Miklashevsky²⁰. The first two, while still considering statistics an independent science, indicated that the doctrine of methods constituted the theory of that method. Since it is the theory that determines any science, there remained only one step to acknowledging statistics just as a methodological science (Hodsky 1896) whereas Ster (1898) simply denied the possibility of forming statistics as an independent science and thought that all the factual material with which statistics has to do was distributed among the relevant sciences. He was inclined to define statistics as a methodological science and in essence, Miklashevsky (1901) generally sided with Ster.

Among Russian statisticians who even at the end of the 19th century had promoted the viewpoint that statistics was a doctrine of method, we may name A. I. Chuprov (1895) since he considered the acquaintance with “the means of applying the statistical method when studying social phenomena” as the main goal of that science.

Various arguments had been put forward in favour of recognising statistics as a doctrine of methods and of the impossibility of representing it as an independent material science. However, that statement was mainly justified by indicating that a material science ought to possess its special subject of study whereas statistics actually does not have it since all the data which it investigates belonged to other sciences.

As I indicated, such was the position of Ster and we also find the same idea in Kaufman (1928, p. 12):

The main and decisive argument against acknowledging statistics as an independent science is that statistics undoubtedly has no subject of its own. The advocates of the opposite opinion believe that this subject is the totality of those facts whose significance can only be established by quantitative mass observation. But the same facts also constitute the subject of political economy and criminal law, hygiene and the financial

science [...]. *The subject of statistics is therefore both broader and narrower than that defined by Yanson, Lexis, Mayr, Conrad et al; broader because phenomena not included in the field of social sciences are being studied by the statistical method and in ever higher considerable degree; narrower because all problems on which statistics dwells in that field, such as those concerning the structure, constitution and life of the society, are also studied by a number of other sciences. [...] Each of them more or less makes use of the statistical method which does not prevent them to remain by themselves and does not at all unite them into a single science, into statistical sociology or simply statistics.*

[17] Another eminent Russian statistician and a partisan of the viewpoint that statistics is a methodological discipline was Orzentsky. He (1914, pp. 1 – 2) put forward similar arguments in his textbook on mathematical statistics:

As a material science, statistics should have possessed its own special subject of study. However, the data obtained statistically belong to phenomena included in the field of other sciences. [...] The treatment of that data naturally demands special relevant knowledge, and, when the materials are being collected for practical purposes, their application presupposes special acquaintance with applied issues. It is clear therefore that statistics as a special material science does not exist.

And here is his definition of the scientific contents of statistics:

On the contrary, the study of the method itself is a special and independent problem which does not coincide with studying nature or the laws of some special field of knowledge. It constitutes the subject of a special and independent methodological science, of statistics.

The usual objection to such an understanding of statistics is that it thus becomes some science devoid of an object and, in addition, acquires the features of a *supplementary* discipline serving other sciences. It is in this manner that Kaufman, when identifying the concepts of *doctrine of method* and *supplementary science*, called it. He thought that in that role statistics can be considered a science just as rightfully “as the doctrine of measuring devices which belongs to the sphere of special education”.

Lange et al (1971, p. 57) also call statistics “a supplementary science” for political economy and add that “This [...] also takes place in respect to the relation of statistics to other sciences studying mass phenomena”. This obviously false statement is continuing to serve as a foundation for current criticisms of the methodological concept.

In addition, it is possible to say the following. When keeping to the bourgeois agnostic philosophy, and only recognizing the scientific method as a purely subjective notion, only as a totality of methods and rules arbitrarily created by the mind for *convenient* cognition, – then the possibility of regarding some doctrine of method as an independent discipline ought to be really rejected.

However, according to the dialectical materialistic point of view on the theory of knowledge, a scientific method should be considered as an

objective notion, as a most important part of logic. Method is what is extracted from the science itself in the course of its development. Both the scientific cognition itself and the method created by it are reflecting the essence and connections of objective reality. The forms of the organisation and movement of matter being the object of study of a given science demand that a suitable method be applied as well. On the one hand, the development of the doctrine of method is only possible when advancing the science to which it is being applied. On the other hand, the history of science testifies that the perfection of the method itself is a powerful impulse to develop science. A science poses a problem causing a search for one or another appropriate method [of its solution], and statistics, in addition to the doctrine of such methods, acquires its own subject although mediated by that science.

The Soviet philosopher Rakitov (1965, p. 96), for example, describes the role of method in science:

Any method of scientific cognition is known to be secondary since its structure is determined by the objective essence of the studied processes and phenomena and because this method represents the means of their cognition and reflection. However, scientific methods, in addition, insofar as they are created and really exist, are not arbitrary and contain indications of objectivity. Having to do with methods of cognition checked by experience and ensuing an objective value of knowledge derived by applying them, their contents is objective.

Exactly in this respect scientific methods can serve as an object of cognition and in such cases no elements of subjectivism are contained in the statement that certain methods of cognition are the object of one or another scientific discipline.

[18] Some theoreticians of statistics considered it a methodical discipline mostly belonging to the sciences of society. Thus, even in the 19th century Bruno Hildebrand (Kaufman 1928, p. 3) called statistics the art “of measuring political and social phenomena”. A. I. Chuprov (see above) thought that the doctrine of methods was the main object of the statistical science but he also bore in mind the study of phenomena of social life by these methods. However, the developing scientific knowledge, accompanied by the ever wider penetration of the statistical methods into investigations of the phenomena of both society and nature, inevitably had to lead to the recognition of the universality of that method. In this century, all theoreticians of the statistical science while acknowledging its methodological essence are keeping to this viewpoint. Kaufman (1928, p. 8), for example, wrote:

Statistics or the statistical method is interlaced with political economy and economical politic, criminal law, medicine and hygiene, linguistics, meteorology; the statistical method is being applied in civil and military management, taxation, famine relief, insurance business, drawing-up and critical consideration of electoral laws etc. The sphere of its application has therefore no clearly defined boundaries. ... Statistics is the method of measuring or calculating social and a number of other kinds of mass and in general complicated phenomena.

The celebrated British statistician Bowley (1926/1930, p. 2), wishing to stress this wide-ranging significance of statistical methods in scientific investigations, thus expressed this idea:

Statistics is not a section of political economy and is not restricted to any single science. Its knowledge is similar to the knowledge of foreign languages or algebra, it can prove itself useful at any time and under any circumstances.

The statement that the statistical method is not playing any role in concrete investigations either in social or natural sciences would have nowadays ring hollow. And insofar as social and economic investigations are concerned, Lenin's economical contributions provide classical examples of applying statistical methods²¹. While describing the role of statistics in such studies, he is known to have called it one of the most powerful tools for cognition. And, when stressing that statistical methods themselves are greatly important for the development of science, he (*Polnoe...*, vol. 24, p. 281) indicated that properly compiled statistical tables can "really overturn the science of economics of agriculture".

Notes

1. Quetelet had not left any doctrine; his innovations were bold but not properly thought out (Sheynin 1986).

2. See my Introduction II.

3. This is a very strange principle. Herschel (1817/1912, p. 579), without knowing that the stars belonged to different classes, and that their sizes tremendously differed one from another, reasoned about their mean size, – about a meaningless magnitude.

True, Markov (1908/1924, p. 2) said something similar to Chuprov's statement:

Various concepts are defined not by words, each of which can in turn demand recognition, but rather by [our] attitude towards them ascertained little by little.

4. Druzinin left out Chuprov's important specification to the effect that the new form of the olden qualitative Staatswissenschaft will primarily have to do with numerically described mass phenomena, so that he did not foresee a resolute return to the past. Chetverikov did quote Chuprov in full but somehow paid no attention to that qualifying remark.

5. I do not believe that natural science had been "suffering" because of that attitude. On the contrary, many natural scientists, e. g., Poincaré (Sheynin 1973, p. 296), opposed the application of the statistical method beyond their field.

6. Wagner was hardly acquainted with the appropriate literature. Just one example: that water boiled at various temperatures depending on the air pressure had been known at least since the 18th century. And deviations of the planets from their mean paths are caused by random disturbances; anyway, in this example Drobisch discussed treatment of observations inevitably corrupted by errors, as was known to a certain extent even to Ptolemy.

7. See previous Note. Also, the law of conservation and conversion of energy had been repeatedly specified for each new form of energy studied.

8. This problem deserves more attention. Physicists (and especially Boltzmann) naturally attempted to describe their findings in terms of classical mechanics, i. e., quite rigorously which was impossible by means of the statistical method.

9. Chuprov's description of the "battles" was faulty, cf. my Introduction II.6. In particular, it was Humboldt who had suggested the term *anthropometry* to Quetelet who put it into circulation thus taking away the appropriate sphere from *anthropology*.

10. If at all possible. See also [II, Note 13].

11. Druzinin should have noted that Kaufmann's contribution did not appear *before* Chuprov's writings, as implied by the context.
12. Logicians are distinguishing between mathematical and complete induction. Indeed, Novoselov (1970/1976, p. 257) discussed the latter which occurred patently different from the former and the *Enc. Brit* (14th edition, vol. 12, 1964, p. 278 of the item *Induction* (pp. 278 – 281) indicated that complete induction means complete enumeration of cases.
13. Political arithmeticians had been happy rather than puzzled. Indeed, as they believed, regularities reflected Divine care for mankind.
14. At the time of Graunt and Petty the Staatswissenschaft had not yet been really developed.
15. It was Lexis who first introduced connections into his theory of stability of series.
16. It follows that statisticians ought to work together with mathematicians, or be mathematically qualified themselves. For a long time Soviet statisticians had been, however, quoting that materialistic principle as an excuse to deny mathematics, to preserve Marxism in its initial form (Sheynin 1998, especially §§ 3.5 and 5.1).
17. I am translating Slutsky's paper [V] and indirectly note there that Druzinin's decision that *it* means *statistics* seems wrong; I would say, *logic*. Properties of concepts rather belong to philosophy.
18. The very origin of mathematical statistics is closely linked with Fisher.
19. The stochastic theory of errors had been a most important chapter of probability theory from the mid-18th century to the 1920s but mathematical statistics borrowed its principles of maximum likelihood and minimal variance from the error theory. Today, the stochastic theory of errors is the application of the statistical method to the treatment of observations. Incidentally, when defining statistics, Fisher (see above) forgot about it. There also exists the determinate theory of errors.
20. Chuprov (Bortkevich & Chuprov 2005, Letters 28, 1897, and 54, 1900) had an extremely low opinion of all the mentioned statisticians. Bortkiewicz agreed with his colleague in regard to Ster and Hodsky (Ibidem, Letter 29, 1898).
21. For a sober opinion about these contributions see Kotz & Seneta (1990).

Bibliography

- Bortkevich V. I., Chuprov A. A.** (2005), *Perepiska* (Correspondence), 1895 – 1926. Berlin. Also at www.sheynin.de
- Bowley A. L.** (1926), *Elements of Statistics*. London. Russian transl.: Moscow – Leningrad, 1930.
- Chuprov A. A.** (1903, in Russian), Statistics and the statistical method. Their vital importance and scientific aims. In Chuprov (1960, pp. 6 – 42).
- (1906), Statistik als Wissenschaft. *Arch. f. soz. Wiss. u. soz. Politik*, Bd. 5 (23), pp. 647 – 711. Russian transl.: Chuprov (1960, pp. 90 – 141).
- (1909), *Ocherki po Teorii Statistiki*. (Essays on the Theory of Statistics). Moscow. Later editions: 1910, 1959.
- (1922, in German), Review of F. Zizek (1921), *Grundriss der Statistik*. München – Leipzig. *Nordisk Statistisk Tidskrift*, Bd. 1, pp. 329 – 340.
- (1924), Ziele und Wege der stochastischen Grundlegung der statistischen Theorie. Ibidem, Bd. 3, p. 433 – 493. Russian transl.: Chuprov (1960, pp. 162 – 221).
- (1960), *Voprosy Statistiki* (Issues in Statistics). Moscow. Coll. reprints and translations. Editor, N. S. Chetverkov.
- Chuprov A. I.** (1895), *Statistika* (Statistics). Moscow.
- Drobisch M. W.** (1867), *Die morale Statistik und die menschliche Willfreiheit*. Leipzig. References to Russian transl.: Petersburg, 1867.
- Druzinin N. K.** (1952, in Russian), On the contents of statistics as a science. *Voprosy Ekonomiki*, July.
- (1961, in Russian), On the essence of statistical regularities and on the contents of statistics as a science. *Uchenye Zapiski po Statistike*, vol. 6, pp. 65 – 77.
- (1963), *Khrestomatia po Istorii Russkoi Statistiki* (Reader in the History of Russian Statistics). Moscow.
- (1979), *Razvitie Osnovnykh Idei Statisticheskoi Nauki* (Development of the Main Ideas of the Statistical Science). Moscow.
- Engels F.** (written 1873 – 1882), *Dialektik der Natur*. In Marx K., Engels F. (1985), *Gesamtausgabe*, Bd. 26, pp. 5 – 283. Berlin.

Fedorovich L. V. (1894), *Istoria i Teoria Statistiki* (History and Theory of Statistics). Odessa.

Fisher R. A. (1925), *Statistical Methods for Research Workers*. In author's *Statistical Methods, Experimental Design and Scientific Inference*. Oxford, 1990. Separate paging for each of the three items.

Fortunatov A. (1900, in Russian), Quantitative analysis in sociology. *Narodnoe Khoziastvo*, Bk. 1.

--- (1909, in Russian), Sociology and statistics. *Vestnik Vospitania*. Druzinin's reference is ambiguous: he also mentioned the year 1905.

Herschel W. (1817), Astronomical observations and experiments tending to investigate the local arrangement of celestial bodies in space. *Scient. Papers*, vol. 2, pp. 575 – 591. London, 1912, 2003.

Hodsky L. V. (1896), *Osnovaniya Teorii i Tekhniki Statistiki* (Principles of the Theory and Methods of Statistics). Petersburg.

Kablukov N. A. (1922), *Statistika* (Statistics). Moscow.

Kaufman A. A. (1928), *Teoria i Metody Statistiki* (Theory and Methods of Statistics). Moscow. Posthumous edition with many additions, notably by V. I. Romanovsky, greatly strengthening the theory of probability and mathematical statistics which obviously contradicted the opinion of the late author. German translation of previous edition (1913), *Theorie und Methoden des Statistik*. Tübingen.

Kotelnikov A. N. (1908, in Russian), Is statistics a science and what is its contents? *Vestnik Znania*, No. 8.

Kotz S., Seneta E. (1990), Lenin as a statistician. *J. Roy. Stat. Soc.*, vol. A153, pt. 1, pp. 73 – 94.

Lange O., Banasinsky L. (1971), *Teoria Statistiki* (Theory of Statistics). Moscow. Information on original Polish contribution not provided.

Lenin V. I. (1961), *Polnoe Sobranie Sochinenii* (Complete Works), vol. 24. Information on initial publication not provided.

Miklashevsky I. (1901, in Russian), Statistics. *Enzikopedichesky Slovar Brockhaus & Efron*, halfvol. 62, pp. 476 – 505.

Markov A. A. (1908), *Ischislenie Veroiatnostei* (Calculus of Probability). Petersburg. German transl., Leipzig – Berlin, 1912. Other editions: 1900, 1913, 1924.

Moreau de Jonnés A. (1847), *Eléments de statistique*. Paris. Russian transl.: Petersburg, 1859.

Novoselov M. M. (1970, in Russian), Induction. *Great Sov. Enc.*, vol. 10, 1976.

Orzentsky R. (1914), *Uchebnik po Matematicheskoi Statistike* (Textbook on Math. Statistics). Petersburg.

Poisson S.-D., Dulong P. L. et al (1835), Review of manuscript: Civile, Recherches de statistique sur l'affection calculeuse. *C. r. Acad. Sci. Paris*, t. 1, pp. 167 – 177.

Poroshin V. S. (1838), *Kriticheskoe Issledovanie ob Osnovaniakh Statistiki* (Critical Study of the Principles of Statistics). Petersburg.

Ptukha M. V. (1916, in Russian), Essays on the history of the population and moral statistics. *Zapiski Juridicheskoy Fakultet Petrogradskii Univ.*, No. 4.

Rakitov A. A. (1965, in Russian), The contents and structure of statistics and its role in the theory of knowledge. In *Nekotorye Problemy Metodologii Nauchnogo Issledovania* (Some Issues in the Methodology of Scientific Research). Moscow.

Reichesberg N. M. (1893), *Die Statistik und die Gesellschaftswissenschaft*. Stuttgart. Russian transl.: Petersburg, 1898.

Sheynin O. (1973), Finite random sums. Hist. essay. *Arch. Hist. Ex. Sci.*, vol. 9, pp. 275 – 305.

--- (1986), Quetelet as a statistician. *Ibidem*, vol. 36, pp. 281 – 325.

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

--- (1999), Statistics, definitions of. *Enc. of Stat. Sciences*, 2nd ed., vol. 12, 2006, pp. 8128 – 8135.

Statistika (1969), *Statistika* (Statistics). Moscow.

Ster G. (1898), *Kratkii Kurs Statistiki* (Short Course in Statistics), pt 1, No. 1. Kazan.

Vobly K. G. (1918), *Statistika* (Statistics). Kiev.

Yanson Yu. E., Editor (1879), *Istoria i Teoria Statistiki v Monografiakh Wagnera, Rümelina, Ettingena i Schwabe* (History and Theory of Statistics in the Monographs of ...). Petersburg.

IV
E. E. Slutsky

**Theory of Correlation
and Elements of the Doctrine of the Curves of Distribution***

Introduction

Kiev, 1912,
Izvestia Kiev Kommerchesky Institut, Book 16

During the two latest decades, theoretical statistics has greatly advanced. Perfection of old methods; discovery and development of new ones; appearance of excellent works on biology and social sciences illustrating methods and proving their unquestionable scientific significance; finally, creation of a yet small personnel of scientists systematically applying and developing the new methods further, – all this, taken together, allows us to say that a new era has originated in statistics.

This movement had started and has been developed in England, and it is only beginning to penetrate other nations. Initiated by the recently deceased celebrated Francis Galton, it grew out of the demands of contemporary biology. Galton, however, was not a mathematician, and the merit of theoretically developing new ideas and establishing a school must almost solely be credited to Karl Pearson whose name will remain in the history of our science alongside those of Laplace, Gauss and Poisson¹. In all fairness, the new school ought to be therefore called after Galton and Pearson.

The general awakening of interest in theoretical statistics allows us to expect that not in a very remote future the ideas of the new school will spread over all nations and all fields of their possible application, and I am humbly aiming at fostering that natural and inevitable process. The application of the new methods is comparatively easy and not difficult to learn. For making use of formulas, it is sufficient to understand their meaning and be able to calculate what they indicate, a task simplified by applying special tables also compiled on Pearson's initiative.

However, it is impossible to manage without breaking from routine. Unforeseen details can be encountered in each problem, and the boundaries of the applicability of a method, and the significance of the results obtained can perplex a student. Not only prescriptions for calculation are therefore needed, it is also necessary to comprehend the spirit of the theories and of their mathematical justification. Life itself thus raises a most important demand before those working at statistics: A statistician must be a mathematician because his science is a mathematical science.

It is for this reason that I had paid so much attention to formulas and mathematical proofs; nevertheless, one more point also played a certain role. Dry prescriptions are only good enough for being applied in old and

* A report read 5 April 1912 at the sitting of the Society of Economists

firmly established spheres. I believe that no success can be expected in planting new methods in new soil without justifying them.

The sphere of mathematical knowledge needed for understanding most of the provided derivations and proofs is comparatively small. Most elementary information on analytic geometry and differential calculus as can be acquired in a few days is sufficient for understanding the elements of the theory of correlation. Further generalization in that area as well as the first part of my work dealing with curves of distribution demand somewhat wider mathematical knowledge.

I have attempted to satisfy different groups of possible readers and the proofs are therefore simplified as much as a rigorous description allowed it. Those mathematical derivations which I thought understandable to least prepared readers are provided in more detail than necessary for accomplished mathematicians. Finally, I attempted to elucidate the material in such a way that the reader, even after skipping a difficult place, will be able to pick up the lost thread and understand the meaning of the formulas and the manner of applying them. I do not however flatter myself by hoping to have solved that problem quite satisfactorily.

My main subject is the theory of correlation but I did not feel it possible to avoid the theory of the curves of distribution which I described far, however, from comprehensively and possibly even too concisely, in pt. 1. I advise readers poorly acquainted with mathematics and only mainly interested in the method of correlation, to go over to pt 2 immediately after acquainting themselves with the first four sections of pt. 1.

Note

1. Where are Chebyshev, Markov, Liapunov?

Bibliography

Kolmogorov A. N. (1948, in Russian), Slutsky. *Math. Scientist*, vol. 27, 2002, pp. 67 – 74.

Ondar Kh. O., Editor (1977, in Russian), *Correspondence between Markov and Chuprov on the Theory of Probability and Mathematical Statistics*. New York, 1981.

Sheynin O. (1990, in Russian), *Chuprov: Life, Work, Correspondence*. Göttingen, 1996.

--- (2005), *Theory of Probability. Historical Essay*. Second edition. Berlin, 2009. Also at www.sheynin.de

V
E. E. Slutsky

Statistics and mathematics. Review of Kaufman (1916)
Statistichesky Vestnik, No. 3 – 4, 1915 – 1916, pp. 104 – 120

[1] Kaufman's treatise, now in its third edition, is certainly an outstanding phenomenon in our educational statistical literature, and not only in our as testified by the reviews of its German edition (1913) written by the most notable representatives of the European statistical thought¹. This third edition will also obviously find many friendly readers the more so since in its main parts and especially in its first theoretical part it is entirely recast as compared to 1912.

However, those who attentively followed the evolution of Kaufman's work will not fail to note that at least in one respect this third edition is not a simple development of the previous one but as though some new stage in the author's statistical Weltanschauung. Indeed, the author intended both the second and the third editions as a manual for those wishing to prepare themselves for working in statistics but lacking that mathematical background necessary for *entirely* mastering statistical theory and methods.

The author (1912, p. 235) believed (and believes) that

It is hardly possible to master consciously the principles of the statistical theory [...] without [its] connection with the main principles of the theory of probability.

He therefore devoted sufficient efforts and place to provide his readers with a possibly more distinct idea about both the theory of probability and its application for solving fundamental issues of the theory of statistics. As to the practical application of the formulas and tricks of the higher statistical analysis, the author (p. 236) properly and tactfully warned those insufficiently prepared:

Thoroughly perceive the boundaries of your competence. [...] In particular, certainly abstain from mechanically applying final formulas provided by mathematical statistics without being quite clearly aware of their intrinsic meaning and sense, otherwise misunderstanding can often result.

For consoling his readers he (p. 235) stated that he was sure that "In its current state, statistics still leaves for them an infinitely broad area of activity". True, he (p. 234) apparently did not entirely get rid of his serious doubts about the issue of the interrelations between statistics and mathematics and while acknowledging that "It is *hardly possible to resolve* the difference of opinion among the representatives of the statistical theory (my italics), he even avoided any attempt to clear up this matter in his manual. Given these circumstances, his practical way out, as mentioned above, to which he became inclined, could have only been welcomed. This is all the more so since the general outline of his

introduction to the theory of statistics and a number of other instances (see, for example, his very indicative remarks on pp. 131 and 132) allow to think that it were considerations about the difficulties of mathematical methods rather than doubts about the principles themselves that compelled the author to hesitate.

[2] That practical dualism is not at all specifically peculiar to statistics and is observed in other sciences and reflects the distinction between the individual features of the researchers and the subjects of their work (theoretical and practical astronomy, theoretical and experimental physics etc). However, in this third edition it became transformed into a dualism different in principle, the dualism between statistical theory and practice (p. 148):

As a rule, because of the very properties of this [statistical – E. S.] material statistical analysis does not allow, and because the structure based on that data is coarse and at the same time complicated, does not demand the application of formulas of the calculus of probability. However, this does not at all contradict the fact that each such structure is entirely based on the principle of probabilities.

But the author (p. 153) also keeps to his previous divide between statisticians who “follow and will follow the routes demanding application of more or less complicated forms of mathematical analysis” and others who “while treating [...] statistical material and interpreting its results, may restrict their efforts by elementary methods of calculation”.

This motif now seems rather inconsistent with the previous. Indeed, how is it possible to reconcile the right of a purely practical distinction only founded in essence on the division of labour between the researchers and the abovementioned standpoint negative in principle, or the author’s statements (p. 152) that such procedures as the construction of frequencies of distribution, adjustment of series etc “not only do not help to elucidate the real features of the studied phenomena, but, on the contrary, can provide ideas corrupting reality” and that “the method of correlation does not add anything essential to the results of elementary analysis”.

Choose one or the other: either these procedures and methods are useless and therefore harmful and ought to be altogether abandoned; or, they are useful, but demand an understanding of their essence, meaning and boundaries of application which is at least partly possible even in a treatise intended for readers lacking sufficient mathematical background.

The dualism of the author’s point of view which is not objectively resolved in those texts becomes nevertheless somewhat explained after reading that (p. 147)

The issue of our right to apply [in the area of general statistics – E. S.] the methods of the calculus of probability is in any case left open, or, as he adds, open for me.

Objectively speaking, this pronouncement certainly only confuses the matter since the reader remains ignorant of the basis on which, as the author believes, his own arguments against applying the calculus of

probability are weakened and he is led to adopt it in the practical sense on the one hand and to candidly ignore it on the other.

And still, if I am allowed to express my general feeling, I ought to say that the main and the specific for the author is apparently at present not this previous hesitation and doubts which reflected the former stage of his scientific evolution, but the formed and almost firmly established conviction in that the statistical analysis does not either allow or demand probability theory.

I have thus returned to the quote from p. 148 with which I started to describe the present viewpoint of Professor Kaufman and I think that after all stated above I am compelled at least conjecturally to adopt it as the expression of the real opinion of the author, to assume it as the starting point and main object of my critical remarks below.

However, I have to begin elsewhere. Indeed, I am sure that the indicated dualism between statistical theory and practice is rooted much deeper, i. e., not in the author's understanding of the role of probability theory, but in his ideas about the essence of statistics, and that issue is not yet clarified in contemporary literature in any sufficient measure.

[3] Kaufman adheres here to the now apparently dominant point of view that statistics is a method or methodological doctrine and not at all a science with its own special subject of research. And I personally would have been prepared to adjoin somewhat the critical aspect of his considerations, provided he had sharpened his reasoning to allow for Chuprov's view whose idea of statistics as an ideographic science² he does not regrettably even mention in spite of its certainly being the most powerful argument possessed by the camp which Kaufman criticizes. True, I think that even that argument cannot be upheld, but Kaufman did not prove that. I will not dwell on this difficult point because of lack of space and the more so since here I am not really at any variance with Kaufman.

Distinctions between us start further on, and exactly where Kaufman believes to have concluded the issue, where he recognizes the methodological essence of statistics. Let us ask ourselves whether it is in essence indeed indifferent, as he (p. 17) thinks, whether "to discuss statistics as a *supplementary science*, or simply as a *methodological doctrine*". When allowing for the author's considerations, we, as it seems to me, ought to conclude, first of all, that he does not sufficiently clearly distinguish between the various versions of the term *statistics* (see his pp. 15 – 18) and does not follow up to conclusion the reasoning on the place of the statistical method in the system of logical knowledge.

I begin with the issue of the method itself. As a method, statistics is certainly not a science, but a technique, that is, a system not of reasoning, but of tricks, rules and patterns of practical cognizing work, whether applied systematically or not, conscientiously or unconscientiously, for scientific or practical goals. Just the same as addition and subtraction remain arithmetical operations independently from who is applying them and what for. This will become quite clear after analysing the contents of statistics as a methodological *doctrine*. We will find there, in particular, a number of propositions concerning even the most simple procedure of the statistical technique, enumeration of the elements of a totality and its necessary conditions and forms.

The methodology of enumeration based on the analysis of its very nature, allows us to see how practice is conditioned by the general properties of totalities on the one hand, and the properties of known logical operations on the other. To oppose, as Kaufman (pp. 13 – 14) does, statistical method and statistical art by issuing from indications external in regard to operations themselves, i. e., from the aims of the work, is in essence wrong even without allowing for the difficulty of drawing the necessary boundary which he mentions.

Whether the enumeration of social masses, say, is applied practically (for the aims of administration, say, as *statistical art* according to Kaufman) or for knowledge (*statistical method*) is of no consequence. Not technically, as the author believes, but according to its essence the nature of the operation will be the same as will be the conditions for it to be properly done; consequently, the corresponding reasoning belonging to statistics as a methodological doctrine will also be the same in both cases.

And the last inference: since this reasoning does not change with the aims of the operation, it follows that the location of the boundary depends not on Kaufman's decision or otherwise, but on the essence of the matter. And, incidentally, this means that a discussion of the issues concerning the contents of a science is not idle, is not to be decided by opportunistic considerations of expediency; no, it is important and, if properly formulated, fosters the deepening and the solution of the most general problems of science.

[4] I think that it is just as impossible to agree with Kaufman's arguments about the nature of statistics as a methodological science. As a system of considerations, statistics, understood in that sense is necessarily either a science or its part. Kaufman (p. 16) compares it with the doctrine of measuring devices which allegedly cannot be isolated as a special science. However, if that doctrine is not a *special* science, it is a part of another one, – of which, it ought to be asked, of the science which it provides with the means of research, or of that on which it is logically based?

Both alternatives fall away almost at once; the former, because measuring devices such as clocks and microscopes serve all or many sciences and purely practical needs as well, and the latter, since a complete theory of one and the same device as of an ideographic item demands the application of many sciences, such as mathematics, mechanics, physics, chemistry, psychology (recall the personal equation in astronomy) rather than one, etc.

After thinking it over, the entire issue of attributing the theory of measuring devices to a certain *theoretical* science becomes absolutely mistaken because the peculiar logical structure of such a doctrine is overlooked here. The considerations constituting that doctrine are united into a system in a manner absolutely different than in any theoretical science. Here, the systematic connection is conditioned not by the objective relations between things and their various aspects but by their teleological function with these things being seen as the means for attaining the aims of the researcher.

Hence the natural grouping of separate technical disciplines according to the pursued goals intersected with their partition for the sake of

achieving maximal possible homogeneity of the contents according to the essence of the underlying theoretical doctrines.

And it is now also understandable why in the process of teaching and elucidating some technical disciplines are more closely adjoined to those sciences from which they derive their theoretical elements (for example, the doctrine of physical measurements) whereas others are in the neighbourhood of those sciences which make use of their results (e. g., the doctrine of devices for psychological measurements). Finally, still other disciplines in addition possess external independence (metallurgy or the doctrine of fibrous substances). All this, however, is an issue of teaching and elucidation and has no direct bearing on the logical essence of the relevant doctrines.

[5] These considerations justify the independence of statistics as a technical or practical science which according to some tests admits in addition of separation into statistical methodology and statistical technique and at the same time leads us real earnestly to the problem of statistics as an independent theoretical science. Indeed, any practical doctrine, as Husserl (1900 – 1901) had discovered in an inimitable masterly way, certainly assumes some underlying theoretical doctrine justifying its propositions. Indeed, for proving the possibility of some goal by definite means we ought to perceive the connection between means and goal as between cause and effect. And the study of such connections leads us to a totality of considerations constituting a system whose main point is the essence and properties of the subject rather than of the goals.

We thus arrive at an analysis of the theoretical considerations on which statistical methodology is built. Isolating that which relates to the properties of, first, judgements and concepts, i. e., to logic, and then of the properties of quantitative images upon which it [logic]³ is operating, i. e., of mathematics, we nevertheless obtain some remainder for which no acknowledged sanctuary is in existence, which remains uncoordinated and homeless until we perceive its special theoretical essence and provide it with the missing unity in the system of judgements fully deserving the name of theoretical statistics.

All the existing various propositions of the doctrine of totalities and their general properties only provisionally adjoining methodological problems will belong here. We will thus have first of all the doctrine of the main formal properties of totalities; then, of their quantitative and structural forms (which now constitutes an essential part of the so-called *Kollektivmasslehre*, that is, of the doctrine of frequencies and surfaces of distribution, of means etc); then, also included will be a generalized formal doctrine of population, or, more correctly, of totalities of a changing composition, whose elements emerge, change their state and disappear, be they [individuals of a] population, trees in a forest or atoms⁴.

Finally, here also belongs the doctrine of the machinery of causes determining the frequency of phenomena rather than of separate events. All this is not a logical doctrine of the world of judgement and concepts, but statistical doctrines of the world of phenomena in [the entirety of] their forms and mutual conditionality.

Whether to separate them as a special subject for elucidation and teaching, certainly depends on our arbitrary opinion, but a special science emerges not by arbitrariness but intrinsic ties [of the appropriate

components], cognized as something objectively compelling, as establishing a systematic likeness and unity of the corresponding relations as well as of the considerations expressing our knowledge of their properties and ties [between them].

[6] And now we approach the issue of the relation of the calculus of probability to statistics. It only suffices to compare the contents of some purely mathematical treatise on the former with statistical reasoning on probability; for example, the contribution of Markov with the writings of von Kries [1886] or Chuprov, and the deep intrinsic heterogeneity of the problems, methods and of the very spirit of these writings becomes striking. And further considerations will show that these distinctions are based on the difference between the subjects.

Calculus of probability is a purely mathematical science⁵. How something is occurring is of no consequence to it; it deals not with *factual* but *possible* frequencies, not with their real causes but their *possible* probabilities. And the concept of probability itself is there quite different, is generalized and abstract. As soon as some number is arbitrarily assigned as the weight of each possible event and a number of definitions is made use of, the basis is prepared for building in a purely abstract way infinitely many purely abstract castles of combinations in the air, and of going over from those weights to the weights of various derivative possibilities (for example, of some groups of repeated occurrences of events).

For the *calculus* of probability, any enrichment of the concept of probability as compared with the above is useless, it would have nothing to do with it. Throwing a bridge from that ethereal atmosphere of mathematical speculations to the region of real events is only possible by abandoning the ground of the *calculus* of probability and entering the route of studying the real world with its machinery of cause and effect. Only thus we obtain knowledge about the ties between frequency and probability, justify [experimentally] the law of large numbers and find the basis for applying the calculus of probability to studies of reality.

Chuprov investigates free causal connections⁶; von Kries discovers the causal underpinning of games of chance and the actual justification for the tendency of frequencies to coincide with probabilities; Venn and Edgeworth attempt to build the very notion of probability on the concept of frequency⁷; – but nothing mentioned has any relation to the mathematical science of the calculus of probability. Here [in statistics], the mind operates not with ideal forms and quantities but with real things and phenomena although considered from an extremely general viewpoint⁸.

Above, I did not add anything to the essence of the doctrines of theoretical statistics, I had not even demanded the creation of such a science (always a somewhat dangerous enterprise) and only mentioned a number of existing doctrines and their intrinsic ties [with each other?]. If, however, it occurred that these doctrines constitute the main theoretical contents of statistical methodology, then I will hardly be mistaken when stating that statistics as a theoretical science does exist, that collective items, totalities considered as such, to whichever area they belong, are its subject.

Incidentally, it also follows all by itself that since statistics studies quantitative properties which we cannot ignore because of their part in the relations and ties peculiar to the subject of statistics, statistics should be

indeed considered a mathematical science, i. e., one of those sciences in which mathematical methods are essential and unavoidable⁹.

Then, it is natural that also in practical applications of theoretical statistics and statistical methodology, that is, in the practice of concrete statistical work on empirical data, mathematical methods are also unavoidable, and that there exists no essential boundary between various chapters of statistical practice in regard to the *subject* of study. It is only possible to distinguish in each chapter more elementary and more complicated problems, and, in general, problems of one and another theoretical type.

[7] The study of Kaufman's viewpoint only corroborate, as I believe, these considerations. Indeed, after formulating his essential objections to the application of the theory of probability to statistics he finally arrives at a conclusion whose considerable significance I ought to deny decisively. He (p. 147) assigns as the area of such application the set of *simplest* phenomena of population statistics and a certain part of phenomena in natural sciences, then (true, somewhat hesitatingly, see above) refuses to agree that the calculus of probability is applicable to *general statistics*.

It is impossible to be satisfied by such a decision. Indeed, *general statistics* (an expression that the author himself writes in inverted commas) is obviously a heterogeneous group of problems lacking any intrinsic ties. And, if only the theory of probability can at all be applicable to analysing reality, the necessary boundaries and conditions can depend not on the concrete properties of the totalities, but on their formal properties on the one hand and on the properties of the problems to be solved on the other. It is exactly in this direction that a manual of statistics ought to guide the beginners.

Kaufman (p. 147) expresses himself in the sense that for the areas mentioned (population statistics etc)

The existence of the prerequisites for the [application of] the theory of probability can be considered justified a posteriori, and the application of its methods here does not in principle excite objections anymore.

That the author is hardly in the right here can be already seen by the quotations from Markov and von Kries that he provides there. Indeed, even in the area where "the validity of applying the elements of the calculus of probability is *least* doubtful" (Kaufman, p. 145, his italics), the former denies the right of statisticians to justify *in principle* their practice (tables of mortality whose usefulness he does not deny) "by referring to the formulas of the calculus of probability".

As to the latter, since he is against the application of those formulas, his viewpoint concerns not one or another area of statistics (population or general statistics), i. e., not real objects but formally traced problems.

[8] Turning to the essence of the matter, inasmuch as it is possible in the boundaries of this paper, I am issuing from Markov's demand, that is, from the need to ascertain in each separate case whether the trials were independent, the probability was invariable and [the appropriate cases] equally possible. Under such restrictions, all the applications of the calculus of probability to statistics are partitioned in two main groups: in one of them, the applicability is justified a posteriori by proving that those

conditions are fulfilled at least approximately; and in the other one, applications are substantiated a priori.

In the first instance we are dealing with predicting probable frequencies of some phenomenon by known frequencies of other facts (urn experiments, heredity, insurance etc). In the second case we have to do with comparing reality with a theoretical pattern for which Markov's demands are postulated a priori. Simplest examples here again are experiments with urns, coins etc only considered in their different logical aspect.

Here, we compare the actual frequency with its value expected with one or another probability under the conditions of constancy, independence (or a definite dependence) and equal possibilities. And we do not act differently when studying the fluctuations of the sex ratio at birth or death etc. by the Lexian or any other similar method.

The same standpoint underlies the method that Kaufman discusses under a somewhat unfortunate name *differential*. Thus, when comparing for example the percentage of peasants without a horse of their own in two different localities so as to find out how significant is the difference, the real basis for the comparison is some imagined totality of individual farms, some imagined nation where the conditions determining the number of horseless peasants are assumed to be everywhere the same and the distribution of the farms over the territory is purely accidental¹⁰. And, issuing from that image, we calculate the probability that the difference mentioned could have been not less than in reality.

We are thus able to imagine at least the order of the probability of a correct judgement about whether the observed difference may be explained as being purely accidental, or whether we should assume as its basis either some detectable in principle causes or the insufficient accuracy of the data.

Kaufman correctly states that for such a conclusion it is not necessary to determine invariably the value of the appropriate probabilities, but he fails to notice that practitioners are infinitely many times guilty, also in our zemstvo statistics, of absolutely unfounded decisive inferences made from insufficient data. To oppose such arbitrary conclusions and to train systematically the feeling for the digits (so valued by the author), the calculation of probabilities or estimation of their order by determining mean square or probable errors and other measures of probable deviations ought to be practised incomparably oftener than it is done now.

Incidentally, it should be noted that Kaufman, when referring to von Kries for corroborating his views, hardly noted that that he (p. 244) discussed that very method of applying the calculus of probability to statistics calling it *Untersuchende Methode*. He allowed its application for studying mass social phenomena even in case of large numbers (*in erhebliche Umfange*). Kries very highly appreciates the investigations of Lexis and argues that they simply constitute a variety of that same method.

I believe that the contemporary statistical literature (above, I myself did not say anything essentially new)¹¹ sufficiently justified the application of the method under discussion to statistics and that, according to the train of thought leading to that substantiation, no partition whatsoever of statistics into areas, as Kaufman attempts to accomplish, can hardly be supported by any perceptible logical foundation¹².

[9] Not less shaky is the author's understanding of the application of the patterns and formulas of the theory of probability to sampling. As von Knies rightfully remarked, its embryos in the conjectural statistics¹³ of political arithmeticians developed by mathematicians (by Laplace!) failed to be sufficiently justified by an exhausting criticism of the empirically applied methods of isolating the sample. The work of Kiaer [at the turn of the 19th century] which in a sense marked a new stage suffered from the same shortcoming.

But Kaufman himself (p. 98) demands sampling with mechanical selection, that is, as I understand him, with a purely accidental choice, and he admits that such a procedure "provides a full guarantee of typicality, of *representativeness* of the results of sampling". This, however, is indeed what is needed for a justified application of the calculus of probability to sampling. Yule and Bowley deal only with this method [of sampling] and I am unable to understand how Kaufman (p. 97) could have concluded that, according to the Bowley method, it was indifferent whether to snatch at random 100,000 individual farms of a province or to select as the sample the entire population of its two uyezds [districts], or of an entire longitudinal strip.

This statement, may the author excuse me, is a misunderstanding pure and simple. And when he comes to deny the importance of the sample size (already gaining the upper hand also in our practice) as opposed to its relative size I cannot but perceive here the results of the same misunderstanding¹⁴.

I will dwell, for example, on his (Ibidem) reproach of sampling for extinguishing those qualitative nuances, those varieties of phenomena which exist in real life when considered in large masses, and are exhibited ever more distinctly with their increase. This is of course true, but we must not overlook that, on the other hand, the more considerable is the mass, the simpler and more curtailed usually ought to be the programme [of its investigation] so that in most cases, on the contrary, only sampling can allow us to approach reality from so different sides and therefore to perceive it more or less fully and distinctly in all its variety, see for example Westergaard (1890, pp. 205, 207 and some other places).

Then, for recognizing any *nuance* a corresponding *absolute* size of the sample is needed so that, having formulated beforehand definite cognitive theoretical or practical goals, we will be able to determine the corresponding sample size. No flair will help here since it did not guarantee even such an experienced investigator as Kaufman¹⁵ against an entirely mistaken recognition of the decisive importance of the relative size of the sample. Only a systematic application of tests provided by the calculus of probability, if, certainly, the researcher possesses all the other qualities peculiar for a good worker, can ensure the success of sampling. A critical discussion of the experience already at hand could have indicated all this with indisputable clarity.

I ought to add that this issue is by no means academic only. Exactly for the practitioner the problem of establishing the number of elements to be described, invariably connected with financial considerations, often determines whether the investigation will take place or not.

[10] My paper has already become too lengthy and I cannot consider in detail separate propositions made by the author concerning particular

issues. Although unable to agree with many of them, I must abandon his views on separate methods of mathematical statistics and will only point out some of his shortcomings. Thus, I am inclined to believe that the uninitiated will be put to difficulties by the three definitions of probability (by Laplace, Bortkiewicz and Vlasov [1909]) with which the author (p. 49) begins his exposition, and that such readers will understand them the less the more they will ponder over them. As it seems, this is especially true in regard to Laplace's definition provided out of the author's context.

Just the same (pp. 50 – 51), Czuber's definition is not understandable without a long explanation. I imagine that it would have been more advantageous to expound the principles of the theory of probability by examples with balls etc making use of the most elementary concept of probability as the ratio of the favourable cases to all of them and only to deepen this idea afterwards by indicating other possibilities. Here, however, the "logical foundation of the notion of probability" cannot be avoided since otherwise the reader will be confused by those various definitions rather than assimilate them. Moreover, the discussion of these issues (the viewpoints, say, of Venn, Cournot, von Kries, Chuprov) are much more important for understanding the beginnings of the theory of statistics than many other parts of the author's exposition, and, in addition, they are more readily understood.

Thus, I think that the derivation of the probability integral (pp. 76 – 78) could have been omitted since the reader will not be able to conclude it; it would have been better to explain instead the general train of thought leading to it and its significance and meaning. The author (p. 81) provided the appropriate approximate calculations, but it would have been better to choose an example allowing in addition to calculate the same probability in an elementary way by adding up probabilities of separate cases. For the beginner, this would mean much and it will also clearly indicate that the integral only provides approximations. Then, I think that the generalization of the law of large numbers (pp. 82 – 83) based on the [Bienaymé –] Chebyshev inequality can also be omitted, but, on the contrary, that it would have been apparently better to prove the Bayes theorem and to explain it in more detail. Indeed, it was the source of so many logical sins!

[11] And in general, it seems to me that for the goals attempted by Kaufman the volume of mathematics could have been lessened, but that the selected minimal information should have been worked out in rather more detail. Then, it would have been easier for the reader to learn how to calculate and to use the formulas of the calculus of probability at least at the minimal possible level as well as to apply the table of the probability integral which would have been useful to adduce at least in an abridged form as was the case with Chuprov (1909) or even in a more abridged way as Westergaard did.

Among the minor shortcomings [...]. These, however, are trifles which will hardly dumbfound a shrewd reader. More essential, as it seems to me, is the statement (pp. 121 – 129) that the naturalism of the coincidence of the empirical so-called check of the formulas of the calculus of probability with the theoretically predicted for games of chance is explained because "the law [of random deviations – E. S.] itself was, after all, derived from

the results of such experiments and games”. This, however, seems to be an accidental lapse.

Then, the author unjustly attributes to Bortkiewicz (to the law of small numbers) the ascertaining of the “theoretical distribution of the fluctuations” of small numbers without mentioning that the appropriate main formula is due to Poisson. Finally, I would argue against the use of the expression *method of moments* not in regard to the method competing with least squares for drawing a curve, but to the calculation of means by issuing from data grouped into intervals of equal length (p. 531). This will result in the use of an absolutely definite term in an extraneous manner which is hardly sufficiently justified.

Some probabilities are calculated wrongly [...], 0.995 instead of 0.95, 0.999979 instead of 0.997. The figure on p. 566 is scarcely vivid since only a quarter of the correlation diagram is shown, but the exposition of the calculation of the correlation coefficient itself seems to be sufficiently clear even for a beginner which of course was not easy to attain. I only think that the author with his knowledge of explaining could have included in his lengthy treatise rather more practical advice on, and patterns of calculation and not to refer readers so often to other sources either only helpful to a few because of linguistic difficulties (Yule) or insufficiently suited for the beginners (my own contribution of 1912) or, finally, to those entirely unsuited for his aims because of mistakes made (M. B. Gurevich).

In particular, I bear in mind the calculation of means, index numbers and more elementary methods of smoothing series which the author also admits for certain purposes. The inclusion of the formulas and tables due to Pareto and provided by Benini¹⁶ transforms the application of the method of least squares to calculating smoothing curves of the first four degrees into a childish occupation possible perhaps even for a school student of the third form and it would have compelled many practitioners to thank the author heartily. And this is desirable to see in the likely deservedly soon to appear next edition of his generally speaking excellent treatise.

Notes

1. I mention Lexis (1913). Kaufmann’s contribution “fills an important gap” and occupies “a special place” in the German statistical literature. He manages [makes do] with elementary mathematics which is a favourable circumstance. Then, Lexis believes that the theory of probability assumes equally possible cases and that the law of large numbers ought to be justified by empirical data. In short, I do not discern here a pioneer in the field of statistics.

Slutsky did not say anything about previous studies of the same subject but later he refers to several authors. However, it is opportune to add a few lines (Sheynin 1999). Slutsky mentions, and italicizes the term theoretical statistics but avoids mathematical statistics, a term that appeared at least in 1869 (Zeuner), and he did not define *statistics*. That statistics is a method (see his text a bit below) was stated in 1860 (Fox); and Pearson’s maxim (1892, p. 15) certainly comes to mind: “Unity of all science consists alone in its method, not in its material”. And it was Alphonse DeCandolle who first stated, in 1833, that statistics was a branch of mathematics. I note finally that later scholars, Pearson and Fisher, held that statistics was (“essentially” – Fisher) a branch of applied mathematics.

2. According to Chuprov (1909), who followed the German philosophers Windelband and Rickert, various sciences are either ideographic or nomographic (rather than nomothetic, as those philosophers called it). The former described reality (history), the latter studied regularity.

Late in life, in his reviews of several books, Chuprov again stated that statistics was an ideographic science although mostly having to do with quantitative data (which is not the case with history, allegedly an ideographic science). However, the literature concerning philosophy of probability does not anymore mention those philosophers and anyway even history is not a science without discussing regularities. For more detail see my Introduction to this collection where I refer to Chuprov's German paper of 1905 (also translated here).

3. Judgements and concepts rather belong to philosophy. The *it* in the next sentence is not altogether clear (cf. Druzinin's opinion [III, § 15] who quoted this passage) and is only one example of Slutsky's careless style. And in § 10 Slutsky mentioned definitions of probability offered by four authors whereas they really were either general considerations or comments.

4. Atoms do not *disappear*.

5. Yes, purely mathematical, but, at that time, not yet belonging to pure mathematics.

6. Chuprov (1909/1959, p. 133) set great store by free causal connections but I am not at all satisfied by his considerations. Their existence, as he reasoned, led to an unavoidable recognition of the need for probabilities, but he did not mention either correlation or randomness.

7. At the time, Mises had not yet formulated his frequentist theory of probability.

8. Venn (1866/1888, p. 88) expresses this idea very distinctly:

There is, it seems to me, a broad and important distinction between a material science which employs mathematics and a formal one which consists of nothing but mathematics.

And on p. 40:

During these [...] chapters we have been entirely occupied with laying what may be called the physical foundations of Probability.

See also pp. 41 and 265 – 266. I quote Venn because both von Kries and Chuprov, as it seems to me, were not altogether just in respect to him. His empiricism is not at all as coarse as can be judged by their opinion and in any case he is not guilty of simply identifying probabilities with empirical frequencies. E. S.

9. It seems to me that these considerations answer Kaufman's objection (p. 151) to my statement that statistics "is a mathematical science": "This is certainly not the case. Statistics is not mathematics". I agree with the latter words, but hope that he will also agree that neither is physics the same as mathematics. E. S.

10. Slutsky several times uses this not quite acceptable expression obviously having in mind a uniformly distributed random variable.

11. I believe that it is superfluous to corroborate this statement by quotations and references or name some names since any such attempt may be objected to by saying that all this is a *mathematical school*. In regard to at least this issue the essence consists not at all in opposing a *school*. When the debate is about a substance of something, it would be strange to group authors into schools according to their attitude towards propositions sufficiently clearly established by most authoritative scholars from, let us say, Laplace to leading contemporary figures of statistical thought.

I should hardly qualify this statement by adding that, when referring to experts, I do not wish to doubt that hesitations and debates are justifiable. E. S.

12. Also here I indicate that the description of the *differential method* (pp. 139 – 141) is hardly understandable to a beginner, and in essence hardly correct. The interpretation of the formula [without consulting Kaufman's treatise the following lines will not be clear. In essence, the matter is rather elementary]. My remark (1912) concerning the probable error of the difference of dependent variables Kaufman (pp. 140, 143, 146) interprets to his advantage, but wrongly, without allowing for my statement elsewhere (1912, p. 100). [...] E. S.

13. Bortkiewicz (1904, p. 825) used the same expression in the sense of *sampling*.

14. I take the opportunity to remark that it seems wrong to attribute to Bowley, as became usual apparently because of Chuprov [1912], the principle of composing the sample from purely accidentally snatched elements. The point is that this is the *only* method of sampling prompted by the calculus of probability and it was known long ago; in any case, Laplace had used it. As to the statistical aspect of the problem, it consists not in the principle of randomness as such, but in the technical tricks needed to achieve a

purely accidental selection of observations, and here the last word is far from being pronounced. In 1903, at the Berlin session of the International Statistical Institute, March, as the author of a resolution on Kiaer's report adopted by its demographic section, quite rightly, as I believe, objected to him by connecting the only correct version of sampling with Laplace's investigations rather than with Bowley. E. S. On the history of sampling see You Poh Seng (1951).

15. Kaufman had indeed published many concrete statistical investigations, but I doubt that they were ever seriously reviewed.

16. Chuprov (1925) later also referred to Benini (1906) and noted that he was unable to get hold of the relevant Pareto memoir (which he did not name either).

Bibliography

- Benini R.** (1906), *Principii di statistica metodologica*. Torino.
- Bortkiewicz L. von** (1904), *Anwendungen der Wahrscheinlichkeitsrechnung auf Statistik*. *Enc. math. Wiss.*, Bd. 1/2. Leipzig, pp. 822 – 851.
- Chuprov A. A.** (1909), *Ocherki po Teorii Statistiki* (Essays on the Theory of Statistics). Moscow, 1910, 1959.
- (1912, in Russian), *Sampling*. Translation incorporated in Sheynin (1997).
- (1925, in Russian), Review of Khotimsky V. (1925), *Vyravnivanie Statisticheskikh Riadov po Metodu Naimenshikh Kvadratov* (Adjustment of Statistical Series by the Method of Least Squares). Moscow – Leningrad. *Russk. Ekonomich. Sbornik* (Prague), No. 2, pp. 166 – 168.
- Cournot A. A.** (1843), *Exposition de la théorie des chances et des probabilités*. Paris, 1984.
- Husserl E.** (1900 – 1901, in German), *Logical Investigations* (1913). London, 1973.
- Kaufman (Kaufmann) A. A.** (1909), *Teoria Statistiki* (Theory of Statistics). Moscow.
- (1912), *Teoria i Metody Statistiki* (Theory and Methods of Statistics). Later editions: Moscow, 1916, 1922 (posthumously reprinted edition), 1928 (posthumous, with many material added by actual coauthors stressing the importance of mathematical methods).
- (1913), *Theorie und Methoden der Statistik*. Tübingen.
- Kries J. von** (1886), *Die Principien der Wahrscheinlichkeitsrechnung*. Tübingen, 1927.
- Lexis W.** (1913), Review of Kaufmann (1913), *Schmollers Jahrbuch f. Gesetzgebung, Verwaltung u. Volkswirtschaft in Deutsche Reiche*, Bd. 37, pp. 2089 – 2092.
- Sheynin O.** (1997), Chuprov's early paper on sampling. *Jahrbücher f. Nationalökonomie u. Statistik*, Bd. 216, pp. 658 – 671.
- (1999), Statistics, definitions of. *Enc. of Statistical Sciences*, 2nd edition, vol. 12, 2006, pp. 8128 – 8135.
- Slutsky E. E.** (1912), *Teoria Korreliatsii* (Theory of Correlation). Kiev.
- Venn J.** (1866), *Logic of Chance*. London, 1888. [New York, 1962.]
- Vlasov A. K.** (1909), *Teoria veroiatmostei* (Theory of Probability). Moscow.
- Westergaard H.** (1890), *Grundzüge der Theorie der Statistik*. [Jena, 1928, coauthor H. C. Nybolle.]
- You Poh Seng** (1951), Historical survey of the development of sampling theories and practice. *J. Roy. Stat. Soc.*, vol. A114, pp. 214 – 231. Reprint: Kendall M. G., Plackett R. L. (1977), *Studies in the History of Statistics and Probability*, vol. 2. London, pp. 440 – 457.

VI Oscar Sheynin

Karl Pearson 150 Years after His Birth

*Rossiiskaia i Evropeiskaia Ekonomicheskaia Mysl:
Opyt Sankt-Peterburga, 2005*

(Russian and European Economic Thought: the Experience of Petersburg).
Petersburg, 2007, pp. 97 – 119

1. Youth. Broad Interests

Karl Pearson (1857 – 1936) became an applied mathematician and philosopher, but in the first place he is remembered as the creator of biometry, the main branch of what was later mathematical statistics. In 1875 he obtained a scholarship at King's College Cambridge and took his bachelor's degree with honours in mathematics in 1879.

As a student, he refused to attend divinity lectures under compulsion, then continued voluntarily after regulations were softened. Already then he thus refused to comply with established by-laws. In 1877 K. P. took interest in finding his own way in religion and in studying philosophy, especially (in 1880 – 1883) Spinoza and German authors. True, in 1936 he (ESP vol. 28, p. 196) thought that Spinoza was “the sole philosopher who provides a conception of Deity in the least compatible with scientific knowledge”.

Until 1884 Pearson had also been studying literature, history and politics and came, perhaps independently, without being influenced by Mach, to comprehend science as description of phenomena.

In 1880 Pearson began to consider himself a socialist, entered into correspondence with Marx and even offered to translate (the first volume of) *Das Kapital* into English (Marx did not agree). He spent about a year in the universities of Heidelberg and Berlin (in the former, he read physics), studied the social and economic role of religion, especially in medieval Germany, and decided to deliver a course in German history and literature.

In 1882 – 1883, K. P. indeed gave lectures, in particular, in Cambridge, on the history of Germany during the Middle Ages and Reformation, and on the role of science and religion in society, and in 1884 he continued lecturing, this time in London, on Lassalle and Marx.

Pearson could have quite possibly become an outstanding historian, but his inherent mathematical ability was apparently stronger. Actually, he never forgot about mathematics: in 1881 – 1882, substituting for a staff professor, he taught mathematics in King's College and in 1881 and 1883 he unsuccessfully attempted to gain appointment to a professorship in mathematics. At about the same time K. P. (ESP, vol. 28, p. 200) had been “engaged on his first considerable piece of mathematical work” connected with physics.

In 1884 Pearson was finally appointed professor of applied mathematics at University College London. In the next year or two he gave a few lectures on the *Women's question* and established the *Men and Women's Club* that existed until 1889 for free and unlimited discussions of

everything concerning the relations between the sexes. Thus, Pearson thought that unmarried women should be allowed sexual freedom, and he certainly did not shirk from the eternal problem of combining job and family either. And (Haldane 1957, p. 305/1970, p. 429),

If today association with prostitutes is generally regarded as degrading, while seventy years ago it was generally condoned and not rarely approved, we owe it largely to men like Karl Pearson.

Here also we witness his refusal to accept without question the moral norms of his time.

2. Physics. Philosophy of Science

All through those early years and until about 1893, Pearson actively studied physics on which he expressed some extremely interesting ideas. Thus, “negative matter” exists in the universe (1891, p. 313); “all atoms in the universe of whatever kind appear to have begun pulsating at the same instant” (1887b, p. 114) and “physical variations effects” were perhaps “due to the geometrical construction of our space” (Clifford 1885/1886, p. 202). He did not, however, mention Riemannian spaces whereas it is nowadays thought that the curvature of space-time is caused by forces operating in it.

Remarkable also was Pearson’s idea (1892, p. 217), although subjectively expressed, about the connection of time and space:

Space and time are so similar in character, that if space be termed the breadth, time may be termed the length of the field of perception.

And (Ibidem, p. 103) here is another example of a similar perception of nature:

The law of gravitation is not so much the discovery by Newton of a rule guiding the motion of the planets as his invention of a method briefly describing the sequences of sense-impression which we term planetary motion.

This is correct insofar as that law does not explain the essence of gravitation. Mach (1897, Introduction) mentioned K. P. in the first edition of his book which appeared after 1892:

The publication [of the Grammar of Science] acquainted me with a researcher whose erkenntniskritischen [Kantian] ideas on every important issue coincide with my own notions and who knows how to oppose, candidly and courageously, extra-scientific tendencies in science.

Again in the same contribution we find Pearson’s celebrated maxim (1892, p. 15): “The unity of all science consists alone in its method, not in its material”.

In 1896 K. P. was elected Fellow of the Royal Society, which, in 1898 (ESP vol. 29, p. 194) awarded him, as proposed by Weldon, the Darwin medal. He declined it since the medal “must go to encourage young men” as he explained in a letter of 1912 on another such occasion, a refusal of

the Weldon Memorial Prize in Biometry (ESP vol. 29, p. 194). From 1912 to the end of his life he (Magnello 2001, p. 255) continued to refuse prizes, medals, a knighthood and, finally, the Guy Medal of the Royal Statistical Society.

Newcomb (Sheynin 2002, p. 163, Note 8), who presided at the International Congress of Arts and Sciences (St. Louis, 1904), invited Pearson to speak on *Methodology of Science*, undoubtedly because of the *Grammar of Science*. Pearson declined citing financial difficulties and fear of leaving his Department of applied mathematics under “less complete supervision”. He only became Head of the Department of applied mathematics in 1907. The Congress was successful; among speakers there were Boltzmann and Kapteyn.

In 1916, Neyman (ESP, vol. 28, p. 213) read the *Grammar of Science* on advice of his teacher at Kharkov University, S. N. Bernstein, and the book greatly impressed “us”. Wilks (1941, p. 250) called the same *Grammar*

One of the classics in the philosophy of science. In it, he attacked the dogmatism of the past and stressed the need of eliminating from science any jurisdiction which theology and metaphysics may claim.

Wilks was a most eminent American statistician, and his initials, S. S., were being interpreted as *Statistician Supreme*.

Neither did Lenin (1909/1961, pp. 190 and 174) fail to notice Pearson, calling him a “conscientious and honest enemy of materialism” and “one of the most consistent and lucid Machians”. Pearson’s lectures had also been lucid – and intelligent, and, in turn, he (1887a, pp. 347 – 348) expressed his opinion about revolutions in general and about Lenin in particular (1978, p. 243):

We invariably find that something like the old system springs again out of the chaos [of revolution], and the same old distinction of classes, the same old degradation of labour, is sure to reappear. [...] You may accept it as a primary law of history, that no great change ever occurs with a leap.

Petersburg has now for some inscrutable reason been given the name of the man who has practically ruined it.

In that latter source Pearson (p. 423) also said a few words about Kerensky, the Prime Minister of the Russian Provisional Government in 1917:

Men of liberal ideas, in particular liberal scientists have not the foresight and the strength which are needed to control a revolution. As Kerensky was to Lenin, so was Condorcet to Robespierre.

3. Statistics, Eugenics, Biology

When lecturing on statics, Pearson widely applied graphical methods and began to study the same methods in statistics, perceiving them as a general scientific tool answering his not at all received ideas about the

need to provide a broad mental outlook to students. Soon, however, discussions of the issues of evolution with Weldon as well as the writings of the much older Galton (1822 – 1911), turned his attention to biology and eugenics and to their study by statistical means (in the first place, by applying the nascent correlation theory). All this happened in spite of his being extremely busy: in 1891, without leaving the University College (in which he continued lecturing until 1911) he became professor of geometry in the celebrated Gresham College in London but had to abandon that new position in 1894 because of overwork.

Here are two of his statements on eugenics (Pearson 1887a, p. 375; MacKenzie 1981, p. 86):

Shall those who are deceased, shall those who are nighest to the brute, have the power to reproduce their like? Shall the reckless, the idle, be they poor or wealthy, those who follow mere instinct without reason, be the parents of the future generations? Shall the phtisical father not be socially branded when he hands down misery to his offspring and inefficient citizens to the state? It is difficult to conceive any greater crime against the race.

Do I [...] call for less human sympathy, for more limited charity, and for sterner treatment of the weak? Not for a moment.

The first pronouncement made in 1909 concerns *negative eugenics* which involves *subjective and controversial matter* (*New Enc. Brit.*, 15th edition, vol. 19, 2003, p. 725 of the item on *Eugenics and Heredity*). And in any case I denounce the abominable statement (Boiarsky & Tsyrlin 1947, p. 74) that Pearson's racist ideas "had forestalled the Goebbels department". This is where the influence of the (Soviet) environment in general and of the troglodytes of the Maria Smit stamp (see § 6) had indeed been felt. Finally, a preliminary ascertaining of hereditary illnesses by genetic means can also be attributed to eugenics.

In 1913 – 1914, and then intermittently in 1921 – 1929, Pearson and his collaborators delivered lectures for the general public on subjects of eugenics; he himself also published several related papers on the influence of tuberculosis, alcoholism and mentally illness on heredity. At times, his inferences were surprising and led to embittered debates. In 1925 Pearson established the periodical *Annals of Eugenics* and had been editing it for five years. In an editorial in its first issue, he indicated that the journal will be exclusively devoted to studying race problems and favourably regard the statement (Galton) that eugenics was based on probabilities. It is perhaps significant, however, that in 1954 the periodical changed its name into *Annals of Human Genetics*.

Weldon died in 1906, and Pearson had been compelled to solve biological problems alone. Still, it was Weldon and Galton who established the Biometrical school for statistically justifying natural selection, and Pearson became its head and the chief (for many years, the sole) editor of its celebrated periodical, *Biometrika*.

I insert a passage from the Editorial in its first issue of 1902 but after quoting Weldon (1893, p. 329) as reprinted by ESP (vol. 28, p. 218):

It cannot be too strongly urged that the problem of animal evolution is essentially a statistical problem: that before we can properly estimate the changes at present going on in a race or species we must know accurately (a) the percentage of animals which exhibit a given amount of abnormality with regard to a particular character [three more points are listed]. These are all questions of arithmetic; and when we know the numerical answers to these questions for a number of species we shall know the direction and the rate of change in these species at the present day – a knowledge which is the only legitimate basis for speculations as to their past history and future fate.

The problem of evolution is a problem in statistics. [...] We must turn to the mathematics of large numbers, to the theory of mass phenomena, to interpret safely our observations. [...] May we not ask how it came about that the founder of our modern theory [of ... hypothesis] made so little appeal to statistics? [...] The characteristic bent of C. Darwin's mind led him to establish the theory of descent without mathematical conceptions [...]. But [...] every idea of Darwin – variation, natural selection [...] – seems at once to fit itself to mathematical definition and to demand statistical analysis. [...] The biologist, the mathematician and the statistician have hitherto had widely differentiated fields of work. [...] The day will come [...] when we shall find mathematicians who are competent biologists, and biologists who are competent mathematicians.

Had Weldon not died prematurely, he would have been able to do much more; K. P. understood this well enough and in any case he published a paper (1906) honouring Weldon's memory. For that matter, he compiled a contribution (1914 – 1930) on Galton's life and achievements, a fundamental and most comprehensive tribute to any scholar ever published. It testified to its author's immense capacity for hard work.

The immediate cause for establishing *Biometrika* seems to have been scientific friction and personal disagreement between Pearson and Weldon on the one hand, and biologists especially Bateson, on the other hand, who exactly at that time had discovered the unnoticed Mendel. It was very difficult to correlate Mendelism and biometry: the former studied discrete magnitudes, the latter investigated continuous quantitative variations.

It is somewhat questionable to what extent had Pearson acknowledged Mendelism, but in any case he (1904, pp. 85 – 86) almost at once stated:

In the theory of the pure gamete there is nothing in essential opposition to the broad features of linear regression, skew distribution, the geometric law of ancestral correlation etc. of the biometric description of inheritance in populations. But it does show that the generalized theory here dealt with is not elastic enough to account for the numerical values of the constants of heredity hitherto observed.

To recall, gametes make possible the development of new individuals and the transmission of hereditary traits from the parents to the offspring.

And here is Pearson's statement of 1913 (ESP, vol. 29, pp. 169 – 170):

Mendelism is being applied wholly prematurely to anthropological and social problems in order to deduce rules as to disease and pathological states which have serious social bearing. [...] To extrapolate from theory beyond experience [which was apparently practised by some Mendelians] in nine cases out of ten leads to failure, even to disaster when it touches social problems.

And ESP himself (vol. 28, p. 242) stated:

A myth regarding some essential error in the biometricians' approach has persisted to this day. [...] But Pearson saw clearly, as most of his critics did not, that no theory of inheritance could discredit certain established facts following from a statistical analysis of observational data.

Continental statisticians had not then thought about biology. Much earlier Quetelet (1846, p. 259), who lived until 1874 but never mentioned Darwin, stated that "The plants and the animals have remained as they were when they left the hands of the Creator". Knapp (1872), an eminent German statistician, when discussing Darwinism did not mention randomness and said nothing about statistically studying biological problems. Later K. P. (1923, p. 23) owned that

We looked upon Charles Darwin as our deliverer, the man who had given a new meaning to our life and to the world we inhabited.

The speedy success of the Biometric school had been to a large extent prepared by the efforts of Edgeworth (1845 – 1926), a peculiar scholar whose works have recently appeared in three volumes (1996) and have been described in general terms by Chuprov (1909/1959, pp. 27 – 28), Schumpeter (1954/1955, p. 831) and Kendall (1968). I quote Chuprov:

The pioneers of the new statistical ideas to a large extent owe their rapid success to the fact that the soil for the propagation of their sermons had been prepared: an authoritative ally, [...] Edgeworth, met those pioneers in the bowels of the Royal Society. For two decades he had been popularizing there the mathematical methods of statistics. [...] However, those representatives of the statistical science, to whom he had applied, blocked his efforts by a lifeless wall of inertia. [...] His voice had found no response.

Edgeworth is too special in every way, [...] he is a lone figure [...], disciples he has none. [...] Nevertheless, his activities had not been futile.

Pearson's results in statistics include the development of the elements of correlation theory and contingency; introduction of the *Pearsonian curves* for describing empirical distributions; and a derivation of a most important chi-squared test for checking the correspondence of experimental data with one or another law of distribution, as well as the compilation of many important statistical tables.

K. P. devised his curves for practical application as the solution of some differential equation with four parameters but had not quite properly

justified them theoretically. In one particular case that solution led to the normal law, and, in all, 12 more curves thus appeared and at least some of them really proved themselves useful.

Pearson's posthumously published (by ESP) lectures (1978) examined the development of statistics in connection with religion and social conditions of life. On the very first page we find there the statement about the importance of the history of science: *I do feel how wrongful it was to work for so many years at statistics and neglect its history.* However, in spite of his historical studies, K. P. had not mentioned the Continental direction. True, he only discussed previous events, but it would have been easy and opportune to say a few pertinent words. Then, he (1925, p. 210) provided a patently false appraisal of the Bernoulli law of large numbers:

It is somewhat a perversion of historical facts to call [that law] by the name of the man who [...] had not gone further than the crude values [...] with their 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!

He had not noticed Bernoulli's solution of his own philosophical problem: the proof that in principle induction was not worse than deduction; K. P. evidently did not set high store on theorems of existence (in this case, of a certain limit), and he inadmissibly compared Bernoulli's result with a false system of the world.

Pearson (1926) reasonably held a high opinion about De Moivre and, as also later in his lectures, stressed the social roots and religious incentives of eminent statisticians and philosophers:

Newton's idea of an omnipresent deity, who maintains mean statistical values, formed the foundation of statistical development through Derham, Süßmilch, Niewentyt, Price to Quetelet and Florence Nightingale. [...]

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 [the memoir of 1733 in which De Moivre proved his limit theorem] or Bayes to his theorem were more theological and sociological than purely mathematical [...].

Maintaining mean statistical values means, as I understand it, regularly readjusting the system of the world which is being gradually corrupted by (random) mutual perturbations.

Pearson's laboratories deserve special notice. He had been head of the Biometric laboratory from 1895, of the eugenic laboratory (established in 1906 by Galton) from 1908. They were amalgamated in 1911 and in 1933 K. P. (ESP, vol. 29, p. 230) submitted his final report to the University of London in which he noted the "development in the last ten years" of Continental laboratories "on the lines" of that amalgamated entity, i. e., the work on "the combination of anthropometry, medicine, and heredity, with a statistical basis". Since Pearson mentioned medicine, it is opportune to add that physicians had recognized his merits by electing him, in 1919, Honorary Fellow of the Royal Society of Medicine (ESP, vol. 29, p. 206).

During World War I “the laboratories under your [his] charge” rendered “very valuable assistance [...] to the Ministry in general, and to this Department in particular” [...]. This is a quote from a letter of 14th February, 1918, to K. P. by Vice-Admiral R. H. Baker of the Munitions Inventions at the Ministry of Munitions (ESP, vol. 29, p. 244).

Biometry had been created by the efforts of Pearson and his school (and he himself had invented that term) and paved the way for the development of mathematical statistics.

I add a few words on Pearson’s later (much later than the appearance of his *Grammar of Science*) attitude towards science in general and religion. In 1922, in a rare source, he (ESP, vol. 29, p. 237) argued that

New phases of philosophy, new phases of religion will grow up to replace the old. But [?] the cultivated mind can never regard life and its environment in the same way as men did before those days of Darwin and before these days of Einstein. The ‘value’ of words, the ‘atmosphere’ of our conceptual notions of phenomena, has been for ever changed by the movement which began with Darwin and at present culminated in Einstein.

ESP (vol. 28, p. 194) left a few relevant, but, regrettably, too general words:

In the life of Karl Pearson we may trace all the signs of a struggle for freedom, of a period of uncertainty and trial, of the development of a new faith and of the blending of this faith with his outlook on science.

I can only add Pearson’s statement (1936, p. 33 note 2) made at the very end of his life which once more testifies to his independent outlook:

Stripped of its formalism and tribalism, the Jewish Unitarianism seems to me personally a higher form of religious faith than the Gentile Trinitarianism.

4. Other Branches of Science

Pearson attempted, often successfully, to apply the statistical method, and especially correlation theory, in many other branches of science; he had not, however, dwelt on the kinetic theory of gases, apparently because it demanded stochastic rather than statistical underpinning. Here is his interesting pronouncement (1907, p. 613):

I have learnt from experience with biologists, craniologists, meteorologists, and medical men (who now occasionally visit the biometricians by night!) that the first introduction of modern statistical method into an old science by the layman is met with characteristic scorn; but I have lived to see many of them tacitly adopting the very processes they began by condemning.

Later, in 1920, K. P. compiled (and possibly disseminated) a note (ESP, vol. 29, p. 164) explaining the aims of the Biometric school, cf. the Editorial in the first issue of *Biometrika* in my § 3:

To make statistics a branch of applied mathematics [...], to extend, discard or justify the meagre processes of the older school of political and social statisticians, and, in general, to convert statistics in this country from being the playing field of diletanti and controversialists into a serious branch of science. [...] Inadequate and even erroneous processes in medicine, in anthropology [anthropometry], in craniometry, in psychology, in criminology, in biology, in sociology, had to be criticized [...] with the aim of providing those sciences with a new and stronger technique.

The battle has lasted for nearly 20 years, but there are many signs now that the old hostility is over and the new methods are being everywhere accepted.

Pearson studied almost all of these sciences (disciplines). I have mentioned the correlation theory above, and ought to say that he attempted to apply it in astronomy (1907, pp. 517 – 518):

Astronomers have been guilty of a considerable amount of circular reasoning. They start from the hypothesis that [star] magnitude is very closely related to parallax, and when the statistician shows that the [...] parallaxes show no continuous relationship between parallax and magnitude, they turn around and say: Yes, but our stars were selected because they had big proper motions. They thereby screen entirely the fact that the fundamental hypothesis that the brighter stars are much the nearer as yet awaits statistical demonstration.

Nevertheless, by that time astronomers had been doubting the connection between magnitudes and distances (or parallaxes). Or, more precisely, doubting that the mean proper motion of stars of a given magnitude (which indirectly indicated their mean distance) had a certain meaning. Newcomb (Sheynin 2002, pp. 160 – 161), in a letter of the same year (1907), politely criticized Pearson (I have not found the latter's answer). He noted that

The known relations between magnitudes, distances and parallaxes must be taken as the basis of the investigation. [...] No general result [...] can be reached by pure induction.

Unlike statistics, the theory of errors has to do with constants, and Pearson (1920/1970, p. 187) apparently had considered it rather one-sidedly:

There is not a word in their innumerable treatises [on the method of least squares and adjustment of observations] that what is really being sought are the mutual correlations of a system of correlated variables. The mere using of the notation of the correlational calculus throws a flood of light into the mazes of the theory of errors of observation. [...] The Gaussian treatment leads (i) to a non-correlated surface for the directly observed variates, (ii) to a correlation surface for the indirectly observed variates. This occurrence of product terms arises from the geometrical

relations between the two classes of variates, and not from an organic relation between the indirectly observed variates appearing on our direct measurement of them. [...] There is no trace in Gauss' work of observed physical variables being – apart from equations of condition – associated organically which is the fundamental conception of correlation.

The notion of connection (dependence) in the theory of errors is indeed different. From the time of Gauss (possibly earlier than that, and without reference to him, but mostly intuitively) the dependence between two empirical magnitudes is being estimated there by the presence/absence of common initial observations. Kapteyn (1912) stressed this point, and, having been dissatisfied with (statistical) correlation, had introduced his own correlation coefficient. He had not mentioned Gauss either, and his paper had apparently been barely noticed. In any case Eisenhart (1978, p. 382), supplementing Pearson's opinion cited above, stated:

When Karl Pearson and G. Udny Yule began to develop the mathematical theory of correlation in the 1890s, they found that much of the mathematical machinery that Gauss devised [...] was immediately applicable in correlation analysis in spite of the fact that the aims of correlation analysis are the very antithesis of those of the theory of errors.

A curious result (Pearson 1902) related to many types of measurement can be cited: For two persons, the results of halving a segment with naked eye were not independent.

Not later than by mid-19th century meteorologists had established that the densities of the daily distribution of meteorological elements were often asymmetrical and for that reason Meyer (1891, p. 32) stated that the theory of errors was not applicable to that science. Pearson (1898), however, made use of Meyer's data for illustrating his theory of asymmetric curves.

Finally, Pearson (1928) studied Laplace's sample determination of the population of France. Laplace was the first to estimate the precision of such attempts, and Pearson's paper was apparently his only incursion into population statistics.

Let N and n be the known yearly births in France and in some of its regions, and m , the population of these latter. Laplace naturally assumed that the population of France was $M = (m/n)N$; important, however, was his estimation of the precision of that estimate. For his part, Pearson indicated that $(m; n)$ and $(M; N)$ were not independent samples from one and the same infinite totality (as Laplace tacitly thought) and that the very existence of such a totality remained doubtful.

5. Pearson As Seen by Others

Kolmogorov (1948/2002, p. 68) criticized the Biometric school:

Notions held by the English statistical school about the logical structure of the theory of probability which underlies all the methods of mathematical statistics remained on the level of the eighteenth century. [...] Rigorous results concerning the proximity of empirical sample characteristics to theoretical related only to the case of independent trials. [...] In spite of the great [...] work done [...], the auxiliary tables used in

statistical studies proved highly imperfect in respect to cases intermediate between small and large samples.

Anderson (1914, p. 269; English translation Sheynin 1990 (1996), p. 121), see my § 6, had indicated the first of the deficiencies mentioned above, and a few years later Chuprov himself (1918 – 1919, 1919, pp. 132 – 133; English translation: Sheynin, *Ibidem*), seconded the opinion of his student:

The disinclination of English researchers for the concepts of mathematical probability and mathematical expectation caused much trouble [...], obscured the stochastic statement of problems; on occasion, it even directed the attempts to solve them on a wrong track. If, however, this attire [this approach], so uninviting to the Continental eye, is shed and the discarded is picked up, it will be distinctly seen that Pearson and Lexis often offer different in form but basically kindred methods for solving essentially similar problems.

Fisher (1922, p. 311) expressed similar feelings:

Purely verbal confusion has hindered the distinct formation of statistical problems, for it is customary [for the Biometric school] to apply the same name, mean, standard deviation, correlation coefficient, etc., both to the true value which we should like to know, but can only estimate, and to the particular value at which we happen to arrive by our methods of estimation.

To recall, the development of the Continental direction of statistics had begun in the 1870s with the appearance of the work of Lexis on the stability of statistical series (on the constancy of the probability of the studied event and on the independence of the separate trials).

Kolmogorov could have added that neither had Laplace based his investigations on the notion of random variable. Even its heuristic introduction (hesitatingly made by Poisson) and the pertinent notation, ξ (say), which Poisson had *not* adduced, could have methodologically converted densities of distribution and characteristic functions into mathematical objects and thus prepared a transformation of the theory of probability.

Possibly because of his Machian outlook, Pearson had not done anything of the sort either, and criticism did not change anything. True, Chuprov (Sheynin 1990/1996, pp. 54 and 55) had privately informed him about mistakes in *Biometrika*, and Pearson (1919a) finally acknowledged them, but this apparently had no bearing on the problem under discussion.

Chuprov (*Ibidem*) also informed his correspondents that Continental statisticians (especially Markov) did not wish to recognize Pearson. Here is one of his letters (undated, written after Markov's death in 1922):

It seems that Pearson is unaware of the extent to which the mathematical forms of his researches hamper an appropriate appraisal of his contributions [...]. Because of Pearson's insufficiently rigorous, to their taste, approaches to mathematical problems, Continental

mathematicians look down on him to such an extent that they do not even bother to study his works. How many lances did I have occasion to brake because of Pearson while substantiating the considerable scientific importance of his oeuvre [...]!

Markov regarded Pearson, I may say, with contempt.

Markov himself, in a letter of 1910 (Ondar 1977/1981, p. 5) stated: “Neither [...] nor [...] nor Pearson has done anything worthy of note”. And, against this background, Slutsky’s foresight (his letter to Markov of 1912, see Sheynin 1990/1996, p. 45) is all the more interesting:

I believe that the shortcomings of Pearson’s exposition are temporary and of the same kind as the known shortcomings of mathematics in the 17th and 18th centuries. A rigorous basis for the work of the geniuses was built only post factum, and the same will happen with Pearson.

Slutsky naturally referred to his own book (1912) which Markov, unlike Kolmogorov (1948), had not understood properly. For Russian readers, exactly that contribution had remained perhaps for 15 years the only serious source of pertinent knowledge. During that time, Markov had been completing his study of dependent magnitudes (the *Markov chains*) and began to fall behind the development of mathematical statistics and even probability theory, cf. Sheynin (2006a, § 5.4), in particular due to his purely mathematical rather than more general scientific outlook (see above).

I am now adducing the opinions of other scientists about Pearson.

1) Fisher, letter of 1946 (Edwards 1994, p. 100):

He was singularly unreceptive to and often antagonistic to contemporary advances made by others in [his] field. [Otherwise] the work of Edgeworth and of Student, to name only two, would have borne fruit earlier.

Anyway, about 1914 Pearson (Sheynin 1990/1996, p. 124) stated, in a letter to Anderson, that Student “ist nicht ein Fachmann”, although Student, had by that time published five papers in *Biometrika*!

Fisher (1937, p. 306) also seriously accused Pearson: Pearson’s

Plea of comparability [between the methods of moments and maximum likelihood] is [...] only an excuse for falsifying the comparison [...].”

2) And now a testimony of a contrary nature: Mahalanobis, letter of 1936 (Ghosh 1994, p. 96):

I came in touch with [Pearson] only for a few months, but I have always looked upon him as my master, and myself, as one of his humble disciples.

3) And Newcomb, who had never been Pearson’s student, in a letter of 1903 to him (Sheynin 2002, p. 160):

You are the one living author whose production I nearly always read when I have time and can get at them, and with whom I hold imaginary interviews while I am reading.

I (Ibidem) have also collected similar pronouncements of that eminent scholar.

4) Hald (1998, p. 651) offered a reasonable general description of one aspect of the Biometric school:

Between 1892 and 1911 he [Pearson] created his own kingdom of mathematical statistics and biometry in which he reigned supremely, defending its ever expanding frontiers against attacks.

5) Fisher (1956/1990, p. 3) again:

The terrible weakness of his mathematical and scientific work flowed from his incapacity in self-criticism, and his unwillingness to admit the possibility that he had anything to learn from others, even in biology, of which he knew very little. His mathematics, though always vigorous, were usually clumsy, and often misleading. In controversy, to which he was much addicted, he constantly showed himself without a sense of justice. In his dispute with Bateson on the validity of Mendelian inheritance he was the bull to a skilful matador. [...] His activities have a real place in the history of a greater movement.

I left out much of Fisher's statement only because the source is readily available. Fisher began his description of the work of Pearson on p. 2, and there, in particular, is the phrase: "Pearson's energy was unbounded".

A genius is hardly able to appreciate properly lesser scholars, and it is scarcely possible to describe the mathematical quality of work (if not downright bad) done during several decades in a single phrase. And though in biology Pearson "new very little", he nevertheless essentially contributed to that science. Thus (ESP, vol. 28, p. 230):

The value of statistical method has been almost universally accepted among biologists, and tools which trace their origin to Pearson's workshop are applied along widely spreading lines of research investigation.

ESP, vol. 28, p. 230, describes Bateson's criticism of Pearson's long article of the same year (of 1901), and quotes his remark:

It is impossible to write of [it] without expressing a sense of the extraordinary effort which has gone to its production and of the ingenuity it displays.

Again, in a letter to Pearson of 1902 the same Bateson wrote (ESP, vol. 28, p. 204 note):

I respect you as an honest man, and perhaps the ablest and hardest worker I have met, and I am determined not to take up a quarrel with you

if I can help it. I have thought for a long time that you are probably the only Englishman I know at this moment whose first thought is to get at the truth in these problems [...].

To conclude with Fisher, I note that he could have well added that Pearson had paved the way for him.

6. Pearson in Russia and the Soviet Union

I begin with his contemporaries. His attitude towards Slutsky proved hardly satisfactory (Sheynin 1990/1996, pp. 46 – 47). In 1913, Pearson rejected both manuscripts submitted by Slutsky whereas the author called the objection raised to one of them “an obvious misunderstanding”. Acting on Chuprov’s advice, Slutsky sent it to the *Journal of the Royal Statistical Society* where it was indeed published (in 1914). ESP (vol. 29, p. 202), stated that Pearson’s own paper (1916) was devoted to a problem “for which the immediate suggestion was no doubt [that last-mentioned paper of Slutsky]”.

I (Sheynin 2004, pp. 222 – 240) published the three relevant letters of Slutsky to Pearson in their original English which Slutsky did not master sufficiently well; his German was perfect, but he obviously had not known that Pearson would have certainly preferred good German.

Anderson published two papers in *Biometrika* (1914; 1923), and one more, in 1926 – 1927, being by that time an emigrant. In the first one he (Sheynin 1990/1996, p. 121) stated that

The English statistical school neglects a method which is often used by Russian and German scientists [...] and which besides being quite rigorous and exact enjoys the advantage of being very elementary – namely, the method of mathematical expectations.

Pearson had objected to this statement in a private letter and then expressed his thoughts publicly (1919b, p. 285): “The remark of Dr Anderson [...] seems based on a misunderstanding of the moment method”. Anyway, Pearson had been indeed applying that method for determining parameters of empirical distributions.

Chuprov’s paper (1918 – 1921) obviously played a serious role in his election, in 1923, to honorary fellowship of the Royal Statistical Society. After his death, the Society passed a *Resolution of Condolence* published in a rare source but reprinted (Sheynin 1990/1996, p. 126). And Pearson had indeed honoured his memory by inserting his portrait in *Biometrika* (vol. 18, 1926, before p. 233); the only other similarly honoured Russian scholars were Chebyshev (Ibidem, vol. 22, 1930) and Markov (vol. 24, 1932). Pearson also intended to publish an obituary of Chuprov (Heymons 1926/2004) which, however, never appeared.

In 1923 – 1936 Romanovsky published six papers in *Biometrika*. His correspondence with Pearson (Sheynin 2008) testifies that Pearson rejected one more paper only because the author, acting in good faith, had published its abstract in the *C. r. Acad. Sci. Paris*. Pearson rejected yet another manuscript related to one of his own papers in spite of his wish because of financial difficulties, see *Biometrika*, vol. 17, 1925, p. 199.

Soviet statisticians denied the significance of Pearson’s work (see below), but Romanovsky continued keeping to an opposite opinion and

called him the head of modern mathematical statistics (*Bolshaia Sov. Encz.*, 1st edition, vol. 38, 1938, p. 409). A conference on mathematical statistics (*Resolutsia* 1948, p. 314/2005, p. 183) denounced “servility and kow-towing to outlandish ideas”, worryingly noted that “methods of bourgeois statistics were sometimes popularized and applied” and put on record that Romanovsky had acknowledged his earlier ideological mistakes.

Lenin’s criticism of Pearson’s philosophical outlook (§ 2) was in itself a sufficient cause of the extremely negative Soviet attitude towards Pearson. Maria Smit’s statement (1934, pp. 227 – 228) was its prime and vulgar example: his curves are based

On a fetishism of numbers, their classification is only mathematical. Although he does not want to subdue the real world as ferociously as it was attempted by [...] Gaus [her spelling], his system nevertheless only rests on a mathematical foundation and the real world cannot be studied on this basis at all.

In 1939, Smit was elected corresponding member of the Soviet Academy of Sciences... Recall also the abominable statement of Boiarsky & Tsyrlin in my § 3.

The second edition of the *Great Sov. Enc.* (vol. 33, 1955) declared that Pearson “advocated reactionary, pseudoscientific “theories” of race and blood” etc, etc, and that Lenin “destructively” criticized him. The tone of the same item, *Pearson*, in the third edition of the same source (vol. 19, 1975/1978, p. 366) was quite different: he “considerably contributed to the development of mathematical statistics” and Lenin criticized his “subjective-idealistic interpretation of the nature of scientific knowledge”.

7. Egon Sharpe Pearson (1895 – 1980)

For his biography see Bartholomew (2001), who however provided a wrong date of his birth. ESP was Karl Pearson’s son. His mother, Maria, neé Sharpe, died in 1928 (and K. P. married for the second time). In 1936, after his father’s death, ESP became Editor of *Biometrika* and continued in that capacity until 1966. He was a very successful statistician and especially remarkable was the Neyman – Pearson theory of testing hypotheses. He also studied the application of statistics to industrial standardization; in 1933 – 1936 he published several articles on this subject and a book (1935), edited statistical tables and had been seriously engaged in the history of statistics. It was ESP who edited the posthumous book *Pearson* (1978). In 1966 ESP was elected to the Royal Society.

Bibliography

Karl Pearson

- (1887a), *The Ethic of Freethought*. London. [London, 1901.]
- (1887b), On a certain atomic hypothesis. *Trans. Cambridge Phil. Soc.*, vol. 14, pp. 71 – 120.
- (1891), Atom squirts. *Amer. J. Math.*, vol. 13, pp. 309 – 362.
- (1892), *Grammar of Science*. London. Recent edition: New York. 2004.
- (1894), On the dissection of asymmetrical frequency curves. *Phil. Trans. Roy. Soc.*, vol. A185, pp. 71 – 110.
- (1896), Regression, heredity and panmixia. *Ibidem*, vol. A187, pp. 253 – 318.
- (1898), Cloudiness. *Proc. Roy. Soc.*, vol. 62, pp. 287 – 290.
- (1900), On the criterion etc. *Phil. Mag.*, Ser. 5, vol. 50, pp. 157 – 175.

- (1902), On the mathematical theory of errors of judgement etc. *Phil. Trans. Roy. Soc.*, vol. A198, pp. 235 – 299.
- (1904), On a generalized theory of alternative inheritance etc. *Phil. Trans. Roy. Soc.*, vol. A203, pp. 53 – 86.
- (1906), W. F. R. Weldon, 1860 – 1906. *Biometrika*, vol. 5, pp. 1 – 52.
- (1907), On correlation and the methods of modern statistics. *Nature*, vol. 76, pp. 517 – 518, 613 – 615, 662.
- (1914 – 1930), *Life, Letters and Labours of Fr. Galton*, vols 1, 2, 3A, 3B. Cambridge.
- (1916), On the application of “goodness of fit” tables etc. *Biometrika*, vol. 11, pp. 239 – 261.
- (1919a), Peccavimus. *Ibidem*, vol. 12, pp. 259 – 281.
- (1919b), On generalized Tchebysheff theorems in the mathematical theory of statistics. *Ibidem*, pp. 284 – 296.
- (1920), Notes on the history of correlation. *Ibidem*, vol. 13, pp. 25 – 45. Reprint: E. S. Pearson & Kendall (1970, pp. 185 – 205).
- (1923), *Charles Darwin*. London.
- (1925), James Bernoulli’s theorem. *Biometrika*, vol. 17, pp. 201 – 210.
- (1926), Abraham De Moivre. *Nature*, vol. 117, pp. 551 – 552.
- (1928), On a method of ascertaining limits to the actual number of marked individuals [...] from a sample. *Biometrika*, vol. 20A, pp. 149 – 174.
- (1936), On Jewish – Gentile relationships. *Biometrika*, vol. 28, pp. 32 – 33.
- (1948), *Early Statistical Papers*. Editor E. S. Pearson. Cambridge.
- (1978), *History of Statistics in the 17th and 18th Centuries against the Changing Background of Intellectual, Scientific and Religious Thought*. Lectures of 1921 – 1933. London. Editor E. S. Pearson.

Other Authors

- Anderson O.** (1914), Nochmals über “The elimination of spurious correlation due to position in time or space”. *Biometrika*, vol. 10, pp. 269 – 279.
- (1923), Über ein neues Verfahren bei Anwendung der “Variate-difference” Methode. *Ibidem*, vol. 15, pp. 134 – 149, 423.
- Bartholomew D. J.** (2001), E. S. Pearson. In Heyde, Seneta (2001, pp. 373 – 376).
- Bortkevich V. I., Chuprov A. A.** (2005), *Perepiska (Correspondence), 1895 – 1926*. Berlin. Also at www.sheynin.de
- Boiarsky A. Ya., Tsyrlin L.** (1947, in Russian), Bourgeois statistics as a means for apologizing capitalism. *Planovoe Khoziastvo*, vol. 6, pp. 62 – 75.
- Chetverikov N. S.**, Editor (1968), *O Teorii Dispersii (On the Theory of Dispersion)*. Moscow.
- Chuprov, Tschuprow A. A.** (1909), *Ocherki po Teorii Statistiki (Essays on the Theory of Statistics)*. Moscow. Third edition, 1959.
- (1918 – 1919), Zur Theorie der Stabilität statistischer Reihen. *Skand. Aktuarietidskr.*, Bd. 1, pp. 199 – 256; Bd. 2, pp. 80 – 133.
- (1918 – 1919, 1921), On the mathematical expectation of the moments of frequency distributions. *Biometrika*, vol. 12, pp. 140 – 149, 185 – 210; vol. 13, pp. 283 – 295.
- (2004), *Statistical Papers and Memorial Publications*. Berlin. Also at www.sheynin.de Compiled/translated by O. Sheynin.
- Clifford W. K.** (1885), *Common Sense of the Exact Sciences*. London, 1886, this being the first posthumous edition essentially extended by K. Pearson. Several later editions, for example, New York, 1946.
- Edgeworth F. Y.** (1996), *Writings in Probability, Statistics and Economics*, vols 1 – 3. Cheltenham.
- Edwards A. W. F. R.** (1994), R. A. Fisher on Karl Pearson. *Notes & Records Roy. Soc. London*, vol. 48, pp. 97 – 106.
- Eisenhart C.** (1974), Pearson. *Dict. Scient. Biogr.*, vol. 10, pp. 447 – 473.
- (1978), Gauss. In Kruskal W., Tanur J. M., Editors (1978), *Intern. Enc. Stat.* New York, vol. 1, pp. 378 – 386.
- Fisher R. A.** (1922), On the mathematical foundations of theoretical statistics. *Phil. Trans. Roy. Soc.*, vol. A222, pp. 309 – 368.
- (1937), Professor K. Pearson and the method of moments. *Annals of Eugenics*, vol. 7, pp. 303 – 318.

- (1956), Statistical methods and scientific inference. In author's *Statistical Methods, Experimental Design and Scientific Inference*. Oxford. 1990. Includes reprint of the edition of 1973 of that book with separate paging.
- Ghosh J. K.** (1994), Mahalanobis and the art and science of statistics: the early days. *Indian J. Hist. Science*, vol. 29, pp. 89 – 98.
- Hald A.** (1998), *History of Mathematical Statistics from 1750 to 1930*. New York.
- Haldane J. B. S.** (1957), Karl Pearson, 1857 – 1957. *Biometrika*, vol. 44, pp. 303 – 313. Reprint: E. S. Pearson & Kendall (1970, pp. 427 – 437).
- Heyde C. C., Seneta E.**, Editors (2001), *Statisticians of the Centuries*. New York.
- Heymons Helene** (manuscript 1926, in German), Letters to Karl Pearson. In Chuprov (2004, pp. 182 – 186). Russian translation: *Voprosy Statistiki*, No. 3, 2001, pp. 62 – 64.
- Kapteyn J. C.** (1912), Definition of the correlation coefficient. *Monthly Notices Roy. Astron. Soc.*, vol. 72, pp. 518 – 525.
- Kendall M. G.** (1968), F. Y. Edgeworth. *Biometrika*, vol. 55, pp. 269 – 275. Reprint: E. S. Pearson & Kendall (1970, pp. 253 – 254).
- Knapp G. F.** (1872), Darwin und die Sozialwissenschaften. *Jahrbücher f. Nationalökonomie und Statistik*, Bd. 18, pp. 233 – 247.
- Kolmogorov A. N.** (1948, in Russian), E. E. Slutsky. *Math. Scientist*, vol. 27, 2002, pp. 67 – 74.
- Lenin V. I.** (1909), *Materialism i Empiriokrititsizm* (Materialism and Empiriocriticism). *Polnoe Sobranie Sochinenii* (Complete Works), 5th edition, vol. 18. Moscow, 1961.
- Mach E.** (1897), *Die Mechanik in ihrer Entwicklung*. Leipzig. Third edition.
- MacKenzie D. A.** (1981), *Statistics in Britain, 1865 – 1930*. Edinburgh.
- Magnello Eileen** (2001), Karl Pearson. In Heyde, Seneta (2001, pp. 249 – 256).
- Merrington M. et al.** (1983), *List of the Papers and Correspondence of Karl Pearson*. London.
- Meyer Hugo** (1891), *Anleitung zur Bearbeitung meteorologischer Beobachtungen*. Berlin.
- Morant G. M. et al.** (1939), *Bibliography of the Statistical and Other Writings of Karl Pearson*. London.
- Ondar Kh. O.**, Editor (1977, in Russian), *Correspondence of Markov and Chuprov on the Theory of Probability and Statistics*. New York, 1981.
- Pearson E. S.** (1935), *Application of Statistical Methods to Industrial Standardization*. Brit. Standards Instn.
- (1936 – 1937), Karl Pearson: an appreciation of his life and work. *Biometrika*, vol. 28, pp. 193 – 257; vol. 29, pp. 161 – 248.
- Pearson E. S., Kendall M. G.** (1970), *Studies in the History of Statistics and Probability*. London.
- Porter T. M.** (2004), *Karl Pearson*. Princeton.
- Quetelet A.** (1846), *Lettres sur la théorie des probabilités*. Bruxelles.
- Resolutsia** (1948), Resolution, *Vtoroe Vsesoiuznoe Soveshchanie po Matematicheskoi Statistike* (Second All-Union Conference on Mathematical Statistics). Tashkent, pp. 313 – 318. Translation: Sheynin (2005, pp. 181 – 185).
- Schumpeter J.** (1954), *History of Economic Analysis*. New York, 1955.
- Sheynin O.** (1990, in Russian), *Chuprov. Life, Work, Correspondence*. Göttingen, 1996.
- (1999), E. E. Slutsky, 50 years after his death. *Istoriko-Matematich. Issledovania*, vol. 3(38), pp. 128 – 137. English transl.: in Sheynin (2004), *Russian Papers on the History of Probability and Statistics*. Berlin. Also at www.sheynin.de See pp. 222 – 240.
- (2002), Newcomb as a statistician. *Historia Scientiarum*, vol. 12, pp. 142 – 167.
- (2005), *Probability and Statistics. Russian Papers of the Soviet Period*. Berlin. Also at www.sheynin.de
- (2006a), Markov's work on the treatment of observations. *Historia Scientiarum*, vol. 16, pp. 80 – 95.
- (2006b), Review of Porter (2004). *Ibidem*, pp. 206 – 209.
- (2008), Romanovsky's correspondence with K. Pearson and R. A. Fisher. *Archives Internationales d'Histoire des Sciences*, vol. 58. No. 160 – 161, pp. 365 – 384.
- Slutsky E. E.** (1912), *Teoria Korreliatsii* (Theory of Correlation). Kiev.
- Smit Maria** (1934, in Russian), Against the idealistic and mechanistic theories in the theory of Soviet statistics. *Planovoe Khoziasstvo*, No. 7, pp. 217 – 231.

Weldon W. F. R. (1893), On certain correlated variations in *Carcinus Moenus*. *Proc. Roy. Soc.*, vol. A54, pp. 18 – 329.

Wilks S. S. (1941), Karl Pearson: founder of the science of statistics. *Scientific Monthly*, vol. 53, pp. 249 – 253.

VII
H. Kellerer, W. Mahr, Gerda Schneider, H. Strecker

Oskar Anderson, 1887 – 1960

O. Anderson, *Ausgewählte Schriften*, Bd. 1.
Tübingen, 1963, pp. XIII – XIX

[1] Oskar Johann Viktor Anderson was born 2 August 1887 in Minsk [capital of Byelorussia] as a son of Baltic German parents. He descended from a family of scientists; his father, Nicolai Carl Adolf Anderson, was Professor of Finno-Ugric languages at Kazan University; his brother Wilhelm later worked as Professor of theoretical physics at Dorpat [Tartu], and his brother Walther, as Professor of comparative folklore at universities in Dorpat, Königsberg and Kiel.

After attending a gymnasium in Kazan, Oskar A. for one term studied mathematics and physics at the celebrated mathematical faculty of Kazan University where the great mathematician Lobachevsky had been teaching in the first half of the 19th century. In 1907 Anderson entered the economic faculty of the Polytechnical Institute in St.-Petersburg which at that time enjoyed an outstanding reputation both in Russia and abroad.

Already during his student years, Anderson began his scientific activities: he became an assistant of his teacher, A. A. Chuprov the younger (1874 – 1926), and custodian of the library of the statistical and geographical room. After defending his diploma, the Faculty accepted his dissertation, *On the application of the coefficient of correlation to time series*, and resolved to publish it at its own expense. At the end of 1912, it bestowed upon him the degree of *Candidate of economic sciences* which roughly corresponded to the German *Doctor of Staatswissenschaften*. Only somewhat later Anderson passed a state examination at the law faculty of the Petersburg University which also granted him the right to become an articulated clerk.

In addition to his duties as assistant, Anderson, beginning in 1912, had for a long time been teacher of economics, economic geography and law at a commercial gymnasium in Petersburg; according to Russia's laws, he was therefore exempted during the war from military service.

[2] In summer 1915, a momentous event occurred in the life of the young scholar. As a participant in a state scientific expedition to the artificially irrigated region of the West Middle Asia, he directed both the preparation and implementation of a sample survey, the first one in the history of statistics¹, – a representative sampling of agricultural work in the basin of Syr-Daria, a region of almost a million hectares. The wide expanse of that land compelled Russian statisticians, long before American, English etc specialists, to involve themselves actively in consolidating the theory of sampling which in those times had been stuck in its infancy, and Anderson achieved pioneering results in that field.

From the end of 1915 Anderson had been appointed to the state fuel management control and remained there until the downfall of the Kerensky Provisional Government. He then moved to Southern Russia and started working at the section of economic research of a large local

cooperative office. At that time, under his direction and with his active participation, a series of monographs on the economic situation of that region had been compiled.

At the end of 1918 A. qualified as professor of methods of research in mathematical statistics at the Kiev Commercial Institute and, at the same time, was appointed assistant head of the Demographic Institute of the Kiev Academy of Sciences². In 1920, the Soviets captured all Southern Russia, and Anderson decided to leave Russia with wife and children. In 1912, he married Margarethe Natalie, née von Hindenburg-Hirtenberg, also a Baltic German.

[3] After a short stay in Constantinople, A. found his first refuge in Budapest. There, he resumed his scientific work, and, in particular, published his paper (1923) in *Biometrika*. There also already appeared (1914) an extract from his dissertation which, in particular, was the first description of the variate-difference method of analysing time series connected in the literature with the names of Oskar Anderson and W. S. Gosset (Student), who developed it independently from each other. In 1919, during the Civil War in Russia, the dissertation itself together with all the calculations involved and various luggage had been left at the quay in Novorossiisk and lost.

In autumn 1924 Anderson was appointed extraordinary professor at the Higher Commercial School in Varna (Bulgaria). There, he gave lectures in theoretical statistics, economic geography, finances, commercial policy, banking and monetary systems as well as the cooperative system. In 1929 he became ordinary professor of economics and statistics.

[4] Scientifically those years had been especially fruitful. A number of his important monographs and papers in periodicals had appeared, and, among them, further contributions to the difference method one of which we are mentioning (1926 – 1927). Above all, however, during the same period A. published a large number of indicative writings on the method of sampling quite decisive for his general recognition. We only mention the larger contribution (1929a) translated from Bulgarian into German in 1949 according to the desire of the Deutsche statistische Gesellschaft.

Anderson became generally known because of his basic investigations in which he separated himself from the entirely empirically oriented studies of conjuncture. He especially rejected the uncritical decomposition of time series as practised at the end of the 1920s by Warren M. Persons and his collaborators and known as the *Harvard Barometer*. He supported the idea of constructing an adequate model for the connection between the involved classes of causes *before* any such decomposition (1927; 1929a; 1929b). Not least because of these the Christian Albrechts University in Kiel bestowed upon him the title of Doctor of Staatswissenschaften.

He also was member and correspondent for Bulgaria of the International Conference of Agricultural Economists (O. A.).

And, finally, during the same period Anderson published one of the first ever econometric study, a pioneer work in the field of checking economic theory by statistical methods (1931). In 1930, A., as member of the International Conference of the Agrarian Commission at Cornell University in Ithaca, was invited to review the subject of *Theory of probability and economic research*.

[5] In the following years the first contacts with American scientists led to Anderson's participation in the standard American *Encyclopedia of Social Sciences* (1934). Together with Irving Fisher, Ragnar Frisch, Tinbergen, Schumpeter and others, he was an initiator and founder member of the Econometric Society established on 29 December 1930. This Society, and the papers in its periodical, *Econometrica*, decisively contributed to the development of econometrics, a bordering field between theoretical economics, mathematical statistics and mathematics.

At the end of 1933 Anderson received a grant from the Rockefeller Foundation which he made use of for visiting Germany and above all England where he had been able to work in many libraries. As a result, he, among other achievements, compiled the manuscript of his book (1935) in which he put down his most important findings and through which he became the successor of the late A. A. Chuprov and L. von Bortkiewicz as the most eminent representative of the Continental school of mathematical statistics.

In 1935 Anderson was appointed full-time director of the just established and partly financed by the Rockefeller Foundation Statistical Institute for Economic Research at the Sofia University. At the same time he became the scientific advisor of the Bulgarian General Direction of Statistics; from 1924 he had been member of the national Supreme Statistical Council. In these capacities he was the Editor of the *Publications of the Statistical Institute for Economic Research* (appearing in Bulgarian with translation into a foreign language, mostly into English) and co-editor of the *Revue trimestrielle de la Direction générale de la statistique* (appearing in Bulgarian with translation into a foreign language, mostly into French).

49 important monographs had appeared in the *Publications* until the outbreak of World War II as well as a number of contributions on various problems of the statistical method and the *Struktur* (1936). The great praise which Anderson had earned in Bulgarian official statistics and economic research were acknowledged a few years later in a special way when Czar Boris decorated him with a Commander Cross for Civil Merit.

[6] Anderson's activities reached far beyond Bulgaria. In 1936 he gave guest lectures at the London School of Economics; in 1935 – 1939 he had been member of the International Union of Conjuncture Institutes, and, from 1936 to 1939, associated member of the Committee of Statistical Experts of the League of Nations. Accordingly, he published his first critical comments on the theory of indices. Especially notable there were many statements about the problems of the so-called scissors of prices, of the measurement of real rates of exchange in foreign trade, compilation of "seasonal indices of the cost of living" and of constructing internationally comparable indices of industrial production.

In 1942 Anderson received an offer of professorship from the Christian Albrecht University in Kiel. Connected with that work was the leadership of a section of Ostforschung (study of Eastern countries) at the Institut für Weltwirtschaft. Already in Kiel he began turning to the problem which during the next decade became near to his heart, the elevation of statistical teaching at the economic faculties of German colleges to the level of international standard. He is meritorious for imparting significance to the application of the methods of mathematical statistics in the theory and

practice of social sciences in Germany thus ending a standstill of many decades. He devoted many reports and papers to this subject.

In 1947 Anderson was appointed to the new chair of statistics at the Ludwig Maximilian University in Munich. Until receiving the emeritus status in 1956, he had successfully taught the theory of statistical methods and its application to thousands of economists and business managers³. In addition, he surrounded himself by a closer circle of students a number of which nowadays represent modern statistics in colleges, statistical offices, economic research, in economy and administration. Professional statisticians have coined and are using the expression *Munich school*.

During Anderson's years in Munich, he published contributions to the field of probability theory, theory of index numbers, propagation of systematic errors, and, first of all, to distribution-free methods of testing. He devoted many papers to the development of a stochastic foundation in social sciences starting from his earlier notion of social-statistical probability as the "relative frequency in the totality of a higher order" (1947; 1949). He especially turned his attention to the often overlooked fact that the data on social statistics was usually corrupted by systematic errors. And, time and time again, he recommended to take into account the so-called propagation of errors so as to avoid untenable conclusions (1951; 1954b).

[7] During his last years, he once more devoted himself to problems of statistical investigation of causality in social sciences (1953). Nevertheless, most of all he aimed at developing distribution-free methods of testing. In social statistics, it is only rarely possible to adopt the hypothesis of normal parent distribution; just the same, the characteristics of samples taken from a totality of a higher order cannot always be assumed normal, so that distribution-free procedures are very important (1954a; 1955; 1956).

In an astounding and fruitful manner Anderson's statistics and economics unite to form a synthesis aspired to by contemporary research. He described his scientific statistical creed in a textbook (1954a, 1957, 1962) which we may perceive as a culmination of his work in Munich. This fundamental contribution ought to⁴

play the role of a pilot in mathematical statistics, and its goal is to provide the reader with the possibility of making his first independent steps in applying modern higher methods of investigation in mathematical statistics.

There, as in each of his earlier works, Anderson attributed great importance to clarity of the main notions and of an exact formulation of the assumptions stipulated by the applied methods. He especially and repeatedly indicated that

mathematical methods of statistics not only make it possible to conclude better and more precise inferences from the data, but in addition quite often provide brakes by indicating that the collected material is not yet sufficient for leading to more or less certain results.

That Anderson's scientific work has found recognition both in Germany and abroad is also expressed by the numerous honours bestowed upon him. He was

Honorary Doctor of Vienna University
Honorary Doctor of the Wirtschaftshochschule Mannheim
Honorary Fellow, Royal Statistical Society
Honorary member and member of the board, Deutsche statistische Gesellschaft
Fellow and founder member, Econometric Society
Member, International Statistical Institute
Fellow, American Statistical Association
Fellow, Institute of Mathematical Statistics
Fellow, American Association for the Advancement of Science
Member, Bulgarian Economic Society – O. A.

Oskar Anderson died on February 12, 1960, in the 73rd year of his life, after six weeks of serious illness.

Notes

1. Kiaer is known to have practised sampling in Norway from the turn of the 19th century, and in 1906 Kapteyn initiated the study of the starry heaven by stratified sampling.
2. More precisely, the Ukrainian Academy of Sciences.
3. Anderson's students obviously included future statisticians as well.
4. Both quotations below are apparently extracted from the book just mentioned.

Bibliography (O. Anderson)

Abbreviation: his *Ausgewählte Schriften*, Bde 1 – 2. Tübingen, 1963: AS

- (1914), Nochmals über "The elimination of spurious correlation due to position in time or space". *Biometrika*, vol. 10, pp. 269 – 279. AS, Bd. 1, pp. 1 – 11.
- (1923), Über ein neues Verfahren bei Anwendung der "Variate-Difference" Methode. *Biometrika*, vol. 15, pp. 134 – 149. AS, Bd. 1, pp. 12 – 27.
- (1926 – 1927), Über die Anwendung der Differenzenmethode [...] bei Reihenausgleichungen, Stabilitätsuntersuchungen und Korrelationsmessungen. *Biometrika*, vol. 18, pp. 293 – 320; vol. 19, pp. 53 – 86. AS, Bd. 1, pp. 39 – 100.
- (1927), On the logic of the decomposition of statistical series into separate components. *J. Roy. Stat. Soc.*, vol. 90, pp. 548 – 569. AS, Bd. 1, pp. 101 – 122.
- (1929a, in Bulgarian), *Über die repräsentative Methode und deren Anwendung auf die Aufarbeitung der Ergebnisse der bulgarischen landwirtschaftlichen Betriebszählung vom 31. Dezember 1926*. Sofia, 1929. German translation: München, 1949 and in AS, Bd. 1, pp. 302 – 376.
- (1929b), Zur Problematik der empirisch-statistischen Konjunkturforschung. Kritische Betrachtung der Harvard-Methoden. *Veröff. Frankfurter Ges. Konjunkturforschung*, No. 1. AS, Bd. 1, pp. 123 – 165.
- (1929c), Die Korrelationsrechnung in der Konjunkturforschung. Ein Beitrag zur Analyse von Zeitreihen. *Ibidem*, No. 4. AS, Bd. 1, pp. 166 – 301.
- (1931), Ist die Quantitätstheorie statistisch nachweisbar? *Z. f. Nationalökonomie*, Bd. 2, pp. 523 – 578. AS, Bd. 1, pp. 415 – 470.
- (1934), Statistics: Statistical method. *Enc. Social Sciences*, vol. 14, pp. 366 – 371. AS, Bd. 2, pp. 539 – 544.
- (1935), *Einführung in die mathematische Statistik*. Wien.
- (1936), *Die sozialökonomische Struktur der bulgarischen Landwirtschaft*. Berlin.
- (1947), Zum Problem der Wahrscheinlichkeit a posteriori in der Statistik. *Schweiz. Z. f. Volkswirtschaft u. Statistik*, Jg. 83, pp. 489 – 518 + Ergänzung. AS, Bd. 2, pp. 683 – 714.

- (1949), Die Begründung des Gesetzes der großen Zahlen und die Umkehrung des Theorems von Bernoulli. *Dialectica*, vol. 3, pp. 65 – 77. AS, Bd. 2, pp. 727 – 739.
- (1951), Über den Genauigkeitsgrad wirtschaftsstatistischer Daten. *Weltwirtschaftliches Archiv*, Bd. 67, pp. 8* – 15*. AS, Bd. 2, pp. 836 – 843.
- (1953), Moderne Methoden der statistischen Kausalforschung in den Sozialwissenschaften. *Allg. statistisches Archiv*, Bd. 37, pp. 289 – 300. AS, Bd. 2, pp. 878 – 889.
- (1954a), Ein exakter nicht-parametrischer Test der sogenannten Null-Hypothese im Falle von Autokorrelation und Korrelation. *Bull. Inst. Intern. Stat.*, t. 34, No. 2, pp. 130 – 143. AS, Bd. 2, pp. 864 – 877.
- (1954b), Über den Umgang mit systematischen statistischen Fehlern. *Statistische Vierteljahresschrift*, Bd. 7, pp. 38 – 44. AS, Bd. 2, pp. 890 – 896.
- (1954, 1957, 1962), *Probleme der statistischen Methodenlehre*. Würzburg.
- (1955), Eine “nicht-parametrische” (verteilungsfrei) Ableitung der Streuung (variance) des mutiplen [...] und partiellen [...] Korrelationskoeffizienten [...]. *Mitteilungsbl. f. math. Statistik und ihre Anwendungsgebiete*, Jg. 7, pp. 85 – 112. AS, Bd. 2, pp. 897 – 924.
- (1956), Verteilungsfrei (nicht parametrische) Testverfahren in den Sozialwissenschaften. *Allg. statistisches Archiv*, Bd. 40, pp. 117 – 127. AS, Bd. 2, pp. 927 – 937.

VIII S. Sagoroff

Oskar Anderson. Obituary

Metrika, Bd. 3, 1960, pp. 89 – 94

[1] On 12 February 1960 Oskar Anderson died in Munich. With him *Metrika* lost one of its founders and editors. Silent mourning fell upon our scientific community and obliged us, loving friends, to commemorate our great teacher and honoured colleague. Death released him from his long physical suffering and allowed us to appreciate his work without any suspicion of wishing to flatter or please him.

Anderson was born of German parentage in Minsk, Byelorussia, in 1887 and grown up in the vast Russian world in the traditions of thorough German scientific circles. His father was a university professor of Finno-Ugric languages, his brother, also a university professor of philology. He spent his youth in Kazan where he finished a gymnasium top of the class in 1906, then studied physics and mathematics at the university there and later, in 1912, graduated from the Economic Faculty of the Petersburg Polytechnical Institute as a student of Chuprov, a famous representative of the Continental school of mathematical statistics, and defended his dissertation in economics on the application of the correlation theory to time series. In 1914 he passed a state examination in jurisprudence at Petersburg University.

[2] Being 25 years old, Anderson began his scientific work and professorial life as a teacher. In 1912 he became assistant at the Statistical Institute [?] of the Polytechnical Institute and at the same time teacher of economics, economic geography and law at a Commercial school in Petersburg.

The world war drew the young scientist into economic life and formulated for him his first research tasks. At first, in 1915, he was appointed a leading post in the state board of fuel, then, in 1917, in a head office of a cooperative in southern Russia.

In 1915 Anderson experienced something that especially enriched his statistical practice and revealed him the pleasure of research: participation in a statistical scientific expedition to Turkestan in West Middle Asia. He directed one of the first representative surveys in the history of statistics, the statistical observation of agricultural works in the artificially irrigated oases in the upper and middle reaches of Syr-Daria.

In the autumn of 1918 Anderson qualified in statistics at the Kiev Commercial Institute and began teaching there mathematical statistics as Privat-Dozent. In 1920 radical political changes in Russia compelled him, however, to emigrate together with his family. Constantinople and Budapest were stages of a difficult section of his life on which, already during his escape from Russia, he lost his only daughter.

His life experienced a happy turn when in 1924 he received an invitation to the Commercial High School in Varna, Bulgaria, as extraordinary professor of statistics and economics. Five years later he became full professor, and, after five years more, he was called to Sofia

University. There, from 1935 to 1942 he directed its State Institute for Economic Research.

[3] In 1942 Kiel University invited him as ordinary professor of statistics. There, at the same time he directed the sector of Ostforschung (study of East European countries) at its Weltwirtschaft institute. The devastation of the city by air raids and the death of one of his sons in a battle in Tunisia shook him morally and in general weakened him corporally so that in 1947 he left Kiel with a feeling of deliverance to follow an invitation from Munich University.

[4] The eighteen years that Anderson spent in Bulgaria constituted the richest period of his theoretical work. His creative spirit followed two directions: he sometimes built, and sometimes proceeded critically. His numerous contributions about the notions and methods of mathematical statistics were of the first type. They began with the creation of the so-called differential method, see especially his papers (1926 – 1927), and reached its peak in his main work (1935). That method independently worked out by Anderson and Student (Gosset) became rightfully recognized in science, although it did not yield to a generalization into a universally applicable analytical method as for example correlation analysis.

His book (1935) represented an extremely rare example of statistical literature as being both original and systematically composed in a unified manner. Formally and roughly speaking, it may be compared with Laplace's *Théorie analytique des probabilités* or Mises' *Wahrscheinlichkeitsrechnung*. In spite of its merit, this contribution did not ensure him a proper place in the development of statistical theory and in my opinion this may be put down to two circumstances.

First, the unfavourable time of its appearance in the dominant twilight atmosphere on the eve of and during the world war in Europe in which science found itself. Second, its main concept, the attempt to develop the general statistical theory on a special case, on the notion of *social statistical probability*.

May we reproach Anderson therefore? Criticize him? I would decisively say no, with a capital *n*. Statistics is a bridge between the world of the feelings of material sciences and the imaginary realm of mathematics. Anderson wished to be nearer to the finite and discrete social and economic reality than to mathematics with its infinity and continuum. His incorruptible love of truth as he saw it and his yearning for independent judgement directed him to the path which, although not wrong, was not as broad and open as the route followed by modern mathematical statistics.

[5] The contributions written by Anderson, all of them without exception belonging to statistics, had been sufficient to make him internationally famous. And his critical works were of fundamental importance for science in general, and especially for econometrics. He was happy to work at the time when studies of conjuncture had emerged and, in addition, to belong to those few scholars who possessed all the important qualities for being an econometrician: the mastery of the morphology of economics, economics itself and mathematics as well as of pure and applied statistics.

The study of conjuncture which originated in the 1920s in the United States and came into blossom in the 1930s in Europe¹, attempted to insert mathematics into it. It was thought at first that that goal could be reached in a purely empirical inductive manner, but the downfall of the Harvard conjuncture barometer led to sobering. It was Anderson who discovered by applying mathematics the great deficiency of that approach.

In a number of publications following one another (1927; 1929a; 1929b) he showed the arbitrariness of the then prevailing methods of partitioning series. The ground from under the *mechanistic* conception of the essence of conjuncture was thus cut clearing the path for radically comprehending the perceptive reality; namely, for recognizing that without theoretical hypotheses a scientific understanding of that reality was impossible.

From the new standpoint there emerged the modern theory of economic process, the building of econometric models. Today, it is self-evident to distinguish between theoretical and empirical values of the magnitudes inserted into a model. This difference, which, incidentally, takes place for all material sciences, appeared in the 1940s as a discovery connected with the names of Haavelmo and Koopmans although Anderson had worked at the issue of empirically formulating the theoretical relations some ten years earlier (1931).

[6] During his years in Bulgaria, Anderson's role as a theoretician had been founded whereas his *German years* had been decisive for his part in the development of statistics in that country. After the death of Bortkiewicz mathematical statistics taught as a doctrine in German universities became for a decade extinct and it was Anderson who then carried the banner of the Continental school of mathematical statistics. His appointment after the war to the largest German university offered him many possibilities to break lances for the dissemination of mathematical statistics by reports and publications. Out of Munich he influenced the teaching and the theoretical thought in German universities as well as the German official statistics. If that discipline in Germany is today most important as a doctrine, it is to a large extent his merit.

When determining Anderson's place in statistical science, it is possible to say that he was the last representative of the Continental school of mathematical statistics brought forth by Lexis, Bortkiewicz and Chuprov. Considering it together with the old English school founded by Pearson, Bowley and Yule² as the classical school of mathematical statistics, Anderson must be called one of its last classics. He was really prepared for the advances in statistics. He especially admired the contributions of Fisher and incorporated them into his lectures in their entirety [?], but he did not stay as near to the direction taken by Neyman and E. S. Pearson.

The only restriction which Anderson introduced into modern mathematical statistics consisted in that the statistical methods proper for natural sciences cannot be immediately transferred to social sciences.

In each nation in which Anderson worked, he cultivated tight ties with official statistics. This was in accordance with his opinion about the nature of statistics as partly being pure theory, a formal science, and partly, in the theory of producing data (*Betriebslehre*), a doctrine of statistical production, a material science.

[7] That the stochastic representative method became very early introduced into the Russian official statistics was especially favourable for him. Indeed, it is too little known that a sample census of population was first made in Russia, in 1916 – 1917. (Its materials had not been processed and lost in the turmoils of the revolution.)

Anderson brought the Russian tradition of sampling to Bulgaria. As a member of the Supreme Statistical Council (1926 – 1942) and scientific advisor of the Bulgarian General Direction of Statistics, he had been able to introduce sampling into official statistics. Under his leadership the 1926 general census of population and industry was processed both in its entirety and following the representative method so as to test the reliability of mathematical methods. The results of the comparison were amazingly good. Later, in 1931/1932, a sampling investigation of agricultural industry and production was carried out. After moving to Munich Anderson collaborated with the Bavarian statistical board; that the mathematical methods of statistical observation and processing had been introduced there also is to a large extent due to him and his successor, Professor Kellerer.

Anderson's successful work over many years both in theory and practice [of statistics] brought him many honours. [The author lists his fellowships, see [VII, § 7], and in addition reports the following:] Shortly before he died, the Law and Statecraft (Staatswissenschaftliche) Faculty of Vienna University bestowed upon him the degree of Honorary Doctor of Staatswissenschaft. He had not received the diploma, but had been able to dictate a thank-you letter and express his feelings which inspired him all his life³.

Notes

1. The author apparently did not know anything about Kondratiev; see Schumpeter (1954/1955, p. 1158) who favourably mentioned him, provided references to available commentaries but did not describe the essence of Kondratiev's work.

2. The Biometric school was not old, at least not if compared with the Continental direction of statistics. And Bowley and Yule were not its cofounders.

3. Sagoroff did not mention that Anderson had greatly influenced statistical thought in Bulgaria (not only directing it to sampling investigations). However, his note is important as providing a somewhat new dimensions as compared with similar publications.

Bibliography

O. Anderson

(1926 – 1927), Über die Anwendung der Differenzmethode [...] bei Reihenausgleichungen, Stabilitätsuntersuchungen und Korrelationsmessungen. *Biometrika*, vol. 18, pp. 293 – 320; vol. 19, pp. 53 – 86.

(1927), On the logic of the decomposition of statistical series into separate components. *J. Roy. Stat. Soc.*, vol. 90, pp. 548 – 569.

(1929a), Zur Problematik der empirisch-statistischen Konjunkturforschung. Kritische Betrachtungen der Harvard-Methoden. *Veröff. Frankfurter Ges. Konjunkturforschung*, No. 1.

(1929b), Die Korrelationsrechnung in der Konjunkturforschung. Ein Beitrag zur Analyse von Zeitreihen. *Ibidem*, No. 4.

(1931), Ist die Quantitätstheorie statistisch nachweisbar? *Z. f. Nationalökonomie*, Bd. 2, pp. 523 – 578.

(1935), *Einführung in die mathematische Statistik*. Wien.

Other Authors

Schumpeter J. A. (1954), *History of Economic Analysis*. New York, 1955.

IX V. I. Riabikin

Oskar Anderson, Chuprov's Student

Uchenye Zapiski po Statistike, vol. 26, 1976, pp. 161 – 174

[1] The outstanding role of Chuprov's statistical school in the development of statistics the world over is determined above all by a rational combination of precise mathematical propositions and empirical investigations. The typical feature of Russian statistics is that theory and practice are mutually supplementing and enriching each other rather than contrasting one another¹.

The German scholar Oskar Anderson (1887 – 1960) was Chuprov's eminent student and follower. In 1907 – 1915, both had been working at the Petersburg Polytechnical Institute where Anderson was Chuprov's assistant and custodian of the library of the geographical & statistical room². Anderson was a professionally trained mathematician (he studied mathematics and physics at Kazan University), statistician and economist and he had been teaching the appropriate disciplines in various universities in Russia and Europe³. He also carried out practical statistical investigations. Thus, in 1915 he directly participated in preparing and implementing a scientific expedition for an agricultural census in the basin of Syr-Daria. For many years its scale, and, what is the main point, its representativeness left behind similar work in the USA and Europe.

Along with I. Fisher, Frisch, Tinbergen, Schumpeter and others, he was the initiator and founder member of the celebrated Econometric Society whose member he had remained until the end of his life. And the problems with which he busied himself are topical to this very day.

Anderson's dissertation, *On the application of the coefficients of correlation to dynamical series*, stemmed from Chuprov's correlation theory as applied to time series⁴. Anderson had been developing the methods of correlation analysis in conjuncture investigations, see e. g. his writing (1929a). For example, he stated that, when analysing dynamical series and calculating the appropriate coefficients of correlation, the problem had usually been reduced to eliminating the trend either by preliminary adjusting the series and calculating the deviations, or by introducing a variable (time). These methods can lead to differing results and their reasonableness entirely depends on the successful choice of the initial function for the adjustment or description of the mutual connections between dynamical series. Anderson stressed that at the same time any function selected as a basis of a dynamical series was only an ersatz of reality so that the most important point here was a correctly formulated and economically justified aim of the study corresponding to the initial conjuncture indications⁵. [...]

Anderson pointed out a special structure of time series, their total lack of correspondence with non-variational series which demanded special methods of their analysis. To achieve that, he developed an original way, the variate-difference method, making use of the classical theory of interpolation⁶. He proved that for two series the function [chosen for

preliminarily adjusting them] includes, as a rule, three types of correlation coefficients two of which are heterogeneous. One of these types is most of all important for describing a time series and his method allows to calculate it.

Assumptions and theoretical mistakes made when choosing a function for dynamical series lead to mistakes of principle like those made by Persons who had applied methods of mathematical statistics for conjunctural predictions. Anderson critically considered the *Harvard Barometer*, and, applying the variate-difference method, proved that it was not suitable for the aims which the Harvard school had attempted to achieve before the world crisis of 1929. His work and conclusions made testify to the theoretical correctness of his methods and to his intuition in economics, see Anderson (1929b). And we ought to agree with him in that, when describing a statistical idea in economics, formulas can only serve as supplementary means.

[2] Another direction of his work was a theoretical and practical justification of the representative method in economics and sociological investigations, see in particular his paper (1947) on posterior and prior probabilities. The problem actually consisted in combining mathematical statistics with the practice of statistical studies. It was twofold:

1) Was it necessary to study sample frequency and mean issuing from statistical probability and expectation derived from the parent distribution?

2) Was it possible to conclude, on the contrary, from sample to parent population?

Anderson showed that these points were not contradictory; and that, quite the opposite, new theoretical indicators were derived when issuing from the data. And the conditions arising in practical statistical investigations (for example, sampling without replacement), compel to develop the theory. Indeed, along with the work of the German statistician Keller [H. Kellerer], Anderson published a number of papers on the theory and practice of sampling⁷. And he was one of the first to indicate that systematic errors were seriously dangerous for statistical calculations. Nowadays this issue is sufficiently studied by Morgenstern (1950). [...]

Anderson (1951) also believed that the influence of systematic errors ought to be allowed for when calculating indices after a representative selection of the appropriate data. For our [Soviet] statistics this issue is not especially typical⁸, but it is impossible to deny its elements, and therefore actuality. It is not out of the question that systematic errors corrupt the indices of the cost of living, international comparisons and chain indices⁹. [...]

[3] Anderson very critically appraised the Anglo-American statistical schools not only, as indicated above, in connection with the *Harvard Barometer* and its underlying statistical concepts. He invariably indicated the special structure of the totalities occurring in economics and sociology, and it is for this reason that during his last years he (1956) turned his attention to non-parametric statistical methods.

In that report, as in all of his publications in general, in which the influence of Chuprov's school is clearly perceived, he criticized the contemporary state of mathematical statistics and econometrics. Here are his main propositions made there which show him as the follower of that school. He mentioned his direct participation in the establishment of the

Econometric Society and its significance for the application of mathematical methods in economics. He then indicated that many representatives of that Society had been leading econometrics along a route quite different from the one envisaged by its founder members, along the path of mathematical abstractions only nominally concerning economics.

If suchlike investigations, even, as Anderson expressed it, were supported by a parade of figures, they have little in common with practice and turn out to be either useless or harmful. As an example, he cited the works of Modigliani [to a work of Modigliani & Sauerlender] in which the correlation coefficients were applied for analysing stock exchange indicators and the t -distribution was made use of for describing totalities in which there were no, and could not have been any normal distributions or stochastic estimates [in general].

Anderson listed and explained a number of points necessary for those mathematicians who attempt to apply their methods in economics and sociology. They include the finiteness and, not rarely, the small size of the general totalities encountered by the economist and sociologist. Then, wars and crises corrupt cycles and tendencies in the economic development of nations so that a more or less reasonable mathematical statistical analysis is only possible for 15 – 20 years; general totalities are asymmetric, multimodal and have a number of gaps in frequencies.

Even such a gala example provided by many statisticians as demographic indications testifies to corruptions in the structure of the general totality, of the population, when considered dynamically. Finally, along with random errors which statisticians easily cope with, there are systematic errors. [...]

[4] Anderson therefore concludes that non-parametric methods were needed in economic studies, He compared the σ rule which follows from the normal distribution with the [Bienaymé –] Chebyshev inequality and its strengthened form according to the formulas of Guldberg – Pearson and Cramer – Chebyshev¹⁰. Assuming the normal distribution, the σ rule provides confidence intervals of $[-\alpha\sigma; \alpha\sigma]$ with $\alpha = 1.96, 2.58$ or 3 . The [Bienaymé –] Chebyshev inequality for the deviation of a random variable from its expectation is

$$P(|\xi - E\xi| \leq t\sigma) \geq 1 - (1/t^2)$$

whereas, according to the formulas mentioned,

$$P(|\xi - E\xi| \leq t\sigma) \geq 1 - (\beta_2/t^4), \quad (1)$$

$$P(|\xi - E\xi| \leq t\sigma) \geq 1 - \frac{\mu_4 - \sigma^4}{\mu_4 - 2t^2\sigma^2 + t^4\sigma^4}. \quad (2)$$

Here, β_2 is the excess and μ_4 the fourth moment.

Anderson considers formulas (1) and (2) as [practically] distribution-free. In economics and sociology the probabilities and the corresponding confidence intervals are sometimes the main indicators for a correct decision. [...] Thus, both the [Bienaymé –] Chebyshev inequality and formulas (1) and (2) are more robust with respect to the normal

distribution than the Laplace formula and are more preferable. The further development and application of the statistical methods corroborated that inference.

Very remarkable is Anderson's opinion about the choice of probabilities in the applications of one or another test. The representatives of the Anglo-American school¹¹ (Yates 1949) argue that the 5% level of significance is sufficient in all cases which is not however altogether true. [...]

Anderson devoted special attention to the issue of modelling in econometrics. The errors in the model itself, i. e. its deviation from reality, together with errors in the data can corrupt the appropriate "law" of distribution beyond recognition. The correct way out here is the application, as the least evil, of the classical methods of Markov and Chuprov¹².

[5] One more of the Anderson's contributions ought to be mentioned¹³ [...].

A synthesis of sorts of Anderson's achievements is his book (1954). It is methodologically advantageous on a number of points, and in essence it provides sailing directions for applying methods of mathematical statistics in economics. [...]

Anderson describes the main methods and indicators (including parameters of distribution and indices, also the index of the cost of living), statistical errors, the concepts of probability and expectation (with a discussion of the propositions made by Kolmogorov, Khinchin and Gnedenko), the law of large numbers, statistical hypotheses and their appraisal, induction and deduction in economics and sociology¹⁴. [...]

The notion of *equipossibility* of events is very subjective and the demand of infinitely many trials for approaching mathematical probability is unrealistic. Anderson therefore believed that the theory of probability was justifiably applied for studying social economic phenomena of being considered from the point of view of axiomatics¹⁵. [...]

Notes

1. It was preposterous to elevate Russian statistics which had for several decades been governed by ideology (Sheynin 1998).

2. This is misleading; Anderson had only been a student, and, for that matter, not a trained mathematician at all: he only studied mathematics for one term.

3. In Russia, Anderson only taught at a commercial school (Petersburg) and commercial institute (Kiev), see [VII, § 2].

4. At the time, Chuprov had no such published theory, but see [X, § 8].

5. The author describes the treatment of time series but specialists are acquainted with it whereas uninitiated will not understand him.

6. The variate-difference method was at least to some extent also developed by Student (Gosset).

7. I have not found them in Anderson's *Selected Works* (see Bibliography).

8. Indeed, various indices, even if calculated by Soviet authorities, had not been known to the public.

9. There follows a quite elementary discussion of systematic errors.

10. Since formulas (1) and (2) are not distribution-free, they only strengthen the Bienaymé – Chebyshev inequality under a certain restriction, see below.

11. Above, the author mentioned Anglo-American schools (plural). Anyway, Anderson was certainly better acquainted with this point, see [XIII, § 2.2].

12. Where did the author find these methods?

13. The author discusses the book *Das Konjunkturtest-Verfahren und sein Beiträge zur empirischen Untersuchung der ex ante – ex post Problematik*. München, 1957. It is not

mentioned in the list of Anderson's publications provided at the end of Bd. 2 of his *Ausgewählte Schriften* but included in the *Gesamtverzeichnis des deutschsprachigen Schrifttums 1911 – 1965*. I think that its author was Oskar Anderson junior.

14. Here follows a discussion on the lines of Chuprov's *Essays* (1909), see my Introduction.

15. Those interested in this issue ought to look up Anderson himself (but where exactly?). I, for my part, am certainly doubtful since statisticians hardly ever followed him. See also the very end of [XVIII].

Bibliography

O. Anderson

Abbreviation: his *Ausgewählte Schriften*, Bde 1 – 2. Tübingen, 1963: AS

(1929a), *Zur Problematik der empirisch-statistischen Konjunkturforschung. Kritische Betrachtung der Harvard-Methoden*. Veröff. Frankfurter Ges. Konjunkturforschung, Bd. 1, No. 1, pp. 123 – 165.

(1929b), *Die Korrelationsrechnung in der Konjunkturforschung. Ein Beitrag zur Analyse von Zeitreihen*. Ibidem, No. 4, pp. 166 – 301.

(1947), Zum Problem der Wahrscheinlichkeit a posteriori in der Statistik. *Schweiz. Z. f. Volkswirtschaft und Statistik*, Bd. 83, pp. 489 – 518 + Ergänzung. AS, Bd. 2, pp. 683 – 714.

(1951), Über den Genauigkeitsgrad wirtschaftsstatistischer Daten. *Weltwirtschaftliches Archiv*, Bd. 67, pp. 8* – 15*. AS, Bd. 2, pp. 836 – 843.

(1954), *Probleme der statistischen Methodenlehre in der Sozialwissenschaften*. Würzburg.

(1956), Verteilungsfreie (nichtparametrische) Testverfahren in den Sozialwissenschaften. *Allg. stat. Archiv*, Bd. 40, pp. 117 – 127. AS, Bd. 2, pp. 927 – 937.

(1959), Mathematik für marxistisch-leninistische Volkswirte. *Jahrbücher f. Nationalökonomie u. Statistik*, 3. Folge, Bd. 171, pp. 293 – 299.

Other Authors

Morgenstern O. (1950), *On the Accuracy of Economic Observations*. Princeton. Later editions: 1963, 1965.

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

Yates F. (1949), *Sampling Methods for Censuses and Surveys*. London. Later editions: 1953, 1960.

X
O. Anderson

To the Memory of Professor A. A. Chuprov Junior

Archiv za Stopanska i Sozialna Politika
2nd year, No. 3, 1926, pp. 291 – 299
Also in author's *Ausgewählte Schriften*, Bd. 1.
Tübingen, 1963, pp. 28 – 38

[1] Professor Aleksandr Aleksandrovich Chuprov died in Geneva on April 19, after a serious heart disease lasting six months¹. His death in full flowering of his mental powers is a powerful blow not only for the *Russian statistical school* but for the entire statistical science in which the deceased occupied one of the first places along with Pearson and Bortkiewicz.

[2] Chuprov's life was not especially rich in external events and can be described in a few words. The only son of a well-known Moscow professor of political economy and statistics, he received a splendid education at home, then for a comparatively short period attended the last classes of a gymnasium. His main scientific interests had been in the field of social sciences, but he graduated from the mathematical faculty of Moscow University because of being convinced that the phenomena belonging to them should be studied by the statistical method which needed, however, a deeper philosophical and in the first place mathematical basis.

After very successfully finishing the university, Chuprov went abroad for extending his economic and philosophical education, at first to Berlin, then to Strasburg, to Professor Knapp. There, he defended his dissertation (1902) which in a natural way turned to him the attention of specialists. Chuprov then passed with distinction an examination in Moscow for the Russian master degree in political economy and statistics. By the autumn of the same year he was appointed Dozent of statistics at the economic department of the just established Polytechnical Institute in Petersburg.

In that capacity he actively participated in the life of the really remarkable educational institution, the favourite creation of Count Witte. In 1909 Chuprov published his *Essays* whose first edition of 1200 copies was sold out in less than a year. For a scientific contribution this was strange and already during the next year, 1910, a second still larger edition became necessary.

On 2 December 1909 Chuprov brilliantly defended his dissertation in a public discussion at Moscow University, and what was even more seldom in Russian circumstances, the faculty skipped the master degree and at once bestowed upon him the degree of Doctor of political economy and statistics. After that Chuprov from being a Dozent was directly appointed full professor at the Polytechnical Institute.

He became the generally acknowledged head of a scientific direction and the circle of his followers and admirers became ever wider and a group of his own students formed around him. And thus the so-called

Chuprov school had come into being, but war and revolution put an end to this development. In 1917 Chuprov left Russia.

Some savings allowed him to live several years extremely modestly as a *private researcher* at first in Stockholm, then in Dresden. For Chuprov, those years were a period of unusual creative scientific achievement but they also finally undermined his probably always delicate health.

Separated by shortest intervals of time, his remarkable monographs had been appearing in *Biometrika*, the *Journal of the Royal Statistical Society*, *Metron* (Padua, Italy, edited by Gini), *Nordisk Statistisk Tidskrift*, *Skandinavisk Aktuarietidskrift*, *Ekonomicheskyy Vestnik*, *Russkii Ekonomicheskyy Sbornik* [both edited by Prokopovich], *Trudy Russkikh Uchenykh za Granitsei* [Berlin] etc. and a book (1925).

[3] These publications became the foundation of Chuprov's wide international reputation. The Royal Statistical Society honoured him, the Corresponding Member of the Petersburg Academy of Sciences and a long-standing member of the International Statistical Institute, in a manner extremely rare when concerning a foreigner, by electing him Honorary Fellow.

[4] In the summer of 1925 Chuprov's personal savings had finally dried up and he decided to take a professorship in a foreign university. Oslo (Christiania), Heidelberg [in Germany], Prague, Riga and even the Soviet Union offered him chairs. He chose Prague, but it seems that this decision was unfortunate. Some bureaucratic delays and difficulties, so dangerous for his weakened heart, and perhaps overwork brought about the first attacks of the fateful disease. As he wrote me in mid-August of 1925²,

These last weeks I am feeling extremely bad. Acute insomnia wore me completely down. My head does not want to work and, in responding to my attempts to compel it, it answers by rushes of blood and rises of temperature up to 38° and higher. And as though on purpose, much work has piled up. Teubner had finally energetically begun printing my book [1925] so that during the three last weeks I have been overloaded with the proofs. I dream very much of a holiday and hope to return back to my normal track.

And Chuprov went South to his beloved Italy but there he did not feel any better either. At the end of December, already quite ill, he managed to come to Geneva. Heart specialists pronounced his situation hopeless and the end came in the morning of April 19: Chuprov died in his sleep without waking up.

[5] Chuprov's scientific works are unusually numerous even if we only take into account what he himself allowed to publish. The same amount or even more is contained in his sketches and manuscripts and we ought to hope that they will be arranged and published. Chuprov worked, in his own words, "on the border between statistics, mathematical theory of probability and logic".

However, he permanently wandered in all directions from there making more or less remote journeys. That happened especially often in the beginning of his scientific activities with the excursions being made into the field of logic and philosophy, but at the end of his life he felt himself most powerfully attracted by the theory of probability and mathematics in

general. And, as he himself admitted, work in mathematics allowed him to forget all that formerly distressed him.

[6] The main goal of Chuprov's scientific work consisted in constructing a unified scientific system encompassing the entire theory of mass phenomena and in uniting the different directions of statistical thought not rarely hostile to each other due to misunderstanding in an item of a higher order; namely, the Anglo-Saxon direction headed by Galton and Pearson, the Continental direction of Lexis and Bortkiewicz, the philosophical direction connected with Windelbrand and Rickert etc³.

And exactly because of that he devoted his *Essays* to a deeper philosophical justification of the statistical method, then at once beginning to inspect all its main procedures. The theory of stability of statistical series brilliantly popularized (and deepened at the same time) in the *Essays* was developed anew and logically completed in a number of his monographs published during his last years.

In this connection he derived many new, often important indeed formulas, formulated and solved many most difficult problems belonging to probability theory, for example finding the expectation of a quotient of two [random] variables, investigating the so-called problem of moments⁴ etc.

Chuprov's final inferences were in a certain respect sensational and strongly changed, and left over pretty little of the traditional Lexian doctrine with its celebrated dispersion coefficient. He undertook a detailed and deep investigation of stochastic interrelation and its distinction from functional relation, or, less precisely but therefore more easily understandable, of the problem of comparing two statistical series and clarifying their interrelation.

[7] In a number of (sometimes extremely difficult in the mathematical sense but therefore especially interesting) monographs completed in his last book (1925) Chuprov had been able to unite the Pearsonian correlation theory with the standpoint of the Continental mathematical statisticians. These specialists previously held a sceptical if not altogether negative opinion on the somewhat mathematically confused and not invariably sufficiently rigorous considerations of the English school.

Chuprov deepened, enriched and specified their doctrine in such a way that even Bortkiewicz, not at all a follower of Pearson and his school, perceived in that last book a direct supplement and conclusion of the *Essays*⁵.

Then, we are obliged to Chuprov for his perceptive investigations of other problems of statistical methodology, some of them most highly significant. I recall, for example, the really important for the practising statistician theory of sampling, some problems of insurance statistics, of the theory of calculating mortality, of the sex ratio at birth.

Purely economic subjects were not alien to him either although he only treated them incidentally. There is only one subject which Chuprov, as far as I know, never touched in his published writings although being interested in it, the partition of complicated series pertaining to economics into components. I am not sure although have every reason to suspect that this restraint was caused by Chuprov's delicacy in scientific relations. Two of his students including me had chosen that subject for their research

and Chuprov apparently did not wish to interrupt undesirably the ripening fruit of their investigations⁶.

[8] In Chuprov's writings it should be stressed first of all the crystal clarity and harmony of his thought coupled with an unusually rich acquisition of knowledge. He felt himself equally at home in philosophy as in exact sciences, in economics or theoretical or practical statistics. He did not like the last-mentioned discipline, but definitely respected it. Because of his remarkable ability to scientific work he easily appropriated any branch of human knowledge which for any reason awakened his interest.

The eminent Ostwald is known to have separated all scientists in his work about great personalities into two diametrically opposed types of equal value, classics and romantics⁷. Chuprov unquestionably belonged to the former (and just as undoubtedly P. B. Struve, for example, was among the latter). However strange it may seem, Chuprov, in spite of his unusual gift, worked very slowly. He dealt with his scientific ideas for years since he thought them over from all sides, examined and analysed them in every possible way. Unripe ideas not entirely thought out even if fruitful which is not rare but typical for a *romantic*, Chuprov did not stand either in his own works or not and pursued them especially stubbornly in cases concerning his students.

And he detested, as he himself formulated it, "mixtures of different styles" because he regarded printed words unusually strictly and expected from each author a clear picture of for whom and what does he write. For himself, Chuprov only allowed two styles: either the traditional scientific of the type of the *Essays*, understandable by each intelligent reader and outstanding by elegance, or a compressed description of the results of his own research (mostly mathematical symbols) only understandable by specialists but also delicate in a special way, possessing logical clarity, consistency and perfection causing all the initiated to admire it. A monograph of this kind on the moments [of a random variable] compelled even Pearson, the proud Englishman, to publish an Editorial in his *Biometrika* under the significant title *Peccavimus!* [we were guilty]⁸. Chuprov's work of this kind about the coefficient of dispersion inspired Keynes, who even explained that among German and Russian mathematical statisticians only the form of Chuprov's papers can give aesthetic pleasure.

How carefully Chuprov prepared his publications can be perceived from the following. Two thirds of the main contents of his *Essays*, as I myself had convinced myself, could be found in his Moscow student *Composition* written at least ten years previously⁹, and we, his students, had already heard the significant part of his last monograph (1925) in his lecture of 1910. On the other hand, the remarkable course in insurance statistics, about which Chuprov had written me already a few years ago that it was almost prepared for publication, did not yet appear.

And his outstanding lectures on the general theory of statistics, which he had been reading each year at the Petersburg Polytechnical Institute from 1902 to 1907, incessantly extending and perfecting it, had not been published. It alone would have probably ensured for Chuprov the first place among all contemporary theoreticians of our science. In my opinion, he would have needed at least ten years more to prepare for publication

that what he already started and completed in its main features. It would be a crime pure and simple if those now in possession of the manuscripts of the deceased will not compile and publish his literary heritage. Even if the form will not be as accomplished as it would have been if done by him himself, it is still important that the contents of his creative thought are not lost.

[9] In spite of the method of his scientific work and the rigid boundaries which he set for himself in this sense, Chuprov had not at all resembled the dry German *scientists* who isolated themselves in the narrow confines of their speciality. On the contrary, he always was a cheerful, friendly and fascinating interlocutor (at least for those selected) and the field of his interests was unusually large, and just as unusually large was the region of his learning in all the branches of human knowledge. As one of the managers of the *professorial* newspaper *Russkie Vedomosti*, Chuprov attentively followed Russian political and economic life and not rarely compiled his own meaningful articles¹⁰. As far as I remember, he never discussed his pertinent activity with his students entirely separating his public writings from the field of work of an eminent scholar.

Chuprov had been greatly interested in music and belles lettres; he is known to have been the first who acquainted Russia with the Spaniard Bl. Ibanez. We even suspected that he secretly wrote verses but I do not know whether this is true¹¹. In any case, since being responsible for acquisitions for the large library of the statistical-economic Room managed by Chuprov, and therefore having had permanent contacts with ordinary and second-hand bookshops, I very often located for him rare editions of classics, for example of Pushkin whom he especially loved. Once I even bought for Chuprov Belyi's *Symbolism* [collected articles, 1910] and was properly reprimanded by him for my insufficiently respectful assessment of that, as I thought at the time, apparently hardly sensible writing.

[10] When looking at that bearded and imposing man with light-coloured hair combed smooth to the right side of his wide forehead; seeing how he kindly glanced at the audience through his glasses; noticing his cheerful smile often accompanying his words; hearing his exemplary fluent speech in the perfect Moscow accent with a merry joke inserted now and then invariably followed by a noisy and merry echo from the huge audience, – we, students, believed that we had before us a rarely robust man born by the grace of God to be a speaker able to lecture effortless.

In truth, however, these lectures gave him a lot of trouble and his health was always really delicate. Only after some years had passed, I understood how much the preparation of those lectures must have cost him and how even the three or four weekly lecture hours strained him. Chuprov himself confessed to me that during his first years of pedagogical work he had felt himself completely worn out after a one-hour lecture and had to rest on a divan all day long, and rather often thought of quitting teaching altogether. Each public speech deeply excited him and there were cases when this outstanding speaker just read out the always prepared beforehand speech to his usual audience.

It is evident that such a delicate and nervous organism, which could have perhaps go on living several decades more, had soon burnt out under the rough and relentless conditions of life of an emigrated professor.

[11] Chuprov was unusually kind and warm. Anyone requesting his advice or help on some personal occasion could have reckoned on an attentive and friendly participation. For his students and for all those whom he considered scientists, actual or prospective, Chuprov literally spared neither time nor efforts without likes and dislikes¹². I remember, for example, how he helped one of the students, Kushin [Kuchin?], to find a job and saved him from an entirely justified as I ought to add, police accusation, although during one of our stupid student strikes that very Kushin had led a group of rowdy students who *removed* Chuprov from his lectern and unashamedly abused him. I remember also Chuprov's unchanged benevolent attitude to me myself after I had twice extremely radically argued with him about his, as I believed then, unnecessary strict scientific demands.

[12] Again, Chuprov, who had in his paternal home met many Bulgarians, knew Bulgaria fairly well and loved it, led me to accept an invitation to a chair in the Higher Commercial School in Varna in that country¹³. I am still keeping his letter in which he expressed trust in the future of that nation and asked me to disregard unfavourable rumours about it. I also remember that as a specialist he held a really high opinion about the Bulgarian official statistics and believed that under K. Popov their General Direction of Statistics had reached the level of Western Europe.

[13] In my time, that is, between 1907 and 1915, among those teaching at the Economic faculty, were professors M. Kovalevsky, N. I. Kareev, P. B. Struve, M. A. Diakonov (academician), Yu. S. Gambarov, I. I. Ivanniukov, A. S. Posnikov, B. E. Nolde, B. M. Gessen and many other scholars and lecturers of the first rank. And still, rarely had anyone of them for a long time an audience as well packed as Chuprov had. I think that no other professor with the possible exception of Struve was beloved as much as Chuprov. Almost all the economists of the [graduated from the?] Polytechnical Institute had attended his lectures, and a considerable part of them met with him afterwards in connection with their own scientific work, or at the sittings of the optional seminar for those qualified, in his apartment on the fourth floor of the professorial house, or on the marvellous premises of our *statistical-economic* Room. Therefore, so many previous polytechnicians until now consider themselves students of the teacher devoted to that dry, and in addition mathematical statistics!

Nevertheless, Chuprov always had only a small number of students in the narrow sense of that word. He selected them very cautiously among those who already had some previous experience in mathematics or were prepared to deepen themselves earnestly in the study of higher mathematics. Being ascetically strict in his own scientific work, Chuprov was relentlessly demanding in respect of his students, and it was not at all easy to satisfy him.

However, he never encroached upon our scientific freedom and did not spell out any scientific activity in accordance with some definite programme. And it was a misfortune, pure and simple, if the Council decided to publish a certain statistical dissertation at the Institute's expense. Without bothering about time, whether his own or not, he compelled the unfortunate author to rewrite his work several times over since he corrected and discussed not only its general outline and manner of

presentation, but literally each of its thought and phrase. Even that was insufficient: he personally read each of the proofs from beginning to end and conclusively corrected them.

I recall how Chuprov in his study had once entered into a heated debate with his beloved student, M. M. Vinogradova (who died much prematurely from typhoid shortly after the Revolution) about whether one single digit, an eight, belonged to the same script as the entire text or should it be replaced by the same digit of the appropriate typeface. I was also called in the discussion which really could have hardly been resolved without a magnifying glass¹⁴. On the other hand, anyone who had experienced such a test became accustomed for all his/her life to the habit of exact and serious scientific work and kept feeling deep admiration for the unerring and selfless teacher.

[14] Many of us, former students and admirers of Chuprov, today find ourselves on this or that side of the chasm separating Soviet and anti-Soviet Russians and much bitterness and mutual misunderstanding between us has built up. But I still believe that all of us, White and Red, are united today by a common feeling of deep mourning for the premature death of our esteemed teacher in the prime of his life, who left us in full power of his mental capacities without concluding even a half of what he could and should have done.

The entire scientific world grieves together with us.

Notes

1. The author mistakenly added “being 53 years old”; actually, 52.
2. At about the same time Chetverikov (Sheynin 1990/1996, p. 104) and Bortkiewicz (Bortkevich & Chuprov 2005, Letters NNo 207 – 209) received similar letters.
3. There was no *etc.*
4. Possibly: study of densities by derived moments.
5. Anderson possibly had in mind Bortkiewicz’ letters to him about which nothing is known.
6. Chetverikov also studied that subject although perhaps after graduation (Manellia 1998, p. 95).
7. Ostwald (1853 – 1932) is mostly remembered as the initiator of the series *Ostwald Klassiker der exakten Wissenschaften*.
8. The Editorial, indeed published in *Biometrika* (vol. 12, 1919, pp. 259 – 281), corrected factual mistakes and inaccuracies mostly indicated by Chuprov whom Pearson acknowledged.
9. That *Composition* (1896) is kept by an archive in Moscow. I have read it, and believe, contrary to Anderson, that by 1909 Chuprov should have changed much.
10. I have traced more than 60 of Chuprov’s articles there (Sheynin 1990/1996, pp. 129 – 131), but, apart from Anderson, no-one mentioned that Chuprov was “one of the managers”.
11. One of his verses was recently published by Dmitriev (Eliseeva et al 1996, p. 62).
12. Apparently in 1922 Anderson sent Chuprov a manuscript for the *Journal of the Royal Statistical Society*, obviously of his future paper (1923) or even for (1927). Chuprov (letter of 1922 to Chetverikov; Sheynin 1990/1996, p. 60) “spent a good two weeks on his calculations”, informed Anderson about the changes done after which he “completely lost heart”.
13. Did Anderson know that Chuprov “at last” had been able to “fix him up with a post” [in Varna]? See *Ibidem* Chuprov’s letter of 1924 to Chetverikov.
14. I do not understand this at all, and the less so since late in life Chuprov allowed himself to use monstrous notation difficult for printing and horrible for the reader.

Bibliography

A. A. Chuprov

(1896), *Matematicheskie Osnovania Teorii Statistiki. Teoria Veroiatnostei i Statisticheskoy Method* (Mathematical Foundations of the Theory of Statistics. Theory of Probability and Statistical Method). Composed by student Chuprov of the phys. & math. dept, seventh term. Gorky Library, Moscow University. The Chuprovs' Fond, 9/1.

(1902), *Die Feldgemeinschaft*. Strassburg. Editor, G. F. Knapp.

(1909), *Ocherki po Teorii Statistiki* (Essays on the Theory of Statistics). Moscow. Later editions: 1910, 1959.

(1925), *Grundbegriffe and Grundprobleme der Korrelationstheorie*. Leipzig – Berlin. Russian editions: Moscow, 1926, 1960. English edition: London, 1939.

Other Authors

Anderson O. N. (1923), Über ein neues Verfahren bei Anwendung der variatē-Difference Methode. *Biometrika*, vol. 15, pp. 134 – 149, 423.

--- (1927), On the logic of the decomposition of statistical series into separate components. *J. Roy. Stat. Soc.*, vol. 90, pp. 548 – 569.

Bortkevich V. I., Chuprov A. A. (2005), *Perepiska* (Correspondence), 1895 – 1926. Berlin. Also www.sheynin.de

Eliseeva I. I., Dmitriev A. L., Storchevoi M. A., Editors (1996), *A. A. Chuprov. Materialy Konferentsii* etc (A. A. Chuprov. Conf. 70 Years after His Death). Petersburg.

Manellia A. (1998, in Russian), Life and scientific work of N. S. Chetverikov. *Voprosy Statistiki* No. 10, pp. 94 – 96.

Sheynin O. (1990, in Russian), Chuprov: *Life, Work, Correspondence*. Göttingen, 1996.

XI Oskar Anderson

Ladislaus von Bortkiewicz¹

Z. f. Nationalökonomie, Bd. 3, 1932, pp. 242 – 250
Also in author's *Ausgewählte Schriften*, Bd. 2.
Tübingen, 1963, pp. 530 – 538

[1] On 15 July 1931, in Berlin, Professor Ladislaus von Bortkiewicz died from a heart disease. His sudden death carried him off in the prime of his mental abilities and at the time of his fruitful scientific activity, when his capacity for work had hardly decreased. The international science, which lost in his person an outstanding economist and one of the really eminent scientists in the field of mathematical statistics, experienced a powerful blow.

Ladislaus von Bortkiewicz was born 7 August 1868 in Petersburg. Of Polish descent, he nevertheless grew up in an entirely Russian cultural milieu², and there also, in Petersburg, he studied in the university. The first significant contributions of the young scholar appeared in the beginning of the 1890s (1890a; 1890b; 1891) and somewhat later (1894 – 1896). In those years, he signed his name in quite the Russian manner: Bortkevich.

Having been supported by Lexis and Knapp, he was able, in 1895, to become a Dozent at Strasburg University and for two years taught there insurance of labourers and theoretical statistics. The same period witnessed the beginning of his close scientific relations with another eminent Russian statistician, Chuprov, six years his junior, who then received a doctorate under Knapp. Their friendship ended when Chuprov prematurely died in 1926.

Upon returning to Russia, Bortkiewicz became, in 1899 – 1901, Dozent at the prestigious Aleksandrovsy Lyceum in Petersburg out of which a number of most prominent Russian statesmen had come. In 1901 Bortkiewicz received an invitation to fill the position of extraordinary professor of economics and statistics at Berlin University to which he had been remaining loyal all the 30 years until death. Nevertheless, he only became full professor in 1920.

[2] Bortkiewicz' scientific lifework can be described thus. In theoretical statistics he was an acknowledged master and head of the school, or more precisely, of a stream known as *Continental*. It originated with a few papers published by Lexis in the 1870s, but certainly would not have acquired its present significance without Bortkiewicz' innovative studies. Our (younger) 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 then managed to discover.

We are very much obliged to Bortkiewicz for clearing up the philosophical and cognitive fundamentals of the theory of the statistical method. He is meritorious for clearly indicating the essential significance of the Poisson form of the law of large numbers for statistics, and he also

completed in a certain logical way the theory of the coefficient of dispersion Q^2 thus enabling its further development.

[3] In addition, Bortkiewicz essentially improved the methodology of mathematical statistics and introduced a number of new effective procedures. Here, the method of so-called mathematical expectations should be mentioned, a tool whose paramount importance is now being ever more widely recognized. Among his separate contributions the law of small numbers [1898]³ had in its time attracted special attention, but its practical value occurred lower than originally thought. Then, exceptionally significant were his deep investigations in the theory of index numbers as well as his last work on the mathematical analysis of the statistics of income (1930) presented the same year to the Tokyo session of the International Statistical Institute. Bortkiewicz also thoroughly and very commendably engaged in mathematical insurance and population and moral statistics.

His ideas essentially enriched statistical studies in Italy, Scandinavia, Russia and France. Even in the Anglo-Saxon statistical world which went ahead along its own route under Karl Pearson seemingly contradicting the *Continental* direction, Bortkiewicz' influence had been unquestionable. Only among the anti-mathematical German statisticians he did not excite any serious response. Nevertheless, it appears that there also a new livening of mathematical statistics is approaching so that they will perhaps recall him again.

[3] In the field of economics I ought to mention first of all Bortkiewicz' fruitful debates with Walras – Pareto, his consideration of the work of Marx, Bohm-Bawerk and Knapp in which he discussed all the main theoretical issues to such a considerable extent. Although not at all an eclectic and combative in other areas, Bortkiewicz kept here to a more conciliatory standpoint and calmly acknowledged the sensible kernel in the doctrines of both debating sides concerning such serious disagreements as between objectivism and subjectivism in the theory of value, between the nominalistic and metallistic theories of money etc.

He is also meritorious in fostering the economic theory in Germany at the time when it was seriously neglected in almost all universities. For mathematics, Bortkiewicz is important first of all as a first-rate researcher in the field of the theory of probability and his book (1913) is of lasting importance for physics.

[4] Striking in personal contacts with Bortkiewicz and also perceived in all his writings was his unusually sharp, and, as it is almost possible to say, merciless analytical mind which did not stand scientific errors or blunders, either his own or not. He quite surprisingly endured checks of numerical examples and derivations of mathematical formulas. For him, even a smallest [inaccuracy] in his own work was never insignificant, so that his examples and formulas are absolutely certain and at the same time the main features and ties are there sufficiently worked out⁴.

The breadth of Bortkiewicz' knowledge was immense. He felt himself equally sure in all areas of theoretical statistics and economics as well as in insurance, mathematics and some chapters of physics. The style of his work was peculiar and resembled Edgeworth's manner and in any case it occurred entirely unusual for some German economists.

During his scientific activities which lasted more than 40 years, Bortkiewicz published a lot of separate investigations but did not create any *system*, had not systematically summarized in a single contribution the results of his own work and the writings of other authors in any wide area of science, and did not compile even a single extensive book [except (1913) and (1917)].

However, when attentively considering his research as a single whole, it is easily seen, as I mentioned above, that he investigated almost all those important problems which interested theoretical economics and statistics in our time. From the point of view of Ostwald's⁵ generally known classification who separated all talented scholars into romantics and classics, Bortkiewicz certainly belonged to the first group.

In spite of his love of approaching each problem from his own special side, he needed a certain prompting for starting up the incomparable machinery of his mind. Not rarely the scientific investigations of other authors served as such a point of departure. He first thought them over, then wove further, reconstructed, and sometimes completely rejected the original structure. And it is not at all accidental that his best and deepest works he always began as usual reviews. Thus, *Die Iterationen* (1917), for him a book of quite an exceptional volume of 205 pages, emerged from a review of Marbe (1916) and simply crushed that author. In a similar way, three papers on index numbers (1923 – 1924) resulted from a review of Fisher (1922) who was nevertheless much more mildly damaged⁶.

In his reviews Bortkiewicz not rarely refuted those very ideas that had prompted him and initiated his response, but later became unacceptable. To a large extent this explains the strictness and sharpness of his judgement which sometimes alienated and in any case offended the authors.

According to the general opinion, Bortkiewicz was a harsh and acerbic judge and even the most prominent scientists took into consideration his statements. It was maintained not altogether jokingly that his scientific significance consisted in that he not only put forward one or another theory or influenced this or that researcher, but that he also led authors of quite a lot of weaker contributions to fear his destructive criticism and avoid publishing them which was certainly very advantageous for science.

However, we should never forget that Bortkiewicz' judgement had always been factual and remained impartial; no scientist had been personally closer to him than Chuprov, but there occurred scientific duelling between them with exchanges of very painful blows. In his immediate surroundings Bortkiewicz could have been charming and his home in Berlin remained a place of pilgrimage for scientists the world over coming to express themselves and receive advice.

[5] As Chuprov's student, I belong however to the younger generation in which eyes Bortkiewicz seemed a distant Olympian, but I also could have reported much about the patient kindness of the *strict master* and the many valuable stimulations in his letters to me⁷.

Bortkiewicz did not write for a broad circle⁸ and was not at all good at popularizing his own ideas. Then, he made great demands on the educational background and intelligence of his readers. With stubbornness conditioned partly by his asceticism and partly indeed by the *romantic* type of his scientific mind he refused to accept the advice of the classic

Chuprov and to choose an easier form for his publications. The applied mathematical machinery with which Bortkiewicz occasionally all but shone, made it especially difficult for German economists inclined against mathematics to penetrate the deeper essence of his theories.

Add to this that the titles of, and sources for his publications definitely were not always in accordance with what the reader would have justly searched for. For example, which statistician would have expected to find important theorems about the Lexian coefficient of dispersion and, really at the same time, the theory of the Pearsonian chi-squared test in Bortkiewicz' paper (1922) published, moreover, in a mathematical periodical? On the other hand, which mathematician will search for valuable articles on the theory of probability in the *Jahrbücher für Nationalökonomie und Statistik*?

His contributions are scattered over a lot of German and foreign periodicals a part of which now became difficult to come by. Out of his 54 lengthy statistical monographs only four are books or booklets, which had long ago disappeared from the book market. The rest 50 contributions are contained in 27 different German, Austrian, Russian, Swedish, Italian, Swiss etc journals and series.

Almost the same happened with his 23 long economic contributions. And I have not even included either a load of annotations and reviews which Bortkiewicz had written over the years, or his shorter items in the great Russian *Brockhaus & Efron Encyclopedic Dictionary*⁹ or similar reference books.

[6] I never had the advantage of being present at his lectures, but, as far as I know, Bortkiewicz had been preparing them studiously and diligently. Nevertheless, he was unable to gather a circle of students around himself either in Berlin or, for that matter, in Germany. Perhaps, as Altschul (1931) believes, this disconcerting fact can be explained by acknowledging that pedagogic activities were "not to his liking" which, however, should rarely be the case with *romantics*.

Another possible cause is that, although Bortkiewicz had spent almost half of his life at Berlin University, he invariably remained there almost an *alien element* which did not properly fit the conditions of work or traditions of German economic faculties. On the face of it, he had been indeed coldly approached with deep respect but was internally rejected. And Bortkiewicz' death certainly stronger impressed the scientific world in Italy, Scandinavia, Russia¹⁰ etc than inside Germany where the younger generation seems even to know him but little.

[7] Everyone aware of the real significance of the life work of this great researcher will therefore lively welcome at least a first step towards his *discovery* and, as it occurred with Edgeworth¹¹, a publication of his statistical and stochastic papers in a volume of collected works. If not in Germany, could not a publisher be found among the rich Scandinavian institutions?

A few years ago, a German publishing house had offered Bortkiewicz an opportunity to put out a collection of his more important investigations, but nothing came of it because he preferred new research to processing prior work. As he informed me at about the same time, he staked everything on leaving nothing half-finished. And still, dying unexpectedly, he left behind probably the most [of such work].

For making easier to acquaint the reader with Bortkiewicz' work, I adduce a preliminary list of his larger scientific publications. Annotations and reviews are not included and the list can also have many gaps¹².

Notes

1. I also mention the author's previous paper (1929) as well as his very short obituary which I (2001) reprinted. There also, I quoted archival documents from the present Humboldt University.

2. For Bortkiewicz, German was a mother tongue along with Russian which is seen in his publications and perceived as well since he delivered lectures in German and obviously experienced no linguistic difficulties while living in Germany. In 1905 he wrote to Chuprov (Bortkevich & Chuprov 2005, Letter No. 79) that he was feeling himself "perfectly well in regard to the kind, the conditions and the place of work" in Berlin.

It is probable that German was the language spoken in Petersburg in the family and that the main language in the gymnasium which he attended was also German.

3. Anderson was one of only a few statisticians apparently even then dissatisfied with the law of small numbers, cf. my paper (2008).

4. I strongly disagree, see for example Sheynin (2008).

5. Concerning Ostwald see [X, Note 7].

6. Woytinsky (1961, pp. 452 – 453) reported that publishers had quit asking Bortkiewicz to review their books.

7. Nothing is known about these letters.

8. Winkler (1931, p. 1030) quoted a letter from Bortkiewicz (without adducing its date): he was "glad" to find in Winkler one of his five expected readers!

9. Only one contribution (1897) to that *Enz. Slovar* is known.

10. Bortkiewicz certainly corresponded with Soviet statisticians; their and his extant letters are kept in his papers at Uppsala University (Sweden) and partly published in Russian. However, no obituary appeared in Russia; *Vestnik Statistiki*, the leading statistical periodical, had been suppressed for many years (including 1931). True, the first edition of the *Bolshaia Sovetskaia Enziklopedia* (Great Sov. Enc.), see vol. 7, 1927, p. 198, acknowledged his work in various areas of statistics but in 1933, the same edition (vol. 63, pp. 279 – 283) called both the Biometric school and the Continental direction of statistics *bourgeois*. In the second edition (vol. 5, 1950, p. 605) the appraisal was also extremely negative.

11. In 1925, his collected papers on economics had appeared (one volume). In our time (in 1996), his *Writings* were published in three volumes.

12. I do not reproduce it. My own Bibliography of Bortkiewicz' writings is in Sheynin (2001) to which I (2007, pp. 276 – 282) added many of his reviews.

Bibliography

L. von Bortkiewicz

(1890a), Auseinandersetzung mit Walras. *Rev. d'econ. politique*, t. 4.

(1890b, in Russian), Mortality and longevity of the male Orthodox population of the European part of Russia. *Zapiski Imp. Akad. Nauk*, vol. 63, Suppl. 8. Separate paging.

(1891, in Russian), Same for female population. *Ibidem*, vol. 66, Suppl. 3. Separate paging.

(1894 – 1896), Kritische Betrachtungen zur theoretischen Statistik. *Jahrbücher f. Nationalökonomie u. Statistik*, Bd. 8, pp. 641 – 680; Bd. 10, pp. 321 – 360; Bd. 11, pp. 701 – 705).

(1897), Accidents. *Brockhaus & Efron Enz. Slovar*, halfvol, 40, pp. 925 – 930.

(1898), *Das Gesetz der kleinen Zahlen*. Leipzig.

(1913), *Die radioaktive Strahlung als Gegenstand wahrscheinlichkeitstheoretischer Untersuchungen*. Berlin.

(1917), *Die Iterationen*. Berlin.

(1922), Das Helmerische Verteilungsgesetz für die Quadratsumme zufälliger Beobachtungsfehler. *Z. f. angew. Math. u. Mech.*, Bd. 2, pp. 358 – 375.

(1923 – 1924), Zweck und Struktur einer Preisindexzahl. *Nordisk Statistisk Tidskrift*, Bd. 2, pp. 369 – 408; Bd. 3, pp. 208 – 251 and 494 – 516.

(1930), Die Disparitätsmasse der Einkommensstatistik. *Bull. Intern. Stat. Inst.*, t. 25, No. 3, pp. 189 – 298 and 311 – 316.

Other Authors

Altschul E. (1931), Ladislaus von Bortkiewicz. *Magazin der Wirtschaft*, 7. Jg, pp. 1183 – 1184.

Anderson O. (1929, in Bulgarian and French), Professor V. Bortkevich. *Revue trimestrielle de la Direction générale de la statistique*, année 1, No. 1, pp. 7 – 9.

Bortkevich V. I., Chuprov A. A. (2005), *Perepiska (Correspondence) 1895 – 1926*. Berlin. Also at www.sheyinin.de

Fisher I. (1922), *Making of Index Numbers*. Boston, 1997.

Marbe K. (1916 – 1919), *Die Gleichförmigkeit in der Welt*, Bde 1 – 2. München.

Sheynin O. (1970), Bortkiewicz von, Ladislaus. *Dict. Scient. Biogr.*, vol. 2, pp. 318 – 319.

--- (1990, in Russian), A. A. Chuprov: *Life, Work, Correspondence*. Göttingen, 1996.

--- (2001), Anderson's forgotten obituary of von Bortkiewicz. *Jahrbücher f. Nationalökonomie u. Statistik*, Bd. 221, pp. 226 – 236.

--- (2007), *Chetvertaia Khrestomatia po Istorii Teorii Veroiatnostei i Statistiki* (Fourth Reader in History of Theory of Probability and Statistics). Berlin. Also at www.sheyinin.de

--- (2008), Bortkiewicz' alleged discovery: the law of small numbers. *Hist. Scientiarum*, vol. 18, pp. 36 – 48.

Winkler W. (1931), Ladislaus von Bortkiewicz. *Schmollers Jahrbuch f. Gesetzgebung, Verwaltung u. Volkswirtschaft in Deutschen Reiche*, 5. Jg, pp. 1025 – 1033.

Woytinsky W. S. (1961), *Stormy Passage*. New York.

XII Oskar Anderson

On the Notion of Mathematical Statistics

Einführung in die mathematische Statistik. Wien, 1935, pp. 1 – 2

Mathematical statistics is not a robot, and it is not always possible to stick statistical material into it, and, after some mechanistic manipulations, to read the result as though from a calculator
Charlier (1920, p. 3)

Little experience is sufficient to show that the traditional machinery of statistical process is wholly unsuited to the needs of practical researches. Not only does it take a cannon to shoot a sparrow, but it misses the sparrow!
Fisher (1932, p. vii)

Mathematicians playing statistics can only be overcome by mathematically armed statisticians
Chuprov (1922, p. 143)

The aim of statistical work is to turn the collected observational material into numbers and number series. And one of the most important problems of the statistical methodology consists in showing how to process these numbers by summing, abstracting, dividing etc so as to arrive at certain conclusions about the observed mass phenomena. Since the four arithmetical operations likewise obviously belong to the field of mathematics, it is hardly possible from a purely logical viewpoint to draw a clear and incontestable line between the *general* and the so-called *mathematical statistics*.

The latter term is as a rule understood as the chapter of statistical methodology that is represented by the aid of the calculus of infinitesimals, higher algebra and analytic geometry. It remains therefore incomprehensible for the average economist and statistician whose mathematical education ended with graduation from the secondary school. From this point of view and presuming the usual forms of presentation of the theory we ought to attribute to mathematical statistics a number of really heterogeneous sections and in the first place most applications of the theory of probability, the [Lexian] theory of dispersion, the theory of means and analysis of variance, the Pearsonian and other frequency curves, sampling investigations, theory of correlation, partition of series, harmonic analysis, theory of interpolation and adjustment, of index numbers, the formal theory of population, insurance mathematics and similar subjects.

Still, this opinion is considerably disadvantageous in that the essential contents of a scientific discipline depends on how is the mathematical school curriculum composed; moreover, it can change from country to country and from year to year. Considering this, it is exactly among the aims of this book to show how a number of the received theorems of

mathematical statistics can be easily proved or at least made understandable only by means of elementary school mathematics.

It would be much more likely to agree that the term *mathematical statistics* denotes that part of the statistical theory which is closely connected with the theory of probability and can almost be considered as its applied part. And this is, roughly speaking, the viewpoint chosen in this book. We think, however, that it is senseless to include the so-called formal theory of population with the closely connected insurance mathematics even if it is (as it does not always happen) tightly associated with the theory of probability. In respect to the subject of study and training the profession of actuary has already separated their young generation from other mathematical statisticians so that it is practically useless to pay special attention to the former in a book mainly intended for economists and statisticians.

Thus outlined, in different countries the methodology of mathematical statistics in its entirety occupies very differing positions in the science of statistics. In the English speaking world it is seen under the influence of the Pearson school as almost the statistical theory, but apparently in the Soviet Union in spite of the initial success of the Chuprov school statisticians are ever more turning away from it, see e. g. the Introduction to the latest Russian textbook Boiarsky et al (1931)¹. And concerning the German nations, today mathematical statistics finds itself in general in a really difficult position.

The reaction of real statisticians concerning mathematicians *playing statistics* is described by the three epigraphs.

Note

1. Anderson apparently had not wished to add the political dimension here. He (1959, p. 294), see also Sheynin (1998, p. 533), did so later by quoting that source thus showing the horrible situation then prevailing in Soviet science in general. He also remarked that the textbook was not up to contemporary statistical knowledge.

Bibliography

Anderson O. (1959), Mathematik für marxistisch-leninistische Volkswirte. *Jahrbucher f. Nationalökonomie u. Statistik*, Bd. 171, pp. 293 – 299.

Boiarsky A, Starovsky V., Khotimsky V., Yastremsky B. (1931), *Teoria Matematicheskoi Statistiki* (Theory of Math. Statistics), 2nd edition. Moscow – Leningrad.

Charlier C. V. L. (1920), *Vorlesungen über die Grundzüge der mathematischen Statistik*. Hamburg.

Chuprov A. A. (1922), Lehrbücher der Statistik. *Nordisk Statistisk Tidskrift*, Bd. 1.

Fisher R. A. (1932), *Statistical Methods for Research Workers*. Edinburgh – London.

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

XIII O. Anderson

Mathematical Statistics

Handwörterbuch der Sozialwissenschaften, Bd. 10, 1959, pp. 46 – 52

Also in author's *Ausgewählte Schriften*, Bd. 2.

Tübingen, 1963, pp. 945 – 951.

1. The Concept

Opinions differ about what should be understood as *mathematical statistics*. Neither the statistician, nor the economist can get along in his theoretical or practical work without applying the four arithmetical operations. Arithmetic, however, represents a mathematical science par excellence so that strictly speaking statistics in its entirety is *mathematical*. In that sense some modern Anglo-Saxon authors only perceive statistics as a part of applied mathematics which means that they certainly leave out the theory of business statistics in its proper sense (planning, investigating, preparing the work) which belongs to social statistics. Others, like M. G. Kendall, do not use the term *mathematical statistics* at all and obliquely describe its subject as *advanced theory of statistics*. If this expression becomes established, the situation in statistics will be to a certain extent similar to circumstances in physics, where *theoretical physics* was previously called *mathematical*.

Another important question is, where then should the border between *elementary* and *advanced* theories of statistics be. Some authors, especially German, perceive it to be already in the application of relatively simpler algebraic formulas and symbols whereas Anglo-Saxon textbooks of *Elementary statistics*, of its *Elements*, and of *Introductions* to the theory of statistics do not shy away from substantially applying algebra, analytical geometry and the main notions of the calculus.

If the not unquestionable expression *mathematical statistics* will hold its ground at all, it could be understood as meaning that part of the general statistical theory

1) Whose comprehension demands a certain *higher* level of mathematical knowledge

2) Which is closely connected with the theory of probability and from a certain standpoint can simply be regarded as applied probability theory

In the first case the subject of a scientific discipline will directly depend on the contemporary school curriculum, in the second instance certain *higher* parts of the statistical theory will be left aside.

2. History

2.1. The beginnings. Insofar as mathematical statistics is built on the *theory of probability*, the history of their emergence cannot be separated, and the names of Jakob Bernoulli (1654 – 1705), Laplace (1749 – 1827), Gauss (1777 – 1855) and Poisson (1781 – 1840) will belong to one of them just as to the other.

Certain elements of mathematical methods in social and especially population statistics already existed by the end of the 17th century in

political arithmetic, in the works of Graunt and Halley, and later Süßmilch (1707 – 1767). For the first edifice of mathematical statistics in the proper sense we are grateful to a student of Laplace, Quetelet (1796 – 1874)¹. Mostly because of the harsh reaction against the deterministic exaggerations made by the *Queteletians*, to which the young Adolph Wagner also belonged [and of the ensuing downfall of Queteletism], there appeared in the 1870s a German school of mathematical statistics whose most important representatives were Lexis (1837 – 1914) and Bortkiewicz (1868 – 1931). Their preferable field of study formed the theory of *stability of statistical series* and its practical re-examination.

This school has played an important role, especially in the building of population statistics and the theory of insurance. In spite (and, partly, also because) the most eminent German economists, Schmoller and Knapp in the first place, had actively participated in the struggle against Queteletism, the mental climate in our universities had been hardly favourable to the further development of the Lexis – Bortkiewicz direction. In a similar way, the statistically based investigations made by the psycho-physiologist Fechner, whom Lexis had only mentioned in passing, did not influence the theory of social statistics to any considerable extent.

The work of Knapp's student Chuprov (1874 – 1926)² brought to a certain conclusion the theory of the German school of mathematical statistics. Incidentally, he strove for some synthesis of that theory with the Biometric school, and at the same time he was the founder of the Russian school of mathematical statistics.

2.2. The development of the modern theory of mathematical statistics. It essentially began with investigations of heredity by the English biologist Galton (1822 – 1911)³ and the lifework of his younger colleague Pearson (1857 – 1936) with the previous work of Edgeworth making the speedy acknowledgement of their theories in England considerably easier. In the course of time an ever increasing circle of students and people allied to them had been gathering around Pearson and *Biometrika*, the periodical established in 1901 by Weldon, Galton and Pearson, which he edited.

At first, their field of studies was restricted to *mass phenomena in biology*, but very soon it expanded to include other scientific spheres, in particular social sciences (Bowley, Yule). In many cases there certainly existed valuable prior work made partly by older mathematicians and physicists and partly by the followers of Quetelet, but at that point it was the train of thought of the Pearsonian school that determined the further development of the appropriate disciplines. The investigations of Fisher ensured a new stage of the progress of mathematical statistics with wide theoretical prospects⁴.

Nowadays the direction taken by Pearson and Fisher is dominating in mathematical statistics but its certain *national versions* can differ. Next to the English school and the closely related to it American and Indian (Mahalanobis, Rao) direction it is especially necessary to mention the Italian (Gini), Scandinavian (Charlier, Cramér, Wold), Russian (Chuprov, Slutsky, Romanovsky)⁵ and French (Darmois, Fréchet) schools. In Germany, during the 20th century the theoretical achievements of the Lexis – Bortkiewicz school became forgotten to such an extent, that some of

them had to be *discovered* in England once more. Thus, as Kendall correctly noted, Lexis (1879)⁶ contained the essence of the Fisherian analysis of variance. Then, Helmert (1876) had derived both the law of distribution of the squared sum of random errors of observation⁷ which Pearson (1900) discovered in connection with his *chi-squared test*, and the precise expression for the variance of the mean error in case of the normal distribution found once more by Fisher in 1920.

Bortkiewicz (1898) had applied the Poisson formula for the frequency of random events in various directions, but then Student (1907) “discovered” it anew. Bortkiewicz certainly did not see many possible applications of the Poisson formula in the field of biological sciences and especially bacteriology, but, for example, the priority of statistically treating the counts of blood corpuscles belongs to Abbe (1876) rather than Student (1905) to whom the Anglo-Saxon literature usually attributes it.

The relatively modest attempts to revive mathematical statistics in investigations belonging to social sciences, which became noticeable in Germany after World War I, were again destroyed by the events of 1939 – 1945 when a number of its representatives had been forced to leave Germany and Austria for good⁸. A new period of gradually adopting the body of thought in mathematical statistics established meanwhile abroad has begun after 1945. An essential part of the related scientific *terminology* was borrowed in the English wording or in German back translation from English whereas for a number of terms there had already existed expressions coined and applied by German scientists, – mathematicians, physicists and statisticians, – for a whole century. This is especially true for the modern *opinion research* and partly for the statistical control of industrial production. In applications of statistics to medicine, biology and physics the break with tradition is less pronounced and there also the resistance to the mathematical methods of statistics was never as radical as in social sciences. Nowadays a special subcommittee under the German Committee of Norms is being occupied with standardizing terminology in mathematical statistics.

3. Fields of application

The essence of statistical investigations consists in

- 1) Defining the object of study
- 2) Establishing the features to be ascertained
- 3) Fixing the exact limits in time and space
- 4) Choosing the suitable objects

The appropriate statistical method is everywhere applicable when *statements about whole collectives or totalities* are needed so that it is not at all restricted to mass social phenomena. This conclusion is especially true in regard to the stochastic part of the theory of statistics.

The theory of probability emerged in the mid-17th century as some kind of a gentleman’s science. At first, its subject was only the estimation of the chances of winning in games of chance, then came, – and, incidentally, met with very little success, – the application to the evaluation of testimonies⁹.

Gradually, however, the stochastic point of view, in the closest possible manner connected with the statistical thought, fought its way in most various sciences: insurance and the appropriate part of population statistics, theory of astronomical, physical and geodetic measurements

(theory of errors), physics (kinetic theory of gases, thermodynamics, quantum mechanics, atomic theory), meteorology, biology, agronomy, anthropology and anthropometry, medicine, experimental psychology and psycho-physiology, linguistics and folklore, sociology, economic theory (econometrics), and various other parts of social sciences. One of the latest conquests made by the stochastic statistical methodology is the control of the quality of industrial production and the ensuing Wald *sequential analysis*.

Seen from this perspective, mathematical statistics certainly does not belong to social sciences and scientists working in that field and applying *higher* mathematical means are carrying on the *study of fundamentals* thus providing a theoretical foundation for a number of disciplines. However, the mass social phenomena possess some properties which distinguish them and make it desirable to develop a special applied theory of mathematical statistics.

First of all, these phenomena are special in that they are more or less intensively developing in time so that the obtained relative numbers, means and other mass indicators, are only stable for a short time and over a small region. In addition, the distributions of the masses are very complicated, often unsteady and of a mixed form and only rarely allow to assume that the parent distribution is *normal* which still is the dominant hypothesis in the modern theory of mathematical statistics.

The law of large numbers on which the theory of statistics is built, is not at all a purely mathematical theorem. According to Cournot and Chuprov, it can be derived in the following way:

1) For a sufficiently large number of trials, larger deviations of the observed relative frequency or mean from their expectations, which are calculated by issuing from the stochastic model in question, possess low probabilities.

2) Events with low probabilities occur seldom.

3) Therefore, the larger deviations are seldom.

Theorem 1) is a proposition of pure mathematics whereas 2) is after all founded on the collected evidence about the real events having occurred in our *macroworld* during several centuries¹⁰. However, the figures with which biologists and especially modern physicists are dealing, are incomparably larger than those that can be recorded for the objects of social statistics. [...] Therefore, when a modern physicist applying statistics bases his theorem on the law of large numbers, the statistician working in social statistics really has, on the contrary, every reason to complain about *the law of insufficiently large numbers*.

It follows from the described difference that a number of modern methods of mathematical statistics and tests, that quite satisfyingly prove their worth in the field of natural sciences, are more or less useless when applied to *mass social phenomena*. On the other hand, various special methods can be developed only for social sciences. Taken together with mathematical statistics of natural sciences, the general stochastic structure persists as well as the possibility to assume, given a large statistical mass, an approximate normal distribution for at least the relative frequency and mean.

This circumstance concerning the mean is especially significant for the application of the representative method. On the contrary, the possibility

of applying the *theory of small samples* as well as the usual nowadays form of the *analysis of variance* and *design of experiments* remains doubtful.

4. Contents

Looking from the standpoint of the theory of insurance [?], the essential contents of the methodology of modern mathematical statistics in the field of mass social phenomena can be summarized as follows.

4.1. The theory of business statistics. It has to do with planning and execution of statistical observations as well as with treating and presenting the obtained data, demands *sampling* and the application of the theory of probability and is therefore included in mathematical statistics.

4.2. The doctrine of totalities. Following Fechner, we understand this theory as that part of the general statistical methodology which has as its subject the description of various *indicators of the collective*, for example relative numbers and means. Therefore, the theory of certain complicated measures of scatter and covariance (the Pearsonian moments, Fisherian cumulants, Thiele's semi-invariants, the general theory of frequency distributions and their parameters, – skewness, excess, – as well as the formal theory of indices in the sense of Irving Fisher and Bortkiewicz), belong to the field of mathematical statistics when it is simply understood as the mathematically *higher* theory.

In this connection we may mention those various statistical constructions which attempt to meet the requirements of modern econometrics: the theory of empirical curves of supply and demand, problems of identification and aggregation etc. If, on the other hand, mathematical statistics is only understood as the stochastic part of the theory of statistics, then in any case the entire area of calculating the best or most effective¹¹ measures of the collective will belong to it. This means, when deriving by sampling, such estimators which relatively least deviate from their respective expectations corresponding to the parent distribution, the deviations being measured by their variance. These problems directly belong to estimation.

4.3. Estimation. Bearing in mind the true and the estimated values of statistical figures, two cases ought to be distinguished first of all from each other.

1) The differences between them, or the errors of estimation are not seen to be random but result, for example, from the incompleteness of the material, poor investigation or deliberate falsification of the *object*.

2) The deviations may be considered random, as though obeying the assumptions of the stochastic theory of errors.

The first case, with which the German *non-mathematical* theorists deal almost exclusively, may still, if need be, be considered *mathematical* when the statistical figures corrupted by error are connected one with another by addition, subtraction, multiplication or division because of the *propagation* of errors, a phenomenon which in Germany is not usually taken into account. [...]

An experienced statistician who really understands the appropriate mass phenomena is usually able to estimate approximately the possible interval of error of the estimate and not rarely the sign of its error. It is therefore possible, under certain circumstances, to choose the method of calculation or the respective indicator in a manner that prevents the often occurring

snowballing of the relative error after each additional arithmetic operation [...].

In the second case, when the errors are *random*, the application of partly involved methods of mathematical statistics is at once necessary. It ought to be pointed out, however, that the concept of randomness in mathematical statistics is stricter, i. e. defined in a more restricted way than according to its usual understanding in life¹². [...]

As to the choice of the most reliable or presumable estimate, this should be either the value which appears when the trials are repeated many (or infinitely many) times under the same arrangement and the appropriate arithmetic mean is taken, or the value which will occur then most often. Under certain additional assumptions the second alternative leads to the Fisherian likelihood concept. The principle of maximal likelihood in its proper sense amounts to choosing among various hypotheses or models that, which possesses, given the observed event or their combinations, the highest mathematical probability.

4.4. The study of causality. Among the various methods of such studies in mathematical statistics we should name

4.4.1. Various methods for establishing whether the difference between two numbers describing a collective or two distributions etc might still be only attributed to chance, or is it too large so that we ought to suspect an essential difference between the underlying complexes of causes. The appropriate methods of checking or tests are, according to Kendall, subdivided (not exhaustively) in the following way.

a) Tests for ascertaining whether the presumed value of a number can still be adopted as the true value of a parameter. The value 0 is often ascribed to the appropriate hypotheses (a null hypothesis as for example when judging a coefficient of correlation)¹³.

b) Tests of the agreement between a theoretically expected and an observed numerical series (the Pearsonian chi-squared test).

c) Tests of the homogeneity of two or more statistical masses or totalities. The opposite case concerns the discriminate analysis.

d) Tests for distinguishing between a random and a non-random appearance of a series of statistical numbers. This can be done, for example, by some measures or coefficients of autocorrelation. The tests themselves can be either parametric or distribution-free. In the first case the law of distribution of the parent population and its parameters are entirely or partly known. In the second case the derived tests do not depend on such assumptions.

4.4.2. Certain aspects of contingency and correlation theory. The modern theory attempts first of all to discover the causal, or functional dependence between the observed statistical number series, that is, the appropriate equation of regression and only applies the coefficients of correlation for estimating the influence of the disturbing random or non-random remaining complex of causes. The analysis of variance and covariance pursues similar aims.

4.4.3. Various methods for decomposing or adjusting statistical series. Here, the aim is to decompose a given series into separate parts each reflecting the influence of certain groups of causes.

5. Significance

Mathematical statistics is only a part of the general statistical theory. As such, its immediate goal is to develop mathematically *higher* or stochastic methods allowing to obtain more complete and precise knowledge from given numerical data as compared with *elementary* methods. We must not forget, however, that mathematical statistics has yet another not less important aim, to serve as a brake, to provide by proper tests timely warnings about the insufficiency of the data at hand for arriving at some important conclusion.

Notes

1. I believe that Fourier, the Editor of the fundamental *Recherches statistiques ...* published in 1821 – 1829, influenced Quetelet incomparably more than Laplace.
2. In this context, it was not Knapp but Markov with whom Chuprov corresponded in 1910 – 1917.
3. Galton was indeed a biologist, certainly not by education but since he successively studied heredity.
4. Even in Anderson's context Fisher deserved much more attention.
5. The *Russian school* included many more representatives.
6. References such as Lexis (1879) are not included in the appended Bibliography.
7. Anderson only mentioned Abbe later and in another connection.
8. I do not understand this mild statement. Anderson himself (1954/1962, p. 3n) discussed various estimates of the number of exterminated Jews and mentioned *unsere deutsche Schande* (our German shame) which did not lessen even if it *only* amounted to four million.
9. But where is insurance (Halley, De Moivre)? A few lines above Anderson mentioned a *gentleman's science*, but had not qualified his statement: from the very beginning (Pascal, Huygens) scientists had perceived there the emergence of a new branch of mathematics.
10. Yes, statisticians may insist that practice had justified the law of large numbers, but its rigorous mathematical proof is not therefore made superfluous. Then, sufficiency of data (cf. the last lines of Anderson's paper) is estimated, among other possibilities, by the mathematical (certainly not by the *statistical*) law of large numbers.
11. Here and in § 4.4.1 *effective* bears no relation to *effective estimators* in the strict sense.
12. Statistics only attempts to distinguish randomness from necessity by special stochastic tests.
13. I would say, *symbol 0* is ascribed etc.

Bibliography

I have arranged it in alphabetical order, and, except for the five first entries, I do not understand the principle of the initial arrangement of the separate items

- Buros Oscar K.** (1938), *Research and Statistical Methodology. Books and Reviews 1933 – 1938*. New Brunswick.
- (1941), *The Second Yearbook of Research and Statistical Methodology. Books and Reviews*. New Jersey.
- Bibliography** (1951), *Bibliography of Basic Texts and Monographs on Statistical methods*. The Hague.
- Bibliographie** (1952), *Bibliographie sur la méthode statistique et ses applications. Liste des principaux ouvrages en langue française*. Paris.
- Bibliographie** (1955), *Bibliographie der seit 1928 in Buchform erschienenen deutschsprachigen Veröffentlichungen über theoretische Statistik und ihre Anwendungsgebiete*. München.
- Anderson Oskar** (1935), *Einführung in die mathematische Statistik*. Wien.
- (1954), *Probleme der statistischen Methodenlehre in den Sozialwissenschaften*. Würzburg, 1957. Also Würzburg, 1962.
- Anderson Richard L., Bancroft Theodore A.** (1952), *Statistical Theory in Research*. New York.

- Baranow Lotty von** (1950), *Grundbegriffe moderner statistischer Methodik*, Tle 1 – 2. Stuttgart.
- Borel Emil** (1933), *Traité du calcul des probabilités et de ses applications*. Paris.
- Bowley Arthur L.** (1901), *Elements of Statistics*. London, 1948.
- Charlier Carl V. L.** (1911 – 1914), Contributions to the mathematical theory of statistics. *Arkiv for Matematik*, Bde 7 – 9.
- (1920), *Vorlesungen über die Grundzüge der mathematischen Statistik*. Lund, 1931.
- Chuprov A. A.** (1909), *Ocherki po Teorii Statistiki* (Essays on the Theory of Statistics). Moscow, [1959].
- (1918 – 1919), Zur Stabilität statistischer Reihen. *Skand. Aktuarietidskrift*, Bde 1 – 2.
- Cramér Harald** (1946), *Mathematical Methods of Statistics*. Princeton.
- Czuber Emanuel** (1920), *Die statistischen Forschungsmethoden*. Wien, 1938. Reprint 1956.
- Darmois George E.** (1928), *Statistique mathématique*. Paris.
- David Florence N.** (1949), *Probability Theory for Statistical Methods*. London.
- Dixon Wilfrid J., Massey Frank J.** (1951), *Introduction to Statistical Analysis*. New York.
- Doob Joseph L.** (1953), *Stochastic Processes*. New York.
- Fechner Gustav Th.** (1897), *Kollektivmaßlehre*. Editor G. F. Lipps. Leipzig.
- Fisher Ronald A.** (1925), *Statistical Methods for Research Workers*. Edinburgh – London, 1954.
- Fréchet Maurice** (1946 – 1951), *Leçons de statistique mathématique*, No. 1 – 5. Paris.
- Freund John E.** (1952), *Modern Elementary Statistics*. London.
- Gebelein Hans** (1943), *Zahl und Wirklichkeit, Grundzüge einer mathematischen Statistik*. Leipzig. Heidelberg, 1949.
- Gebelein Hans, Heite Hans-Joachim** (1951), *Statistische Urteilsbildung*. Berlin.
- Gini Corrado** (1954), *Corso di statistica*. Roma.
- Goulden Cyril H.** (1939), *Methods of Statistical Analysis*. New York, 1952.
- Hald Anders** (1948), *Statistiske Metoder*. Stockholm. *Statistical Theory with Engineering Application*. New York, 1952.
- Hoel Paul G.** (1954), *Introduction to Mathematical Statistics*. New York – London.
- Kendall Maurice G.** (1943 and 1952), *Advanced Theory of Statistics*, vol. 1; 1946, 1951, vol. 2. London.
- Kenney John E.** (1939), *Mathematics of Statistics*, vols 1 – 2. New York, 1947, 1951.
- Linder Arthur** (1945), *Statistische Methoden für Naturwissenschaftler, Mediziner und Ingenieure*. Basel, 1951.
- Mills Frederick C.** (1924), *Statistical Methods Applied to Economics and Business*. New York, 1955.
- Mood Alexander McFarlene** (1950), *Introduction to the Theory of Statistics*. New York.
- Morice Eugène, Chartier F.** (1954), *Méthode statistique*, tt. 1 – 2. Paris.
- Neyman Jerzy** (1938), *Lectures and Conferences on Mathematical Statistics and Probability*. Washington, 1952.
- (1950), *First Course in Probability and Statistics*. New York.
- Rao C. Radhakrishna** (1952), *Advanced Statistical Methods in Biometric Research*. New York – London.
- Richardson Clarence H.** (1934), *Introduction to Statistical Analysis*. New York, 1944.
- Rietz Henry L.** (1924), *Handbook of Mathematical Statistics*. Boston – New York.
- Romanovsky Vsevolod I.** (1938), *Matematicheskaja Statistika* (Math. Statistics). Moscow – Leningrad.
- Schmetterer Leopold** (1956), *Einführung in die mathematische Statistik*. Wien.
- Snedecor George W.** (1937), *Statistical Methods Applied to Experiments in Agriculture and Biology*. Ames, 1950.
- Tippet Leonard H. C.** (1931), *The Methods of Statistics*. London, 1952.
- Wald Abraham** (1947), *Sequential Analysis*. New York.
- (1950), *Statistical Decision Functions*. New York.
- Weatherburn Charles E.** (1946), *First Course in Mathematical Statistics*. Cambridge, 1949.
- Weber Erna** (1948), *Grundriss der biologischen Statistik für Naturwissenschaftler, Landwirte und Mediziner*. Jena, 1956.

Wold Herman (1938), *Study in the Analysis of Stationary Time Series*. Stockholm, 1954.

Yule G. Udny (1911), *An Introduction to the Theory of Statistics*. Editor M. G. Kendall. London, 1953.

Periodicals

Journal of the Royal Statistical Society, ser. A and B.

Journal of the American Statistical Association.

Allgemeines statistisches Archiv.

Biometrika.

Metron.

Annals of Mathematical Statistics.

Revue de l'Institut International de Statistique.

Sankhya. Indian Journal of Statistics.

Mitteilungsblatt für mathematische Statistik und ihre Anwendungsgebiete.

Statistische Vierteljahresschrift.

XIV
P. P. Permjakov

**From the History of the Combinatorial Analysis.
On the Development of the Method of Generating Functions
until the Mid-19th Century**

Istoria i Metodologia Estestvennykh Nauk, vol. 25, 1980, pp. 121 – 131

The methods of enumeration comprise an essential and earliest part of the combinatorial analysis. Main among them is the method of generating functions¹ and a vast literature is devoted to its description and application, but, nevertheless, its origin and development are not yet sufficiently studied. It is usually held that that method goes back to Euler and Laplace, see for example Berge (1968, p. 6), although some authors of contributions on the history of probability (Seal 1949, p. 209/1977, p. 67) state that De Moivre was the first to apply it.

I am briefly outlying the prehistory, origin and development of the method of generating functions up to the mid-19th century from the standpoint of its applications to solving combinatorial problems in various branches of mathematics. I shall use a shorter, although not altogether precise expression, *method of generating functions in the combinatorial analysis*.

1. The prerequisites for creating the method of generating functions

The origin of that method was determined in a natural way by the accumulation and development of various methods of solving enumerative combinatorial problems which for a long time had been embracing almost the entire contents of combinatorics. Elementary combinatorial propositions are known to have emerged very long ago (Kutlumuratov 1964, Chapter 8; Rybnikov 1972, Chapters 1 and 2); their appearance occurred approximately during the first stages of the forming of theoretical mathematical notions. For a long time combinatorial facts were being collected and applied thus making possible the establishment of the theory of combinatorics. During the 17th century, isolated combinatorial problems solved by particular methods became considered from a general theoretical point of view. The forming of combinatorics had been going on via many paths, and I am studying this process by following a typical example, i. e. by examining problems of counting the number of possible outcomes of a throw of several dice, an appropriate model for many combinatorial problems in probability.

The first known pertinent calculations belong to the 10th – 11th centuries (Maistrov 1967/1974, p. 15)². From that time onward, many authors had been solving isolated problems on the possible number of one or another outcome of a throw of an ordinary die or two or three such dice. Galileo (date of his note is unknown) was apparently the first to provide a complete systematic solution of this problem³. His table (see Table 1)

**Table 1. Galileo (1718, posthumous/1962, p. 194)
Outcomes of throws of three dice. Explanation in text**

1																
3																
6	10		9		8		7		6		5		4		3	
10	631	6	621	6	611	3	511	3	411	3	311	3	211	3	111	1
15	622	3	531	6	521	6	421	6	321	6	221	3				
21	541	6	522	3	431	6	331	3	222	1						
25	532	6	441	3	422	3	322	3								
27	442	3	432	6	332	3										
108	433	3	333	1												
108		27		25		21		15		10		6		3		1
216																

shows the possible number of points (columns) and the corresponding number of ways of the outcomes (lines). As Galileo noted, the second half of the Table (for the outcomes from 11 to 18) was “similar” to the first one.

Later, and independently from him, Huygens in 1657 had published a similar solution reproduced in Bernoulli (1713/1899, pt 1) who offered a better method of dealing with that problem by means of a special table which I adduce here in an abbreviated form (Table 2). He offered an

Table 2. Jakob Bernoulli (1713/1899, p. 27)

- 1) Outcomes of throws of dice. Number of points (columns) and number of dice (lines)
 - 2) Number of ways for various throws (columns) and number of dice (lines)
- Extreme right not reproduced

I.	1	2	3	4	5	6											
II.	2	3	4	5	6	7	8	9	10	11	12						
III.	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
IV.	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
V.	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
VI.	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
I.	1	1	1	1	1	1											
		1	1	1	1	1	1										
			1	1	1	1	1	1									
				1	1	1	1	1	1								
					1	1	1	1	1	1							
						1	1	1	1	1	1						
II.	1	2	3	4	5	6	5	4	3	2	1						
III.	1	3	6	10	15	21	25	27	27	25	21	15	10	6	3	1	
IV.	1	4	10	20	35	56	80	104	125	140	146	140	125	104	80	56	
V.	1	5	15	35	70	126	205	305	420	540	651	735	780	780	735	651	
VI.	1	6	21	56	126	252	456	756	1161	1666	2247	2856	3431	3906	4221	4332	4

example: 4 dice. 1 way to throw 4 (or 24) points, 4 ways to throw 5 (or 23) points etc. He constructed his table inductively proceeding from the case of $(n - 1)$ to n dice, gave some explanation and noted the symmetry of his table. I stress that Bernoulli

1) Applied his table not only as a means for systematizing and keeping the results obtained, as was done by Galileo, but mainly for deriving them,

and in this respect it was similar, for example, to the generally known Pascal triangle.

2) The recurrence method of constructing the table exactly conforms to consecutively calculating the coefficients of the expansion of $(x + x^2 + \dots + x^6)^n$ by issuing from the expansion of $(x + x^2 + \dots + x^6)^{n-1}$ and only a step was left from the idea of a generating function.

Bernoulli however did not take it. His method made unnecessary a direct enumeration of the possible combinations, which the preceding authors had not been able to avoid; nevertheless, it demanded the compilation of a rather awkward table containing a whole mass of results even if only one number was needed.

Further progress could have been achieved by introducing generating functions since everything was already prepared both in the combinatorial and general mathematical sense: methods for solving of combinatorial problems were collected and developed, letter symbols were widely used, methods of dealing with formal power series promoted, the binomial theorem generalized.

2. The first applications of the method of generating functions

To the best of my knowledge, the earliest formulation of the idea of a generating function is due to Leibniz, to the founder of combinatorics as a separate science. In a manuscript of 1676 or 1700 we find the following reasoning (Knobloch 1973, p. 229):

The sought number of permutations of a given kind coincides with the value of the coefficient [of the term] of the same kind of a homogeneous polynomial raised to a [certain] power. If, for example, $x = a + b + c + \dots$, then $x^6 = a^6 + 6a^5b + 15a^4b^2 + \dots$ and 6 and 15 are the numbers of permutations of the kind of a^5b and a^4b^2 .

Montmort (1708/1713, p. 34) independently arrived at the same result:

The coefficients of some polynomial q raised to the power of p [$(a + b + c + \dots)^p$ with q terms – P. P.] coincide with the number of various combinations of some number p of dice having q faces; this is a new and very important theorem which I extensively apply [...] to calculate the number of cases in dice [games] and to the theory of combinations and of polynomials raised to an arbitrary power.

Note that the “new and very important theorem” was indeed its general case of an arbitrary polynomial. Its particular case for the binomial was apparently known much earlier, and not only to Montmort, but, for example, to Arbuthnot (1712) and De Moivre (1712), see Sheynin (1973, p. 276).

In modern language the case dealt with applying the usual generating function of many arguments; Montmort did not go any further, he had not sufficiently grasped the possibilities of that method. He (p. 46) offered a general formula for the number of times p points to appear in a throw of d dice with f faces, obtaining it by a witty but awkward and laborious method. He wrote out all the possible combinations for three dice in a special table and noted the general regularities for that particular example.

Only 20 years later De Moivre (see § 3) provided a much simpler derivation of that formula by means of a generating function of one variable. And Montmort (p. 62) was not at all able to solve a more difficult problem of determining the number of cases for the outcome of p points when choosing three cards out of 30, of three decks containing 10 cards each numbered from 1 to 10, and only offered a method for compiling a list of every possible combination.

The solution of this problem follows as a particular case of the results achieved 30 years later by Euler who applied a generating function of two variables (§ 3). I note finally that the *Leibniz – Montmort* theorem can be applied both “to calculate the number of cases in dice, to the theory of combinations” (that is, as some analytical combinatorial method) and to the theory of “polynomials raised to an arbitrary power” (that is, as a combinatorial method applied to formal analytical calculations).

The first direction corresponds to the development of the method of generating functions. The second culminated by the end of the first quarter of the 19th century in the Hindenburg combinatorial school whose approach is known (Kutlumuratov 1964, Chapter 3; Rybnikov 1972, Chapter 8) to be special in that general formulas were derived enabling to avoid recursions when raising series to a power or inverting them, etc.

When searching for these formulas, Hindenburg applied combinatorial methods; when, for example, multiplying series, the coefficients of the product were derived as combinatorial items formed from the coefficients of the factors. Depending on the kind of series involved, those items occurred to be combinations of various types: with repetitions, having a definite sum, etc. I do not consider this second direction.

3. The further development and application of the method of generating functions in the 18th century

De Moivre (1730, pp. 191 – 197) applied that method to prove the formula for obtaining p points in a throw of d dice with f faces. In modern notation it is

$$\sum_{i=0}^{\alpha} (-1)^i C_d^i C_{p-1+if}^{d-1}, \quad \alpha = \left[\frac{p-d}{f} \right].$$

De Moivre proved that the value of that expression coincided with the coefficient of r^{p-d} in the expansion

$$(1 + r + \dots + r^{f-1})^d = (1 - r^f)^d (1 - r)^{-d}.$$

He then applied the binomial formula to both multipliers obtained. Interesting is his interpretation (p. 192) of the variable r :

Let us imagine a die with only one face marked 1, and as many faces marked 2 as there are units in r ; as many marked 3 as units in r^2 etc.

De Moivre thus thought that r was a natural number although he never allowed for this restriction and considered r as an abstract algebraic symbol.

Following him, Simpson (Seal 1949, pp. 211 – 213/1977, pp. 69 – 71) applied the method of generating functions in the theory of errors for the

discrete case⁴ and made use of the similarity of that problem with studying a given outcome of a throw of a certain number of dice having an appropriate number of faces.

I ought to dwell especially on Euler's *Introduction* (1748/1988, Chapter 16). On 27 August 1740 Philippe Naudé the junior⁵ proposed two problems to Euler:

- 1) *To find the number of different ways in which a given number can be obtained by summing a given number of unequal integral numbers.*
- 2) *To find the number of different ways in which a given number m can be partitioned in μ equal or unequal parts or integral numbers.*

Euler (Bashmakova et al 1972, p. 108) gave an analytical solution in his reply of September of same year and his method of solution was based on applying generating functions. Namely, the number of ways for representing a number n as a sum of m different terms of the sequence $\alpha, \beta, \gamma, \delta, \dots$ was sought as the coefficient of $x^n z^m$ of the expansion of

$$(1 + x^\alpha z) (1 + x^\beta z) (1 + x^\gamma z) (1 + x^\delta z) \dots$$

The sequence can be assumed to be that of natural numbers. Similarly, the number of ways in which n can be represented by m equal or unequal parts was the coefficient of the same product in the expansion of

$$(1 - x^\alpha z)^{-1} (1 - x^\beta z)^{-1} (1 - x^\gamma z)^{-1} (1 - x^\delta z)^{-1} \dots$$

If, however, the number of parts was not specified, the sought number of ways was equal to the coefficient of x^n in the appropriate expansion where, as a preliminary, $z = 1$ was assumed. Euler's results easily lead to the solution of the Montmort problem (§ 2): the sought number of ways is equal to the coefficient of $x^n z^3$ in the expansion of

$$[(1 + xz) (1 + x^2z) \dots (1 + x^{10}z)]^3.$$

Euler applied his method for proving a series of theorems on partitioning of numbers into summands and compiled appropriate tables showing the number of the partitions. He also indicated the connection between the numbers in his tables and figurate numbers, but he did not comment on the essence of the variables x and z . Indirectly, he assumed them to be formal symbols with defined formal operations on them (addition, multiplication etc) possessing natural algebraic properties. Issues concerning the convergence of series were here of no consequence.

And so, Euler apparently arrived at the method of generating functions independently from his predecessors, essentially developed it, and systematically applied it for solving rather difficult and previously unyielding problems.

Lagrange (1776) applied the method of generating functions to the theory of errors and made use of such functions not only, like Simpson, of one variable, but also, in Euler's spirit, of two variables as well. With the exception of the Newton binomial theorem, he had not referred to his predecessors but he very often wrote *as is known*. Like Simpson, he made

use of the similarity with throwing dice, and the style of his reasoning is essentially Eulerian. The appropriate contributions of Euler and Simpson were then already published, and we can be almost sure that he was acquainted with them. Apparently novel (1776/1868, p. 209) was the application of a formal differentiation for determining the coefficients of expansions. I note that Euler (1788) commented on the Lagrange contribution but did not say anything about the connection of that memoir with his own previous studies (see above)⁶.

During the same period other mathematicians had also applied the method of generating functions. Thus, Lambert (1771) made use of generating functions

$$\sum_{n=1}^{\infty} \frac{x^n}{1-x^n}, \quad \sum_{n=1}^{\infty} \frac{nx^n}{1-x^n}$$

for determining the number of divisors of a given n and their sum respectively.

From this and the preceding section we may infer that

1) The method of generating functions had originated long before Laplace (1782) introduced the term itself, and essentially earlier than the appearance of the works of Euler and even De Moivre (1712).

2) During the 18th century, that method had been applied rather often and was developed to a comparatively high level.

4. The Laplacean theory of generating functions

Laplace first formulated his theory in his memoir (1782). He (1782/1894, p. 1) began it by stating that the theory of series was a most essential subject of analysis since all the problems which are reduced to approximations and therefore almost all applications of mathematics to nature depend on it. He proposed a method for dealing with series

Based on considering that which I call generating functions. It is a new kind of calculus which might be called calculus of generating functions.

Laplace went on to indicate that it was possible to apply that calculus for interpolating series, integrating difference and differential equations as well as for the “analysis of chances”. And here is his definition (Ibidem, pp. 5 – 6):

Let y_x be some function of x ; form an infinite series

$$y_0 + y_1 t + y_2 t^2 + \dots + y_x t^x + y_{x+1} t^{x+1} + \dots + y_{\infty} t^{\infty}$$

and designate its sum, or what is reduced to the same, the function whose expansion forms that series, by u . Then that function will be what I call the generating function of the variable y_x . Thus, a generating function of some variable y_x is some function of t , and, after being expanded into a power series of t , has this variable y_x as the coefficient of t^x . Similarly, the variable corresponding to some generating function is the coefficient of t^x in the expansion of that function in powers of t .

He defined the generating function of two or more arguments, that is, for sequences with many indices, in a similar way and considered the transformations of that function when the variable y_x was being transformed. I adduce his results in a tabular form.

**Table 3. Laplace (1782/1894, pp. 5 – 6)
Dictionary of the generating function for different variables**

Variable	Generating function
y_x	u
y'_x	u'
y''_x	u''
$y'_x + y''_x$	$u' + u''$
$\Delta y_x = y_{x+1} - y_x$	$u \left(\frac{1}{t} - 1 \right)$
$\nabla y_x = ay_x + by_{x+1} + \dots + qy_{x+n}$	$u \left(a + \frac{b}{t} + \dots + \frac{q}{t^n} \right)$
y_{x-r}	ut^r
$\Delta^i \nabla^s y_{x-r}$	$ut^r \left(a + \frac{b}{t} + \dots + \frac{q}{t^n} \right)^s \left(\frac{1}{t} - 1 \right)^i$

It is not difficult to convince ourselves in that the variables in the left column are the coefficients of t^x in the corresponding expressions of the right column.

Laplace stated that those relations were also valid for fractional and even irrational values of i , s and r . He offered no explanation, but apparently for non-integral i , s , r the last expression in the left column should be understood as the corresponding term in the development, if this is possible, of the right side. He mostly devoted his memoir to the theory of series.

In his lecture *On probability* Laplace (1795/1912, pp.154 and 157 in 1912) explained how he understood the subject of the theory of generating functions and formulated a new and wider interpretation of the term *calculus of generating functions*:

The subject of this theory is the correspondence between the coefficients of the powers of some indefinite variable in the expansion of some function of that variable and the function itself.

The theory of generating functions and the theory of approximation for the formulas which are functions of very large numbers⁷ may be considered as two branches of a single calculus which I call the calculus of generating functions.

Then, Laplace (1809/1912, p. 178), in a section of that memoir called *On the calculus of generating functions*, wrote:

The subject of that calculus is the reduction of all the operations touching differences, and especially integration of equations in finite and

partial differences, to simple expansions of functions. This is the main idea.

Later, Laplace (1811/1898, p. 360), after repeating his earlier pronouncements, dwelt especially on the application of generating functions to the theory of probability:

The calculus of generating functions is the basis of some theory of probability which I intend to publish soon. Problems concerning random events are most often easily reduced to difference equations, and the first branch of that calculus provides their most general and simplest solutions.

Laplace included all his studies of generating functions in his *Théorie analytique des probabilités* (1812); see Sheynin (1973) for considering to what extent had the generating functions become the basis of his theory of probability.

From the viewpoint of the general combinatorial theory, the significance of the Laplacean theory of generating functions consists in that he was the first to formulate explicitly the concept of a generating function, considered some of its general properties, introduced the appropriate terminology and a symbolic machinery. His work concluded the first stage of the development of the method of generating functions, but

1) He did not aim at developing that method as a means for solving combinatorial problems. He considered his theory almost as a basis for the entire mathematical analysis (and, later, for the theory of probability) and only assigned an illustrative role to its combinatorial applications.

2) As formulated by him, the theory of generating functions was very vulnerable: in mathematical analysis, the variable t could not certainly be anymore considered as a purely abstract symbol. The convergence of the appropriate series ought to be studied which Laplace had not done.

Wronski (1819), Laplace's junior contemporary, devoted an entire book to criticizing the Laplacean theory of generating functions. He correctly, for those times, indicated that the knowledge of series was insufficient and that the case of negative, fractional and irrational values of x was still vaguer and unjustified, see the Laplacean definition of the generating function. However, as Wronski argued, a theory claiming to form the basis of the entire analysis ought to allow maximal generality. As to specific discrete combinatorial problems solved by means of generating functions, Wronski considered them too unimportant for analysis and mathematics in general.

5. On the method of generating functions in the combinatorial analysis in the beginning and mid-19th century

The algebra of usual generating functions of one variable is named after Cauchy (Bell 1923; Riordan 1958; Rybnikov 1972). These authors apparently bear in mind his theorems on the convergence of sums and products of series (Cauchy 1821/1897, p. 140):

Theorem 3. Suppose that at some value of x two series⁸

$$a_0, a_1x, a_2x^2, \dots, a_nx^n, \dots; b_0, b_1x, b_2x^2, \dots, b_nx^n, \dots$$

converge and that their sums are S_1 and S_2 . Under these conditions

$$a_0 + b_0, (a_1 + b_1)x, (a_2 + b_2)x^2, \dots, (a_n + b_n)x^n, \dots$$

will be a new convergent series and its sum will be $(S_1 + S_2)$.

Theorem 4. Let the conditions of the previous theorem be valid.[...] Then

$$a_0b_0, (a_0b_1 + a_1b_0)x, (a_0b_2 + a_1b_1 + a_2b_0)x^2, \dots, \\ (a_0b_n + a_1b_{n-1} + \dots + a_nb_0)x^n, \dots$$

will be a new convergent series and its sum will be S_1S_2 .

I ought to add that these theorems are obviously related to the issue of isomorphism between the algebra of sequences and of their generating functions; Cauchy, however, did not construct any algebra of generating functions, and, although he applied them to a certain extent, did not even make use of the appropriate term. For example (Ibidem, p. 434):

The coefficients of consecutive powers of x in the expansions of the expressions

$$(1 + x)^{-2}, (1 + x)^{-3}, (1 + x)^{-4}, \dots$$

are called figurate numbers du premier, du second, du troisième ordre etc.

In other words, these expressions are generating functions of the figurate numbers mentioned.

Application of the method of generating functions is also seen in the writings of other mathematicians of the 19th century. Thus, Abel (?/1839, p. 224)⁹ referred to Euler and effectively wrote out the exponential generating function for the Bernoulli numbers:

If we expand the function $[1 - (u/2) \cot(u/2)]$ into a series of integral powers of u , and set

$$1 - \frac{u}{2} \cot \frac{u}{2} = A_1 \frac{u^2}{2} + A_2 \frac{u^4}{2 \cdot 3 \cdot 4} + \dots + A_n \frac{u^{2n}}{2 \cdot 3 \cdot 4 \cdot \dots \cdot 2n} + \dots,$$

the coefficients A_1, A_2, A_3, \dots , as is known, will be the Bernoulli numbers.

And in one of his posthumous memoirs Abel (?/1839, p. 77) mentioned the term *generating function* although in respect to other matters:

The generating function φ is connected with the defining function f by the integral transformation

$$\varphi(x; y; z; \dots) = \int e^{xu + yv + zp + \dots} f(u; v; p; \dots) dudvdp \dots$$

Legendre (1830) briefly expounded Euler's results from the *Introductio* concerning the partition of numbers into summands derived by means of generating functions. He did not apply the term itself.

Buniakovsky (1846) systematically applied generating functions but not the term itself, and only mentioned their Laplacean theory (and his contribution to the theory of probability in general) in a historical essay at the very end of his book (p. 273) and abstained from any comment.

The further development of the method of generating functions in combinatorial analysis began in the mid-19th century. It was essentially stimulated by the application of the methods of the theory of decomposition to enumerating invariants and covariants in the theory of algebraic forms, especially in the works of Cayley and Sylvester.

I am only considering the very beginning of that process by following Cayley. His works are very indicative in this respect and in addition they strongly influenced the further development of the theory of generating functions in combinatorial analysis. In particular, we find the following reasoning in his work (ca. 1856/1889, p. 260):

The number of terms of the degree θ and of the weight q is obviously equal to the number of ways in which q can be made up as a sum of θ terms with the elements $(0, 1, 2, \dots, m)$, a number which is equal to the coefficient of $x^q z^\theta$ in the development of

$$[(1 - z)(1 - xz)(1 - x^2z) \dots (1 - x^mz)]^{-1};$$

and the number of the asyzygetic [independent] covariants of any particular degree for the quantic $(x; y)^m$ can therefore be determined by means of this development. In the case of a cubic, for example, the function to be developed is [equal to the previous fraction with $m = 3$], which is equal to

$$1 + z(1 + x + x^2 + x^3) + z^2(1 + x + 2x^2 + 2x^3 + 2x^4 + 2x^5 + x^6) + \dots$$

His other contribution of about the same year (ca. 1855/1889, p. 235) was specifically devoted to the partition of numbers:

I propose to discuss the following problem: To find in how many ways a number q can be made up of the elements a, b, c, \dots each element being repeatable an indefinite number of times. The required number of partitions is represented by the notation

$$P(a; b; c; \dots)q,$$

and we have, as is well known,

$$P(a; b; c; \dots)q = \text{coefficient } x^q \text{ in } [(1 - x^a)(1 - x^b)(1 - x^c) \dots]^{-1}$$

where the expansion is to be effected in ascending powers of x .

Cayley then expounds his method of solution which included the decomposition into partial fractions.

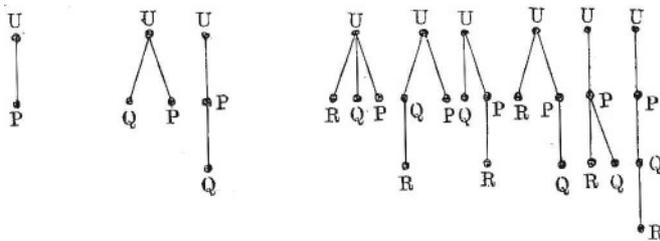
A little later he also applied the method of generating functions to enumerative problems of the not yet existing graph theory, namely for enumerating trees which is similar to enumerating decompositions. I consider the first of his relevant writings (1857/1890). He introduced trees for carrying out analytic calculations. Suppose that there are operators

P, Q, R, \dots of the type of $A\partial x + B\partial y + \dots$

where A, B, \dots are functions of x, y, \dots , i. e. operands (*operandators*, as Cayley called them) and the operators are also operands. Let U be a usual operand; it is easy to show that

$$QPU = (QxP)U + (QP)U.$$

However, it is rather difficult to derive a similar expression for $RQPU$. Cayley indicated the action of the operands by marked trees although he did not formally define tree, root, branches, etc, he thought that their meaning is obvious given the figures appended.



The problem was thus reduced to constructing the trees, but for being sure that all of them are constructed, a method for determining their number beforehand is needed.

Let A_n be their number having n branches. Cayley compiled an equation for the generating function

$$(1-x)^{-1}(1-x^2)^{-A_1}(1-x^3)^{-A_2}(1-x^4)^{-A_3} \dots = 1 + A_1x + A_2x^2 + A_3x^3 + A_4x^4 + \dots$$

and derived A_n for $n = 1, 2, \dots, 10$. And, in the same way, he found the number of trees B_r with r free branches from the equation for the generating function

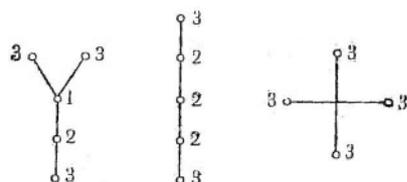
$$(1-x)^{-1}(1-x^2)^{-B_2}(1-x^3)^{-B_3}(1-x^4)^{-B_4} \dots = 1 + x + 2B_2x^2 + 2B_3x^3 + 2B_4x^4 + \dots$$

Finally, Cayley indicated that Sylvester's studies in differential calculus touching on substitution of independent variables had prompted him to consider the explicated theory of trees. Note that for a long time he had not been using the term *generating function*. Later he (Cayley 1873 – 1874/1896, p. 188) did use it and wrote out the (exponential) generating function

$$u = u_1x + u_2 \frac{x^2}{2} + \dots + u_n \frac{x^n}{1 \cdot 2 \cdot \dots \cdot n} + \dots = \frac{e^{x/2+x^2/4}}{\sqrt{1-x}}.$$

Caley also discovered ever more applications for his theory of trees. Here is a small extract (Cayley ca. 1875/1896, p. 427):

As regards the paraffins C_nH_{2n+2} , we have n atoms of carbon connected by $n - 1$ bands, under the restriction that from each carbon-atom there proceed at most 4 bands (or [...] we have n knots connected by $n - 1$ branches), in the form of a tree; for instance, $n = 5$, such forms (and only such forms) are



The numbers are those of the hydrogen atoms connected with the appropriate atom of carbon. In the same paper Cayley applied the method of generating functions and the term itself denoting it by GF.

And so, we see that the further development of the method of generating functions as a means for solving combinatorial problems was called forth by the appearance of new fields of application. And in this connection we ought to stress the role of the graph theory whose enumerative problems have been a wide area for such applications and, at the same time, a powerful stimulus for the further development of the theory of generating function. It is opportune to note that the classical Polya theorem of that theory of generating functions is also a classical theorem of the graph theory (and was formulated for solving its problems).

Notes

1. The usual generating function of a sequence $a_0, a_1, \dots, a_n, \dots$ is the formal sum

$$A(t) = a_0 + a_1t + \dots + a_n t^n + \dots$$

An exponential generating function of the same sequence (Riordan 1958, p. 19) is the formal sum

$$E(t) = a_0 + a_1t + \dots + a_n t^n/n! + \dots$$

From the very outset of their writings, many authors consider t a real or complex variable and specify that the appropriate series are convergent. P. P.

2. Maistrov did not substantiate this statement.

3. English translation of his note: Galilei (1718, posthumous/1962). I copied Table 1 from that source rather than from Maistrov (as the author did).

4. The theory of errors did not exist then (and neither in 1776, see below the author's discussion of the work of Lagrange). Next year (in 1757) Simpson in a general analytical way considered the limiting continuous case.

5. The author quoted from F. Rudio's comment inserted on p. 310 of the Russian edition of 1961 of Euler's *Introductio*.

6. Lagrange's memoir contained findings of general mathematical interest. He was the first to use integral transformations, derived (in Problem 6) the equation of the multivariate normal distribution, introduced the term *courbe de la facilité des erreurs* and anticipated characteristic functions by considering a "generating integral" (Seal 1949, p. 214/1977, p. 72; Freudenthal & Steiner 1966, p. 170; Sheynin 1973, § 2).

7. These were functions of the type

$$\int x^p(1-x)^q dx$$

where p and q were large numbers (e. g., the number of boys and girls born in France during many years) and the limits of integration were defined by the Bayesian approach to problems in probability.

8. As written here and below, the “series” are actually sequences.

9. In many cases, here also, the author did not provide the date of the original publication. Abel’s *Oeuvres Complètes* (1839) were reprinted in 1881, perhaps with some change, and I was unable to find the appropriate place there. The same happened in the second instance (see below).

Bibliography

- Abel N. H.** (1839), *Oeuvres Complètes*, t. 2. Christiania, 1881.
- Arbuthnot J.** (1712), An argument for divine Providence taken from the constant regularity observed in the birth of both sexes. Reprint: Kendall M. G., Plackett R. L. (1977, pp. 30 – 34).
- Bashmakova I. G., Ozigova E. P., Youshkevich A. P.** (1972), Theory of numbers. In *Istoria Matematiki s Drevneishikh Vremen do Nachala XIX Stoletia* (History of Mathematics from the Most Ancient Times to the Beginning of the 19th Century), vol. 3. Editor, Youshkevich. Moscow, pp. 101 – 125.
- Bell E. T.** (1923), Euler algebra. *Trans. Amer. Math. Soc.*, vol. 25, pp. 135 – 154.
- Berge C.** (1968), *Principes de combinatoire*. Paris.
- Bernoulli J.** (1713), *Ars conjectandi. Werke*, Bd. 3. Basel, 1975, pp. 107 – 259. German transl. (1899): *Wahrscheinlichkeitsrechnung*. Frankfurt/Main, 1999.
- Buniakovsky V. Ya.** (1846), *Osnovania Matematicheskoi Teorii Veroiatnostei* (Principles of the Mathematical Theory of Probability). Petersburg.
- Cauchy A. L.** (1821), Course d’analyse de l’Ecole Royale Polytechnique, pt. 1. *Oeuvr. Compl.*, sér. 2, t. 3. Paris, 1897.
- Cayley A.** (ca. 1855), Researches on the partition of numbers. In author’s book (1889, pp. 235 – 249).
 --- (ca. 1856), Second memoir upon the quantics. *Ibidem*, pp. 250 – 275.
 --- (1857), On the theory of the analytical forms called trees. In author’s book (1890, pp. 242 – 246).
 --- (1873 – 1874), On the number of distinct terms in a symmetrical or partially symmetrical determinant. In author’s book (1896, pp. 185 – 190).
 --- (ca. 1875), On the analytical forms called trees, with application to the theory of chemical combinations. *Ibidem*, pp. 427 – 460.
 --- (1889, 1890, 1896), *Collected Mathematical Papers*, vols 2, 3, 9. Reprinted: New York, 1963.
- De Moivre A.** (1712, in Latin), De mensura sortis, or, On the measurement of chance. *Intern. Stat. Rev.*, vol. 52, 1984, pp. 236 – 262. Comment (A. Hald): pp. 229 – 236.
 --- (1730), *Miscellanea analytica*. London.
- Euler L.** (1748, Latin), *Introduction to Analysis of the Infinite*, Book 1. New York, 1988.
 --- (written 1777; 1788), Eclaircissements sur le mémoire de La Grange [...]. *Opera Omnia*, ser. 1, t. 7. Leipzig – Berlin, 1923, pp. 425 – 434.
- Freudenthal H, Steiner H.-G.** (1966), Die Anfänge der Wahrscheinlichkeitsrechnung. In *Grundzüge der Mathematik*, Bd. 4. Editors, H. Behnke et al. Göttingen, pp. 149 – 195.
- Galilei G.** (1718, posthumous, in Italian), Sopra el scoperte dei dadi. English transl. (without English title) by E. H. Thorne in David F. N. (1962), *Games, Gods and Gambling*. London, pp. 192 – 195.
- Kendall M. G., Plackett R. L.** (1977), *Studies in the History of Statistics and Probability*, vol. 2. London.
- Knobloch E.** (1973), *Die mathematischen Studien von G. W. Leibniz zur Kombinatorik*. Wiesbaden.
- Kutlumuratov G.** (1964), *O Razvitií Kombinatornykh Metodov Matematiki* (On the Development of Mathematical Combinatorial Methods). Nukus.
- Lagrange J. L.** (1776), Sur l’utilité de la méthode de prendre le milieu entre les resultats de plusieurs observations. *Oeuvres*, t. 2. Paris, 1868, pp. 173 – 236.
- Lambert J.-H.** (1771), *Anlage zur Architektonik*, Bd. 2. *Phil. Schriften*, Bd. 4. Hildesheim, 1965.

- Laplace P. S.** (1782), Sur les suites. *Oeuvr. Compl.*, t. 10. Paris, 1894, pp. 1 – 89.
 --- (read 1795; 1800), Leçons de mathématiques professes à l'Ecole Normale en 1795. *Oeuvr. Compl.*, t. 14. Paris, 1912, pp. 10 – 177. The lecture on probability is on pp. 146 – 177.
 --- (1809), Sur divers points d'analyse. *Ibidem*, pp. 178 – 214.
 --- (1811), Sur les intégrals définies. *Oeuvr. Compl.*, t. 12. Paris, 1898, pp. 357 – 412.
 --- (1812), *Théorie analytique des probabilités*. *Oeuvr. Compl.*, t. 7. Paris, 1884.
- Legendre A. M.** (1830), *Théorie des nombres*, t. 2. Paris.
- Maistrov L. E.** (1967, in Russian), *Probability Theory. A Historical Sketch*. New York – London, 1974.
- Montmort P. R.** (1708), *Essay d'analyse sur les jeux de hazard*. Paris, 1713. Reprint: New York, 1980.
- Riordan J.** (1958), *Introduction to Combinatorial Analysis*. New York.
- Rybnikov K. A.** (1972), *Vvedenie v Kombinatorny Analis* (Introduction to Combinatorial Analysis). Moscow.
- Seal H. L.** (1949), The historical development of the use of generating functions in probability theory. *Bull de l'Association des Actuairees Suisses*, t. 49, pp. 209 – 228. Reprint: Kendall & Plackett (1977, pp. 67 – 86).
- Sheynin O. B.** (1973), Finite random sums. A historical essay. *Arch. Hist. Ex. Sci.*, vol. 9, pp. 275 – 305.
- Wronski H.** (1819), *Critique de la théorie des fonctions génératrices de Laplace*. Paris.

XV
Bierman K.-R.

**Problems of the Genoese Lotto
in the Works of Classics of the Theory of Probability**

Istoriko-Matematicheskie Issledovania, vol. 10, 1957, pp. 649 – 670

1. Among problems occurring in games of chance, such as games of dice, heads or tails, cards, and considered in fundamental works of the classics of the theory of probability (see e. g., Biermann 1955), those concerning numerical lotteries had very soon occupied a certain place. The game of chance described here had first been called *lotto* whereas *lottery* meant the drawing of lots for a determined number of winners among an established number of participating tickets. Gradually, however, it became usual also to call a numerical lotto a numerical lottery, and I do not distinguish these two terms.

The exact date of the origin of the numerical lotto is not known; it is only certain that even before 1620, which is believed to be the year when that game had first occurred (Cantor 1892/1900, p. 336; Wahschauer 1885, p. 7)¹, bets were being made in Italy, *the cradle of the lotto*, during elections and other events attracting the attention of the public.

Lotto was first mentioned in a document on February 9, 1448 (Weidlich 1922)². In the second half of the 16th century the arrangement of betting in Genoa (sponsiones Genuenses) became a business of sorts. The occasion for them was the yearly elections by lot of five out of the hundred senators to the Serenissimo Collegio (The Highest Board)³. The gambler named the senators who, in his opinion, will be thus elected, and, if one or several of them were indeed chosen, he got a firmly established sum, a multiple of his stake; otherwise, he lost his money.

In 1620 the 100 names of senators were replaced by numbers from 1 to 90, and that game became known as the *Genoese lotto*. At first, the rich merchants (mercatores opulenti) had been carrying out the lotto; soon, however, the state had perceived that source of revenue and took over the business. After 1620 the numerical lotto had spread extremely widely over Italy; for details see Weidlich (1922), and after some time it also became rooted beyond that country⁴.

On August 18, 1751, a concession of the first numerical lotto was granted in Vienna (Sieghart 1898, p. 11). In 1757 that lotto was introduced in Paris, on February 8, 1763, in Berlin, and by 1771 there already were 26 numerical lottos in Germany (Warschauer 1885, p. 12).

At the same time, opposition against that game had been strengthening. Thus (Rodotà 1769, p. 27), the origin of its “already unquenchable fire” was attributed to the pernicious talent of a certain arithmetician whose name and nationality remained unknown. In many localities life became centred around the lotto. Thus, in 1795 (Warschauer 1912, p. 59ff) the game rush in Luckenwalde⁵ was so intensive, that the populace did nothing except interpreting and discussing the lotto numbers. Drawing on an official source, that author reports that wives had been getting into debt

behind their husbands' back so as to buy lotto tickets whereas men spent money and excused themselves by stating that the Holy Spirit had appeared to them at night and wrote the next winning numbers on the wall.

A special weekly called *Lottologie* had been put out in Hamburg in 1770 and 1771 (cf. Endemann 1899, p. 75).

Pope Benedict XIII, 1649 – 1730, forbade Catholics from participating in lottos, but his successor, Pope Clement XII, 1652 – 1740, established his own numerical lottery. In spite of the spread of the game rush, common people understood the essence of lotteries. In Italy, it was the talk of the town (Sieghart 1898, p. 5) that the devil carried off a city counsellor Benedetto Gentile from Genoi for having invented the lotto.

In Berlin, a song was on everyone's lips (Warschauer 1885, p. 61):

*Nature provided the East with the plague,
Never is it unjust,
So it provided numerical lotto
To the West.*

In 1810, in Preussia, the numerical lotto was closed down after 799 drawings (Warschauer 1912, p. 52), and in 1861 the last German numerical lotto in Bavaria was abolished. The same occurred in Hungary in 1897⁶, but it continued to exist in Italy and Austria. Under new conditions, numerical lottos have recently been established in Germany anew, two of them in Berlin and one, very popular, in the German Democratic Republic. There, about 15 mln indications of future winning numbers are made weekly.

2. It is necessary to distinguish between three different types of conditions of the game as described below. Note, however, that my mathematical study only concerned the first two of them, whereas the third one is mentioned for the sake of comprehensiveness the more so since it is being applied until now.

A. According to the initial conditions, a gambler bets on five chosen by him candidates or numbers. His winnings depend on whether his choice coincides with the drawings on one or several (not more than five) occasions and on the established beforehand multiplier of the stake which is the greater the more coincidences he scores. Nothing and nowhere was it stated about restricting the number of indications made by a single gambler. Such bets were common during the first period of the Italian lottery.

B. Later, the following rules, roughly speaking, had been established everywhere. A gambler may choose one or several of the following versions of indications. Participation in a certain version did not prevent him from taking part in other ones.

B1. He states that his chosen number will be included in the five to be drawn. This version of the game was called, in Italian, *estrado semplice* (simple extraction; French: *extrait simple*; German: *einfacher Auszug*, or *unbestimmter Auszug* or *Ruf*).

B2. The conditions are as above, but the gambler also names the place of his chosen number among the five to be drawn. Suppose he chooses 3 and the fifth place, and numbers 4, 28, 15, 88 and 63 are drawn, – he wins.

This version was called *estrado determinato, extrait déterminé, bestimmter Auszug* or *Ruf*, respectively, and also *Nominate*.

B3. The gambler chooses two numbers. In Latin, this version is called *ambo* (both); in France, there also existed an *ambo déterminés*, i. e., a fixed *ambo*.

B4, B5, B6. Three, four or five numbers were chosen (*terno, quaterno, quinterno*). Not each country allowed the last two mentioned versions, but the least possible stake was established everywhere just as in some places the maximal stake. The winnings had been fixed, roughly speaking, as follows

Versions of participation	Times the stake
Estrado semplice	14 – 15
Estrado determinato	70 – 75
Ambo	240 – 270
Terno	4,800 – 5,500
Quaterno	60,000 – 75,000
Quinterno	1,000,000

I have not included either the extremely low winnings, as those, for example, in the Papal State, or the extremely high winnings in Sardinia⁷.

C. In the abovementioned contemporary German numerical lotteries the gambler indicates the five numbers to be drawn out of 90 and is free to repeat his indications without any limit. So far, the condition is the same as it was in the original lottery. Other conditions are, however, different in principle. The stake for an indication is now not higher than 1/2 of the German mark; there is no established rate of taxation so that the winnings depend on the number of indications. They are distributed according to four classes from the first to the fourth: coincidences of all five numbers; of four, three, or two numbers and there is no risk for the business. Keeping to the boundaries of my subject, I am not discussing the distinction existing now between the lotteries in Germany and I do not provide any arguments in favour of them or against the rules which had existed in the past.

3. Now I go over to the studies concerning the Genoise lotto made by the classics of the theory of probability. In the first place, the work of Niklaus I Bernoulli, 1687 – 1759 (1709)⁸ should be named since he touches there, in particular, on the “extremely celebrated Genoise lotto”. He begins by stating the conditions A: the gambler indicates five from the 100 candidates and deposits his stake. After the drawing, if one or more of his chosen senators is/are elected, he receives his winnings from the *rich merchants* and he wins the more the more coincidences occurs.

Niklaus determines the just winnings in the following way. Suppose that the stake is a ($= 1$, as he assumes), $g_i, i = 1, 2, \dots, 5$ are the just winnings in case of one, two, ..., five coincidences, e_0 is the number of cases with no coincidences, and $e_i, i = 1, 2, \dots, 5$ are the number of cases in which there are 1, 2, ..., 5 coincidences. Thus, there are e_5 cases in which someone can win g_5 , e_4 cases in which someone can win g_4 , etc. and e_0 cases in which there is no winner at all. The [just] expected winning is therefore

$$\frac{e_5 g_5 + e_4 g_4 + \dots + e_1 g_1 + e_0 \cdot 0}{e_5 + e_4 + \dots + e_1 + e_0} = a. \quad (1)$$

Bernoulli then sets

$$e_5 + e_4 + \dots + e_1 + e_0 = e. \quad (2)$$

The values of e_5 etc, as he indicates, are calculated “according to the known rules of combinations” but he does not directly indicate how. He knew those rules since his uncle had provided them (1713). Indeed, in pt. 2⁹ we find the solution of this problem: Combinations are formed from n elements, r at a time, without repetition. Among the elements m are somehow marked and required is the number of combinations containing p and only p elements out of m . His answer was $C_m^p C_{n-m}^{r-p}$. In our case, $n = 100$, $m = 5$, $r = 5$ and $p = 0, 1, 2, \dots, 5$. Therefore,

$$\begin{aligned} p = 5: e_5 &= C_5^5 C_{95}^0 = 1, & p = 4: e_4 &= C_5^4 C_{95}^1 = 475, \\ p = 3: e_3 &= C_5^3 C_{95}^2 = 44,650, & p = 2: e_2 &= C_5^2 C_{95}^3 = 1,384,150, \\ p = 1: e_1 &= C_5^1 C_{95}^4 = 15,917,725, & p = 0: e_0 &= C_5^0 C_{95}^5 = 57,940,519. \end{aligned}$$

And, according to equality (2), $e = 75,287,520$. I check: $e = C_{100}^5$. Formula (1) provides

$$e_5 g_5 + e_4 g_4 + \dots + e_1 g_1 + e_0 \cdot 0 = e_5 a + e_4 a + \dots + e_1 a + e_0 a$$

and, since $a = 1$, equality (2) leads to the left side being equal to e .

The winning ought to be proportional to the number of coincidences achieved, so that g_5 is the maximal, and g_1 , the minimal winning. If g_5 is unknown, the other g_i can be expressed in the following way:

$$e_5 g_5 = e_4 g_4 = \dots = e_1 g_1; 5e_5 g_5 = e, \quad g_4 = \frac{e_5 g_5}{e_4} \text{ etc.} \quad (3)$$

Therefore

$$g_5 = \frac{e}{5e_5} = \frac{75,287,520}{5 \cdot 1} = 15,057,504 = e_5 g_5 \text{ since } e_5 = 1.$$

[The author explains in detail the calculation of g_4 , g_3 , g_2 and g_1 getting

$g_4 = 31,7004/475$, $g_3 = 3375,227/22,325$ “rather than 6,227 [in the fractional part] as N. Bernoulli”, $g_2 = 10608,002/692,075$, $g_1 = 15,057,504/15,917,725$ of the stake.]

Niklaus then applies his results to show how the merchants carrying out the Genoese lottery had been duping the public: as a rule, they paid out 10,000, 1,500, 300, 10 and 1 gold coin(s) for five, four, ..., one

coincidence respectively¹⁰. He substituted the derived magnitudes e_5, e_4 , etc and g_5, g_4 , etc as established by the merchants into equation (1) and got [...] 2,925,115/5,019,168 of the unit stake, i. e., instead of the gold coin. The merchants had thus been misappropriating [the difference].

Laplace (1814/1995), 1749 – 1827, another classic of the theory of probability, also calculated the just winnings in his celebrated, and, as we call it nowadays, popular scientific exposition of the principles and results of the theory. He assumed a lottery with 90 numbers carried out under conditions B and concluded that the just winnings were

18, 400.5, 11,478, 511,038 and 43,949,268

times the stake for the versions B1, B3, B4 and B5 respectively.

His explanation was very concise; as throughout the *Essai*, he avoided formulas which never simplifies the exposition and does not invariably assist its understanding, so that I describe his considerations (1814/1995, pp. 15 – 16) in a mathematical language.

For $r = 1, 2, \dots, 5$ there can be

$C_{90}^r = 90, 4,005, 117,480, 2,555,190, 43,949,268$

combinations for the number of estrado simplice, ambos, ternos, quaternos and quinternos respectively¹¹. Out of the five numbers drawn there are, respectively,

$C_5^r = 5, 10, 10, 5, 1$ combinations.

Calculating the ratio of the favourable cases to all the possible ones, we get, according to the Laplace definition of probability, the following probabilities of the just winnings, and, at the same time, the just winnings themselves [for a unit stake]:

1/18, 1/400.5, 11/11,748, 1/511, 038, 1/43,949,268.

Laplace (1814/1995, p. 92 and 93) returned to the numerical lottery and the disadvantages of the gamblers' ratio of their chances to those of the lottery holders¹². There also he mentions the widespread delusions:

When one number has not been drawn for a long time in the French lottery, the mob is eager to bet on it. [...] Under an illusion contrary to the preceding ones, one may look in previous draws of the French lottery for the numbers that have most often been drawn to form combinations on which one believes one's stake may advantageously be placed. [...] I have analysed several of these drawings and I have constantly found that they fall within limits about which, under the supposition that all numbers are equally likely to be drawn, there can be no doubt.

Another series of problems concerning the numerical lotto is connected with determining the probability of two or more consecutive numbers (of a sequence) in the same drawing. For example, 17, 3, 76, 35, 16 include a

sequence of two numbers, 16 and 17; or, if 61, 62 and 63 are among the five numbers, they form a sequence of three numbers etc.

Euler, 1707 – 1783, investigated that problem, very complicated, as he (1767) himself remarked, because it involved greatest difficulties. By that time, only two years had passed since the number lottery was established in Berlin, but its conditions were generally known so that he did not describe them.

Euler began by calculating with $r = 2$, went on until $r = 6$, provided the general solution and applied it for the case of $r = 7$. He also gave a numerical answer for $r = 5$, i. e., for the numerical lottery itself. Here are the general solutions inductively derived by Euler and the pertinent numerical values for the lotto.

The probability that no sequence will occur among the r numbers of a drawing is

$$P_s = \frac{(n-r)(n-r-1)(n-r-2)\dots(n-2r+2)}{n(n-1)(n-2)\dots(n-r+2)} = \frac{C_{n-r+1}^r}{C_n^r}.$$

For $n = 90$ and $r = 5$ $P_s = \frac{404,957}{511,038}$

and the probability of at least one sequence is

$$P_c = 1 - P_s. \tag{5}$$

The probability that among the r numbers of a drawing there will be α sequences of a numbers, β and γ sequences of b and c numbers etc, is, as Euler stated, introducing the notation $\alpha(a)$, $\beta(b)$, $\gamma(c)$, etc,

$$P = \frac{(n-r+1)(n-r)(n-r-1)\dots(n-r-k+2)}{\alpha!\beta!\gamma!\dots} : C_n^r, \tag{4}$$

$$k = \alpha + \beta + \gamma + \dots, \alpha(a) + \beta(b) + \gamma(c) + \dots = r.$$

It followed that [the author provides a table of the number of various sequences and their probabilities. The sequences are 1(5); 1(4) + 1(1); 1(3) + 1(2); 1(3) + 2(1); 2(2) + 1(1); 1(2) + 3(1); and 5(1).].

For explaining that table I adduce calculations for case 4, i. e. for 1(3) + 2(1), which means one sequence of 3 numbers and 2 isolated numbers, for example, 5, 6, 7, 65, 83; the order of these numbers is of no consequence. Here $\alpha = 1$, $\beta = 2$ and $k = 3$. Substituting these values in formula (4), we get

$$\frac{86 \cdot 85 \cdot 84}{1 \cdot 2} \cdot \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}{90 \cdot 89 \cdot 88 \cdot 87 \cdot 86} = \frac{3,570}{511,038}.$$

And the sum of all such probabilities for the appearance of at least 1(2), of one sequence of two numbers, is 106,081/511,058 which is equal to P_c , see formula (5).

[The author similarly deduces the probabilities of the occurrence of at least one sequence 2(2), 1(3), 1(4) or 1(5).]

We ought to explain why exactly cases 1 – 7 correspond to $r = 5$. The question is obviously identical with the one concerning the number of possible partitions of number 5 into natural summands and Euler indicates his “studies of partitioning the numbers”.

It is opportune to mention Jakob Bernoulli’s *Ars Conjectandi* where the appropriate rule was already provided in pt. 1 [in his commentary to Huygens’ Proposition 9]. It adjoins there the considerations of Huygens on the number of possible combinations appearing in a definite result of a throw of dice.

It is sufficient to formulate the problem as follows: In how many ways can we obtain 5 points in a throw of 1 die, or 2, ..., or 5 dice without distinguishing between them? Following Bernoulli’s indication, we find that there are only seven such possibilities [for three dice see Bernoulli (1713/1899, Table on p. 27)].

Then Euler examines the number of possibilities occurring among all the types of sequences and here is his answer:

$$A = \frac{(r-1)(r-2)\dots(r-k+1)}{1 \cdot 2 \cdot \dots \cdot (k-1)} \cdot \frac{(n-r+1)(n-r)(n-r-1)\dots(n-r-k+2)}{1 \cdot 2 \cdot 3 \cdot \dots \cdot k}.$$

For the numerical lottery we have

a) If $k = 1$, i. e., for 1(5), there are 86 possibilities of drawing 5 consecutive numbers out of 90.

b) $k = 2$, i. e., 1(4) + 1(1) and 1(3) + 1(2): $A = 14,620$ possibilities.

c) $k = 3$, i. e., 1(3) + 2(1) and 2(2) + 1(1); $A = 614,040$ possibilities.

d) $k = 4$, i. e., 1(2) + 3(1): $A = 8,494,220$ possibilities.

e) $k = 5$, i. e., 5(1): $A = 34,826,302$ possibilities.

Thus all the possibilities for the case of combinations of 90 elements 5 at a time are exhausted. If the calculations are correct, the sum of the possibilities for all the sequences from $k = 1$ to $k = r = 5$ should be equal to $C_n^r = 43,949,268$. Indeed,

$$86 + 14,620 + 614,040 + 34,826,302 + 43,949,268 = 43,949,268.$$

In the introduction to his memoir Euler indicated that the occurrence of numbers 90 and 1 in one and the same series of drawings might be considered as a sequence of two numbers; then, however, he adds: “It is more natural to exclude [such cases] and only keep to the natural order of numbers”.

Johann III Bernoulli, 1744 – 1807, and N. de Beguelin, 1714 – 1789, considered the probability of the appearance of sequences also if 90 and 1 were thought to constitute a sequence. Beguelin (1767, p. 233) remarked:

This means, to imagine that all the numbers are arranged along a circumference so that one more sequence is added, – that, which is formed by the greatest and the smallest numbers closing the circumference.

Johann III Bernoulli (1771, p. 235) justified the same assumption by noting that “here is no number preceding 1 or following after 90”.

Beguelin's memoir was published directly after Euler's study, and Bernoulli's work appeared two years later. In a footnote added at the time of publication, Bernoulli remarked: "This memoir was read in 1765 after Euler's memoir on the same subject which was included in the *Histoire de l'Académie* [...] for the same year".

At the time when Beguelin wrote his paper consisting of two parts, he was already acquainted with Bernoulli's study. An examination of his results derived under the assumption mentioned would have led us too far afield; in addition, Cantor (1908/1965, pp. 235ff) had described Beguelin's *involutorische* method.

For the sake of completeness and to compare their conclusions with the result achieved by Euler, I indicate the Beguelin – Bernoulli formula as provided by Cantor for the absence of sequences among r numbers of a drawing:

$$P_s = \frac{n}{r} \frac{C_{n-r-1}^{r-1}}{C_n^r}.$$

In the footnote mentioned above, Bernoulli indicated that the lotto had become more popular than ever before. Warschauer (1885, p. 38) confirmed this remark. He noted that on June 10, 1769, the leaseholders of a lottery petitioned the King for raising the rent and prolonging the term of their contract for a longer period than previously. With each drawing ever more people participated in the lottery, and the leaseholders' profit continuously increased. No wonder that the numerical lotto repeatedly offered occasions to studies in the theory of probability and Bernoulli, just like Euler, stressed the difficulties which he encountered.

Another problem presents itself in connection with the numerical lottery: To find the probability that after i drawings of r numbers each all the n initial numbers will have appeared. Euler (1785), as quoted by Cantor (1908/1924, p. 236), busied himself with this problem as well. I am only reporting his result which can be expressed as

$$P = \frac{[C_n^r]^i - n [C_{n-1}^r]^i + C_n^2 [C_{n-2}^r]^i - \dots}{[C_n^r]^i}.$$

Performing that division, we arrive at the expression derived otherwise by Meyer (1874/1879, p. 46ff) in his lectures of 1849 – 1857 at Liège:

$$P = 1 - \frac{n}{1} \left(\frac{n-r}{n}\right)^i + \frac{n(n-1)}{2!} \left[\frac{(n-r)(n-r-1)}{n(n-1)}\right]^i - \frac{n(n-1)(n-2)}{3!} \left[\frac{(n-r)(n-r-1)(n-r-2)}{n(n-1)(n-2)}\right]^i + \dots$$

For example, assuming $n = 90$, $r = 5$ and $i = 100$, for 5 coincidences we have $P = 0.7410$. The expression above enables to calculate P approximately. Since

$$(1-x)^n = 1 - C_n^1 x + C_n^2 x^2 - C_n^3 x^3 + \dots,$$

we have

$$1 - C_n^1 \left(\frac{n-r}{n}\right)^i + C_n^2 \left(\frac{n-r}{n}\right)^{2i} - C_n^3 \left(\frac{n-r}{n}\right)^{3i} + \dots = \left[1 - \left(\frac{n-r}{n}\right)^i\right]^n.$$

For large values of n the omitted terms may apparently be neglected so that approximately, for the numerical lotto¹³,

$$P = \left[1 - \left(\frac{90-5}{90}\right)^i\right]^{90}. \quad (6)$$

Given P , i can now be directly calculated; for $P = 1/2$, $i \approx 85$ drawings.

Had the probability of extracting a number remained constant in all the drawings, formula (6) would have been exact. However, P varies during the drawings; initially, it is equal to r/n , then it becomes $(r-1)/(n-1)$ etc, and when the fifth number is to be drawn, it will be $(r-4)/(n-4) = 1/(n-4)$, see Meyer (1874/1879, p. 54).

Meyer (pp. 33ff and 41ff) also examined another problem: To determine the probability that at least one drawn number will be expressed by one digit. In our notation, his formula, which he derived by solving an urn problem, was

$$P = \frac{1}{n(n-1) \dots (n-r+1)} \cdot \sum_{f=1}^r \left\{ \frac{r!}{f!(r-f)!} [a(a-1) \dots (a-f+1)][b(b-1) \dots (b-r+f+1)] \right\}.$$

Here, a and b are the numbers of one- and two-digit numbers (= 9 and 81), and f is the number of one-digit numbers in one drawing ($f = 1, 2, 3, 4, 5$).

The probability that at least one number in one of the drawings will consist of one digit (equal to the sum in that formula) is 0.417019. Separate summands, each of them multiplied by the fraction indicated, provide the probabilities that all the five numbers, or four, three or two of them, or one, will consist of one digit. Thus, for $f = r = 5, 4, 3, 2, 1$

$$P = 0.000003, 0.000232, 0.006193, 0.069888, 0.340703.$$

It is interesting to compare these with the actual results of 201 drawings made in Berlin in 1794 – 1805 (*Nachricht* 1806?, pp. 21ff)¹⁴. A one-digit number won 70 times; expected number of times, 68.5. Such a number occurred twice 11 times, and three times in two drawings (expected number of times, 14 and 1.3). In all, at least one one-digit number thus occurred 83 times out of 201; expected number of times, 83.8. Bearing in mind, that only 201 drawings were made, the coincidences were very good¹⁵.

My aim was to indicate selected results of those studies prompted by games of chance over three centuries and made by classics of the theory of probability. And the vivid interaction of theory and life becomes apparent in the example of the numerical lottery.

Notes

1. Sighart (1898, p. 5) mistakenly names the year 1720. K.-R. B.
2. The author adduces a long list of Italian literature on the number lottery. K.-R. B.
3. According to Nina as quoted by Weidlich (1922), there were 120 candidates. Weidlich also states that bets have been made since ca. 1576. K.-R. B.
4. For more details see Weidlich (1922). In Austria, Count O. di Cataldi, an Italian, got the first patent on a numerical lotto. Later E. A. Calcabigi [spelling uncertain] obtained the same in Prussia. The date Nov. 13, 1751 (Endemann 1899, p. 74) concerned the patent *Codex austriacus*, Bd. 3, p. 66ff. K.-R. B.
5. South of Berlin, latitude ca. 52°.
6. The lotto became the subject of many accusing and defending works, interpretations and astrological writings. Poets glorified it, winning formulas were being published. Members of parliaments ardently spoke for and against the lotto etc. The indicated sources contain many further references. K.-R. B.
7. According to Sieghart (1898) and Warschauer (1885). I do not dwell on such particulars as connecting the drawings with raffles for five complete sets of trousseaus for orphan girls (Prussia and Austria) or sequestering stakes, or decreasing the winnings if more than three coincidences occurred in a terno. The sources indicated consider these issues in detail. K.-R. B.
8. I am following the original publication (*Acta Eruditorum Supplementa*, t. 4, 1711, pp. 159ff, 167ff). [See Jakob Bernoulli 1975, pp. 320ff.]
Here and in the sequel I am using my own notation. I aimed at keeping to the original as closely as possible, but, for facilitating the understanding, I allowed myself to develop partly N. Bernoulli's considerations. K.-R. B.
9. On p. 105ff. K.-R. B. Chapter 4, Corollary 5, without formulas.
10. Warschauer (1885, p. 6) reports, without stating his source, that for a stake of 1 gold pistole the winning was 20 thousand, 5 – 6 thousand and 500 – 600 pistoles for 5, 4 and 3 coincidences respectively. K.-R. B.
11. Coste (1933) mistakenly mentions 8,789,832 for a quinterno. See also Vega (1838). K.-R. B.
12. Sieghart (1898) and Warschauer [where exactly?] report that the leaseholders or the state, for example, in Austria and Prussia, received great profits from numerical lotteries. K.-R. B.
13. The author had not estimated the error of his approximation. For the numerical lotto, his formula (6) provides $P = 0.7427$ instead of $P = 0.7410$ as above.
14. This source contains the most comprehensive relevant data available. K.-R. B.
15. In such cases two hundred trials are quite sufficient.

Bibliography

- Beguelin N. de** (1767), Sur les suites ou séquences dans la loterie de Genes. *Hist. Acad. Roy. Sci. et Belles-Lettres Berlin*, année 1765, pp. 231 – 280.
- Bellhouse D. R.** (1991), The Genoese lottery. *Statistical Science*, vol. 6, pp. 141 – 148.
- Bernoulli Jakob** (1713), *Ars Conjectandi*. Basel. Reprinted in author's book (1975, pp. 107 – 259). German transl. (1899): Frankfurt/Main, 1999.
--- (1975), *Werke*, Bd. 3. Basel.
- Bernoulli Jean III** (1771), Sur les suites ou séquences dans la loterie de Genes. *Hist. Acad. Roy. Sci. et Belles-Lettres Berlin*, année 1769, pp. 234 – 253.
- Bernoulli N.** (1709), De usu artis conjectandi in iure. Reprinted in Jakob Bernoulli (1975, pp. 287 – 326).
- Biermann K.-R.** (1955), Über eine Studie von G. W. Leibniz zu Fragen der Wahrscheinlichkeitsrechnung. *Forschungen und Fortschritte*, Bd. 29, pp. 110 – 113.
- Cantor M.** (1892), *Vorlesungen über Geschichte der Mathematik*, Bd. 2, Abt. 3. Leipzig, 1900. [New York, 1965.]
--- (1908), Same source, Bd. 4. Leipzig, 1924. [New York, 1965.]
- Coste P.** (1933), *Les loteries d'état en Europe et la loterie nationale*. Paris.
- De Moivre A.** (1718), *Doctrine of Chances*. Third edition, 1756, reprinted: New York, 1967.
- Endemann Fr.** (1899), *Beiträge zur Geschichte der Lotterie*. Berlin.
- Euler L.** (1749), Letter to Friedrich II concerning an Italian lottery. *Opera omnia*, ser. 4A, t. 6. Basel, 1986, pp. 317 – 319.

- (1767), Sur la probabilité des séquences dans la loterie Génoise. *Opera omnia*, ser. 1, t. 7. Leipzig – Berlin, 1923, pp. 113 – 152.
- (1785), Solutio quarundam quaestionum difficiliorum in calculo probabilitium. *Ibidem*, pp. 408 – 424.
- (1862, read 1763), Réflexions sur une espèce singulière de loterie nommée Génoise. *Ibidem*, pp. 466 – 494.
- Laplace P. S.** (1814, in French), *Philosophical Essay on Probabilities*. New York, 1995. Transl. A. I. Dale.
- (1868, read 1819), Sur la suppression de la loterie. *Oeuvr. Compl.*, t. 14, 1912, pp. 375 – 378.
- Meyer A.** (1874, in French), *Vorlesungen über Wahrscheinlichkeitsrechnung*. Leipzig, 1879. Transl. E. Czuber.
- Nachricht** (1806?), *Nachricht von den Kgl Preuß. Zahlen- und Classen-Lotterien auf das Jahre 1806*. Berlin.
- Poisson S.-D.** (1837), *Recherches sur la probabilité des jugements etc.* Paris. [Paris, 2003.]
- Rodotà P. P.** (1769), *De' giuochi d'industria, di sorte, e misti; di quello in particolare, che si denomina Lotto di Genova*. Roma.
- Sheynin O.** (2007), Euler's work in probability and statistics. In *Euler Revisited. Tercentenary Essays*. Editor R. Baker. Heber City, Uta, pp. 281 – 316.
- Sieghart R.** (1898), *Geschichte und Statistik des Zahlenlottos in Österreich*. Freiburg a. o.
- Todhunter I.** (1865), *History of the Mathematical Theory of Probability*. New York, 1949, 1965.
- Vega G.** (1838), *Vorlesungen über die Mathematik*, Bd. 1. Wien.
- Warschauer O.** (1885), *Die Zahlenlotterie in Preussen*. Leipzig.
- (1912), *Lotteriestudien*. Berlin.
- Weidlich C.** (1922), *La questione del lotto*. Palermo.

XVIa

V. Ya. Buniakovsky

Principles of the Mathematical Theory of Probability

Osnovania Matematicheskoi Teorii Veroiatnostei. Petersburg, 1846
Extracts (Prokhorov 1999, pp. 863 – 869)

1. From the Writer

The analytical theory of probabilities is included in the realm of applied mathematics and essentially differs from the other applications of pure analysis. In geometry, and in the subjects of natural philosophy, such as for instance in the phenomena of universal gravitation, in the theories of light, heat, sound, electricity, all research is based partly on our notions about various really existing or only imagined by us magnitudes, and partly on the laws derived from experiments, or, should such an experimental foundation be lacking, on more or less likely hypotheses. On the contrary, the analysis of probabilities studies and numerically assesses phenomena that depend on causes which are not only completely unknown to us, but which, owing to our ignorance, do not even yield to any assumptions¹.

Subtle and profound deductions leading us to this goal, make up in their totality the most reliable route to at least a possible approach to the truth if not to revealing it undoubtedly. And upon taking into consideration that, being entrusted with this important purpose, the applications of the mathematical doctrine of probabilities embrace the subjects of the physical and moral world, we may affirm that this theory, the most elevating

creation of our mind, as though indicates the boundaries of knowledge beyond which it is impossible to pass².

The book which I am offering now, is the first Russian composition including a detailed description of both the mathematical principles of the theory of probability and of its most important applications to social life and natural philosophy as well as to political and moral sciences. The last chapter is devoted to historical details on the gradual development of the analysis of probabilities. Ten purely mathematical Supplements are appended to the book and they will spare some readers the trouble of looking elsewhere for an explanation of the various theories often encountered in the calculus of probability. After them follows an Explanation of the two useful tables [also] appended to my book, and, finally, an Addendum contains a solution of a curious problem. However, I refer readers to the Contents themselves where a detailed indication of the subjects included in the book can be found.

A few words about the execution itself of my work. Laplace's immortal creation, the *Théorie analytique des probabilités*, invariably served me as a specimen both by the elegance of the analysis made use of and profundity of reasoning. At the same time, however, while offering many theories created by him, I have always attempted to simplify as much as possible both their description and proof and the analysis itself. I fearlessly hope that mathematicians will do me justice by noting that I have essentially facilitated the study of Laplace's book, which, owing to its conciseness and the special difficulty peculiar to its subject, is only intelligible to very few readers.

Scientific research accomplished by other celebrated geometers, and in the first place by Euler, Lagrange and Poisson, had also been useful for me. From Poisson I have borrowed the description of the mathematical theory of legal proceedings. Regarding my other results, I restrict my comments to referring to some critical remarks and to the changes in the generally received analytical methods, useful in my opinion, and made in various cases. In this respect, I am turning the readers' attention in the first place to chapters 7 and 10.

The reader himself will note these changes when attentively reading many sections of my book and comparing them with the description in well-known other sources. More extensive research of my own are accompanied by indications in the text. One more remark. Since until now we had no separate work and not even a translation of a work on the mathematical theory of probability, I was compelled to describe in Russian a subject for which established turns of phrases and expressions were lacking. I do not dare hope that I have quite satisfactorily created a simple and definite language for the analysis of probabilities but am pleased to be sure that in any case I have exerted every effort to approach this goal as closely as possible³.

I conclude by expressing my wish that the offered work will foster the spread of sensible notions and useful practical truths among my compatriots. And even if some of my readers are not mathematically educated to a sufficient level and are unable to follow the analytical exposition of all the theories comprising the subject of the doctrine of probability, – even for them an attentive reading of my book will not be

useless. They will derive various results of general usage showing in their true light many entertaining issues and truths concerning our social life.

2. Conclusion. In What Sense Ought We To Understand the Corollaries Provided by the Mathematical Theory of Probability

Having consistently described the mathematical principles of the analysis of probabilities and its main applications to social life, natural philosophy and moral sciences, I am now able to give a concise account about what may we expect and demand from this theory which in all fairness can become on a par with the most important branches of our knowledge. Except a very few indisputable truths which had become a treasury of mankind, everything in nature and in the moral world is based on more or less likely conjectures.

The doctrine of probability therefore actually encompasses almost the entire sphere of intellectual activities. Such a vast purpose of that science is undoubtedly essentially restricted by lack of data furnished by observing physical and moral phenomena and their inadequacy, and, on the other hand, although to a lesser degree, by the imperfection of mathematical analysis. Nevertheless, the achievements of the theory of probability attained until now places it on the level of a most important intellectual tool for revealing truths and protecting the mind from delusions into which it often falls after superficially viewing a subject.

In cases in which an acute-minded person is only able to foresee approximate results, the theory often leads us to precise and numerically expressed conclusions. Such definite estimation of the measure of confidence in some conjectural truth, impossible for usual logic, undoubtedly deserves full attention of thinkers. However, these numerical conclusions should not be understood unquestionably as is done by some empiricists who have not comprehended the real spirit of the analysis of probabilities. Thus, for example, if the probability of some event is very close to certainty, or to unity, it does not follow that that event will occur without fail, or, to the contrary, that the theory leads to wrong inferences. The result mentioned ought to be understood in another sense which is quite justified by the well-known general proposition due to Jakob Bernoulli.

A high probability only indicates that, if we were able to repeat trials many times over under identical circumstances, the number of occurrences of the event will be incomparably greater than the number of its failure to occur, and the ratio of the first number to their sum will ever more closely approach the derived value of the probability. As to an isolated trial, the analysis of probabilities is unable to provide any definite conclusions because the conditions [of the problem] are not clear-cut as I have explained at the very beginning of this book. And in general, when explaining the various results of the analysis of probabilities, we ought always to bear in mind the law expressed by the Jakob Bernoulli theorem. In case of events depending on chance because of our ignorance regularity in the number of their occurrences only takes place if the number of trials is very large. For this reason, any decision concerning an isolated case should be only understood as a *mean conclusion* which in many instances might essentially deviate from the result revealed by posterior events. If,

however, it would have been possible to repeat the same trial indefinitely many times under the same conditions, the mean result thus obtained will approximate the sought ratios between the number of occurrences of various events the closer the more trials are made.

In concluding, I offer the readers a concise historical essay on the gradual development of the mathematical theory of probability.

3. A Concise Historical Essay on the Gradual Development of the Mathematical Theory of Probability

118. The time to which the initial speculative notions on probability belong, is as indefinite as are the beginnings of most branches of our knowledge. Under various circumstances, as in games [of chance], when betting etc, and long before the first attempts at a mathematical theory of that science had been made, man resorted to comparing the numbers of favourable and unfavourable cases, and more or less luckily derived appropriate conclusions.

Suchlike considerations as well as some rules occurring in the works of philosophers of old, certainly belong to the doctrine of chances; there even exist definite testimonies that some remarkable applications of the science of probabilities had not been alien to very remote times. Thus, for example, Libri (1845) noted that the *Digest* contains a law on foodstuffs which clearly indicates that even the Romans had been determining mean life for various age groups. Libri went on to discuss societies of marine insurance which had already existed in the medieval Italian republics. This compels us to suggest that it was known in those times how to determine approximately the probability of shipwreck.

It is also known that later, in the beginning of the 17th century, the illustrious Galileo busied himself with a very important problem in the theory of probability, with deriving errors and estimating their influence on the results of observation. His examination of such a difficult issue naturally had not led to desired success. It seems that, again, the first idea about turnover of capital based on probabilities of human life originated in the first half of the same century. The Neapolitan Lorenzo Tonti suggested a special pertinent institution which until now has been called after him (the *tontine*, § 73).

All the attempts mentioned above undoubtedly belonged to the theory of probability, but, nevertheless, having been fragmentary and imperfect, they were unable to satisfy the demands of a science. Only Pascal and Fermat in the mid-17th century laid the foundation of the mathematical theory of chances. Chevalier De Méré had proposed to Pascal the first problem which they solved. It dealt with a fair division of stakes in an interrupted game, see details in §§ 32 and 38.

And so, in fairness and justice, we may state that the calculus of probability owes its mathematical beginnings and independence to these two celebrated scholars. Soon afterwards their contemporary Huygens busied himself with the same subject. He collected the already solved problems, supplemented them by his own research and composed a tract on reasoning in games subjected to chance. His writing, the first to appear on the theory of probability, came out incorporated in Schooten in 1657. It

is also included in pt. 1 of the *Ars Conjectandi*, which I discuss below, enriched by Jakob Bernoulli's commentary.

In the second half of the 17th century Sauveur also studied the theory of probability (*J. des Savans* for February 1679; Todhuner 1865, p. 46). He examined the chances in a game resembling Pharaon and known as *bassète*. In addition, Montucla (an X, 1802, t. 3, p. 391) mentions a booklet entitled *Of the Laws of Chance* anonymously published in London in 1692. He believes that the author had been Benjamin Motte⁴.

Towards the end of the same century there appeared the works of van Hudden [Hudde]⁵, Witt [De Witt], the *Pensionnaire d'Hollande*, and Halley on the probabilities of human life. Halley (1694) published a mortality table, the first known to us (§ 60)⁶. Without mentioning other, less remarkable acquisitions made by the calculus of probability during the same period, I pass over to the work of the celebrated Jakob Bernoulli. Already in 1685 he proposed to mathematicians a rather difficult question concerning a game of dice. Since no answer had been offered, he himself published his solution, although without proof, in the *Acta Eruditorum* for 1690⁷. This prompted Leibniz to examine the posed problem. He solved it at once and published a detailed description of his method of solution in the same *Acta* [De Mora-Charles 1986].

The main service that Jakob Bernoulli rendered to the mathematical theory of probability was undoubtedly the compilation of his remarkable *Ars Conjectandi* on which he had deliberated for many years. His nephew, Niklaus Bernoulli, published it in Basel in 1713⁸, seven years after Jakob had died. This work, distinguished by correct viewpoints and clever analytical methods, is separated into four parts. The first one, as mentioned above, is made up by explanatory comments on Huygens' treatise. The second part includes an extensive theory of various kinds of combinations. Many problems dealing with various games are discussed in the third part and, finally, the fourth part contains the use and application of the rules explicated in the previous parts to issues in everyday life and moral and political sciences⁹.

This part deserves special attention because the Newtonian binomial¹⁰, being so important in the calculus of probabilities, is applied for solving the problems there. But the most remarkable item in pt. 4 is unquestionably the proof of the well-known theorem which has retained its author's name and which I had the opportunity to mention so many times, see details in §§ 20, 22, 24, 26, ..., 117. After the fourth part follows a treatise on infinite series and a curious investigation of an unknown author¹¹.

Niklaus Bernoulli, who published the *Ars Conjectandi*, had been himself somewhat successfully studying the theory of probability. In 1709, in Basel, he defended his dissertation for a doctor's degree in jurisprudence on the application of the calculus of probability to the administration of justice. Among the curious problems solved there I may indicate, in the first place, that which constitutes the subject of pt. 3 of Niklaus' thesis: how much time should pass before an absentee about whom nothing is known must be declared legally dead.

119. The 18th century marked by so splendid successes achieved by pure mathematical analysis essentially improved the theory of probability as well. In the very beginning of that period Montmort in France, and De

Moivre of French extraction living in England most diligently examined the calculus of probability. The former published a book (1708) in which he offered solutions of many curious problems belonging to various card and dice games etc. In the second edition of his book, corrected and enlarged on many points, he included his curious correspondence with Niklaus Bernoulli, the nephew of Jakob and Johann Bernoulli.

Special notice here is merited by Niklaus' clever solution of many problems and by the description of the one known as the Petersburg problem which he proposed for Montmort, see details in § 45. Among difficult problems whose solution occupied Montmort, I may also point out the division of stakes between gamblers when, according to the essence of that problem, the ending of the game remains indefinite. De Moivre also studied the same subject, but the solutions of these scholars lacked adequate completeness (§§ 33 and 40).

De Moivre submitted his first work on the theory of probability to the London Royal Society, and it appeared in the *Philosophical Transactions* (1712). He had been then publishing his research gradually perfecting it (1716 [1718!], 1738, and 1756). Concerning analytical methods, this writing is much preferable to all the previous works. In general, problems are solved there more generally by means of the Newtonian binomial. The Jakob Bernoulli theorem was very importantly developed there: De Moivre established the probability that the difference between the ratio (otnoshenie) of the actual number of the occurrences of events and the ratio of their probabilities is contained within given bounds¹², and he was the first to apply here the Stirling theorem (§ 21). But the book is especially remarkable for describing his invented theory of *recurring series* which he highly successfully applied for solving various problems about probabilities. Actually, his theory includes a method of integrating equations in finite differences with constant coefficients, so fruitful in its applications to the analysis of probabilities.

At about the same time many other mathematicians have been more or less successfully working at the theory of probability. Among them was Mairan, see § 37 (Todhunter 1865, pp. 200 – 201) and Nicole (1732a; 1732b; Todhunter 1865, pp. 201 – 203) who solved various problems on the fate [the expectations] of gamblers of unequal skill after some of them wins more rounds than others.

In the second half of the 18th century many scientists had been very carefully collecting various data on population in general, on mortality, the number of births, marriages etc. These numerical indications, after having been adequately examined, served for compiling many extremely useful tables and solving highly practical and various problems on probabilities of human life, on life annuities, savings banks, tontines, insurance of any kind etc. Historical details concerning this subject are to be found in Montucla (an X, 1802, t. 3), and I am restricting my description by concisely describing the main works of the scientists.

Very close to the mid-18th century remarkable writings on the same subject were published by Thomas Simpson in England, Kersseboom and Struyk in Holland and Deparcieux (1746) in France. Curious studies of mortality by the Swedish astronomer Wargentín are included in the periodical of the Swedish academy (1754 – 1755). In Germany, the mathematician Lambert examined the same subject (§ 60) as did Euler and

some others. During the last years of the century, Deparcieux (1781), a nephew of the man mentioned just above, had published a treatise on annuities and soon afterwards appeared a very remarkable book on financial turnovers by Duvallard (1787), see § 60. At about the same time the writing of Price (1771) on various issues of political arithmetic deserved general attention.

I shall describe now as concise as possible the most important acquisitions attained by the calculus of probability in the 18th century. Daniel Bernoulli, the son of Johann Bernoulli, whose discoveries enriched higher geometry and mechanics, was the first to suggest that the mathematical and the moral expectations be distinguished one from another and to introduce a measure of the latter (Chapter 4). Almost at the same time the illustrious natural scientist Buffon (1777)¹³ described his thoughts on the same subject (§ 42). The readers will also find a letter from Daniel Bernoulli to Buffon dated 19 March 1762 which testifies that he regarded Buffon's viewpoint on moral probability absolutely sound although did not quite agree with its proposed measure [proposed value]. There also Buffon included mathematical solutions of a few problems from the analysis of probabilities and an application of that theory to issues in human life, births, marriages, mortality tables etc.

I return now to the works of Daniel Bernoulli. The analysis of probabilities owes him also an original idea, so fruitful because of its application to considering posterior probabilities of events, i. e., by issuing from observed phenomena (Chapter 7)¹⁴. Later Bayes and Price (1764 and 1765) and then Laplace adequately generalized them. Daniel (1766) also applied the calculus of probability to the problem of inoculation (§ 64) which led to a rather heated discussion between him and D'Alembert (1761; 1768a; 1768b).

D'Alembert's other contributions to the theory of probability are in his various works and partly in the *Enc. méthodique* (Mathématique). That excellent source also contains articles by Condorcet belonging to the analysis of probabilities; the most important of them both by its length and contents is *Probabilité*. His other contributions to that science were published in the *Mémoires* of the Paris Academy of Sciences for the years 1781 – 1783¹⁵, and his most remarkable work (1785) treats majority decisions. I referred to some of its places in Chapter 11.

Euler enriched almost every branch of pure and applied mathematics, and he also studied different parts of the theory of probability. He left rather many relevant memoirs and I indicated some of them in §§ 36, 65 and 72. There are also his manuscripts [1862a; 1862b] as well as his curious correspondence with Friedrich II, also unpublished, on a special lottery [apparently (1749)]. His main merit is, however, the perfection of the integral calculus which to the highest degree fostered the speedy progress of the analysis of probabilities.

Lagrange (1777) offered a simple and handy method of integrating equations in partial finite differences and showed how to apply it for solving difficult and at the same time curious problems of the calculus of probability. I considered this subject in detail in Chapter 3 and Note 7, and in §§ 78 and 79 I mentioned another of his works devoted to the determination of most advantageous results of observations (1776).

I also indicate a work by Lacroix belonging to the theory of probability. In 1781, the Paris Academy of Sciences proposed a prize question on marine insurance. Since it had not received satisfactory solutions, the question was twice repeated, after which, two from the eight answers taken together were judged to warrant half the prize. Lacroix wrote one of these answers and got 1800 francs, the other, written by Bicquille, was awarded by 1200 francs. Lacroix had also published a very satisfactory book (1816; 1828; 1864; German translation 1818) and Bicquille (1783) has a book to his credit as well.

I will not discuss the works of Legendre and Gauss on the determination of the most probable results of observation, see Chapter 10 (§ 92)¹⁶. There also I provide other historical details on the most beneficial combination of the initial equations and, among other items, I mention the method of the English mathematician Cotes (end of § 85).

Nevertheless the analytical theory of probability owes Laplace more than anyone else. I had occasion to discuss his works so often that it seems sufficient only to offer here as concisely as possible the main merits of that great geometer. In addition to many memoirs which appeared in the periodical of the Paris Academy of Sciences on the analytical theory of probability, he (1812) published a great work on the same subject which covered its complete theory and all of its main applications. Laplace's profound mind, subtle viewpoints and might of mathematical analysis is not seen in any of his other writings as powerfully as there.

Elegance and generality of his methods applied for solving the most difficult problems of the analysis of chances elevated this theory to a high level of perfection. From the most remarkable of his studies which most of all enriched the doctrine of probabilities I may mention in the first place

1) The *theory of generating functions* for integrating equations in finite differences, a procedure so often encountered in such issues.

2) Then, approximate calculations of various integrals of functions of large numbers; particular cases of such formulas were also treated previously, see for example the Stirling approximation of $n!$ (§ 21) the exact value of which is represented by the definite integral

$$\int_0^{\infty} x^n e^{-x} dx.$$

3) General formulas for posterior probabilities derived from observations made (Chapter 7) and the calculation of the probabilities of future events when prior chances are considered, but actually are not equally possible (Chapter 5).

4) Various applications of the calculus of probability to phenomena observed in the solar system; thus, for example, the determination of the probability that there existed an initial cause forcing all the planets and their satellites to rotate about their axes and to move along their orbits from West to East, that is, in the same direction as the sun's rotation and almost in one and the same plane with its equator¹⁷.

5) The theory of most beneficial results of observations (my Chapter 10) so important by its applications to observational sciences, owes to Laplace its present perfection. He also indicated its application to geodetic work and developed it.

6) Finally, in a separate publication (1814) we find a complete compendium and description of the facts of the theory and applications of the analysis of probabilities without formulas or calculations.

I have presented a cursory list of Laplace's most important works [results] in the analysis of probabilities and it can be concluded that this theory, which originated in France by the hands of Pascal and Fermat, also owes its speedy perfection to a French geometer.

120. To our century, apart from Laplace's main works and Gauss and Legendre whom I mentioned above, also belong various investigations of many astronomers and mathematicians. Bessel, Plana, Encke, Struve, Poisson, Lindenau, Bonenberger and others theoretically and practically busied themselves with the issue of the most beneficial results of observations (§§ 89, 91, 92, 95). In addition to Poisson's writing mentioned in § 94, he published a few memoirs on the calculus of probability including memoir (1837b). In this curious writing he expounded the mathematical theory of probability of target shooting. From the obtained formulas he derived rules for comparing the accurateness of firearms and skill of shots. Experiments made by French artillery men completely corroborated his theory and proved that the obtained formulas were perfectly useful.

Poisson rendered his main service to this science by a separate treatise (1837a) on the mathematical theory of legal proceedings. It is separated into five chapters with the first four being devoted to the general principles of the calculus of probability and its most common applications. The last chapter, however, exclusively deals with the analytical theory of legal proceedings.

In that book, Poisson extended the Jakob Bernoulli theorem to cover variable chances and named a certain proposition the *law of large numbers* which I mention in a footnote on p. 35. In addition to the mathematicians mentioned above who studied the theory of probability during recent years, many others can be named, in particular Amper, Fourier, Puissant, Hansen, Quetelet, Littrow, Moser.

After having expounded the mathematical principles and main applications and offering a short essay on the progress of the theory of probability, I am concluding my book by Laplace's words (1814/1995, p. 124) on the importance of that science for human knowledge:

[...] *The theory of probability is basically only common sense reduced to calculus*¹⁸. [...] *There is no science at all more worthy of our consideration* [...].

Author's Footnotes

1. The Digest is known to be a compendium of decisions made by the most celebrated Roman lawyers and composed as a code by command of Emperor Justinian. Its appearance is attributed to year 528¹⁹.

2. P. N. Fuss, the perpetual secretary of the [Petersburg] Imperial Academy of Sciences, had informed me about these manuscripts. He is also keeping other memoirs written by Euler on various mathematical subjects and we may hope that these valuable works will be eventually published.

3. The Brussels edition of those collected articles on the art of artillery was published in 1839. In the same periodical Poisson published contribution (1830).

Contents

Chapter 1. On the Laws of Probability in General (§§ 1 – 16)

General rules for determining probability. Definition of probability.

Calculation of probability when chances are not equal to one another.

Application to the pitch-and-toss game. D'Alembert's mistake

The probability of compound events both when the constituting simple events are independent one from another or not. The relation between probabilities of a compound, observed and future events. Relative probability. The rule for its determination

General formulas for calculating probabilities of repeated events or events combined in whichever way. Determination of probability of repeated occurrences of an event or of a certain combination of two events in a given number of trials. Extension of the derived rules to three or, in general, to any number of events. Formulas for determining the probability that one or a few of simple events will be repeated not less than a certain number of times in a given number of trials

Application of the previous formulas to numerical solution of some problems. As an exercise, a solution of seven simple problems is provided; the last of them, proposed by the Chevalier De Méré to Pascal, is remarkable for being one of the first investigated in the theory of probability

Chapter 2. On the Laws of Probability for Infinitely Many Repeated Trials (§§ 17 – 30)

On the most probable compound events. On the most probable compound event composed of two repeated simple events. Proof of various properties of the development of the binomial $(a + b)^n$ on which depends that determination. [?] An example showing that the absolute probabilities of the most likely events lower with the increase in the number of trials. On the contrary, the relative probability of such an event as compared with that of any other heightens with the increase of that number. Extension of these properties to the general case

The Jakob Bernoulli theorem. Its explanation by examples and its exposition. This general law concerning two events can be expressed thus. Given infinitely many trials each leading to one of the two simple events the ratio of the number of their occurrences becomes ever closer to the ratio of their simple probabilities, and, finally, given an adequate number of trials, differs from it as little as desired

The Stirling formula for an approximate calculation of the product $1 \cdot 2 \cdot 3 \dots x$. A detailed proof of the Jakob Bernoulli theorem. Examination of the various corollaries appearing in the process of an analytical proof of the Jakob Bernoulli proposition

Application of that theorem to numerical examples

Its generalization on an arbitrary number of events

A study of a particular case in which the chances vary during the trials. An urn contains A white balls and B black ones. Balls are randomly taken out one after another and set aside. To find the probabilities of various compound events possible for a given number of extractions. The determination of the most probable event in this case. Factorial binomial. Application of the derived formulas to solving some

numerically expressed examples

Chapter 3. On Mathematical Expectation (§§ 31 – 40)

On mathematical equality or fairness of games of any kind and on the measure of mathematical expectation. The notion of mathematical equality of a game. Mathematical expectation or mathematical benefit
Analytical proof of the general condition of fairness or mathematical equality of a game

Exposition and explanation by an example of the rule of fair division of stakes

Application of the condition of fairness to the problem of points in a game that can continue indefinitely

The extension of the rule of mathematical equality of a game to an arbitrary number of events whose occurrence provide stipulated gains for one of the two gamblers

Application of the previous rules to solving various problems concerning the dice game, lotteries, betting and to fair division of stakes in an interrupted game

Chapter 4. On Moral Expectation (§§ 41 – 46)

Notion of the benefit called moral. Impossibility of its precise definition

Measure of moral expectation suggested by Buffon

Another measure applied by Daniel Bernoulli and received up to now by almost every mathematician. Determination of moral benefit according to the Bernoulli formula when it depends on several expected events
Application to insurance and justification, under certain conditions, of the usefulness of insurance institutions

Exposition of a more general hypothesis concerning the measure of moral expectation. Issuing from it, mathematicians rigorously prove that

- 1) Any game or betting [even] mathematically just, is detrimental
- 2) Entirely fair lotteries are detrimental
- 3) When a property [a stock] ought to be subjected to some danger [risk], better to separate it in parts than to expose it to chance in its entirety. On the contrary, when issuing only from mathematical expectation, the choice of an alternative is indifferent

The Petersburg problem; historical remarks about it

Solution based on considering mathematical expectation. An imaginary contradiction appearing in this solution. Its explanation by Condorcet²⁰

Solution based on the Daniel Bernoulli formula

Another solution by Poisson

Some thoughts about the use of moral rather than mathematical expectation

Chapter 5. On the Influence of Unequal Chances Considered Equal on the Results of the Calculus of Probabilities and Investigation of a Special Kind of Combinations Leading to the Discussion of Infinitely Many Chances (§§ 47 – 48)

How the unequally possible chances when considered equally possible change the results of the calculus of probability. Application to the pitch-and-toss game. The solution shows that in two throws of a coin the appearance of the same side without indicating which one is more probable than the appearance of differing sides. Generalization of this result on any events. General formula expressing the influence of inequality of the possibility of chances when supposed equally

possible always heightens the probability of repetition of the same events

On prior calculated probabilities when there are infinitely many chances

Chapter 6. Solution of Some special Problems

of the Analytical Theory of Probability (§§ 49 – 51)

Given, an equation $x^2 + px + q = 0$ whose coefficients p and q are supposed integral and varying from $-m$ to m ; also, for the sake of simplicity, it is supposed that neither of them vanishes. Required is the probability that a randomly written equation [of this kind] has real roots

Given a plane of indefinite dimensions covered by a system of adjoined [congruent] equilateral triangles. A very thin cylinder of a given length is randomly thrown on that plane. Required is the probability that it falls on at least one side of these triangles

Given two squares of a chessboard. Required is the probability that a castle standing on one of them reaches the other square in x moves

Chapter 7. On the Laws of Probability

When There Are Infinitely Many Chances (§§ 52 – 59)

General notions on posterior determination of probability. A numerical example. Rules for determining probabilities of one or more causes or assumptions

The probability of some assumption is equal to the probability of the observed event calculated under the same supposition and divided by the sum of the probabilities of the same event over all possible assumptions

The probability of several assumptions considered in their entirety is equal to the sum of the probabilities of the events over this total divided by the sum of probabilities of the events over all possible assumptions

General formulas for determining

- 1) The probability that the possibility of a simple event is contained within known bounds given the observed compound event
- 2) The probability of a future event given its [present] observation

Extension of the Jakob Bernoulli theorem on the case in which the probability is only determined a posteriori

General formulas for determining the probability under the assumption that the observed event depends on simple phenomena of two or more kinds

Explication of general rules by several simple examples

Chapter 8. On the Probability of Human Life (§§ 60 – 69)

Compiling mortality tables. Graphical representation of the process of mortality. Index of mortality. Equations of the line of mortality suggested by Lambert and De Moivre

Explaining the use of mortality tables for solving various problems concerning the probabilities of human life. Probable life. The measure of longevity. Mean life, formulas for its determination. Numerical results for various states and cities. Determination of the number of inhabitants of a country by a mortality table and distribution of population by age

Influence of the distinction of sex on mortality. Permanent numerical superiority of male births over female births. Numerical results for some states

Some remarks on the growth of population

Determining the length of mean life if a certain cause of mortality is done away with or at least weakened. Vaccination of smallpox lengthens mean life by more than three years. Laplace's formula for calculating the measure of the decrease in the number of deaths when some cause of mortality is done away with

Notes on the movement of population. The measure of increase or fertility. Measure of mortality. Coefficient of increase. Formulas for solving various problems on the movement of population

Solution of some problems on the movement of population based on the indications of mortality tables

Determining the probable and mean duration of marriages, or, in general, of any brotherhoods or companies (tovarishchestvo) or societies. The probability of the existence of a society after the lapse of a given number of years. Given a considerable number of brotherhoods or companies of the same kind, determine the most probable number of the survived after a given number of years

Analytical determination of probability that the possibility of a male birth is higher than the possibility of a female birth. Application of general formulas to births in Petersburg

Determining the population of a great state given the number of yearly births and the population in its various places. Calculating the probability that the error of this determination is contained within given bounds. Numerical application to the population of France

Chapter 9. On Life Annuities, Widow Funds, Tontines, Savings Banks and on Insurance Institutions in General (§§ 70 – 76)

The subject of this chapter. Formulas for reducing deposits and payments to the present time. General remarks on keeping to possible fairness in stipulations between a society and its new members and various obligations concerning any societies

General notions on societies known as tontines. Solution of some problems concerning this kind of mutual insurance. Given the annuity, determine the initial payment made by a member of a tontine, and, given the payment, determine the annuity. An approximate calculation of the annuity of a member of a tontine after one, two, three, ... years. Solution of same problem if the society is only paying some part of the annuity allowing for the number of deceased members

Notion about savings banks in general. Solution of problem: Each of N investors of same age deposited S at the same time. Required is their life annuity after n years. Determining that annuity if the depositors make in addition yearly payments S_1, S_2, \dots

On the insurance of property in general. An insurance premium. For being mathematically fair, the premium ought to equal the cost of the insured item multiplied by the probability of its loss or damage. Actually, the premium is invariably higher; given its moderate excess, the insurance society with a rather wide range of activities makes a sure profit whereas the client will be insured in terms of moral expectation which justifies the mutual benefit of suchlike societies

Application of mathematical analysis to the solution of the following problem. A merchant insures m ships for a each and pays mb . Assess the conditions of such insurance concerning both

the insurer and the insured

Analytical solution of the first part of the problem. Numerical examples

Analytical solution of its second part. Notes on insurance societies in general and on the advantages of societies of mutual insurance

Assuming a smallest excess over mathematical fairness to the benefit of the society and its wide range of activities, it must almost certainly expect essential profit increasing proportional to the number of transactions

Analytical proof of this proposition

Chapter 10. On the Most probable Results of Observations (§§ 77 – 96)

On observations in general

A numerical example. The mean error

A numerical example. Explanation of the contradiction concerning the probability of the mean error

Initial equations. In each application of the method of most probable deduction their linearity is assumed. Various combinations of the initial equations. A system of equations ensuring the least possible [absolute] maximal error. Such a combination of these equations was called the *méthode des situations*²¹. The system of equations leading to the least sum of errors. The mean result of observations. The Cotes rule

Determination of the measure of precision of the mean result of observations

A detailed proof [justification] of the method of least squares. The *erreur moyenne à craindre*. Its comparison with the mean error under the Cotes rule. Determination of the constant introduced in the formulas by an unknown law of probability of the errors. This constant depends on the squared sum of the observational errors. Determination of the *erreur moyenne à craindre* given the coefficients of the initial equations

The weight (*poind*) of the result

If the probability that the error is contained between given bounds is the same, the weight increases as the bounds become closer to each other
The errors are inversely proportional to the square roots of the corresponding weights. The condition for increasing weights. The probable error of the results

The rule for determining an unknown when several series of observations of different kind are available. Here also, just as in the case of one series, it leads to the method of least squares. The resemblance of this rule to the theory of the centre of gravity

Remark on the case in which the method of observation leads to a preponderance of either positive or negative errors. On constant errors
Historical information on the method of least squares. The relevant works of Legendre, Gauss and Laplace

Formulas derived from the method of least squares in the case of two or three unknowns determined by the initial equations

On a special kind of mean error used by German astronomers

Numerical example of the determination of two unknowns from the initial equations

Chapter 11. Application of the Analysis of Probabilities to Testimonies, Legends, Various Kinds of Choice between Candidates Or Opinions and to Majority Decisions in Law Courts (§§ 97 – 117)

General remarks on the subject of this chapter and on the application of mathematical analysis to moral issues

On the probabilities of testimonies

On unusual events

Laplace's solution of a problem concerning testimonies. He states that in such problems two elements ought to be considered, the honesty and the experience of the witness

On probability of legends

On election of candidates. General remarks on this subject. The case of one or two candidates. When there are three or more candidates, the relative majority of votes does not invariably reveal who of them should be preferred

Exposition of the method of voting suggested by the mathematician Borda. An analytical proof [justification] of this method for the case of three candidates. A numerical example

Another kind of voting when the mean merits of the candidates is not allowed for

Extension of the Borda method on an arbitrary number of candidates
Practical inconvenience of the method

On the choice of the most probable proposition or cause. The relevant guiding rule. Its analytical proof. A numerical example

Application of the analysis of probability to administration of justice
Preliminary details and general remarks about its application to judicial decisions. Resemblance of this subject to problems on testimonies

Indications about the work of Condorcet, Laplace, Ostrogradsky and Poisson. The proper viewpoint for considering judicial decisions; what should the sentences *guilty* and *innocent* mean. Mathematical theory of administration of justice only provides mean results of essentially many decided cases but does not apply to isolated sentences

Conclusion. In what sense should we understand the inferences provided by the mathematical theory of probability

Chapter 12. A Concise Historical Essay on the Gradual Development of the Mathematical Theory of Probability (§§ 118 – 120)

Notes

Note 1. The derivation of the Euler [summation?] formula for transforming an integral in finite differences into a usual integral

Note 2. The development of the sine into infinitely many multipliers. The Wallis expression for a quarter of a circumference. Summing infinite series

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots, 1 + \frac{1}{2^4} + \frac{1}{3^4} + \dots, 1 + \frac{1}{2^6} + \frac{1}{3^6} + \dots \text{ etc.}$$

Note 3. On the convergence of infinite series

Note 4. Various investigations concerning the definite integrals

$$\int_0^t \exp(-t^2) dt, \int_t^\infty \exp(-t^2) dt, \text{ etc.}$$

Note 5. Proof of the factorial binomial

Note 6. Proof of the identity ($m < s/n$)

$$C_s^m - sC_{s-1}^m + \frac{s(s-1)}{1 \cdot 2} C_{s-2}^m - \dots + (-1)^{s-n} C_s^n C_n^m = 0.$$

Note 7. Explication of the theory of integrating equations in finite differences

Note 8. Derivation of the general term $p^t y_{-t,0}$ of the equation

$$1 = p^t y_{-t,0} + t p^{t-1} q y_{-t+2,0} + \frac{t(t-1)}{1 \cdot 2} p^{t-2} q^2 y_{-t+4,0} + \dots$$

Note 9. On definite integrals considered in their connection with arithmetical means

Note 10. Summing the series

$$1 + 2(\cos \varphi + \cos 2\varphi + \cos 3\varphi + \dots + \cos n\varphi).$$

Explanation of Tables

Table 1. Contains the numerical values of the integral

$$\frac{2}{\sqrt{\pi}} \int_0^t \exp(-t^2) dt$$

for all the values of the argument t from $t = 0$ to $t = 2$ for each hundredth [$t = 0(1/100)2$]

Table 2. Contains numerical values of the integral

$$\int_T^\infty \exp(-t^2) dt$$

from $T = 0$ to $T = 3$ also for each hundredth. In addition, the table includes logarithms of the same integral for same values of the argument T

Notes

1. Buniakovsky never supplemented this statement by any examples, and he himself (see below beginning of **Conclusion**) went back on his word.

2. This opinion seems to contradict the wrong statement above about the unrestricted realm of the theory of probability.

3. In those times, stochastic terminology was not yet developed. Thus, the expressions *limit theorem* and *random variable* (or, in Russian, *random magnitude*) did not exist. And Buniakovsky used different expressions for the theory of probability calling it in addition *doctrine* (also *doctrine of chances*) and *analysis of probabilities*.

4. Actually, the booklet was a translation of Huygens' treatise (Todhunter 1865, p. 48).

5. On Hudde see Haas (1956).

6. Buniakovsky had not mentioned Graunt.

7. For his problem and his own solution see J. Bernoulli (1975, p. 91; 1993, pp. 160 – 163).

8. This is a mistake. Niklaus only supplied a short Introduction. And, in his own earlier dissertation of 1709, see below, he borrowed separate passages from the *Ars* and even from Jakob's Diary never meant for publication (Kohli 1975, p. 541).

9. Bernoulli only thought of describing this subject; part 4 of the *Ars* contains nothing of the sort.

10. Newton is credited for generalizing the development of the binomial on rational numbers of the exponent whereas Bernoulli had only applied the binomial for the case of natural exponents.

11. Bernoulli himself was the author of both the *Treatise* and the “curious investigation” of a version of tennis.

12. This is a hardly understandable description (repeated in the **Contents**, Chapter 2, see below).

13. See Buffon (1777, § 15), where the author's letter to Gabriel Cramer of 1730 is appended. There also, in § 8, is Bernoulli's letter to him mentioned just below.

14. This is left unclear; moreover, according to the context Buniakovsky referred to a contribution published before 1764, and I am unable to corroborate him.

15. See Kendall & Doig (1968).

16. Invariably following Laplace, Buniakovsky had obviously regarded Gauss quite insufficiently. Here, for example, he did not say “most reliable” results, although Gauss replaced in 1823 *most probable* (as he expressed himself in 1809) by that term.

17. Exceptions have been since discovered.

18. This phrase is generally known, but it seems that in 1814 it could have been applied to the entire realm of mathematics.

19. Somewhat differing dates have also been mentioned.

20. See a much later similar but independent explanation (Freudenthal 1951).

21. Actually, this expression (Laplace 1818) denotes the choice of the median.

Bibliography

Bayes T. (1764 – 1765), An essay towards solving a problem in the doctrine of chances. Second part of Essay: *Phil. Trans. Roy. Soc.*, vol. 54 for 1764, pp. 296 – 325. First part of the Essay: in *Biometrika*, vol. 45, 1958, pp. 293 – 315 and in Pearson & Kendall (1970, pp. 131 – 153).

Bernoulli D. (1766), Essai d'une nouvelle analyse de la mortalité causée par la petite vérole etc. *Werke*, Bd. 2. Basel, 1982, pp. 235 – 267.

Bernoulli J. (1690), Questiones nonnullae de Usuris cum solutione problematis de sorte aleorum propositi [in 1685]. In Bernoulli J. (1993, pp. 160 – 163). See the problem itself in Bernoulli J. (1975, p. 91).

--- (1975, 1993), *Werke*, Bde 3 – 4. Basel.

Bernoulli N. (1709), De usu artis conjectandi in jure. In Bernoulli J. (1975, pp. 289 – 326).

Bicquille C. F. (1783), *Du calcul des probabilités*. Paris.

Buffon G. L. L. (1777), Essai d'arithmétique morale. *Oeuvr. Philosophiques*. Paris, 1954, pp. 456 – 488.

Condorcet M. J. A. N. (1765), Probabilité. *Enc. raisonné des arts et des métiers*, t. 13. Stuttgart – Bad Cannstatt, 1966, pp. 393 – 400. Appeared anonymously.

--- (1785), *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. New York, 1972.

D'Alembert J. Le Rond (1761), Sur l'application du calcul des probabilités à l'inoculation de la petite vérole. *Opusc. math.*, t. 2. Paris, pp. 26 – 95.

--- (1768a), Sur un mémoire de M. Bernoulli concertant l'inoculation. *Opusc. math.*, t. 4. Paris, pp. 98 – 105.

--- (1768b), Sur les calcul relatifs à l'inoculation. *Ibidem*, pp. 310 – 341.

--- (1768c), Sur la calcul relatifs à l'inoculation, addition. *Opusc. math.*, t. 5. Paris, pp. 508 – 510.

De Moivre A. (1712, in Latin), De mensura sortis, or, the measurement of chance. *Intern. Stat. Rev.*, vol. 52, 1984, pp. 236 – 262.

--- (1718), *Doctrine of Chances*. London. Later editions 1738, 1756; reprint of last edition: New York, 1980.

- De Mora Charles Maria Sol** (1986), Leibniz et le problème des parties. *Hist. Math.*, vol. 13, pp. 352 – 369.
- Deparcieux Antoine** (1781), *Traité des annuités*. Paris.
- Deparcieux Antoine** (1746), *Essai sur la probabilité de la durée la vie humaine*. Paris.
- De Witt J.** (1671, in vernacular), Value of life annuities in proportion to redeemable annuities. In Hendriks F. (1852), Contributions to the history of insurance. *Assurance Mag.*, vol. 2, pp. 232 – 249.
- Duvillard de Durand E. E.** (1787), *Recherches sur les rentes, les emprunts, les remboursements* etc.
- Euler L.** (1749), Letter to Friedrich II. *Opera omnia*, ser. 4A, t. 6. Basel, 1986, pp. 317 – 320.
- (1862a), Vera aestimatio sortis in ludis. *Opera omnia*, ser. 1, t. 7. Leipzig – Berlin, 1923, pp. 458 – 465.
- (1862b), Reflexions sur une espèce singulière de lotterie nommée Génoise. *Ibidem*, pp. 466 – 494.
- Freudenthal H.** (1951), Das Peterburger Problem in Hinblick auf Grenzwertsätze der Wahrscheinlichkeitsrechnung. *Math. Nachr.*, Bd. 4, pp. 184 – 192.
- Haas K.** (1956), Die mathematischen Arbeiten von J. H. Hudde. *Centaurus*, Bd. 4, pp. 235 – 284.
- Halley E.** (1694), An estimate of the degree of mortality of mankind. Baltimore, 1942.
- Huygens C.** (1657), De calcul dans les jeux de hasard. *Oeuvr. Compl.*, t. 14, 1920, pp. 49 – 91.
- Kendall M. G., Doig A. G.** (1968), *Bibliography of Statistical Literature*, vol. 3. Edinburgh. Covers the period up to 1940.
- Kohli K.** (1975), Kommentar zur Dissertation von N. Bernoulli. In Bernoulli J. (1975), pp. 541 – 556).
- Lacroix S.-F.** (1816), *Traité élémentaire du calcul des probabilités*. Paris. Later editions: 1828, 1833, 1864. German transl. 1818.
- Lagrange J. L.** (1776), Sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations. *Oeuvr.*, t. 2. Paris, 1868, pp. 173 – 236.
- (1777), Recherches sur les suites récurrentes. *Oeuvr.*, t. 4. Paris, 1869, pp. 151 – 251.
- Laplace P. S.** (1812), *Théorie analytique des probabilités*. *Oeuvr. Compl.*, t. 7, pt. 1 – 2. Paris, 1986.
- (1814, in French), *Philosophical Essay on Probabilities*. New York, 1995. Transl. A. I. Dale.
- (1818), *Théor. anal. prob.*, Supplément 2. *Oeuvr. Compl.*, t. 7, pt. 1, pp. 531 – 580.
- Libri G.** (1845), Fermat. *Rev. des deux mondes*, 15 mai 1845.
- Markov A. A.** (1914, in Russian), Bicentennial of the law of large numbers. In Ondar (1977/1981, pp. 158 – 163).
- Montmort P. R.** (1708), *Essay d'analyse sur les jeux de hazard*. Paris, Second edition (1713) reprinted: New York, 1980.
- Montucla J. E.** (an X, 1802), *Histoire des mathématiques*, t. 3. Paris.
- Nicole F.** (1732a), Examen et résolution de quelques questions sur les jeux. *Hist. Acad. Roy. Sci. Paris* pour 1730, pp. 45 – 56.
- (1732b), Méthode pour déterminer le sort de tant de joueurs que l'on voudra etc. *Ibidem*, pp. 331 – 344.
- Ondar Kh. O.** (1971, in Russian), *Correspondence between A. A. Markov and A. A. Chuprov*. New York, 1981.
- Pearson E. S., Kendall M. G.** (1970), *Studies in the History of Statistics and Probability*. London.
- Poisson S.-D.** (1830), Formules des probabilités [...] qui peuvent être utiles dans l'artillerie. *Mémorial de l'Artillerie*, No. 3, pp. 141 – 156.
- (1837a), *Recherches sur la probabilité des jugements* etc. Paris, 2003.
- (1837b), Sur la probabilité du tir à la cible. *Mémorial de l'Artillerie*, No. 4, pp. 59 – 94.
- Price R.** (1771), *Observations on Reversionary Payments* etc, vols 1 – 2. London, 1783. Many later editions.
- Prokhorov Yu. V., Editor** (1999, in Russian), *Veroiatnost i Matematicheskaia Statistika. Enzikopedia* (Probability and Math. Statistics. An Encyclopedia). Moscow.
- Sheynin O. B.** (1991), On the work of Buniakovsky in the theory of probability. *Arch. Hist. Ex. Sci.*, vol. 43, pp. 199 – 223.

Steklov V. A. (1924, in Russian), A. A. Markov. *Izvestia Ross. Akad. Nauk*, ser. 6, vol. 16 for 1922, pp. 169 – 184.

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

Wargentin P. (1754 – 1755), Om nyttan af Forteckningar pa födda och döda. *Kgl. Vetenskaps Acad.*, tt. 15 – 16.

XVIIb

A. Ya. Boiarsky, E. M. Andreev

The Buniakovsky Method of Constructing Mortality Tables

Demografichesky Enziklopedichesky Slovar

(Demographic Encyclopedic Dictionary). Moscow, 1985, pp. 50 – 51

The method is based on using information on the number of deaths during a certain period of time grouped according to age and year of birth, and on the number of yearly births. Buniakovsky suggested it in 1864 and, using it, calculated a mortality table for the male, then for the female Orthodox population of Russia by issuing from the data on deaths in 1862 and births in 1796 – 1862.

The Buniakovsky method is a further modification of the method of lists [bills] of mortality and ensures an adequate measurement of the level of mortality of so-called closed populations (without migration) having an invariable order of extinction.

The initial indicator for calculating mortality tables according to the Buniakovsky method is the number of deaths d_x assumed equal to the ratio of deaths at a given age x to the number of births x years ago.

Migration essentially corrupts such calculations which, furthermore, cannot be attributed either to a certain period or to a certain generation, and there is no guarantee that the equality $\sum d_x = 1$ holds. The Buniakovsky method had been therefore only applied in the absence of censuses for calculating a few mortality tables.

For younger children the data on their deaths are sometimes more easily comparable with the data on births than with the results of censuses which are used for constructing mortality tables by the demographic method. The Buniakovsky method is therefore sometimes applied in such mortality tables calculated by the demographic method for the earliest ages (for example, up to five years of age). When necessary data are available, and independently from the method of construction of mortality tables, the Buniakovsky method is used for determining the coefficient of infant mortality.

XVII
A. M. Liapunov

Pafnuty Lvovich Chebyshev

First published 1895

P. L. Chebyshev, *Izbrannye Matematicheskie Trudy* (Sel. Math. Works).
Moscow – Leningrad, 1946, pp. 9 – 21.

[1] Academician P. L. Chebyshev died 26 November 1894 and in his person science has lost one of the greatest geometers of this expiring century. Chebyshev's investigations explained many difficult issues in analysis, established connections between differing heterogeneous theories and paved the way for solving many important problems not yielding to usual treatment. In a note (Markov & Sonin 1895) devoted to the memory of the late scholar, we find, in particular, an excellent testimonial about his works:

Chebyshev's works bear the imprint of greatness. He derived new methods for solving many difficult problems proposed long ago and had been remaining unsolved. And he also formulated a series of new important issues with whose development he busied himself to his last hour.

Owing to the originality of Chebyshev's investigations, he rarely had to mention research made by others¹, whereas other scientists ever oftener cite our glorious fellow member [of the Academy] and are drawing their ideas from that rich treasury of Chebyshev's works.

It is impossible to assess properly the significance of the great scholar without deeply analysing all of his works, and I do not make so bold as to take on this goal which [moreover] nowadays cannot be done in any satisfactory manner.

The great ideas scattered over Chebyshev's works are undoubtedly not only not exhausted in all their conclusions, they can only bear adequate fruit in the future, and only then a possibility will present itself for properly comprehending the great significance of the scholar recently lost by science. Here, I only wish to explicate the known to me facts from Chebyshev's life and scientific activity, to indicate his most important investigations and to make public some of my personal recollections of Chebyshev as professor.

Chebyshev, who belonged to an old noble family, was born 14 May 1821 in the small village Okatovo, Borovsky uyezd (district) of Kaluga province, in his mother's estate. He studied at home, then entered Moscow university where, in 1841, he achieved his candidate's degree in mathematical sciences. When continuing his scientific learning under Professor Brashman, his talented and original investigations soon attracted attention.

[2] If I am not mistaken, Chebyshev's name appeared for the first time in print in 1843, when he published a note on multiple integrals. Two years later he published his *Essay on an Elementary Analysis of the*

Theory of Probability and defended it in 1846 at Moscow University as a master's dissertation. Thus, during this early period, Chebyshev's attention was drawn to that theory. It had not failed to interest him later, and owes him very important acquisitions.

Soon after the defence, the Petersburg University had invited Chebyshev to the chair of mathematics, and he moved there. His most important investigations began to appear and it was not long before the entire mathematical world became acquainted with his name. In 1848 he submitted a very important memoir on the number of primes not exceeding a given boundary to Petersburg University; it appeared four years later (1849b). In 1849 he published his *Theory of Congruences* (1849a) and defended it as a doctoral dissertation. This excellent writing later became a manual for a large number of generations of the studying young men and still remains the best source of information on the branch of number theory expounded there.

In 1850 Chebyshev submitted to the Academy his celebrated *Mémoire sur les nombres premiers* (1852) where he solved many important and very difficult problems of the number theory. He described a method of determining the higher and the lower bounds for the sum of the logarithms of primes existing between given limits and indeed indicated such bounds which he made use of, in particular, for deriving some conclusions about the number of primes contained between given limits. There also he solved problems on the convergence and approximate summation of series whose terms were determined as values of given functions corresponding to the values of the independent argument taken from a series of primes.

[3] Chebyshev had attracted attention even earlier, and now he achieved fame of a first-class geometer by his research into number theory. It finally consolidated after the publication of a series of important memoirs on algebraic and logarithmic integrability of differentials containing irrational functions.

The first of these memoirs (1853) was devoted to the determination of the logarithmic part of the integral

$$\int \frac{f(x)}{F(x) \sqrt[m]{\theta(x)}} dx$$

when it can be expressed in a closed form. Here, the functions involved are any integral rational functions and m , any integral positive number.

But especially important work of the same kind (1860) belonging to the integration of elliptical differentials appeared somewhat later, at first in an Academic edition, then in the *J. math. pure appl.* (1860). There, Chebyshev very importantly supplemented the works of Abel. He showed that if the integration of any elliptical differential is possible in a closed form, it is reduced to integrating differentials of the type

$$\frac{(x+A)dx}{\sqrt{x^4 + \alpha x^3 + \beta x^2 + \gamma x + \delta}}. \quad (1)$$

Already Abel proved that the possibility of integrating these taken with a proper value for the constant A depended on the periodicity of the

continued fraction into which the denominator of (1) is expanded. Chebyshev, however, did not stop here. Noting that, if the answer was in the negative, the described method cannot lead to any conclusion and wishing to make up this deficiency in Abel's work, he (1861) offered a new method which allowed to solve the problem either in the positive, or negative sense by means of a finite number of algebraic operations if only α , β , γ and δ were real rational numbers.

Among non-elliptical differentials Chebyshev considered in more detail those that included a cubic root (1865) and he devoted some of his later memoirs to their integration.

[4] At the same time Chebyshev's memoirs concerning another field began to appear. There, he proposed and solved quite special problems connected with approximate representation of functions and, having hardly any predecessors, he was entirely original. I am only acquainted with the first part of the memoir (1854) and do not know whether its concluding part was ever published. This first part is only introductory and contains a solution of an analytical problem, very important for the theory which Chebyshev thought of developing further. It concerned such an approximate representation of a function between two given bounds by a polynomial of a given degree that the maximal error between these bounds was as small as possible. His results are very remarkable, but even more so is the method that yielded them. Possibly excluding one proposition on which it is based, it is entirely his own. Indeed, no general methods for solving such problems existed, and, apart from several simplest particular cases considered by Poncelet, Chebyshev was unable to find any examples of solution of such problems in the writings of his predecessors.

Chebyshev did not stop after solving the described analytical problem. Those concerning approximate representation of one or another kind of a given function considered between given bounds under the same condition, can obviously have very diverse and important applications. In addition, their solution demands special methods and they are therefore extremely interesting from an analytical point.

All this compelled Chebyshev to undertake more general studies in the same direction and thus appeared his remarkable memoir (1858) in which he put forward general conditions for solving such problems. And so it occurred that Chebyshev laid the foundation of his theory of functions "least deviating from zero" [of constructive theory of functions].

The memoir of 1855 belongs to the same kind of his writings. There, Chebyshev indicated important properties of continuous fractions as applied to developing functions in series and provided a general formula for interpolating by the method of least squares. To the same direction also belong memoirs on interpolation including two (1859a; 1859b) the first of which contains special formulas for interpolating given vast observational data and the second one expounds in detail the methods of interpolation following from the considerations provided earlier (1855).

For Chebyshev, all these memoirs served as points of departure for many further studies. There, he more fully developed his methods and discovered their diverse applications both to some related problems of kinematics of mechanisms which especially attracted the illustrious scholar's attention.

[5] I will not list the items of the long series of his subsequent memoirs but I ought to mention some of them, especially important either for the problems dealt there or because of the methods applied. Chebyshev's memoir (1867a) expounds the points of departure for solving the same kind of problems concerning sums as those which involve integrals and are treated by calculus of variations.

Another memoir of the same year (1867b) contains a rigorous elementary proof of an important proposition of the theory of probability which includes as a particular case the so-called law of very [!] large numbers. This short contribution is especially remarkable owing to the idea underlying the proof which later led Chebyshev to propose an important analytical problem.

In his memoir (1874) Chebyshev posed an entirely special problem, very remarkable on many counts, about maximal and minimal values. He could have been prompted here by an investigation due to Bienaymé although he was also led to it in a natural manner by many of his own studies. Here it is. To find the extreme values which the integral

$$\int_a^b f(x)dx$$

can possibly take for $a < b$ given the values of the integrals

$$\int_A^B f(x)dx, \int_A^B xf(x)dx, \int_A^B x^2 f(x)dx, \dots, \int_A^B x^m f(x)dx$$

considered in wider intervals $A < a, B > b$ with the function $f(x)$ not being negative between $x = A$ and $x = B$.

There also Chebyshev indicated an important proposition revealing the connection of the mentioned problem with developing the integral

$$\int_A^B \frac{f(x)dx}{z-x}$$

into a continuous fraction and formulated the final conclusion for an interesting particular case.

Many of his later memoirs were devoted to that problem concerning integrals and to related problems. One of them contained an application of his conclusions to the proof of an important theorem in the theory of probability (1887). Chebyshev's last memoir (1895) deals with the same kind of problems; he presented it to the Academy on 16 February of the past year, and it appeared only recently and posthumously.

Chebyshev's work on number theory and integral calculus is important both because of the difficult problems treated there and for the methods which he devised for their solution. There, he followed his great predecessors, Euler, Legendre and Abel who proposed these problems. However, as regards approximate representation of functions and various original problems on maxima and minima, Chebyshev is the inventor of methods for solving very important and difficult problems which he

himself had proposed. Here, he opened up many entirely new issues and indicated new paths and a new direction in science. Although in general his deep ideas are far from being exhausted, it is especially true as regards those contained in the works belonging to the last-mentioned kind.

[6] The problems with which the mentioned most important research had been concerned were not the only ones attracting the attention of the great scholar. He was interested in very many issues of pure and applied mathematics. Thus, he busied himself with construction of geographical maps, special kind of problems on deformation of surfaces, with many problems of practical mechanics. Theoretical mechanics also attracted him, and I, for example, know that he investigated some issues in hydrostatics. In 1884, during one of our last encounters, he told me that, among other issues, he had studied the problem of the ring-shaped form of equilibrium of fluid rotating mass whose particles were mutually attracted according to the Newtonian law. And, as far as I remember, he wished to publish his related investigations. Other work probably prevented him and it can only be hoped that something concerning these interesting studies will be found in his posthumous papers.

Various scientific institutions and societies, both Russian and foreign, have long ago recognized and highly appreciated Chebyshev's scientific merits. Apart from many scientific societies considering him their member, I only mention the Petersburg and the Paris academies of sciences. In 1853, the former elected him adjunct (junior scientific assistant), chair of applied mathematics, and from 1859 he had been full academician. In 1860 he became corresponding member, and in 1874, *membre associé étranger* of the latter. This last-mentioned circumstance clearly shows how highly did that scientific institution appreciate Chebyshev's merits; indeed, very few of the most distinguished foreign scientists have been honoured by election to *membre associé étranger*.

[7] Until now, I spoke about Chebyshev the mathematician, but his name is also very honourably known in another field, in that of inventions of mechanisms. From his youth up, he had been especially inclined to such inventions and all his life they did not fail to interest him. In the first place they were mechanisms transforming circular motion into rectilinear motion. Such mechanisms invented by him have already been practically applied in various ways. His parallelograms do not provide precise rectilinear motion, but they only insignificantly deviate from it and successfully replace "precise" mechanisms of the same kind and, being simpler, are preferable to them.

Chebyshev explicated the theory of his mechanisms in a series of memoirs especially interesting because they reveal the connection between his invention of this kind and analytical research on approximate representation of functions.

He made many more inventions one of which ought to be especially mentioned. About 1878 he invented the arithmometer which may be considered the most perfect of all the existing machines of that kind. Its model is being kept at the Conservatoire des arts et métiers in France. Chebyshev himself described many of his inventions and delivered many reports about them both in Russia and in Paris and London.

[8] A scholar and inventor of genius, Chebyshev had also been an exemplary professor. His professorial activities began, as I mentioned

above, in 1847 and had continued without interruption until 1882 when he left the university and became until his last days exclusively engaged in scientific research.

During different periods of his professorial activities, Chebyshev delivered differing courses. When, at the end of the 1870s, I had been a student, he gave lectures in number theory, the theory of definite integrals and calculus of finite differences for third-year students and the theory of probability for those of the fifth year. His courses were not extensive, and he took care not so much about the amount of reported material as rather about clearing up the treated issues in principle.

His lectures were lively and captivating and accompanied by many interesting remarks on the significance and importance of some problems or scientific methods. Sometimes these were expressed in passing and referred to a particular case, but they invariably left a deep impression. His lectures therefore brought about a lofty development and after each of them his listeners acquired a wider opinion and new viewpoints.

Chebyshev hardly missed a lecture. I had been his listener for two years, and at least I do not remember that during that time he failed to come even once. He had always been appearing exactly on time, beginning at once, without wasting a single second, to continue the deliberations of the previous lecture. He calculated extremely speedily and, in spite of being an excellent calculator, he therefore often made mistakes so that it was necessary to follow his computations very attentively and to warn him about the mistakes on time which he always asked us to do.

And, when finally arriving at the desired goal, Chebyshev sat down in an armchair that was always placed for him near the first row of desks. Then it was that he expressed various remarks which all the listeners were awaiting impatiently and which ensured that his lectures had been especially interesting. Very often Chebyshev made known his opinion about some related works, sometimes he recalled incidents occurring during his travels abroad and related his conversations with some foreign scientist.

After a more or less long talk of such nature which served him as a respite, Chebyshev, who was quite quick both in his speech and in all of his actions, jumped up, took the piece of chalk and began to continue his lecture. He dropped the chalk as soon as the bell rang at whichever point of his deliberations.

His professorial activities at Petersburg University, where he had been teaching for 35 years, could not have failed to effect most beneficially the entire staff of the mathematical faculty whose chairs had become occupied by his most talented former students. The elevated position which that faculty had long ago occupied at the University is therefore understandable.

However strongly had Chebyshev influenced the university, his main merit as professor was the creation of that mathematical school which is known under his name and is distinguished by a special direction of research. Chebyshev's students have continued to develop his invented methods and, while solving the problems posed by him, are proposing new problems of the same kind. Thus, gradually, new sections with which his name will be forever connected are being created in science. At the same

time, the work of his followers are ever wider spreading those viewpoints to which the great scholar had been keeping in all his investigations.

[9] The partisans of Riemann's extremely abstract ideas delve ever deeper into function-theoretic research and pseudo-geometric investigations in spaces of four, and of more than four dimensions. In their investigations they sometimes go so far that any present or future possibility of seeing the meaning of their work with respect to some application is lost, whereas Chebyshev and his followers invariably stick to reality and are guided by the viewpoint that only those investigations are valuable that are called forth by applications, whether scientific or practical, and only those theories are really useful which are occasioned by considering particular cases².

An elaboration of issues especially important for applications and at the same time presenting unusual theoretical difficulties, demanding invention of new methods and elevation to the principles of science, followed by generalization of conclusions obtained and creation of a more or less general theory – this is the direction of most of the work of Chebyshev and of the scientists who have adopted his viewpoints.

The entire scientific activity of Chebyshev, who had proposed and solved quite new and important issues of analysis by starting from applied and sometimes purely practical problems, clearly shows how fruitful can such a direction be in a purely scientific aspect. Such, however, is the path of many important discoveries in the mathematical field.

I am concluding my note by wishing that the preparations of Chebyshev's complete works be started as soon as possible. Their study is now very difficult because of their scatter over various periodicals, some of which are rather rare. Indeed, the acquaintance with these works can be necessary in many different cases and their study can show the way for many new discoveries because the ideas of the great scholar can become extremely important for solving many difficult problems now awaiting their turn.

Notes

1. This is not the whole story. In a private letter of 1885 to A. P. Karpinsky (a geologist and public figure, 1847 – 1936), Markov made known that Chebyshev (1885) had failed to refer to him, and “continues in the same vein” (Grodzensky 1987, pp. 62 – 63).

2. These investigations were recently often connected, but have nothing in common with Lobachevsky's geometric research. Like Chebyshev, the great geometer always remained on solid ground and would have hardly seen these transcendental investigations as a development of his ideas. A. L.

And here is a modern comment (Novikov 2002, p. 330): “In spite of his splendid analytical talent, Chebyshev was a pathological conservative”. He corroborated his opinion by a reference to V. F. Kagan, an eminent geometrician (1869 – 1853). The latter, “when being a young Privat-Docent”, had listened to Chebyshev's scornful statement on the “trendy disciplines like the Riemann geometry and complex-variable analysis”.

Bibliography

P. L. Tchébichev (Chebyshev)

With reference to the volumes of his *Oeuvres* (1899 – 1907) and his Russian *Complete Works* as listed in its volume 5 (1951), in that order. Thus, [– ; 5, pp. 26 – 87] means that the appropriate work was not included in the *Oeuvres*, but was reprinted in vol. 5 of the *Complete Works*. Indication *In Russian* certainly does not apply to reprints in French

journals or in the *Oeuvres*. Finally, I provide the year of the first publication in accord with the *Complete Works* rather than with Liapunov's essay. Abbreviation: *J. de math. pures et appl.* = JMPA

(1843), Note sur une classe d'intégrales définies multiplies [1, pp. 3 – 6; 2, pp. 5 – 7]. JMPA, sér. 1, t. 8, 1843, pp. 235 – 239.

(1845), *Opyt Elementarnogo Analiza Teorii Veroiatnostei* (Essay on an Elementary Analysis of the Theory of Probability). Moscow. [–; 5, pp. 26 – 87].

(1849a), *Teoria Sravnenii* (Theory of Congruences). Petersburg. Third edition, 1901. German transl., Berlin, 1888; Italian transl., Roma, 1895. [–; 1, pp. 10 – 172].

(1849b), Sur la fonction qui détermine la totalité des nombres premiers inférieurs à une limite donnée. In (1849a, pp. 209 – 229), also in later editions of that source, also in *Mém. présentés à l'Acad. Imp. des Sci. de St.-Pétersbourg par divers savants*, t. 6, 1851, pp. 141 – 157 and in JMPA, sér. 1, t. 17, 1852, pp. 341 – 365. [1, pp. 29 – 48; 1, pp. 173 – 190].

(1852), Mémoire sur les nombres premiers. JMPA, sér. 1, t. 17, 1852, pp. 366 – 390. [1, pp. 51 – 70; 1, pp. 191 – 207].

(1853), Sur l'intégration des différentielles irrationnelles. JMPA, sér. 1, t. 18, 1853, pp. 87 – 111. [1, pp. 147 – 168; 2, pp. 52 – 69].

(1854), Théorie des mécanismes connus sous le nom de parallélogrammes (première partie). *Mém. présentés à l'Acad. Imp. des Sci. de St.-Pétersbourg par divers savants*, t. 7, 1854, pp. 539 – 568. [1, pp. 111 – 143; 2, pp. 23 – 51].

(1855), Sur les fractions continues. *Uchenye Zapiski Imp. Akad. Nauk*, t. 3, 1855, pp. 636 – 664; JMPA, sér. 2, t. 3, 1858, pp. 289 – 323. [1, pp. 203 – 230; 2, pp. 103 – 126].

(1858), Sur les questions de minima qui se rattachent à la représentation approximative des fonctions. *Bull. Cl. phys.-math. Acad. Imp. Sci. St.-Pétersbourg*, t. 16, 1858, pp. 145 – 149. [1, pp. 705 – 710; 2, pp. 146 – 150].

(1859a), Sur l'interpolation dans le cas d'un grand nombre de données fournies par les observations. *Mém. Acad. Imp. Sci. St.-Pétersbourg*, sér. 7, t. 1, No. 5, 1859, 81 pp. [1, pp. 387 – 469; 2, pp. 244 – 313].

(1859b), Sur l'interpolation par la méthode des moindres carrés. *Ibidem*, No. 15, 24 pp. [1, pp. 473 – 498; 2, pp. 314 – 334].

(1860), Sur l'intégration des différentielles irrationnelles. *C. r. Acad. Sci. Paris*, t. 51, 1860, pp. 46 – 48. JMPA, sér. 2, t. 9, 1864, pp. 242 – 247. [1, pp. 511 – 514; 2, pp. 342 – 344].

(1861, in Russian), Sur l'intégration de la différentielle... *Bull. Cl. phys.-math. Acad. Imp. Sci. St.-Pétersbourg*, t. 3, 1861, 12 pp. [1, pp. 517 – 530; 2, pp. 345 – 357].

(1865), Sur l'intégration des différentielles les plus simples parmi celles qui contiennent une racine cubique. *Prilozenie (Supplement) No. 5*, t. 7, 1865, *Zapiski Akad. Nauk*. [1, pp. 563 – 608; 2, pp. 375 – 411].

(1867a, in Russian), Des maxima et minima des sommes composées de valeurs d'une fonction entière et de ses dérivées. *Prilozenie (Supplement) No. 3*, t. 12, 1867, *Zapiski Akad. Nauk*, 47 pp.; JMPA, sér. 2, t. 13, 1868, pp. 9 – 42. [2, pp. 3 – 40; 2, pp. 438 – 466].

(1867b), Des valeurs moyennes. JMPA, sér. 2, t. 12, 1867, pp. 177 – 184. [1, pp. 687 – 694; 2, pp. 431 – 437].

(1874), Sur les valeurs limites des intégrales. JMPA, sér. 2, t. 19, 1874, pp. 157 – 160. [2, pp. 183 – 185; 3, pp. 63 – 65].

(1885, in Russian), Sur la représentation des valeurs limites des intégrales par des résidus intégraux. *Prilozenie (Supplement) No. 4*, t. 51, 1885, *Zapiski Imp. Akad. Nauk; Acta Math.*, t. 9, 1886, pp. 35 – 56. [2, pp. 421 – 440; 3, pp. 172 – 190].

(1887, in Russian), Sur deux théorèmes relatifs aux probabilités. *Prilozenie (Supplement) No. 6*, t. 55, 1887, *Zapiski Imp. Akad. Nauk; Acta Math.*, t. 14, 1890 – 1891, pp. 305 – 315. [2, pp. 481 – 492; 3, pp. 229 – 239].

(1895), Sur les sommes qui dépendent des valeurs positives d'une fonction quelconque. *Zapiski Imp. Akad. Nauk*, ser. 8, t. 1, 1895, No. 7. [2, pp. 681 – 698; 3, pp. 373 – 390].

(1899 – 1907, Russian and French), *Oeuvres*, tt. 1 – 2. New York, 1962.

(1951, in Russian), *Polnoe Sobranie Sochinenii* (Complete Works), vol. 5. Moscow – Leningrad.

Other Authors

- Grodzensky S. Ya.** (1987, in Russian), *A. A. Markov*. Moscow.
- Markov A. A., Sonin N. Ya.** (1895, in Russian), On Chebyshev's collected works. *Izvestia Peterburg. Akad. Nauk*, vol. 2, No. 1, pp. X – XI.
- Novikov S. P.** (2002, in Russian), The second half of the 20th century and its result: the crisis of the physical and mathematical associations in Russia and in the West. *Istoriko-Matematich. Issledovania*, vol. 7 (42), pp. 326 – 356.

XVIII
A. A. Konüs

On the Definition of Mathematical Probability

Problemy Teorii Statistiki. (Issues in the Theory of Statistics).
Moscow, 1978, pp. 64 – 77

[1] The classical definition of probability is formulated in the following way (Druzinin 1970, p. 19; see also Kantorovich 1946):

Mathematical probability is expressed by a proper fraction representing the ratio of the number of cases favourable for the occurrence of event A to the general number of equally possible and exclusive cases.

At the beginning of the 1950s there occurred an opinion (Arley & Buch 1950, p. 10) that that definition

Is logically a circle definition, since equally likely can be defined only as equally probable, which is to be defined.

Such reasoning first stated at the beginning of the century by the author of the *frequentist* definition of probability (Mises 1928) became widespread (Gottinger 1974), but it cannot be considered well-founded. Indeed, if probabilities of random events are magnitudes subject to measurement, then, when comparing them with each other, it is not needed to establish the unit in which they are measured, it is sufficient to be able to judge whether one is higher or lower than, or equal to the other. This is known to be ensured by the preparation of the experiment or observation (Konüs 1954, p. 12).

Khinchin (1961/2004, p. 397) and Glivenko¹ had been our main proponents of the idea of equipossibility in the theory of probability:

We may consider this definition [of probability by means of equally possible cases] as a reduction of the problem of finding a quantitative measure of probability in general to a preceding notion of equiprobability of events; the vicious circle thus disappears and the definition itself acquires some scientific meaning.

Glivenko (1939, p. 18) expressed himself more resolutely:

The notion of equiprobable is comprehensible to all and, like, for example, the concept of real space, does not need a logical definition. We discuss the equiprobability of two or several events if we have sufficiently objective pertinent grounds resting in data provided by practice, just like we mention for example the linearity of a ray of light.

Critics of the classical definition of probability indicate (Khinchin 1961/2004, p. 397) that it has

An extremely restricted sphere of application. Having originated and been developed due to games of chance and simplest insurance operations, the old theory of probability had built for itself a basis fit to a certain degree for treating these simplest problems; however, once the sphere of its problems had extended in connection with the requirements of physical and social statistics, and, later on, of biology and technology, the initially adopted foundations became too narrow. In problems reaching beyond the realm of games of chance, those equally possible cases, without which the classical concept cannot even speak about probabilities, just do not exist. Mises' celebrated example of an irregular die is unsurpassed in validity and simplicity of argumentation².

However, in some important theoretical problems and in many practical applications the classical definition of probability entirely preserves its significance. Therefore, it is indeed included with some reservations in most treatises on the theory of probability. True, the axiomatic approach to the theory as developed by Kolmogorov (1933) is compulsory for contemporary mathematical treatises intended for university students.

[2] The initial notion in the axiomatic theory of probability is *event*. It makes sense to say that it occurs or does not occur (at present, in the past, or in the future). Abstract prototypes of operations with events are contained in the set theory (Gnedenko 1950) and the Boolean algebra (Glivenko 1939). The most important features of these operations and a number of necessary definitions are (in a simplified representation):

A group (a *set*) of events A, B, C, \dots is considered. The letters here are not magnitudes, but the events themselves. For example, when throwing a die [...] the events will be “the appearance of one”; “the appearance of two” etc. A second example concerns the state of the weather in a certain locality at 12 o'clock on Nov. 1, 1980: the events are rain, snow, [...].

In each of these examples, one of the events listed will occur without fail. Concerning such events, we say that they constitute a *complete group* and it is with them that the theory of probability has to do. [The author introduces the notions of complete group; certain event; multiplication and addition of events; product of events; exclusive events; sum of events; elementary events; complete group of exclusive events; impossible event.]

Operations of multiplication and addition can be extended to any, and even to a *countable set* [...]. An *ordering* relation can be established for two events of a given group. It is designated by $A \subset B$ which means that *A precedes B* or *B follows* after *A* (Glivenko 1939, p. 207). Under various circumstances this relation is also interpreted as “A involves B” or “A is a particular case of B”. [...] If event *A* precedes *B*, i. e. if $B \subset A$, and at the same time event *B* precedes *A*, i. e., if $B \subset A$, these events are called *equivalent* which is designated as $A = B$. [...]

Mathematical courses on the theory of probability also list a number of other important definitions and relations between events. Their exhaustive and precise description predetermines the appropriate formulation of the necessary axioms.

Axioms founded on the set theory can be formulated thus (Gnedenko 1950/1973, p. 48).

Axiom 1. With each random event A in the Borel field of events F [of the system of events satisfying some special relations partly mentioned above

– A. K.] *there is associated a nonnegative number $P(A)$ called its probability.*

Axiom 2. [The probability of a *certain* event is 1.]

Axiom 3 (Axiom of addition). *If events A_1, A_2, \dots, A_n are pairwise mutually exclusive, then*

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

This last axiom, the most important for the theory of probability, can be extended to infinitely many events:

Extended axiom of addition (Ibidem, p. 51). If event A can be decomposed into a finite or countable set of pairwise mutually exclusive events $A_1, A_2, \dots, A_n, \dots$, its probability is represented by the sum of the probabilities of events $A_1, A_2, \dots, A_n, \dots$:

$$P(A) = P(A_1) + P(A_2) + \dots + P(A_n) + \dots$$

The axiomatic approach is remarkable in that the theory of probability, being the calculation of probabilities of any events by issuing from the probabilities of the initial events, is solely based on these axioms (actually, only on the addition axiom): However, the listed axioms do not determine those initial probabilities without which it is nevertheless impossible to manage, and mathematicians have to turn to the Mises frequentist theory “when describing the necessary assumptions for applying the theory of probability to the world of real events” (Kolmogorov 1933/1936, p. 11).

[3] Mises proposed his theory as a denial of classical tradition in the theory of probability and it “defines probability of an event as the limit of frequency under an infinitely increasing number of trials (Khinchin 1952, p. 527). However, without additional assumptions it is impossible to prove that a number sequence with randomly defined terms has a limit.

Therefore, mathematical treatises on the theory of probability do not refer to the Mises definition in its exact formulation, but mention results of observations and experiments and thus restrict their discussion to less binding principles. The following reasoning is typical for contemporary treatises (Yaglom & Yaglom 1973, p. 18):

In many cases, when one and the same trial is repeated many times over under the same conditions, the frequency of the occurrence of the studied result [...] remains all the time approximately invariable, close to some constant magnitude. Thus, for a given shot and under given conditions the frequency of his hitting the target almost always as a rule remains approximately invariable and only seldom deviates essentially at all from some mean figure. This figure can certainly change with time: the person either improves his performance or forgets how to shoot. [...] It is therefore concluded that in each case there exists a certain constant magnitude objectively describing the very process of shooting, of throwing a die, of manufacturing articles, etc, around which the mean frequency of the pertinent result [...] is fluctuating during the long series of trials without deviating from it in any essential measure. This constant magnitude is called the probability of the considered event.

In other words (Kolmogorov 1956, p. 270), it is presumed that there exist such random phenomena

Peculiar in that the frequencies of their occurrence are stable, i. e. that, given a large number of recurrences of certain circumstances, they tend to occur with frequencies grouped around some normal level, around the probability.

It is not difficult to see that in essence such reasoning anticipate the law of large numbers whose proof the classics of the theory of probability considered so important. And indeed Kolmogorov (1956) argues that the law of large numbers is only necessary because “the need to specify quantitatively the statement that in “large” series of trials the frequencies of an event are “close” to its probability is quite natural”³.

At the same time, the frequentist theory, as Mises himself remarked, leads to the needlessness of axiomatisation of the theory of probability regarded as an applied science. He (Khinchin 1961/2004, p. 401)

Labels as nihilists those who want to perceive a mathematical doctrine in the theory of probability; this is why, filled with disgust and horror-stricken, he struggles against the proposition, accepted without hesitation by all advanced scientists of our day, that the theory of probability is a part of the general doctrine of functions and sets.

[4] All this leaves the student of the theory of probability with a feeling of dissatisfaction. It is impossible to agree that, “when actually analysing the notion of probability, it is not at all obligatory to aim at its formal definition” (Kolmogorov 1956, p. 275). Indeed, we have to provide three different approaches to that definition in one and the same course in the theory of probability. Here is an example (Prokhorov & Rozanov 1975, pp. 10, 11 and 132):

First approach. Consider some trial or phenomenon in which, depending on chance, the studied event A occurs or not. Suppose that the conditions of the trial (under which the considered phenomenon takes place) can be represented time and time again, so that in principle an entire series of the same trials independent one from another can be made and in each of them the event A occurs or not depending on chance.

Let n be the number of trials in such series and n(A), the number of trials leading to the occurrence of A. Then the ratio n(A)/n is called the [relative – A. K.] frequency of event A in the given series of trials. Practice indicates that for large values of n the frequencies n(A)/n in different series of trials are approximately equal to each other. There exists some value P(A) around which these frequencies are grouped:

$$P(A) \approx \frac{n(A)}{n}.$$

Second approach. If the outcomes of [if the favourable and unfavourable cases in] a considered trial are equally probable, the

probability $P(A)$ of event A connected with that trial can be calculated according to the simple formula

$$P(A) = \frac{N(A)}{N}$$

where N is the total number of equiprobable and mutually exclusive outcomes, and $N(A)$ is the number of those which lead to the occurrence of A .

The term *equiprobable* is lacking in Kolmogorov (1933; 1956) and Neyman (1950) so that the above definition should have perhaps been formulated differently or altogether abandoned.

Third approach. Each stochastic pattern is based on the so-called space of elementary events $(\Omega; \tilde{A}; P)$, a measurable space of elements ω called elementary events or elementary outcomes with a probability measure $P(A)$ given on the σ algebra \tilde{A} :

$$P(\Omega) = 1.$$

The sets of space Ω are called events and the measure $P(A)$ of the set $A \subset \tilde{A}$ is called the probability of event A .

Three definitions of mathematical probability appearing in one and the same book means that neither is of full value. This is the reason why Gnedenko (1950/1973, p. 43), when describing the frequentist definition of probability, specifies that it is not the mathematical, but the *statistical* probability. He also considers it necessary to indicate that the problem of defining probability is still open:

The statistical definition of probability given here is descriptive rather than formally mathematical in character. It is deficient in yet another aspect as well: it does not lay bare the actual peculiarities of those phenomena for which the frequency is stable. This is to stress the necessity of further investigations in the indicated direction.

[5] For ascertaining *the actual peculiarities of those phenomena for which the frequency is stable*, it is natural to provide practical examples which are usually mentioned when justifying the frequentist definition of probability. [...] It is not difficult to see that, when a man is target shooting, he naturally wishes that his next attempt will not be worse than the previous one. However, for the frequency of hitting the target to be approximately the same for a given shot and under given conditions (Yaglom & Yaglom, see above), and for the frequentist definition of probability to remain valid, the man should not level his gun each time better than before. This means that the probability of hitting the target ought to remain invariable. When manufacturing goods, all measures are certainly taken for their quality to be not worse than previously. But, again, for the frequentist definition of probability to be valid, the manager should not allow any improvement of quality and thus ensure an equal probability of each item produced to be substandard.

As to the *remarkable* (Mises) example of an irregular die, whose centre of gravity does not coincide with its geometric centre, its importance for justifying the frequentist theory of probability consists exactly in that it is easy to ensure here a constant probability of a certain outcome. The idea of a constant probability is seen in the proof of the law of large numbers as provided, for example, by Glivenko (1939, p. 111):

Let μ be the number of occurrences of event A in n independent trials and p – the constant probability of A in one separate trial⁴. Then

$$P\left(\frac{\mu}{n} \rightarrow p\right) = 1.$$

The axiomatisation of the theory of probability corroborated the considerations above. The axioms founded on the set theory allow us to derive as a corollary the statement that (Gnedenko 1969, pp. 21 and 51)

$$\text{If } A = B, \text{ then } P(A) = P(B)^5.$$

When elucidating axiomatics based on Boolean algebra, Glivenko (1939, p. 27) thought it necessary to isolate that corollary as a special axiom. [...]

[6] Such differing attitudes to one and the same statement is explained by the different meanings of the symbol $A \subset B$ which is the starting point here: in set theory it denotes that “the occurrence of the event B necessarily follows from the occurrence of A” (Kolmogorov 1936, p. 13) whereas in the Boolean algebra it means that “A precedes B” or “B follows after A” (Glivenko 1939, p. 207).

Yaglom & Yaglom (1973, p. 65) believe that

The connection of the theory of probability with Boolean algebras can serve as the foundation of the very definition of its subject. Indeed, we may say that The theory of probability studies the totality of objects forming a normed Boolean algebra. These objects are called events, and the norm $P(A)$ of event A is called probability.

I will keep to the Boolean interpretation and consider A and B equivalent, if, at the same time, A precedes B and B precedes A. The statement that “equivalent events are equiprobable” and the extended addition axiom allow us to extend the classical definition of probability to a countable set of events and at the same time to justify firmly the frequentist theory⁶.

Suppose that a countable set

$$E_1, E_2, \dots, E_n, \dots \tag{3}$$

constituting a complete group of pairwise exclusive events equivalent to one another is given. We will consider an event A which can be subdivided into a countable set of particular cases

$$G_1, G_2, \dots, G_m, \dots \tag{4}$$

included in the group (3).

Because of the extended addition axiom of addition and axiom 2

$$P(E_1) + P(E_2) + \dots + P(E_n) + \dots = 1. \quad (5)$$

It follows now from the statement above that

$$P(E_1) = P(E_2) = \dots = P(E_n) = \dots = 1/n \quad (6)$$

and, again on the strength of that extended axiom,

$$P(A) = P(G_1) + P(G_2) + \dots + P(G_m) + \dots \quad (7)$$

Since events (4) are particular cases of the complete group (3), equalities (6) and (7) lead to

$$P(A) = \frac{m}{n}, \quad m = n(A), \quad n \rightarrow \infty. \quad (8)$$

[7] Thus, mathematical probability can only be defined when it is possible to indicate a complete group of exclusive and equivalent events. In other words: axiomatics allows us to conclude that mathematical probability is a proper fraction whose denominator is the number of all and only possible exclusive and equivalent particular cases (a finite or countable set of equivalent elementary events constituting a complete group) and the numerator is the number of those particular cases of the denominator in which the given event is occurring.

The reservation “equivalent elementary events”, as compared with the classical definition of probability, means that the order of those events can change in any manner. Consider a die, a regular cube made from homogeneous material. We may believe that the face “one” is the same as and, since the numbering of the faces is arbitrary, precedes face “two”. Then the outcomes “one”, “two”, etc are equivalent in the sense of a Boolean algebra and constitute a complete group of six pairwise exclusive events. It will therefore be necessary to apply the axiom of addition for a finite number of events and substitute in formula (8) values $n = 6$ and $m = 1$.

The faces of an irregular die in the *celebrated* Mises example can be arranged in a certain order of their distance from its centre of gravity; this order is not arbitrary and for a given die it cannot be changed. The different outcomes are therefore not equivalent and it is now impossible to determine their probabilities directly, before experimenting. However, it is not difficult to establish that for numerous and like throws of an irregular die the throws themselves are equivalent elementary events.

The first throw, E_1 , precedes the second one, E_2 , so that $E_1 \subset E_2$. The numbering of the events, of the like throws, is arbitrary and the second one may be considered as preceding the first one, so that $E_2 \subset E_1$. These elementary events, the throws, are pairwise exclusive and, for a countable set, they constitute a complete group. Therefore, $E_1 = E_2$ and $P(E_1) = P(E_2)$. The same reasoning applies to any pair of throws; in formula (8) n

will represent the total number of throws, and m , the number of them with the given outcome.

Theoretically, the number of throws is infinitely large but practically it is always finite, so that a complete group does not occur here. The probability of a given outcome determined from the experiment will always only approximate the true probability:

$$\frac{m}{n} \approx P(A), m = n(A). \quad (9)$$

The bounds characterizing the precision of this estimate of the probability are determined for various values of n by formulas of the theory of probability pertaining to the reciprocal law of large numbers as formulated, for example, by Bernstein (1946), see also Benedetti (1976, pp. 344 – 493 [?])⁷.

In the example of one and the same person shooting at a target equivalence can only be understood conditionally because he improves his performance with time and it not possible to exchange a man after only a few of his shots for another person who had made several hundred of them. And the number of shots is not arbitrary for the gun being used either. This means that the estimation of the probability (9) of hitting the target is less precise than stipulated by the appropriate formulas of the theory of probability.

Considering the example of forecasting the weather for determining the probability of rain etc. in a given locality on Nov. 1, 1980, we may rest on previous data if 1980 might be exchanged for 1960 or 1940 etc, i. e., if tendencies towards climatic changes are lacking. [...]

As mentioned above, the frequentist definition of mathematical probability originated as a denial of the classical definition. However, the axiomatic direction of the theory of probability made the latter suitable for establishing that very frequentist definition. This is seen as a manifestation of dialectical development of science (Konüs 1970; Kravets 1976).

Acknowledgement. I am deeply grateful to Ms N. A. Tolmacheva, the head of the sector of mathematical programming at the research institute NIEM, State Planning Committee, for valuable comments on the initial version of this paper.

Notes

1. The author many times refers to Glivenko. Kolmogorov (1941) highly appraised his work in mathematics in general and appended a list of his publications.

2. Newton (manuscript 1664 – 1666; 1967, pp. 58 – 61) was the first to discuss throws of an irregular die. He indirectly remarked that the probabilities of the various outcomes can be determined by trials. In the same manuscript he introduced a thought experiment with its result depending on calculation of geometric probability, its first indirect and unpublished appearance in the theory of probability.

3. The wording is unfortunate: the law of large numbers is “so important”, but “only” needed etc.

4. In contemporary treatises, as it seems, *constant* is simply implied.

5. Gnedenko continues:

However, and this is particularly important, in the given definition we retain the objective character of probability that is independent of the investigator. The fact that only after performing certain preliminary observations we can judge that some event has a probability does not in the least detract from our conclusions, for a knowledge of

regularities is never derived from nothing; it is always preceded by experiment and observation. Of course, these regularities existed prior to the intervention of the experimenting thinking person, but they were simply unknown to science.

6. This is a very strong declaration (although at the very end of his paper Konüs only mentions justification of the frequentist definition of probability). See my Introduction [XVIII].

7. I have not seen Benedetti (1976). Bernstein (1946, pp. 220 – 221) introduced what he called “the reciprocal of sorts of the Bernoulli theorem”. For a constant probability p and a sufficiently large number n of independent trials the probability P that the frequency of the studied event m/n satisfies the inequality

$$\left| p - \frac{m}{n} \right| \leq \varepsilon, \varepsilon > 0$$

becomes arbitrarily close to 1 if only the prior probability R of m is higher than t/n with positive t not depending on n . Bernstein then derived an estimate of P for a finite n :

$$P > 1 - \frac{3}{16n\varepsilon^4 t}.$$

I do not understand, however, why had Bernstein (and Konüs) called this the reciprocal etc. Then, obviously, for any n , $t \leq n$.

The 1973 English translation includes the following statement (p. 20):

In all considerations of probability theory, equivalent events can replace one another. [...] Any two equivalent events [are] simply identical.

Bibliography

Arley N., Buch K. R. (1950), *Introduction to the Theory of Probability and Statistics*. New York, 1953 (second printing). Original Danish edition: 1946. [1951.]

Benedetti C. (1976), *Istituzioni di statistica*. Roma.

Bernstein S. N. (1946), *Teoria Veroiatnostei* (Theory of probability). Moscow. Fourth edition.

Diewert W. E. (1987), Konüs. *The New Palgrave. Dictionary of Economics*, vol. 3, p. 62.

Druzinin N. K. (1970), *Vyborochny Metod i Ego Primenenie v Sotsialno-Ekonomicheskikh Issledovaniakh* (Sampling and Its Application in Socio-Economic Studies). Moscow.

Glivenko V. I. (1939), *Teoria Veroiatnostei* (Theory of Probability). Moscow.

Gnedenko B. V. (1950, in Russian), *Theory of Probability*. Moscow, 1973 (second printing of English edition of 1969). Edition of Russian original not indicated. [1969.]

Gottinger H. W. (1974), Review of concepts and theories of probability. *Scientia*, No. 1 – 4.

Kantorovich L. V. (1946), *Teoria Veroiatnostei* (Theory of Probability). Moscow.

Khinchin A. Ya. (1952, in Russian), Method of arbitrary functions and the struggle against idealism in the theory of probability. In *Filosofskie Voprosy Sovremennoi Fiziki* (Philosophical Issues of Contemporary Physics), Moscow, pp. 522 – 538.

--- (1961, in Russian), R. Mises' frequentist theory and the modern concepts of the theory of probability. *Sci. in Context*, vol. 17, 2004, pp. 391 – 422.

Kolmogorov A. N. (1933), *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin. English transl.: *Foundations of the Theory of Probability*. New York, 1956. [1936.]

--- (1941, in Russian), Valerii Ivanovich Glivenko (1897 – 1940). *Uspekhi Matematich. Nauk*, vol. 8, pp. 379 – 383.

--- (1956, in Russian), Theory of probability. In *Matematika, Ee Soderzanie, Metody i Znachenie* (Mathematics, Its Contents, Methods and Significance), vol. 2. Moscow.

Komlev C. L. (1991, in Russian), The Conjecture Institute. In *Repressirovannaiia Nauka* (Suppressed Science). Leningrad, pp. 163 – 180.

Konüs A. A. (1954), *Kratkie Svedenia po Teorii Veroiatnostei* (Concise Information about the Theory of Probability). Moscow.

--- (1970), Contribution to the Herman Wold Festschrift. In *Scientists at Work. Festschrift in Honour of Herman Wold*. Editors, T. Dalenius et al. Uppsala, pp. 75 – 77.

--- (1984), Definition of mathematical probability in connection with its application in economics. In *Metody Ekonomicheskikh Obosnovaniy Planovogo Razvitiya Vneshneekonomicheskikh Sviazei* (Methods of Economic Justification of a Planned Development of External Economic Ties). Moscow.

Kravets A. S. (1976), *Priroda Veroiatnosti* (The Essence of Probability). Moscow.

Krengel U. (1990), Wahrscheinlichkeitstheorie. In *Ein Jahrhundert Mathematik 1890 – 1990. Festschrift zum Jubiläum der Deutsche Mathematiker-Vereinigung*. Braunschweig/Wiesbaden, pp. 457 – 489.

Mises R. von (1928), *Wahrscheinlichkeit, Statistik und Wahrheit*. Wien. [1930.]

Newton I. (1967), *Mathematical Papers*, vol. 1. Cambridge.

Neyman J. (1950), *First Course in Probability and Statistics*. London. [1968.]

Prokhorov Yu. V., Rozanov Yu. A. (1975), *Teoria Veroiatnostei. Osnovnye Poniatia* etc. (Theory of Probability. Main Notions etc). Moscow.

Sheynin O. (2005), *Theory of Probability. Historical Essay*. Berlin, 2009. 2nd edition. Also at www.sheynin.de

Uspensky V. A., Semenov A. L., Shen A. Kh. (1990, in Russian), Can an (individual) sequence of zeros and unities be random? *Uspekhi Matematich. Nauk*, vol. 45, pp. 105 – 162. Since 1960, this periodical is being translated as *Russ. Math. Surveys*.

Yaglom A. M., Yaglom I. M. (1973), *Veroiatnost i Informatsia* (Probability and Information). Moscow. Several German editions, for example, Berlin, 1984.

Dates in square brackets denote Russian translations to which the author referred rather than to the original foreign editions (five cases); in one case (Khinchin), he referred to the original Russian edition.

XIX Oscar Sheynin

**Review of Ivar Ekeland, *The best of all possible worlds. Mathematics and destiny.* Chicago and London: University of Chicago Press, 2006, 207pp.
Unpublished**

A somewhat differing version of this review appeared in Russian (*Voprosy Istorii Estestvoznania i Tekhniki*, No. 2, 2009, pp. 211 – 213)

The main story begins with Leibniz who stated that everything is possible if not contradictory and that God had created the world by choosing the most perfect alternative. In 1740, Maupertuis explained the choice (true, only of the course of some natural physical processes) by the principle of least action (of least product of distance travelled by the velocity of motion and mass which remains constant or the least value of the appropriate integral) and applied it to justify (mistakenly) the Snell law of refraction. Euler applied the same principle for studying important problems in mechanics and physics (partly even preceding Maupertuis), introduced it into mathematics and thus, along with Lagrange, initiated the calculus of variations.

The author then describes the work of Hamilton and C. G. Jacobi (Ostrogradsky is not mentioned) who showed that the Maupertuis principle was doubtful (what is possible motion? And how to calculate the appropriate action of forces?), transferred it to the phase space (position + velocity), and finally replaced it by the principle of stationary action (the quantity of action should be insensitive to small changes in the appropriate path).

Ekeland does not here recall the earlier mentioned Fermat principle according to which light travelled along the fastest possible route. Religious and philosophical views prevailing in the 18th century were forgotten; instead, according to Poincaré and Mach, a theory had only to be fruitful but necessarily true. Regrettably, the author had not explained all this clearly enough although he obviously intended his book for a broader circle of readers. Thus, in 1752 Chevalier d'Arcy discovered that in a certain case light did not pick the shortest path, but Ekeland did not connect this mentioned fact with the new principle.

Turning his attention to randomness and rejecting its usual interpretation as intersection of two (or a few) chains of determinate events, the author suggests that reality “lies somewhere between” order and dependence of everything on everything (p. 86). He thus refuses to study randomness, and he never mentions its regularity in case of mass random events.

Instead, he considers the example of the motion of a ball on a non-elliptical billiard table. Owing to unavoidable small uncertainty of its initial conditions, the path of the ball becomes a cloud which fills a certain region. This chaos, which the author (p. 125) unfortunately compares with a game of chance, actually defies quantitative definition, and, unlike Brownian motion, cannot be stochastically studied.

Ekeland attributes the foundation of the chaos theory to Poincaré who started from the principle of stationary action distorted by perturbations, and he concludes (p. 128) that randomness (contrary to Einstein's opinion) exists at the subatomic level with the most likely paths of elementary particles corresponding to stationary action (Feynman, p. 120) and chaos governing at our scale with the principle "caught somewhere in the middle". But where can that middle exist?

The following chapters are devoted to the theory of evolution and the existing situation in the world. He somehow understands evolution as a tendency towards an equilibrium between species (not as a stochastic process, as I suggested in 1980) and does not mention Mendel. Moreover, there is a suggestion that biological evolution is chaotic, and the author should have commented on it. It is perhaps permissible to add that Lamarck (*Histoire naturelle des animaux sans vertèbres*, t. 1. Paris, 1815, p. 169) stated that the equilibrium between "universal attraction" and "L'action repulsive des fluides subtiles" was the cause of all observed facts and especially those concerning living creatures.

It would have been opportune to mention the mistaken theory of spontaneous generation of the simplest organisms which had been yet received by Lamarck, i. e., the most serious significance attributed to randomness in biology even long before Darwin.

As to our situation, "God had receded, leaving humankind alone in a world not of its choosing" (p. 180). This quote also shows Ekeland's style, as does the very first phrase of the book: "The optimist believes that this is the best of all possible worlds, and the pessimist fears that this might be the case".

The book is interesting and instructive. A special example concerns the actually not so well-known trial of Galileo: he was accused of believing that a mathematical hypothesis reflected reality, "something that mathematicians would never do". Copernicus, or rather his publisher had indeed denied this connection, but had there been other such instances? Another statement (p. 25) is however doubtful: Descartes unified geometry and algebra "thereby creating modern mathematics".

The contents of the book are not presented clearly enough and bibliographic information is simply poor. Even the "second uncertainty principle in classical mechanics" that states, that in some sense the uncertainty in the initial data of motion cannot be lessened, is without any further details attributed to Gromov, 1980. The author could have surely done much better. He is Director of the Pacific Institute of Mathematical Studies, and he put out several books including *Mathematics and the Unexpected* (1988) and *The broken Dice* (1993), both issued by the same publisher.

XX
Oscar Sheynin

Antistigler
Unpublished

Stigler is the author of two books (1986; 1999) in which he dared to profane the memory of Gauss.

I had unavailingly criticized the first one (1993; 1999a; 1999b), but not a single person publicly supported me, whereas several statisticians, only justifying themselves by arguments *ad hominem*, urgently asked me to drop that subject. The appearance of Stigler's second book showed that they were completely wrong but the same general attitude is persisting. One of those, apparently believing that a living dog was more valuable than a dead lion, is the present President of the International Statistical Institute. But to go into detail.

1) A few years ago Stigler was elected President of that same Institute (and had served in that capacity). He is now member of the Institute's committee on history to which I was also elected (chosen?) without my previous knowledge or consent. I refused to work together with him (and with Descrosières, – of all members of the Institute, see below!).

2) A periodical (*Intern. Z. f. Geschichte u. Ethik (!) der Naturwissenschaften, Technik u. Medizin*, NTM) refused to consider my proposed subject, – the refutation of Stigler. The Editor politely suggested that I should apply to a statistical periodical.

3) The Gauss-Gesellschaft-Göttingen is silent and had not even answered my letter urging them to support me.

4) Healy (1995, p. 284) indirectly called Stigler the best historian of statistics of the 20th century, and Hald – yes, Hald (1998, p. xvi) even called Stigler's book (1986) *epochal*. Epochal, in spite of slandering Gauss, of humiliating Euler (below), and of its being an essay rather than THE HISTORY (!) of statistics, as Stigler had the cheek to name it.

So much is absent in THE HISTORY, – cf. my book Sheynin (2005/2009), – in spite of which it became the statisticians' Bible, that I shall extrapolate this phenomenon by reducing it with Lewis Carroll's help *ad absurdum*:

*Other maps are such shapes, with their islands and capes:
But we've got our brave Captain to thank
(So the crew would protest) "That he's bought us the best –
A perfect and absolute blank!"*

Stigler is regarded as a demigod. *Historia Mathematica* had published a review of his book (1999). Instead of providing its balanced account, the reviewer (an able statistician; H. M. vol. 33, No. 2, 2006) went out of his way to praise, *to worship* both the book and Stigler himself.

5) *Centaurus* rejected the manuscript of my paper (1999a) initially submitted to them since the anonymous reviewer, contrary to facts and common sense, did his damndest to exonerate Stigler.

In addition to my papers mentioned above, I can now add two more publications (2005; 2006, see their Indices), but I ought to add several points here.

1. Stigler (1986, p. 145): *Gauss solicited reluctant testimony from friends that he had told them of the method [of least squares, MLSq] before [the appearance of the Legendre memoir in] 1805.*

And in 1999, p. 322, repeating his earlier (of 1981) statement of the same ilk: *Olbers did support Gauss's claim ... but only after seven years of repeated prodding by Gauss.* Grasping at straws, Stigler adds an irrelevant reference to Plackett (1972).

So what happened with Olbers? On 4.10.1809 Gauss had asked him whether he remembered that he had heard about the MLSq from him (from Gauss) in 1803 and again in 1804. Olbers apparently did not answer (or answered through a third party). On 24.1.1812 Gauss asked even more: Was Olbers prepared to confirm publicly that fact? And Olbers answered on 10.3.1812: *gern und willig* (with pleasure), and at the first opportunity. However, during 1812 – 1815 Olbers had only published a few notes on the observation of comets (*Catalogue of Scientific Literature*, Roy. Soc. London), and he therefore only fulfilled Gauss' request in 1816. (Much later Gauss, who became sick and tired of the whole dispute, mentioned that his friend had acted in good faith, but that he was nevertheless displeased by Olbers' testimony made public.)

2. Again in 1999, Stigler had deliberately omitted to mention Bessel's statement on the same subject. I discovered it while being prompted by Stigler's attitude and quoted Bessel in a paper (1993) which Stigler mentioned in 1999. Bessel's testimony, all by itself, refutes Stigler's accusation described above.

3. Stigler (1999, pp. 322 – 323) mentions von Zach, his periodical (*Monatl. Corr.*) and some material published there in 1806 – 1807 which allegedly (indirectly) proved that von Zach had not considered Gauss as the inventor of the MLSq. Stigler leaves out a review published in the same periodical in 1809 whose anonymous author (von Zach?) described the actual history of the discovery of the MLSq, see p. 191. Incidentally, I (1999a, p. 258) found von Zach's later statement in which he repeated Gauss' explanation to the effect that he, Gauss, discovered the MLSq in 1795.

4. Stigler (1986, p. 57): "It is clear [...] that Legendre immediately realized the method's potential". And, on p. 146: "There is no indication that [Gauss] saw its great general potential before he learned of Legendre's work". Stigler thus denies Gauss' well-known statement that he had been applying the MLSq since 1794 or 1795, denies simply because he is inclined to dethrone Gauss and replace him by Legendre.

5. Stigler (1986, p. 143): Only Laplace saved Gauss' first justification (in 1809) of the MLSq from joining "an accumulated pile of essentially ad hoc constructions". And how about Legendre? Stigler (1986, p. 13): *For stark clarity of exposition the presentation [by Legendre in 1805] is unsurpassed; it must be counted as one of the clearest and most elegant introductions of a new statistical method in the history of statistics.* His work (Stigler, p. 57) revealed his "depth of understanding of his method". All this in spite of two mistakes made by Legendre and lack of any demonstration of the method. Legendre alleged that the MLSq agreed with

the minimax principle, and he mentioned errors instead of residual free terms of the initial equations. And can we believe that Stigler did not know that the Gauss' proof of 1809, which allegedly almost joined "the accumulating pile" of rubbish, had been repeated in *hundreds* of books on the treatment of observations? Was it only due to Laplace?

6. Stigler (p. 146): *Although Gauss may well have been telling the truth about his prior use of the method, he was unsuccessful in whatever attempts he made to communicate it before 1805.* The first part of the phrase was appropriate in respect to a suspected rapist, but not to Gauss. As to his "attempts", Gauss had communicated his discovery to several friends and colleagues but did not proclaim it through a public crier or by a publication in a newspaper.

Other pertinent points.

7. Stigler (1986, p. 27) denounced Euler as a mathematician who did not understand statistics. After I (1993) had refuted that pernicious statement, Stigler (1999, p. 318) declared that, in another case, Euler *was acting in the grand tradition of mathematical statistics.* He did not, however, renounce his previous opinion. More: in that second case, Euler had rejected the method of maximum likelihood, because, as he put it, the result should not change whether an outlying observation be rejected or not (read: the treatment should be such that ...). Euler suggested to keep to the known and reliable method, to the mean; he had not mentioned the median although it (but not the term itself) had actually been earlier introduced by Boscovich.

8. Descrosières (1998, transl. from French) believes that Poisson had introduced the strong law of large numbers and that Gauss had derived the normal distribution as a limit of the binomial law, see my review in *Isis*, vol. 92, 2001, pp. 184 – 185. And Stigler (1999, p. 52)? He called Descrosières *a scholar of the first rank!*

9. There also, Stigler named another such high ranking scholar, Porter, and he (p. 3) also called Porter's book of 1986 *excellent.* I reviewed it (*Centaurus*, vol. 31, 1988, pp. 171 – 172) and declared an opposite opinion. In 2004 Porter published Pearson's biography, see my review in *Hist. Scientiarum*, vol. 16, 2006, pp. 206 – 209. I found there such pearls of wisdom as (p. 37) *Even mathematics has aspects that cannot be proven, such as the fourth dimension.* In my opinion, that book is barely useful.

10. In 1983, issuing from a biased stochastic supposition, Stigler declared that another author rather than Bayes had actually written the Bayes memoir. In 1999, while reprinting his 1983 paper, in spite of his sensational finding being stillborn and forgotten, Stigler got rid of its criticisms in a tiny footnote (p. 391).

11. Stigler (1986) is loath to mention his predecessors. On pp. 89 – 90 he described the De Moivre – Simpson debate forgetting to refer to me (1973a, p. 279). And on pp. 217 – 218 he discussed the once topical but then completely forgotten conclusion concerning statistics of population without citing his only possible source of information, Chuprov's letter to Markov of 10.3.1916 (Ondar 1977/1981, No. 72, pp. 84 – 85).

Long before that Stigler (1977) dwelt on Legendre's accusation of Gauss concerning number theory without naming me (1973b, p. 124, note 83).

So why does Stigler remain so popular? Because the statistical community is crassly ignorant of the history of its own discipline; because it pays absolutely no attention to the slandering of Gauss' memory (even if realizing that fact, as the reviewer for *Hist. Math.* did, see above, – I personally informed him about it in 1991, but he had known it himself); because it possesses a narrow scientific Weltanschauung; and because the tribe of reviewers does not feel any social responsibility for their output. And of course there is a special reason: Stigler published his book (1986) when there was hardly anything pertinent except for papers in periodicals. The same happened to a lesser extent with Maistrov's book of 1974 which is still remembered!

To end my pamphlet, I quote, first, the most eminent scholar and historian of science, the late Clifford Truesdell (1984, p. 292), whom I will never forget and whose alarm bell apparently fell on deaf ears, and, second, Einstein's letter of 1933 to Gumbel, a German and later an American statistician (Einstein Archives, Hebrew Univ. of Jerusalem, 38615, in translation):

1) *No longer is learning the objective of scholarship. [...] By definition, now, there is no learning, because truth is dismissed as an old-fashioned superstition.*

2) *Integrity is just as important as scientific merits.*

Bibliography

- Descrosières A.** (1998), *The Politic of Large Numbers*. Harvard Univ. Press, Cambridge (Mass.) – London.
- Hald A.** (1998), *History of Mathematical Statistics etc.* Wiley, New York.
- Healy M. G. R.** (1995), Yates, 1902 – 1994, *Intern. Stat. Rev.*, vol. 63, pp. 271 – 288.
- Ondar Kh. O., Editor** (1977, in Russian), *Correspondence between A. A. Markov and A. A. Chuprov etc.* Springer, New York, 1981.
- Plackett R. L.** (1972), The discovery of the method of least squares. *Biometrika*, vol. 59, pp. 239 – 251. Reprinted: Kendall M. G., Plackett R. L., Editors (1977), *Studies in the History of Statistics and Probability*, vol. 2. Griffin, London, pp. 279 – 291.
- Porter T. M.** (2004), *Karl Pearson*. Princeton Univ. Press, Princeton – Oxford.
- Sheynin O. B.** (1973a), Finite random sums. *Arch. Hist. Ex. Sci.*, vol. 9, pp. 275 – 305.
- (1973b), Mathematical treatment of astronomical observations. *Ibidem*, vol. 11, pp. 97 – 126.
- (1993), On the history of the principle of least squares. *Ibidem*, vol. 46, pp. 39 – 54.
- (1999a), Discovery of the principle of least squares. *Hist. Scientiarum*, vol. 8, pp. 249 – 264.
- (1999b), Gauss and the method of least squares. *Jahrbücher f. Nationalökonomie u. Statistik*, Bd. 219, pp. 458 – 467.
- (2005), *Theory of Probability. Historical Essay*. NG Verlag, Berlin, 2009. Also www.sheynin.de
- (2006), *Theory of Probability and Statistics As Exemplified in Short Dictums*. NG Verlag, Berlin. Also www.sheynin.de
- Stigler S. M.** (1977), An attack on Gauss published by Legendre in 1820. *Hist. Math.*, vol. 4, pp. 31 – 35.
- (1986), *The History of Statistics*. Harvard Univ. Press, Cambridge, Mass.
- (1999), *Statistics on the Table*. Harvard Univ. Press, Cambridge, Mass.
- Truesdell C.** (1984), *An Idiot's Fugitive Essays on Science*. Springer, New York.