Alexandr Chuprov

Statistical Papers and Memorial Publications

Compiled andTranslated by Oscar Sheynin

Berlin 2004

Contents

0. Foreword …

Part 1. Chuprov's Contributions

1. Moral Statistics. *Энциклопедич. словарь Брокгауза и Ефрона* (Brockhaus & Efron Enc. Dict.), vol. 21, 1897, pp. 403 – 408 …

2. [On Teaching Statistics in High School], 1904. Reprinted by A.A. Rusov in *Voprosy Statistiki*, No. 9, 2003, pp. 75 – 76 together with a summary of the opinions of a number of other statisticians on the same issue (pp. 74 – 75). Material prepared by A.G. Volkov …

3. Main Issues of the Theory of Mass Phenomena. Published as a manuscript ca. 1908, incorporated in author's *Очерки по теории статистики* (Essays in the theory of Statistics). Moscow, 1909, 1910, 1959, pp. 9 – 16 of the last edition …

4. Bortkiewicz. *Нов. энциклоредич. словарь* (New Enc. Dict.), Eds Brockhaus & Efron, vol. 7, 1912, p. 647. Signed "Ch." …

5. On the Expectation of the Coefficient of Dispersion. *Izv. Imp. Akad. Nauk*, vol. 10, No. 18, 1916, pp. 1789 – 1798 …

6. On the Mean Square Error of the Coefficient of Dispersion. Manuscript (end of 1916 or early 1917). Archive, Russian Acad. Sci., Fond 173, Inventory 1, No. 50 …

7. The Theory of Stability of Statistical Series. *Nord. Stat. Tidskr.*, Bd. 5, No. 2 – 3, 1926, pp. 195 – 212. In Swedish. Reported in 1918. Translated from its Russian translation in author's collected articles *Вопросы статистики* (Issues in Statistics). Moscow, 1960, pp. 224 – 239. …

8. La décomposition du bolshevisme. N.p., n.d., no title page. Signed on last page (p. 16) A. Tchouprov, Professeur dÉconomie politique à l'Université de Moscou. Reprinted from copy kept at the Bibliothéque Nationale de France. Extremely rare …

8a. Dmitriev, A.L., Semenov, Chuprov and the Bolshevik Revolution (a fragment). Published as Foreword to their Russian translation of [8] in *Voprosy Istorii*, No. 10, 2003, pp. 3 – 18 …

8b. Sheynin, L.B. Chuprov: Socialism, Money and Leninist Russia. Manuscript …

9. On the Expectation of the Ratio of Two Mutually Dependent Random Variables. *Trudy Russk. Uchenykh za Granitsei*, vol. 1. Berlin, 1922, pp. 240 – 271 …

Part 2. Other Authors

10. Dmitriev, A.L., Sheynin, O., Chuprov: Additional Bibliography. Prepared for this collection …

11. Struve, P., Lappo-Danilevsky, A., Diakonov, M. Note on Professor Chuprov's Scientific Work (1917), this being the author's successful nomination of Chuprov to Corresponding Membership of the Russian Academy of Sciences. Reprinted in Chuprov, A.A. *Материалы конференции к 70-летию со дня кончины*(Materials of Conf. Observing 70 years after His Death). Petersburg, 1996, pp. 56 – 59 …

12. Anonymous [A Document from Chuprov's Personal Records] [1918]. Ibidem, pp. 59 – 60 …

13. Romanovsky, V.I. Review of Chuprov, Zur Theorie der Stabilität statistischer Reihen. *Skand. Aktuarietidskr.*, Bde 1 – 2, 1918 – 1919, pp. 199 – 256 and 80 – 133. *Vestnik Statistiki*, No. 1 – 3, 1923, pp. 255 – 260 …

14. Obituaries of Chuprov; the first two were published in Swedish and translated here from their private Russian translations by A.A. Muravyev

14a. Gulkevitj, K. *Nord. Stat. Tidskr.*, vol. 5, 1926, pp. 167 – 170 …

14b. Bortkiewicz, L. von. Ibidem, pp. 163 – 166…

14c. Kaminka, A. Newspaper *Rul*, 21 April 1926, pp. 1 – 2 …

15. Heymons, Helene (née Chuprov), Letters to Karl Pearson. Manuscript. (1926). Pearson Papers 719/9, University College London …

16. Eliseeva, I.I., The Chuprov Statistical School. *Voprosy Statistiki*, No. 2, 1995, pp. 40 – 43 …

17. Eliseeva, I.I., Chuprov's Heritage (a Fragment). *Материалы конференции* ... (see Item 11), pp. 15- 19)
18. Eliseeva, I.I., Dmitriev, A.L., Chuprov's Letters to Lutokhin. (Includes Lutokhin's previously unpublished recollections of Chuprov.) *Izvestia Petersb. Univ. Ekon. & Finansov*, No. 2, 1997, pp. 112 – 118
19. Eliseeva, I.I., Dmitriev, A.L., Chuprov (Fragments). *Статистики русского зарубежья* (Statisticians of the Russian Diaspora).Petersburg, 1998, pp. 77 – 89.

Foreword

Aleksandr Aleksandrovich Chuprov (1874 – 1926) was the most influential Russian statistician of his time. His contributions included in pt. 1 below are, with a single exception (a reprint of his extremely rare French political pamphlet), my translations from Russian. Leaving aside several reviews, all of Chuprov's contributions on mathematical statistics are now available in English or German. True, there remains his Russian paper of 1923 on correlation in case of two variables, but an English paper of the same title for the case of three variables appeared in 1928 (*Trans. Cambridge Phil. Soc.*, vol. 23, No. 12, pp. 337 – 382). The second part of this book consists of my translations of papers concerning Chuprov, again with a single exception (letters of Chuprov's sister to Karl Pearson written in German).

Chuprov's biography is well known; I myself (1996) extensively described his life and work and several items in pt. 2 below are also relevant, but two pertinent points ought to be mentioned here. First, concerning Chuprov's refusal to return to Russia after the Bolshevik coup d'état in 1917, it is instructive to adduce several lines from a letter of Sergei Speransky, the husband of one of Chuprov's sister, to Chuprov's student Boris Karpenko (1926; ЦГА СПб, Fond 9960, Inventory 1, Delo 255, No. 66 reverse):

> *Did the deceased* [Chuprov] *desire to return? Indeed, a return would have been inseparably linked with the need to work* (*and, most naturally, in the sphere of statistics*) *with the Soviet power. But A.A. disapproved of such joint work until the very end of his life. This is seen by his refusal to meet with Popov during the session of the International Statistical Institute in Rome* {in 1926}. *Popov*[…] *wanted to employ Chuprov at the Directorate. This fact became known in Moscow through Litoshenko.*

Lev Litoshenko (1886 – 1936) was an economist and worked then at the Soviet Central Statistical Directorate. Speransky stated that Litoshenko had served Popov, the Head of the Directorate, as an interpreter at the Session. In 1930, Litoshenko was "subjected to repressive measures" (Kornev 1993, p. 84).

The second point: there exist widely differing opinions about the significance of Chuprov's *Essays* (1909). Thus, Eliseeva [16] (such numbers are references to the respective items in this collection) sets great store by them whereas I am keeping to my previous contrary opinion (Sheynin 1996, pp. 98 – 99 and 114). There, on p. 114, I quoted Markov's pertinent and highly unfavorable comments. Chuprov's student Chetverikov (1968, p. 51) testified that (after 1910) his teacher did not agree to reprint the *Essays* and suggested that Chuprov was dissatisfied with his exposition of the theory of dispersion. I venture to add: at least after having exchanged, in 1910 -1917, many letters with Markov, Chuprov came to regard critically the entire mathematical contents of his *Essays*. On the other hand, this contribution showed Chuprov's profound knowledge of the literature on statistics, philosophy and logic. Other relevant examples are his acquaintance with the works of Süssmilch (see [1]), with Lambert's attempts to study randomness (Chuprov 1909, p. 188) and his extremely high and apparently warranted opinion of Cournot (Sheynin 2004a).

Chuprov's writings before 1916 focussed on "non-mathematical" statistics and agricultural economics but these writings are not sufficiently known even in Russia. Even his English contribution (Chuprov 1912), in spite of Keynes' high opinion about it (Sheynin 1996, p. 25), is likely forgotten.

After 1916, without abandoning these fields, he began actively working on important problems of the not yet really existing mathematical statistics. Commenting on his entire heritage, the eminent physicist Ioffe (Sheynin 1996, p. 15) stated that Chuprov was inspired by statistics as much as physics inspired Einstein.

Chuprov is especially meritorious for his (only partly successful) attempts to unite the two contemporaneous currents of statistics, the Continental direction and the Biometric school. True, I believe that, before appreciating Chuprov's role in originating mathematical statistics, the transition from the Pearsonian biometry to the Fisherian statistics ought to be thoroughly investigated. Anyway, after Chuprov's death the Royal Statistical Society passed a resolution of condolence stating that his contributions were "admired by all" and "did much to harmonize the methods of statistical research developed by continental and British workers" (Sheynin 1996, p. 126 or 2004b, p. 231). But in 2001 *Biometrika* (vol. 88) published five essays devoted to its centenary without mentioning any continental scientists. On the other hand, Markov's correspondence with Chuprov published in Russian was translated with an Introduction by Neyman (Ondar 1981) who documented Chuprov's achievement of 1923 in studying stratified sampling which he, Neyman, had overlooked in his contribution of 1952. I ought to add that the publication of this correspondence was corrupted by many mistakes corrected by me (Sheynin 1996, §8.3) and that I discovered and published there, in §8.2, 13 more letters exchanged between Markov and Chuprov.

Finally, Chuprov was the founder of the Russian statistical school among whose representatives was Oskar Anderson (1887 – 1960) who later became the most influential statistician of Bulgaria, then, from 1942 to his death, of Germany and West Germany. Chuprov's attitude towards his students and other statisticians (Romanovsky, Slutsky) was akin to that of a loving and caring father towards his offspring. He devoted much time to teaching and to corresponding with his colleagues. One of his students, referring to Chuprov, noted that his teacher had been spending a third of his working hours on correspondence (Sheynin 1996, p. 32). Eliseeva & Dmitriev [18] put on record that, shortly before Chuprov died, the students of the Polytechnical Institute(who never saw, but obviously heard enough about him) had urged him to return home.

In Soviet Russia, from ca. 1928, when the Stalinist regime began to be felt in earnest, and perhaps to the 1960s, Chuprov had been passed over in silence (Sheynin 1998, pp. 530 – 531 and 540, n24). Nowadays, however, Russian statisticians are acknowledging his merits. And Andersen (1957, p. 97) apparently believed that Lexis, Bortkiewicz and Chuprov had originated "die alte deutsche mathematisch-statistische Schule". It is opportune to add that Anderson's student, Professor Heinrich Strecker, told me several years ago that his teacher had time and time again mentioned Chuprov in his lectures and that he, Strecker, considered himself "Chuprov's grandson".

Several explanatory remarks are in order.

1. Chuprov formulated and studied problems of the most general nature and inevitably derived unwieldy formulas, see my Foreword to [6]. I ought to explain that my unsigned remarks and forewords in the main text are enclosed in curly brackets.

Furthermore, Chuprov apparently had not paid due attention to notation which I had to change, notably in [9], also see Foreword to [7]. Especially frustrating is his often use of subscripts upon subscripts (and superscripts upon superscripts). In 1923, answering Slutsky's pertinent remark, Chuprov (Sheynin 1996, p. 126) belatedly stated that notation ought to be "attentively discussed". Finally, Chuprov did not apply the notation $n!$.

2. In many cases Chuprov did not provide exact bibliographic description of his sources and I had to supplement them.

3. He did not number his formulas adequately and in some cases I changed their numeration accordingly.

4. I also somewhat changed Chuprov's dated terminology. His *squared* (or *mean*) *error* became *mean square error* and only the three first words of the expression *law of distribution of the values of* … are now retained. Then, Chuprov's *variable* means *random variable* and I note finally that he had not used the term *normal distribution*.

5. In the references appended to the contributions below I have replaced the spelling *Tschuprow* by *Chuprov*.

6. In several cases I have excluded some parts of the memorial publications mainly to avoid repetition.

Acknowledgements. I am indebted to Dr. A.L. Dmitriev (Petersburg) for reprints of his papers and copies of many archival materials unavailable outside Russia. And I am eight years late in thanking Professor Strecker without whose favorable opinion and editorship of my work on Chuprov (Sheynin 1996), and without his, strongly suspected by me, financial support of its publisher, it would not have appeared. A large portion of my translations below are available in microfiche collections in the Deutsche Hochschulschriften series (DHS) put out by Hänsel-Hohenhausen (Egelsbach), namely in DHS 2514 (1998), 2656 (1999), 2696 (2000) and 2799 (2004), but I have retained the copyright to ordinary publication.

Abbreviations in main text

JNÖS = *Jahrb. f. Nat. Ökon. u. Statistik*
L = Leningrad
M = Moscow
NST = *Nordisk Statistisk Tidskr.*
Psb = Petersburg
R = In Russian

References

Anderson, O. (1957), *Probleme der statistischen Methodenlehre in den Sozialwissenschaften.* Würzburg.
Chetverikov, N.S. (1968), Remarks on Lexis' work. In *О теории дисперсии* (On the Theory of Dispersion).Coll articles compiled & translated by Chetverikov. Moscow, pp. 39 – 54.
Chuprov, A.A. (1909), *Очерки по теории статистики* (Essays on the Theory of Statistics). Moscow. Subsequent editions, 1910, 1959. References in text to last edition.
--- (1912), The break-up of the village communities in Russia. *Econ. J.*, vol. 22, 1912, pp. 173 – 197.
Kornev, V.P. (1993), *Видные деятели отечественной статистики* (Eminent National Statisticians). Moscow.
Ondar, Kh.O., Editor (1981), *The Correspondence between A.A. Markov and A.A. Chuprov on the Theory of Probability and Mathematical Statistics* (1977, in Russian). New York.
Sheynin, O. (1996), Chuprov. *Life, Work, Correspondence* (1990, in Russian). Göttingen.
--- (1998), Statistics in the Soviet Epoch. *Jahrb. f. Nat. Ökon. u. Statistik*, Bd. 217, pp. 529 – 549.
--- (2004a), On Cournot's heritage in the theory of probability (2002, in Russian). In Deutsche Hochschulschriften 2799, pp. 209 – 221.
--- (2004b), *History of the Theory of probability to the beginning of the 20*th *Century.* Berlin

Part 1

1. Moral Statistics

Энциклопедич. словарь Брокгауза и Ефрона (Brockhaus & Efron Enc. Dict.), vol. 21, 1897, pp. 403 – 408

Foreword

{If Chuprov's student composition of 1896 [1, Chapt. 9] is disregarded, the essay below is his first scientific work. It was Bortkiewicz who agreed to compile such an article for the *Brockhaus & Efron EnziklopedicheskySlovar* ; feeling, however, that he had no time for this task, he asked the Editors of the *Slovar* to offer it toChuprov [1, p. 39]. And, just as his student composition, this writing surprises the reader: it seems to be composed by a master having at least a dozen contributions to his credit.

I omitted Chuprov's references to a number of other articles in the same source (for example, to Quetelet and to the mathematical direction in statistics). Apparently conforming to the format of the *Slovar*, Chuprov had not provided the initials of the first names of the authors included in his Bibliography (to which he did not refer in the text itself, again evidently conforming to the same format). In some instances, and especially in the case of Fischer, this restriction proved frustrating.}

* * *

Moral statistics studies mass manifestations of human activities going on beyond the sphere of economic relations. Together with demography that considers biological processes in which man is only passive; and with economic statistics whose subject is man's activity directed towards material welfare for society, moral statistics represents a science of mass phenomena in social life. Marriages and legitimate births (as far as these are not brought about by physiological causes), lechery and illegitimate births, crime and suicide, school and church find their place in the investigations carried out by moral statistics which is thus a combination of many heterogeneous elements. Historians attribute all these, as far as they belong to the past, to the sphere of cultural history. And it would perhaps be more correct, following Rümelin, to rename moral statistics calling it *cultural statistics* and thus to abandon definitively the dated idea about a close connection between moral statistics and ethics.

The term *moral statistics* is comparatively recent. The French scientist Guerry was the first to use it in 1833, but the science itself is older than its name. Its origin is usually attributed to 1741 – 1742, to the time when the first edition of Süssmilch's *Göttliche Ordnung* had appeared. Issues belonging to moral statistics were being touched even earlier. Süssmilch himself dealt above all with demography, but his right to be called the father of moral statistics cannot be questioned. The rate and the fertility of marriages, prostitution, illegitimate births, crime, suicides, – these are the phenomena of moral statistics that interest Süssmilch. He also touches some relatively subtle issues, as the connection between the rate of marriage and family status. He discovers in the studied phenomena "a stable, universal, great, perfect and splendid order". For each given place the numbers expressing them change but little from year to year, and everywhere the ratios remain constant and expediently fit for the purpose of mankind, the inhabiting of the Earth.

Süssmilch's contemporaries acknowledged the significance of his works, but later on they were forgotten. When the direction, to which he kept on, was revived in the first quarter of this century, it had not adjoined his investigations. Only after the new doctrines achieved predominance in science, their first prophet was discovered and appraised anew.

It was mostly the general reanimation of theoretical thinking, the first impetus to which was given by the brilliant school of French mathematicians, that led to the rebirth of moral statistics in the 1820s. The appearance of rich materials pertaining to some spheres of moral statistics, chiefly to criminal statistics, no less influenced the situation. France showed the way: from 1825 onwards, the *Compte général de l'administration de la justice* is being published there.

Two names mark this epoch in the history of moral statistics, – Guerry and Quetelet. According to the former's definition,

> *The subject of statistics is the spiritual life of man; it studies his abilities, customs, feelings, passions. It thus covers moral philosophy, politics, religion, legislation, history, literature and art.*

Statistics, or, as Guerry prefers to call it, *analytics*, aims at concentrating the initial data, and, by a consecutive treatment, reduces them to a small number of general abstract propositions. Guerry mistrusts speculations about the result [thus] obtained, but he does not restrict analytics by simply establishing the facts. On the contrary, he tries to discover "general and invariable laws of moral physiology", and he only desires to save science from serving politics, to protect it from the invasion of the *esprit de système* which he hates. The regularity of the phenomena in the issues of ethics, that is revealed when passing from isolated cases to the mass, surprises Guerry and he admits to being unable of finding some explanation for it. Neither does he consider it possible to base any propositions restricting free will on this regularity.

Quetelet, whose investigations are less careful and precise than Guerry's, introduced more new points of view into statistics and is undoubtedly in the forefront of the epoch's scientific movement with regard to influencing his contemporaries. He underpins his system by the principle that the phenomena concerning man obey regularities as invariable as the laws of nature. Moral statistics, that studies phenomena influenced by man's free will, proves this. Each time when we take sufficiently numerous data, a surprising constancy in the repetition of phenomena (for example, in contracting marriages) comes to light. What is the relation between this regularity of mass phenomena and the conditions for the course of separate cases? Does not this regularity presuppose the absence of free will? Not at all, is Quetelet's answer to the second question. But his response to the first one sounds by far not as decidedly; it can be reduced to the following.

A man is, first of all, an independent unity; however, as a social being, he forgoes some of his independence and enters a complicated social organism living its own life and obeying its own laws. This connection with the whole indeed introduces order in the course of social phenomena; the free will that expresses the independence of an individual is effaced as soon as we pass on to considering numerous groups. Being turned, now in one, now in another direction, it is a random cause and its influence in the final reckoning is reduced to zero. Mass phenomena are conditioned not by it, but by causes beyond the power of separate persons, and external with regard to the will of an individual. It should not be thought, however, that these factors [causes] are only physical; factors of ethical nature are just as important. The influence of customs, prejudices, decorum on social life is great, – but with regard to a man all these mighty powers are as external as climate, soil, etc. Thus, the personal freedom of an individual and the external conditions, natural and social, are the three categories of causes that Quetelet rightfully advances for explaining the regularities of moral statistics. Only he is not always able to combine them correctly, and he especially often exaggerates the role of the external causes.

Quetelet's influence on the development of statistics was exceptionally great. The direction, that most decidedly expressed its sympathy with him, was therefore called *Queteletism*. Keeping to this designation, we ought to remember, that, although it adjoins Quetelet, he himself belongs to this school not more than to the opposite one. Queteletism borrows from Quetelet the proposition that "moral phenomena obey the same invariable laws as the natural phenomena", but it interprets this principle in the sence of external causation leaving without attention the intrinsic factors; the system of Queteletism is therefore named *mechanical*.

Buckle in England; the materialistic school of the 1850s – 1860s; and, later, Ad. Wagner in Germany, – these are the most prominent representatives of Queteletism. Buckle's originally conceived and vividly and sharply written work rendered considerable services to moral statistics in spite of the amateurism of its statistical sections on which the learned German critics like to dwell: he attracted even more attention to the issues of moral statistics than the work of Quetelet himself did. Buckle issues from the idea that our actions are determined by motives that in turn are not groundless; perceiving all the conditions, external and intrinsic, we can foresee man's line of action. In other words, the actions of men "and therefore of societies" obey invariable laws.

Statistics shows that the same order, the same regular connection with various given circumstances is observed in murders, suicides and marriages as in the movement of the tide or in the alternation of the seasons. All these social phenomena represent the product of the general conditions of the life of a group: an individual only fulfils that, which is a necessary corollary of causes lying beyond him. Not the man's vices, but the state of the society in which the criminal is living engenders crimes; not the temperaments and desires of separate men determine the yearly number of contracted marriages, but the price of food and the level of wages. Buckle searches out the justification of all these propositions in moral statistics. Issuing from a correct assumption that "if the actions of individuals obey invariable laws, the actions of societies are also subject to them", he states that "if the action of societies obey invariable laws, the actions of individuals are also subject to them". In order to prove this idea, he resorts to two logical leaps: "our actions are determined when all the conditions, external and intrinsic, are given – they are determined when only the external conditions are given" and "our actions depend on external causes – they depend only on them".

The German materialists Fischer and Loewenhardt obtain the same conclusions and in the same way. Not less resolutely but with infinitely less talent they come out against the defenders of free will; they easily destroy the stronghold of the freedom of arbitrariness upheld by no-one; and then with the same easiness they make the fateful leap: By abandoning the intrinsic causes they reduce everything to the external causes alone.

Ad. Wagner, in his first purely statistical work, which he later repeatedly renounced, had kept to the Queteletism's point of view. True, he censures the onesidedness of the conclusions derived by the materialists from Quetelet's investigations, but this does not hinder him himself from representing the issue in such a light as though necessity, and purely external necessity at that, is the only legitimate conclusion from the regularities discovered by moral statistics. He is convinced that the conditions, lying beyond the acting persons, insurmountably demand the accomplishment of a definite number of certain actions irrespective of the moods and feelings of these persons. In justifying this proposition, he refers to the invariability of statistical ratios under constant conditions and to the connection between the changes of these ratios with various external circumstances.

The movement of the price of bread, for example, explains for him the change in the number of marriages, which, in separate cases, the brides, the bridegrooms and their relatives "accounted for by various purely personal reasons". We hesitate, we deliberate, but after all our actions are determined by the fact that the law should be complied with: a few cases are still lacking for the regularity to be accomplished, – so these cases indeed ought to occur. Owing that he personally *cannot clearly imagine himself the compatibility of statistical regularities and the doctrine of free will, but that* he is *unable to be reconciled with the recognition of unconditional necessity either*, Wagner throws down the guntlet to the German philosophers:

> *It is time for you to abandon the ostrich policy; you will not save your ideas by keeping silent about the conclusions of moral statistics.*

The challenge had not disappeared without trace. Even before that, the resolute blows, which moral statistics, according to the prophets of Queteletism, delivered on the doctrine of free will, began to worry the opponents of the mechanical view on the nature of moral issues and caused their uncoordinated attempts of rebuff. Now, the antagonists of Queteletism took up the work in concert, and for some time turned the course of statistical thinking to a new channel. Drobisch's profound investigations and the subtle, sometimes fanciful articles of Knapp; Rümelin's essays, elegant and peculiar with respect both to idea and form; Schmoller's discourses full of zeal but not rich in originality; Vorländer's and Siebeck's flat notes; the work of Öttingen, vast, ponderous, but, in spite of all his erudition, not free from smattering, – all this expresses the protest against the inferences of Queteletism in the name of ethics and philosophy.

At the same time Rhenisch sharply criticizes the factual basis of these conclusions. All this group of scientists decidedly rejects the notion of external determination of our actions. Intrinsic causes are put in the forefront and the role of personal morals is stressed. This direction may be called *ethical*. Its most prominent representatives are Drobisch, Rhenisch and Öttingen. Drobisch agrees that the dreams of the defenders of absolute freedom do not stand a comparison with moral statistics. But may we confuse freedom with arbitrariness? Private freedom of a man is his ability to determine his desires and, consequently, actions in accord with the requirements of the mind.

Moral statistics does not affect this freedom. On the contrary, this discipline even testifies in its favor: the possibility of separating the general mass into groups, not equally often acting in a certain way, directly indicates the part played by the personal traits of the individual. In considering the logical foundations of the theory of probability more successfully than anyone before him, Drobisch establishes a general principle: Everywhere, when constant causes act together with changing, random causes; when, in addition, only two incompatible events can happen, the numbers of the occurrences of each of them gradually tend, as they increase, to some invariable ratio. In connection with this principle, the constancy of statistical numbers indicates that, in a large social union, the desires and the grounds for action expressed in these numbers remain invariable from year to year, and that, taken as a whole, the number of people seeing no causes for suppressing their inclinations or of those whose moral stability is too weak, hardly changes. Feeling that his explanations are vague and incomplete, Drobisch admits that the study of moral statistics leads to doubts concerning the possibility of upholding any form of free will: too often does this discipline reveal the action of external causes. With more precise and comprehensive data there would have perhaps been no remainder for attributing it to free will.

Rhenisch proceeds to criticize Queteletism from an absolutely another point of view. He does not argue about the conclusions justified by the constancy of statistical ratios, he denies the very constancy, reconsiders the data that surprised Quetelet by their invariability and convinces himself in that the variations are actually not so small at all. In one case he was able to show that Quetelet had overlooked the change in the register {of crimes} and, when attempting to reveal constancy, compared heterogeneous data. Rhenisch's sharp remarks spread considerable discord in the notion of constancy, but already because of their indefiniteness they were unable to solve the issue. Which variations are large, and which are small? Where are the criteria of stability? Rhenisch gives no answer to these questions; however, if stability is estimated by eye, then, of course, one will be surprised by a constancy, whereas another one will be astonished by mutability.

Öttingen is a representative of the same wave of the reaction against Queteletism, but his views are very peculiar. He is sure from the very beginning that there cannot be any contradictions between the empirical data of moral statistics and the ethical postulates of freedom and responsibility; statistics can only formulate these more definitely. It compels us to abandon the notion about an arbitrary and indefinite freedom; such a notion however cannot hold out in any case. The essence of ethical freedom includes an element of necessity: a man is the freeer the more consistently he puts into practice his ethical ideal. Moral statistics corroborates this concept of freedom understood as an intrinsic determinism of actions; how else can we explain that, under the same external conditions, mass actions are not always the same, and the separate actions making up the mass are performed by those rather than by other individuals?

On the contrary, moral statistics refutes the ideas of the representatives of the mechanical direction as well as of the point of view of the "atomists" who admit purely personal freedom of a quite independent origin. Had the society been an aggregation of autonomous atoms only more or less often colliding one with another, the regularities revealed by {moral} statistics would have been unthinkable. We ought to recognize that a person, for all his ethical freedom, is most closely connected with the society and is a product of the social system. A peculiar feature of the Öttingen's construction is the significance that he attaches to the close connection between the individual and the society, and his system is therefore called *social-ethical*. This connection is most clearly revealed in his views on crime. According to Öttingen, a crime is not a private business of one person: the entire society is always participating in it.[1] It is also the society that is responsible for the crime although neither is the criminal free from guilt: not compulsion, not external necessity, but his own vicious will led him to the crime. After Öttingen, all the most important issues of moral statistics remain just as intricate as they were before him. It occurs that the attempt to arrange the facts in accord with a pattern constructed in advance is far from equivalent to their scientific explanation.

The other representatives of the ethical direction provide few new ideas. Each of them offers his own definition of freedom; sets off the difference between the disputes about freedom and about the determinism of action, stresses, to a greater or lesser extent, the close connection of the individual and the society; and all of them recognize, as one man, that the inferences made by moral statistics are incompatible with the idea of freedom as arbitrariness. We see originality of theoretical thought only in the works of one of the first representatives of the reaction against Queteletism, G.-C. Lewis, the author of the anonymous review of Buckle's book in the *Edinburgh Review* for 1857. Without upholding arbitrariness or groundlessness of man's actions, Lewis shows that moral statistics gives no right even for denying these concepts, and still less for recognizing external necessity.

All the conclusions reached by Buckle and Quetelet are quite compatible *with that, what any advocate of freedom had been stating at any time,*

Lewis categorically declares and justifies his proposition by a subtle analysis of the logical connection of the statistical regularities with the principles of the theory of probability. His interesting ideas remained hardly noticed and were not subsequently reproduced in the polemic against Queteletism.

Although the ethical direction played a prominent part in the history of moral statistics, it barely influenced the development of that discipline. The main merit of this direction is purely negative. It showed with such obviousness the irrelevance of the claims of moral statistics to play the decisive part in the sphere of philosophical issues that the rebirth of the same notions in the Italian anthropological [anthropometric] criminology only stirs up perplexity. Some of the representatives of the ethical direction noticeably advanced the investigation of concrete statistical problems, but they did little to solve the general theoretical issues and neither did they attempt to achieve much [in this sense].

The new school is concentrating its attention on these very issues. Unlike the mechanical or the ethical schools, this one can be called *mathematical* because it is looking for support towards mathematical disciplines, and mainly to the theory of probability. It may also be called *logical* because it most closely adjoins logic. Unlike the previous school, this one does not run after applications. It is interested in science for the sake of science, but, at the same time, it arrives at more precise solutions of those issues with which the representatives of the ethical direction fruitlessly struggled. Similar to the school that, in the 1820s, reanimated moral statistics, this one obtained its first impetus from mathematics. The theory of probability, that Laplace, Gauss and Poisson developed thus allowing it to become close to statistics, played the main, although not an exclusive part here.

The coming together was sketched out by Cournot; and Lexis, in the second half of the 1870s, accomplished it in a number of writings rich in deep original ideas. At the same time Sigwart offered the first clear essay on the relations between statistics and logic which Rümelin then supplemented. After them, already in the 1880s – 1890s, appeared Kries, who subtly and thoroughly analyzed the logical grounds of probability theory, and Edgeworth, Lehr, Westergaard and Bortkiewicz developed the methods of statistical investigation. This movement is issuing from the theory of probability; at the same time, the development of the mathematical issues of demography is going on. It leads to the ascertaining of the main idea of statistics, of the notion of group and of its changes. The work of the new school is still continuing. Much is not done yet; the accomplished parts are very disconnected, and there is no summary, no completed system, – but the animating spirit of the new ideas is already felt in each branch of statistics. It also carries off moral statistics and I therefore ought to dwell on some propositions of the new school, although the mathematical direction has no such direct ties with moral statistics as have the mechanical and the ethical directions.

The application of the theory of probability to statistics is underpinned by the notion of objective possibility introduced by Sigwart and developed by Kries. Only on the face of it this concept contradicts the law of causality. The point is that any phenomenon that a scientific investigation is dealing with is very involved; and its cause is also intricate. Once the cause A is given in the entirety of its complexity and definiteness, the event is necessary and anything else is impossible. If, however, not the entire cause is taking place but only some part of it, the action under consideration is not necessary anymore; it becomes only possible because apart from it all the other phenomena resulting from some other supplement of the given conditions that does not lead to A, are also possible. We therefore have to consider not only the relations of necessary causal connections but also those of possible ties whose various forms Kries had brilliantly and thoroughly discussed. Their peculiar feature is that action is only correlated with a part of the cause. The analytic investigation of the relations of a *possible* causal connection is a logical function of the statistical method whereas a similar study of the relations of a *necessary* connection falls to the share of induction.

Objective possibility can be greater or less; a special numerical measure called mathematical probability serves for expressing it quantitatively. Its significance is revealed by the celebrated Bernoulli theorem: in a large series of phenomena partly subject to the action of the same general causes and partly exposed to causes peculiar to each of them separately and unconnected with the general causes, the numbers of the repetitions of all the possible events are always roughly proportional to their probabilities under given general conditions.

This proposition logically represents the main form of the so-called law of large numbers which also has applications to issues of moral statistics. If the intervals of time and the boundaries of regions are not too wide, the phenomena of social life are subject to a number of invariable conditions, both natural and social. When we collect a definite mass in accord with some indication and for given conditions of time and place, we may consider all the phenomena peculiar to this mass as possible corollaries of the complex of invariable causes supplemented by the chosen indication. Under the selected conditions all these corollaries possess their own definite mean probabilities. If the general conditions are the same for two masses, the relative numbers of the repetitions of all the possible corollaries (which, in accord with the law of large numbers, ought to be roughly proportional to their mean probabilities) will be almost the same. In this pattern, statistical regularities only tell us that the general conditions remain invariable from year to year and that, together with them, the system of probabilities remains invariable. That is, the various groups (for example, age groups) are included in the constituted mass in about the same ratios, and, within these groups, the causes leading to the actions remain almost invariable.

The Bernoulli theorem, when applied to explaining statistical regularity, led to the recognition of the mean as a characteristic of the separate cases; the Poisson theorem, showing that the mass can also be heterogeneous, reveals the unlawfulness of such a transition. Rümelin is right when he exclaims:

> *If statistics, guided by its mean figures, will dare tell me that, with a probability of one over so and so much, I will, during a year, become an object of prosecution, I shall unhesitatingly answer: ne sutor ultra crepidam* {Cobbler! Stick to your last!}.

But he adds here:

> *If statistics tells me that, during this year, I ought to die with probability 1/49, I shall humbly yield to this severe truth.*

This statement betrays that not the power of logic but the depth of his revolted feeling led Rümelin to a correct conclusion. A person who exceeds the level of the crowd by his vital power as much as Rümelin does by his moral strength, would have been no less right in becoming indignant because of the attempt of statistics to condemn him to death than Rümelin did become of the threat of bringing him to court.

The law of large numbers throws light also on the relations between the regularity of mass phenomena and the course of separate cases. True, common causes determine the final result, they are not however revealed before these cases occur, but stand out against the entire series [of observations] and depend on the phenomena comprising the series. It is senseless to say that I ought to perform a certain action because the general conditions require a definite number of such actions, and somebody else if not I myself should meet the lack. Indeed, by failing to perform, I change the general conditions together with the final result [which is impossible]. That my line of actions does not noticeably influence the general course of the phenomenon is nothing but a corollary of the simple truth which is that a unity becomes an ever lesser part of a group the more unities it includes. This fact does not at all corroborate the {wrong} inference that the course of social life is allegedly determined not by the attributes of the individuals {involved} but by the properties of some mysterious nation differing from the sum of its separate representatives. External determinism does not find support for itself in moral statistics, and neither does the intrinsic determinism.

Suppose that the role of the external and intrinsic motives is reduced to excluding some decisions and that the will selects a decision from among those that remain with the same irregularity as exists in the alternation of *white* and *red* in the game of roulette. The lawfulness of the mass result will still be the same as when assuming a complete determinism of the volitional decisions. Here, the fruitlessness of the attempts to transfer a metaphysical debate on empirical grounds is revealed.

Such is the interpretation of the issues of moral statistics from the point of view of probability theory. But does this approach have any solid grounds? We are explaining statistical regularities whereas their very existence was denied. We discuss general conditions, series of independent phenomena, separate trials, – but do these notions, developed on the simple and verifiable basis of games of chance, conform to those intricate relations of social life to which we transfer them? May we attribute reality to our patterns? Statistics is obliged to Lexis for the methods of solving these questions and for the first more or less precise answers to them. He established subtle and convenient tests for judging the stability of statistical ratios. By their means he discovered that the stability was hardly ever as high as it was thought to be, but that its existence could not be denied. He proved that in the greatest majority of cases we have to do with general conditions that are not invariable but changing in accord with known laws. Lexis was able to show that the constructions of the theory of probability were very important for statistics and at the same time to outline the boundaries of their cautious scientific usage. His works, in which originality of ideas competes with clearness of exposition, opened up the newest epoch in the development of moral statistics. In our country, Russia, moral statistics fared worse than the other branches of science. Our literature is not rich even in concrete pertinent researches. True, we have the incompleted monographic investigation of Yanson, but it is the least successful work of the late professor: the ideas are hardly original, the argumentation is insufficiently rigorous, and the explication is heavy and unclear. General textbooks touching also the issues of moral statistics restrict their attention to dated theories, and even so insufficiently and not really rigorously.

Note

1. {This is what Quetelet advocated.}

References

1. (To the Foreword) Sheynin, O. *Chuprov* (1990, in Russian). Göttingen, 1966.
1. Buckle, H.T. *History of Civilization in England*, vol. 1. London, 1857.
2. Drobisch, M.W. Moralische Statistik. *Leipziger Repertorium*, 1849.
3. - - - *Die moralische Statistik und die menschliche Willensfreiheit*. Leipzig, 1867.
4. Fischer, J.C. *Über die Freiheit des menschlichen Willens*. Leipzig, 1858.
5. Guerry, A.M. *Essai sur le statistique morale de la France*. Paris, 1833.
6. --- *Statistique morale de l'Angleterre etc*. Paris, 1864.
7. Knapp, G.F. Die neueren Ansichten über Moralstatistik. JNÖS, 1871.
8. Lexis, W. Moralstatistik. In *Handwörterb. d. Staatswissenschaften*, Bd. 2. Jena, 1891.
9. --- Statistik. Ibidem.
10. --- Les causes des régularités statistiques. *Bull. Inst. Intern. Stat*., t. 9, 1896.
11. Loewenhardt, S.E. *Die Identität der Moral- und Naturgesetze*. Leipzig, 1863.
12. Öttingen, A. *Die Moralstatistik und die christliche Sittenlehre*. Erlangen, 1868 – 1869. Third edition: *Die Moralstatistik in ihrer Bedeutung für eine Socialethik*. Erlangen, 1882.
13. Quetelet, A. All his statistical writings.
14. Rehnisch, E. Zur Orientierung über die Untersuchungen und Ergebnisse der Moralstatistik. *Z. f. Phil. u. phil. Kritik*, Bd. 69, 1876.
15. Rümelin, G. Über den Begriff eines socialen Gesetzes (1867). [16, pp. 1 – 31].
16. --- *Reden und Aufsätze*. Tübingen, 1875.
17. --- Statistik. In *Handb. der polit. Ökonomie*, Bd. 3. Tübingen, 1885.
18. Schmoller, G. Über die Resultate der Bevölkerungs- und Moralstatistik. Berlin, 1871.
19. Siebeck, Das Verhältniss des Einzelwillens zur Gesammtheit im Lichte der Moralstatistik. JNÖS, 1879.
20. Süssmilch, J.P. *Die Göttliche Ordnung etc*. Berlin, 1741 – 1742.
21. Vorländer, Die Moralstatistik und die sittliche Freiheit. *Tübinger Z*., 1866.
22. Wagner, A. *Die Gesetzmässigkeit in den scheinbar willkürlichen menschlichen Handlungen vom Standpunkte der Statistik*. Hamburg, 1864.
23. Yanson, Yu.E. *Направления в научной обработке моральной статистики* (Directions in the Scient. Treatment of Moral Statistics). Psb, 1871.

2. [On Teaching Statistics in High School], 1904.

Reprinted by A.A. Rusov in *Voprosy Statistiki*, No. 9, 2003, pp. 75 – 76 together with a summary of the opinions of a number of other statisticians on the same issue (pp. 74 – 75).
Material prepared by A.G. Volkov

Foreword

{In 1901, the Statistical Commission of the Imperial Free Economic Society began collecting the opinions of statisticians, members of the professorial staff of several universities, concerning the feasibility of introducing statistics into high school. By 1904, most differing answers had come from fourteen scientists including Chuprov, and A.A. Rusov, a statistician himself, published a pertinent general résumé now reprinted together with and preceding the reprint of Chuprov's answer.

About ten years later Nekrasov initiated attempts to introduce the theory of probability into high school. His efforts proved unsuccessful, above all because he recommended a faulty program prepared by another mathematician. He had not mentioned statistics at all which evidently proves that it was not included either. Chuprov participated in the discussions of Nekrasov's proposal; he expressed doubts and suggested carrying out a trial in several schools (which, as it seems, was not implemented), see Sheynin, O. (1996), *Chuprov. Life, Work, Correspondence*. Göttingen, p. 21.}

* * *

Not having sufficient teaching experience ([and] I had no occasion to teach in the high school) I cannot bring myself to answer the raised question unconditionally. I think, however, that both statistics and the school will gain only little from the addition of statistics to the subjects taught in the high school. Insofar as statistics in the first place is concerned, the risk is too high that, at the hands of *ill-prepared* teachers not interested in statistics as a science, its teaching in the high school will only lower the interest of the school students to our science. *It is not difficult to present statistics in a boring way*; but within the boundaries of the possible it is very difficult, as I see it, to arouse the school students' love to it. And where to get scientifically prepared and sufficiently talented teachers? In Russia, statistics is only read at law faculties, but lawyers very seldom devote themselves to teaching in the high school. To hope of success, even to some extent, it is first necessary to make sure to prepare teachers. Until then, it is dangerous, if taking into account the interests of statistical science, to introduce compulsory teaching of statistics in the high school. And practical statistics will not benefit much if, instead of a *tabula rasa*, it gets a school graduate filled with the same indifferent contempt if not hate for statistics as the gymnasium student taught by a bad teacher of classical languages feels towards Latin and Greek. And it would hardly be expedient to let practical interests determine the school curriculum. Other subjects will [then] follow statistics. [Suppose that] I shall now express my wish that the high school gives its students the information necessary in case they will have to participate in carrying out a census. But then, why not prepare them for work in land or quantity surveying? Why not acquaint them with methods of fire-fighting? We may address such demands to special schools preparing village constables and volost [small rural districts] clerks, even to theological seminaries since they are not high schools but schools preparing clergymen. The representatives of those interests, that are affected by the sphere of the students' future work, are naturally entitled to express their wishes that the students of such schools be acquainted with everything necessary for fulfilling this work. But the high school that does not prepare its students for definite duties, should be left free from suchlike infringements. The only way in which practical statistics would be able to enlist the school's help is to follow the German custom mentioned in Rusov's report: to acquaint school students just before a census with its course and pertinent forms.

An attempt is made to reconcile the interests of statistics and the school by indicating the educational importance of statistics and its method. Here, above all, we should ascertain what exactly is meant here, an acquaintance with the theory of the statistical method; with the techniques of a statistical investigation; or with descriptive statistics. The possibility of acquainting a student of a high school more or less thoroughly with the theory of the statistical method seems highly questionable. Suffice it to recall what do the school graduates gain from the course in logic. Do many of them have any distinct impression at least about the methods of induction? Absolutely exceptional talent is required for teaching this chapter of statistics { ! } in school. The issue with regard to the techniques of statistical investigations is undoubtedly easier. It is possible to impart information about the methods of collecting statistical data, the principles of their treatment, about tables and graphical representation without great difficulties. However, and certainly without denying the usefulness of all this, I cannot agree that this information is of outstanding general educational importance beyond the general theory of the statistical method. Finally, as to the introduction of the principles of descriptive statistics into the school curriculum, this would unquestionably be very desirable. However, it is already being done although imperfectly. The main elements of demography and some facts about economic statistics are, or at least should be included in courses on geography. And I would indeed think that it is quite sufficient to expand somewhat this chapter of the program and to join to it such scant information concerning the statistical techniques as is appropriate in the high school. At the same time such a reform will be able to heighten the level of teaching geography: being compelled to introduce a considerable amount of statistical information, the teacher of geography will turn to studying descriptive statistics and stop demanding the learning of facts about bygone ages by heart as it is now often done with respect to the geography of Russia. And the introduction of the elements of statistics into the geographic curriculum will also compel to alter the preparation of the teachers of geography which is now going on absolutely abnormally.

3. The Main Issues of the Theory of Mass Phenomena

Published as a manuscript ca. 1908, incorporated in author's *Очерки по теории статистики* (Essays in the theory of Statistics). Moscow, 1909, 1910, 1959, pp. 9 – 16 of the last edition …

Foreword

{In 1909, Chuprov published his book [1] intended as a Master dissertation to be defended at the Law faculty of Moscow University. Actually, the Faculty conferred on him the degree of Doctor (of political economy and statistics). Below, I am translating the first pages of this contribution that constituted his introductory speech at the defence of his dissertation. They were published separately (apparently in a small number of copies) and then prefixed to the second (and then to the third) edition of his book. The speech naturally had not contained any list of references, but Chuprov provided it at the end of his book, and my own list at the end of the translation is extracted therefrom. I had not however included better known sources (Mill, Lexis).

Chuprov had apparently come to regard his writing critically enough; and in any case Markov was utterly dissatisfied with it, see Foreword. In particular, he [2, Letter 3] was surprised that Chuprov had failed to mention him or Liapunov. There, he apparently bore in mind his extension of the law of large numbers onto dependent random variables. On the other hand, Chuprov hardly applied the central limit theorem and did not therefore need to refer to Liapunov's investigations.

Chuprov's style is repulsive. His original text translated below contained many badly constructed sentences occupying seven lines, a few of them even nine or twelve.}

References

1. Chuprov, A.A. *Очерки по теории статистики* (Essays on the Theory of Statistics). M., 1909, 1910, 1959.

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

* * *

After many years of stagnation, the theoretical thought in the field of statistics is now entering a period of reanimation. The time of exclusive attention to problems of statistical technique is apparently coming to an end. The interest in theoretically comprehending the general elements of the science, in rationally justifying the methods of work applied in statistics, is flaring up anew. The nation, from where the renovation of the statistical theory is coming, is England. As it was during the first conception of the statistical method at the time of Graunt and Petty, natural sciences are providing the impulse. But there is a difference as compared with the 17th century: then, the creators of social empirical science, borrowing the general elements of the scientific Weltanschauung from natural sciences, independently built new forms of scientific thought by issuing from new observational data on social life. Now, however, natural scientists [themselves] transfer to the field of their work the methods of research that took shape in social science and actively begin to develop them systematically. In Darwin's mother country, in the circle of biologists who continue to develop the evolution theory, the conviction in the necessity of mass observations for successfully investigating the immediate issues is ripening; those peculiar methods of work that for so long were being considered as a specific feature of the sciences of society under the name of statistics, now find for themselves wide application in the sciences concerned with organic nature. However, it occurs that those methods of statistical research, that satisfy statisticians working in social science, hardly meet the requirements of natural scientists who are accustomed to greater precision and thoughtfulness in the methodical arrangement of research. In this connection the slogan "No salvation without mass observations" proclaimed more than a quarter century ago by Francis Galton, finds a complete effect only at the end of the 19th century when in Karl Pearson's person a man of science combining a considerable mathematical gift with delicate taste for tedious statistical-empirical studies goes to the stage.

Searching for the *New Organon* {Fr. Bacon's *Novum Organum* }, Pearson turns to the theory of probability[1] and this mathematical discipline, that hardly moved forward from the times of Poisson[2], offers him the necessary assistance. More subtle, more precise and rationally justified methods of research gradually replace the traditional methods gropingly discovered by statisticians working in social science. A circle of students carried away by his methodological ideas soon closed in on Pearson. At first the substance of their investigations was mainly biological but soon the new methods were transferred to the fields of other subjects, statistical in the narrow sense, and a mathematical school of statisticians-theoreticians, that now sets the fashion for the movement of statistical thought in England, is being born.

Another strong current in theoretical statistics, that is strengthening year after year, adjoins the works of Lexis. For a long time his original and deep investigations on the stability of statistical series constituted the only source of vivid theoretical thought in our science. Everyone, unsatisfied with the empiricism that had been established in statistics, turned to them. However, even in Germany the number of those, craving for a statistical freshening from among the generation that experienced the defeat of Queteletism, was not large. Only rather recently Lexian ideas became able to influence more actively the general course of the scientific thought, and a lively work in the direction outlined by him had begun. The most prominent representative here is our compatriot Borkiewicz, Professor at Berlin University.

The third current in modern science enlivening theoretical statistics far oversteps the limit of statistics and is of philosophical nature. The protest of Windelband and Rickert against the scornful attitude towards knowledge not confined within the bounds of a "natural-scientific formation of concepts" most closely touches the interests of statistics. This discipline suffers more than any other one from a unilateral attention of the logicians to the issues of cognizing the general and eternal in the phenomena, to the "nomographic" issues, as Couturat called them. The constructions of these German philosophers liberate the thought from the hypnosis of the traditional clichés and challenge the representatives of the special branches of knowledge to a participation in the revision of the general theory of science. This challenge is all the more insistent because in many respects they, the constructions, barely answer the real requirements of those "idiographic" (Windelband's term) disciplines that they attempt to set on a new philosophic basis.

There are many points of internal contact between the three mentioned currents of modern thought. Regrettably, however, there are hardly any contacts between their representatives. Philosophers are afraid of the mathematical nature of the new statistical theories whereas statisticians are only slightly inclined to convert independently their constructions into abstract forms of a general logical system. However, the work of the logicians leads to conclusions that make it possible to formulate more precisely the problems studied by statisticians. And, on the other hand, the results achieved by the statisticians infuse vivid substance into the sketchy patterns of the logicians. The peculiar forms of the idiographic science, only obscurely outlined in the imagination of the philosophers, acquire definite and clear contours in the works of statisticians, but the wealth and the diversity of these forms are still waiting for the touch of the hand of a systematizer. At the same time, the two branches of the theoretical work in statistics are hardly making use of that potential support which they might have been rendering each other. My *Очерки* (Essays) are an attempt, or, rather, the first stage of such an attempt, at joining together these three directions. Those general issues of the theory of science where the statistician turns to the logician for help, are naturally pushed into the foreground in such a synthetic construction. In this field [of immediate attention] two subjects required revision with a special insistence: The logical essence of that primordial dualism in statistics which usually manifests itself as a contraposition between statistics as a science and statistics as a method; and the interrelation between the statistical method and the methods of induction by whose formulas most theoreticians of statistics attempt to underpin statistical investigations.

According to Zhuravsky's well-known expression[3], statistics is a calculus of categories, a counting-up in each category or group. The statistician unites the studied phenomena into classes or totalities and then more or less precisely establishes the number of the elements in such totalities. What aims are here pursued? What for are the totalities counted? Statisticians have long since recognized that these aims are various, but the degree of the variety is not clearly enough expressed in usual formulations. The Windelband antithesis between the two types of scientific problems, idiographic and nomographic, allows us to comprehend the entire depth of the issue with an incomparably greater distinctness. It enables us to show that, although statistics as an idiographic science and the statistical methods of nomographic studies are equally well expressed in the forms of the calculus of categories, the attitude of the statistician to his totalities; his point if view from which he regards them; and those logical operations that he performs on them, are essentially different in the two cases.

When systematically describing that which exists and occurs around us, this calculus along with selection in accord with value as put forward in the Rickert system, is a way out of that contradiction between the restricted force of the cognizing mind and the boundlessness of the universe considered as an object of cognition, which Rickert so expressively depicted.

When analyzing the goals of the idiographic system of our knowledge we convince ourselves in that the unrealizable requirement of an exhaustive acquaintance with the exact location of all the individual objects in space and time actually goes too far. In many cases we may be satisfied by a summary statistical picture only indicating the number of the objects for comparatively wide boundaries in space and time. This is especially true if we are able to grasp the causal conditionality of the changes in the statistical pictures without following those unit processes that the statistical point of view is actually required to replace owing to the unfeasibility of such an approach. The last-mentioned circumstance advances the problem of statistical pragmatism[4] to a prominent place among the fundamental issues of the statistical theory.

As to the nomographic attempts at discovering general and perpetual laws, the formation of totalities serves here as one of the means for unravelling the interconnections in the tangle displayed before our eyes. Being extremely simple, the methods of induction usually recommended for this goal are fit for solving rarely occurring problems; their assumptions are inconsistent with the actual formulation of the problems, and they are therefore hardly ever applied. These methods assume that the casual connections that we ought to untangle are invariably indissoluble. They state that if A is not always present when B occurs, then there is no connection between A and B. This assumption is clearly inconsistent with the realities of scientific work. The children of tall parents are not always tall, but a researcher does not conclude that the stature of the offsping does not at all depend on the stature of their parents. On the contrary, he diligently counts how often do the cases of similarity and dissimilarity between children and parents with respect to stature occur and only on the basis of such a reckoning does he decide whether there exists a connection between the phenomena interesting him, and how close is it.

Researchers have to do with various forms of such "more or less close" connections in each field of scientific work. The actually applied rules for scientific conclusions about the presence or absence of a mutual dependence between phenomena adjoin these very forms. On the face of it, all such rules sharply contradict the usual notions about causal ties, which, being indissoluble, cannot be "more or less close". At the same time we cannot forgo the concept of causal tie as a quite definite, one-valued and indissoluble connection without deserting the fundamentals of our scientific Weltanschauung. The first aim of the theory of the methods intended for picking out such kind of free ties should therefore consist in clearing up this inadmissible situation. Adjoining in the main the works of Mill and his school and partly agreeing with Lossky, I attempt to solve this paradox and to reconcile the need to reckon in scientific work with various forms of mutual dependence not to be revealed by the methods of induction and requiring special methods of treating the data for its empirical study, – to reconcile this with an assumption of a total causal determinism of the course of events in the universe.

A peculiar feature of such special methods as taking shape in practical scientific work is the very unification of separate objects into groups, the "calculus of categories". My opinion being developed here about the logical aims of these methods explains their peculiarity. At the same time my view on the relations between the inductive and the statistical methods of nomographic studies allows me to outline, with a greater than usual definiteness, the part played by the mathematical theory of probability in the field of the statistician's work. In the concept of objective mathematical probability we find the clue to the understanding of the intrinsic meaning of that calculation of the frequencies of events, to which directly or obliquely (through the so-called expectations) the methods of statistical nomographic work are essentially reduced. The common basis of these methods, not always however recognized as such, is the law of large numbers in its diverse mathematical robes which connects the objective probabilities of events with their statistical frequencies.

The first three essays of my book are devoted to these common boundaries of the theory of statistics, theory of probability and logic. They aim at preparing the ground for the construction, on the basis of the theory of probability, of a connected system of statistical methodology. It is necessary however to strengthen the foundation before erecting the building. The mathematical theory of probability numbers many representatives of our science fundamentally opposed to it. The attempts at enlisting its services for justifying the theory of statistics meet with a considerable obstacle in the not yet erased from memory recollections of the downfall experienced by the theoretical system of Quetelet from whose *Lettres* many [statisticians] are to this day drawing information about the elements of this mathematical discipline as applied to statistics. Consequently, it seems to be most important to ascertain the difference between the formulation of the issue today and in the times of Quetelet; to show that not the theory of probability was guilty of his mistakes, but a hardly successful application of its principles, and that without the theory even Quetelet's opponents are unable to refute convincingly those conclusions that stir up their embittered attacks. This leads us to the issue of stability of statistical numbers, the main scene of that noisy battle between the admirers of Quetelet and his opponents that ended by a quarter-of-a century period of theoretical apathy.

Many of the numbers established in statistics for one and the same social milieu possess the amazing property of remaining without considerable change for more or less long intervals of time. From year to year the birth-rate and the rate of marriages in a given country; the sex ratio at birth and at death; the portion of people of various ages among those entering marriage, etc, *in infinitum*, manifest not very large variations. Irrespective of its interpretation, this is a fact indisputably established by direct observation, and it is of great interest, in spite of its being barely visible, as one of the foundations of our culture. It adjoins our calculations concerning the future in any sphere of social life. It upholds the modern economic system supported by a wide-spread division of labor and by work for an indefinite demand. What, however, causes such a constancy of the numbers characterizing mass phenomena of social life? Why does it become noticeable only when the field of our observation is sufficiently wide? Look at separate families, Süssmilch indicated already more than one and a half centuries ago, – you will find only boys in some of them; or, only girls; or, boys and girls in most various numerical proportions, – but we need only to take more considerable social groups and the sex ratio at birth will display a striking constancy. What exactly is the matter here? Then, how to explain that separate and as though quite free acts reveal a regularity in the mass?

Everyone concerned with statistical material formulates similar questions for himself and nowadays they may be answered validly and precisely enough. The vast literature of special investigations adjoining Lexis accumulated an abundant stock of facts and the leading ideas drawn from the theory of probability allow us to throw a bright light on this information. Owing to Lexis, the problem of stability of statistical numbers that provoked such hot debates even in the preceding generation, now represents one of the few more or less completed sections of theoretical statistics.

The fourth essay of my book is indeed aimed at acquainting the reader with the modern status of this problem in science. This essay is the first attempt at a systematic summary of the work of the Lexian school. To some degree, the theory of stability of statistical series is a crucial test of the scientific expediency of the constructions provided in the first three essays. At the same time it serves as their concluding link because only on its basis it is possible to approach the solution of many problems outlined but left without an answer in the preceding exposition. As an example, I indicate the issue of statistical pragmatism that occupies a central position in the general theory of mass phenomena as an independent object of scientific study. A s a whole, all the four essays closely linked one with another constitute a general introduction to a detailed theoretical study of the methods of scientific work expressed in the forms of the calculus of categories.

Such are the motives that guided me when I isolated for immediate development the subject from among the issues raised by the modern theoretical movement in statistics. The problem that I formulate for myself would have been however incompletely outlined without my indicating one more circumstance. A scientific work is usually addressed to a quite definite circle of comrades in speciality and bears them in mind in the first place. Work in the field chosen by me is in this respect subject to special conditions. It goes on along a boundary path between statistics, theory of probability and logic. It is compelled to address specialists of one of these three disciplines, because, given the present division of scientific labor, persons working in one of them are seldom specially trained in the other two. I consider my investigation as mainly statistical which forced me to be somewhat cautious when explicating issues crossing over to the fields of the theory of probability or logic. I was compelled to dwell on many points that present no interest to a specialist in mathematics or a logician; on the contrary, not infrequently I had to avoid subjects most curious for them but providing no direct interest for a statistician. The composition of my book considered in detail is also defined by the fact that it bears in mind a statistical readership. I attempt to make full use of the statistical literature and include sometimes rather long historical and bibliographical digressions. Neither do I avoid references to the most important polemics. On the contrary, from among the philosophical literature I draw only on those sources that left a more or less noticeable stamp on the development of statistical thought or can at present serve as an especially convenient point of departure for the reflections of a statistician. Here, I avoid polemics and do not furnish historical or bibliographical information. This is why I readily adjoin Mill and Venn and am often guided by Sigwart but do not at all mention many other authors enjoying no less authority in their science. For the same reason, having chosen the Windelband – Rickert classification of scientific issues as my point of departure, I leave aside its relation with other similar constructions, for example with the classification of science in accord with Comte, Spenser and Kareev.

I can now precisely define my goal. I attempted to fuse together in a single whole the most important results of that intensive but subdivided work that is at present going on in various spheres of theoretical statistics, and, isolating for my immediate attention the problems of the general introduction to the theory of mass phenomena, to shape their exposition in accord with the interests of those workers in statistics who would desire to follow the movement of the theoretical thought in their science but do not possess the necessary mathematical training. If I shall be able to facilitate their acquaintance with the new attempts, difficult for understanding also for most of their West-European comrades, with a rational justification of the methods of their everyday work as developed by long-standing practice, to attract to these subtle and complicate problems, that require intensive efforts of reasoning but at the same time deeply excite the mind, a somewhat greater than at present portion of that scientific force which is at disposal of the Russian society, – then I shall consider my aim fulfilled even if new workers will pass the constructions proposed by me and abandon them.

Notes

1. {This turn was hardly sufficient.}

2. {Although statisticians had not then been interested in his achievements, Chuprov should have mentioned Chebyshev, cf. Chuprov's failure to refer to Markov and Liapunov (my foreword).}

3. {D.P. Zhuravsky, 1810 – 1856.}

4. I naturally apply this term in its previous meaning, equally close to historians and statisticians. A.C. {The term pragmatism is discussed in the *New Enc. Brit*., vol. 25, 1987, p. 980, but I have not found there its "previous meaning".}

References

1. Couturat, L. Les principes des mathématiques. *Rev. Metaphys. Morale*, t. 12, 1904.
2. Lossky, N. *Основание интуитизма* (Justification of Intuitivism). Psb, 1906.
3. Quetelet, A. *Lettres … sur la théorie des probabilités*. Bruxelles, 1846.
4. Rickert, H. *Die Grenzen der naturwissenschaftlichen Begriffsbildung*. Tübingen – Leipzig, 1902.
5. Sigwart, Chr. *Logik* (1873). Tübingen, 1904.
6. Windelband, W. *Über die Gewissheit der Erkenntnis*. Leipzig, 1873.
7. --- *Geschichte und Naturwissenschaft*. Strassburg, 1894.
8. --- *Über Willenfreiheit*, 2. Aufl. Tübingen, 1905.

4. Bortkiewicz

Нов. энциклоредич. словарь (New Enc. Dict.), Eds Brockhaus & Efron, vol. 7, 1912, p. 647. Signed "Ch." …

Foreword

{This note in an encyclopaedic dictionary (*Slovar*)was signed "Ch." Because of Chuprov's stature, and bearing in mind that its author indicated hardly known facts about Bortkiewicz' lectures of 1895 – 1897, I attribute it to Chuprov, who had certainly been close to the latter. The years of publication of the separate volumes of the *Slovar* were not provided. However, the *Introduction* to vol. 1 stated that not less than six volumes will appear yearly, and it is known that vol. 1 was published in 1911, and that, in all, 29 volumes appeared in 1911 – 1916, see article Brockhaus and Efron Encyclopedic Dictionary. *Great Sov. Enc.*, vol. 4, 1974, this source being the translation of the same volume (1971) of *Bolshaia Sovetskaia Enziklopedia*, 3rd edition.

An earlier, shorter and unsigned biography of Bortkiewicz is in *Brockhaus & Efron Enziklopedichesky Slovar*, Supplement Halfvol. 1, 1905, p. 301. There also, Bortkiewicz' lectures of 1895 – 1897 on "insurance of workers and theory of statistics" are mentioned.}

Bortkevich, Vladislav Iosifovich – statistician. Born in 1868, graduated from the Law Faculty of Petersburg University. From 1895 to 1897 read lectures as Privat – Dozent on insurance of workers and on statistics at Strasbourg University. Later on, in Petersburg, a clerk of the administrative department at the pension fund of the employees of state owned railways and teacher at the Imperial Aleksandrovsky Lyceum. From 1901, Professor at Berlin University where he teaches mathematical statistics and related disciplines. Already as a student, Bortkevich submitted an investigation on mortality and longevity of the Orthodox population of the European part of Russia to the Imperial Academy of Sciences [1; 2]. Apart from tables of mortality, it contained a general essay on the principles of the theory of measuring mortality. In an extended version, Bortkevich submitted that essay to Gцttingen University as a Doctor's thesis [3].

His works on the issues of the theory of stability of statistical series are of great importance. Closely adjoining the contributions of Lexis, he essentially developed that problem in many ways. His long paper [4] and especially his law of small numbers [5] belong to the main writings in statistics. In addition to the abovementioned investigations, Bortkevich published in various periodicals and collected papers many reviews and articles on theoretical statistics, calculus of probability, demography, insurance and economic theory.

Literature (Bortkevich's contributions mentioned above)

1. Sterblichkeit und Lebensdauer der mдnnlichen griechisch-orthodoxen Bevцlkerung des Europдischen Russlands. *Zapiski Imperatorskoi Akademii Nauk,* vol. 63, Supplement 8, 1890. Separate paging. (R)

2. Same title for weibliche Bevцlkerung. Ibidem, vol. 66, Supplement 3, 1891. Separate paging. (R)

3. *Die mittlere Lebensdauer, die Methoden ihrer Bestimmung und ihr Verhдltniss zur Sterblichkeitsmessung.* Jena, 1893.

4. Kritische Betrachtungen zur theoretischen Statistik. JNЦS, Bd, 8 (63), 1894, pp. 641 – 680; Bd. 10 (65), 1895, pp. 321 – 360; Bd. 11 (66), 1896, pp. 671 – 705.

5. *Das Gesetz der kleinen Zahlen.* Leipzig, 1898.

5. On the Expectation of the Coefficient of Dispersion

Izv. Imp. Akad. Nauk, vol. 10, No. 18, 1916, pp. 1789 – 1798 …

1. Denoting the expectation of a variable by E, we have

$$\mathrm{E}\, x_k = a_k,\ \mathrm{E}\,(x - a_1)^k = \mu_k.$$

Let r series of trials with n trials in each be carried out and assume a mutual independence of the trials and an invariable law of distribution of x.[1] Denote the value of variable x in trial j from series i by $x_{j,i}$ and suppose that

$$x_{(n),\, i} = (1/n) \sum_{j=1}^{n} x_{j,i},$$

$$x_{(nr)} = (1/nr) \sum_{i=1}^{r} \sum_{j=1}^{n} x_{j,i} = (1/r) \sum_{i=1}^{r} x_{(n),i}.$$

Replacing for the sake of convenience symbol $x_{(j,i)}$ in the second formula by x_i, we can also denote

$$x_{(nr)} = (1/nr)\,(x_1 + x_2 + \ldots + x_{nr}).$$

Considering all nr trials as a single totality, we derive, as is known,

$$\mu_2 = [1/(nr - 1)]\, \mathrm{E} \sum_{i=1}^{nr} [x_i - x_{(nr)}]^2.$$

On the other hand, issuing from the values of the arithmetic mean of each of the r series, we have

$$\mu_2 = [n/(r - 1)]\, \mathrm{E} \sum_{i=1}^{r} [x_{(n),\, i} - x_{(nr)}]^2.$$

Therefore,

$$\mathrm{E}\, [1/(r - 1)] \sum_{i=1}^{r} [x_{(n),i} - x_{(nr)}]^2 = \mathrm{E}\, \{1/[n(nr - 1)]\} \sum_{i=1}^{nr} [x_i - x_{(nr)}]^2$$

or $\mathrm{E}\, z = \mathrm{E}\, y$. The ratio $z/y = Q^2$ plays a very prominent role in the contemporary theory of statistics. In the investigations carried out by the Lexian school, Q is the main criterion for ascertaining the nature of fluctuations in a series under study. If Q is sufficiently close to 1, the stability of the series is considered normal; *i.e.*, it is assumed that the conditions of mutual independence of the trials and of the invariance of the law of distribution of x are indeed fulfilled. If Q is larger (smaller) than 1, the stability of the series is lower (higher) than the norm. The magnitude Q is therefore called the coefficient of stability, or of dispersion.

The Lexian theoretical constructions are based on the assumption that the expected value of the coefficient of dispersion is equal to 1. In its initial form this supposition is wrong: $\mathrm{E}\, Q < 1$. After Bortkiewicz [1] had indicated this fact, the role originally played by Q began to pass on to Q^2. At the same time, however, it is assumed without proof, just as it was done before with respect to EQ, that $\mathrm{E}Q^2 = 1$. The expectation of the numerator of Q^2 is equal to that of its denominator, but it does not at all follow that $\mathrm{E}Q^2 = 1$. In the general case the deviation of $\mathrm{E}\,(x/y)$ from $\mathrm{E}x / \mathrm{E}y$ can be arbitrarily large in either direction, and, in particular, if $\mathrm{E}\,(x/y) = 1$, $\mathrm{E}\,(y/x)$ cannot be equal to 1, but will certainly be larger than 1 if xy remains positive for all of its possible values.[2]

I have provided the first proof of the assumption that $\mathrm{E}Q^2 = 1$ for the case of r series of n trials each. My derivation was based on the proposition that $\mathrm{E}(x/y) = 1$ if

$$\mathrm{E}\, xy^k = \mathrm{E}y^{k+1},\ k = 0,\ 1,\ 2, \ldots, \infty. \qquad (1)$$

In its initial form the derivation was rather involved.[3] Markov, to whom I had conveyed my result, found a more direct and more simple proof, extending it onto the case in which the number of trials in separate series did not remain constant [2]. However, my method of deriving formula (1) can be somewhat modified. In its new form it furnishes an extremely simple proof of the proposition for the general case of any variable possessing any law of distribution.

2. Suppose that $z/y = 1$ (§1) for such values of x for which both z and y vanish. We shall show that then, for any k, formula (1) holds. Noting that

$$\sum_{i=1}^{nr} [x_i - x_{(nr)}]^2 = \sum_{i=1}^{nr} (x_i - a_1)^2 - nr\,[x_{(nr)} - a_1]^2$$

we have

$$\mathrm{E}\,y^{k+1} = \{1/[n(nr - 1)]\}\,\mathrm{E}\,y^k\{\sum_{i=1}^{nr} [x_i - x_{(nr)}]^2\} =$$

$$\{1/[n\,(nr - 1)]\}\,\{nr\,\mathrm{E}\,y^k\,(x_i - a_1)^2 - nr\,\mathrm{E}\,y^k\,[x_{(nr)} - a_1]^2\}.$$

But

$$\mathrm{E}\,y^k\,[x_{(nr)} - a_1]^2 = (1/n^2r^2)\ \mathrm{E}\,y^k\,[\sum_{i=1}^{nr} (x_i - a_1)]^2 =$$

$$(1/nr)\,\mathrm{E}\,y^k\,(x_i - a_1)^2 + [(nr - 1)/nr]\,\mathrm{E}\,y^k\,(x_i - a_1)\,(x_j - a_1),$$

therefore

$$\mathrm{E}\,y^{k+1} = (1/n)\,\mathrm{E}\,y^k\,(x_i - a_1)^2 - \mathrm{E}\,y^k\,(x_i - a_1)\,(x_j - a_1).$$

On the other hand,

$$\sum_{i=1}^{r} [x_{(n),i} - x_{(nr)}]^2 = \sum_{i=1}^{r} [x_{(n),i} - a_1]^2 - r\,[x_{(nr)} - a_1]^2$$

and

$$\mathrm{E}\,zy^k = [1/(r - 1)]\,\mathrm{E}\,y^k\ \{\sum_{i=1}^{r} [x_{(n),i} - x_{(nr)}]^2\} =$$

$$[1/(r - 1)]\{\,r\mathrm{E}y^k\,[x_{(n),i} - a_1]^2 - r\mathrm{E}y^k\,[x_{(nr)} - a_1]^2\} =$$

$$[r/(r - 1)]\ \{\frac{1}{n}\ \mathrm{E}y^k\,[x_i - a_1]^2 + \frac{n-1}{n}\ \mathrm{E}y^k\,(x_i - a_1)\,(x_j - a_1) -$$

$$(1/nr)\,\mathrm{E}y^k\,(x_i - a_1)^2 - \frac{nr-1}{nr}\ \mathrm{E}y^k\,(x_i - a_1)\,\mathrm{E}y^k\,(x_i - a_1)\,(x_j - a_1)\} =$$

$$(1/n)\,\{\mathrm{E}y^k\,(x_i - a_1)^2 - \mathrm{E}y^k\,(x_i - a_1)\,(x_j - a_1)\}.$$

Thus, for any k,

$$\mathrm{E}\,zy^k = \mathrm{E}\,y^{k+1}.$$

In particular, it follows that, for $k = -1$,

$\mathrm{E}\,(z/y) = \mathrm{E}\,Q^2 = 1$.[4]

3. Suppose that r series of mutually independent trials are made with variable x having a constant law of distribution, the first series constituting of s_1 trials, the second one, of s_2 series, etc. If

$$s_1 + s_2 + \ldots + s_r = s$$

we have, on the one hand, as in §1,

$$\mu_2 = [1/(s-1)]\,\mathrm{E}\sum_{i=1}^{s} [x_i - x_{(s)}]^2.$$

On the other hand, if

$$(1/s_i)\sum_{i=1}^{s_i} x_{j,i} = z_i,\ (1/r)\sum_{i=1}^{r} z_i = z_{(r)},$$

where $x_{j,i}$ is the value of x in the j-th trial of series i, then

$$\mu_2 = \mathrm{E}\,[1/(r-1)]\sum_{i=1}^{r} s_i\,[z_i - x_{(s)}]^2,$$

$$\mu_2 = \mathrm{E}\,[r/(r-1)]\sum_{i=1}^{r} [z_i - z_{(r)}]^2 \Big/ \sum_{i=1}^{r} (1/s_i).$$

Denote further

$$y = [1/(s-1)]\sum_{i=1}^{s} [x_i - x_{(s)}]^2,$$

introduce such w and u that the two expressions for μ_2 become $\mu_2 = \mathrm{E}\,w$ and $\mathrm{E}\,u$ respectively. Then, supposing that

$$Q^2 = w/u,\quad Q'^2 = u/y,$$

we shall show that $\mathrm{E}Q^2 = \mathrm{E}Q'^2 = 1$.

Like above, we ascertain that

$$\mathrm{E}\,y^{k+1} = \mathrm{E}\,y^k\,[(x_i - a_1)^2 - (x_i - a_1)(x_j - a_1)].$$

On the other hand, noting that

$$\sum_{i=1}^{r} s_i\,[z_i - x_{(s)}]^2 = \sum_{i=1}^{r} s_i\,(z_i - a_i)^2 - s\,[x_{(s)} - a_1]^2,$$

we have

$$\mathrm{E}\,wy^k = [1/(r-1)]\,\mathrm{E}\,y^k\,\{\sum_{i=1}^{r} s_i\,[z_i - a_1]^2 - s\,[x_{(s)} - a_1]^2\} =$$

$$[1/(r-1)]\,\{\sum_{i=1}^{r} s_i\,\mathrm{E}\,y^k\,[z_i - a_1]^2 - s\,\mathrm{E}\,y^k\,[x_{(s)} - a_1]^2\}.$$

However,

$$s\,\mathrm{E}\,y^k\,[x_{(s)} - a_1]^2 =$$

$$\mathrm{E}\,y^k\,[x_i - a_1]^2 + (s-1)\,\mathrm{E}\,y^k\,(x_i - a_1)(x_j - a_1)$$

and

$$(z_i - a_1)^2 = (1/s_i^2)\,[\sum_{i=1}^{s_i} (x_{j,i} - a_1)]^2 =$$

$$(1/s_i^2)\,[\sum_{j=1}^{i} (x_{j,i} - a_1)^2 + \sum_{i=1}^{s_i}\sum_{h\neq j} (x_{j,i} - a_1)(x_{h,,i} - a_1)]$$

so that

$$s_i\,\mathrm{E}\,y^k\,[z_\mathrm{i} - a_1]^2 = \mathrm{E}\,y^k\,[x_i - a_1]^2 + (s_i - 1)\,\mathrm{E}\,y^k\,(x_i - a_1)(x_j - a_1)$$

and the sum of such terms from $i = 1$ to $i = r$ is

$$r\,\text{E}\,y^k\,[x_i - a_1]^2 + (s - r)\,\text{E}\,y^k\,(x_i - a_1)\,(x_j - a_1).$$

Therefore

$$\text{E}\,wy^k = [1/(r - 1)]\{r\,\text{E}\,y^k\,[x_i - a_1]^2 +$$

$$(s - r)\,\text{E}\,y^k\,(x_i - a_1)\,(x_j - a_1) -$$

$$\text{E}\,y^k\,[x_i - a_1]^2 - (s - 1)\,\text{E}\,y^k\,(x_i - a_1)\,(x_j - a_1)\} =$$

$$\text{E}\,y^k\,[x_i - a_1]^2 - \text{E}\,y^k\,(x_i - a_1)\,(x_j - a_1)$$

or

$$\text{E}\,wy^k = \text{E}\,y^{k+1},\ \text{E}\,(w/y) = \text{E}\,Q^2 = 1.$$

In a similar way

$$\text{E}\,y^k \sum_{i=1}^{r}\ [z_i - z_{(r)}]^2 =$$

$$[(r - 1)/r] \sum_{i=1}^{r} \frac{1}{s_i}\,\text{E}\,y^k\,\{\,[x_i - a_1]^2 - (x_i - a_1)\,(x_j - a_1)\}$$

and

$$\text{E}\,uy^{\,k} = \text{E}\,y^{\,k+1},\ \text{E}\,(u/y) = \text{E}\,Q'^{\,2} = 1.$$

Noting that

$$\mu_2 = \text{E}\sum_{i=1}^{r}\ (z_i - x_{(s)}]^2\ /\ \sum_{i=1}^{r}\ [(1/s_i) - (r/s)]$$

we may construct a third modification of the coefficient of dispersion for series having unequal numbers of observations. Denote the right side of the last expression without the symbol E by v and $v/y = Q''^{\,2}$. Then, it is not difficult to ascertain that the expectation of $Q''^{\,2}$ is again unity:

$\text{E}\,Q''^{\,2} = 1.$[5]

4. Suppose that the variable x is connected with some event in such a way that $x = 0$ if the event happens and $x = 0$ otherwise. Then $(x_1 + x_2 + \ldots + x_s) = z_{(s)}$ is equal to the number of occurrences of the event in s trials, and $(1/s)\,(x_1 + x_2 + \ldots + x_s) = t_{(s)} = x_{(s)}$ is the frequency of the event.

Denote the probability of the event by p. If it is invariable and the trials are mutually independent, then

$$a_1 = \text{E}x = p,\ a_k = \text{E}x^k = p,\ \mu_2 = p - p^2 = p\,q,$$

$$\mu_3 = p\,q\,(q - p),\ \mu_4 = p\,q\,(q^3 + p^3),\ \mu_k = p\,q\,[q^{k-1} + (-1)^k\,p^{\,k-1}.$$

Suppose that

$$\mu_{k,(s)} = \text{E}\,[x_{(s)} - a_1]^k = \text{E}\,[t_{(s)} - p]^k,$$

then

$$\mu_{2,(s)} = p\,q/s,\ \mu_{3,(s)} = p\,q\,(q - p)/s^2,$$

$$\mu_{4,(s)} = 3p^2\,q^2/s^2 + p\,q\,(1 - 6p\,q)\,/s^3.$$

Noting that $x_i^2 = x_i$, we have

$$\sum_{i=1}^{s} [x_i - x_{(s)}]^2 = \sum_{i=1}^{s} x_i^2 - s\,x_{(s)}^2 = \sum_{i=1}^{s} x_i - s\,x_{(s)}^2 =$$
$$s\,[x_{(s)} - x_{(s)}^2] = s\,t_{(s)}[1 - t_{(s)}],$$
$$Q^2 = [1/(r - 1)] \sum_{i=1}^{r} s_i\,[t_i - t_{(s)}]^2 / [s/(s - 1)]\,t_{(s)}\,[1 - t_{(s)}]^2,$$
$$Q'^2 = [r/(r - 1)] \sum_{i=1}^{r} [t_i - t_{(r)}]^2 \div [s/(s - 1)]\,t_{(s)}\,[1 - t_{(s)}] \sum_{i=1}^{r} (1/s_i),$$
$$t\,(r) = \sum_{i=1}^{r} (1/r)\,t_i,$$
$$Q''^2 = \sum_{i=1}^{r} [t_i - t_{(s)}]^2 \div [s/(s - 1)]\,t_{(s)}\,[1 - t_{(s)}]\,[\sum_{i=1}^{r} (1/s_i) - r/s].$$

For the case in which the number of observations does not change from series to another one and remains equal to s, we have

$$Q^2 = [1/(r - 1)] \sum_{i=1}^{r} [t_i - t_{(sr)}]^2 \div \{1/[s - (1/r)]\}\,t_{(sr)}\,[1 - t_{(sr)}].$$

In accordance with the above,

$$\mathrm{E}Q^2 = \mathrm{E}Q'^2 = \mathrm{E}Q''^2 = 1.$$

Notes

1. The law of distribution of a variable is, as I call it, the system of all of its possible values and their corresponding probabilities.

2. This follows from the fact that

$$x/y + y/x = (x^2 + y^2) / xy = 2 + (x - y)^2 / xy$$

and that, therefore, $\mathrm{E}(x/y) + \mathrm{E}(y/x) > 2$ because $\mathrm{E}\,[(x - y)^2 / xy] > 0$ if only xy does not take negative values.

3. {Chuprov refers to his contribution, then in print. It had not appeared and its manuscript was lost.}

4. The derivation is based on the relations

$$\mathrm{E}y^k\,(x_i - a_1)^2 = \mathrm{E}y^k\,(x_h - a_1)^2,$$

$$\mathrm{E}y^k\,(x_i - a_1)\,(x_j - a_1) = \mathrm{E}y^k\,(x_h - a_1)\,(x_g - a_1)$$

which, under the stipulated conditions, do not demand any special proof, but $\mathrm{E}\,(y/z)$ cannot be calculated in the same way since $\mathrm{E}z^k\,(x_i - a_1)\,(x_j - a_1)$ depends on whether x_i and x_j belong to the same or to different series.

5. It seems really essential to note the difference between the construction of Q' and Q'' consisting in that, while deriving the former, the deviations of the magnitudes z_i are taken from their simple arithmetic mean

$$z_{(r)} = (1/r)\,(z_1 + z_2 + \ldots + z_r)$$

whereas in the latter case the deviations are calculated from their weighted mean

$$x_{(s)} = (1/s)\,(s_1 z_1 + s_2 z_2 + \ldots + s_r z_r).$$

When dividing by

$$(1/s_1) + /1/s_2) + \ldots + (1/s_r),$$

which is not infrequently recommended in the case of unequal numbers of observations in the series, it is therefore necessary to adhere to the following rule: For the numerator, calculate the deviations from the usual mean of z_i rather than from the general mean of all trials, whereas the deviations from the latter are entered in the denominator of Q'.

References

1. Bortkiewicz, L. Der wahrscheinlichkeitstheoretische Standpunkt im Lebensversicherungswesen. *Österr. Rev.*, Jg. 31, 1906, NNo. 24 – 28, pp. 149 – 150, 155 – 156, 161 – 165, 167 – 168, 173 – 174.

2. Markov, A.A. On the coefficient of dispersion (1916). In *Избранные труды* (Sel. Works). N.p., 1951, pp. 523 – 535. Transl.: DHS 2514, 1998, pp. 195 – 204.

6. On the Mean Square Error of the Coefficient of Dispersion

Manuscript (end of 1916 or early 1917). Archive, Russian Acad. Sci., Fond 173, Inventory 1, No. 50 …

Foreword

{I am publishing a manuscript written by Chuprov in 1916 or in the very beginning of 1917. Indeed, he referred there to his related paper of 1916 [3], and, on 28 Jan. 1917, he [16, p. 70, Letter 88c] answered Markov's question certainly occasioned by his new work. And, in general, the manuscript was more or less discussed in a number of letters from the Markov – Chuprov correspondence [11; 16]. The manuscript is kept at the Archive of the Russian Academy of Sciences (Fond 173, Inventory 1, No. 50).

In 1916 or somewhat earlier, as a result of his scientific intercourse with Markov, mathematical statistics became for Chuprov no less important than economics or statistics in general. In particular, during 1918 – 1919 he [15] continued his investigations in mathematical statistics and partly made use (in Essay 2, or, more precisely, in its Chapt. 2, §§2 – 5; Chapt. 3, §1; and in its Conclusions) of the results obtained in the manuscript.

Statistical series became an object of research in the last quarter of the 19th century, mainly due to Lexis [10; 9; 14; 17, Chapt. 3; 16, §11]. He proposed to estimate the stability of such series, i.e., of the underlying probability of success of the (Bernoulli) trials, by a certain *coefficient of dispersion* Q. The desire to specify Lexis' ideas compelled statisticians and mathematicians including Markov [1] and Chuprov [2; 3] to determine the expectation and the variance of of Q (or Q^2) and Chuprov's manuscript was devoted to the second problem; here, in the same book, the reader will find his investigation of the first problem. Chuprov also qualitatively studied the distribution of Q^2 and, at the last moment [16, p. 72, Letter 91a], attempted to prove that it might be described by a Pearson curve of type III (thus going beyond the boundaries as designated by the title of his manuscript). Markov [16, p. 70, Letter 88a], however, had refuted his proof and Chuprov never returned to this subject.

Chuprov's formulas are very involved and not without reason did Markov [11, Letter 105] state: "Your work frightens me with its abundance of complicated calculations…" It is well worth to recall Romanovsky's remark [12, p. 416] that Chuprov's formulas concerning correlation theory "being of considerable theoretical interest" were "almost useless" due to complicated calculations involved. In our case complication was occasioned by the subject of investigation: Chuprov studied series whose terms were mean values of differing numbers of observations and could have made a few mistakes either in calculations or when writing down his results. For my part, I attempted to render his formulas without error..

I introduced some notation, notably D and F (between formulas (5) and (6), $f_1(n; r)$ and $f_2(n; r)$ after formula (9), G (formula (10)) and $s_{[r]}$ (between formulas (16) and (17)).}

* * *

1. Denoting the expectation of a variable by E, we shall set

$$\mathrm{E}\,x^k = a_k,\ \mathrm{E}\,(x - a_1)^k = \mu_k.$$

Suppose that r series of trials, n trials in each, are made under the conditions of mutual independence of the trials and an invariable law of distribution of x. Denote the value of the variable x at the j-th trial of the i-th series by x_{ij} and set

$$x_{(n),i} = (1/n)\sum_{j=1}^{n} x_{ji},\quad x_{(nr)} = (1/nr)\sum_{i=1}^{r}\sum_{j=1}^{n} x_{ji} = (1/r)\sum_{i=1}^{r} x_{ni}.$$

Replacing for the sake of convenience symbol x_{ji} by x_i in the second expression we may also have

$$x_{(nr)} = (1/nr)\sum_{i=1}^{nr} x_i.$$

Introduce now

$$[1/(r-1)]\sum_{i=1}^{r} [x_{(n),i} - x_{(nr)}]^2 = z,\quad [1/n(nr-1)]\sum_{i=1}^{nr} [x_i - x_{(nr)}]^2 = y,$$

$$z/y = Q^2.$$

I [3] have shown that for any k

$$\mathrm{E}\, zy^k = \mathrm{E}\, y^{k+1}$$

and, if z/y is assumed to take the value 1 at such values of x for which both z and y vanish, $\mathrm{E}Q^2 = \mathrm{E}(z/y) = 1$. It is not difficult to show in a similar way that, for any k,

$$\mathrm{E}y^{k+2} = (1/rn^3)\,\mathrm{E}y^k\,[(x_1-a_1)^4 - 4(x_1-a_1)^3(x_2-a_1)$$

$$-3(x_1-a_1)^2(x_2-a_1)^2 + 12(x_1-a_1)^2(x_2-a_1)(x_3-a_1)$$

$$-6(x_1-a_1)(x_2-a_1)(x_3-a_1)(x_4-a_1)] +$$

$$[(nr+1)/(nr-1)n^2]\,\mathrm{E}\,y^k$$

$$\cdot\,[(x_1-a_1)^2(x_2-a_1)^2 - 2(x_1-a_1)^2(x_2-a_1)(x_3-a_1) +$$

$$(x_1-a_1)(x_2-a_1)(x_3-a_1)(x_4-a_1)]; \qquad (1)$$

$\mathrm{E}\,z^2y^k$ = {same expression as (1), only the multiplier $(nr+1)/[(nr-1)n^2]$ of the second term becomes $(r+1)/[(r-1)n^2]$}.

Denote $(1/n^2)\,\mathrm{E}\,(1/y^2)\cdot${the last square bracket in (1)} by A, $(1/n^2)\,\mathrm{E}\,(1/y^2)\cdot${the first square bracket in (1)} by B and assume that $z/y = 1$ at such values of x for which both z and y vanish, then, at $k = -2$,

$$1 = (1/rn)\,B + [(nr+1)/(nr-1)]\,A,$$

$$\mathrm{E}\,(z^2/y^2) = \mathrm{E}Q^4 = (1/rn)\,B + [r+1)/(r-1)]\,A$$

or

$$\mathrm{E}\,Q^4 = 1 + \frac{2r(n-1)}{(r-1)(nr-1)}\,A, \qquad (2)$$

$$\mathrm{E}\,Q^4 = 1 + \frac{2r(n-1)}{(r-1)(nr+1)}\,[1 - (1/nr)]\,B.$$

Assume also that

$$(1/n^2)\,\mathrm{E}\,\{(1/y^2)\,[(x_1-a_1)^4 - 4(x_1-a_1)^3(x_2-a_1) + 3(x_1-a_1)^2(x_2-a_1)^2]\} = C,$$

then $B = C - 6A$ so that

$$A = \frac{nr(nr-1)}{(nr-2)(nr-3)}\,[1 - (1/nr)\,C], \qquad (3)$$

$$1 - (1/nr)\,B = \frac{(nr+1)nr}{(nr-2)(nr-3)}[1 - (1/nr)\,C],$$

$$\mathrm{E}\,Q^4 = 1 + \frac{2r(n-1)nr}{(nr-2)(nr-3)}\,[1 - (1/nr)\,C]. \qquad (4)$$

The problem of determining

$$\mathrm{E}\,(Q^2 - 1)^2 = \mathrm{E}Q^4 - 1 \qquad (5)$$

is thus reduced to finding out one of the magnitudes A, B or C which in turn are easily written down as

$$A = (nr - 1)^2\,\mathrm{E}\{[x_1^2 x_2^2 - 2x_1^2 x_2 x_3 + x_1 x_2 x_3 x_4]\,/\,D\} =$$

$$4n^2r^2(nr - 1)^2\,\mathrm{E}\{[\text{ same numerator as just above}]\,/\,F\},$$

$$\mathrm{B} = (nr - 1)^2\,\cdot$$

$$\cdot\mathrm{E}\,\{[x_1^4 - 4x_1^3 x_2 - 3x_1^2 x_2^2 + 12x_1^2 x_2 x_3 - 6x_1 x_2 x_3 x_4]\,/\,D\} =$$

$$4n^2r^2(nr - 1)^2\,\mathrm{E}\{[\text{ same numerator as just above}]\,/\,F\},$$

$$C = (nr - 1)^2\,\mathrm{E}\{[\,x_1^4 - 4x_1^3 x_2 + 3x_1^2 x_2^2]\,/D\} =$$

$$4n^2r^2(nr - 1)^2\,\mathrm{E}\,\{[\text{ same numerator as just above}]\,/\,F\},$$

where {in my notation}

$$D = \{\sum_{i=1}^{nr}\,[x_i - x_{(nr)}]^2\}^2,\; F = \{\sum_{i=1}^{nr}\,\sum_{j\neq i}\,(x_i - x_j)^2\}^2.$$

Noting that the first term, x_1^4, of the sum in the numerator of C can be replaced by x_2^4, we determine

$$C = 2n^2r^2(nr - 1)^2\,\mathrm{E}\,[(x_1 - x_2)^4\,/\,F].$$

In a similar way

$$C = 4n^2r^2(nr - 1)^2\,\mathrm{E}\,\{[(x_1 - x_2)^2(x_1 - x_3)^2\,/\,F\}, \qquad (6)$$
$$A = n^2r^2(nr - 1)^2\,\mathrm{E}\,\{[(x_1 - x_2)^2(x_3 - x_4)^2\,/\,F\},$$

$$B = 2n^2r^2(nr - 1)^2\,\mathrm{E}\,\{[(x_1 - x_2)^4 - 3\,(x_1 - x_2)^2\,(x_3 - x_4)^2\,/\,F\} =$$

$$n^2r^2(nr - 1)^2\,\mathrm{E}(\{[(x_1 - x_2)^2 - (x_3 - x_4)^2]^2 - 4\,(x_1 - x_2)^2\,(x_3 - x_4)^2\,\}/F).$$

2. The exact values of A, B and C and, therefore, of $\mathrm{E}Q^4$, depend on the law of distribution of x. However, boundaries not depending on this law can be indicated for these magnitudes.

1) Noting that $C > 0$ we determine

$$A < \frac{nr(nr-1)}{(nr-2)(nr-3)},$$

$$1 - (1/nr)\,B < \frac{(nr+1)nr}{(nr-2)(nr-3)},$$

$$\mathrm{E}\,Q^4 < 1 + \frac{2r(n-1)nr}{(nr-2)(nr-3)},$$

or, for $r \geq 5$ (cf. Markov [1, 1951, p. 532]), $EQ^4 < 1 + 2/(r - 1)$.

2) Noting that

$$E[(x_1 - x_2)^4 / F] > [E(x_1 - x_2)^2 / \sqrt{F}]^2,$$

$$E[(x_1 - x_2)^2 / \sqrt{F}] = [1/nr(nr - 1)]\, E(\sqrt{F}/\sqrt{F}) = [1/nr(nr - 1)],$$

we find that $C > 2$. Hence

$$A < \frac{nr-1}{nr-3} < 1 + 2/(nr - 3),\ 1 - (1/nr)B < \frac{nr+1}{nr-3},$$

$$EQ^4 < 1 + \frac{2r(n-1)}{(r-1)(nr-3)},$$

or, for $r \geq 3$, $EQ^4 < 1 + 2/(r - 1)$.

3) Noting that (cf. (6))

$$n^2r^2(nr - 1)^2 E\{[(x_1 - x_2)^2 + (x_1 - x_3)^2 + \ldots + (x_1 - x_{nr})^2]^2 / F\} =$$

$$(nr - 1)C/2 + (nr - 1)(nr - 2)C/4 = nr(nr - 1)C/4$$

we find that

$$nr(nr - 1)C/4 > n^2r^2(nr - 1)^2 \cdot$$

$$\{E[(x_1 - x_2)^2 + (x_1 - x_3)^2 + \ldots + (x_1 - x_{nr})^2]/\sqrt{F}\}^2$$

and $C > 4(nr - 1)/nr$ or $C > 4 - 4/nr$. Therefore

$$A < \frac{(nr-1)(nr-2)}{nr(nr-3)} < 1 + \frac{2}{nr(nr-3)}, \qquad (7)$$

$$1 - (1/nr)B < \frac{(nr+1)(nr-2)}{nr(nr-3)},$$

$$EQ^4 < 1 + \frac{2(n-1)(nr-2)}{(r-1)n(nr-3)}$$

or, if $r \geq 1 + 2/n$, $EQ^4 < 1 + 2(r - 1)$.

According to the conditions of the problem, r and n are integers and none of them can be less than 2. Therefore, we convince ourselves in that, for all possible values of n and r,

$$EQ^4 < 1 + 2/(r-1). \qquad (8)$$

3. Noting that $1/a = 1/b - (a - b)/ab$, we find for any variables α and β [2, Essay 2]

$$E(\alpha/\beta) = E\alpha/E\beta - (1/E\beta)\,E\{[\alpha(\beta - E\beta]/\beta\} =$$

$$E\alpha/E\beta - [1/(E\beta)^2]\,E[\alpha(\beta - E\beta] + [1/(E\beta)^2]\,E\{[\alpha(\beta - E\beta)^2]/\beta\}.$$

It follows that, if α/β is always > 0 for any of its possible values,

$E(\alpha/\beta) > E\alpha/E\beta - [1/(E\beta)^2][E\alpha\beta - E\alpha E\beta]$.

This inequality provides a computable lower boundary for $E(\alpha/\beta)$. If there exists no connection between α and β, or if the connection is inverse, then, for whichever law of distribution of α and β, $E(\alpha/\beta) > E\alpha/E\beta$.

1) Set $\alpha = 2n^2r^2(nr - 1)^2(x_1 - x_2)^4$, β = { Chuprov writes out the expression that I denoted by F}, so that $E(\alpha/\beta) = C$. We have

$$E\alpha = 4n^2r^2(nr - 1)^2(\mu_4 + 3\mu_2^2),$$

$$E\beta = 4nr(nr - 1)[(nr - 1)\mu_4 + (n^2r^2 - 2nr + 3)\mu_2^2], \qquad (9)$$

$$E\alpha/E\beta = \frac{nr(nr-1)[\mu_4 + 3\mu_2^2]}{(nr-1)\mu_4 + f_1(n;r)\mu_2^2} =$$

$$\frac{nr}{1+[(nr-2)(nr-3)/(nr-1)]}[\mu_2^2/(\mu_4 + 3\mu_2^2)].$$

{Here, in my notation, $f_1(n; r)$ is the trinomial in formula (9).}

For $\mu_4 = \mu_2^2$ we have

$$E\alpha/E\beta = 4 - 8/(n^2r^2 - nr + 2).$$

{Denote this trinomial $f_2(n; r)$.} If $\mu_4 > \mu_2^2$, then $E\alpha/E\beta > 4 - 8/f_2(n; r)$. For the Gauss distribution, assuming {therefore} that $\mu_4 = 3\mu_2^2$,

$$E\alpha/E\beta = 6(nr - 1)/(nr + 1) = 6 - 12/(nr + 1).$$

The fraction $E\alpha/E\beta$ increases with μ_4/μ_2^2 if $nr > 3$. If $\mu_4/(nr\mu_2^2)$ tends to 0 as n increases, then, under the same condition, $E\alpha/E\beta$ tends to $\mu_4/\mu_2^2 + 3$. And if $\mu_4/(nr\mu_2^2)$, as n increases, tends to limit k differing from 0, then, under the same condition, $E\alpha/(nrE\beta)$ tends to $k/(k + 1)$.

Separating β into two parts, β', independent from x_1 and x_2, and β'', dependent on these two magnitudes, and noting that

$$E\alpha\beta - E\alpha E\beta = E\alpha\beta'' - E\alpha E\beta'',$$

we easily see that the expansion of the left side in powers of nr does not include terms higher than n^7r^7 whereas $(E\beta)^2$ includes terms with n^8r^8. Therefore, as n increases, the fraction

$$(E\alpha\beta - E\alpha E\beta)/(E\beta)^2 = G \qquad (10)$$

tends to 0 except for the case considered below and the lower boundary of C tends at the same time to the value taken by $E\alpha/E\beta$ at $n = \infty$, i.e. to $\mu_4/\mu_2^2 + 3$.

Neither does the calculation of the exact value of G (10) present any difficulties. After transformations demanding nothing except attention and patience, we find

$$G = 1/[(nr - 1)\mu_4/\mu_2^2 + f_1(n; r)]^2 \cdot$$
$$\{2n^3r^3[\mu_6/\mu_2^3 + 5\mu_4/\mu_2^2 - 4\mu_3^2/\mu_2^3 - 6] + n^2r^2[\mu_8/\mu_2^4 - 16\mu_5\mu_3/\mu_2^4 +$$

$$5\mu_4^2/\mu_2^4 - 32\mu_4/\mu_2 + 12\mu_3^2/\mu_2^3 + 42] - 2nr[\mu_8/\mu_2^4 - 7\mu_6/\mu_2^3 -$$
$$4\mu_5\mu_3/\mu_2^4 - 3\mu_4^2/\mu_2^4 - 3\mu_4/\mu_2^2 - 14\mu_3^2/\mu_2^3 + 36] + [\mu_8/\mu_2^4 -$$

$$16\mu_6/\mu_2^3 - 8\mu_5\mu_3/\mu_2^4 + \mu_4^2/\mu_2^4 + 24\mu_4/\mu_2 - 40\mu_3^2/\mu_2^3 + 54]\}.$$

If $\mu_{2i} = \mu_2{}^i$ and $\mu_{2i-1} = 0$ for $i = 1, 2, 3$ and 4, we have

$$C > 4 - \frac{8}{f_2(n;r)} - \frac{16(n^2r^2 - 3nr + 4)}{[f_2(n;r)]^2}$$

and (cf. (3)) $A < 1$ at $nr > 10$. For the Gauss distribution

$$C > 6 - \frac{12}{nr+1} - \frac{48(nr+2)}{(nr+1)^2},\ A < 1 \text{ for } nr \geq 30.$$

If, as n increases,

$$\mu_8/(n^2r^2\mu_2{}^4),\quad \mu_6/(nr\mu_2{}^3),\quad \mu_5\mu_3/(n^2r^2\mu_2{}^4),$$
$$\mu_4/(nr\mu_2{}^2),\ \mu_3{}^2/(nr\mu_2{}^3) \to 0 \qquad (11)$$

then the lower boundary for C tends to $\mu_4/\mu_2{}^2 + 3$.

Suppose that the variable x can only take values 1 and 0 with probabilities p and $q = 1 - p$ (cf. my article [3, §4]). Then

$$\mu_2 = pq,\ \mu_k = pq[q^{k-1} + (-1)^k p^{k-1}].$$

For $p = q = 1/2$, $\mu_{2i} = \mu_2{}^i$ and $\mu_{2i-1} = 0$ for any positive integer i. Therefore, also in this case, for $nr > 10$

$$\mathrm{E}\,Q^4 < 1 + 2/(r-1).$$

For p and q not being infinitesimals of the order of $1/n$, the determined lower boundary of C tends to $\mu_4/\mu_2{}^2 + 3 = 1/pq$ as n increases. If p or q is an infinitesimal of the order of $1/n$, – if, for example, $p = m/n$, – then, as n increases,

$\mu_4/(nr\mu_2{}^2)$ and $\mu_3{}^2/(nr\mu_2{}^3)$ tend to $1/(rm)$; $\mu_5\mu_3/(n^2r^2\mu_2{}^4)$ and $\mu^6/(n^2r^2\mu_2{}^3)$ tend to $1/(m^2r^2)$ and $\mu_8/(n^3r^3\mu_2{}^4)$ tends to $(1/m^3r^3)$.

At the same time

$G/(nr)$ tends to $(1 + 2rm/[rm(rm+1)^2]$ and $\mathrm{E}\alpha/(nr\mathrm{E}\beta)$, to $1/(rm+1)$.

In this case $C/(nr)$, even at $n = \infty$, remains larger than

$$(r^2m^2 - rm - 1)/[rm(rm+1)^2],$$

$$1 - C/nr < \frac{rm}{rm+1} + \frac{2rm+1}{rm(rm+1)^2},$$

$$\mathrm{E}\,Q^4 < 1 + [2/(r-1)]\left[\frac{rm}{rm+1} + \frac{2rm+1}{rm(rm+1)^2}\right],$$

$$\mathrm{E}\,Q^4 < 1 + [2/(r-1)]\left[1 - \frac{r^2m^2 - rm - 1}{rm(rm+1)^2}\right].$$

If $r^2m^2 - rm - 1 \geq 0$ or $rm \geq (1 + \sqrt{5})/2$, $\mathrm{E}\,Q^4$ remains less than $1 + 2/(r-1)$ also at $n = \infty$. Thus, for $p = m/n$ the limit of $\mathrm{E}(Q^2 - 1)^2$ as n increases is not equal to $2/(r-1)$, but, being dependent on m, remains certainly less than $2/(r-1)$ for all values of m if only $rm \geq 1.62$.

2) Set $\alpha = n^2r^2(nr-1)^2(x_1 - x_2)^2(x_3 - x_4)^2$ and $\mathrm{E}(\alpha/\beta) = A$. We have

$$\mathrm{E}\,\alpha = 4n^2r^2(nr-1)^2\mu_2{}^2,$$

$$\mathrm{E}\,\beta = 4nr(nr-1)[(nr-1)\mu_4 + f_1(n;r)\,\mu_2^2],$$

$$\mathrm{E}\,\alpha / \mathrm{E}\,\beta = \frac{nr(nr-1)}{f_1(n,r)+(nr-1)\mu_4/\mu_2^2},$$

$$G = \frac{2}{\{[(nr-1)\mu_4/\mu_2^2]+f_1(n;r)\}^2}\{2n^3r^3[(\mu_4/\mu_2^2)-1] +$$

$$n^2r^2[\mu_6/\mu_2^3 - 9\mu_4/\mu_2^2 + \mu_4^2/\mu_2^4 - 8\mu_3^2/\mu_2^3 + 11] -$$

$$2nr[\mu_6/\mu_2^3 + 2\mu_5\mu_3/\mu_2^4 - 14\mu_4/\mu_2^2 + \mu_4^2/\mu_2^4 - 17\mu_3^2/\mu_2^3 + 18] +$$

$$[\mu_6/\mu_2^3 + 4\mu_5\mu_3/\mu_2^4 - 33\mu_4/\mu_2^2 + 3\mu_4^2/\mu_2^4 - 50\mu_3^2/\mu_2^3 + 57]\}.$$

If $\mu_4 = \mu_2^2$, $\mu_6 = \mu_2^3$, $\mu_3 = \mu_5 = 0$, then

$$\mathrm{E}\,\alpha / \mathrm{E}\,\beta = 1 - 2/f_2(n;r), \quad G = \frac{8}{f_2(n;r)} - \frac{8(2nr-5)}{[f_2(n;r)]^2},$$

$$A > 1 - 10/f_2(n;r), \qquad (12)$$

$$\frac{r(n-1)}{nr-1}A > 1 - \frac{r-1}{nr-1} - \frac{10}{f_2(n;r)},$$

$$\mathrm{E}\,Q^4 > 1 + [2/(r-1)][1 - \frac{r-1}{nr-1} - \frac{10}{f_2(n;r)}].$$

For the Gauss distribution

$$\mathrm{E}\,\alpha / \mathrm{E}\,\beta = 1 - 2/(nr+1), \; G = 4/(nr+1) + 12/(nr+1)^2.$$

$$A > 1 - \frac{6}{nr+1} - \frac{12}{(nr+1)^2},$$

$$\mathrm{E}\,Q^4 > 1 + [2/(r-1)][1 - \frac{r-1}{nr-1} - \frac{6}{nr+1} - \frac{12}{(nr+1)^2}].$$

Under the conditions stated above $\mathrm{E}\,\alpha / \mathrm{E}\,\beta$ tends to 1 as n increases and G tends to 0. At the same time the lower boundary for A tends to 1, and, on the strength of (7), its higher boundary, and, consequently, A itself, also tend to 1. Thus, as $n = \infty$,

$\mathrm{E}\,Q^4 = 1 + [2/(r-1)]$ if $\mu_4/(nr\,\mu_2^2)$, etc (cf. (11)) tend to 0 as n increases.[1]

If the variable x takes value 1 with a very low probability $p = m/n$ and value 0 with probability $q = 1 - p$, then $\mathrm{E}\alpha/\mathrm{E}\beta$ tends to $rm/(rm+1)$ as n increases, and G tends to $4/(rm+1)$. In this case, also at $n = \infty$,

$$A > rm/(rm+1) - 4/(rm+1). \qquad (13)$$

As to $\mathrm{E}\,Q^4$, it, again at $n = \infty$, obeys the inequalities

$$1 + [2/(r-1)][\frac{rm}{rm+1} - \frac{4}{rm+1}] < \mathrm{E}\,Q^4 <$$

$$1 + [2/(r-1)][\frac{rm}{rm+1} + \frac{2rm+1}{rm(rm+1)^2}]$$

so that

[2/(r − 1)] [same expression as in similar bracket above] <

$E(Q^2 - 1)^2 < [2/(r - 1)]$ [same expression as in similar bracket above].

For a small m and not a very large r the boundaries for $E(Q^2 - 1)^2$ are here not sufficiently close one to another. A somewhat more advantageous lower boundary can be found by another method (§5).

4. When the law of distribution of x is known, the symbol E in the expressions for A, B, C and EQ^4 can easily, although in general without a tangible benefit, be replaced by summation. Suppose that the variable x can take values t_1 and t_2 with probabilities p and $q = 1 - p$. Assuming that

$(x_1 - x_2)^4 / F = 1 / [n^2r^2(nr - 1)^2]$,
when both the numerator and the denominator vanish if $x_1 = x_2 = \ldots = x_{nr}$ and noting that if l of the values of x are t_1 and $(nr - l)$ are t_2,

$$\sqrt{F} = 2l(nr - l)(t_1 - t_2)^2,$$

we have

$$C = 2n^2r^2(nr - 1)^2 E[(x_1 - x_2)^4 / F] = 2n^2r^2(nr - 1)^2 \cdot$$

$$\{\frac{p^{nr}+q^{nr}}{n^2r^2(nr-1)^2} + 2pq(t_1 - t_2)^4 \sum_{l=0}^{nr-2} \frac{C^l_{nr-2}p^lq^{nr-l-2}}{[2(l+1)(nr-l-1)(t_1-t_2)^2]^2}\} =$$

$$2(p^{nr} + q^{nr}) + n^2r^2(nr - 1)^2 pq \sum_{l=0}^{nr-2} \frac{C^l_{nr-2}p^lq^{nr-l-2}}{[(l+1)(nr-l-1)]^2} =$$

$$2(p^{nr} + q^{nr}) + nr(nr - 1) \sum_{l=0}^{nr-2} \frac{C_{nr}^{l+1}p^{l+1}q^{nr-l-1}}{(l+1)(nr-l-1)}.$$

In a similar way

$$A = n^2r^2(nr - 1)^2 E[(x_1 - x_2)^2(x_3 - x_4)^2 / F] = \quad (14)$$

$$n^2r^2(nr - 1)^2 \cdot \{(p^{nr} + q^{nr}) / [n^2r^2(nr - 1)^2] + 4p^2q^2 \cdot$$

$$\sum_{l=0}^{nr-4} \{C^l_{nr-4} p^l q^{nr-l-4} / [2(l + 2)(nr - l - 2)]^2\} =$$

$$(p^{nr} + q^{nr}) + n^2r^2(nr - 1)^2 \cdot \sum_{l=0}^{nr-4} \frac{C^l_{nr-4}p^{l+2}q^{nr-2-l}}{[(l+2)(nr-l-2)]]^2} \quad (15)$$

$$(p^{nr} + q^{nr}) + nr(nr - 1) / [(nr - 2)(nr - 3)] \cdot$$

$$\sum_{l=0}^{nr-4} \{C_{nr}^{l+2} p^{l+2} q^{nr-l-2} / [(l + 1)(nr - l - 3) / (l + 2)(nr - l - 2)]\}.$$

Hence, cf. (4) and see Markov [1, pp. 198 – 199 of translation]),

$$EQ^4 - 1 = \frac{2(n-1)nr}{(r-1)(nr-2)(nr-3)} \cdot$$

$$\{(nr - 2)(nr - 3) / [nr(nr - 1)] (p^{nr} + q^{nr}) +$$

$$\sum_{l=0}^{nr-4} \{\frac{(l+1)(nr-l-3)}{(l+2)(nr-l-2)} C_{nr}^{\,l+2}\, p^{\,l+2} q^{\,nr-2-l}\}.$$

Neither is it difficult to obtain similar expressions for A, C and EQ^4 for the general case in which x takes k different values $t_1, t_2, \ldots, t_k$ with probabilities $p_1, p_2, \ldots, p_k$. Thus, we have

$$A = \sum_{i=1}^{k} p_i^{nr} + n^2r^2(nr-1)^2 \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} p_i^2 p_j^2 (t_i - t_j)^4 \cdot$$
$$\sum \{(nr-4)!\, p^\alpha / \alpha!\, \{nr\,[2t_i^2 + 2t_j^2 + \alpha_f t_f^2] - [[2t_i^2 + 2t_j^2 + \alpha_f t_f]^2\}^2\}.$$

{Here, in my notation, p^α actually means the product of p_i to the power of α_i, $i = 1, 2, \ldots, k$; $\alpha!$ stands for the product of $\alpha_i!$; and $\alpha_f t_f^2$ and $\alpha_f t_f$ are the sums of such products for $f = 1, 2, \ldots, k$.} The last sum is extended onto all integer non-negative values of α_i satisfying the condition $\alpha_1 + \alpha_2 + \ldots + \alpha_k = nr - 4$ and {again in my notation}

$$\sum (nr-4)!\, p^\alpha / \alpha! = (p_1 + p_2 + \ldots + p_k)^{nr-4} = 1.$$

5. We have determined above that if the variable x takes values t_1 and t_2 with probabilities p and $q = 1 - p$, then {Chuprov repeats formula (15)}. Noting that

$$(l+2)(nr-l-2) \le n^2r^2/4,$$

we have

$$A > [(nr-1)^2 / n^2r^2] \cdot 16p^2q^2, \quad EQ^4 > 1 + \frac{2(n-1)(nr-1)}{(r-1)n\, nr} 16p^2q^2$$

If the variable x takes more than two different values, and denoting its least and greatest values by t_1 and t_2 respectively, we find that

$$\sqrt{F} \le (n^2r^2/2)(t_k - t_1)^2.$$

But {Chuprov repeats formula (14)} so that

$$A > \frac{(nr-1)^2}{n^2r^2} \cdot \frac{16\mu_2^{\,2}}{(t_k - t_1)^4},$$

$$EQ^4 > 1 + \frac{2(n-1)(nr-1)}{n(r-1)nr} \cdot \frac{16\mu_2^{\,2}}{(t_k - t_1)^4}.$$

If x only takes values 1 and 0 with probabilities p and $q = 1 - p$, then $x_i^2 = x_i$ and

$$\sqrt{D} = (1/nr)\,[\sum_{i=1}^{nr} x_i] \sum_{i=1}^{nr} (1 - x_i),$$

$$A = [n^2r^2(nr-1)^2/4]\, E\,[(x_1 - x_2)^2 (x_3 - x_4)^2 / F] =$$
$$(1/4)(p^{nr} + q^{nr}) + n^2r^2(nr-1)^2 p^2q^2 \cdot$$
$$E\,\{1 / \{(2 + \sum_{i=5}^{nr} x_i)\,[2 + \sum_{i=5}^{nr} (1 - x_i)]\}^2\}.$$

If variable z takes only positive values, then

$E(1/z^2) > [E(1/z)]^2 > 1/(Ez)^2$,

so that

$$A > \frac{n^2r^2(nr-1)^2p^2q^2}{[2(nr-2)+(nr-4)(nr-5)pq]^2}. \qquad (16)$$

For $p = q = 1/2$ this becomes

$$A > [\frac{nr(nr-1)}{n^2r^2-nr+4}]^2 > [1 - \frac{4}{n^2r^2-nr+4}]^2.$$

The boundary for A is here somewhat more advantageous than the one (12) determined above.

If pq does not tend to 0 as n increases, then the lower boundary of A tends to 1. And if $p = m/n$, this boundary, under the same condition, tends to $r^2m^2/(rm + 2)^2$ which is more advantageous than before, see (13), but nevertheless for a small m and not a very large r the boundaries for $E(Q^2 - 1)^2$ are still rather widely spaced.**2** If $p = m/n$ and $n = \infty$ the methods applied above are unable to establish the precise value of $E(Q^2 - 1)^2$. Assuming that also under these conditions $E(\alpha/\beta)$ tends to $E\alpha/E\beta$ as n increases, the limit of EQ^4 and $E(Q^2 - 1)^2$ will be respectively

$$1 + \frac{2rm}{(r-1)(rm+1)} \quad \text{and} \quad \frac{2rm}{(r-1)(rm+1)}.$$

6. Our methods of investigation can also be applied in the case of differing numbers of trials in the series of observations. Without dwelling on all the structures of the coefficient of dispersion possible here [3, §3], I only consider the main cases. Suppose that the first series of observations consists of s_1 trials, the second series, of s_2 trials, …, and denote by s the total number of trials. Let, as before, x_{ji} be the value taken by the variable at the j-th trial of the i-th series and denote also

$$(1/s_i)\sum_{i=1}^{s_i} x_{ji} = z_i,\ (1/s)\sum_{i=1}^{s} x_{ji} = x_{(s)},\ [1/(s - 1)]\sum_{i=1}^{s}[x_i - x_{(s)}]^2 = y,$$

$$[1/(r - 1)]\sum_{i=1}^{r} s_i[z_i - x_{(s)}]^2 = w,\ Q^2 = w/y.$$

As I proved earlier [3], $EQ^2 = 1$. Keeping to the above meaning of the symbols A, B and C, we shall obtain by the same way as in §1 {I denote $(1/s_1) + (1/s_2) + \ldots + (1/s_r)$ by $s_{[r]}$}

$$EQ^4 = [1/(r - 1)^2][s_{[r]} - r/s - (r - 1)/s]B + [(r + 1)/(r - 1)]A =$$

$$[1/(r - 1)^2][s_{[r]} - r^2/s]B + (1/s)B - [(r + 1)/(r - 1)]A.$$

Noting that

$$(1/s)B + [(s + 1)/(s - 1)]A = 1$$

we have

$$EQ^4 = 1 + \frac{2(s-r)}{(r-1)(s-1)}A - \frac{s}{(r-1)^2}[s_{[r]} - r^2/s][\frac{s+1}{s-1}A - 1].$$

If all the s_i are equal one to another we return to formula (2), otherwise

$$s_{[r]} - r^2/s > 0.$$

The sign of the difference between $\mathrm{E}\,Q^4$ for series consisting of differing numbers of trials and the same magnitude for the same total number of trials s equally distributed among the r series will therefore be determined by the sign of

$$[(s+1)/(s-1)]\,A - 1. \qquad (17)$$

If $A < 1 - 2/(s+1)$ then $\mathrm{E}\,Q^4$ as well as $\mathrm{E}\,(Q^2 - 1)^2$ will be larger in the first case; and they will be smaller otherwise, when $A > 1 - 2/(s+1)$. Finally, if $A = 1 - 2/(s+1)$, then neither $\mathrm{E}\,Q^4$ nor $\mathrm{E}\,(Q^2 - 1)^2$ will depend on the distribution of the trials among the series.

If $\mu_4 = \mu_2^2$, $\mu_6 = \mu_2^3$ and $\mu_3 = \mu_5 = 0$, then $A > 1 - 10/(s^2 - s + 2)$, cf. (12), and, for $s \geq 7$, expression (17) will be positive. If variable x takes the values 1 and 0 with probabilities p and $q = 1 - p$, then, cf. (16),

$$A > \frac{s^2(s-1)^2p^2q^2}{[2(s-2)+(s-4)(s-5)pq]^2}$$

and, for a sufficiently large s and $pq \geq 2/9$, expression (17) will again be positive. For $p = q = 1/2$ the same holds if $s \geq 5$.

In all these cases $\mathrm{E}\,Q^4$ and $\mathrm{E}\,(Q^2 - 1)^2$ are *the smaller* the more irregular are the trials distributed among the series. For $p = q = 1/2$

$$\mathrm{E}\,(Q^2 - 1)^2 < 1 + \frac{2(s-r)}{(r-1)(s-1)}A - .$$

$$\frac{2}{(r-1)^2}\,[s_{[r]} - r^2/s]\,\frac{s(s^3-5s^2+4s-4)}{(s^2-s+4)^2}.$$

Suppose that all the $(1/s_i)$ tend to 0 as s increases; for example, let $s_i = \gamma_i s$ where γ_i are positive and not infinitesimal numbers {unfortunate expression} less than 1. Then, also when $s = \infty$,

$$s\,[s_{[r]} - r^2/s] = \gamma_{[r]} - r^2$$

remains positive, is not infinite, and differs from zero if γ_i are not equal one to another. If, at the same time, A tends to 1 as s increases, then

$$[(s+1)/(s-1)]\,A - 1 = A - 1 + [2/(s-1)]\,A$$

tends to 0 and $\mathrm{E}Q^4$ as in the case of series of equal numbers of trials tends to $1 + 2/(r-1)$. On the contrary, if A tends to a limit less than 1, for example to $rm/(rm+1)$, then $\mathrm{E}Q^4$ tends to

$$1 + \frac{2rm}{(r-1)(rm+1)} + \frac{1}{(r-1)^2(rm+1)}\,[\gamma_{[r]} - r^2] =$$

$$1 + \frac{2rm}{(r-1)(rm+1)}\{1 + \frac{1}{2rm(r-1)}\,[\gamma_{[r]} - r^2]\}.$$

In this case $\mathrm{E}\,(Q^2 - 1)^2$ occurs to be, as $s = \infty$, *the larger* the more irregular is the distribution of the trials among the series. If at the same time

$$[1/(r-1)]\,[\gamma_{[r]} - r^2] < 2, \qquad (18)$$

then

$$\mathrm{E}\,(Q^2 - 1)^2 < 2/(r-1);$$

if, however, the trials are distributed so irregular that the left side of (18) is larger than 2, then the sign of that inequality is reversed.

If not all the $(1/s_i)$ tend to 0 as s increases, then $[s_{[r]} - r^2/s]$ remains finite even as $s = \infty$. If at the same time $s(A - 1)$ tends to 0, as it does for example if $p = q = 1/2$, then

$$s\{[(s + 1)/(s - 1)]A - 1\}$$

tends to 2, and E Q^4, to

$$1 + \frac{2}{r-1} - \frac{2}{(r-1)^2} s_{[r]}.$$

Under these conditions E Q^4 and E $(Q^2 - 1)^2$ even as $s = \infty$ remain for $p = q = 1/2$ less than when the trials are uniformly distributed among the series.

7. Bortkiewicz [3] was the first to formulate the question about the squared error of Q^2. Issuing from considerations not having the nature of a convincing proof, he [3] conjectured that E $(Q^2 - 1)^2 = 2/r$. Markov [1] established that for series consisting of the same numbers of trials, for a finite n and $r \geq 5$, E $(Q^2 - 1)^2 < 2/(r - 1)$, and that $2/(r - 1)$ was the limit to which E $(Q^2 - 1)^2$ tended as n increased. I have shown above that Markov's first statement persisted for any r and for any variable having whichever law of distribution. As to the limit to which E $(Q^2 - 1)^2$ tends as n increases, it equals $2/(r - 1)$ only if the law of distribution obeys certain conditions, see the lines preceding Note 1.

Bortkiewicz' formula is not infrequently applied in practice when studying the stability of statistical series. So, to what extent does its theoretical groundlessness and inaccuracy reflect on the results of its use? In the case of series having the same number of trials, if E $(Q^2 - 1)^2$ tends to $2/(r - 1)$ as n increases, the difference between these magnitudes is, for a finite n, of the order of $1/(nr)$. For the values of n and r usually encountered in statistical practice, E $(Q^2 - 1)^2$ is situated between $2/(r - 1)$ and $2/r$ nearer to the former. We may disregard the error of assuming that it is equal to either of them without essentially corrupting the result.

The matter is somewhat different when E $(Q^2 - 1)^2$ tends to a limit less than $2/(r - 1)$ as n increases. The error connected with the value of n can again be neglected, but the error connected with the limit can be considerable. If, for example, the limit is $[2/(r - 1)]\,[rm/(rm + 1)]$, then, when assuming, for $m < 1$ and $r < 10$, that E $(Q^2 - 1)^2 = 2/(r - 1)$, we are increasing this magnitude not less than by 10%; for $m = 0.2$ and $r = 10$ the increase will equal 1/3, and, for $m = 0.1$ and $r = 10$, the relative increase will be twofold. If now we assume that E $(Q^2 - 1)^2 = 2/r$, then, for $m < (r-1)/r$, we increase it, and we decrease it otherwise. Therefore, for small values of m, $2/r$ is nearer to the true value than $2/(r - 1)$ but the increase can still be rather large.

For a not too small m $[m > (r - 1)/r]$, E $(Q^2 - 1)^2$ is situated between $2/(r - 1)$ and $2/r$ and in practice it may be assumed equal to either of these magnitudes. An irregular distribution of the trials among the series can reflect on E $(Q^2 - 1)^2$ in different ways. It can *decrease* the squared error of Q^2, – this is the case in which the variable x can take values 1 and 0 with probabilities close to 1/2; the distribution of the newly born among the sexes can serve as an example. On the contrary, under other conditions irregularity can *increase* E $(Q^2 - 1)^2$. In particular, if this magnitude tends to $2/(r - 1)]\,[rm/(rm + 1)]$ as n increases, irregularity can not only increase E $(Q^2 - 1)^2$ up to $2/(r - 1)$, but heighten it still more. Special care is therefore needed when calculating the squared error of Q^2 in all those cases when we find ourselves in the domain of the so-called *law of small numbers*: supposing that E $(Q^2 - 1)^2 = 2/(r - 1)$ or $2/r$ we can considerably deviate from reality in either direction. However, keeping to the conditions that ensure the equality E $(Q^2 - 1)^2 = 2/(r - 1)$ as $n = \infty$, it is possible, in spite of the irregularity, to assume in addition, without undue risk, the same equality for such finite values of n that are usual for statisticians.

8. The theoretical value of E $(Q^2 - 1)^2$ is made use of in three directions. The magnitude of the squared error of Q^2 is taken into account when estimating the observed deviation of Q^2 from 1. Supposing that the law of distribution of Q^2 is close to the Gaussian and issuing from

E $Q^6 = 1 + 45/[4(r - 1)^2]$.

5. Pearson [8] provided a convenient table for calculating $\Gamma[(r - 1)/2]$.

References

1. Markov, A.A. On the coefficient of dispersion (1916). *Избранные труды* (Sel. Works). N.p., 1951, pp. 525 – 535. Transl.: DHS 2514, 1998, pp. 195 – 204.

2. Chuprov, A.A. On the mathematical foundations of the theory of stability of statistical series. (R) Unpublished and lost.

3. --- On the expectation of the coefficient of dispersion (1916). Translated in this book.

4. Bortkiewicz, L. Über den Präcisionsgrad der Divergenzcoefficienten. *Mitt. Verb. Österr. u. Ungar. Versicherungstechniker*, No. 5, 1901, pp. 1 – 3.

5. --- *Die radioaktive Strahlung als Gegenstand wahrscheinlichkeitstheoretischer Untersuchungen*. Berlin, 1913.

6. --- Realismus und Formalismus in der mathematischen Statistik. *Allg. stat. Archiv*, Bd. 9, 1915, pp. 225 – 256.

7. Elderton, W. Palin, *Frequency Curves and Correlation* (1906, date of Preface). Fourth ed., London, 1953.

8. Pearson, K., Editor, *Tables for Statisticians and Biometricians* (1914). London, 1924.

9. Bauer, R.K. Die Lexische Dispersionstheorie u.s.w. *Mitt.- Blatt math. Stat. Anwendungsgeb.*, Bd. 7, 1955, pp. 25 – 45.

10. Lexis, W. Über die Theorie der Stabilität statistischer Reihen (1879). In author's *Abh. zur Theorie der Bevölkerungs- und Moralstatistik*. Jena, 1903, pp. 170 – 212.

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

12. Romanovsky, V.I. *Математическая статистика* (Math. Statistics). M. – L., 1938.

13. Chetverikov, N.S., Compiler & Translator, *О теории дисперсии* (On the Theory of Dispersion). M., 1968. Incorporates translations of [9; 10 and 15].

14. --- Comment on [9]. [13, pp. 39 – 54].

15. Chuprov, A.A. Zur Theorie der Stabilität statistischer Reihen. *Skand. Aktuarietidskr.*, Bd. 1, 1918, pp. 199 – 256; Bd. 2, 1919, pp. 80 – 133.

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

17. Heyde, C.C., Seneta, E. *I.J. Bienaymé.* New York, 1977.

18. Kruskal, W. Formulas, numbers, words: statistics in prose (1978). In *New Directions for Methodology of Social and Behavioral Science*. Editor, D. Fiske. San Francisko, 1981, pp. 93 – 102.

19. Pearson, K. On the dissection of assymetrical frequency curves. *Phil. Trans. Roy. Soc.*, vol. A185, pt. 1, 1894, pp. 71 – 110.

7. The Theory of Stability of Statistical Series

Nord. Stat. Tidskr., Bd. 5, No. 2 – 3, 1926, pp. 195 – 212. In Swedish.

Reported in 1918. Translated from its Russian translation in author's collected articles *Вопросы статистики* (Issues in Statistics). Moscow, 1960, pp. 224 – 239.

Foreword

{This text only appeared in print, in Swedish, about eight years after Chuprov had read his report, and about the same eight years after he had published his main contribution (1918 – 1919) on the same subject (mentioned in Note 7). Nevertheless, I consider Chuprov's report important; for one thing, it enables the reader to grasp the author's main ideas without having to delve into his numerous and involved formulas.

Chetverikov (again see Note 7) corrected the original text corrupted by misprints by checking it against Chuprov (1918 – 1919). In addition, I remark that the values of random variables were denoted in the text by x_i, then by x_i', then again by x_i; I have used x_i throughout.}

* * *

Ladies and gentlemen: Allow me, first of all, to express my deep gratitude for your flattering invitation to read a lecture at the Society of Swedish Actuaries who are greatly meritorious for developing the theory of statistics in Sweden [1], and to beg you to be lenient towards my insufficient knowledge of the Swedish language in which I wish to explicate my thoughts. During my stay in your country [2], I aimed at acquainting myself, as thoroughly as it was possible for a foreigner, with the culture and the economic life of Sweden. I do not doubt that the knowledge of the people's language is the most important condition for understanding the spirit of a nation. However, having lived here, in Stockholm, for only a few months, I was unable, even with the best will in the world, to master the so developed, and in many aspects the so peculiar language as the Swedish is. I hope that you will answer my attempts by a kind *Ut desint vires* [3].

For the reason that I want to indicate in my short introduction, I especially highly appreciate the possibility offered me to describe some of my probability-theoretic views to a section of advanced foreign colleagues. At present, after a prolonged stagnation, the theory of statistics is entering a period of lively progress, which is, however, being delayed because of the splitting up of the scientific force along the lines of both speciality and nation.

Statistical methods have gradually penetrated many fields of scientific investigations; they began to be applied ever oftener not only in sociology, but also in natural sciences. At present, there exist statistical physics and statistical mechanics just as anthropological {anthropometric}, meteorological, biological and even astronomical {stellar} statistics. The subject-matter of problems being solved by statistics in different scientific fields certainly differ much one from another, but the methods are often very similar, and the conclusions are the same or almost so. Scientists, sometimes not even knowing it, follow the same paths, sometimes the same roundabout ways, or even the same mistaken paths. They believe that they are setting foot on virginal soil, but, instead, they keep going after one or several forerunners, who, with great difficulties, paved the way to the goal, common to all of them.

Let us for example consider the totality of formulas which Professor Bortkiewicz called "law of small numbers". Poisson, who derived the main pertinent formulas, developed that problem from a purely mathematical point of view without thinking about their various applications. Only after five decades, Abbe (1878) encountered the same statistical - mathematical problem when the Zeiss shop charged him with developing a haemocytometer. His formulas were also applied to the study of plankton, but no-one from the outside world paid attention to his works.

About 20 years hence Bortkiewicz (1898)published his well known investigation on the "law of small numbers" (1898), systematically expounded all the relevant issues from the viewpoint of the Lexian theory of dispersion, and adduced examples with appropriate calculations belonging to social-economic applications of his constructions. After him, physicists began working on the same problem without any knowledge of the findings of Abbe, or Bortkiewicz, or of others. The same happened with Pearson's gifted pupil Student (a pen name), who published a short but substantial article (1907) on the same subject, - on the counting of blood corpuscles, - without knowing anything about his predecessors. What a waste of efforts! How much most precious energy could have been spared by a better management of the common work in various scientific fields! The benefits of an agreed cooperation would have apparently be great, but the accompanying difficulties would not be less. It seems that, in accord with this point of view, the struggle against national separation is nevertheless more promising.

It is during the contemporary stage in the development of statistical science that insufficient contacts between statisticians of different countries interested in the theory of their discipline greatly influences its progress. The most important work in our field is now going on in England. Edgeworth, Pearson with his numerous students, and others are working during the latest decade with an extreme intensity that I highly appreciate. However, those efforts are presented in such a mathematical form, that seems uninviting to Continental investigators used to more rigorous demonstrations. Allow me to provide an example. Markov, an eminent Russian scientist working in the field of calculus of probability, confessed that he had been utterly unable to overcome his aversion to Pearson's mathematical deliberations. And I know many colleagues, who, like Markov, shelve the English investigations without reading them. I ought to own, that I always force myself somewhat when I have to study the insufficiently thought out English formulas. I feel as though reading verses in an alien tongue: for understanding the real content of their deliberations, I ought to transpose them first into the language of [sound] formulas. Not rarely does this English manner harm English scientists themselves; mistakes only caused by insufficiently rigorous methods are revealed. Recent English investigations, developing more complicated issues, are especially often corrupted. I have discovered many absolutely unreliable formulas not only in the papers of minor workers, but in Pearson's own writings as well, and I ought to attack volume 11 of *Biometrika* with special force.

The rapprochement of the English and Continental methods would have greatly benefited our science. {The heritage of} Lexis, on the one hand, and, on the other hand, Chebyshev's students from among Russian mathematicians, should influence, much more than they do now, the methods of developing those issues, whose formulation we owe to the shrewdness of the English. The method of expectations, developed by the Chebyshev school, and applied with remarkable skill by Bortkiewicz, the guest of last year of the Swedish Society of Actuaries, would have reflected on mathematical derivations especially fruitfully.

For a long time now, the tempting problem of preparing a more solid unification of the various stochastic directions followed by theoreticians of statistics, especially interested me. An unexpected spell of spare time allowed me, working for a number of years, to carry this work through to a certain completion. [4] Those parts of it which directly adjoin the investigations of the English statisticians, will be published in the next issues of *Biometrika*. [5] Those, however, which are most closely connected with the Lexian direction, I shall have the honor of reporting to you this evening. "On the theory of stability of statistical series", - this is how I entitled my work. From the very beginning of the development of modern statistics, the issues, connected with the amazing permanency of materials founded on mass observations, constitute the real focus of all the statistical theories that claim to represent something more important than a gathering of technical prescriptions and methods of collecting and summarizing the data. The merit of discovering proper ways for rationally justifying that permanency is due to those, who took the stochastic stand in the understanding of statistical regularities. I hope that you will allow me to use the beautiful Bernoulli term *stochastics* that had regrettably remained forgotten until Bortkiewicz (1917, p. x) recently resurrected it and applied it for designating the explanation of empirical data founded on the theory of probability and the law of large numbers.

Two centuries have passed since Bernoulli had derived the main relation between statistically established recurrences of events and their probabilities. At first, only mathematicians had become interested in his brilliant idea, but then it began its triumphal procession that gradually shaped it into a most important foundation of modern science. En route, De Moivre and Laplace, while keeping to the Bernoulli concept, provided the next incentive. Poisson made a very important step forward by proving that the interconnection between probability and observed frequencies did not demand the existence of one and the same probability in all cases. Cournot, the well known French philosopher and the real founder of the modern philosophy of statistics, justified his entire system of theoretical propositions by the Poisson finding, and Bienaymé, his contemporary, mathematically developed all that, which was contained in the initial {Bernoulli?} concept. For statisticians, however, the importance of that step {made by Poisson?} had remained incomprehensible. Even Bienaymé himself, although manifesting most subtle understanding of the theory, did not know how to apply it when acting as a statistician and getting down to work with concrete statistical data.

It was the French actuary Dormoy, almost a contemporary of Lexis, who threw the bridge from the abstract constructions of the mathematicians to the concrete problems of the statisticians. The formal priority is due to Dormoy, but his contemporaries have not noticed his contribution to the theory of stability. Lexis it was from whom science obtained this new method {of studying statistical series} and who discovered his predecessor and indicated his deserved honorable place in the history of statistics. Strictly speaking, we perpetrate some injustice when discussing the "Lexian theory of dispersion"; it would be more proper, more correct from the historical point of view, to call it after Dormoy and Lexis. [6]

The ideas of Lexis essentially influenced the contemporary development of statistics. For the next quarter of the century his works had remained the only source, from which a statistician interested in the theory of his discipline, could have quenched his spiritual thirst. In all countries, statisticians began to construct diligently the theory of statistics on that foundation. Bortkiewicz most prominently participated in that work. To him, the Lexian theory of dispersion owes a far-reaching generalization in the formulation of the issue as well as an essential improvement of the mathematical methods of its development.

The construction that grew little by little on the Lexian basis amazes spectators by the soundness of the internal connections between all of its parts, but its external appearance retains clear traces of gradual development. Indeed, no plans thought out beforehand were available when that stately edifice was being built, and it still lacks a common front. The stochastic theory of dispersion can find a draft for that façade by studying Markov's masterly investigation of the law of large numbers.

Let us consider a variable X that can take different values $\xi_1, \xi_2, \ldots, \xi_s$ with probabilities $p_1, p_2, \ldots, p_s$ respectively. We shall call the system consisting of $\{ \xi \}$ and $\{ p \}$ the law of distribution of the values of the random variable X, or, in short, the distribution of X. We denote the expectation of X by EX so that

$$EX = \Sigma p_i \xi_i.$$

Suppose now that n trials are made concerning $X_1, X_2, \ldots, X_n$, variables which can possess differing laws of distribution and be anyhow connected one with another, each of them participating once in each trial. Denote also [7] the value of X_h in a given trial by x_h. In addition, I introduce the following notation:

$$m_{1h} = Ex_h; \quad m_{2h} = Ex_h^2; \quad \mu_{2h} = m_{2h} - (m_{1h})^2;$$

$m_{[1,n]} = (1/n)[m_{11} + m_{12} + \ldots + m_{1n}]$;

$x_{(n)} = (1/n)(x_1 + x_2 + \ldots + x_n)$,

the mean of the observed values. We find at once that

$$\mathrm{E}x_{(n)} = (1/n)(\mathrm{E}x_1 + \mathrm{E}x_2 + \ldots + \mathrm{E}x_n) =$$

$$(1/n)(m_{11} + m_{12} + \ldots + m_{1n}) = m_{[1,n]}.$$

Thus, the expectation of the mean of the observed values is always equal to the mean of the expectations of the respective variables.

In the same way we obtain the mean square error of the mean

$$m^2(x_{(n)}) = \mathrm{E}(x_{(n)} - m_{[1,n]})^2 = \mathrm{E}x^2_{(n)} - m^2_{[1,n]} =$$

$$(1/n^2)\{\mathrm{E}(x_1 + \mathrm{E}x_2 + \ldots + \mathrm{E}x_n)^2 - (m_{11} + m_{12} + \ldots + m_{1n})^2\} =$$

$$(1/n^2)[\mathrm{E}x_1^2 + \mathrm{E}x_2^2 + \ldots + \mathrm{E}x_n^2 + 2\mathrm{E}x_1 x_2 + \ldots + 2\mathrm{E}x_{n-1} x_n -$$

$$(m_{11})^2 - (m_{12})^2 - \ldots - (m_{1n})^2 - 2m_{11}m_{12} - \ldots - 2m_{1n-1}m_{1n}] =$$

$$(1/n^2)\sum (m_{2h} - m_{1h}^2) + (1/n^2)\sum_h \sum_g (\mathrm{E}x_h x_g - m_{1h}m_{1g}) =$$

$$(1/n^2)\sum_h \mu_{2h} + (1/n^2)\sum_h \sum_g (\mathrm{E}x_h x_g - m_{1h}\, m_{1g}) =$$

$$(1/n^2)\sum \mu_{2h} + (1/n^2)\sum_h \sum_g (\mathrm{E}x_h x_{h+g} - m_{1h}\, m_{1h+g}). \qquad (1)$$

This is the main formula reproducing, with an unimportant change, one of Markov's formulas [8]. In spite of its simplicity, it contains the entire modern theory of stability. When issuing from it, it is only necessary to ascertain exactly the meaning of "trial" for obtaining all the other essential formulas of the theory. Assume at first that the variables can only take values 1 and 0 with probabilities $p_1, q_1, p_2, q_2, \ldots, p_n, q_n$, respectively. Keeping to this traditional point of view, consider some random variable A that occurs when one of the [initial] variables takes value 1, and does not occur when the latter is 0. Then $x_{(n)}$ will be the relative number of cases in which A occurs in n trials so that we have

$$\mathrm{E}x_h = m_{1h} = p_h;\quad \mathrm{E}x_h^2 = m_{2h} = p_h;$$

$$\mu_{2h} = m_{2h} - (m_{1h})^2 = p_h q_h. \qquad (2)$$

Denote, further, the probability that in some trial both x_h and x_g are equal to 1 by p_{hg}. Then

$$\mathrm{E}x_h x_g = p_{hg}.$$

Introducing these magnitudes in our main formula we get

$$m^2(x_{(n)}) = (1/n)\sum p_h q_h + (1/n^2)\sum_h \sum_g (p_{hg} - p_h p_g).$$

This is the well known Bohlmann formula (1909, p. 262) that he derived in a rather complicated manner by applying the Laplace method of generating functions.

Before considering the derivation of the other formulas of the theory by issuing from our general formula I wish to dwell on some of the latter's direct corollaries. I bear in mind the assumptions leading to the decrease of the error of the mean as the number of the trials increases, to its tending to zero as that number increases unboundedly. Most statisticians are apt to rely blindly upon the proposition that the random fluctuations of statistical numbers must decrease when the number of trials increases and that we are thus always able to attain any degree of certainty. For the practitioner, the theoretical law of large numbers is changed into an indisputable directive: Collect numerous observations and forget to care about anything being not in full order. Our formula shows, nevertheless, that the matter is not at all as simple as that.

The error of the mean consists, as we saw, of two parts. The first only depends on the laws of distribution of the variables, and, to be sure, decreases as n increases because its denominator then increases in proportion to n^2 whereas its numerator is proportional to n. But the second term of the error, determined by the degree of interdependence between separate variables, behaves differently. Here again, the number n increases the denominator proportional to its second degree, but the double sum in the numerator is also proportional to n^2. And it can easily happen that the fraction will not decrease as n increases. The validity of the law of large numbers is thus not at all obvious in itself; on the contrary, it depends on the difference

$$\mathrm{E}x_h x_g - m_{1h} m_{1g} \qquad (3)$$

and should be checked each time anew.

If the separate observations are independent one from another and all the differences (3) are zeros, the second term of the right side in formula (1) vanishes and the law of large numbers is applicable. The same also happens when all the differences are negative. If, however, they are positive, the matter is different. Suppose that the totality [the sum] of the differences is larger than some given magnitude ζ^2; then the mean square error (1) cannot become smaller than ζ^2 even if the number of trials becomes infinitely large.

If all the magnitudes (2) are equal one to another, and equal to μ_2, and all the differences (3) are also equal, and equal to r ($r > 0$), then the error (1) will decrease to

$$\mu_2 / n + [1 - (1/n)]\, r = r + (\mu_2 - r]/n.$$

It will therefore remain constant and be larger than r because $(\mu_2 - r)$ cannot become negative. And so, special assumptions for the applicability of the law of large numbers are required for the case in which all the differences (3) are positive. Two points deserve to be specially considered. If the difference (3) does not depend on h but depends on j, i.e., on the interval between the respective trials, it might be supposed equal to r_j, so that

$$m^2(x_{(n)}) = (1/n^2)\sum_h \mu_{2h} + (2/n^2)\sum_{j=1}^{n-1} (n - j)\, r_j. \qquad (4)$$

Thus, all depends on the mathematical type of the dependence between r_j and j. If, for example, the intensity of the interconnection decreases with an increasing j in a geometric progression, the second term in the right side of (4) tends to zero with an increasing n and the law of large numbers is therefore applicable. It is also applicable when the difference (3) vanishes for all the values of j exceeding a given finite magnitude; in other words, when a separate trial ceases to influence other trials as soon as the appropriate interval begins to exceed a definite magnitude.

Many instances in which the law of large numbers is valid in spite of all the differences remaining positive can indeed be ascertained. Of special interest is the case when the differences are partly positive and partly negative and their sum might be both positive and negative or equal to zero if these differences exactly compensate each other.

Thus, in spite of the existence of a close connection between trials, $m^2(x_{(n)})$ can be exactly equal to that magnitude, which corresponds to their mutual independence. This conclusion is of special importance in the light of the main ideas of the Lexian theory of dispersion, and I shall return to it after inviting your attention for a short time to the derivation of various formulas of the theory by issuing from our general formula.

Let us consider the following case. An urn contains n_1 tickets with number ξ_1, n_2 tickets numbered ξ_2, …, n_k tickets with number ξ_k and U tickets are drawn without replacement. Suppose now that

$$n_1 + n_2 + \ldots + n_k = N$$

and denote the number occurring at the h-th drawing by x_h. Then, as it is easy to show, the expectations of each of these numbers are

$$\mathrm{E}x_h = (n_1/N)\,\xi_1 + (n_2/N)\,\xi_2 + \ldots + (n_k/N)\,\xi_k = m_1,$$

and, in addition,

$$\mathrm{E}x_h^2 = (n_1/N)\xi_1^2 + (n_2/N)\xi_2^2 + \ldots + (n_k/N)\xi_k^2 = m_2.$$

The estimate μ_2 will thus remain the same for each drawing and be equal to $(m_2 - m_1^2)$. All the differences (5) are also equal to each other, and equal to $-\mu_2/(N-1)$. Our main formula therefore provides

$$\mathrm{E}x_{(n)} = m_1,$$

$$m^2(x_{(n)}) = \mu_2/U - [(U-1)/U]\,[\mu_2/(N-1)] =$$

$$(\mu_2/U)[1-(U-1)/(N-1)] =$$

$$[(N-U)/(N-1)]\,\mu_2/U.$$

For drawings without replacement the mean square error thus decreases in the ratio $\sqrt{N-U}/\sqrt{N-1}$. In this simple case we obtain a confirmation and generalization of the well known Pearsonian formula derived by him when studying hypergeometric series (1899; 1928). This generalization has practical importance in that our formulas enable the estimation of the reliability of the results of the so-called sampling method, and of that of its version which practitioners quite justifiably prefer. Thus, in Russia, during an agricultural census in 1916, about 18 *mln* forms carrying very detailed data were completed. The census was repeated in 1917 in accord with a still more detailed program. It was only possible to make use of that excessively vast material in full by applying a sample investigation answering an expedient pattern. One of my students provided responsible indications. Under my supervision he compiled a plan of the sampling based on these formulas. He directed a part of the necessary preliminary work and the already begun investigation promises to become a good example of the importance that the mathematical theory can have for practice, at least with regard to work input and financial expenses. The statistical organization that was commissioned to carry out {or, perhaps, to process the materials of ?} those censuses had regrettably vanished in the whirlwind of revolutionary events, but we can only be surprised at the good fortune of the appearance, notwithstanding, in the autumn of 1917, of a little part of our investigation, of the salvation of something else in addition to its theoretical introduction [9].

I shall not bother you now with a detailed derivation of the other formulas of the theory of dispersion by issuing from our main formula. The well known Lexian expression for the supernormal dispersion showing how the mean square error is separated into an "essential" and a "random" components; the Bienaymé – Bortkiewicz formula for the case of the so-called "acute solidarity of separate trials [10]; the formula for the probability of a constant structure, – we derive all of them without exception from our simple formula that represents a quite sufficient and an all-embracing mathematical basis of the entire theory of dispersion. The derivations are quite simple, but an oral presentation can prove tiresome. My time will be better used up if I thoroughly expound the logical side of the issue rather than dwell on mathematical demonstrations that can be simpler and easier examined from short published sources.

I am thus turning to the principal conclusions resulting from the already ascertained possibility that in the case of mutually dependent magnitudes the mean square error of the mean occurs to be the same as when these magnitudes are mutually independent. Lexis is known to have defined normal stability as such that occurs when the probability of the event remained constant and all the trials were independent one from another. His theory of dispersion recognizes the stability of a given series of statistical numbers as normal as soon as the actual mean square error occurs, under the abovementioned assumptions, to be equal to its expectation. Such a method of investigation is not quite proper because the statement that the expectation of that error, under normal dispersion, has a given value, does not admit, as we saw, an unconditional inversion: the dispersion is normal when the expectation of the mean square error has a given value. There remains a possibility that the dependence existing between separate trials is such that the positive and negative differences (5) indeed compensate each other. Utmost caution is required so as to allow for that case, and always to prove rigorously the existence of a normal dispersion [11].

As you see, the usually applied method for ascertaining whether the dispersion is normal or not should be essentially improved. That method is based on the calculation of the so-called coefficient of dispersion defined as a fraction having as its numerator the mean square error of the actual data, and, for its denominator, the expectation of that error under normal dispersion. If the dispersion is indeed normal, the expectation of the coefficient of dispersion is unity. A sufficiently large deviation from 1 shows that the dispersion is not normal. Until now, all the deliberations are irreproachable. In practice, however, the coefficient of dispersion is rarely calculated in accord with that definition: the theoretical values in its denominator are replaced by those empirically derived. The expectations in both cases are the same if all the assumptions of the normal dispersion are realized, but it does not at all mean that the expectation of the fraction is unity. Thus, $\mathrm{E}x$ can take the same value as Ey, but, in spite of that, $\mathrm{E}(x/y)$ can have any value, both greater and smaller than unity. That also $\mathrm{E}(x/y) = 1$ ought to be additionally proved in each case; until then, we have no right to make such a statement.

For a long time I have been engaged in ascertaining the conditions for the expectation of the coefficient of dispersion to equal unity. In 1914, I finally discovered the answer based on a general lemma

$$\mathrm{E}(x/y) = \mathrm{E}x/\mathrm{E}y \text{ if } \mathrm{E}(xy^h) = \mathrm{E}\, y^{h+1} \text{ for } h = 0, 1, 2, 3, \ldots$$

I communicated my result to Markov who offered its direct proof. Later on I was able to find an extremely simple third proof of a very general nature [12].

Denote the square of the mean square error calculated in accord with the "physical" method, as Lexis called it, by m_f, and the same magnitude derived "combinatorially", by m_k. Then the square of the Lexian coefficient of dispersion will be m_f/m_k. It will then be easy to show, by a method which I derived, that for a normal dispersion and any integral h, both positive and negative,

$$\mathrm{E}\, m_f\, \mathrm{E}\, m_k^h = \mathrm{E}\, m_k^{h+1}.$$

For $h = -1$, we have $\mathrm{E}(m_f/m_k) = 1$ which indeed means that the expectation of the coefficient of dispersion is unity.

It followed that the Lexian assumption was absolutely correct; I ought to add, however, that the matter would have been different should have he divided, instead, m_k by m_f. His choice was especially lucky; or, rather, it was his scientific intuition that prompted him to make the proper choice.

The method, that I applied for discovering that the expectation of the coefficient of dispersion in case of normal dispersion is equal to unity, can also be used for deriving the mean square error, $m(Q^2)$, of that coefficient. This more rigorous method shows that, although the exact value of $m(Q^2)$ depends on the law of distribution of the relevant variable and on the number of observations in separate series, it is still possible, without making a large error, to apply the Bortkiewicz approximate formula which only allows for the number of series, call it r. In accord with that formula, we assume that

$$m(Q^2) = 1/\sqrt{r-1}.$$

Under the most often encountered conditions, the mean square error will be somewhat exaggerated but nevertheless usually not great. However, when the so-called law of small numbers is valid, Bortkiewicz' approximate formula becomes absolutely useless; the real error might be either greater or less than the value it indicates. The difference can be very significant, so that, in such cases, we should hardly trust the conclusions based on the calculated mean square error. There is regrettably nothing by which to replace that unreliable formula: neither my, nor Markov's method is applicable under such conditions. The problem of estimating the mean square error of the coefficient of dispersion, when the law of small numbers is valid, is fraught with mathematical difficulties still awaiting to be overcome.

Since I have not surmounted them, the best way is to decline from the derivation of the coefficient of dispersion in all such cases. Instead of the fraction m_f/m_k it is then possible to calculate the difference $(m_f - m_k)$ and to consider that the dispersion is normal when it vanishes; that it is supernormal (subnormal), when the difference is positive (negative).

In essence, this method should also be preferred when the mean square error of the fraction m_f/m_k can be calculated. The reasons why Dormoy and Lexis chose the fraction are not more valid now then they were in their time. They, the reasons, might be reduced, as I think, to two arguments.

1. The deviations of the fraction from 1, and especially those towards 0, are more expressive than the positive and negative differences whose magnitudes depend to a considerable degree on the number of observations.

2. Dormoy, and, in the beginning of his work, Lexis, thought that the magnitude of the deviations of the coefficient of dispersion from 1 will serve as a convenient measure of the difference between the actual and the normal dispersion.

The first argument lost its previous demonstrative power as soon as it became clear that in both cases, before coming to any conclusions, the deviations should be divided by the respective mean square error. As to the second argument, Lexis himself, in a later writing, showed that the magnitude of the coefficient of dispersion cannot serve as a measure of the deviations from normal dispersion. And it was none else than Lexis who proposed to calculate the difference ($m_f - m_k$), the so – called *essential component of variability*, which, in the typical case of supernormal dispersion, enables to isolate the mean square deviation of the basic probabilities themselves.

Nowadays, this method alone of investigating dispersion is justified. Indeed, it can be applied to any kind of statistical numbers, -- to such variables that obey any law of distribution, or take one of only two possible values. In the latter case, there exists a formula for the mean square error of the essential component, applicable under practically encountered conditions. It should be borne in mind first of all that almost always we have to do with series consisting of very different numbers of observation and that a very small number of the observations exist for most of the studied series. In such cases, the approximate formulas that assume a very large number of observations are useless. On the contrary, we require formulas valid for any number of observations, large or small.

Both problems can be solved without great difficulties by the same method with which we acquainted ourselves when deriving the main formula for the mean square error of the mean. And the formula that enables the calculation of the essential component also provides the possibility of obtaining its mean square error in all cases that can be encountered in practice. Let us ask, for example, can the sex of a newly born depend on the ages of its parents? A countless number of investigations had been at one time devoted to that issue, but, regrettably, we now know not more about it than during the times of Hofacker (1788 – 1828) and Sadler (1780 – 1835) [13]. The reason why the numerous researchers engaged in that problem achieved such poor success depended mostly on the lack of a suitable method. Those, with which they were satisfied, were too primitive. The influence of the parents' age on the probability, that a newly born will be a boy, cannot be revealed by simple means; more subtle methods are required. It is the calculation of the essential component that should be applied here. This method can solve a problem in all cases in which we have sufficiently subdivided groups of data. Such, for example, are the tables for Berlin providing the yearly data for the last 50 years on newly born boys and girls for about 1,500 age groups [of parents]. The mean square error of the essential component for such a material is so insignificant that a smallest discrepancy between the probabilities would not have escaped notice. The method is only accompanied by one difficulty: it demands so much calculations that a single investigator would have hardly reached the solution.

It is possible to develop the idea underlying the Lexian method. The essential component provides the magnitude of the mean square deviation [between] the main probabilities. A still deeper analysis is possible by calculating the third, the fourth, … moment. The calculation of the third moment might be recommended even in those cases in which the essential component vanishes, and the dispersion seems normal. If it occurs that the third moment differs from zero, the actual dispersion cannot be normal; it is then pseudo-normal, and assumes such an appearance exactly because the positive and the negative differences discussed above have indeed compensated each other.

Issuing from this general essay on the new Lexian method of investigating dispersion, we are to make a most important step towards English statisticians. Investigations of dispersion can pursue both a qualitative and a quantitative aim. It is possible to restrict our attention to forming a general idea for ourselves about the stochastic pattern suitable for explaining the actual dispersion of the numbers under investigation. But it is also possible to aim at determining those constants which characterize that pattern. The calculation of the coefficient of dispersion only bears in mind a qualitative characteristic because the values of that coefficient depend on the number of observations and are therefore of little use for describing the conditions themselves of dispersion. On the contrary, the calculation of the essential component ascertains the most important quantitative characteristic of the stochastic pattern, the mean square deviation of the main probabilities.

As far as the English statisticians are concerned, they, from the very beginning, were engaged in the quantitative side of the issues of theoretical statistics. The qualitative side, – I bear in mind the stochastic assumptions that serve as the basis of quantitative observations, – mostly remains without such notice as it deserves. Consequently, English constructions often absolutely lack a sound foundation that is notable for the Lexian theory of dispersion. Then, it is necessary to stress that, contrary to the Lexian school oriented towards sociology, English statisticians, oriented towards natural sciences, only pay attention to problems in interpolation.

But as soon as English statisticians enter this purely stochastic field, the acute contradiction between the English and the Lexian directions, stressed for example, by Bortkiewicz, will, as it seems to me, become absolutely groundless. 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 empiricism nowadays predominating in statistics, and crave for a rational theory constructed on a sound stochastic basis.

Notes

1. {Two years after Chuprov had read his report, Cramér became an actuary in Sweden, see beginning of his autobiography (1976).}

2. {Chetverikov (1926, p. 318) stated that Chuprov, after declining to return to Soviet Russia from his scientific trip to Europe, lived for three years in Oslo and Stockholm.}

3. The end of the saying is *tamen est laudanda voluntas*, - desire worthy of praise is present.

4. See Note 2.

5. {See Chuprov (1918 – 1919, 1921).}

6. {Later on Bortkiewicz (1930) stated that the theory of dispersion was mostly due to Lexis.}

7. {The text was corrupted by many misprints. I have corrected it in accord with Chuprov's main relevant work (1918 – 1919) published just after he had read his report.}

8. {Chuprov never provided an exact reference. See some relevant considerations in Sheynin (1996, p. 115).}

9. The student mentioned above was Stanislav Salsievich Kon. A summary of his manuscript was published (1922); also see Kon (1917). {Chetverikov (1920) reviewed this latter writing. I have not found Kon (1922) in the source indicated by Chetverikov.}

10. {I suspect that no such formula existed; read "Bienaymé – Bortkiewicz pattern": probability of success remains constant within a series but differs from one series to the next one. The two terms (the second one is mentioned just below), as Bortkiewicz (1894 – 1896, 1895, p. 332 and 1894, p. 650) called them, were *Solidarität der Einzelfälle gewissermaßen in acuter Form* and *konstant zusammengesetzte Durchschnittswahrscheinlichkeit.*}

11. {Chuprov (1923) and his student Mordukh (1923) proved that in the absence of *prior information* about the precise value of the expectation or about the nature of the connection between the observations the Lexian test Q was unsatisfactory.}

12. {See Chuprov (1916, p. 1791; 1921, p. 270).}

13. {I found these dates in the Index of Names appended to Chuprov (1960).}

References

Abbe, E. (1878), Über Blutkörpernzählung. *Ges. Abh.*, Bd. 1.Hildesheim, 1989, pp. 173 –180.

Bohlmann, G. (1909), Die Grundbegriffe der Wahrscheinlichkeitsrechnung in ihrer Anwendung auf die Lebensversicherung. *Atti IV Congr. Intern. Matem. 1908.* Roma, 1909, t. 3, pp. 244 – 278.

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. 701 – 705.

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

--- (1917), *Die Iterationen.* Berlin.

--- (1930), Lexis und Dormoy. *Nord. Stat. J.*, vol. 2, pp. 37 – 54.

Chetverikov, N.S. (1920), Review of Kon (1917). *Vestnik Statistiki*, No. 5 – 8, pp. 1 – 6. (R)

---(1926), Al. A. Tschuprow. *Metron*, t. 6, No. 3 – 4, pp. 314 – 320.

Chuprov (Tschuprow), A.A. (1916), Sur l'espérance mathématique du coefficient de divergence. Translated in this collection.

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

---(1918 – 1919, 1921), On the mathematical expectation of the moments of frequency distributions. *Biometrika*, vol. 12, pp. 140 – 169 and 185 – 210; vol. 13, pp. 283 – 295.

---(1922), On the mathematical expectation of the ratio of two mutually

dependent random variables. Translated in this collection.
--- (1923), Ist die normal Stabilität empirisch nachweisbar? *Nord. Stat. Tidskrift*, Bd. 1, No. 3 – 4, pp. 369 – 393.
--- (1960), *Вопросы статистики* (Issues in statistics). M. Reprints / translations. In addition to the translation of the present work, contains translation of Chuprov (1923).
Cramér, H. (1976), Half a century with probability theory. *Annals Probability*, vol. 4, pp. 509 – 546.
Kon, S.S. (1917), *К вопросу о применении выборочного метода при подготовке сельскохозяйственных переписей* (On the application of sampling to agricultural censuses). Petrograd.
--- (1922), Dispersion of the elements of agriculture as covered by the census of 1916. Summary. *Vestnik Statistiki.*
Mordukh, Ja. (1923), On connected trials complying with the condition of stochastic commutativity. *Trudy Russkikh Uchenykh za Granitsei*, vol. 2. Berlin, pp. 102 – 125. (R) Transl.: DHS 2696, 2000, pp. 209 – 233.
Pearson, K. (1899), On certain properties of the hypergeometrical series, etc. *London, Edinburgh and Dublin Phil. Mag.,* ser. 5, vol. 47, pp. 236 – 246.
--- (1928), On a method of ascertaining limits to the actual number of marked members from a sample. *Biometrika*, vol. 20A, pp. 149 – 174.
Sheynin, O. (1996), *Chuprov.*(1990, in Russian). Göttingen.
Student (Gosset, W.S.) (1907), On the error of counting with a haemacytometer. *Biometrika*, vol. 5, pp. 351 – 360.

8. La décomposition du bolchevisme

N.p., n.d., no title page. Signed on last page (p. 16) A. Tchouprov, Professeur dÉconomie politique à l'Université de Moscou. Reprinted from copy kept at the Bibliothéque Nationale de France. Extremely rare.

Foreword

{See below comments on the subject of this pamphlet whose copy I discovered in the *Bibliothèque National de France* where it is kept as a rare edition and recorded as *Don 188205*, but I was unable to find out who had donated it or when was it received. This lack of information is really regrettable because the pamphlet was apparently put out in a very few copies (I am unaware of any other of them). Moreover, the author's name, only provided at the very end of the pamphlet, is either a mistake or a deception: the Professor of political economy in Moscow was Aleksandr Ivanovich Chuprov (rather than his son), who, furthermore, died in 1908.

I [1, pp. 251 – 252; 2, p. 23] have also quoted Chuprov's denouncement of Lenin and his declaration that Bolshevism as a political idea was dead. Now, the real author of the pamphlet was A.A. Indeed:

1) A.A. is the author of many reviews of various sources and essays on economic issues and in 1919 he lived in Stockholm.

2) Gulkevich (see his note in this collection) testified that Chuprov had been *attentively* following the events in post-revolutionary Russia.

3) Chuprov left many friends, colleagues and students in Russia and he would not have published such a pamphlet in the usual way. It is even possible that its copy was deposited in Paris after his death.

4) In 1919, Chuprov had no political opponents (and hardly ever any enemies) so that no-one would have wished to compromise him in the eyes of the Soviets. In addition, any provocateur would have abstained from spending so much time needed for composing such a pamphlet.

I also believe that Chuprov signed the pamphlet by his father's name out of caution but that the Soviet authorities, had they seen the pamphlet, would have undoubtedly understood the truth.

References

1. Sheynin, O. (1993), Chuprov, Slutsky and Chetverikov: some comments. *Hist. Math.*, vol. 20, pp. 247 – 254.
2. --- (1996), *Chuprov: Life, Work, Correspondence*. Göttingen.}

* * *

1.1. La Russie des Soviets est séparée du monde extérieur par des cloisons plus étanches encore qu'au temps de la guerre. Seuls quelques numéros dépareillés des journaux bolchevistes franchissent de temps en temps la frontière, et complètent assez facilement les communiqués officieux de *Rosta* {Russian Telegraph Agency} avec leurs silences et leurs exagérations malhabiles. Pour reconstituer le tableau de la véritable situation, il faut recourir principalement aux récits des émigrants. La source est abondante, elle permet de se faire une idée assez exacte de certaines choses. En Suède, par exemple, on peut se tenir assez bien au courant des conditions de la vie en Petrograd, grâce aux informations appuyées sur l'expérience personnelle du narrateur. Mais dès qu'on veut élargir le cadre et franchir les limites des impressions directes, les dépositions deviennent confuses, et, malgré l'assurance du ton, peu sûres. Les réflexions des gens sur les événements ne correspondent guère à la réalité, même en temps normal. Or, dans la Russie des Soviets, la presse est bâillonnée; il n'y a pas de communications postales régulières; le service des trains de voyageurs est ralenti – pour ne pas dire plus: l'infatigable inquisition policière oppresse les habitants des capitales qui ne se risquent pas à continuer les relations d'amitié dans les formes habituelles d'auparavant. Impossible, dans ces conditions, de savoir ce qui se passe de l'autre côté de la barrière. Même sur place on ne vit que de nouvelles sensationnelles et contradictoires. Ce qui filtre par la frontière n'est également qu'un ramassis de bruits inconsistants, qui s'est déposé dans la tête de gens dépourvus de critique, épuisés de souffrances et dont les nerfs sont à bout. Faire le départ dans les dépositions et témoignages enregistrés par nous, entre les faits bien établis et les couleurs ajoutées par l'imagination est une tâche malaisée; rien d'étonnant donc si les nouvelles "de source autorisée" sont parfois une invention et si nous sommes surpris par le cours réel des événements.

1.2. Le hasard m'a procuré une assez riche collection de publications économiques bolchevistes – *Izvestia* (Bulletins) de commissariats divers et de *Glavki* (Directions) et *Centres* de toute espèce. Malheureusement, il est impossible de se faire d'après eux une idée tant soit peu claire de la situation de l'économie nationale dans la République des Soviets. Les pages des revues que j'ai sous les yeux sont remplies de décrets, de projets, d'hypothèses et de considérations, de procès-verbaux des commissions, conférences et congrès, mais d'une pauvreté lamentable en informations de fait sur la vie économique de la Russie, si l'on excepte ce qui est puisé aux sources dont les "dictateurs" actuels sont redevables à la période "bourgeoise" de l'histoire de la pensée et de l'étatisme russes. En feuilletant toutes ces publications on se convainc d'une chose: c'est que les directeurs mêmes de la politique économique des Soviets, "qui organisent d'un centre unique" la production et la répartition sur tout le territoire de la Sovdépie {degrading abbreviation of [the country of the] Soviets (= of councils) of deputies} russe, agissent tout le temps à l'aveuglette, faute de savoir recueillir d'une manière passable les informations relatives à la véritable situation dans le pays. L'intérêt principal des matériaux parcourus par moi réside moins dans les indications fragmentaires qu'on peut attraper cà et là sur la marche objective des affaires dans le pays que dans le tableau des variations survenues dans l'état d'esprit de cette partie des "sphères dirigeantes bolchevistes" qui n'a pas complètement perdu la faculté de penser et se rend compte de ce qui se passe.

1.3. Jusqu'au début de l'été 1918, on sent une fièvre universelle de création sur le papier, "un enthousiasme administratif" devant lequel pâlissent les charges les plus savoureuses de Chtchédrine {Saltykov-Shchedrin}. Le coup d'Etat nous a donné le pouvoir, c'est-à-dire nous a installés dans les ministères devant les tables vertes. Et dans la première ivresse des perspectives qui s'ouvrent à eux, les "hommes nouveaux" placés au gouvernail pour "créer une vie nouvelle" se perdent avec un enthousiasme presque touchant dans la paperasserie. On voit renaître un "bureaucratisme" d'une envergure que nos ministères d'affaires avaient su oublier depuis longtemps, et dont une faible image ne survivait "sous le régime tsariste" que dans les coins perdus du mécanisme administratif. Pour des raisons d'ordre privé, les hôtes nouveaux des chancelleries de "départements" n'avaient pas encore eu l'occasion de se faire à la cuisine de la créa on gouvernementale; tout ici est nouveau pour eux; tout les intéresse, comme un joli jouet qu'on a reçu en cadeau et qu'on ne sait pas faire fonctionner. On presse à l'envi un bouton, puis l'autre, dans l'espoir anxieux qu'à la fin un miracle s'accomplira et que tout marchera.

Le miracle pourtant ne s'accomplit pas; rien ne marche, malgré l'ardeur au "travail" des faiseurs de décrets et le zèle des copistes qui ne peuvent venir à bout des matériaux accumulés. Les esprits plus rassis commencent à s'alarmer; et la foi n'est plus aussi forte en la possibilité, pour une poignée de personnes, d'imposer un cours arbitraire à la vie d'un peuple de cent millions d'habitants à l'aide de décrets, même les plus menaçants.

L'expérience éveille la conscience qu en accumulant les paragraphes de décrets, on ne fait pas la moitié du travail, qu'il faut dans la vie des engrenages pour que les fantaisies fixées en paragraphes laissent des traces ailleurs que sur le papier, que la création bureaucratique fait songer terriblement à un moteur d'où on n'a pas fait passer la courroie de transmission au mécanisme travailleur. Depuis le milieu de l'été 1918, un courant de réflexion qui s'éveille commence à se manifester de plus en plus nettement dans les discours et les écrits d'une partie des économistes soviétistes; il grandit peu à peu, se transforme en un sentiment de dégrisement amer et une sorte de peur d'avoir à pâtir pour l'issue d'expériences si légèrement entreprises.

On peut noter clairement deux sources principales de désillusions et d'inquiétudes: l'une politique, peut-on dire, l'autre technique. Les deux se ramènent à ceci: il faut faire des concessions; en tenant compte de la réalité, il faut renoncer à ce programme "prolétarien", qui, dans l'ivresse de la victoire, paraissait si aisément réalisable.

2. Les bolcheviks ont pris le pouvoir pour mettre en vigueur une dictature de classe, celle du prolétariat. La Russie est un pays agricole, mais le paysan a peu de terre et il n'a pas encore oublié l'affront fait à sa classe quand, lors de l'émancipation {in 1861}, il n'a pas reçu la part qu'il espérait: avec une foi inébranlable dans son droit, il réclame instamment la remise des domaines des propriétaires fonciers entre les mains des paysans. Le prolétariat des usines et des fabriques est, en Russie, numériquement faible, sa culture est incomplète, il ne dispose pas d'organisations professionnelles solides, il n'a pas de chefs ayant passé par l'école de la lutte de classe quotidienne. Un prolétariat des campagnes ayant une conscience de classe prolétarienne manque absolument; l'ouvrier agricole, qui par sa psychologie est le paysan même, ne rêve pas du socialisme, mais de la bonne terre, du lopin sur lequel il s'installera et qu'il exploitera. On conçoit difficilement la psychologie de ceux qui, se débrouillant tant soit peu dans les rapports économiques, peuvent avoir une foi sincère dans la dictature du prolétariat en tant que menant au triomphe du socialisme en Russie. On ne peut trouver d'explication que dans la fureur révolutionnaire, dans un tempérament déchaîné, enlevant toute la capacité de réfléchir. Naturellement ce sont des particularités de tempérament qui expliquent encore la majorité écrasante des Israélites dans les premiers rangs du bolchevisme mondial. L'esprit russe, naturellement sain, n'est pas enclin à se bercer d'illusions, à s'abandonner pour longtemps à de pareils mirages, dès qu'il a un certain minimum de connaissances. Lénine lui-même, – il ne peut y avoir de doute à ce sujet pour ceux qui l'ont suivi au cours de sa "carrière" et qui se représentent quelque peu sa tournure d'esprit, – n'a pas cru une minute à la possibilité d'instaurer en fait la dictature du prolétariat en Russie, et ce n'est pas pour l'établir qu'il s'est emparé du pouvoir. En octobre 1917 {the coup d'Etat 25 October 1917, old style}, comme au cours de toute sa vie orageuse, Lénine a eu soif du pouvoir pour le pouvoir sans penser ni à la Russie ni au prolétariat russe: il s'intéressait uniquement à une expérience grandiose pratiquée *in corpore vili* {on a body of small value}, sur le peuple russe. L'échec final était clair pour lui: il fut, comme toujours, indifférent au sort des hommes qui le suivaient. "Quand on coupe le bois, les copeaux volent", dit un proverbe russe. Lénine est à la tête du groupe des intellectuels et demi-intellectuels qui travaillent dans les *souterrains*: des unités souffrent. Lénine est sur le trône de l'autocrate de la Grande Russie: ce sont des dizaines de millions qui souffrent. De telles bagatelles ne troublent pas la sérénité de "politiques réalistes" et ne leur font pas perdre leur sang-froid. Lénine a besoin de tout le sien pour tenir ferme le pouvoir en face de la haine croissante du pays et du manque enfantin de compréhension parmi ses émules grisés de leurs fonctions ministérielles, incapables de reconnaître leur impuissance et la profondeur de l'abîme où mène irrévocablement le désir de satisfaire sérieusement en Russie les soi-disant revendications de classe du prolétariat. Le prolétariat a pu conférer le pouvoir, mais ce n'est pas assez de son appui pour le conserver: il ne suffit pas non plus de décrets sur le papier et de discours flatteurs pour conserver les sympathies dudit prolétariat. Lénine l'a vu et très vite, mais il a longtemps été seul à le voir dans le cercle de ses collègues. Napoléon, idole d'une armée, a pu oser un coup d'Etat. Il n'y a pas derrière Lénine de force armée qui lui soit personnellement dévouée: les janissaires lettons-chinois ne le considèrent pas comme leur chef. Au moindre faux pas, le prolétariat qui a élevé sur le pavois le camarade Lénine, déchirera en pièces le traître Oulianof – et il sera dans son droit, car la ligne de conduite réaliste de Lénine – Oulianof repose incontestablement sur la trahison de la révolution bolchevico- prolétarienne. Lénine n'ambitionne pas la couronne du martyre, il ne rêve pas du sort de Rosa Luxembourg; aussi se conduit-il avec toute la prudence possible, découvrant un don enviable de jouer des hommes et des masses humaines. Lénine est persuadé que le temps travaille pour lui, car il a escompté les "prémisses objectives" du processus historique de la Russie. En attendant que les leçons de choses de la vie ramènent à la raison les enfants politiques qui l'entourent, il se permet seulement de temps en temps de leur administrer des douches froides dosées de main de maître, dans l'espoir de hâter la maturité de ceux qui n'ont pas perdu irrémédiablement la tête dans la tourmente révolutionnaire.

Dans quelle mesure réussit-il dans cette tâche, c'est ce qu'attestent les événements qui se déroulent dans la Russie des Soviets et c'est ce que disent éloquemment les revues étalées sur ma table. Je vais m'efforcer de caractériser maintenant le dégrisement graduel des bolcheviks "réalistes", dégrisement qui prépare un revirement dans la politique du pouvoir soviétiste.

3. Je commencerai à partir du moment politique où les hommes des Soviets ont pris peu à peu conscience de l'impossibilité pour eux d'être maîtres de la Russie, sans l'appui des paysans.

> *Avant tout, il faut noter ce qui explique et facilite les victoires des Tchéco-slovaques et des gardes blancs* {in the Civil War}: *ce sont les tendances toujours de plus en plus réactionnaires dans la masse des paysans propriétaires*

dit l'article de tête du *Izvestia Narodnogo Komissariata po Prodovolstviu* No. 14 – 15 {that continues thus}

> *La politique de ravitaillement du pouvoir des Soviets va à l'encontre des intérêts des paysans riches* (*koulaki*). *La campagne riche ne désire pas nourrir les ouvriers d'industrie. La fixation des prix du blé, la lutte contre les porte-sacs et autres intermédiaires nuisibles, les réquisitions des réserves de blé, le crainte de ne pas réussir à réaliser la bonne récolte prochaine comme on a réalisé l'ancienne, mais d'avoir à la livrer à l'Etat à prix ferme, tout cela* {two or three words undecipherable} *paysan-propriétaire les sentiments les plus malveillants à l'égard de la République soviétiste* […] {Chuprov's omission points} *Une fois mis en possession de la terre, le paysan propriétaire* (*et même une partie importante des paysans non propriétaires, demi-journaliers*) *ne s'intéresse plus aux ouvriers ni davantage à la ville, sinon dans la mesure où il peut en recevoir les marchandises indispensables. Si ces marchandises manquent ou si on lui en promet d'un autre côté à un prix semblable* (*et nous croyons que le gouvernement sibérien avec l'aide de ses hauts protecteurs d'au-delà des frontières entreprend déjà quelque chose dans ce sens*), *on peut envoyer promener la ville russe. Si cette même ville, au nom de quelques exigences incompréhensibles, essaîé de réquisitionner le bien des paysans riches* (*koulaki*), *alors, non seulement qu'on la trahisse, mais qu'on la détruise! En présence de ce nouvel ennemi qui prend davantage chaque jour conscience de lui-même et qui s'organise, le prolétariat doit décupler son travail d'organisation et son activité économique.*[1]

On commence, comme nous le voyons, à prendre conscience du danger. "La réaction" grandit chez les paysans, – on reconnaît le fait, mais on n'en tire pas encore la conclusion fatale. L'auteur s'efforce en outre d'esquiver des conclusions extrémistes: il cherche à rassurer avec les phrases ordinaires, et, devant l'orage menaçant, il n'ose rien proposer d'autre que le décuplement de l'énergie de ce prolétariat. Cependant la vie ne donne pas de répit, et dans le numéro suivant, des *Izvestia* No. 16 – 17, sous la signature en toutes lettres du rédacteur N. Orlov, avec une intrépidité et une clarté de pensée qui ne le cèdent pas aux meilleures productions du talent de publicité de Lénine, les idées "de lâchage" se développent déjà ouvertement:

> *Après le 25 octobre, l'armée, comme force révolutionnaire, s'est évanouie. Les paysans ont terminé leur mission révolutionnaire du jour où ils se sont emparés de la terre. Désormais, seule la crainte de perdre cette terre* [et] *l'absence de points visibles de divergence leur ont fait accepter ce pouvoir ouvrier, indocile et turbulent. Dès qu'apparaît une force nouvelle qui, en apparence ou en réalité, garantit sa conquête au moujik, la campagne trahit le prolétariat comme, en la personne de l'armée, elle a trahi la République sous Dvinsk, Pskov, Revel* {Tallinn}.
> *Seule la classe ouvrière qui a le moins reçu dans la nuit du 25 octobre, reste fidèle à l'ancien drapeau. Mais jusque dans ses rangs, il n'y a pas cette unanimité et cette énergie qu'exige le besoin du moment. Et moins il reste de possibilités objectives de développer le programme économique adopté par lui, plus il y aura de désordre et d'apostasie. Voyons par exemple la position des masses ouvrières dans les rayons occupés et séditieux: l'Ukraine, le Don, la Volga, l'Oural, la Sibérie … Même l'Oural rouge n'a pas voulu mobiliser*

la totalité de ses forces et a rendu sa capitale aux ennemis des Soviets. De tous ces faits malheureux et de beaucoup d'autres, il faut hardiment tirer les conséquences appropriées. Une fatigue mortelle, le désenchantement s'installent dans l'âme de cette classe, à qui personne, sinon elle-même, ne peut garantir l'intégrité de ses conquêtes d'octobre … Avec quelles revendications la classe ouvrière-industrielle russe est-elle arrivée au pouvoir? La paix, la terre, le pain, telles étaient les devises des journées d'octobre. Les représentants du pouvoir des Soviets à Brest-Litovsk {the peace treaty of 1918 is meant; its history is rather complicated} *ont annoncé que la Russie se retirait de la guerre et ont développé leur programme de paix… Cette sortie de la guerre fut comprise comme une possibilité de s'enrichir aux dépens de l'Etat. Les paysans n'ont pas voulu demeurer un seul jour sur le front, les ouvriers étaient trop faibles pour s'opposer aux armées allemandes. Et voilà que nous avons ce que nous avons… Il dépendait de nous d'avoir une paix meilleure, mais la trahison de la petite bourgeoisie ne nous a pas permis d'obtenir même un mauvais traité. Le motif est que l'autorité des Soviets n'a pas tenu sa première et principale promesse … La République des Soviets a rempli sa deuxième promesse, elle l'a remplie contre les intérêts réels du prolétariat: la terre est passée en fait à la libre disposition des petits propriétaires. Cette terre a-t-elle été payée par les paysans? Point! La petite bourgeoisie trahit de nouveau la République par le refus du blé. En égard aux circonstances, c'est un crime beaucoup plus grave que Dvinsk, Pskov, Mohilev, Oesel… Il faut regarder la vérité en face et parler net: la coalition ouvrière-paysanne d'octobre n'existe plus; il y a dictature du prolétariat attaquée de tous les côtes… Notre situation est dangereuse maintenant, surtout, par suite de l'opposition de la petite bourgeoisie, et de la possibilité pour des fractions étrangères infimes, de spéculer sur la haine du moujik contre nous et sur le mécontentement des masses affamées. Il en résulte que la politique de la logique à outrance, la politique d'isolement, et, par suite, du martyre volontaire, tout en étant funeste à la classe ouvrière de Russie, ne servira en rien à la révolution sociale mondiale. La coalition ouvrière- paysanne, désorganisée ces temps derniers, doit être reconstituée. Et ici, le Commissariat du ravitaillement a une tâche particulièrement importante: prendre l'initiative… Notre œuvre urgente, c'est d'écarter les dissentiments antérieurs avec les paysans, d'unir avec nous la campagne, d'abord notre campagne productrice de blé, puis celles de Sibérie et du Don… Mais comment y arriver? Quelles mesures prendre ici?... Fait à signaler: le jour où ces lignes sont écrites, pas un wagon de denrées de ravitaillement n'est arrivé à Moscou. Cela ne doit pas être. Evidemment la question se pose ainsi: ou bien il faut nous porter tous vers la campagne et arracher au paysan ses réserves, mais c'est la guerre ouverte et la provocation à la révolte de la garde blanche… ou bien mourir de faim…ou bien changer radicalement notre système d'approvisionnement. Dans quel sens? D'abord réviser les prix fermes pour le blé; ensuite, attirer à l'œuvre de l'approvisionnement en blé des hommes nouveaux sur des bases nouvelles…On peut remettre n'importe quel document à nos comités d'approvisionnement: avec du bon ils feront du mauvais; le contraire, jamais.*

Rappelons maintenant la marche réelle des événements: vaines expéditions de détachements armés dans les campagnes; tentative de rétablissement privé de l'industrie *du sac*: *échange des marchandises* tant vanté avec les efforts ridicules et maladroits pour donner à la campagne, sous une étiquette quasi-socialiste, les marchandises qu'elle demande à des conditions acceptables pour elle; relèvement vertigineux des prix fermes, dont la fixité était reconnue jusqu'alors comme la pierre angulaire de la politique de ravitaillement prolétarienne; et, pour couronner le tout, capitulation en paroles devant le *moyen propriétaire* sous la forme d'*explication* des comités de miséreux, dans la circulaire – aussi sage que pitoyable – de Lénine et de

Tsourioupa, commissaire du people au ravitaillement: “Qu’est-ce que les comités de miséreux des campagnes? ”, adressée à tous les Soviets de gouvernement et comités de ravitaillement. D’après les informations qui arrivent de province sur l’organisation des comités de miséreux, Lénine et Tsourioupa se plaignent naïvement:

> *Il est manifeste que très souvent les organisations des miséreux lèsent les intérêts des paysans de moyenne condition. En beaucoup d’endroits, la devise de l’organisation des miséreux est comprise, à tort comme si les miséreux devaient être dressés contre tout le reste de la population paysanne … Le pouvoir des Soviets n’a encore jamais lutté contre le moyen paysan. Le pouvoir des Soviets s’est toujours assigné pour but l’union du prolétariat urbain avec le prolétariat rural et avec le demi-prolétariat, et même avec les paysans-travailleurs de moyenne condition… L’union des ouvriers et des paysans a triomphé des propriétaires et de la bourgeoisie au mois d’octobre de l’année dernière. Cette union seule garantira la terre aux paysans, la manufacture et les usines aux mains des ouvriers et affermira le pouvoir des ouvriers et paysans. L’union des ouvriers et paysans conduira au triomphe final du socialisme.*

Je me représente le sourire cyniquement satisfait de lui-même du président du Soviet des Commissaires du peuple, signant *Oulianov-Lénine* le galimatias si clair pour lui des dernières phrases.

4. Le pouvoir d’Etat en Russie ne peut s’appuyer sur le prolétariat – tel est le sens politique de la circulaire “Que sont les comités ruraux de miséreux?” On ne l’avoue pas encore avec une pleine franchise, mais on le reconnaît nettement. La côté, ainsi, est franchie: après, c’est la descente. A la place de la dictature du prolétariat s’avance le règne du moujik. Pour un *politique réaliste*, la conclusion est claire: chercher un appui dans la classe paysanne, de quelque prix qu’il faille le payer.

En même temps, la vie impose un autre aveu, encore plus pénible: la bande de *demi-fanatiques, de demi-aventuriers*, qui mène les affaires dans la *Russie des Soviets*, depuis le coup d’Etat d’octobre, est incapable de réaliser le pouvoir d’Etat. Le temps passe et la vie s’obstine à ne pas se laisser organiser. Au contraire, plus cela va, plus c’est désespérant. Au centre, le désordre; dans les provinces, l’écroulement. Un changement radical s’impose, ou la catastrophe est inévitable.

Quand on a assumé une tâche importante et lourde, il est pénible de s’avouer à soi-même son incapacité. S’entraîner à rédiger des décrets, à y ajouter des commentaires enthousiastes, à apporter dans les meetings et dans la presse son énergie révolutionnaire et chanter des louanges mensongères à l’avant-coureur de l’affranchissement universel, – le prolétariat russe, – en fermant la bouche aux adversaires et en collant *au mur* les récalcitrants: ce sont là toutes tâches ordinaires et d’une technique peu compliquée. Mais de pareils talents ne sont pas suffisants pour administrer un peuple de cent millions d’hommes.

C’est affligeant, mais incontestable: il a fallu malgré soi s’en convaincre par expérience. On peut entasser des montagnes de décrets; les chancelleries de *départements* travaillent sans répit; c’est une fièvre d’organisation; comités, commissions, *glavki*, centres naissent, s’élargissent, se transforment, appellent des centaines de collaborateurs, dépensent l’argent du peuple au point qu’on ne suffit pas à imprimer des billets de banque; et – hors le bruit et le tapage, hors la ruine croissante du pays – aucun résultat. La situation est vraiment *pire que celle du temps des gouverneurs*, comme on disait jadis. Où donc est la racine du mal? A qui la faute? Et quelle issue?

On commence par le plus simple et le plus agréable: rejeter la responsabilité sur ses adversaires. Ils boycottent le pouvoir des Soviets ou font du sabotage; d’autres, sous un faux air de collaboration, sapent avec perfidie les meilleures initiatives des Soviets. Voici quelques tableaux qui caractérisent bien la façon dont ces idées se reflètent dans le cerveau des penseurs soviétistes.

1) Nationalisation des industries de pêche de la Caspienne (d’après les *Izvestia Narodnogo Komissariata po Prodovolstviu* No. 18 – 19):

> *La révolution d’octobre a mis fin au conflit séculaire du travail et du capital sur les bords de la Caspienne, cependant, les charrois d’automne ayant été négligés, les consommateurs russes sont privés de la plus grande partie du poisson attendu; et ce sont surtout les gros centres de consommation qui, au cours de cet hiver, ont le plus souffert de cette disette. L’entêtement des*

entrepreneurs de pêcheries à Astrakhan n'était pas encore définitivement brisé… Le capital privé n'a pas encore mis bas les armes. Convaincu de son impuissance à vaincre dans une lutte ouverte les masses travailleuses, il a recouru au procédé éprouvé de la calomnie et de la corruption, cherchant à pousser une partie des ouvriers contre l'autre, à soulever les ouvriers industriels contre les pêcheurs, à acheter les travailleurs de transports et à demander l'aide des cosaques du pays. La guerre civile à Astrakhan une fois apaisée, le Soviet de région a pris, en février, la résolution de nationaliser toutes les industries avec leur matériel et leurs capitaux. On a soustrait ainsi à l'industrie privée les petites entreprises artisanes éparpillées sur des dizaines de verstes $\{\text{versta} = 1.06km\}$. *La mise en pratique de la nationalisation – inventaire rigoureux des animaux et du matériel, large contrôle ouvrier – a fait défaut dès le début, une grande partie des petits patrons, occupés au dur travail des filets, au transport du poisson, à la tonnellerie, n'avaient aucune préparation économique ni technique. Ca et là les industries commencèrent à fonctionner normalement; mais la plus grande partie se morcela en petite groupes syndicalistes qui poursuivaient leurs petits intérêts au détriment de l'œuvre. Sous l'étiquette de la nationalisation, l'argent qui restait de l'organisation commerciale privée a été mis au pillage par chaque groupe en particulier, et les millions envoyés par le Comprod (Comité d'approvisionnement) ont été vite épuisés. Le travail salarié a été appliqué aux industries après la socialisation aussi bien qu'avant la révolution d'octobre; tous les anciens modes d'exploitation ont reparu, comme chez les anciens patrons. Pour les filets, les matériaux, l'alimentation même les nouvelles industries nationalisées ont pris aux ouvriers pêcheurs la paiement en nature: la moitié ou le tiers de la pêche. Le rôle principal était tenu ici par la petite et la moyenne bourgeoisie, qui s'est rapidement adaptée à la nouvelle situation, a organisé des artels de travail, des coopératives qui, a fond, ne cherchent qu'à réduire à néant les conquêtes de la Révolution; elles n'ont en pour cela qu'à partager les bénéfices des industries avec la partie des ouvriers qui était entrée dans l'artel. Toutefois, la bourgeoisie d'Astrakhan a su ici sauver les formes et acquérir du capital: la plus grande partie du dividende employée à soudoyer les ouvriers membres de l'artel, a été reportée sur le dos des consommateurs sous forme de hausse des prix sur le poisson. Certaines institutions officielles des autorités des Soviets ne sont pas demeurées en reste avec la bourgeoisie d'Astrakhan. Ainsi la section de la Banque Populaire, à Astrakhan, a socialisé de sa propre autorité quelques dizaines des meilleures industries de la pêche non sans créer autour d'elle une atmosphère de mercantilisme. Profitant de leur organisation financière, les manitous de la Banque Populaire se sentaient à l'abri de toute concurrence. L'organisation militaire locale s'est réservé quelques industries pour sa coopération: de nombreux Soviets de canton et de village où s'étaient solidement retranchés tous les anciens propriétaires d'industries, des dizaines d'artels, d'unions, de sociétés, de comités d'usines, poussés comme des champignons après la pluie créés sous la direction et avec la participation des anciens patrons… sont devenus rapidement des nids de spéculation avec toutes les funestes conséquences… La nationalisation des industries, réalisée par les capitalistes et leurs écornifleurs dans les conditions d'Astrakhan, devait entraîner forcément une masse de crises. La première – crise financière, n'a pas troublé particulièrement les financiers d'Astrakhan. Après avoir épuisé toutes les ressources liquides des banques locales, avalé à titre de prêts et de subventions les sommes liquides recues du pouvoir central, Astrakhan s'est pourvue ingénieusement d'une planche à imprimer du papier-monnaie l'opération a été fructueuse pour quelques-uns: par suite de l'absence de Bourse des valeurs dans la République des Soviets, les billets de banque d'Astrakhan ont prêté à un agiotage lucratif. Pour* *maintenir le cours de ces billets, on a eu recours à divers moyens de répression, jusques et y compris la confiscation des billets de la Banque d'Etat aux personnes qui les exportaient ou les rapprochaient des capitales au détriment des agioteurs…La difficulté suivante a été la vieille querelle entre les pêcheurs et les ouvriers d'industrie, que l'introduction du syndicalisme ou le seul changement de forme de l'exploitation patronale privée n'ont pu supprimer. Etant donné nos conditions de vie et le faible développement de la conscience prolétarienne des masses travailleuses d'Astrakhan, les contradictions nouvellement créés entre deux catégories de travail menacent d'arrêt et de ruine toute l'industrie…La nationalisation intronisée par en haut, sans préparation, a amené à l'état aigu ces contradictions. Il a fallu rejeter à la mer sur place une grande quantité du précieux poisson, par suite du refus de l'ouvrier de le recueillir… Le prolétariat d'Astrakhan et ses meilleurs éléments doivent lutter de toutes leurs forces pour leurs intérêts et l'épuration incessante de leurs organisations. Ils doivent se souvenir que l'arme favorite de la bourgeoisie dans sa lutte contre la classe ouvrière a été sa certitude que les ouvriers ne sauraient pas manier l'appareil économique.*

2) Socialisation de la terre dans le gouvernement de Kostroma (d'après un rapport sur la question agraire au cinquième congrès de gouvernement):

Beaucoup de domaines ont été mis sous séquestre uniquement sur le papier; en nombre de cas les domaines se trouvaient de nouveau dans les mains de poméchtchiks (gros propriétaires fonciers) qui les affermaient comme avant… Dans beaucoup d'endroits les koulaki avérés se sont retranchés dans les sections agraires et opposés par tous les moyens à la socialisation de la terre. Par exemple, dans le canton {volost} *d'Ilinskoé (district de Kostroma), par la faute des koulaki qui composaient la section agraire, aucun des douze domaines seigneuriaux placés sous séquestre n'a été partagé entre la population, et tous restent aux mains des poméchtchiks.*

3) Situation générale dans le gouvernement de Kostroma (d'après la chronique A travers la Russie dans le *Vestnik Narodnogo Komissariata po Torgovle i Promyshlennosti* No. 7 – 8):

Le gouvernement de Kostroma est une des régions de la Russie où le régime soviétiste s'est établi graduellement, sans violences tumultueuses, ni effusion de sang. Un tel état de choses aurait dû, semble-t-il, favoriser et hâter le travail d'organisation pour la mise en pratique des principes nouveaux de l'économie nationale. Néanmoins, le gouvernement de Kostroma, loin de dépasser dans cette œuvre d'autres régions qui se trouvent dans des conditions plus difficiles est resté longtemps en arrière. L'exemple du gouvernement de Kostroma prouve avec quelle peine on arrive, dans la vole du développement prudent et susceptible d'éviter les conflits à retrouver la force et la profondeur de pression révolutionnaire qui étaient nécessaires pour le coup d'Etat d'octobre. Avant l'été, les campagnes dans ce gouvernement étaient encore le rempart de l'ésérisme (de S.R., social-révolutionnaire) sous le drapeau duquel réussissaient à se dissimuler la bourgeoisie rurale et même les poméchtchiki; quant à la ville, elle était jusqu'à ces derniers temps sous l'influence du menchévisme {a faction among Russian Social-democrats (1903), opponents of Lenin's demand for establishing a party of *the new type*. After October 1917 went underground} *qui cherchait à entraver le travail soviétiste.*

Une explication si simpliste de tous les insuccès par l'opposition criminelle d'agitateurs malintentionnés ne peut tout de même rassurer longtemps l'esprit d'un bolchevik lui-même. La vie, de plus, met en face de faits dont même la conscience bien trempée des Soviets ne peut pas se dégager si simplement. *Vu l'absence de blé*, communique-t-on au commissariat du peuple du ravitaillement de Riazan, la population riveraine de l'Oka et celle qui est voisine des lignes de chemins de fer, vote une motion sur la saisie des bateaux et trains d'approvisionnements.

Il y a eu des cas où, sur la base de leurs propres décisions, les habitants ont pillé les bateaux qui transportaient du blé. Les mesures, même rigoureuses, prises par le Soviet ne donnent pas de résultats, car le pillage a lieu non par mauvaise intention, mais *par suite de la faim*, dit le rapport avec une naïveté inimitable. *Le Commissaire du Peuple à l'intérieur*, lisons-nous dans la chronique En Province du numéro 16 – 17 du *Izvestia Narodnogo Komissariata po Prodovolstviu*, était encombré de demandes, par les provinces, d'envoi d'argent pour l'entretien des hôpitaux et des asiles. Les institutions locales, faute de versement d'impôts par la population, manquent de fonds. En conséquence, dans beaucoup de gouvernements, on meurt de faim dans les hôpitaux et les asiles. Selon les renseignements du *Commissariat de l'intérieur* du gouvernement de Tambov, un grand nombre de malades dans les établissements hospitaliers et dans les asiles d'enfants, est mort de faim. "Ces faits, ces victimes sont pour nous, pour la révolution, une honte et un crime", conclut l'auteur de cette chronique, Alexeev.

Pour se reconnaître directement coupable de honte et de crime, il faut un courage qu'on ne saurait demander chez tous. La première étape sur cette voie douloureuse est la reconnaissance ouverte d'un lien étroit entre le scandaleux désordre qui règne dans la République des Soviets et la structure bizarre du pouvoir des Soviets.

> *Dans beaucoup de provinces*, – geind Tsourioupa au cinquième Congrès des Soviets, – *la devise <u>Tout le pouvoir aux Soviets</u> a été comprise comme si toute l'autorité devait appartenir aux Soviets locaux, sans s'occuper de distinguer lesquels: de district, de gouvernement, de province ou de toute la Russie. Chaque Soviet se considère en droit d'agir indépendamment des directives du centre, de faire* {two words undecipherable} *sa politique à lui. Conséquemment, nous n'avons pas reçu de renseignements sur les chargement et les expéditions d'ou le retard des chargements sur les lieux; on ne nou informait pas de l'exécution de nos ordres d'expédition, etc. Les organisations locales ont proclamé le séparatisme également dans l'administration des transports. Les réquisitions opérées par les Soviets locaux ont joué un rôle particulièrement nuisible pour l'exécution des plans d'approvisionnement. Cette mainmise pure et simple sur les chargements de blé a désorganisé le ravitaillement.*

Voici une petite illustration de la *désorganisation* dans le département de M. Tsourioupa, d'après le rapport de l'agent envoyé dans le gouvernement de Tambov:

> *Dans le gouvernement de Tambov, le ravitaillement empire d'une manière rapide, menaçante et sombre…Ni le collège de ravitaillement du gouvernement à Tambov, ni celui du district à Kozlov ne sont capables de remplir leur tâche. Les gens qui sont à la tête des organisations ne se rendent aucun compte de ce qui doit constituer leur besogne…Celle-ci se réduit en fait à recevoir les visiteurs. Quant à un plan de travail quelconque, inutile d'en parler. Le collège de gouvernement ne sait ce que fait le district; le district ignore ce que fait le canton. Aucun recensement statistique détaillé. Le Soviet des députés du district de Kirsanov a confisqué, conformément à l'accord conclu avec le comité régional, la réserve de pommes de terre. Il demande, en sus, au profit du sovdep trois roubles par poud [16kil.] comme droit d'envoi. Le sovdep de canton de Injavin a réquisitionné pour ses propres besoins, sur la réserve de 2,000 pouds de pommes de terre du comité supérieur de ravitaillement, 700 pouds à un prix inférieur… le commissaire au ravitaillement de Kozlov a interdit à tous les chefs de station de charger des pommes de terre pour l'approvisionnement du collège de ravitaillement du gouvernement exigeant chaque fois une autorisation spéciale donnée par lui.*

Comme on le voit, il y a de quoi désespérer. Cependant le gouvernement de Tambov n'est pas une exception. C'est partout le même tableau.

Tout ce qui a été dit nous dépeint la plus complète incapacité des organes locaux de ravitaillement, leur totale et enfantine inaptitude à s'adapter et leur ignorance des problèmes élémentaires qui rentrent dans la sphère de leur activité.

Tel est l'aveu de l'auteur de la chronique En province dans *Izvestia Narodnogo Komissariata po Prodovolstviu* No. 14 – 15, après un voyage à travers toute la Russie soviétiste.

Il va de soi qu'un pays qui vivait hier encore sous le joug de l'autocratie ne peut se transformer en un clin d'œil…Mais, sans oublier cet axiome, on ne peut tout de même pas ne pas sans oublier cet axiome, on ne peut tout de même pas ne pas éprouver un sentiment pénible au spectacle de la vie présente de notre province…Le pouvoir des Soviets…a posé devant le pays des problèmes précis et des buts qui sont des problèmes historiques pour le prolétariat. Si ces buts se réalisent avec lenteur et timidité dans un pays, hier encore esclave, c'est tout naturel; il n'y a pas lieu d'en éprouver un sentiment particulièrement pénible. Mais si avec le temps il apparaît clairement que ce travail, non seulement n'est pas fondé organiquement sur la vie, mais que, comme dans les premiers moments de la République soviétiste, il suscite une hostilité nuisible, tantôt ouverte, tantôt cachée, l'impression pénible a toute sa raison d'être. La République soviétiste pendant toute son existence est rongée sans cesse et traîtreusement…par un ennemi caché et secret…qui se fait sentir partout: c'est cette incapacité d'organisation, cette étroitesse de vues dont sont, sauf de rares exceptions, imbue les organes locaux, en particulier ceux de ravitaillement, dans la République soviétiste.

Qui se ressemble s'assemble. Les sovdeps locaux, les goubprodkoms {goubernia = province; prodovolstvie = foodstuffs; com = committee}, etc, etc…, sont faits à l'image des organes soviétistes centraux. Mêmes hommes, mêmes chansons. Ce n'est pas dans les mécomptes des organisations qu'est la racine du mal, mais dans les travailleurs qui feront mal leur partie, où qu'on les place. Chez les bolcheviks – les hommes à la fois de doctrine et de travail sont des unités – il y en a trop peu pour pourvoir les fauteuils de commissaires. Tout le reste – au centre et en province – sert la République des Soviets par peur, par cupidité et non par conscience, ne s'intéressant aucunement à sa tâche, – mettons les choses au mieux: faisant des heures et recevant un traitement convenable – le plus souvent inconvenant. Voici le tableau brossé, par un homme compétent – puis qu'il y participe – des travaux du pouvoir des Soviets et des résultats obtenus (d'après l'article de P. Fedorov, Pensées Bourgeoises, *Izvestia Narodnogo Komissariata po Prodovolstviu* NNo. 14 – 15, 16 – 17):

Le lecteur sait dans quelle situation se trouve l'industrie de la République des Soviets. Chaque jour on nous annonce, non l'ouverture de nouveaux établissements, non l'exploitation de nouvelles carrières, mines, forêts, tourbières ou terres cultivées, mais la réduction ou l'arrêt du travail des anciennes exploitations. Les fabriques et les usines s'arrêtent faute de combustible, les chemins de fer sont réduits, on diminue la distribution du courant destiné à l'éclairage et au mouvement des tramways. Nous attendons l'hiver avec effroi: comment chaufferons-nous nos habitations? Il n'y a pas de bois, quoique on aperçoive des forêts de toutes les banlieues de nos villes. Mais il ne s'agit pas seulement de combustible. Il n'y a pas de matières premières pour alimenter notre industrie. Les Sartes [2] *ensemencent leurs plantations de coton avec du blé et réduisent les textiles de Kinechma, Chouia, Moscou, Voznesensk à une part de famine … Des régions fertiles en lin, parviennent des nouvelles annonçant la réduction des ensemencements, car les vieux stocks pourrissent. En même temps on réduit la production des* *filatures de lin et des tissages … Les hauts-fourneaux qui travaillaient nos minerais grands-russiens s'arrêtent. Mais non seulement nous n'avons dans le domaine de notre production aucun chiffre indiquant une prompte guérison; c'est la même chose également dans le domaine de la répartition … Avons- nous recensé, ne fût-ce qu'un produit, comme il convient? Non … Quel que soit le produit qui tombe sous nos yeux parmi ceux qui sont décomptés officiellement, nous osons dire qu'il n'en passe par les mains des préposés au recensement que 1/10 ou 1/100. Et puis toute la quantité recensée de tel ou tel produit passe-t-elle par toutes les étapes marquées de distribution ? Non. Toutes nos expéditions régulières s'exécutent plutôt mal, comme on sait. La pratique de la répartition hors série, de ce qu'on appelle expéditions hors tour avec lesquelles nous avons l'habituée de boucher les trous, lieurt sur une large échelle. Le système des cartes existe, boiteux et chancelant, là où on ne peut presque rien donner par coupon; là où il y a des produits, la carte n'est considérée que comme du papier … Nous ajouterons encore à ce tableau, qui n'a rien de gai, un seul trait connu de tous: le grand amour du pot-de-vin et l'arbitraire sans limites. Cependant, de quoi s'agit-il? Avant tout, partout, quelle que soit la branche de travail économique gouvernemental que nous prenions, une chose nous frappe: beaucoup de gens et pas d'hommes …Combien y en a-t-il chez nous maintenant qui écrivent et transcrivent! Quelles nuées d'entrées et de sorties quotidiennes s'accumulent sans aucune utilité!.. Des Akaki Akakievitch* {Gogol's hero, a small painstaking clerk hardly perceiving anything beyond his desk} *ont retourné leur veste et se sont mis avec le zèle habituel à un nouveau travail, un gribouillage grossier, stupide, doublement nuisible pour l'Etat ouvrier …Les fabriques sont vides et des hôtels à nombreux étages sont bondés de ronds de cuir. C'est une honte! Si la République envoie dans l'autre monde des filous de petite envergure, elle ne peut tolérer les administrateurs inintelligents derrière lesquels, comme en serre chaude, se multiplient les fonctionnaires … Nos commissaires voient peu la vie, séparés qu'ils sont d'elle par des montagnes de paperasses, mais nos dirigeants, administrateurs, secrétaires, les hommes qui, en fait, sont à la tête de toutes les branches du gouvernement, ne voient pas du tout la réalité. Aussi l'ignorance fonctionnariste et la stupidité sont déjà devenues proverbiales … Nous observons avec effroi l'écroulement le plus complet de l'industrie, dont la première cause est notre inexpérience en matière d'organisation et notre optimisme béat. La condition nécessaire pour redonner de la vie à notre industrie de fabrication est que toutes les directions, centres, comités et soviets (sous leur forme présente) renoncent à intervenir directement dans la production … Aucun ingénieur sachant le prix de ses connaissances, aucun ingénieur expérimenté ne viendra chez nous maintenant, car toujours et partout une bande d'ouvriers ignorants, pénétrée de la compréhension étroite de ses intérêts, liera l'initiative d'un tel homme … Nous avons à notre disposition si peu d'hommes de doctrine, adonnés au travail et instruits, que leur nombre n'est même pas suffisant pour occuper les sièges de commissaires. Qui donnera la vie aux beaux décrets de la République? Naturellement, ce sont des hommes d'un autre genre, inconstants au point de vue politique, qui, souvent même, nous sont hostiles, ne reconnaissant aucun idéal, travaillant seulement pour le salaire, parfois pour la carrière. Travaillez parce que votre intérêt l'exige, voila ce qu'il faut dire à toute la phalange des spécialistes saboteurs, persifleurs, qui créent à chaque pas des difficultés. C'est précisément par là que la bourgeoisie a pris ces messieurs: elle les a achetés. Nous aussi nous devons les acheter: nous pouvons les payer plus que ne les a payés la clique avide des exploiteurs.*

La recette du bolchevik qui a appris à penser en bourgeois est extrêmement simple: il n'y a qu'à acheter les spécialistes nécessaires. Mais qu'arrivera-t-il si l'on n'achète pas avec de l'argent? Si le prix de la collaboration n'est pas le rouble, ni même la ration alimentaire, mais le pouvoir – le renoncement à la dictature du prolétariat, au despotisme des sovdeps – le rétablissement des principes de la démocratie et la capitulation sans conditions devant la bourgeoisie paysanne, si méprisée il y a peu de temps?

5. L'abandon du programme donné comme une incarnation des intérêts de classe du prolétariat et l'aveu d'un retour nécessaire de la dictature prolétarienne des sovdeps à la bourgeoise collaboration des classe sur la base d'une organisation politique plus ou moins démocratique, – telles sont les conclusions auxquelles, en octobre 1918, *l'expérience sociale* amène les expérimentateurs qui commencent à réfléchir. En même temps le bolchevisme, en tant qu'idée politique, disparaît de la scène en Russie. Il subsiste encore ça et là, comme état d'esprit; une bande d'ambitieux s'accroche aux épaves de l'appareil administratif et policier qui leur est tombé fâcheusement entre les mains. Les décombres fument, mais l'incendie touche déjà à sa fin. Naturellement, de nouvelles complications sont possibles. Les chefs qui conduisent consciemment le bolchevisme à une capitulation peuvent par exemple avoir les reins brisés avant de se rendre. Mais si même les phraseurs casse-cous l'emportent de nouveau sur les gens qui font de la politique réaliste, si les Peters, les Trotski, les Zinoviev renversent et piétinent les Lénine, et les Krasine, cela n'infusera pas un sang nouveau au bolchevisme en décomposition.

Pour la Russie, ce serait là, à proprement parler, la meilleure solution. Ce crochet ne retardera pas longtemps la liquidation finale. Régénérer le pays, tous ces expérimentateurs-alchimistes qui se sont en partie brûlé les mains, en partie rempli les poches, en sont absolument incapables: le sang versé et les vies détruites les séparent de la future Russie libre.

A. Tchouprov, Professeur d'Economie politique à l'Université de Moscou
Stockholm, Février 1919

Notes

1. A.N. Orlov, a social-democrat of old, an active worker in the cooperative movement and an economist. In 1918 he edited the periodical *Izvestia Narodnogo Komissariata po Prodovolstviu*. Wrote an *excellent* (Lenin) book *Продовольственная работа советской власти* (Foodstuffs: the Work of the Soviet Power). From summer 1921, head of the economic section of the journal *Novy Mir* put out by the Berlin embassy. At the same time, however, he indicated, in his secret diary, that he wished to "expose" the Bolsheviks who had ravaged a great county, to accuse them of "all their meanness, cheating, servility", of the "destruction of our generation, of violating everything we had believed in". N.N. Krestinsky, the Ambassador, in his report sent to Moscow, wrote that in 1923 Orlov had "not only ideologically, but also formally left the Russian Communist Party", refused to return to the Soviet Union and was therefore fired. Upon settling down near Berlin, Orlov had been working on a fictional novel *Dictator* (Dmitriev & Semenov, in a note to their translation of the pamphlet, with an appropriate reference).

2. {The traditionally settled branch of the Uzbek people.}

8a. A.L. Dmitriev, A.A. Semenov. Chuprov and the Bolshevik Revolution (a Fragment)

Published as Foreword to their Russian translation of [8] in *Voprosy Istorii*, No. 10, 2003, pp. 3 – 18.
{This fragment is from the authors' foreword to their translation of Chuprov's *Décomposition du Bolchevisme* (reprinted in this collection).}

The pamphlet endeavors to analyze the changes in the Soviet policy during 1917 – 1919. Drawing on a comparatively few and uncoordinated sources, Chuprov was able to provide a surprisingly accurate characteristic of the economic situation in the first post-revolutionary years. He described the mixed background of the Bolshevik economic policy and its zigzags, indicated the extent of red tape and made an interesting attempt to accustom the reader with the psychology of the representatives of the new power. Chuprov concluded that the strategy of an accelerated construction of socialism had failed and provided glaring examples of its inefficiency. Like many contemporaneous analytics, he expected that the slogan *Dictatorship of the Proletariat* will soon be rejected and that such an ideological volte-face would necessarily mean that the élite of the governing party will aim at a gradual turn of the conducted policy into the channel of normal capitalist development.

At the same time, an important moment that does him credit for his insight, and to some extent offers a clue for the understanding of the possibility of an essentially different development of the events in future, is traced in Chuprov's work. When revealing the total regularization of the economic life in Soviet Russia, the author established that it was paradoxically combined with an absolute lack of inertia in the approach to economic issues. This was manifested by a considerable part of the leaders and ideologists of economics and made it possible to transform the chosen course. He even admiringly shades the intellectual brilliance and uncommon straightforwardness with which eminent Bolsheviks revealed the defects of the conducted policy. It is indicative that the opinions of N.A. Orlov, who soon broke off with the Bolsheviks, had attracted Chuprov's attention.

The cynicism of Russian revolutionaries indicated by Chuprov was specific. The aims of the intellectual layer of the Soviet leadership could not have been confined to a vulgar pursuit of power and money. The mania for realizing the Communist Utopia dominated these people to such an extent that they had been prepared to destroy ruthlessly not only the previous society but also their own constructions if only they had occurred inefficient. The high adaptability of the Russian communism, of which there were no precedents, was indeed caused by this fact.

In the 1920s, being a consistent opponent of the Soviet regime and, unlike the *Smenovekhovtsy* {those émigrés who *changed their poles*; in land surveying, poles are set up at intersected points for sighting at them from other stations. They believed that the Soviet New Economic Policy heralded the restoration of capitalism} perfectly well realizing the brutal essence of the Bolshevik dictatorship, Chuprov apparently had been ever more thoroughly imbued with the idea that the latter's speedy decomposition all by itself was not to be expected. It followed that the decision of the émigré circles to boycott totally all the Soviet structures and institutions was wrong. This most important conclusion was considerably ahead of its time and apparently predetermined the tragedy of the scholar's last years.

8b. L.B. Sheynin. Chuprov: Socialism, the Market and Leninist Russia

Foreword

{The following is a translation of my brother's unpublished comment on Chuprov's pamphlet. The text is not sufficiently documented which is partly justified because, as its author explained to me, the material concerning the general situation in Russia was more or less discussed by previous Russian contributors.
Was Lenin indifferent to the fate of Russia (cf. end of this comment)? I think that he, as well as later rulers of Russia and the Soviet Union, really cared for their nation *provided* that the Communist party (and they personally, as especially appropriate to note concerning Stalin) remained in power and that Russia's national interests were understood as being identical with those of communism. The Chernobyl catastrophe furnishes a horrible recent example. In order to avoid possible public disorder directed against the Communist leaders, the population of the affected regions (even pregnant women and children) were not evacuated; on the contrary, open-air festivities were held in Kiev at the most hazardous time.}

*   *   *

Chuprov concluded that Bolshevism was being "decomposed" by issuing from the Russian experience stored during approximately the first year of Bolshevik rule. Unlike many other authors, who concentrated their attention on social phenomena, Chuprov was also interested in the ideological and political motives that guided the Bolsheviks after coming to power. Perceiving hardly any considered planning and a lot of improvisation in their actions, he (§5) called them *expérimentateurs-alchimistes*. He also distinguished between those who followed a *politique réaliste* (Lenin, Krasin) and *phraseurs casse-cous* (Peters, Trotsky, Zinoviev). The former, as Chuprov thought, were prepared to *capitulate* since they understood that the Bolshevik methods of managing the national economy did not work and that it was necessary to return to the old and proven instruments that we would now call marketing.

We see now that Chuprov's conclusion was ahead of its time; Russia's return to the market system only occurred in the spring of 1921 when the New Economic system (NEP) was introduced. However, during 1919 and 1920 the Soviet state had been carrying out ever more sweeping attacks upon market relations. Thus, the sphere in which money payments were possible had been steadily reduced. And, just before NEP was proclaimed, a decree had been passed demanding the establishment of Sowing Committees in the rural districts. These were to *force* the peasants to plow and to sow their fields without really bothering to discriminate between using their own or their neighbors' implements or seeds.

Lenin's attitude towards the officially prohibited but widely spread private trade is instructive. In Moscow, this trade was concentrated on the Sukharevskaia Square and Lenin repeatedly said and wrote that the *Russian Sukharevka* was the main enemy of the new social order. It is not surprising therefore that the decision about establishing NEP was taken not so much because of the economic dislocation (an economic motive) as under the pressure of peasant revolts and redoubtable military mutinies excited by the same peasants dressed in Red Army overcoats (a political necessity). It is thus wrong to state that the leaders of Russia *abandoned* (§5) Bolshevism, that during the end of 1918 and the beginning of 1919 it was experiencing a crisis or was being decomposed.

The issue should not however be reduced to specifying a date; it is more profound. Chuprov wrote about Bolshevism, including Bolshevism in economics, as about a well-known phenomenon. For his contemporaries, who were seeing the facts of economic life of those times, this term hardly gave rise to questions. Nevertheless, Bolshevism, including its application in economics, was not stationary. What substance should be inserted into that notion? This is an important question, above all for us because we are using this term along with *capitalism, socialism, market, plan*, often even without defining them beforehand. This issue is complicated because the Marxist doctrine of socialism, officially professed by the Bolsheviks (and not only by them) *lacks* any, whether rigorous or not, description of the future socialist order.

It is usually thought that Marx and Engels (and Lenin) were *state socialists* believing in the primacy of the state over the individual. Socialism, as they understood it, was the management of all the means of production at the disposal of the society by the state. Quite a number of German Social Democrats regarded as the *best successors* of Marx and Engels conceived socialism as a system under which the government will also manage *labor battalions* and send them wherever, and for fulfilling whatever task it deems necessary for the society. However, Marx himself abstained from any definite pronouncements on this subject. Moreover, a little known manifesto that he wrote in 1864 on the occasion of the establishment in London of the *Working Men's International Association* (the International) included a phrase [1] about *cooperative factories* on a national scale to be founded with the assistance of the state of the victorious proletariat. In his correspondence with Vera Zasulich Marx assumed that, in Russia, land communities of the peasants might become the basis of socialism. It follows that the founder himself of the Great Doctrine was not altogether certain to which trend of socialism will the future belong.

Proudhon in France; Lassale in Germany; Owen in England; and Kropotkin (living in emigration) advocated the idea of *cooperative factories*. Attempts were repeatedly made to implement this idea; in Russia a factory of metallic articles in Pavlovo, Nizhegorodskaia province, can for example be mentioned [2]. Enterprises belonging to their workers are nowadays called *Narodnye* (People's); in Moscow, such an enterprise was recently established on the basis of Dr. Fedorov's eye clinic. In a "refined" way the same idea was implemented in Russia in the beginning of the 1990s, apparently on the initiative of G.Kh. Popov, the then Moscow mayor, when a part of state-owned enterprises was being de-nationalized. Companies whose shares belonged to their workers were being established but the shares are known to have been mostly bought up by the managements; managers had indeed become owners of those companies.

It is not by chance that Molotov, for a number of years the head of the Soviet government, repeatedly stated, after his career had ended, that only a few understood the essence of socialism and that it was still necessary to find out exactly how did various figures interpret its meaning (Chuev 1991).

But to return to Lenin. In 1918 and before that he mainly came to perceive socialism under the influence of the national economies of the belligerent powers; this follows from his well-known work of 1917 *The State and the Revolution* and from a number of his related articles. The idea of a war-time economy as introduced in various countries consisted in reducing private consumption as much as possible and in directing the thus saved resources to cover military necessities. Theoretically, this could have been achieved by an adequate tax system, but not by taxes which existed in time of peace. The countries at war had therefore introduced ration cards and made workers liable to call-up. Military contracts were being allotted to enterprises and military- industrial committees were, to the exclusion of the market, charged with the distribution of raw materials, fuel, etc among productive entities. States also began to take control of the banking systems; the significance of prices (and, along with them, of money) lowered and natural economy became more important. Germany was especially successful in all that.

In militarized Germany Lenin indeed perceived a specimen of the future socialism. For it to appear there remained, as he thought, hardly anything more; it was only necessary to remove the ruling clique and to replace it by lovers of the people, – by those like Lenin himself and his friends. And they will be able to attract ordinary people for managing large-scale manufacture who will certainly cope with this task because it is largely reduced to simple rules of stock-taking and control. It should be thought that money, taxes and the market itself under such a system, if they persist, must play a secondary role. And such a system was indeed implemented in Russia in 1918, the year described by Chuprov, which was made easier by the fact that many features of Russia's nationalized economy had already existed during previous war-time years.

We should now turn our attention to the (often overlooked by the economists) issue of *the influence of finance on national economy*. Under "normal" conditions this influence is barely noticeable, but if, and as soon as the state fails to collect the necessary taxes, the picture abruptly changes. For satisfying its needs, the state begins to issue paper money and the prices inevitably rise. This is not really dreadful if large-scale enterprises dominate the manufacturing of commodities since natural economy might then become more important than before. This is what happened in belligerent Germany where natural economy had been affecting even the peasants who were obliged to sell definite farming produce to the state at fixed prices. In Russia, however, the market of farming produce mainly persisted to the end of 1916 and the nation was being still more vulnerable owing to its weak tax system.

Like other countries at war, Russia was issuing paper money and both prices (especially those of foodstuffs) and wages were rising. Attempts at fixing the prices of some of the farming produce bought from peasants and landlords were being made at the end of 1916 onwards but they never succeeded. Fixed prices caused the countryside to hold back thus dooming the urban population to half-starved existence. It became necessary either to collect taxes in kind or to stop issuing paper money. Neither the Czarist, nor the Provisional, nor, finally, the Leninist government was able to achieve the latter aim and had to keep to the former policy.

In 1918 the commercial societies were ordered to sell industrial commodities only to those peasants who sold their grain at fixed prices. This, however, hardly helped to victual city and army and the government mainly resorted to irregular requisitions. In the beginning of 1919 these were arranged as the so-called *Prodrazverstka* (a surplus appropriation system). And only in 1921 a foodstuffs tax (the core of NEP) was established; later on, when money became stronger, it was replaced by a money tax.

If political issues are left aside, the economic crisis of 1918 – 1921 in Russia, and above all the shortage of foodstuffs, can to a large extent be attributed to the unsound tax system that had indeed led to the *clandestine working of the printing press*.

However, contemporaries had hardly perceived this dependence and Chuprov was no exception; he (§3) remarked in passing that the economic situation had led to a rise *vertigineux* in fixed prices [of foodstuffs]at the farm level. This measure was however *superficial*. The issue of paper money had not been restricted and the new fixed prices had to be adjusted once more. The peasants, expecting further depreciation of money, were demanding prices for their produce over and above the current level; sometimes they could have even refused to accept depreciated money. And the shortage of foodstuffs persisted. Such was, roughly speaking, the situation in Russia during 1918 – 1920. Thus, the weakness of the state financial institutions can damage essentially the national economy.

It is not amiss to note that a few years later Russia (or, rather, the Soviet Union) had to experience a similar ordeal […]

The laws of the market are the same both under *capitalism* and *socialism*. When the situation becomes extraordinary, the state, in either case, restricts or abolishes the market. But the peculiarities of such measures depend not only on the dominating ideology, but, in addition, on objective circumstances such as the intellectual level of management of the state affairs and the stage of development of the financial institutions, – above all, of the tax system. As to Chuprov, he, like many others, apparently believed that for Bolshevism the market was altogether contra-indicated. However, later events proved that Bolshevism in politics should have been distinguished from its variety in economics. During the years of NEP the economy became market-oriented, but the political structure remained invariable.

Chuprov's pamphlet is unusual in that he attempted to reveal the "economic philosophy" of Russia's contemporaneous rulers. He explained that this philosophy was largely chaotic and that, as far as practical politics was concerned, it led to improvisations replacing one another (which seems to be correct). Chuprov had not however raised another interesting issue: How did the abolition of the market, of money and wage-rates influence the working effort of the millions of producers [3]? But the fate of the nation indeed depended on this circumstance.

Above, I described how the peasants reacted on the abolition of the market in 1918; understandably, the workers' response was no better. When payment is irrespective of the productivity of labor, the incentive to work lowers [4] and this is indeed what happened. Thus, Lenin (and others) is (are) known to have compared the productivity before and after the abolition of the market, and that the results of the calculations were always in favor of former times. Chuprov (§4, Item 1) adduced facts concerning the formally nationalized Astrakhan fisheries; they indirectly confirm that producers had indeed attempted to keep to market relations.

It seems that Chuprov (§2) unjustly thought that Lenin was absolutely indifferent to the fate of Russia. Had that be true, Lenin would not have planned to *electrify* Russia, not have attempted to attract concessionary capital, or to reorganize the *Rabkrin* (the Worker-Peasant Inspection Agency), – that is, to improve the work of government institutions. True, all these events happened later and in 1919 Chuprov could not have known about them. He undoubtedly became acquainted with all this afterwards, and he was opposed to the boycott on, and favored cooperation with Soviet Russia. His pamphlet is not only a valuable still of the contemporaneous life in Russia, but also a penetrating pre-indication of the forthcoming retreat of Bolshevism as understood by Russian economists.

Notes

1. As Franz Mehring described it, see his book on Marx. {The author refers to its Russian translation (Moscow, 1934, p. 274).}

2. Tugan-Baranovsky (1916) described the mainly unsuccessful experience of such factories and agricultural communities the world over.

3. At the end of 1920 workers began to receive free rations. Such were the times…

4. As Molotov (Chuev 1991) indicated, the main point of socialism was "fulfilment of production quotas". This is hardly in line with him being formerly the head of government who never attempted to abolish money, wages, or piece-rate payment.

References

Chuev, F. (1991), *Сто сорок бесед с Молотовым* (A Hundred and Forty Conversations with Molotov). M.
Tugan-Baranovsky, M.I. (1916), *Социальные основы кооперации* (Social Basis of Cooperation).

9. On the Expectation of the Ratio of Two Mutually Dependent Random Variables

Published as Foreword to their Russian translation of [8] in *Voprosy Istorii*, No. 10, 2003, pp. 3 – 18 .

1.1. A number of the main problems of the theory of statistics reduce to the determination of the expectation of the ratio of two variables. The inherent formal mathematical difficulties constitute an essential obstacle to a definitive justification of the foundation of science: the investigator, groping for a full rationality of his applied methods, is compelled to take them into account in the theory of dispersion, correlation theory, etc. There is naturally no lack for attempts to surmount this obstacle but most of them follow one and the same path promising only partial success, success not under any {reasonable} conditions. Other approaches to that goal are however also conceivable. Accordingly, it is urgently important to formulate the problem in a general way, and, after considering the possible methods of its solutions, to outline the conditions for one or another of them to be preferable.

1.2. The terminology is not sufficiently established and I have to begin with a precise definition of the required concepts. I call a variable u taking k differing values $u_1, u_2, \ldots, u_k$ with probabilities $p_1, p_2, \ldots, p_k$ a *random variable of the k-th order*. The system of magnitudes u_i and p_i I call the *law of distribution* of the random variable u. This law can be given not directly, as a number of the abovementioned magnitudes; it might be determined by a number of parameters linked with them and allowing to derive them uniquely. Thus, denoting the expectation of the h-th power of u by m_h, it is not difficult to show that for an unique derivation of the law of distribution of u it is sufficient to have the values of $m_1, m_2, \ldots, m_{2k-1}$.

Similarly, denoting the expectation by E and supposing that

$$\mu_h = \text{E}(u - m_1)^h,$$

we see that, for such a derivation it is sufficient to have $m_1, \mu_2, \mu_3, \ldots, \mu_{2k-1}$; nevertheless, a series of magnitudes μ, however many of them are provided, cannot by itself fully determine that law.

For the case of two variables, u, taking values $u_1, u_2, \ldots, u_k$ with probabilities $p_1, p_2, \ldots, p_k$ and w with values $w_1, w_2, \ldots, w_l$ and probabilities $\pi_1, \pi_2, \ldots, \pi_l$, we designate the probability that they simultaneously take values u_i and w_j respectively by p_{ij}. I call the totality of magnitudes $u_1, u_2, \ldots, u_k$; $w_1, w_2, \ldots, w_l$; and $p_{11}, p_{12}, \ldots, p_{kl}$ the *law of connection* of the variables u and w. As in the case of one variable, the law of connection can also be given not by the magnitudes above, but as some number of parameters connected with, and enabling to establish them. Thus, denoting the expectation of the product $u^a w^b$ by m_{ab} it is possible to show that the law of connection is fully determined by the series $m_{01}, m_{02}, \ldots,$

m_{0k}, m_{10}, m_{11}, …, m_{1k}, m_{20}, …, $m_{l-1\,j}$, m_{l0}, …, $m_{l\,k-1}$. It is also fully determined by the system of m_{10}, m_{01} and $(k\,l + k + l - 3)$ magnitudes μ_{20}, μ_{11}, μ_{02}, … where $\mu_{ab} = E\,[(u - m_{10})^a\,(w - m_{01})^b]$. Nevertheless, a series of μ_{ab}, however many of them are given, cannot all by itself fully determine the law of connection.

It is of course conceivable that there exist an indefinite number of systems of such parameters. From among those proposed in our literature[1] we ought to indicate, in addition to those mentioned above, the system of magnitudes

$$r_{ab} = \frac{\mu_{ab}}{\mu_{20}^{(1/2)a}\,\mu_{02}^{(1/2)b}}\,. \qquad (1)$$

The first in that series of coefficients, namely r_{11}, is called the correlation coefficient; in particular, in case of the so-called *normal* law of connection its value determines all the other coefficients (1) and, therefore, together with the values of m_{10} and m_{01}, it also determines the very law of connection.

If the law of distribution of one of the variables fully persists whichever possible value the other variable takes, the variables are called *mutually independent*.[2] A necessary and sufficient condition for the mutual independence of variables u and w is that $p_{ij} = p_i\,\pi_j$ for any possible values of i and j, or $m_{ab} = m_{a0}\,m_{0\beta}$ for the values of a and b indicated above.

1.3. The problem about the expectation of the ratio of random variables u and w assumes that it is in essence solvable. It is clear however that $E\,(u/w)$ takes a definite value not under any conditions. If variable w can take the value 0 without u vanishing at the same time, then ∞ is included in the series of the possible values of u/w. If the probability of that value is not infinitely low, then also $E\,(u/w) = \infty$ and the formulated problem becomes senseless.

Let us assume, on the other hand, that variable w can take value 0, but that the variable u also vanishes then. In this case the series of possible values of u/w includes indefinite values of the type 0/0. For the calculation of $E\,(u/w)$ to have sense, an additional condition is necessary concerning the value of u/w when both the numerator and the denominator of this fraction vanish. Bohlmann (1913, pp. 386 – 387) suggests that, when calculating the expectations, the indefinite values of the variables should be completely neglected by equating them to zero. Suppose that variable x takes values x_1, x_2, …, x_{s-1}, 0/0 with probabilities p_1, p_2, …, p_{s-1}, p_s whose sum is unity. Denote

$$\sum_{i=1}^{s-1} p_i x_i \,/ \sum_{i=1}^{s-1} p_i = E_{o/o}\,x \qquad (2)$$

and agree that, in case of indefiniteness, the variable takes value (2). Then

$$E\,x = \sum_{i=1}^{s-1} p_i x_i + p_s\,E_{o/o}\,x = E_{o/o}\,x\,[\sum_{i=1}^{s-1} p_i + p_s] = E_{o/o}\,x.$$

Since we only calculate the sum in the numerator of (2), Bohlmann's suggestion leads to equating Ex to that sum. And, when assuming that the variable in cases of indefiniteness equals (2), we equate Ex to (2). My proposal with respect to the coefficient of dispersion, to the only magnitude of the type u/w for which the statistical literature possesses a rigorously established expected value (Markov, 1916; Chuporv, 1916; Chuprov, 1918 – 1919, vol. 1), leads in such cases, in accord with Markov and myself, to assuming value 1 for the variable. On the other hand, Bohlmann's suggestion to assume value 0 would have disturbed all the calculations leading to the solution. Just the same, neither does Bohlmann's suggestion provide a satisfactory solution of problems in cases considered below (§§2.1.A, 2.1.B, cf. Note 16 to §5).

2.1. The most direct way of calculating $E\,u/w$ immediately issues from the definition

$$E\,u/w = \sum_{i=1}^{k} \sum_{j=1}^{l} p_{ij}\,u_i/w_j.$$

This is how Markov (1916) approached his proof that the expectation of the coefficient of dispersion Q^2 is equal to 1 when the probability was invariable and successive trials were independent. The fruitfulness of this method is determined by the expression to which the double sum can be reduced. If however the derivations not only issue from the definition but do not go ahead, do not transform that sum further, then the solution of the formulated problem is certainly impossible. Without repeating Markov's calculations pertaining to the coefficient of dispersion, I illustrate this method by a simpler example.

2.1.A. Suppose that n mutually independent successive trials are performed on variables u and w having a constant law of connection described in §1.2. Let u and w take values u_i and w_j n_i and υ_j times respectively, and let n_{ij} denote the number of times when both these events occur at the same time. Then, let P_h be the probability that $n_i = h$ and $E^{(h)} n_{ij}$ be the conditional expectation of n_{ij} if $n_i = h$. Assuming that the fraction $n_{ij}/n_i = E_{o/o} n_{ij}/n_i$ when $n_{ij} = n_i = 0$ and noting that

$$E^{(h)} n_{ij} = h\, p_{ij}/p_i,$$

we find that

$$E\,(n_{ij}/n_i) = E_{o/o}\,(n_{ij}/n_i) = (1/\sum_{h\neq 0} P_h) \sum_{h\neq 0} (1/h) P_h E^{(h)} n_{ij} = p_{ij}/p_i.$$

2.1.B. In a similar way, noting that magnitudes n_{ij} and n_{gf}/n_g are mutually independent at a given value of n_i, we find that

$$E\,[(n_{ij}/n_i)(n_{gf}/n_g)] = [1/\sum_{h\neq 0} P_h] \sum_{h\neq 0} (1/h) P_h E^{(h)} n_{ij} E^{(h)} n_{gf}/n_g = (p_{ij}/p_i)(p_{gf}/p_g).$$

Therefore, the correlation coefficient between n_{ij}/n_i and n_{gf}/n_g is zero ($r = 0$). Under the assumptions made, our relations above seem rigorously precise.

2.1.C. Denote the conditional probability of the value u_i of u given that w takes value w_j by $p_{i(j)}$:

$$p_{i(j)} = p_{ij}/\pi_j.$$

Then {I delete the next few lines making difficult reading; they are repeated just below in a form of mathematical equations.}

$$\eta(j; r) = \sum_{i=1}^{k} p_{i(j)} u_i^2 = (1/\pi_j) \sum_{i=1}^{k} p_{ij} u_i^r,$$

$$\lambda(j; r) = \sum_{i=1}^{k} (n_{ij}/\upsilon_j) u_i^r.$$

In accord with the above $E\,(n_{ij}/\upsilon_j) = p_{ij}/\pi_j$ so that

$$E\,\lambda(j; r) = \eta(j; r).$$

As its derivation proves, this is a quite precise rather than an approximate relation.[3]

2.1.D. In a similar way we derive

$$E\,[\lambda(j_1; r_1)\,\lambda(j_2; r_2)] = \sum_{a=1}^{k} \sum_{b=1}^{k} u_a^{r_1} u_b^{r_2} \cdot E\,[(n_{a j_1}/\nu_{j_1})(n_{b j_2}/\nu_{j_2})] = \eta(j_1; r_1)\,\eta(j_2; r_2)$$

and, denoting the correlation coefficient between $\lambda(j_1; r_1)$ and $\lambda(j_2; r_2)$ by r, we have $r = 0$. These relations are also quite precise rather than approximate.[4]

2. 2. In some cases the value of $E(u/w)$ can be derived without tiresome calculation of the double sum. Thus, my proof, that the expectation of the coefficient of dispersion when the law of distribution of the variable is constant and the trials are mutually independent exactly equals 1, is reduced to establishing that, under the conditions of this problem, $E(u/w) = E(w/w)$.[5] Without reproducing the proof, I provide as an illustration a calculation, by an actually similar method, of the expectation and squared error of the correlation coefficient for the case of mutually independent variables.

2.2.A. Suppose that N mutually independent successive trials are performed on variables α and β connected by a constant law. Let α take possible values $\alpha_1, \alpha_2, \ldots, \alpha_k$ with probabilities $p_1, p_2, \ldots, p_k$ and β, – $\beta_1, \beta_2, \ldots, \beta_l$ with probabilities $\pi_1, \pi_2, \ldots, \pi_l$. Denote the probability that α and β take at the same time values α_i and β_j respectively by p_{ij}; then, let α_f and β_f be random values of α and β at trial f; $\alpha_{(N)}$ and $\beta_{(N)}$, the mean of α_f and β_f; n_i and υ_j, the number of trials in which α takes value α_i and β becomes β_j; and, finally, n_{ij} the number of times when α_i and β_j occur at the same time. Assuming as in §1.2 that

$$m_{ab} = E(\alpha^a \beta^b) = \sum_{i=1}^{k} \sum_{j=1}^{l} p_{ij} \alpha_i^a \beta_j^b,$$

$$\mu_{ab} = E[(\alpha - m_{10})^a (\beta - m_{01})^b] = \sum_{i=1}^{k} \sum_{j=1}^{l} p_{ij} [(\alpha_i - m_{10})^a$$
$$(\beta_j - m_{01})^b],$$

introduce new notation:

$$m'_{ab} = (1/N) \sum_{i=1}^{k} \sum_{j=1}^{l} n_{ij} \alpha_i^a \beta_j^b,$$

$$\mu'_{ab} = (1/N) \sum_{i=1}^{k} \sum_{j=1}^{l} [n_{ij} (\alpha_i - m_{10})^a (\beta_j - m_{01})^b].$$

Defining the correlation coefficient between α and β as

$$r = \mu_{11} / \sqrt{\mu_{20} \mu_{02}}$$

we consider the magnitude

$$\rho = \frac{\sum_{j=1}^{N} \{[\alpha_j - \alpha_{(N)}][\beta_j - \beta_{(N)}]\}}{\{\sum_{j=1}^{N} [\alpha_j - \alpha_{(N)}]^2 [\beta_j - \beta_{(N)}]^2\}^{1/2}} = \frac{m'_{11} - m'_{10} m'_{01}}{\{[m'_{20} - (m'_{10})^2][m'_{02} - (m'_{01})^2]\}^{1/2}} =$$

$$= \frac{\mu'_{11} - \mu'_{10} \mu'_{01}}{\{[\mu'_{20} - (\mu'_{10})^2][\mu'_{02} - (\mu'_{01})^2]\}^{1/2}}.$$

{In the sequel, Chuprov denotes the sums in the denominator of the first fraction in a special way; I introduce a somewhat better notation and do it right now: I denote these sums by $\Sigma[1; 2]$ and $\Sigma[2; 2]$ respectively.}

In practice, the approach to the estimation of r by empirical data is reduced to the calculation of ρ and it is tacitly assumed that $E\rho = r$, i.e., that

$$\mathrm{E}\,\frac{[1/(N-1)]\sum_{f=1}^{N}[\alpha_f - \alpha_{(N)}][\beta_f - \beta_{(N)}]}{\{[1/(N-1)]\Sigma[1;2]\}^{1/2}\{[1/(N-1)]\Sigma[2;2]\}^{1/2}} =$$

$$\frac{E[1/(N-1)]\sum_{f=1}^{N}[\alpha_f - \alpha_{(N)}][\beta_f - \beta_{(N)}]}{\{E[1/(N-1)]\,\Sigma[1;2]\}^{1/2}\{E[1/(N-1)]\,\Sigma[2;2]\}^{1/2}} = \frac{\mu_{11}}{\sqrt{\mu_{20}\mu_{02}}}.$$

This assumption, as the representatives of the Pearson school have also gradually come to understand, is wrong. Below, I will prove that $\mathrm{E}\rho$ can be either greater or less than r. However, in the particular case of mutual independence of the variables α and β both $\mathrm{E}\rho$ and r vanish so that $\mathrm{E}\rho = r$. Under the same condition

$$\mathrm{E}\rho = \mathrm{E}\frac{\sum_{f=1}^{N}[\alpha_f - \alpha_{(N)}][\beta_f - \beta_{(N)}]}{\{\Sigma[1;2]\,\Sigma[2;2]\}^{1/2}} =$$

$$\sum_{f=1}^{N} E\,\frac{\alpha_f - \alpha_{(N)}}{\sqrt{\Sigma[1;2]}}\;E\frac{\beta_f - \beta_{(N)}}{\sqrt{\Sigma[2;2]}}.$$

However, given the conditions of the problem under consideration, – an invariable law of connection and independence of successive trials, – we have

$$\mathrm{E}\frac{\alpha_f}{\sqrt{\Sigma[1;2]}} = \mathrm{E}\frac{\alpha_g}{\sqrt{\Sigma[1;2]}} = \mathrm{E}\frac{\alpha_{(N)}}{\sqrt{\Sigma[1;2]}}$$

and therefore

$$\mathrm{E}\frac{\alpha_f - \alpha_{(N)}}{\sqrt{\Sigma[1;2]}} = 0$$

and in the same way

$$\mathrm{E}\frac{\beta_f - \beta_{(N)}}{\sqrt{\Sigma[2;2]}} = 0.$$

Thus, if α and β are mutually independent,[6] $\mathrm{E}\rho = 0 = r$ for any laws of distribution of these variables.

2.2.B. In a similar way, noting that

$$\mathrm{E}\frac{[\alpha_f - \alpha_{(N)}][\alpha_g - \alpha_{(N)}]}{\Sigma[1;2]} = \frac{1}{N-1}\mathrm{E}\frac{[\alpha_f - \alpha_{(N)}]\sum_{g\neq f}[\alpha_g - (N-1)\alpha_{(N)}]}{\Sigma[1;2]} =$$

$$-\frac{1}{N-1}\,\frac{[\alpha_f - \alpha_{(N)}]^2}{\Sigma[1;2]},$$

$$\mathrm{E}\,\frac{[\beta_f - \beta_{(N)}][\beta_g - \beta_{(N)}]}{\Sigma[2;2]} = -\frac{1}{N-1}\mathrm{E}\frac{[\beta_f - \beta_{(N)}]^2}{\Sigma[2;2]},$$

we derive

$$\mathrm{E}\rho^2 = \sum_{f=1}^{N} \mathrm{E}\frac{[\alpha_f - \alpha_{(N)}]^2}{\Sigma[1;2]}\ \mathrm{E}\frac{[\beta_f - \beta_{(N)}]^2}{\Sigma[2;2]} +$$

$$\sum_{f=1}^{N} \sum_{g\neq f} \frac{[\alpha_f - \alpha_{(N)}][\alpha_g - \alpha_{(N)}]}{\Sigma[1;2]}\ \mathrm{E}\ \frac{[\beta_f - \beta_{(N)}][\beta_g - \beta_{(N)}]}{\Sigma[2;2]} =$$

$$\frac{N^2}{N-1}\ \mathrm{E}\frac{[\alpha_f - \alpha_{(N)}]^2}{\Sigma[1;2]}\ \mathrm{E}\ \frac{[\beta_f - \beta_{(N)}]^2}{\Sigma[2;2]} =$$

$$\frac{1}{N-1}\mathrm{E}\frac{\sum_{f=1}^{N}[\alpha_f - \alpha_{(N)}]^2}{\Sigma[1;2]}\ \mathrm{E}\frac{\sum_{f=1}^{N}[\beta_f - \beta_{(N)}]^2}{\Sigma[2;2]} = \frac{1}{N-1}.$$

Thus, when α and β are mutually independent, the squared error of ρ is exactly equal to $\sqrt{1/(N-1)}$ whichever are the laws of distribution of these variables.[7]

2.2.C. By similar considerations we derive without difficulties the expectations of the higher degrees of ρ:

$$\mathrm{E}\rho^3 = \frac{N}{(N-1)(N-2)}\ \mathrm{E}\frac{\sum_{f=1}^{N}[\alpha_f - \alpha_{(N)}]^3}{\Sigma[1;3]}\ \mathrm{E}\frac{\sum_{f=1}^{N}[\beta_f - \beta_{(N)}]^3}{\Sigma[2;3]},$$

$$\mathrm{E}\rho^4 = \frac{1}{(N+1)N(N-1)(N-2)(N-3)}\{[N(N+1)\ \mathrm{E}\ \frac{\sum_{f=1}^{N}[\alpha_f - \alpha_{(N)}]^4}{\Sigma[1;4]} -$$
$$3(N-1)][N(N+1)\ \mathrm{E}\frac{\sum_{f=1}^{N}[\beta_f - \beta_{(N)}]^4}{\Sigma[2;4]} - 3(N-1)] + 3N(N^2 + N + 6\}.$$

{Chuprov had not explained his symbol Σ [1; 3] (or Σ [2; 3]); below, in §2.3, and only indirectly, he stated that

$$\Sigma\,[1;4] = \sum_{f=1}^{N} [\alpha_f - \alpha_{(N)}]^4 + \sum_{f=1}^{N} \sum_{g\neq f} [\alpha_f - \alpha_{(N)}]^2\,[\alpha_g - \alpha_{(N)}]^2.\} \qquad (3)$$

We convince ourselves that both $\mathrm{E}\rho^3$ and $\mathrm{E}\rho^4$ depend on the laws of distribution of the variables even if they are not connected with each other.

2.3. Issuing from the definition

$$\mathrm{E}\,(u/w) = \sum_{i=1}^{k} \sum_{j=1}^{l} p_{ij}\,(u_i/w_j)$$

it is sometimes possible to judge E (u/w) even if we are unable to reduce the double sum to a form ensuring a complete solution of the problem at hand. Thus, Markov reduces the expectation of the squared error of the coefficient of dispersion to an ordinary sum. However, the obtained expression

$$\mathrm{E}\,(Q^2 - 1)^2 = \frac{2r^2 n(n-1)}{(r-1)(nr-2)(nr-3)}\ .$$

$$\sum_{m=1}^{nr-1} \frac{(m-1)(nr-m-1)(nr)!}{m(nr-m)m!(nr-m)!} p^m (1 - p)^{nr-m}$$

where r and n are the number of series and of trials in a series respectively and p is the probability of the event, remains too involved for judging the magnitude of the squared error. Nevertheless, it allows to perceive without difficulty that

$$E (Q^2 - 1)^2 < \frac{2n(nr-2)}{(r-1)(n-1)(nr-3)}$$

and that, if p is not too low, and the series consist of the same number of trials, the left side tends to $2/(r - 1)$ as nr increases.

Just the same, my own method, although not enabling to establish precisely the expected squared error of the same coefficient, allows to indicate that (again for series containing the same number of trials) its upper bound persists for whichever law of distribution of the variable.

For the Markov case we can also point out the lower bound, namely

$$E (Q^2 - 1)^2 \geq \frac{2(n-1)(nr-1)}{(r-1)n\, nr} 16p^2(1 - p)^2.$$

My method of calculation allows to discover that an irregular distribution of the trials among the series can influence the squared error of the coefficient of dispersion in either direction depending on the law of distribution of the variables (Chuprov 1918 – 1919, Bd. 1, pp. 230, 236, 238, 246). In a similar way, the expression for $E\rho^4$ derived above allows to establish its boundaries for whichever law of distribution of mutually independent α and β. Noting that

$$E \frac{\sum_{f=1}^{N}[\alpha_f - \alpha_{(N)}]^4}{\Sigma[1;4]} = E \text{ \{here Chuprov indirectly introduces notation (3)\}},$$

that, consequently, the left side is less than unity, and that, on the other hand, it is greater than

$$\{ E\frac{[\alpha_f - \alpha_{(N)}]^2}{\Sigma[1;4]} \}^2, \text{ or greater than } (1/N^2),$$

we find that

$$\frac{3N^3 + 7N^2 + 2N + 16}{(N+1)N(N-1)(N-2)(N-3)} < E\rho^4 <.$$

$$\frac{(N^2 - 2N + 3)^2 + 3N(N^2 + N + 6)}{(N+1)N(N-1)(N-2)(N-3)}$$

so that

$$[E\rho^4 / (E\rho^2)^2] - 3 > (4/N) \frac{4N^3 - 2N^2 - N - 4}{N^3 - 4N^2 + N + 6}.$$

We thus convince ourselves that in case of mutually independent α and β and not a very considerable number of trials N the distribution of ρ noticeably deviates from the Gauss law.

3.1. if the sign of all possible values of u/w does not vary and remains either ≥ 0 or ≤ 0, the boundaries of $E(u/w)$ can be outlined by simple considerations. Issuing from the identity

$$(1/w) = (1/Ew) - (w - Ew)/(wEw),$$

we derive

$$E(u/w) = (Eu/Ew) - (1/Ew)\,E\frac{u(w-Ew)}{w} = (Eu/Ew) -$$

$$\frac{Eu(w-Ew)}{(Ew)^2} + [1/(Ew)^2]\,E\frac{u(w-Ew)^2}{w} = \ldots = (Eu/Ew) +$$

$$\sum_{h=1}^{t}\frac{(-1)^h Eu(w-Ew)^h}{(Ew)^{h+1}} + \frac{(-1)^{t+1}}{(Ew)^{t+1}}\,E\,\frac{u(w-Ew)^{t+1}}{w}.$$

If u/w always remains non-negative, then

$$E\frac{u(w-Ew)^{2t}}{w} > 0,$$

$$E(u/w) > Eu/Ew + \sum_{h=1}^{2t-1}\frac{(-1)^h Eu(w-Ew)^h}{(Ew)^{h+1}}. \qquad (4)$$

In particular, under the stipulated conditions,

$$E(u/w) > (Eu/Ew) - \frac{Eu(w-Ew)}{(Ew)^2}, \qquad (4a)$$

$$E(u/w) > (Eu/Ew) - \frac{E(u-Eu)(w-Ew)}{(Ew)^2},$$

$$E(u/w) > (Eu/Ew)\{1 - r_{11}[(Eu^2/(Eu)^2) - 1]^{1/2}[(Ew^2/(Ew)^2) - 1]^{1/2}.$$

If u/w always remains greater than some magnitude γ^2 then

$$E(u/w) > Eu/Ew + \sum_{h=1}^{2t-1}\frac{(-1)^h Eu(w-Ew)^h}{(Ew)^{h+1}} +$$

$$(1/(Ew)^{2t}\,E\,\gamma^2(w - Ew)^{2t}. \qquad (5)$$

If γ^2 is here constant rather than a randomly variable magnitude, the last term of (5) will be

$$(\gamma^2/(Ew)^{2t}\,E\,(w - Ew)^{2t}.$$

In a similar way, if u/w always remains less than some δ^2, then

$$E(u/w) < Eu/Ew + \sum_{h=1}^{2t-1}\frac{(-1)^h Eu(w-Ew)^h}{(Ew)^{h+1}} +$$

$$(1/(Ew)^{2t}\,E\,\delta^2(w - Ew)^{2t}. \qquad (6)$$

In many cases interesting for the statistician u/w, whose expectation is calculated, cannot exceed 1. Assuming that $\delta^2 = 1$, we have (6a) {Chuprov rewrites formula (6) substituting 1 instead of δ^2 and calls it expression (6a).}. How tight do the inequalities (4), or (5), and (6) enclose $E(u/w)$ depends, on the one hand, on the value of $2t$, the highest exponent in these formulas, and, on the other hand, on the values of δ^2 and γ^2 which we were able to choose. If we were unable to find a sufficiently large value of δ^2, the inequality (6) loses practical interest, but it can nevertheless be applied for deriving the value to which $E(u/w)$ tends as the number of trials increases.

3.2. We note that, in the notation of §2.2,

$$1 - \rho^2 = (1/2)\,\frac{\sum_{f=1}^{N}\sum_{h\neq f}\{[\alpha_f-\alpha_{(N)}][\beta_h-\beta_{(N)}]-[\alpha_h-\alpha_{(N)}][\beta_f-\beta_{(N)}]\}^2}{\sum[1;2]\sum[2;2]}$$

and convince ourselves that $\rho^2 \le 1$.

Therefore, assuming that

$$u = [1/(N-1)^2]\,\{\sum_{f=1}^{N}\ [\alpha_j - \alpha_{(N)}]\,[\beta_j - \beta_{(N)}]\}^2,$$

$$w = [1/(N-1)^2]\sum_{f=1}^{N}\ [\alpha_j - \alpha_{(N)}]^2\,[\beta_j - \beta_{(N)}]^2,$$

we can use inequalities (4) and (6a) for calculating $E\rho^2$. Without difficulties, but after rather laborious calculations we derive

$$Eu = \mu_{20}\,\mu_{02}\,\{r^2_{11} + (1/N)\,[r_{22} - r^2_{11}] + [1/N(N-1)]\,[1 + r^2_{11}]\},$$

$$Ew = \mu_{20}\,\mu_{02}\,\{1 + (1/N)\,[r_{22} - 1] + \{2/[N(N-1)]\}\,r^2_{11}\},$$

$$E\,uw = \mu^2_{20}\,\mu^2_{02}\,\{(1/N^3)r_{44} + \frac{N^2-2N+6}{N^3(N-1)}\,(r_{42} + r_{24}) -$$

$$\frac{2(N^2-2N+2)}{N^3(N-1)^2}\,[r_{41}\,r_{03} + r_{14}\,r_{30}] + [1/N^3(N-1)]\,r_{40}\,r_{04} +$$

$$\frac{N^2-5N+6}{N^3(N-1)^2}\,(r_{40} + r_{04}) + \frac{2(N^2-2N+8)}{N^3(N-1)}\,r_{33}\,r_{11} -$$

$$\frac{2(5N^2-10N+12)}{N^3(N-1)^2}\,(r_{32}\,r_{12} + r_{23}\,r_{21}) +$$

$$\frac{2(N^4-4N^3+11N^2-14N+8)}{N^3(N-1)^3}\,r_{31}\,r_{13} +$$

$$\frac{2(N-2)(N^4-4N^3+13N^2-28N+24)}{N^3(N-1)^3}\,r_{11}\,(r_{31} + r_{13}) -$$

$$\frac{2(N-2)(N^3-6N^2+13N-12)}{N^3(N-1)^3}\,(r_{21}\,r_{03} + r_{12}\,r_{30}) -$$

$$\frac{2(N-2)(N^3-4N^2+9N-8)}{N^3(N-1)^3}\, r_{11}\, r_{30}\, r_{03} \;+\; \frac{N^4-4N^3+19N^2-30N+18}{N^3(N-1)^3}\, r_{23}{}^2 \;+$$

$$\frac{(N-2)(N^4-4N^3+17N^2-42N+36)}{N^3(N-1)^3}\, r_{22} \;+$$

$$\frac{(N-2)(N^4-4N^3+25N^2-78N+72)}{N^3(N-1)^3}\, r_{22}\, r^2{}_{11} \;-$$

$$\frac{2(N-2)(5N^3-28N^2+69N-72)}{N^3(N-1)^3}\, r_{21}\, r_{12}\, r_{11} \;-$$

$$\frac{4(N-2)(N^3-4N^2+9N-9)}{N^3(N-1)^3}\, (r^2{}_{12} \;+\; r^2{}_{21}) \;+$$

$$\frac{2(N-2)(N-3)(N^2-6N+12)}{N^3(N-1)^3}\, r^4{}_{11} \;+$$

$$\frac{(N-2)(N-3)(N^4-4N^3+15N^2-42N+72)}{N^3(N-1)^3}\, r^2{}_{11} \;+\; \frac{(N-2)(N-3)^3}{N^3(N-1)^3}\,\}.$$

$$\mathrm{E}w^2 \;=\; \mu^2{}_{20}\,\mu^2{}_{02}\,\{(1/N^3)\, r_{44} \;+\; \frac{2(N^2-2N+3)}{N^3(N-1)}\,(r_{42} \;+\; r_{24}) \;-$$

$$(4/N^3)\,(r_{41}\, r_{03} \;+\; r_{14}\, r_{30}) \;+\; [(N-1)\,/N^3]\, r_{40}\, r_{04} \;+$$

$$\frac{(N-2)(N^2-2N+3)}{N^3(N-1)}\,(r_{40} \;+\; r_{04}) \;+\; \frac{16}{N^3(N-1)}\, r_{33}\, r_{11} \;-$$

$$\frac{8(N^2-2N+3)}{N^3(N-1)^2}\,(r_{32}\, r_{12} \;+\; r_{23}\, r_{21}) \;+\; \frac{16}{N^3(N-1)}\, r_{31}\, r_{13} \;+$$

$$\frac{16(N-2)(N-3)}{N^3(N-1)^2}\, r_{11}\,(r_{31} \;+\; r_{13}) \;-$$

$$\frac{8(N-2)(N^2-2N+3)}{N^3(N-1)^2}(r_{21}\, r_{03} \;+\; r_{12}\, r_{30}) \;+\; \frac{16(N-2)}{N^3(N-1)}\, r_{11}\, r_{30}\, r_{03} \;+$$

$$\frac{2(N^2-2N+3)^2}{N^3(N-1)^3}\, r^2{}_{22} \;+\; \frac{4(N-2)(N^2-2N+3)^2}{N^3(N-1)^3}\, r_{22} \;+$$

$$\frac{8(N-2)(N-3)^2}{N^3(N-1)^3}\, r_{22}\, r^2{}_{11} \;+\; \frac{16(N-2)(N-3)^2}{N^3(N-1)^3}\, r_{21}\, r_{12}\, r_{11} \;-$$

$$\frac{4(N-2)(N-3)(N^2-2N+3)}{N^3(N-1)^3}(r^2{}_{21} \;+\; r^2{}_{12}) \;+\; \frac{24(N-2)(N-3)}{N^3(N-1)^3}\, r^4{}_{11} \;+$$

$$\frac{8(N-2)(N-3)^3}{N^3(N-1)^3}\, r^2_{11} + \frac{(N-2)(N-3)(N^2-2N+3)^2}{N^3(N-1)^3}\}.$$[8]

Therefore

$$\mathrm{E}\,uw - \mathrm{E}u\,\mathrm{E}w = \mu^2_{20}\,\mu^2_{02}\,\{(2/N)[r_{11}\,(r_{31} + r_{13}) - 2r^2_{11}] +$$

$$(1/N^2)\,[r_{42} + r_{24} - 2r_{22} - 4(r^2_{21} + r^2_{12}) - 2r_{22}\,r^2_{11} -$$

$$10r_{21}\,r_{12}\,r_{11} + r^4_{11} + 18r^2_{11} + 2r_{33}\,r_{11} - 2r_{21}\,r_{03} - 2r_{12}\,r_{30} +$$

$$2r_{31}\,r_{13} - 6r_{11}(r_{31} + r_{13}) - 2r_{11}\,r_{30}\,r_{03}] + \ldots\}.$$

$$\mathrm{E}(w - \mathrm{E}w)^2 = \mu^2_{20}\,\mu^2_{02}\,\{(1/N)\,[r_{40} + r_{04} + 2r_{22} - 4] +$$

$$(1/N^2)[r_{40}\,r_{04} + r^2_{22} - 3r_{40} - 3r_{04} - 10r_{22} + 14 +$$

$$2r_{42} + 2r_{24} - 4r^2_{21} - 4r^2_{12} + 4r^2_{11} - 4r_{12}\,r_{30} - 4r_{21}\,r_{03}] + \ldots\}.$$

If

$$Nr_{44}/N^4,\ Nr_{42}/N^3,\ Nr_{24}/N^3,\ Nr_{41}/N^3,\ Nr_{14}/N^3,\ Nr_{22}/N^2,\ \text{etc},$$

tend to vanish as N increases, – this condition is indeed fulfilled if

$$N\mu_{40} / (N\mu_{20})^2,\ N\mu_{04} / (N\mu_{02})^2,\ N\mu_{30} / (N\mu_{20})^{3/2} \text{ and } N\mu_{03} / (N\mu_{02})^{3/2}$$

also tend to 0 as N increases, – then

$$\mathrm{E}(w - \mathrm{E}w)^2/(\mathrm{E}w)^2$$

behaves in the same way and the boundaries for $\mathrm{E}\rho^2$, calculated in accord with inequalities (4) and (6a) for $t = 1$, close in at $N = \infty$ upon r^2_{11}.

Thus, under the restrictions formulated above, $\mathrm{E}\rho^2$ tends to r^2_{11} as N increases and the random empirical value of ρ^2, established by a large but finite number of observations, can with some confidence be considered as the approximate representation of the theoretical magnitude r_{11} that characterizes the law of connection between α and β.

Inequalities (4) and (6a) allow us to outline, as tightly as desired, those boundaries within which the value of $\mathrm{E}\rho^2$ is enclosed given a no matter how small number of trials N as well as to enable the establishment of terms of any order in the expansion of $\mathrm{E}\rho^2$ in powers of $(1/N)$. Thus, supposing that $t = 1$, we find in this expansion the term free from $(1/N)$ because the expression $\mathrm{E}(w - \mathrm{E}w)^2 / (\mathrm{E}w)^2$ does not include such terms. Supposing that $t = 2$, we also determine the term of the order of $(1/N)$, etc. However, for greater values of t this method is so involved that another method (§4.2) seems more practical.

3.3. Keeping to the notation of §2.1.A and noting that $0 \le n^2_{ij} / \upsilon^2_j \le 1$, we convince ourselves that

$$\mathrm{E}n^2_{ij}/\mathrm{E}\,\upsilon^2_j + \sum_{h=1}^{2t-1} \frac{(-1)^h En^2_{ij}[\nu^2_j - E\nu^2_j]^h}{(E\nu^2_j)^{h+1}} < \mathrm{E}\,(n^2_{ij}/\upsilon^2_j) < [\mathrm{E}n^2_{ij}/\mathrm{E}\upsilon^2_j] +$$

$$\sum_{h=1}^{2t-1} \frac{(-1)^h En^2_{ij}[\nu^2_j - E\nu^2_j]^h}{(E\nu^2_j)^{h+1}} + \frac{E[\nu^2_j - E\nu^2_j]^{2t}}{(E\nu^2_j)^{2t}}. \qquad (7)$$

However, neither the expansion of the last fraction just above in powers of $(1/N)$, nor the expansion of

$$\frac{En^2_{ij}[\nu^2_j - E\nu^2_j]^{2t-1}}{(E\nu^2_j)^{2t}}$$

contain terms of the order $(1/N^{t-1})$, $(1/N^{t-2})$, …, $(1/N)$, $(1/N^{o})$.

Considering the expression in the left side of the first inequality (7) with summation extended from $h = 1$ to $2t - 2$, we can therefore precisely establish the terms up to order $(1/N^{t-1})$ inclusive in the expansion of $E\, n^2_{ij} / \upsilon^2_j$ in powers of $(1/N)$. As t increases, calculations rapidly become ever more complicated. Supposing that $t = 3$ and restricting our calculations in both parts of the fraction

$$\{1/[E\upsilon^2_j]^5\}\{5En^2_{ij}[E\upsilon^2_j]^4 - 10\, En^2_{ij}\upsilon^2_j[E\upsilon^2_j]^3 + 10\, En^2_{ij}\upsilon^4_j[E\upsilon^2_j]^2 -$$

$$5En^2_{ij}\upsilon^6_j E\upsilon_j^2 + En^2_{ij}\upsilon^8_j\}$$

by terms of the order N^{10}, N^9, N^8, we derive $E\,(n^2_{ij}/\upsilon^2_j)$ to terms of the order $(1/N^2)$ inclusively:

$$E\,(n^2_{ij}/\upsilon^2_j) = \frac{p^2_{ij}}{\pi^2_j} + \frac{1}{N\pi^3_j} p_{ij}(\pi_j - p_{ij}) + \frac{1}{N^2\pi^4_j} p_{ij}(\pi_j \;\; p_{ij})(1 - \pi_j) + \ldots$$

In a similar way we derive[9]

$$E\frac{n_{ij}n_{fj}}{\nu^2_j} = \frac{p_{ij}p_{fj}}{\pi_j^2}\,[1 - \frac{1}{N\pi_j} - \frac{1}{N^2\pi_j^2}(1 - \pi_j) + \ldots],$$

$$E\frac{n^2_{ij}}{n_i\nu_j} = \frac{p^2_{ij}}{p_i\pi_j}\{1 + \frac{1}{Np_{ij}}[1 - (p_{ij}/p_i)]\,[1 - (p_{ij}/\pi_j)] -$$

$$\frac{1}{N^2p_{ij}}[1 - (p_{ij}/p_i)]\,[1 - (p_{ij}/\pi_j)]\,[1 - (2p_{ij}/p_i p_j)] + \ldots\},$$

$$E\frac{n_{ij}{}^4}{n_i^2\nu_j^2} = \frac{p^4_{ij}}{p^2_i\pi^2_j}\{1 + \frac{1}{Np_{ij}}[6 - 5\,(\frac{p_{ij}}{p_i} + \frac{p_{ij}}{\pi_j}) + 4\frac{p^2_{ij}}{p_i\pi_j}] +$$

$$\frac{1}{N^2p^2_{ij}}[7 - 6p_{ij}[1 + (3/p_i) + 3/\pi_j] + p^2_{ij}[(5/p_i) + (5/\pi_j) +$$

$$(49/p_i\pi_j) + (11/p^2_i) + (11/\pi^2_j) - ((p^3_{ij}/p_i\pi_j)[4 + (30/p_i) +$$

$$(30/\pi_j) + (18p^4_{ij}/p^2_i\pi_j^2)] \ldots \},$$

$$E\frac{n^2_{ij}n^2_{if}}{n^2_i\nu_j\nu_f} = \frac{p^2_{ij}p^2_{if}}{p^2_i\pi_j\pi_f}\{1 + (1/N)[(1/p_{ij}) + (1/p_{if}) - (1/\pi_j) - (1/\pi_f) -$$

$$(5/p_i) + \frac{2p_{ij}}{p_i\pi_j} + \frac{2p_{if}}{p_i\pi_f}] + \ldots\},$$

$$E\frac{n^2_{ij}n^2_{fh}}{n_i\nu_j n_f\nu_h} = \frac{p^2_{ij}p^2_{fh}}{p_i\pi_j p_f\pi_h}\;\{1 + (1/N)[(1/p_{ij}) + (1/p_{fh}) - (1/p_i) -$$

$(1/\pi_j) - (1/p_f) - (1/\pi_h) + \frac{p_{ij}}{p_i\pi_j} + \frac{p_{ih}}{p_i\pi_h} + \frac{p_{fj}}{p_f\pi_j} + \frac{p_{fh}}{p_f\pi_h}] + \ldots\}.$

3.4.A. In the notation of §2.1.A we denote

$$\varphi^2 = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{(p_{ij} - p_i\pi_j)^2}{p_i\pi_j}.$$

For mutually independent u and w we have $p_{ij} = p_i\pi_j$ for any i and j and $\varphi^2 = 0$; inversely, φ^2 can only vanish if $p_{ij} = p_i\pi_j$ for any i and j. Thus, if $\varphi^2 = 0$, variables u and w are mutually independent. When there exists an exact functional connection between u and w, definite values of w correspond to definite values of u and $p_{ij} = 0$ for $j \neq i$, $p_{ij} = p_i = \pi_i$. Then $\varphi^2 = k - 1$.

We thus have $[1/(k - 1)]\,\varphi^2$ as a measure of the closeness of the connection between the variables. It vanishes in case of their mutual independence, – and only in this case, – and takes value 1 when there exists an exact functional dependence. This measure has essential advantage over the correlation coefficient r_{11} since the equality $r_{11} = 0$ is a necessary but not sufficient condition for the mutual independence of the variables. Furthermore, the coefficient r^2_{11} takes value 1 only in the case of their linear functional connection; when, however, the connection is exactly functional but not linear, $r^2_{11} < 1$.

3.4.B. Representing φ^2 in the form

$$\varphi^2 = (1/N^2)\sum_{i=1}^{k} \sum_{j=1}^{l} \frac{[En_{ij} - (1/N)En_iE\nu_j]^2}{En_iE\nu_j}$$

we denote

$\varphi^2_t =$ {here Chuprov repeats the double sum written just above. The same is true with regard to φ^2_τ and φ^2_a below.} =

$$= (1/N^2)\sum_{i=1}^{k} \sum_{j=1}^{l} \frac{(n_{ij} - Np_i\pi_j)^2}{p_i\pi_j},$$

$$\varphi^2_\tau = [\ \ldots\ ] = (1/N^2)\sum_{i=1}^{k} \sum_{j=1}^{l} \frac{[n_{ij} - (1/N)n_i\nu_j]^2}{p_i\pi_j},$$

$$\varphi^2_a = [\ \ldots\ ] = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{[n_{ij} - (1/N)n_i\nu_j]^2}{n_i\nu_j}.$$

Unlike the coefficient φ^2 whose numerical value can be established only when we know beforehand the law of connection between the variables, the values of φ^2_t as well as of φ^2_τ can be determined empirically, – if, before the trials we know nothing about the connection between the variables, but if, independently from the trials, we know the law of distribution of each variable separately. As to φ^2_a, its numerical value is completely determined by the N trials performed to this end.

Pearson intuitively introduced {the Russian verb is incomprehensible, my translation is only likely} the magnitude φ^2 when he constructed a measure of closeness of connection that he called *mean square contingency* but only later did he grasp the difference in the structure of the coefficients φ^2, φ^2_t and φ^2_a; just as the difference in the construction of the coefficients r and ρ (§2.2.A), until recently it had not remained quite evident to the representatives of the English school. This was the natural result of their disinclination to the concepts of mathematical probability and expectation without which it is hardly possible to formulate clearly such subtle and, at the same time, such essential distinctions. Nowadays Pearson (Young & Pearson 1916, vol. 11, p. 220) distinguishes between φ^2 (*the mean square contingency for the whole population*), φ^2_a (*the approximate value of the mean square contingency*) and φ^2_t (*its true value).*[10] Russian statistical terminology has no established term for φ^2; I call it the indicator of the mutual contingency of variables u and w.

3.4.C. The magnitude $\varphi^2{}_t$ as well as $\varphi^2{}_\tau$ is easily investigated. We find without difficulty that

$$E\,\varphi^2{}_t = \varphi^2 + (1/N)\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p_{ij}(1-p_{ij})}{p_i\pi_j}$$

and convince ourselves that always $E\,\varphi^2{}_t > \varphi^2$. When the variables are mutually independent, $\varphi^2 = 0$ and $E\,\varphi^2{}_t = [(kl - 1)/N]$. In case of exact functional dependence

$$E\,\varphi^2{}_t = (k - 1) + (1/N)\sum_{i=1}^{k}\frac{1-p_i}{p_i}.$$

Then, since

$$E\,\varphi^4{}_t = (1/N^4)\{\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{E(n_{ij}-Np_i\pi_j)^4}{p^2{}_i\pi^2{}_j} +$$

$$\sum_{i=1}^{k}\sum_{j=1}^{l} S\,\frac{E(n_{ij}-Np_i\pi_j)^2(n_{fh}-Np_f\pi_h)^2}{p_i\pi_j p_f\pi_h}\}$$

where S stands for a double sum over all the possible values of f and h excepting the case of simultaneous equalities $f = i$ and $h = j$, we derive without special difficulties

$$E\,[\varphi^2{}_t - E\,\varphi^2{}_t]^2 = (4/N)\{\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{(p_{ij}-p_i\pi_j)^3}{p^2{}_i\pi^2{}_j} + \varphi^2 - \varphi^4\} +$$

$$(1/N^2)\{6\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{(p_{ij}-p_i\pi_j)^2}{p^2{}_i\pi^2{}_j} - 12\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{(p_{ij}-p_i\pi_j)^3}{p^2{}_i\pi^2{}_j} +$$

$$(8 - 4\varphi^2)\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p_{ij}-p_i\pi_j}{p_i\pi_j} + 10\,\varphi^2 - 16\,\varphi^4 - 4kl\,\varphi^2 + 2(kl - 1)\} +$$

$$(1/N^3)\{8\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{(p_{ij}-p_i\pi_j)^3}{p^2{}_i\pi^2{}_j} - 6\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{(p_{ij}-p_i\pi_j)^2}{p^2{}_i\pi^2{}_j} +$$

$$\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p_{ij}-p_i\pi_j}{p^2{}_i\pi^2{}_j} - [\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p_{ij}-p_i\pi_j}{p_i\pi_j}]^2 - (2kl + 8 - 4\varphi^2)\cdot$$

$$\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p_{ij}-p_i\pi_j}{p_i\pi_j} + \sum_{i=1}^{k}\sum_{j=1}^{l}[1/p_i\,\pi_j] - 6\,\varphi^4 + 12\,\varphi^2 + 4kl\,\varphi^2 -$$

$$k^2l^2 - 2kl + 2\}.$$

If the variables are mutually independent, then, cf. Pearson (1919, pp. 259 – 260),

$$E\,[\varphi^2{}_t - E\,\varphi^2{}_t]^2 = (2/N^2)(kl - 1) +$$

$$(1/N^3)\{\sum_{i=1}^{k}\sum_{j=1}^{l}[1/p_i\,\pi_j] - k^2l^2 - 2kl + 2\}.$$

3.4.D. Representing $\varphi^2{}_a$ as

$$\varphi^2{}_a = \sum_{i=1}^{k}\sum_{j=1}^{l}\frac{n^2{}_{ij}}{n_i\nu_j} - 1$$

we derive (§3.3)

$$\text{E}\,\varphi^2{}_a = \varphi^2 + (1/N)\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p_{ij}(p_i - p_{ij})(\pi_j - p_{ij})}{p^2{}_i\pi^2{}_j} +$$

$$(1/N^2)\{\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p_{ij}(p_i - p_{ij})(\pi_j - p_{ij})}{p^2{}_i\pi^2{}_j}\frac{p_{ij} - p_i\pi_j}{p_i\pi_j} +$$

$$\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p^2{}_{ij}(p_i - p_{ij})(\pi_j - p_{ij})}{p^3{}_i\pi^3{}_j}\} + \dots$$

Thus, $\text{E}\,\varphi^2{}_a \geq \varphi^2$. In case of mutually independent variables

$$\text{E}\,\varphi^2{}_a = \frac{(k-1)(l-1)}{N} + \frac{(k-1)(l-1)}{N^2} + \dots$$

and, when they are connected by an exact functional dependence, $\text{E}\,\varphi^2{}_a = k - 1 = \varphi^2$.

Noting that

$$[\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{n^2{}_{ij}}{n_i\nu_j}]^2 = \sum_{i=1}^{k}\sum_{j=1}^{l}\frac{n^4{}_{ij}}{n^2{}_i\nu^2{}_j} + \sum_{i=1}^{k}\sum_{j=1}^{l}\sum_{h\neq j}\frac{n_{ij}^2 n_{ih}^2}{n_i^2\nu_j\nu_h} +$$

$$\sum_{i=1}^{k}\sum_{j=1}^{l}\sum_{f\neq i}\frac{n_{ij}^2 n_{fj}^2}{n_i\nu_j^2 n_f} + \sum_{i=1}^{k}\sum_{j=1}^{l}\sum_{f\neq i}\sum_{h\neq j}\frac{n_{ij}^2 n_{fh}^2}{n_i\nu_j n_f\nu_h},$$

we derive without any special difficulties but performing rather laborious calculations,

$$\text{E}\,[\varphi^2{}_a - \text{E}\,\varphi^2{}_a]^2 = (1/N)\{4\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p_{ij}^3}{p_i^2\pi_j^2} -$$

$$3\sum_{i=1}^{k}(1/p_i^3)[\sum_{j=1}^{l}(p_{ij}^2/\pi_j)]^2 - 3\sum_{j=1}^{l}(1/\pi_j^3)[\sum_{i=1}^{k}(p_{ij}^2/p_i)]^2 +$$

$$2\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p_{ij}}{p_i^2\pi_j^2}[\sum_{f=1}^{k}(p_{fj}^2/p_f)][\sum_{h=1}^{l}(p_{ih}^2/\pi_h)]\} + \dots$$ [11]

3.5. In notation of §2.1.C, and taking into account the relations of §3.3, we have

$$\text{E}[\lambda^2(j;r)] = \eta^2(j;r) + (1/N\pi_j)[\eta(j;2r)] - \eta^2(j;r)]$$

$$[1 + (1/N\pi_j)(1 - \pi_j) + \dots].$$

In the same way, denoting

$$\mu (j; r) = (1/\pi_j) \sum_{i=1}^{k} p_{ij} [x_i - \eta (j; 1)]^r,$$

$$\tau (j; r) = (1/\upsilon_j) \sum_{i=1}^{k} n_{ij} [x_i - \eta (j; 1)]^r,$$

we obtain

$$E \tau^2 (j; r) = \mu^2 (j; r) + (1/N \pi_j) [\mu (j; 2r) - \mu^2 (j; r)]$$

$$[1 + (1/N \pi_j) (1 - \pi_j) + \dots].$$

Similarly

$$E \tau^3 (j; r) = \mu^3 (j; r) + (3/N \pi_j) [\mu (j; 2r) - \mu^2 (j; r)] \mu (j; r) \cdot$$

$$[1 - (1/N) + \dots] + (1/N^2 \pi_j^2) \mu (j; 3r) - \mu^3 (j; r)] + \dots,$$

$$E \tau^4 (j; r) = \mu^4 (j; r) + (6/N \pi_j) [\mu (j; 2r) - \mu^2 (j; r)] \mu^2 (j; r) \cdot$$

$$[1 - (1/N) + \dots] + (1/N^2 \pi_j^2) [4 \mu (j; 3r)] \mu (j; r) + 3 \mu^2 (j; 2r)] -$$

$$12 \mu (j; 2r)] \mu^2 (j; r) + 5 \mu^4 (j; r)] + \dots$$

and therefore

$$E [\tau (j; r) - \mu (j; r)]^2 = (1/N \pi_j) [\mu (j; 2r) - \mu^2 (j; r)] \cdot$$

$$[1 + (1/N \pi_j) (1 - \pi_j) + \dots],$$

$$E [\tau (j; r) - \mu (j; r)]^3 = (1/N^2 \pi_j^2) [\mu (j; 3r) - 3\mu (j; 2r) \mu (j; r)] +$$

$$2\mu^3(j; r)] + \dots,$$

$$E [\tau (j; r) - \mu (j; r)]^4 = (3/N^2 \pi_j^2) [\mu (j; 2r) - \mu^2 (j; 2r)]^2 + \dots$$

If $r = 1$ then (Pearson 1919, p. 260ff)

$$E \tau (j; 1) = 0,\ E \tau^2 (j; 1) = (1/N \pi_j) \mu (j; 2) [1 + (1/N \pi_j) (1 - \pi_j) + \dots],$$

$$E \tau^3 (j; 1) = (1/N^2 \pi_j^2) \mu (j; 3) + \dots,\ E \tau^4 (j; 1) = (3/N^2 \pi_j^2) \mu^2 (j; 2) + \dots$$

3.6.A. Suppose that the variables, whose connection with each other is investigated, can take only two different values, and denote

$$Q = \frac{p_{11}p_{22} - p_{12}p_{21}}{p_{11}p_{22} + p_{12}p_{21}},\quad Q_1 = \frac{n_{11}n_{22} - n_{12}n_{21}}{n_{11}n_{22} + n_{12}n_{21}}.$$

Yule (1900, p. 272; 1912, p. 593) calls Q_1 the coefficient of association and does not distinguish it from Q with due clearness. For mutually independent variables $Q = 0$; inversely, if $Q = 0$, the variables are mutually independent. In case of exact functional dependence $Q = \pm 1$. Inversely, if $Q = \pm 1$, the dependence is exactly functional. As to Q_1, after representing it as

$$Q_1 = 2\frac{n_{11}n_{22}}{n_{11}n_{22} + n_{12}n_{21}} - 1$$

and noting that the fraction just above is non-negative and less than, or equal to 1, we may apply inequalities (4) and (6a

) for calculating EQ_1 and EQ_1^2 to any degree of precision. Calculating to terms of the order (1/*N*), we derive without special difficulties

$$E\,\frac{n_{11}n_{22}}{n_{11}n_{22}+n_{12}n_{21}} = \frac{p_{11}p_{22}}{p_{11}p_{22}+p_{12}p_{21}} +$$

$$(1/N)\,\frac{p_{11}p_{22}p_{12}p_{21}[p_{12}+p_{21}-p_{11}-p_{22}]}{[p_{11}p_{22}+p_{12}p_{21}]^3} + \ldots,$$

$$E\frac{n_{11}^2 n_{22}^2}{[n_{11}n_{22}+n_{12}n_{21}]^2} = \frac{p_{11}^2 p_{22}^2}{[p_{11}p_{22}+p_{12}p_{21}]^2} + (1/N)\cdot$$

$$\frac{p_{11}p_{22}p_{12}p_{21}[(p_{11}+p_{22})(p_{11}p_{22}+p_{12}p_{21})+3p_{11}p_{22}(p_{12}+p_{21}-p_{11}p_{22})]}{[p_{11}p_{22}+p_{12}p_{21}]^4} + \ldots$$

Noting that

$$1-Q^2 = \frac{4p_{11}p_{22}p_{12}p_{21}}{[p_{11}p_{22}+p_{12}p_{21}]^2},$$

$$EQ_1 = Q + (2/N)\frac{p_{11}p_{22}p_{12}p_{21}[p_{12}+p_{21}-p_{11}-p_{22}]}{[p_{11}p_{22}+p_{12}p_{21}]^3} + \ldots =$$

$$Q + [(1-Q^2)/2N]\frac{p_{12}+p_{21}-p_{11}-p_{22}}{p_{11}p_{22}+p_{12}p_{21}} + \ldots,$$

$$E(Q_1 - EQ_1)^2 = (1/N)\,[(1-Q^2)/2]^2\cdot$$

$$[(1/p_{11}) + (1/p_{22}) + (1/p_{12}) + (1/p_{21})] + \ldots$$

In the first approximation, to within terms of the order (1/*N*), $E(Q_1 - Q)^2$ is equal to the same expression (Yule 1900, p. 284; 1912, p. 593).

For mutually independent variables we have

$$E(Q_1 - EQ_1)^2 = (1/4N)[(1/p_{11}) + (1/p_{22}) + (1/p_{12}) + (1/p_{12})] + \ldots =$$

$$(1/N)\,\frac{1}{4p_1p_2\pi_1\pi_2} + \ldots$$

and, if $Q = \pm 1$, we obtain[12] $E(Q_1 - EQ_1)^2 = 0$.

3.6.B. Denoting

$$k_1 = \frac{n_{12}n_{21}}{n_{11}n_{22}},\quad k = \frac{p_{12}p_{21}}{p_{11}p_{22}}$$

and once more applying the same method of calculation, we would have obtained

$$Ek_1 = k\{1 + (1/N)\,[(1/p_{11}) + (1/p_{22})] + \ldots\}$$

and, in complete agreement with Yule (1900, p. 284; 1912, p. 593),

$E(k_1 - Ek_1)^2 = (1/N)\, k^2\, [(1/p_{11}) + (1/p_{22}) + (1/p_{12}) + (1/p_{21})] + \dots$

This result, however, would have been useless. The values that k_1 can take vary from 0 to $+\infty$ and the approximation provided by these formulas is absolutely deceptive: retaining terms of the order $(1/N)$ we neglect not only those of the order $(1/N^2)$, but also the term of order (∞/N^2). Actually, both Ek_1 and $E(k_1 - Ek_1)^2$ are infinitely large magnitudes and it is impossible to estimate the coefficient k by random empirical values of k_1. And at the same time the coefficient k_1 itself, constructed by Yule, loses any importance.

4.1. The most widely used approach to $E(u/w)$ is the approximate formula established by Pearson (1897; 1910) and since then many investigators derived it while considering the same problem. I adduce its derivation keeping to the newest explication (Czuber 1920) almost identical to Pearson's course of consideration.

Represent u/w in the form

$$\frac{Eu + u - Eu}{Ew + w - Ew} = (Eu/Ew)\,\frac{1 + [(u - Eu)/Eu]}{1 + [(w - Ew)]/Ew}$$

and expand it in powers of $[(u - Eu)/Eu]$ and $[(w - Ew)/Ew]$ assuming that the deviations $(u - Eu)$ and $(w - Ew)$ are very small with respect to Eu and Ew:

$$u/w = (Eu/Ew)\,\{1 + \frac{u - Eu}{Ew} - \frac{w - Ew}{Eu} + \frac{(w - Ew)^2}{(Ew)^2} -$$

$$\frac{(u - Eu)(w - Ew)}{EuEw} + \dots\}.$$

Restricting the expansion by the terms of the "second" order and turning to expectations we find that

$$E(u/w) = (Eu/Ew)\,\{1 + \frac{E(w - Ew)^2}{(Ew)^2} -$$

$$(1/Eu\,Ew)\,E\,(u - Eu)\,(w - Ew)\} = (Eu\,/\,Ew)\{\frac{Ew^2}{(Ew)^2} -$$

$$\frac{E(u - Eu)(w - Ew)}{EuEw}\}.$$

The Pearson approximate formula is therefore distinguished from inequality (4a) only in that $Ew^2 / (Ew)^2$, that is always greater than unity, is there included instead of 1 (Chuprov 1918 – 1919, Bd. 1, p. 240). Two essential shortcomings are inherent in the derivation of the Pearson formula just above. It ensures no idea about the degree of approximation to $E(u/w)$; even the sign of the difference between its calculated magnitude and real value remains indefinite, and it is unknown whether the former is greater, or, to the contrary, less than the latter. Then, it may be stated that the restriction concerning the deviations $(u - Eu)$ and $(w - Ew)$ is impracticable. Thus, in the examples provided by Czuber the possible deviations are not only equal to, but in some cases several times greater than the expectations.[13] Had the conditions of the derivation been fulfilled, it would have hardly been possible to apply the obtained approximate formula.

However, it is not difficult to convince ourselves that the indicated requirement is excessively severe. For the Pearson formula to provide a certain approximation to the sought magnitude $E(u/w)$, it is not necessary to restrict the deviations of the empirical values of the variables from their expectations; it is sufficient that the *expectations* of the successive powers of the deviations are small magnitudes of various orders. Given a considerable number of trials, this condition is in many cases fulfilled and then the Pearson formulas can be applied for calculating the successive terms of the expansion of $E(u/w)$ in powers of $(1/N)$. However, each time it is here necessary to make thoroughly sure that the neglected expectations do not include terms of the same order as those retained. It is also necessary to remember that, as a rule, the expectation of an even power includes terms of the same order as the expectation of the preceding odd power. Inattention to this circumstance introduced most essential errors into some recent investigations carried out by the English school for heightening the precision of previously established relations.[14]

4.2. Bearing in mind the indicated reservations, the Pearson approximate formulas may find application for calculating E(u/w). It seems however that in most cases it is more expedient to perform the calculations otherwise, by directly issuing from the expansions of the studied fraction in powers of the differences between the empirical values of the variables and their expectations. I illustrate this method by two examples.

4.2.A. Keeping to the notation of §2.2.A, denote

$$m'_{ab} - m_{ab} = dm'_{ab},\ \mu'_{ab} - \mu_{ab} = d\mu'_{ab}.$$

Since $\mu_{10} = \mu_{01} = 0$ so that

$$\mu'_{11} - \mu'_{10}\mu'_{01} = \mu_{11} + d\mu'_{11} - d\mu'_{10}\, d\mu'_{01},$$

$$\mu'_{20} - (\mu'_{10})^2 = \mu_{20} + d\mu'_{20} - (d\mu'_{10})^2,$$

$$\mu'_{02} - (\mu'_{01})^2 = \mu_{02} + d\mu'_{02} - (d\mu'_{01})^2$$

and we represent ρ in the form

$$\rho = r_{11}\,[1 + (d\mu'_{11}/\mu_{11}) - (d\mu'_{10}\, d\mu'_{01} / \mu_{11})] \cdot$$

$$\{1 + (d\mu'_{20}/\mu_{20}) - [(d\mu'_{10})^2 / \mu_{20})]\}^{-1/2} \cdot$$

$$\{1 + (d\mu'_{02}/\mu_{02}) - [(d\mu'_{01})^2 / \mu_{02})]\}^{-1/2}.$$

Expanding in powers of $d\mu'_{11}$, $d\mu'_{10}$, $d\mu'_{01}$, $d\mu'_{20}$, $d\mu'_{02}$ and then going over to expectations, we obtain, to within terms of the order $(1/N^2)$,

$$E\rho = r_{11} + (1/N)\,[(1/4)\, r_{22}\, r_{11} + (3/8)\, r_{11}\,(r_{40} + r_{04}) -$$

$$(1/2)\,(r_{31} + r_{13})] + (1/N^2)\{(1/4)\, r_{33} - (5/16)\, r_{11}\,(r_{60} + r_{06}) -$$

$$(3/16)\, r_{11}\,(r_{42} + r_{24}) + (3/8)\,(r_{51} + r_{15}) - (1/4)\, r_{22}\, r_{11} +$$

$$(1/2)\,(r_{31} + r_{13}) + (15/32)\, r_{11}\, r_{22}\,(r_{40} + r_{04}) - (3/8)\, r_{11}\,(r_{40} + r_{04}) +$$

$$(9/64)\, r_{11}\, r_{40}\, r_{04} + (9/32)\, r_{22}^2\, r_{11} - (15/16)\,(r_{31}\, r_{40} + r_{13}\, r_{04}) -$$

$$(3/16)\,(r_{31}\, r_{04} + r_{13}\, r_{40}) - (3/8)\, r_{22}\,(r_{31} + r_{13}) +$$

$$(105/128)\, r_{11}\,(r_{40}^2 + r_{04}^2) - (1/4)\, r_{30}\, r_{03} + (15/8)\, r_{11}\,(r_{30}^2 + r_{03}^2) -$$

$$(9/4)\,(r_{21}\, r_{30} + r_{12}\, r_{03}) - (5/4)\, r_{21}\, r_{12} + (3/8)\, r_{11}\,(r_{21}^2 + r_{12}^2) +$$

$$(3/4)\, r_{11}\,(r_{21}\, r_{03} + r_{12}\, r_{30}) - (1/2)\, r_{11} + (1/2)\, r_{11}^3\} + \ldots$$

Depending on the law of connection between the variables, the difference between $E\rho$ and r_{11} can thus be positive, negative, or equal to zero. If $r_{11} = 0$, but the variables are not mutually independent (cf. §1.2, Note 2 and §2.2.A, Note 6), we find that

$$E\rho = -(1/2N)\,(r_{31} + r_{13}) + (1/N^2)\,[(1/4)\, r_{33} + (3/8)\,(r_{51} + r_{15}) +$$

$$(1/2)\,(r_{31} + r_{13}) - (15/16)\,(r_{31}\, r_{40} + r_{13}\, r_{04}) - (3/16)\,(r_{31}\, r_{04} + r_{13}\, r_{40}) -$$

$$(3/8)\, r_{22}\,(r_{31} + r_{13}) - (1/4)\, r_{30}\, r_{03} - (9/4)\,(r_{21}\, r_{30} + r_{12}\, r_{03}) -$$

$$(5/4)\, r_{21}\, r_{12}] + \ldots$$

In case of "normal" correlation we have

$$E\rho = r_{11}\{1 - \frac{1-r_{11}^2}{2N} - \frac{3(1-r_{11}^2)}{8N^2}[1 + 3r_{11}^2] + \ldots$$ [15]

4.2.B. When calculating squared errors, representatives of the English school readily apply the Pearson (Pearson & Filon 1898, p. 245) method of preliminary taking the logarithms of the considered expressions. Indeed, this method can sometimes facilitate calculations, but its application demands extreme caution. Representing ρ/r in the form

$$1 + [(\rho - r)/r] = [1 + (d\mu'_{11}/\mu_{11}) - (d\mu'_{10}\, d\mu'_{01})/\mu_{11})] / \sqrt{D},$$

$$D = \{1 + (d\mu'_{20}/\mu_{20}) - [(d\mu'_{10})^2/\mu_{20}]\}\{1 + (d\mu'_{02}/\mu_{02}) - [(d\mu'_{01})^2/\mu_{02}]\},$$

take the logarithms of both parts of this equality and apply to them the expansion of $\log(1 + x)$ in powers of x:

$$\log\{1 + [(\rho - r)/r]\} = \log[1 + (d\mu'_{11}/\mu_{11}) - (d\mu'_{10}\, d\mu'_{01})/\mu_{11})] -$$

$$(1/2)\log\{1 + (d\mu'_{20}/\mu_{20}) - [(d\mu'_{10})^2/\mu_{20}]\} -$$

$$(1/2)\log\{1 + (d\mu'_{02}/\mu_{02}) - [(d\mu'_{01})^2/\mu_{02}]\},$$

$$[(\rho - r)/r] - (1/2)[(\rho - r)/r]^2 + \ldots = [(d\mu'_{11}/\mu_{11}) - (d\mu'_{10}\, d\mu'_{01})/\mu_{11})] -$$

$$(1/2)[(d\mu'_{11}/\mu_{11}) - (d\mu'_{10}\, d\mu'_{01})/\mu_{11})]^2 + \ldots -$$

$$(1/2)\{(d\mu'_{20}/\mu_{20}) - [(d\mu'_{10})^2/\mu_{20}]\} +$$

$$(1/4)\{(d\mu'_{20}/\mu_{20}) - [(d\mu'_{10})^2/\mu_{20}]\}^2 - \ldots -$$

$$(1/2)\{(d\mu'_{02}/\mu_{02}) - [(d\mu'_{01})^2/\mu_{02}]\} + (1/4)\{(d\mu'_{02}/\mu_{02}) -$$

$$[(d\mu'_{01})^2/\mu_{02}]\}^2 - \ldots$$

Squaring, we obtain

$$[(\rho - r)/r]^2 - [(\rho - r)/r]^3 + \ldots = [(d\mu'_{11})^2/\mu_{11}^2] +$$

$$(1/4)[(d\mu'_{20})^2/\mu_{20}^2] + (1/4)[(d\mu'_{02})^2/\mu_{02}^2] - \frac{d\mu'_{11}d\mu'_{20}}{\mu_{11}\mu_{20}} -$$

$$\frac{d\mu'_{11}d\mu'_{02}}{\mu_{11}\mu_{02}} + (1/2)\frac{d\mu'_{20}d\mu'_{02}}{\mu_{20}\mu_{02}} + \ldots$$

Going over to expectations and assuming that $E[(\rho - r)/r]^3$, $E[(\rho - r)/r]^4$, … do not contain terms of the order $(1/N)$, we obtain in the first approximation (cf. Sheppard 1898, p. 128; Pearson 1905, p. 20)

$$E(\rho - r)^2 = (1/N)\{r_{22}[1 + (1/2)r_{11}^2] - r_{11}(r_{31} + r_{13}] +$$

$$(1/4)r_{11}^2(r_{40} + r_{40})\} + \ldots$$

In the first approximation, to within terms of the order $(1/N)$ inclusively, $E(\rho - E\rho)^2$ is equal to the same expression.

We could have arrived at an identical result even without squaring but directly going over from the expansion of $\log\{1 + [(\rho - r)/r]\}$ to expectations and then substituting the value of $E(\rho - r)$ obtained above into

$$E\,[(\rho - r)/r] - (1/2)\,E\,[(\rho - r)/r]^2 = -\frac{E d\mu'_{10} d\mu'_{01}}{\mu_{11}} - (1/2)\frac{E(d\mu'_{11})^2}{\mu_{20}} +$$

$$(1/2)\frac{E(d\mu'_{10})^2}{\mu_{20}} + (1/4)\frac{E(d\mu'_{20})^2}{\mu_{20}^2} + (1/2)\frac{E(d\mu'_{01})^2}{\mu_{02}} + (1/4)\frac{E(d\mu'_{02})^2}{\mu_{02}^2}.$$

The calculated value of $E(\rho - E\rho)^2$ is correct in the first approximation because the expectations of the neglected summands in either side of the expansion do not indeed include terms of the order $(1/N)$. However, if the same method be applied to calculate $E(\rho - E\rho)^3$, we would have obtained a wrong result already in the first approximation and even when taking into account not only the third, but the fourth degrees of the differences. The reason is, that, first, the terms of the order $(1/N^2)$ in $E(\rho - E\rho)^3$ do not coincide with the terms of the same order in $E(\rho - r)^3$ and, second, $E(\rho - r)^4$ also contains terms of the order $(1/N^2)$ as does $E(\rho - r)^3$. Because of like causes an attempt to calculate $E(\rho - E\rho)^2$ more precisely than to within $(1/N)$ by the same method would have also led to mistaken results. When solving such problems it is more expedient to follow the pattern applied above for calculating $E\rho$ without taking logarithms.

5. As we saw (§2), in a number of cases the expectation of the ratio of two variables is exactly equal to the ratio of their expectations,

$$E\,(u/w) = Eu\,/\,Ew. \qquad (8)$$

The inequality (4a) shows that a necessary although not at all sufficient condition for that is that $E(u - Eu)(w - Ew)$ or r_{11} be positive. If, however, $r_{11} \le 0$, then $E(u/w) > Eu\,/\,Ew$.

It would be very interesting to establish precisely the conditions to be imposed on the law of connection between variables u and w for the equality (8) to hold and to obtain reliable and practical criteria for distinguishing in each separate case, in which $r_{11} > 0$, whether the expectation of the ratio is equal to, or greater, or less than the ratio of the expectations of the numerator and the denominator. In the limit, when the number of trials is infinitely great, this problem can be easily solved in many cases by means of inequalities (5) and (6) which at $N = \infty$ infrequently close in on the value taken by the ratio Eu/Ew when the number of trials is infinitely great. For the case of a finite number of trials it is not difficult to convince ourselves that

$$E\,(u/w) = Eu\,/\,Ew = 1 \text{ if } Euw^{t-1} = Ew^t \text{ at } t = 1, 2, 3, \ldots$$

However, I have yet been unable to find out whether this condition is not only sufficient but also necessary.[16]

Basing oneself on this and similar relations [17] it is sometimes possible to go ahead and solve successively the problem. And it is certainly unnecessary to calculate such expressions as Euw^{t-1} and Ew^t to the very end; it is sufficient to reduce them to a form in which their identity becomes clear. This consideration essentially simplifies the work. My initial proof that $EQ^2 = 1$, based on the derivation of general formulas for these two magnitudes and for any integer positive t, demanded extremely involved calculations whereas a proof in accord with the above remark only requires a few lines.

This method of estimating the expectation of the ratio of two random variables leads us to the formulation of a more general problem about the expectation of a function of two or more random variables given the relations between the expectations of the variables themselves or of other of their functions. I shall venture to return to this curious problem of essential importance for a rational justification of statistical methods in one of the subsequent volumes of these *Proceedings.*

Notes

1. When solving some problems it seems very convenient to determine the laws of distribution and connection by means of *semi-invariants* {cumulants}(Thiele 1889; 1903). Thiele's immediate aim was to construct this system of parameters for studying laws of distribution of one variable, but his *semi-invariants* can be successfully adapted also for investigating laws of connection between variables, see for example Hausdorff's contribution (1901, p. 177) who uses the same system of magnitudes calling them *canonical parameters.*

2. The concepts of connection and independence can be introduced in different ways. For example, it is possible to base the definitions on the magnitude of the difference

$(m_{11} - m_{10} m_{01})$ (*)

and to recognize a mutual connection between the variables if this difference does not vanish, and to call them mutually independent otherwise. The Pearson notions of correlativity and mutual independence (Pearson 1896, pp. 256 – 257; Pearson & Heron 1913, pp. 164 – 165) issue from considering the expectation of one of the variables at various values of the other one: if the expectation of w persists at any value taken by u, there is no connection between u and w; otherwise, they are mutually correlated:

Two organs […] *are said to be correlated, when a series of the first organ of a definite size being selected, the mean of the sizes of the corresponding second organ is found to be a function of the size of the selected first organ. If the mean is independent of this size, the organs are said to be non-correlated.*

Such special notions of connection and independence can possess certain scientific importance but they demand much caution so that the conclusions adjoining one of them will not be transferred to the other ones without check. My definition of independence seems to be rigorous to the utmost: variables independent in my sense remain independent from the point of view of all the other definitions. On the contrary, variables can be not correlated in the Pearson sense but, at the same time, be mutually dependent from the viewpoint of a more strict definition. In a similar way the vanishing of (*) does not at all ensure mutual independence in the sense adopted by me. The Pearson definitions of correlativity and lack of correlation are also deficient in that lack of correlation between u and w does not ensure the same between w and u.

3. See Pearson (1916, vol. 11, p. 240):

Thus to a high order of approximation at least the mean of the array means [my E λ (j; 1) – A.C.] *is the mean of the corresponding array in the sampled population* [η (j; 1) – A.C.] […] *This result cannot be taken as obvious, as the size of the array in the sample varies.*

Also see Pearson (Ibidem, vol. 12, p. 267) where this result is described as *absolutely* precise rather than precise *merely to a high order of approximation.*

4. Cf. Pearson (1916, p. 243) where the relation $r = 0$ is offered as valid *to a high order of approximation at least*.

5. See Chuprov (1916); cf. Chuprov (1918 – 1919, Bd. 1) where I indicated a fourth version of the coefficient of dispersion for series containing differing numbers of trials. The fourth version concludes the cycle of the main varieties of this coefficient.

6. If $r = 0$ but α and β are not mutually independent (see §1.2, Note 2) we cannot come to the definition of Eρ in the way I made use of because in such cases we have no right to assume that

$$\mathrm{E}\frac{[\alpha_f - \alpha_{(N)}][\beta_f - \beta_{(N)}]}{\sqrt{\sum[1;2]\sum[2;2]}} = \mathrm{E}\frac{[\alpha_f - \alpha_{(N)}]}{\sqrt{\sum[1;2]}}\ \mathrm{E}\frac{[\beta_f - \beta_{(N)}]}{\sqrt{\sum[2;2]}}.$$

For the approximate value of Eρ in this case see §4.2.A.

7. Cf. Student (1908): issuing from the empirical data of two skillfully arranged experiments, he put forward a hypothesis that the distribution of ρ, when the variables are not connected and distributed in accord with the Gauss law, was

$$y = y_o (1 - x^2)^{(N-4)/2}$$

that led to the squared error being $\sqrt{1/(N-1)}$. Student, however, hesitated and, as a confirmation, he indicated that his result was sufficiently close to the usual formula for the squared error of the correlation coefficient, $(1 - r)/\sqrt{N}$ at $r = 0$ and a large number of trials (*approximates sufficiently … when* $r = 0$ *and* N *is large*, see his p. 308). Later, Fisher (1915, p. 508) discovered that Student's conjecture about the distribution of ρ was completely correct.

8. If $\beta = \alpha$, Euw = Ew^2 becomes

$$[1/(N-1)^4]\,\mathrm{E}\{\sum_{f=1}^{N}[\alpha_f-\alpha_{(N)}]^2\}^4$$

and we arrive at formula (21) on p. 193 of my paper (Chuprov 1918 – 1919 and 1921, vol. 12). When α and β are mutually independent, $\mathrm{E}w^2$ is reduced to

$$[1/(N-1)^4]\{\mathrm{E}\,(\Sigma[1;2])^2\}^2,$$

cf. formula (19) on p. 192, Ibidem.

9. I indicate only those particular formulas that are necessary for the sequel. Their derivation does not entail special difficulties and can be left out. Neither do I reproduce general formulas obtained by me. Their derivation would have diverted the explication too far from the main subject of the present paper.

10. Cf. Pearson (1915). There, the magnitude φ_t^2 is recognized as a *true value* and the calculation of φ_a^2 is considered as an approach to φ_t^2 from the empirical data (*we use* φ_a^2 *as an approximation to* φ_t^2, see his p. 571). In Young & Pearson (1916, p. 220) φ^2 is however called "the true mean square contingency" (their p. 220).

11. Representing $\mathrm{E}\,[\varphi^2_a - \mathrm{E}\,\varphi^2_a]^2$ in the form

$$(1/N)\{4\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{(p_{ij}-p_i\pi_j)^3}{p_i^2\pi_j^2} - 3\sum_{i=1}^{k}(1/p_i)[\sum_{j=1}^{l}\frac{p_{ij}^2}{p_i\pi_j}]^2 -$$

$$3\sum_{j=1}^{l}(1/\pi_j)[\sum_{i=1}^{k}\frac{p_{ij}^2}{p_i\pi_j}]^2 + 2\sum_{i=1}^{k}\sum_{j=1}^{l}[\sum_{f=1}^{k}\frac{p_{fj}^2}{p_f\pi_j}][\sum_{h=1}^{l}\frac{p_{ih}^2}{p_i\pi_h}] +$$

$$2\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p_{ij}-p_i\pi_j}{p_i\pi_j}[\sum_{f=1}^{k}\frac{p_{fj}^2}{p_f\pi_j}][\sum_{h=1}^{l}\frac{p_{ih}^2}{p_i\pi_h}] + 12\varphi^2 + 4\} + \ldots$$

and assuming, in accord with a known approximate formula, that

$$\mathrm{E}\,[\varphi_a - \mathrm{E}\,\varphi_a]^2 = (1/4\varphi^2)\,\mathrm{E}\,[\varphi^2_a - \mathrm{E}\,\varphi^2_a]^2,$$

we arrive at the formula of Pearson & Blakeman (1906, p. 194). The only difference is that, following an English custom, prior magnitudes p are there replaced by empirical magnitudes n.

12. Yule (1912, p. 593) puts forward a correct in itself idea that the squared error of the coefficient of association vanishes if one of the probabilities p_{11}, p_{22}, p_{12}, p_{21} is zero. However, he does not distinguish sufficiently clearly between prior and empirical values of the considered magnitudes and formulates his idea in an incongruous form: the error "vanishes if anyone {?} of the frequencies n_{11}, n_{22}, n_{12}, n_{22} becomes zero. Hence the standard error of the coefficient is zero, when the coefficient [he means Q_1 – A.C.] itself is ± 1". Cf. Pearson & Heron (1913, Appendix 1).

13. One of Czuber's examples seems to be unfortunate in an especially instructive way. He considers a fraction whose denominator but not numerator can vanish. The expectation and the squared error of the coefficient that he considers are infinitely large and it is absolutely inadmissible to apply the Pearson formula for their approximate estimation. In this case, it would have been very easy to avoid the complication by transposing the numerator and denominator, cf. Czuber, §4, Ibidem.

14. Cf. Pearson's Editorial (1919, pp. 259, 261, 266 – 267) caused by my remarks.

15. Cf. Soper (1913). He calculates $\mathrm{E}\rho$ in the same way as I did when deriving the general formula above, but he at once facilitates his work by assuming a "normal" law of connection. Cf. Soper et al (1917, p. 336).

16. In actual fact, this condition is obeyed not only in the case of the coefficient of dispersion but in all the other instances in which I was able to find out that $\mathrm{E}\,(u/w) = \mathrm{E}u / \mathrm{E}w = 1$. It is not difficult to confirm this for the ratio n_{ij} / υ_j (§2.1.A) by issuing from the general formulas (2) and (5) (Chuprov 1918 – 1919 and 1921, vol. 12, pp. 195 and 196). Taking into account that the $\alpha_{r,\,h}$ included there satisfy the relation

$$\sum_{k=j}^{m+1-i} C_m^{\,k}\,\alpha_{m-k+1,\,i}\,\alpha_{k,\,j} = C_{i+j-1}^{\,i-1}\,\alpha_{m+1,\,i+j}$$

we derive without special difficulties that

$$[\mathrm{E}\,(n_{ij}\,\upsilon_j^t)\,/\,p_{ij}] = [\mathrm{E}\upsilon_j^{t+1}/\pi_j] = \sum_{k=0}^{t} N^{[k+1]}\,\alpha_{t+1,\,k+1}\,\pi_j^k$$

where

$$N^{[k+1]} = N(N-1)(N-2)\ldots(N-k).$$

In a similar way, applying formula (6) from p. 196, we convince ourselves that

$$\frac{En_{ij}n_{fg}\nu_j^t\nu_g^t}{p_{ij}p_{fg}} = \frac{E\nu_j^{t+1}\nu_g^{t+1}}{\pi_j\pi_g},$$

$$\frac{En_{ij}n_{fg}n_{hl}\nu_j^t\nu_g^t\nu_l^t}{p_{ij}p_{fg}p_{hl}} = \frac{E\nu_j^{t+1}\nu_g^{t+1}\nu_l^{t+1}}{\pi_j\pi_g\pi_l} \quad \text{etc}$$

and that, therefore, not only

$$\mathrm{E}\frac{n_{ij}n_{fg}}{\nu_j\nu_g} = \frac{p_{ij}p_{fg}}{\pi_j\pi_g}$$

but also

$$\mathrm{E}\,\frac{n_{ij}n_{fg}n_{hl}}{\nu_j\nu_g\nu_l} = \frac{p_{ij}p_{fg}p_{hl}}{\pi_j\pi_g\pi_l}, \text{ etc.}$$

17. Practically interesting from among such relations can be the generalized and easily derived form of the theorem indicated above: if for $k > 0$ and $i = 1, 2, 3, \ldots$

$$\mathrm{E}uw^{t-1} = k^t\mathrm{E}w^t \pm \sum_{i=1}^{h} k_i^2\,\mathrm{E}w^{t+i} \pm \sum_{j=1}^{h} k_j^2\,\mathrm{E}w^{t-j}$$

then

$$\mathrm{E}\,(u/w) = 1 \pm \sum_{i=1}^{h} k_i^2\,\mathrm{E}w^{i} \pm \sum_{j=1}^{h} k_j^2\,\mathrm{E}w^{-j}.$$

On the other hand, if $\mathrm{E}\,uw^{t-1} = k\,\mathrm{E}w^t$ at $t = 1, 2, 3, \ldots$ then $\mathrm{E}\,(u/w) = k = \mathrm{E}u/\mathrm{E}w$.

References

Bohlmann, G. (1913), Formulierung und Begründung zweier Hilfsätze der mathematischen Statistik. *Math. Annalen*, Bd. 74, pp. 341 – 409.

Chuprov [Tschuprow], A.A. (1916), On the expectation of the coefficient of dispersion. Translated in this book.

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

--- (1918 – 1919 and 1921), On the mathematical expectation of the moments of frequency distributions. *Biometrika*, vols. 12 – 13, pp. 140 – 169, 185 – 210 and 283 – 295 respectively.

Czuber, E. (1920), Über Funktionen von Variablen zwischen welchen Korrelationen bestehen. *Metron*, Bd. 1, pp. 53 – 61.

Fisher, R.A. (1915), Frequency distribution of the values of the correlation coefficient etc. *Biometrika*, vol. 10, pp. 507 – 521.

Hausdorff, F. (1901), Beiträge zur Wahrscheinlichkeitsrechnung. *Ber. Verh. Kgl. Säachs. Ges. Wiss. Leipzig*, Bd. 53, pp. 152 – 178.

Markov, A.A. (1916), On the coefficient of dispersion. *Избранные труды* (Sel. Works). N.p., 1951, pp. 523 – 535. Transl.: DHS, 2514, 1998, pp. 195 – 204.

Pearson, K. (1896), Regression, heredity and panmixia. *Phil. Trans. Roy. Soc.*, vol. A187, pp. 253 – 318.

--- (1897), On a form of spurious correlation, etc. *Proc. Roy. Soc.*, vol. 60, pp. 489 – 498.

--- (1905), *On the General Theory of Skew Correlation*. Draper's Co. Res. Mem., Biometric ser. 2.
--- (1910), On the constants of index-distributions, etc. *Biometrika*, vol. 7, pp. 531 – 541.
--- (1915), On the probable error of a coefficient of mean square contingency. *Biometrika*, vol. 10, pp. 570 – 573.
--- (1916), On the application of "goodness of fit" tables, etc. *Biometrika*, vols 11 – 12, pp. 239 – 261 and 259 – 281 respectively.
--- (1919), Peccavimus! *Biometrika*, vol. 12, pp. 259 – 281. Unsigned Editorial.
Pearson, K., Blakeman J. (1906), On the probable error of mean-square contingency. *Biometrika*, vol. 5, 191 – 197.
Pearson, K., Heron, D. (1913), On the theories of association. *Biometrika*, vol. 9, pp. 159 – 315.
Pearson, K., Filon, L.N.G. (1898), On the probable error of frequency constants, etc. *Phil. Trans. Roy. Soc.*, vol. A191, pp. 229 – 311.
Sheppard, W.F. (1898), On the application of the theory of error to cases of normal distribution and normal correlation. *Phil. Trans. Roy. Soc.*, vol. A192, pp. 101 – 167.
Soper, H.E. (1913), On the probable error of the correlation coefficient to a second approximation. *Biometrika*, vol. 9, pp. 91 – 115.
Soper, H.E., Young, A.W., Cave, B.M., Lee, A., Pearson, K. (1917), On the distribution of the correlation coefficient in small samples. *Biometrika*, vol. 11, pp. 328 – 413.
Student (1908), The probable error of the correlation coefficient. *Biometrika*, vol. 6, pp. 302 – 310.
Thiele, T.N. (1889), *Foreläsninger over almindelig jagttagelsesläre*. Kjöbenhavn.
--- (1903), *Theory of Observations*. London.
Young, A.W., Pearson, K. (1916), On the probable error of a coefficient of contingency without approximation. *Biometrika*, vols. 11 – 12, pp. 215 – 230 and 259 – 260 respectively.
Yule, G.U. (1900), On the association of attributes in statistics, etc. *Phil. Trans. Roy. Soc.*, vol. A194, pp. 257 – 319.
--- (1912), On the methods of measuring association between two attributes. *J. Roy. Stat. Soc.*, vol. 75, pp. 579 – 642.

Part 2

10. A.L. Dmitriev, O. Sheynin. Chuprov: Additional Bibliography

Prepared for this collection

Below, we are listing Chuprov's publications lacking in the book Sheynin, O. *Chuprov*: *Life, Work, Correspondence*. Göttingen, 1996, as well as their new translations and sources concerning him. Materials included in Part 1 of this collection are not entered here.

Abbreviations: Izv. PPI = *Izvestia Peterburgsk. Polytekhnich. Inst.*
NST = *Nordisk Statistisk Tidskrift*

Articles

1. An advice (a poem, written 1893). In Eliseeva et al (1996, p. 62). (R)
2. [On a general European census]. *Bull. Intern. Stat. Inst.*, t. 11, No. 1, 1899, pp. 68 – 74 of first paging, this being a report at the sixth session ogf the Institute (1897, in French). Chuprov (p. 69) advocated a simultaneous census *au moins une fois … dans tous les Etats européens*.
3. On the balance of land for the agrarian reform. *Pravo*, 29, 1906. (R)
4. Resettlements and the agrarian issue. Reprint, 1907, from an unknown source, 19pp. (R)
5. Power for the people and the land. *Poliarnaia Zvezda* weekly, No. 11, 1906, pp. 766 – 778. (R)

Reviews
6. Dmitriev, V.K. *Экономические очерки* (Economic Essays). M., 1904. *Izv. PPI*, vol. 1, No. 3 – 4, 1904, pp. 284 – 287.
7. *Свод отчетов фабричных инспекторов за 1902г.* (Summary of the Accounts of Factory Inspectors for 1902). Psb, 1904. *Izv. PPI*, vol. 2, No. 1 – 2, 1904, pp. 79 – 87.

8. A. Mitscherlich, Die Schwankungen der landwirtschaftlichen Reinerträge berechnet für einige Fruchtfolgen mit Hilfe der Fehlerwahrscheinlichkeits- rechnung. *Z. f.d. ges. Staatswiss.*, Ergzbd. 8. Ibidem, pp. 75 – 79.
9. *Состав служащих в промышленных заведениях в отношении подданства, языка и образовательного ценза* (The Composition of the Employees of Industrial Enterprises with respect to Citizenship, Language and Educational Qualification). Psb, 1904. *Izv. PPI*, vol. 2, No. 3 – 4, 1904, pp. 235 – 238.
10. *Материалы по статистике движения землевладения в России*, вып. 8. (Materials on the Statistics of the Movement of Landownership in Russia). Psb, 1904. Ibidem, pp. 238 – 253.
11. Bowley, A.L. *Statistical Studies relating to National Progress in Wealth and Trade since 1882*. London, 1904. *Izv. PPI*, vol. 3, No. 1 – 2, 1905, pp. 133 – 138.
12. *Ежегодник России* (Yearbook of Russia). Psb, 1906. *Izv. PPI*, vol. 7, No. 1 – 2, 1907, pp. 251 – 256.
13. *Presussische Statistik. Die ländliche Verschuldung in Preussen*. Berlin, 1905 – 1906. Ibidem, pp. 256 – 262.
14. Kaufman, A.A. *Русская община в процессе ее зарождения и роста* (The Russian Land Community in the Course of Its Origin and Development). M., 1908.
15. Jahn, G. *Statistikkens teknik og metode.* Kristiania, 1920. *Deutsches stat. Zentralbl.*, Jg. 13, No. 9 – 10, 1921, pp. 148 – 150.
16. Zur Bedeutung der Mathematik für die Statistik. [Review of March, L. La méthode statistique. *Metron*, No. 1, 1921]. Ibidem, pp. 150 – 151.
17. Charlier, C.V.L. *Vorlesungen über die Grundzüge der mathematischen Statistik.* Lund, 1920. Ibidem, Jg. 14, No.1 – 2, 1922, pp. 22 – 23.
18. Gini, C. Report on the Problem of Raw Materials and Foodstuffs. Gini, C. L'enquéte de la Société des Nations. *Metron*, t. 2, No. 1 – 2, 1922. Mortara, G. *Prospettive economiche*. Citta di Castello, 1922. (Reviews of three sources.) *Krestianskaia Rossia,* No. 2 – 3, 1923, pp. 236 – 240.
19. Volkov, E.Z. *Аграрно.экономическая статистика России* и т.д. (The Agrarian-Economic Statistics of Russia As Reflected in the Results of Its Scientific and Methodological Advances, the Zemstvo Experience and in the Practice of the Last Four Years of the Great Revolutionary Change). *Krestianskaia Rossia*, No. 4, 1923, pp. 193 – 196.
20. Niceforo, A. *Les indices numériques de la civilisation et du progrés*. Paris, 1921. *Arch. f. Sozialwiss. u. Sozialpolitik*, Bd. 50, No. 1, 1922, pp. 260 – 262.
21. Simiand, Fr. *Statistique et expérience*. Paris, 1922. Ibidem, No. 2, pp. 538 – 540.
22. Westnik Statistici 1920 – 1922. NST, Bd. 1, No. 2, 1922, pp. 353 – 360.
23. Zizek, F. *Fünf Hauptprobleme der statistischen Methodenlehre*. München – Leipzig, 1922. NST, Bd. 2, No. 1, 1923, pp. 164 – 167.
24. Mortara, G. *Lezioni di statistica economica e demografica*. Roma, 1920.
Ibidem, pp. 167 – 170.
25. Czuber, E. *Die statistischen Forschungsmethoden*. Wien, 1921. JNÖS, 3. Folge, Bd. 63, No. 4, 1923, pp. 378 – 379.
26. Winkler, W. *Die statistische Verhältniszahlen.* Wien, 1923. *Deutsches stat. Zentralbl.*, Jg. 15, No. 3 – 4, 1923, pp. 57 – 58.
27. Mortara, G. *Prospettive economiche*. Citta di Castello, 1923. Ibidem, pp. 58 – 59.
28. Soper, H.E. *Frequency Arrays*. Cambridge, 1922. NST, Bd. 3, No. 2 – 3, 1924, pp. 414 – 417.
29. Czuber, E. *Mathematische Bevölkerungstheorie*. Leipzig – Berlin, 1923. JNÖS, 3. Folge, Bd. 68, 1925, pp. 130 – 131.
30. Zinserling, D. *Практическое руководство статистики* (Practical Guide to Statistics). L., 1925. *Russk. Ekonomich. Sbornik* (Prague), No. 2, 1925, pp. 175 – 178.
31. Woytinsky, W.S. *Весь мир в цифрах* (The Whole World in Numbers), book 1. Berlin, 1925. Ibidem, No. 4, 1925, pp. 194 – 197.

Translations
(1912) Sampling. JNÖS, Bd. 216, 1997, pp. 658 – 671. Transl. from Russian into English.
(1919). *La décomposition du bolchevisme. Voprosy Istorii*, No. 10, 2003, pp. 6 – 18. Transl. from French into Russian with foreword byA.L. Dmitriev et al. A fragment of the foreword is here translated.
(1922), Westnik Statistiki 1920 – 1922. *Voprosy Statistiki*, No. 1, 1999, pp. 11 – 13. Transl. from German into Russian.

Reprints

(1917), On the plan of reorganizing the statistical establishment in the Empire. Reprinted in orig. Russian: *Voprosy Statistiki*, No. 2, 1995, pp. 43 – 50.

(1919), *La decomposition du bolchevisme*. Reprinted in this collection.

(1922), The problem of indeterminism in the light of statistical physics. Reprinted in orig. Russian in Eliseeva et al (1996, pp. 47 – 55).

Sources concerning Chuprov

Obituaries: Anziferov, A.N. Russian Newspaper *Vozrozhdenie* (Paris), 23 April 1926.

Amfiteatrov, A.V, Ibidem.

F[ediaevsky], G.K. Ibidem, 31 May 1926.

Den, V.E. Newspaper *Krasnaia Gazeta*, 23 April 1926, Evening Issue (R)

Seneta, E. (1985), Sketch of the history of survey sampling in Russia. *J. Roy. Stat. Soc*., vol. A148, pp. 118 – 125.

Eliseeva, I.I., Dmitriev, A.L., Storchevoi, M.A., Editors (1996), *А.А. Чупров. Материалы конференции к 70-летию со дня кончины* (Chuprov. Materials on the Occasion of the 70 Years after His death). Coll. articles. Psb.

Seneta, E. (1997), Chuprov. In Johnson, N.L., Kotz, S., Editors, *Leading Personalities in Statistical Sciences*. New York, pp. 185 – 187.

--- (2001), Chuprov. In Heyde, C.C., Seneta, E., Editors, *Statisticians of the Centuries*. New York, pp. 303 – 307.

11. P. Struve, A. Lappo-Danilevsky, M. Diakonov. Note on Professor Chuprov's Scientific Work (1917),

Reprinted in Chuprov, A.A. *Материалы конференции к 70-летию со дня кончины* (Materials of Conf. Observing 70 years after His Death). Petersburg, 1996, pp. 56 – 59 …

Aleksandr Aleksandrovich Chuprov, the son of the well-known economist A.I. Chuprov, who for many years honorably headed the chair of political economy and statistics at Moscow University, was born on 6 February 1874 in Mosalsk {Kaluga province}; in 1892 graduated with a gold medal from the Moscow gymnasium No. 5, and, in 1896, from the Mathematical Department, Physico-Mathematical Faculty, of Moscow University. He then went to Germany and learned at Berlin and Strasbourg universities. In 1901, in Strasbourg, became Doctor of *Staatswissenschaft* for a portion of a work published next year in full in the proceedings of Knapp's seminar and entitled *Die Feldgemeinschaft. Eine morphologische Untersuchung*. In spring of 1902 had passed his Master examination in political economy at the Law Faculty of Moscow University and, in autumn of the same year, started reading statistics at the Peter the Great Petrograd Polytechnical Institute as a staff Dozent. In 1908 submitted a composition *Essays on the Theory of Statistics* to Moscow University as a Master dissertation on political economy and statistics for which was at once rewarded the degree of Doctor. After that the Economic Department and the Council of the Petrograd Polytechnical Institute elected him ordinary professor of statistics.

Chuprov's scientific work that made him well-known not only here, but also abroad, belongs to two fields: to political economy and theory of statistics. In the former, Chuprov exclusively studied agrarian problems, and, in particular and especially, community landownership. His German contribution devoted to this subject is interesting and valuable as an investigation of the morphology of the land community. Without providing new conclusions bearing on the essence of the pertinent economic problem, A.A. very subtly analyzed the legal substance of a community, and, after establishing the concept and the varieties of a community, described how were the principles of community landownership realized under different technological conditions. One of his remarks (p. 194) might characterize the formal method of his analysis:

> *The formal organizational principle, that we call land-communal, can be applied to most diverse cases, for example in constitutional law when systematizing associations of the type of an allied state. On the other hand* (*there is but one step from sublime to amusing*), *let us imagine as an example the totality of waiters in a café, or of hairdressers in their saloon. Definite tables are appointed to each waiter. If unequally profitable, they might be periodically re-distributed* {for example} *taking into account the waiters'*

ages as {a similar procedure} *is done in the Almende* {German term meaning land in common ownership}. *To some extent individual usage is sometimes done away with when tips are collected in a common money-box and then re-distributed etc*. […]

Chuprov himself does not extend his study so widely, he restricts it by the phenomena of land-usage and landownership. However, appearing indeed as a representative of a formal analysis, he thus makes an original contribution to the literature on communal landownership which is therefore the more valuable since he reviews and treats very heterogeneous material on community regulations not only in Russia but in all countries where a land community can be established in one or another form; in particular, he studied the material pertaining to communities in the colonies, – in India and on the island of Java.

The formal legal schooling that he accomplished as a Strasbourg Doctor when investigating the constructions of such a master as Laband and the philosophical atmosphere of those two German universities where he studied and where philosophy was represented by such subtle thinkers as Windelband and Simmel, undoubtedly helped Chuprov to some extent in his work on the problems of economic morphology. However, this work was only an episode in Chuprov's scientific studies.

Its main field, in which he was especially prepared as having a serious mathematical background has been statistics. At present, it is impossible to develop the theory of statistics without mastering its mathematical premises and methods. As a statistician, Chuprov happily combines in himself a mathematical basis, an intense philosophical interest in clearly formulating the issues of the theory of statistics, and a serious economic erudition acquired by studying problems of political economy. His *Essays on the Theory of Statistics* have therefore indeed become a contribution that consists of several independent studies but at the same time is a generalized exposition of those main issues which were put forward by the entire recent development of the theory of statistics. The author's literary mastery, that he inherited from his father, increased the value of that book as a summarizing contribution already now educating the scientific way of reasoning of Russian statisticians.

The statistical thinking and teaching in Russia are now on a very high level attained abroad only in England and perhaps Italy, and to a large extent this is due to Chuprov as a scientist and teacher. The teaching of statistics at the Petrograd Polytechnical Institute gives very much both to the ordinary student and to those few persons who are distinguished as scientific workers in statistics. They are laying the foundation of a Russian statistical school.

All the above induces us to nominate […] Chuprov for Corresponding Membership {class of historical and political sciences} of the Russian Academy of Sciences.

12. Anonymous.
[A Document from Chuprov's Personal Records]

In June 1917 Prof. Chuprov went abroad to Stockholm for the summer break. He needed to avail himself of the materials at the main statistical bureau there for his vast work on birth-rate and mortality under war conditions. He began this work in Stockholm already a year ago; and he published his preliminary results as a separate paper appended here. In addition, he thought of continuing another study on the application of the so-called sampling in statistical investigations.[1]

As it followed from his letters, he was determined to return to Petrograd by the beginning of the autumn academic half-year, i.e., by September. However, as Chuprov wrote to Petrograd, he had fallen seriously ill in Stockholm and was unable to come. In September 1918 a letter from him arrived. He wrote that he intended to come here in a fortnight and to begin teaching at the Institute. Receiving no money from Petrograd, he finds himself in strained circumstances, as he also wrote in the letter mentioned above. Why had he not returned since then is unknown. A probable cause is the lack of necessary means; the voyage from Stockholm to Petrograd, as the newspapers have reported, costs three thousand rubles.

Prof. Chuprov is one of the most eminent statisticians not only in Russia, but in entire Europe where his contributions on theoretical issues of statistics are well known to all specialists. For the Polytechnical Institute, whose special aim is the preparation of statisticians, it is extremely important to have among its staff such a prominent representative of the Russian statistical science as Prof. Chuprov whose name and work are well known to each Russian statistician.

In April of the current year, 1918, Prof. Chuprov's merits in statistical science caused Comrade Elizarov, the Commissar for insurance, to offer Chuprov the position of head of the then having been established Central

Statistical Directorate of the Soviet Republic. Prof. Chuprov's letters testify that, while living in Stockholm, he has indeed persistently worked on scientific investigations devoted to the abovementioned issues, that they are already prepared to a considerable extent and that he will immediately publish them upon his return to Russia.

Note

1. {No such paper was found. His report on sampling was read in 1910 and published in 1912.}

13. V.I. Romanovsky.
Review of A.A. Chuprov [3]

Zur Theorie der Stabilität statistischer Reihen. *Skand. Aktuarietidskr.*, Bde 1 – 2, 1918 – 1919, pp. 199 – 256 and 80 – 133. *Vestnik Statistiki*, No. 1 – 3, 1923, pp. 255 – 260 …

In 1919 {in 1918 – 1919} there appeared a most important and remarkable work of Chuprov [3] devoted to the theory of stability of statistical series. Every statistician, and all the more every theorist of statistics understands how important, both practically and theoretically, is the fact of stability of statistical series. It is also known, however, that until now its theory constructed by Lexis and his followers was vague and incomplete. It was founded on several particular patterns (I bear in mind the schemes due to Jakob Bernoulli, Poisson, Lexis and Bortkiewicz) without any attempts having been made for generalizing them; there was no precise measure of reliability of the coefficient of dispersion;[1] and, finally, the scientific value itself of that coefficient, – of the central notion of the classical theory of stability of statistical series, – remained unclear and questionable.

In the work under consideration, Chuprov deeply studied all these issues anew and considerably advanced them. We can only regret that, to this day, far from all the Russian scientists working in this field had the opportunity of getting to know this excellent contribution; even now, this is possible only by chance. I was lucky enough to see Chuporov's work and now I wish to acquaint the readers with its main features.

It consists of three rather vast essays: 1) Über den mittleren Fehler des Durchschnittes von gegenseitig nicht unabhängigen Grössen; 2) Über die mathematische Erwartung und den mittleren Fehler der Divergenzkoeffizienten; 3) Über den mittleren Fehler der wesentlichen Schwankungskomponente. This listing already shows that Chuprov devoted his writing to the most important and interesting problems of the theory of stability. Let us consider its essays one after another.

Chuprov begins his **First Essay** by deriving, in the most general case, the mean square error of the arithmetic mean {$x_{(n)}$ of observations $x_1, x_2, \ldots, x_n$} of a number of dependent variables

$$\xi_1, \xi_2, \ldots, \xi_n \quad (1)$$

{respectively}. Then

$$E[x_{(n)} - E x_{(n)}]^2 = (1/n^2) \sum_{h=1}^{n} E[\xi_h - E\xi_h]^2 + (2/n^2) \sum_{h=1}^{n-1} \sum_{g=1}^{n-g} [E x_h x_{h+g} - E\xi_h \xi_{h+g}]. \quad (2)$$

Chuprov established two more equalities:

$$E x_{(n)}^2 = [(1/n) \sum_{h=1}^{n} E\xi_h]^2 + (1/n^2) \sum_{h=1}^{n} E[\xi_h - E\xi_h]^2 + (1/n^2) \sum_{h=1}^{n} \sum_{g\neq h} [E x_h x_g - E\xi_h \xi_g], \quad (3)$$

$$E\sum_{h=1}^{n} [x_h - x_{(n)}]^2 = \sum_{h=1}^{n} [E\xi_h - E x_{(n)}]^2 +$$

$$[(n-1)/n]\sum_{h=1}^{n} \mathrm{E}\,[\xi_h - \mathrm{E}\,\xi_h]^2 - (1/n)\sum_{h=1}^{n}\sum_{g\neq h} [\mathrm{E}\,x_h x_g - \mathrm{E}\,\xi_h \mathrm{E}\,\xi_g]. \quad (4)$$

Relation (2) is especially important: issuing from it and specifying some assumptions about the magnitudes (1), Chuprov derives many other known equalities obtained earlier as particular results by other researchers. Thus, he got the Bohlmann formula

$$\mathrm{E}\,[H - p_{(n)}]^2 = (1/n^2)\sum_{h=1}^{n} p_h q_h + (1/n^2)\sum_{h=1}^{n}\sum_{g\neq h} (p_{hg} - p_h p_g) \quad (5)$$

where H was the number of the occurrences of some {of the studied} event in n dependent trials, $p_1, p_2, \ldots, p_n$, the probabilities of its occurrence respectively, $q_h = 1 - p_h$, $p_{(n)}$, the arithmetic mean of p_h, and p_{hg}, the probability of the event occurring in trials h and g. Issuing from (5), Chuprov

a) Derived the well-known Bortkiewicz formula[2]

$$\mathrm{E}\,[x_{(n)} - c_o]^2 = \frac{c_o(1-c_o)}{mk} + \frac{k-1}{mk}\sum_{h=1}^{s} p_h (c_h - c_o)^2.$$

{There are s variables ξ_h and m series consisting of k trials each are performed with one and the same variable being chosen from among them; p_h is the probability that variable ξ_h is chosen. The variables take values 1 and 0 with probabilities c_h and $(1 - c_h)$ respectively,

$$c_o = \sum_{h=1}^{s} p_h c_h,\quad x(n) = (1/n)\sum_{h=1}^{m}\sum_{j=1}^{k} x_{hj}$$

and x_{hj} is the value taken by ξ_h in series h at trial j.}

b) Established the formula for the variance of magnitudes connected into a chain.

c) Examined the mean square error of the arithmetic mean of the numbers extracted from an urn containing $n_1, n_2, \ldots, n_k$ balls numbered $N_1, N_2, \ldots, N_k$ if λ balls are extracted at once without replacement. He also showed that this problem was relevant to investigations of sampling.

I have only indicated the most important results of Essay 1 leaving aside many other of its issues interesting and important for the theory of stability. It can be seen how sweeping were the generalizations to which Chuprov had arrived and how diversely he was able to apply them.

Essay 2 is devoted to two issues: to investigating the expected square of the coefficient of dispersion and to determining the mean square error of this square. When r series of n trials each are carried out on a random variable ξ (on a variable taking some definite values depending on chance; or, in other words, taking them with some definite probabilities) when the law of distribution of ξ (i.e., both its values and their probabilities) remains invariably constant and the trials are independent from each other, – then, as Chuprov showed by simple and clever calculations, – the expected square of its coefficient of dispersion is unity. Here, he repeated his earlier result [2].

We shall not discuss Chuprov's involved investigation of the mean square error of this square restricting our description by indicating several relevant points. We have mentioned already that Markov [5] had determined the precise value of the mean square error of the coefficient of dispersion (or, rather, the expected square of the difference between the square of this coeffcient and unity) for one particular case, – namely, when the probability of the event under consideration remained constant and the series consisted of one and the same number of trials (r, say) {note the change of notation}. When studying this case, Chuprov established that, as the number of series increased unboundedly, the mean square error tended to $2/(r - 1)$ and remained less than this limit. This, however, cannot be stated in any other instances and it is not known how precise can the number $2/(r - 1)$ be considered for the expectation mentioned. It is therefore clear how shaky is the approximaton $\sqrt{2/(r-1)}$ to the mean square error of the square of the coefficient non-rigorously established by Bortkiewicz.[3]

We also note the interesting inequalities {in Chuprov's Chapt. 2.4 of Essay 2}

$$[2/(r - 1)]\,\{1 - [5/(mr + 1)]\} < \mathrm{E}\,(Q^2 - 1)^2 < 2/(r - 1)$$

that describe the case in which the probability $p = m/n$ of the considered event tends to zero as $n = \infty$.

The revolutionary **Essay 3** is the most interesting and important. It is devoted to studying the mean square error of the essential component of the square of deviations of the values of a certain quantity from its arithmetical mean. In a brief review it is impossible to explicate in any satisfactory manner all its rich and diverse content and we shall therefore only consider its most important conclusions.

Those who are engaged in the theory of stability know how superficial and unsatisfactory is the judgement about the scattering in a given statistical series as estimated by the coefficient of dispersion Q introduced by Lexis and Dormois. If the distribution of the deviations of the values of the studied quantity from its mean is normal (if it obeys the Gauss law), then Q is known to equal 1. The converse judgement is however wrong: Chuprov offered a simple example of a non-normal distribution for which Q was nevertheless unity. Similarly, the case of $Q > 1$ (supernormal dispersion) can take place under most variable conditions; suffice it to recall the Lexian and the Bortkiewicz patterns. Therefore, nothing more can be said here. Chuprov leaves aside the case of $Q < 1$ (subnormal dispersion) because no rational and worthy of attention estimate of the scatter exists here.

After thus critically assessing the service that the coefficient of dispersion can render to the researcher, Chuprov concludes, following Lexis, that it is hardly suited for studying the scatter in statistical series and that, instead, the essential component of the square of deviations in the given series should be directly investigated.

Here, I ought to indicate a fact that is not to be passed over in silence because of its great importance for the history of the theory of stability of statistical series and which Chuprov did not mention. Arriving at the conclusion just above, he chooses that approach to studying the scatter of series which another investigator of the theory of dispersion had already earlier embarked upon. In a number of writings dating back to 1913 Yastremsky treated this theory as a theory of mutability of statistical series and developed methods that enable us not only to put on record a supernormal dispersion of a series, but also to study those fluctuations of the frequencies in the series which cannot be attributed to chance but are caused by some definite evolution of the probability of the studied quantity during the successive trials [6 – 11].

The essence of Yastremsky's method {note the singular instead of the previous plural} consists in that he decomposes the mean square of the deviations of some quantity from its mean value into two components: the mean square of random deviations and the essential part of the mean square of deviations that represents a general characteristic of the physical mutability of the series. For the second component he also discovered a very simple approximate expression convenient for applications. If necessary, the precision of the approximation can be heightened by replacing the expression mentioned by other ones, somewhat more complicated.

Chuprov formulated for himself an absolutely similar problem deepening and generalizing it. Instead of studying the stability of a given statistical series by the coefficient of dispersion, he investigated it by the essential component of the deviations, derived its measure of precision and the methods of determining it in the most general case in which the series consisted of a number of quantities each of them having its own law of distribution and in which a number of series of trials was made. He then specified the results obtained for various particular cases of great theoretical and practical importance. Chuprov, however, solved his problem more rigorously and, as noted above, in a more general setting and adduced many very valuable and interesting remarks. Thus, he indicated that it was possible to apply his formulas

a) For examining the influence of those indicators according to which the individual series were selected, on the serial means of some quantity or on the serial probabilities;

b) For adjusting statistical curves. He illustrated this point by a pattern capable of adjusting the curve of mortality;

c) For cases covered by the law of small numbers, etc.

It is important to note that Chuprov came to a synthesis of the ideas of Lexis and Pearson on the issues of mathematical statistics and showed that they approached, although by differing ways, one and the same main statistical problem of determining the prior pattern underlying the given data. The Lexian way is the determination of the essential component of the square of the deviations and its estimate {estimate of the precision obtained?} and then the derivation of the higher moments of the given distribution so as to ascertain it.

Pearson's approach is however to determine the form of the law of distribution and search for its constants on the basis of empirical data. I indicate once more that Yastremsky had chosen the way of a real synthesis of the Lexian and the Pearsonian methods before Chuprov did, although in a particular and less distinct form. Indeed, in order to ascertain the scatter of a given statistical series he interpolated it by parabolas.

In concluding, I ought to indicate that Chuprov's essays make a difficult reading although all the calculations are there elementary. They demand serious preparation and an acquaintance with, for example, Markov's far from elementary and not at all plain works on the theory of probability. However, it is hardly possible to simplify essentially Chuprov's exposition, – so skillfully it is and such complicated issues he treats. We should only desire that the ideas and conclusions of our famous compatriot be disseminated in full measure and widely applied in statistical practice.

Notes

1. Markov [4] was the first to discover such a measure for a particular case in a hardly applicable form.{I disagree with the author's somewhat negative opinion about the jakob Bernoulli and Poisson schemes.}

2. {Chuprov himself, in Chapt. 2 of Essay 1, called this formula after Bienaymé and Bortkiewicz but had not provided any reference.}

3. The reader can find it in [1].{Actually, according to Bortkiewicz, this is the error of the coefficient itself, see §1 of Essay 2 in the reviewed contribution.}

References

1. Bortkevich, V.I. (Bortkiewicz, L. von), On the measure of precision of the coefficient of dispersion. *Vestnik Stat*., No. 1 – 4, 1921, pp. 5 – 10. (R)

2. Chuprov, A.A. On the expectation of the coefficient of dispersion (1916). Translated in this collection.

3. Chuprov, A.A. Zur Theorie der Stabilität statistischer Reihen. *Skand. Aktuarietidskr*., t. 1, 1918, pp. 199 – 256; t. 2, 1919, pp. 80 – 133. Russian translation with Chuprov's own corrections of the original text: *О теории дисперсии* (On the Theory of Duspersion). Editor and translator N.S. Chetverikov. Moscow, 1968, pp. 138 – 224.

4. Markov, A.A. On the coefficient of dispersion (1916). Transl.: DHS 2514, 1998, pp. 195 – 204.

5. Yastremsky, B.S. The law of sufficiently large numbers when estimating the stability of statistical series (1913) [11, pp. 13 – 32]. (R)

6. --- The theory of dispersion as a theory of mutability of statistical series (Read 1914, publ. 1937) [11, pp. 33 – 58]. (R)

7. --- An indirect method of determining the stability or mutability of statistical series (1919) [11, pp. 58 – 73]. (R)

8. --- The theory of mutability of statistical series (1920) [11, pp. 74 – 80]. (R)

9. --- The law of small numbers in the light of the theory of mutability of statistical series (1922) [11, pp. 80 – 90]. (R)

10. --- An indirect method of determining periodic mutability of statistical series (Report). *Vestnik Stat*., No. 9 – 12, 1922 (1923). Only mentioned but not published (p. 16 of second paging).

11. --- *Избранные труды* (Sel. Works). M., 1964.

14. Obituaries of Chuprov;

the first two were published in Swedish and translated here from their private Russian translations by A.A. Muravyev

14a. K. Gulkevitj

Nord. Stat. Tidskr., vol. 5, 1926, pp. 167 – 170 …

It is painful to speak about a man before whose spiritual purity and extremely sensitive conscience we respectively bow our heads. Painful, because I do not wish to resort to sentimental words although they unintentionally appear all by themselves. I would like to speak plainly, without emphasize.

Although even ten years have not passed since we got acquainted, a most intimate friendship connected Professor Chuprov and me. In those days I was in Christiania {now Oslo}. Due to the war going on, A.A. was unable to pass the time in Italy in accord with his long ago rooted habit and stopped for a while in Norway. His friends in Petrograd {Petersburg – Petrograd – Leningrad – again Petersburg} asked him to visit me, but his ingrained prejudice against bureaucracy hindered him from fulfilling this request. P.N. Savitsky, the then attaché and commercial counsellor at the legation and one of our most gifted young scientists, was never able, in spite of all his efforts, to persuade Chuprov to come over his shyness. Only on the very last day Chuprov plucked up his courage.

Our mutual surprise was great: our relations immediately became so cordial that the Professor spent all his time before the train's departure at the legation and we parted friends at his coach. In addition to his moral purity, I was at once charmed with his boundless goodwill; fairness; patience with respect not only to the points of view of other people, but even to their actions; amazing composure in spite of a hot temper; love of nature; excellent understanding of art; and admiration of music.

Then occurred our brief meeting in May 1917 in Torneo (Tornio) on the Finnish – Swedish border. A.A. went back to Norway whereas I hurried to Petrograd. This meeting was the only bright moment against the background of painful and chaotic impressions of the events taking place in our mother country that accompanied anyone arriving from Russia.

In the autumn A.A. stopped for a short while in Stockholm on his journey back to Russia, but the events {there} were taking on ever more troubled tinges, and I managed to persuade him not to hurry back home, to wait for the situation to clear up. The October {Nov. 7, new style} coup d'état. Mournful days, weeks, months. A.A. moves to my place, attentively follows the events occurring *there*, tries to solve the {Russian} riddle. A perfect unanimity in appraising the developments alleviated our life during that mournful period.

In June 1920 Chuprov moved to Dresden, and I will hardly be mistaken if I say that thus began the happiest time of his life. He entirely devoted himself to science. In Russia, teaching tired him; although it also undoubtedly attracted him, it diverted him from creative scientific work. Here {in Dresden}, on the contrary, he was able to devote himself completely to scientific work. He published a number of writings in various statistical periodicals, read reports for scientific societies in Copenhagen and Stockholm and delivered a special course at Christiania University. Other, more competent authors will write about Chuprov's scientific work. However, after a five-year residence in Dresden, where the worthy and generally respected Fräulein Dietzel, who had let Chuprov her apartment, and was friendly looking after him, his small means began to dry up. It became necessary to search out for a more stable source of income.

In 1925 A.A. moved to Prague. The situation that he found there did not answer either his disposition, or habits and only the goodwill rendered him by E.D. and S.N. Prokopovich and by several others brightened up his life. In summer, some obscure disease accompanied by a fever revealed itself. Physicians wavered in their diagnoses between bronchitis and malaria. A.A. went to Rimini {Italy} to bask in the sun which he passionately loved. A local physician treated him there for bronchitis. At the end of September Chuprov went to the International Statistical Congress in Rome. Was very satisfied with what he there experienced and accomplished. Being afraid of causing me a lot of trouble by coming ill, he took medical advice in Rome. The doctors sent him to hospital. Nothing to be surprised at. Fever did not abate for the fourth month running. Physicians attempted to treat him for various diseases but without success. Professor Shchurovsky, who was then in Rome and ran across him by chance, diagnosed the illness at once: endocarditis lenta. The menacing term did not scare A.A. Shchurovsky had done a good deed by preparing the sick man for the most unpleasant manifestations such as boils usually accompanying his disease. In addition {however} he gave hope to Chuprov that absolute rest will settle everything.

By the end of December Chuprov could not bear his stay in hospital in Rome anymore. With greatest difficulties he moved to Genéva; during his travel, when attempting to leave his coach at the Brieg railway station, he fell down headlong. It was a wonder that his glasses did not break after hitting the sidewalk so that he was not left blind.

In Genéva he again found himself in a hospital ward, but, being true to himself, he did not wish to trouble the doctors just before Christmas and did not allow to send for them immediately. Only after the New Year the head of the medical clinic and Professor at Genéva University, Maurice Roch, a cardiologist, examined Chuprov and found that his condition was almost hopeless. Nevertheless, he continued to keep an eye on the disease with sincere and deep sympathy. The assistant, whom he appointed to look after A.A., Privatdozent at Genéva University, Dr. Katzenellbogen of Russian extraction, displayed a touching tenderness and loving care for the sick person. Chuprov was moved by this attitude and experienced an utmost trust in, and a warmest friendly devotion to Katzenellenbogen.

Meanwhile Chuprov's condition greatly worsened; a threatening symptom, a local temporary paralysis, required a new clinical treatment. On Sunday, 18 April, I and Katzenellbogen took Chuprov to the clinic. At home, he invariably complained every time when the temperature in the room fell below 23° C and objected to the opening of the window. The nurse in the clinic had been opening the great window in his little but neat and cosy ward, but he did not notice it anymore. During the day his face was pale. He breathed hard and his condition deeply worried me. Meanwhile, the nurse soothed me by stating that there was no immediate danger. And indeed, after some time his condition appeared more reassuring: his face became calm and not as pale as before. A.A. drank a cup of coffee and, answering my question whether it was good enough, told me: "So – so". These were his last words. He was given a tranquilizer for the night and we parted until morning.

Yesterday, at 6 a.m., the telephone rang in my bedroom. It was the hospital. At 5^h10^m the nurse, who sat by the sick man, noticed that his breathing had gradually begun to interrupt. The doctor was sent for, and he certified that all was over. Without waking up, A.A. crossed over from being in existence to the next world. With his sister's (Frau Heymons') permission, his body was cremated so as to make possible, when circumstances will allow it, the return of at least his remains to the passionately loved by him Russia.

Genéva, 20 April 1926

14b. L. von Bortkiewicz

Ibidem, pp. 163 – 166…

Chuprov died on 19 April in Genéva in his 53rd year. He was the son of the well-known scientist Aleksandr Ivanovich Chuprov, who for many years had been head of the chair of political economy and statistics at Moscow University and who died 18 years ago, also abroad, in Munich.

From his father, A.A. inherited an interest in economic and social issues and decided, even as a school student, to devote himself to the service of social sciences. Already then, he outlined a peculiar manner of preparing himself for this work. Issuing from his conviction in that social phenomena should be mainly studied by the statistical method that ought to be based on mathematics, Chuprov had entered the physico-mathematical faculty of Moscow University and began to study political economy and statistics "officially" only after graduation from the Mathematical Department.

To this end, he went to Berlin, then to Strasburg, – to Professor Knapp, an outstanding representative of mathematical statistics, of that discipline to which Chuprov experienced a special inclination. By that time, however, Knapp already lost interest in the theory of statistics [1] and exactly for this reason Chuprov came to an arrangement with him choosing a subject for his dissertation from a quite another field. The dissertation was devoted to the problems of village communities and considerably excelled the standard accepted in the German universities for such works with respect to its scope as well as volume. It was published in the *Abhandlungen aus dem staatswissenschaftlichen Seminar zu Strassburg* in 1902 under the title *Feldgemeinschaft*.

In Russia, the Doctor degree obtained abroad opened the way for Chuprov to an examination for the degree of Master of political economy. In the autumn of 1902, after passing the oral part of this examination at Moscow University, Chuprov was elected Dozent of the Faculty of economics at the recently established Petersburg Polytechnical Institute. Although his abilities as a teacher and a lecturer were higher than average, the pedagogic activity was burdening for him, especially in his first years. Indeed, devoting himself with an exceptional conscientiousness to the preparation of lectures and materials for seminars, he was unable to spare enough time or strength for his own scientific work.

One of his long ago planned systematic works on the theory of the statistical method advanced slowly; however, because of external circumstances it was desirable to speed up the publication of a contribution that would have corresponded with respect to volume to a master dissertation and could have ensured him a professorship. Nothing else remained except abandoning his initial plan and publishing, instead of a complete work, a number of articles connected by contents and adjoining one another.

So it happened that Chuprov's *Очерки* (Essays) [1], that constituted a special chapter of sorts of the general theory of the principles of the statistical method, had appeared. In accord with the author's conception, this method is applied not only in social investigations but in the natural sciences as well. There, its aim consists in checking the presence or absence of dependences between the studied phenomena if the inductive reasoning does not help either in causal or stochastic sense and a dependence can only be established by mass observations on the strength of the Bernoulli theorem or the so-called law of large numbers. The statistical method is thus being directly connected with the theory of probability or, in other words, with mathematics. Nevertheless, the *Очерки* are almost free from mathematical proofs and formulas. According to Chuprov's own expression, his exposition goes on "along a boundary path between statistics, {mathematical} theory of probability and logic" and does not assume that the reader has special knowledge of any of these disciplines. At the same time, however, Chuprov impresses the reader with the feeling that he is a first-class specialist in each of them.[2] It might be said that he revealed his gift for popularizing scientific truths, his being a high-class popularizer of which only all-round educated men are capable.

Chuprov submitted his *Очерки* to the Law faculty of Moscow University as a Master dissertation. His official opponents at its public defence were the statistician N.A. Kablukov and B.A. Kistiakovsky as a specialist in philosophy and, especially, methodology of social sciences. The author splendidly defended his dissertation and (a very rare occasion indeed) the faculty at once conferred on him the degree of Doctor of political economy bypassing the {intermediate} master degree required by the charter of the universities.

Immediately afterwards Chuprov acquired the highest academic status of professor and was able to publish his scientific investigations in a format that he considered as the most suitable without bothering about external circumstances. His favorite form was a composition or report on an accurately demarcated subject with a clearly posed problem being definitively solved. Such were for example his papers in *Biometrika* published in 1918 – 1921 crammed with mathematical formulas and making high demands on the readers. Owing to these very writings Chuprov became well-known in England and exactly to them he was indebted for his election to the Royal Statistical Society as Honorary Fellow.

Of a somewhat other nature are Chuprov's essays on certain issues complete with accompanying critical remarks, witness some of his numerous papers in the *Nordisk Statistisk Tidskrift* of 1922 – 1925. As a critic, he was benevolent and tolerant; what he did not bear in scientific work was carelessness and confusion of styles. He himself provided an example of rigor and completeness in his field. Chuprov's last work was devoted to the theory of correlation [2]. To some degree it supplemented and completed his *Очерки*.

I have already emphasized that the post-war years of "forced inactivity" which Chuprov spent in Stockholm and Dresden were marked by the flourish of his scientific work. Refusing to abandon his research, he turned down not less than two requests for heading an university chair. However, not long before his death he changed his opinion on that point. In a letter of 19 January of this year, the last one that I received from him, he wrote about a professorship at Heidelberg:

> *If the arrangement* […] *will be realized, there would be nothing better for me because I do not want to go to Russia from where I have again received a number of inviting proposals.*

He did not want to return to his mother country certainly not because of some apprehensions for his own safety, – on the contrary, he had all grounds for expecting an attentive and even sympathetic attitude from the powers that be,[3] – but because neither his democratic convictions, nor his feeling of solidarity, nor, finally, and quite simply, his moral principles allowed him to secure for himself those privileged rights of which the others were deprived. And, although it should be considered as unnatural that this man, a Russian by blood and in spirit, an ornament and pride of the Russian science, was unable, during the last period of his life, to find a place in Russia for himself, the severe reality of the Soviet regime did not directly touch his sensitive soul. And this may serve as a consolation in the grief occasioned by the passing away of a combative comrade and friend.

Notes

1. {Cf. [1, p. 35]}.

2. {Markov was of a quite another opinion concerning probability, see my Foreword to this collection.}

3. {Bortkiewicz would have hardly repeated these words in 1927, much less in 1928 ([3, p. 30; 4]}.

References

1. Chuprov, A.A. *Очерки по теории статистики* (Essays on the Theory of Statistics). M., 1909, 1910, 1959.

2. --- *Grundbegriffe und Grundprobleme der Korrelationstheorie*. Leipzig, 1925. Engl. transl., 1939. Russian transl., 1926 and 1960.

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

4. --- Statistics in the Soviet epoch. JNÖS, Bd. 217, 1998, pp. 529 – 549.

14c. A. Kaminka

Newspaper *Rul*, 21 April 1926, pp. 1 – 2

[…] His ideology was inspired with high social ideals. Absolutely alien to any intrigues, he possessed considered political ideals and was glad to serve them when perceiving an accruing common good sufficient for justifying the diversion of some of his time from scientific work. […] he was among those few people who […] seriously influenced the elaboration of the ideological substance of the Cadet party.[1] […] neither did he want to return to Russia after the Bolsheviks had established themselves there. The reason was not even because he hated them politically, – he never advocated refusal to return home, – but simply because his scientific conscience told him that, where Bolshevism is reigning, an abomination of ruin is unavoidable not only in political and economic life, but also in cultural and scientific work.

Note
1. { Constitutional Democrats. See Chuprov (1906)}.

References

Chuprov; A.A. (1906), *Конституционно-демократическая партия и социализм* (The Constitutional-Democratic Party and Socialism). M.

15. Letters of Helene Heymons (née Chuprov) to Karl Pearson

Manuscript. (1926). Pearson Papers 719/9, University College London …

Foreword

{Soon after Chuprov's death Helene Heymons wrote two letters to Pearson. They are kept at University College London (Pearson papers 719/9) and I am now publishing them in their original German. A Russian translation of these letters appeared with a short commentary by Prof. Irina Eliseeva (Sheynin 2001). Answering Pearson (whose letters to her and/or to her husband are lost), H.H. provided information about her late brother and expressed her wish to receive some offprints of Pearson's future obituary of Chuprov. Such an obituary had however never appeared perhaps because other authors soon published notices of Chuprov's death, especially Isserlis (1926). Nevertheless, Chuprov's photograph appeared in *Biometrika* (vol. 18, 1926, facing p. 233)[1].

Elena Aleksandrovna Tchouproff was born in Moscow in 1877 and married the zoologist, Professor Richard Heymond in 1903. Her expression, *Frau Professor*, at the end of Letter 1 (now current only when the woman herself is Professor), meant *wife of a Professor*. True, she helped her husband in his work at least when he participated in editing Bd. 2 of the fourth edition of Brehm's classical *Tierleben* (1915). This scant information (that I checked as far as the *Tierleben* was concerned) was provided by H.H.'s granddaughter, Marianne Heymons, an elderly woman living in Berlin. Another granddaughter, Suzanne Whitley, lives in England.

In 1918 – 1919 and 1921 Pearson published Chuprov's lengthy contribution (in two parts) in *Biometrika* and now it might be added that by 1926 Pearson had not changed his high opinion of Chuprov.

H.H. intended to re-bury Chuprov's remains in Russia (also see Chuprov's obituary by Gulkevich in this collection). This, however, had not happened either for political or simply for personal (for example, financial) reasons.}

* * *

Letter No. 1, 3 June 1926

Hochgeehrter Herr!

In Beantwortung Ihres an meinen Mann, Professor Dr. Heymons, gerichteten Schreibens erlaube ich mir Ihnen die Photographie meines verstorbenen Bruders, Professor Dr. Tchouproff, zu übersenden, es ist die gleiche, die ich auch der Zeitung *Rul* zu Verfügung gestellt hatte.[2] Es handelt sich um die letzte Aufnahme meines Bruders, die er im Sommer vorigen Jahres hatte anfertigen lassen, als er ein Passbild für seine Reise nach Rom zum Internationalen Statistischen Kongress brauchte. Ich bitte Sie, mir das Bild nach Benutzung freundlichst wieder zurückschicken zu wollen.

Ich würde Ihnen sehr dankbar sein, wenn Sie mir von Ihrem Artikel über meinen Bruder 20 Separatdrücke zur Verfügung stellen wollten. Ich möchte dieselben gern an meine in Moskau wohnenden Schwestern und an einige Schüler meines Bruders verteilen. Vielleicht darf ich Sie bitten, mir den Betrag für die Separatdrücke mitteilen zu wollen.

Mit vorzüglicher Hochachtung

Frau Professor Helene Heymons geb. Tchouproff

Letter No. 2, 21 June 1926

Hochgeehrter Herr Professor!

Mit vielem Dank bestätige den Empfang Ihres Schreibens und der Photographie und freue mich sehr, dass der Abdruck zu Ihrer Zufriedenheit ausgefallen ist. Ihre Fragen beantworte ich gern und hoffentlich, auch genügend ausführlich! Sollten Sie noch weitere Fragen haben, so stehe ich immer gern zu jeder Auskunft über meinen Bruder zur Verfügung.

Professor Tchouproff ist am 5. Februar (alten Stils) 1874 in Massalsk, einer Stadt im Gouvernement Kaluga in Russland geboren.[3] Sein Vater, ein bedeutender Statistiker, war Professor für National Ökonomie und Statistik an der Universität Moskau.

Tchouproff besuchte ein humanistisches Gymnasium in Moskau,[4] er war immer der beste Schuler, hat das Abiturentenexamen mit Auszeichnung bestanden und dann, schon mit dem Plan später Statistiker zu werden, die mathematische Fakultät in Moskau besucht.[5] In Jahre 1896 ist er nach Deutschland gegangen und hat hauptsächlich in Berlin und Strassburg staatswissenschaftliche Vorlesungen gehört [6] und wissenschaftlich gearbeitet um in Strassburg im Jahre 1901 bei Professor Knapp mit einer umfangreichen Arbeit *Über die Feldgemeinschaft* Summa cum laude den Doktorgrad zu erwerben. 1902 ist er Professor der Statistik am Polytechnischen Institut in Petersburg geworden.[7] Im Jahre 1908 hat er eine grundlegende Arbeit über die Theorie der Statistik – betitelt *Grundrisse der Theorie der Statistik* bei der Universität Moskau eingereicht und daraufhin den Magister und Doktorgrad gleichzeitig bekommen. Es war dies eine große Auszeichnung für ihn, da sonst zu diesem zwecke 2 verschiedene Arbeiten eingerichtet werden müssen.

Von 1902 bis 1917 ist er Professor der Statistik am Polytechnischen Institut gewesen und hat während dieser Zeit mehrere Schüler herausgebildet, die jetzt z. Teil schon Gelehrte von Namen sind.

Im Jahre 1917 weilte er zwecks wissenschaftlich Studiums im Auslande. Von dort ist er nach Einführung des Bolschewismus nicht mehr nach Russland zurückgekehrt. Von 1917 – 19 hat er in Stockholm *Bulletins* über die Weltwirtschaft für Russland herausgegeben.[8] Vom Jahre 1920 lebte er in Deutschland, in Dresden. Er lebte sehr bescheiden, ganz zurückgezogen, nur für seiner Wissenschaft und war sehr glücklich darüber, dass er von keinen Pflichten abgelenkt, nur theoretisch arbeiten konnte. Er sah bei sich kaum Menschen, obwohl er viele Freunde hatte, die er sehr schätzte, und die ihn ebenfalls schätzten, er war sehr musikalisch und verstand viel von der Kunst, gönnte sich aber, vielleicht weil er sein frühes Ende ahnte, keine Zeit für etwas anderes als für seine Arbeit.

Als im Januar 1925 seine Ersparnisse, trotz des äußerst sparsamen Lebens, zu Ende gingen, übersiedelte er auf das Anraten der Freunde nach Prag um dort eine Professur zu übernehmen. Vordem hat man ihm mehrere Professuren angeboten unter anderem in Dorpat {Tartu} und Heidelberg, er wollte aber so lange es irgend ging keine weitere Pflichten haben.

Seine schwere Krankheit (Endocarditis lenta) kam zuerst in Prag zu Ausbruch. Trotz seines Leidens ist er nach Rom zum Internationalen Statistischen Kongress gereist, hat den Kongress mitgemacht. Von dort ist er nach Genf zu seinem besten Freunde, Exzellenz v. Goulkévitch, gefahren. In Genf hat er die letzten 4 Monate schwer krank zugebracht und ist dort am 19. April gestorben. Seine Asche steht im Genfer Kolumbarium und soll später nach Russland überführt werden.

Er war Junggeselle. Von den nächsten Angehorigen sind nur 3 Schwestern am Leben. Tchouproff war Mitglied 1) des Internationalen Statistischen Instituts. 2) der Londoner Königlichen Ökonomischen Gesellschaft.9 3) der Akademie der Wissenschaften in Petersburg (Leningrad)[10] und Ehrenmitglied der Londoner Statistischen Gesellschaft.

Um sein Andenken zu ehren, hat man am 23. Mai in Moskau und am 30. Mai in Leningrad feierliche Sitzungen veranstaltet. Das Polytechnische Institut in Leningrad hat beschlossen, der Abteilung an der er früher als Professor wirkte (den genauen Namen der Abteilung kann ich momentan nicht sagen, weil ich den Brief für kurze Zeit weitergeschickt habe), seinen Namen zu verleihen.[11]

Als Probe seines Handschrift füge ich, Ihrem Wunsche entsprehend, 4 Kärtchen bei und bitte sie mir gelegentlich zurückzuschicken.[12]

Mit vorzüglicher Hochachtung Helene Heymons geb. Tchouproff

Notes

1. {Portraits of two other Russian scientists, Chebyshev and Markov, also appeared in *Biometrika* (in vol. 22, 1930, and vol. 24, 1932, respectively). Pearson received their photographs together with a covering letter of 31 July 1926 from Fersman, the Secrétaire Perpétuel de l'Académie des Sciences de l'URSS (Pearson papers 511). Fersman (a geochemist and mineralogist) answered Pearson's letter of 15 July 1926 and informed him that the Academy had regarded his proposition to publish these photographs *avec un sentiment de vive satisfaction*.}

2. {A Russian newspaper in Berlin. It published Chuprov's obituary written by Kaminka.}

3. {Other sources indicate 6 February. H.H. possibly made a mistake when stating the date of birth in accord with the old style: in the 19^{th} century, the difference between the styles amounted to 12 rather than to 13 days as it is now (and as it was in the 20^{th} century).}

4. {Chuprov had indeed graduated from a *humanistische Gymnasium*. His unpublished student *Composition*, kept at Gorky Library, Moscow University, contains excerpts from many contributions both in Russian translation and in the original languages including Greek and Latin. At the same time, mathematics in the supreme form of his *Gymnasium* was taught then by an eminent mathematician, Mlodseevsky.}

5. {More precisely. Physical and mathematical faculty.}

6. {Staatswissenschaft (university statistics, statecraft) originated in the 17th century and aimed at studying the climate, geographical features, political structure, economics and population of separate countries. It disintegrated into several sciences and for example Knapp (see below) hardly studied anything except economics.}

7. {Chuprov became Professor in 1910 rather than in 1902 after defending his (second) dissertation, see below. True, he held the chair of statistics from 1902.}

8. {Not a single issue of these *Bulletins* was found.}

9. {Chuprov was the Society's Correspondent for Russia and could have become its member. Apparently, he did not really want to.}

10. {He was elected Corresponding Member of that Academy in 1917; in 1914, the city acquired a new name, Petrograd, in 1924 it was called Leningrad, and, from about 1992, it is again Petersburg.}

11. {The Grand Meeting of the Council of the Economic Faculty (of the *Abteilung*, as H.H. called it) was indeed held in 1926 but the Faculty was not called after Chuprov. However, Dmitriev (see Introduction) informs me that he happened to see some books bearing the stamp *Chuprov statistical room* (that belonged to the Faculty). In Moscow, as Eliseeva (Sheynin 2001, p. 64n) noted, the commemorative conference was organized by the "Executive commission of the [national] statistical congresses". The minutes of the conference, unlike those of the Leningrad meeting, were hardly published.}

12. {H.H. also sent Pearson Chuprov's pass (with his signature) given him on 5 September 1921 by the *Staatsbibliothek* in Munich.}

References

I[sserlis, L.] (1926), Obituary of Chuprov. *J. Roy. Stat. Soc*., vol. 89, pp. 619 – 622.

Sheynin, O. (2001), Helene Chuprov's letters to Karl Pearson. *Voprosy Statistiki*, No. 3, pp. 62 – 64. (R)

Letters of Chuprov's Sister Helene Heymons to Karl Pearson

{Soon after Chuprov's death Helene Heymons wrote two letters to Pearson. They are kept at University College London (Pearson papers 719/9) and I am now publishing them in their original German. A Russian translation of these letters appeared with a short commentary by Prof. Irina Eliseeva (Sheynin 2001). Answering Pearson (whose letters to her and/or to her husband are lost), H.H. provided information about her late brother and expressed her wish to receive some offprints of Pearson's future obituary of Chuprov. Such an obituary had however never appeared perhaps because other authors soon published notices of Chuprov's death, especially Isserlis (1926). Nevertheless, Chuprov's photograph appeared in *Biometrika* (vol. 18, 1926, facing p. 233)[1].

Elena Aleksandrovna Tchouproff was born in Moscow in 1877 and married the zoologist, Professor Richard Heymond in 1903. Her expression, *Frau Professor*, at the end of Letter 1 (now current only when the woman herself is Professor), meant *wife of a Professor*. True, she helped her husband in his work at least when he participated in editing Bd. 2 of the fourth edition of Brehm's classical *Tierleben* (1915). This scant information (that I checked as far as the *Tierleben* was concerned) was provided by H.H.'s granddaughter, Marianne Heymons, an elderly woman living in Berlin. Another granddaughter, Suzanne Whitley, lives in England.

In 1918 – 1919 and 1921 Pearson published Chuprov's lengthy contribution (in two parts) in *Biometrika* and now it might be added that by 1926 Pearson had not changed his high opinion of Chuprov.

H.H. intended to re-bury Chuprov's remains in Russia (also see Chuprov's obituary by Gulkevich in this collection). This, however, had not happened either for political or simply for personal (for example, financial) reasons.}

Letter No. 1, 3 June 1926

Hochgeehrter Herr!

In Beantwortung Ihres an meinen Mann, Professor Dr. Heymons, gerichteten Schreibens erlaube ich mir Ihnen die Photographie meines verstorbenen Bruders, Professor Dr. Tchouproff, zu übersenden, es ist die gleiche, die ich auch der Zeitung *Rul* zu Verfügung gestellt hatte.[2] Es handelt sich um die letzte Aufnahme meines Bruders, die er im Sommer vorigen Jahres hatte anfertigen lassen, als er ein Passbild für seine Reise nach Rom zum Internationalen Statistischen Kongress brauchte. Ich bitte Sie, mir das Bild nach Benutzung freundlichst wieder zurückschicken zu wollen.

Ich würde Ihnen sehr dankbar sein, wenn Sie mir von Ihrem Artikel über meinen Bruder 20 Separatdrücke zur Verfügung stellen wollten. Ich möchte dieselben gern an meine in Moskau wohnenden Schwestern und an einige Schüler meines Bruders verteilen. Vielleicht darf ich Sie bitten, mir den Betrag für die Separatdrücke mitteilen zu wollen.

Mit vorzüglicher Hochachtung

Frau Professor Helene Heymons geb. Tchouproff

Letter No. 2, 21 June 1926

Hochgeehrter Herr Professor!

Mit vielem Dank bestätige den Empfang Ihres Schreibens und der Photographie und freue mich sehr, dass der Abdruck zu Ihrer Zufriedenheit ausgefallen ist. Ihre Fragen beantworte ich gern und hoffentlich, auch genügend ausführlich! Sollten Sie noch weitere Fragen haben, so stehe ich immer gern zu jeder Auskunft über meinen Bruder zur Verfügung.

Professor Tchouproff ist am 5. Februar (alten Stils) 1874 in Massalsk, einer Stadt im Gouvernement Kaluga in Russland geboren.[3] Sein Vater, ein bedeutender Statistiker, war Professor für National Ökonomie und Statistik an der Universität Moskau.

Tchouproff besuchte ein humanistisches Gymnasium in Moskau,[4] er war immer der beste Schuler, hat das Abiturentenexamen mit Auszeichnung bestanden und dann, schon mit dem Plan später Statistiker zu werden, die mathematische Fakultät in Moskau besucht.[5] Im Jahre 1896 ist er nach Deutschland gegangen und hat hauptsächlich in Berlin und Strassburg staatswissenschaftliche Vorlesungen gehört [6] und wissenschaftlich gearbeitet um in Strassburg im Jahre 1901 bei Professor Knapp mit einer umfangreichen Arbeit *Über die Feldgemeinschaft* Summa cum laude den Doktorgrad zu erwerben. 1902 ist er Professor der Statistik am Polytechnischen Institut in Petersburg geworden.[7] Im Jahre 1908 hat er eine grundlegende Arbeit über die Theorie der Statistik – betitelt *Grundrisse der Theorie der Statistik* bei der Universität Moskau eingereicht und daraufhin den Magister und Doktorgrad gleichzeitig bekommen. Es war dies eine große Auszeichnung für ihn, da sonst zu diesem Zwecke 2 verschiedene Arbeiten eingerichtet werden müssen.

Von 1902 bis 1917 ist er Professor der Statistik am Polytechnischen Institut gewesen und hat während dieser Zeit mehrere Schüler herausgebildet, die jetzt z. Teil schon Gelehrte von Namen sind.

Im Jahre 1917 weilte er zwecks wissenschaftlich Studiums im Auslande. Von dort ist er nach Einführung des Bolschewismus nicht mehr nach Russland zurückgekehrt. Von 1917 – 19 hat er in Stockholm *Bulletins* über die Weltwirtschaft für Russland herausgegeben.[8] Vom Jahre 1920 lebte er in Deutschland, in Dresden. Er lebte sehr bescheiden, ganz zurückgezogen, nur für seiner Wissenschaft und war sehr glücklich darüber, dass er von keinen Pflichten abgelenkt, nur theoretisch arbeiten konnte. Er sah bei sich kaum Menschen, obwohl er viele Freunde hatte, die er sehr schätzte, und die ihn ebenfalls schätzten, er war sehr musikalisch und verstand viel von der Kunst, gönnte sich aber, vielleicht weil er sein frühes Ende ahnte, keine Zeit für etwas anderes als für seine Arbeit.

Als im Januar 1925 seine Ersparnisse, trotz des äußerst sparsamen Lebens, zu Ende gingen, übersiedelte er auf das Anraten der Freunde nach Prag um dort eine Professur zu übernehmen. Vordem hat man ihm mehrere Professuren angeboten unter anderem in Dorpat {Tartu} und Heidelberg, er wollte aber so lange es irgend ging keine weitere Pflichten haben.

Seine schwere Krankheit (Endocarditis lenta) kam zuerst in Prag zu Ausbruch. Trotz seines Leidens ist er nach Rom zum Internationalen Statistischen Kongress gereist, hat den Kongress mitgemacht. Von dort ist er nach Genf zu seinem besten Freunde, Exzellenz v. Goulkévitch, gefahren. In Genf hat er die letzten 4 Monate schwer krank zugebracht und ist dort am 19. April gestorben. Seine Asche steht im Genfer Kolumbarium und soll später nach Russland überführt werden.

Er war Junggeselle. Von den nächsten Angehörigen sind nur 3 Schwestern am Leben. Tchouproff war Mitglied 1) des Internationalen Statistischen Instituts. 2) der Londoner Königlichen Ökonomischen Gesellschaft.[9] 3) der Akademie der Wissenschaften in Petersburg (Leningrad)[10] und Ehrenmitglied der Londoner Statistischen Gesellschaft.

Um sein Andenken zu ehren, hat man am 23. Mai in Moskau und am 30. Mai in Leningrad feierliche Sitzungen veranstaltet. Das Polytechnische Institut in Leningrad hat beschlossen, der Abteilung an der er früher als Professor wirkte (den genauen Namen der Abteilung kann ich momentan nicht sagen, weil ich den Brief für kurze Zeit weitergeschickt habe), seinen Namen zu verleihen.[11]

Als Probe seines Handschrift füge ich, Ihrem Wunsche entsprechend, 4 Kärtchen bei und bitte sie mir gelegentlich zurückzuschicken.[12]

Mit vorzüglicher Hochachtung Helene Heymons geb. Tchouproff

Notes

1. {Portraits of two other Russian scientists, Chebyshev and Markov, also appeared in *Biometrika* (in vol. 22, 1930, and vol. 24, 1932, respectively). Pearson received their photographs together with a covering letter of 31 July 1926 from Fersman, the Secrétaire Perpétuel de l'Académie des Sciences de l'URSS (Pearson papers 511). Fersman (a geochemist and mineralogist) answered Pearson's letter of 15 July 1926 and informed him that the Academy had regarded his proposition to publish these photographs *avec un sentiment de vive satisfaction*.}

2. {A Russian newspaper in Berlin. It published Chuprov's obituary written by Kaminka reprinted in this collection.}

3. {Other sources indicate 6 February. H.H. possibly made a mistake when stating the date of birth in accord with the old style: in the 19th century, the difference between the styles amounted to 12 rather than to 13 days as it is now (and as it was in the 20th century).}

4. {Chuprov had indeed graduated from a *humanistische Gymnasium*. His unpublished student *Composition*, kept at Gorky Library, Moscow University, contains excerpts from many contributions both in Russian translation and in the original languages including Greek and Latin. At the same time, mathematics in the supreme form of his *Gymnasium* was taught then by an eminent mathematician, Mlodseevsky.}

5. {More precisely. Physical and mathematical faculty.}

6. {Staatswissenschaft (university statistics, statecraft) originated in the 17th century and aimed at studying the climate, geographical features, political structure, economics and population of separate countries. It disintegrated into several sciences and for example Knapp (see below) hardly studied anything except economics.}

7. {Chuprov became Professor in 1910 rather than in 1902 after defending his (second) dissertation, see below. True, he held the chair of statistics from 1902.}

8.{Not a single issue of these *Bulletins* was found.}

9. {Chuprov was the Society's Correspondent for Russia and could have become its member. Apparently, he did not really want to.}

10. {He was elected Corresponding Member of that Academy in 1917; in 1914, the city acquired a new name, Petrograd, in 1924 it was called Leningrad, and, from about 1992, it is again Petersburg.}

11. {The Grand Meeting of the Council of the Economic Faculty (of the *Abteilung*, as H.H. called it) was indeed held in 1926 but the Faculty was not called after Chuprov. However, Dmitriev (see Introduction) informs me that he happened to see some books bearing the stamp *Chuprov statistical room* (that belonged to the Faculty). In Moscow, as Eliseeva (Sheynin 2001, p. 64n) noted, the commemorative conference was organized by the "Executive commission of the [national] statistical congresses". The minutes of the conference, unlike those of the Leningrad meeting, were hardly published.}

12. {H.H. also sent Pearson Chuprov's pass (with his signature) given him on 5 September 1921 by the *Staatsbibliothek* in Munich.}

References

I[sserlis, L.] (1926), Obituary of Chuprov. *J. Roy. Stat. Soc.*, vol. 89, pp. 619 – 622.
Sheynin, O. (2001), Helene Chuprov's letters to Karl Pearson. *Voprosy Statistiki*, No. 3, pp. 62 – 64. (R)

16. I.I. Eliseeva.
The Chuprov Statistical School

Voprosy Statistiki, No. 2, 1995, pp. 40 – 43 …

Any branch of science lives, develops, bears fruitful ideas after a group of like-minded scholars led by a person of precise mind and profound knowledge has emerged. Aleksandr Aleksandrovich Chuprov (1874 – 1926), Professor at Peter the Great Petersburg Polytechnical Institute, who commanded indisputable authority and whose shining light is still visible, was such a figure. The son of the eminent economist and statistician Aleksandr Ivanovich Chuprov, Professor at Moscow University, he was distinguished, in the words of his student Chetverikov, by an "unfailing distinctness of reasoning and inflexible and persistent rigor".

Chuprov developed his innate ability by an excellent education: in 1896 he had graduated from the mathematical faculty of Moscow University, then continued his studies in Germany, in the universities in Berlin and Strasbourg. The last-mentioned university attracted him by the possibility of collaborating with the Petersburg-born statistician Vladislav Iosifovich Bortkevich{= Ladislaus von Bortkiewicz}, 1863 – 1931. Georg Mayr and Knapp, who had been working there at the time, considerably influenced Chuprov. [1] It was the latter who occurred to be instrumental in the development of the special simple and elegant style of exposition that distinguished Chuprov.

During those years Chuprov had been investigating new approaches to the application of mathematics to social sciences. When becoming acquainted with the works of the Austrian economic school (Walras, Pareto a nd others) he gave the palm to statistics (itself considered a social science). Chuprov was one of the first to stress the universality of the statistical method of cognition.

In the autumn of 1902, at the invitation of A.S. Posnikov and M.M. Kovalevsky, Chuprov had moved to Petersburg so as to begin teaching at the Polytechnical Institute, the first Russian academic institute of the new type in which S.Yu. Witte {the Minister for Finance} was able to unite eminent scholars, and Chuprov contributed to the special intellectual atmosphere that was created there.

It was difficult and interesting to set the new undertaking into motion. A Statistical Room as a collection of fundamental works on statistics and as a center for scientific and pedagogic activities was contemplated. A circle of students comprising a school and calling themselves *Chuprovites* had gradually been formed. Among them were N.S. Chetverikov, G.S. Polliak, O.N. Anderson, the sisters M.M. and N.M. Vinogradova, B.I. Karpenko. The students' interest in science was kindled by Chuprov's lectures and supported by work in classes where special statistical cases were being examined, independent work experienced and a taste for statistical figures was being developed. For Chuprov, teaching and scientific work were merged which strengthened his magnetism. Thus, his study of Dmitriev's (1911) just appeared book together with his students led to the compilation of original contributions by them, – by M.M. Vinogradova (1916) and Anderson who generalized the Hooker variance-difference method.

In 1909 Chuprov published his splendid *Essays* where he justified the stochastic nature of the {social} world, the universality of the statistical method of cognition, listed the special features of the statistical science and explicated the main issues of the theory of mass phenomena. They were separated into four parts: Nomographic and idiographic sciences; Nomographic functions of the "categorical calculus" (the method of induction and the statistical method); Mathematical probability and statistical frequency (the law of large numbers); and Stability of statistical series.

His contribution fostered interest in the theory of statistics and in the philosophical substantiation of its ideas and definitively put Chuprov at the head of the statistical science in Russia. By that time, Russian scientists were prepared to perceive Chuprov's ideas. Owing to the works of A.I. Chuprov, Yu.E. Yanson, N.A. Kablukov and others, statistical methods became ever more disseminated. And the development of probability theory led to the appearance of scientific and popular literature on its application to statistical investigations (A.Yu. Davidov, V.Kosinsky, V.Ya. Buniakovsky, P.A. Nekrasov, P.L. Chebyshev {?}, A.A. Markov and others). However, the contents of the theory of statistics, its connection with mathematics were not determined; the question of whether statistics was a science or a method remained unanswered and the same was true with respect to the questions concerning the boundaries of its applications, of its interrelations with other branches of science, etc. The appearance of the *Essays* had thus brought about a considerable effect, suffice it to mention that practically all the leading Russian statisticians (V.K. Dmitriev, A.A. Kaufman, P.A. Nekrasov, A.A. Markov, Bortkiewicz, V.A. Kistiakovsky, A. Woden) reviewed them. Dmitriev (1909, p. 28) called that book an "outstanding achievement not only for our miserable [national – I.E.] literature, but also for the European literature on statistical science". The success of Chuprov's scientific work was followed in 1917 by his election to corresponding membership of the Russian Academy of Sciences, and in 1923 he became Honorary Fellow of the Royal Statistical Society.

Chuprov's main work of the next period of his activity was concentrated on the solution of the fundamental problem of rigorously justifying the statistical method by a synthesis of the ideas of the Continental (Lexis) and the English (Pearson) schools of statistics. In his *Essays*, Chuprov had only been approaching to the perception of this problem. In a later contribution (1924) he revealed the logic of statistical cognition. In the former case, as Bortkiewicz (1910, p. 372) remarked, he was unable to say "what exactly were the nomographic functions of statistics", but in the latter instance he indicated them (these included issues from the prior study of the properties of the general population to their posterior description by sampling and by estimating the conformity between posterior and prior indicators). He thus laid the foundation for the shaping of mathematical statistics, and I completely agree with Sheynin (1989, p. 3) in that "the birth of mathematical statistics as a single discipline that took place in the 1930s was much indebted to Chuprov".

Chuprov issued from the definition of the concepts of *random variable, law of distribution* and *dependence of trials*. The diversity of the conditions determining the stability of a population depends on whether or not the law of distribution persists, and whether or not the {pertinent} trials are mutually dependent. His theoretical constructions were being based on the application of the method of mathematical expectations, i.e., on the law of large numbers. The development of the fundamental problems of mathematical statistics achieved by Chuprov predetermined the subsequent progress of statistics and advanced him to one of the first few figures in the statistical science.

However, the real contribution of a scientist is determined by the existence of his scientific school {but how about Gauss?}. Chuprov's work at the Polytechnical Institute influenced many students and, among them, V.I. Khotimsky, Polliak, and S.G. Strumilin. But it was with Chetverikov and Anderson, to the highest degree congenial in spirit students and helpers, that Chuprov associated most frankly and closely. From among the staff he became the most friendly with Professor Vladimir Eduardovich Den (1867 – 1933) who contributed to economic geography as well as to political economy {economics}, population statistics and finance and who paid due attention to the reliability and completeness of the data, see for example Den (1929). The Statistical Room was closely connected with the Room of Economic Geography (and Anderson was librarian in both of them).

In May 1917, during his break, Chuprov had went abroad (to Sweden and Norway) and did not return home. Den became the custodian of the traditions of the Economic faculty and he (1926) at once responded to Chuprov's death (on 19 April 1926).

Chuprov's name continued to unite his former collaborators, even those finding themselves abroad. In a letter to Den, Anderson (1927b) wrote:

> *I would wish very much to know what is now happening in this Room* [in the Statistical and Geographical Rooms – I.E.] *where so many inventory pages are written in my hand and into which I also put a lot of work and effort. Still more I am certainly interested in the fate of our Polytechnical Institute, its present structure, the state of the economic department, the composition of its staff and the organization of the studies. Had the old spirit prevailing among staff and students persisted* (*at least in a small measure*) *and do the "new" people set as much store by their institute as we, the "old ones" of its first five classes, did*?

Anderson (1926 – 1927; 1927a) continued to study the statistical analysis of time series. In the same letter he mentioned his papers on the variate-difference method (approved by "Bortkiewicz himself") and the decomposition of series (where he argued against the representatives of the Harvard school of conjuncture (Persons) and touched on the Kondratiev theory of long waves). Anderson found it difficult to perceive that his works appeared after Chuprov had died and to accept the impossibility of obtaining his teacher's response. All his life Anderson (1887 – 1960) continued to develop the ideas of the Lexian – Chuprov Continental school as applied to the study of time series. He generalized the variate difference method of their investigation by calculating the first as well as the second, the third, etc difference of the levels themselves and of their deviations from the trend; he developed the stochastic interpretation of observations made for analyzing the series; and contributed to the investigation of distribution-free tests (Fels 1961).

Permanent spiritual and scientific connections had always existed between Chuprov and Nikolai Sergeevich Chetverikov (1885 – 1973). From the beginning of his career, the latter (1915 – 1916) became engaged in the history and application of the index-number method and in the analysis of time series. This direction found further development at the Conjuncture Institute and the Central Statistical Directorate (1919 – 1929). The main subject of Chetverikov's investigations was the crop capacity of cereals, the methods of its measurement and forecasting. By applying correlation methods, he (1963) eliminated the evolutional component from the data and discussed the fluctuations of the crop capacity in space with a preliminary grouping of the provinces into regions of "concordant oscillations". He investigated the connection between crop capacity of various cereals and determined to what extent will a bad harvest of one of them be compensated by a harvest of another one.

After being prevented from continuing his economic research because of the persecution of the staff of the Conjuncture Institute, – and then of geneticists among whom his brother, C.C. Chetverikov, played a prominent part, – N.S. began studying theoretical issues of statistics and the problem of applying radioactive elements in medicine.

After returning to Moscow {!} in 1959 Chetverikov had been wholly engaged in methodological issues of statistics and in the heightening of statistical culture. He prepared a three-volume set of Chuprov's main contributions which included a reprint of his *Essays*, of his book on correlation and his collected articles. Chetverikov (1968) also collected and translated articles of Lexis, Bortkiewicz, Chuprov and Bauer which explained the essence of the theory of stability of statistical series, and, together with Alb.L. Weinstein, translated Cournot's classic (1843). Cournot had not provided rigorous proofs of mathematical theorems, but he minutely described their usage in various branches of knowledge (astronomy, actuarial science, legal proceedings). Chetverikov was attracted to this contribution because Cournot had prophetically foreseen some problems that appeared before mathematicians and statisticians 40, 50 and even 100 years hence (Lexis, Bortkiewicz, and Fisher, respectively). The problems of forecasting really concerned Chetverikov, and, when solving them, he issued from the "inertia" of time series, – of the evolutional trend, or of the long or short "waves".

Even during the most difficult years of isolation {of Soviet scientists} from scientific contacts Chuprov's name connected his students with international science.

Chuprov kept an attentive eye on Isserlis' (1881 – 1966) work, on the man who translated all of his articles on the moments of distribution and correlation in case of three variables into English.

Boris Ivanovich Karpenko (1892 – 1976) was the last Chuprovite. He had experienced exile and concentration camps but did not lose his interest in science. He it was who saved the collections of the Chuprov Statistical Room. Together with Den he (1930) prepared a collection of articles on national economic statistics. He investigated financial statistics and statistics of prices, but his main contribution to statistical science was his posthumously published dissertation (1979) where he studied the forming of our ideas on the nature of mass phenomena and on the methods of their cognition, on the peculiarities of the manifestations of the law of causality in these phenomena, on the distinction between the theoretical (prior) and posterior probabilities and of its meaning for statistics.

Chuprov influenced not only his direct students; all those, who had been lucky to rub shoulders with him, experienced the charm of his whole-hearted personality. Among them were A.F. Ioffe, a physicist of international repute; L.I. Maress, an economist; E.E. Slutsky, a mathematician and economist. Ioffe (1928) stated that Chuprov was above all a scientist and that

His lectures constituted not a systematized description of established knowledge but conversations on the problems which he was studying … From life, he only demanded that it gave him the possibility of working.

Chuprov's school survived the persecutions of the 1930s and those that took place during the post-war period {after 1945} and proved its worth for the development of statistics. It would be wrong to say that it disappeared once its last representatives were gone. On the contrary, the re-kindling of the interest in statistics and its applications, unavoidable under the {new Russian} conditions of a market economy, will necessarily attract an intent attention to methodological problems of statistics and call into a new being the ideas of Chuprov and his students.

Note

1. {The reference to Mayr is hardly warranted. In a letter of 1898 Chuprov (Sheynin 1996, p. 16) wrote that Mayr "is […] unbearable".}

References

Anderson, O. (1926 – 1927), Über die Anwendung der Differenzmethode, etc. Reprinted in Anderson (1963, pp. 39 – 100).

--- (1927a), On the logic of the decomposition of statistical series into separate components. Reprinted Ibidem, pp. 101 – 122.

--- (1927b), Letter to V.E. Den of 25 October. Unpublished {whereabouts not specified}. (R)

--- (1963), *Ausgew. Schriften*, Bd. 1. Tübingen. Hrsg., H. Strecker et al.

Bortkiewicz, L. von (1910), The problems of scientific statistics. *Zhurnal Ministerstva Narodn. Prosveshchenia*, No. 2, pp. 346 – 372 of second paging. (R)

Chetverikov, N.S. (1915 – 1916), The method of index numbers as a means for investigating the change in the value of money. Reprinted in Chetverikov (1975, pp. 20 – 61).

--- (1963), *Статистические и стохастические исследования* (Statistical and Stochastic Investigations). M.

---, Compiler and Translator (1968), *О теории дисперсии* (On the Theory of Dispersion). M.

--- (1975), *Статистические иссдедования* (Statistical Investigations). M.

Chuprov, A.A. (1909), *Очерки по теории статистики* (Essays on the Theory of Statistics). M. Later editions: 1910, 1959.
--- (1924), Ziele und Wege der stochastischen Grundlegung der statistischen Theorie. *Nord. Stat. Tidskrift*, Bd. 3, pp. 433 – 493.
--- (1925), *Grundbegriffe und Grundprobleme der Korrelationstheorie*. Leipzig – Berlin.
Cournot, A.A. (1843), *Exposition de la théorie des chances et des probabilités*. Paris, 1984. Russian transl.: 1970.
Den, V.E. (1926), Obituary of Chuprov. Newspaper *Krasnaia Gazeta*, 23 April. (R)
--- (1929), *Источники важнейших отраслей хозяйственной статистики СССР* (Sources for the Most Important Branches of the National Economy of the Soviet Union). L.
Dmitriev, V.K. (1909), Theoretical statistics. *Kriticheskoe Obozrenie*, No. 6, p. 28. (R)
--- (1911), *Критические исследования о потреблении алкоголя в России* (Critical Investigations of the Consumption of Alcohol in Russia). M.
Fels, E. (1961), Obituary of Anderson. *Econometrica*, vol. 29, pp. 74 – 80.
Ioffe, A.F. (1928), Chuprov. Grand Meeting of the Council of the Econ. Fac. Commemorating Chuprov (1926). *Izv. Ekon. Fak. Leningr. Polytekhn. Inst*., No. 1 (25), pp. 315 – 316. (R)
Karpenko, B.I. (1979), *Развитие идей и категорий математической статистики* (Development of Ideas and Notions of Mathematical Statistics). M.
Sheynin, O.B. (1979), *Материалы о Маркове в архиве Чупрова* (Materials Concerning Markov in Chuprov's Archive). M.
Vinogradova, M.M. (1916), *Потребление водки в России и урожай*

17. I.I. Eliseeva.
Chuprov's Heritage (a Fragment)

Материалы конференции ... (see Item 11), pp. 15- 19)

It was apparently the Chuprov school through which his influence on the further development of statistics was being mainly felt. His students who proudly called themselves *Chuprovites* included N.S. Chetverikov (1885 – 1973), G.S. Polliak (1888 – 1954), O.N. Anderson (1887 – 1960), M.M. Vinogradova (? – 1919), N.M. Vinogradova (1889 – 1975), B.I. Karpenko (1892 – 1976), S.S. Kohn (1888 – 1933), and among those who attended Chuprov's lectures were V.I. Khotimsky (1892 – 1939), S.G. Strumilin (1877 – 1974) and L.V. Nekrash (1886 – 1949).

Chuprov directed his students and encouraged them; he supplemented their first publications (e.g., Polliak 1913; M.M. Vinogradova, 1916) by forewords. After finding himself abroad, he attentively followed the changes in Russian life, in national statistics (e.g., by studying the periodical *Vestnik Statistiki,* see Chuprov 1922) and stressed the importance of the contacts established between the members of the section of probability theory and mathematical statistics at the Moscow Institute of Mathematical Studies at Moscow University with representatives of practical statistics, – with V.M. Obukhov and Chetverikov in the first place. I mention this fact because, for many years from the end of the 1930s, practitioners and mathematicians had been working independently from each other.

Nowadays, our statisticians are faced with the problem of creatively applying the methods of mathematical statistics. The development of computerized treatment and analysis of statistical data is to some extent fraught with the danger of converting statistics into a purely formal branch of knowledge. The traditions of the Russian, and above all of the Chuprov statistical school enrich the statistical technique with logical constructions and secure conformity between the properties of mass phenomena and their reflexion by statistical means.

References

Chuprov, A.A. (1922), Westnik Statistiki 1920 – 1922. NST, Bd. 1, pp. 353 – 360.
Polliak, G.S. (1913), *Профессия как объект статистического учета* (Profession as an Object of Statistical Registration). Psb.
Vinogradova, M.M. (1916), *Потребление водки в России и урожай* (The Consumption of Vodka in Russia and the Harvest). Petrograd.

18. I.I. Eliseeva, A.L. Dmitriev.
Chuprov's Letters to D.A. Lutokhin

(Includes Lutokhin's previously unpublished recollections of Chuprov.)
Izvestia Petersb. Univ. Ekon. & Finansov, No. 2, 1997, pp. 112 – 118

Chuprov's name is especially dear to the scientists of Petersburg. Moving to our city in 1902 on the invitation of A.S. Posnikov and M.M. Kovalevsky for teaching at the then recently established Peter the Great Polytechnical Institute, he was able to create his school and to contribute to statistical science, – to the theory of statistical inferences, correlation, to applied problems of sampling and group building. Chuprov paid much attention to the specific character of mass phenomena, to statistical regularities, and to the importance of the statistical method.

Chuprov's letters to Lutokhin, published below for the first time, were typed on post-cards with handwritten corrections. They are kept in the manuscript section of the Institute for Russian Literature (the Pushkin House), Fond 592, Item 268, where one of us (A.D.) discovered them. Only at first sight, they contain nothing excepting information about everyday life and only testify to the difficulties experienced by a Russian scholar living in emigration.

Chuprov left Russia in May 1917. Since he previously had also spent summers abroad, this was not unusual. He went to Stockholm to continue his examination of birth-rate (at that time, he took great interest in the study of the sex ratio at birth and among still-births and foetuses). In December of the same year the Russian Academy of Sciences was to nominate Chuprov as its Corresponding Member[1] and it seemed that he firmly intended to come back, – but he did not, at first because of an illness, and then, apparently, owing to the events in Russia. His attitude towards Soviet Russia had never been clear. In the 1920s, students at the Polytechnical Institute, who came to believe in the authority of their self-government and earnestly requested him to return, were sincerely distressed when they learned that he died on 19 April 1926.

Sheynin (1996, p. 128) indicated that Chuprov had negatively regarded Soviet Russia. He discovered a political pamphlet kept in the *Bibliothèque Nationale de France* {reproduced in this collection} whose author had declared that the ideas of Bolshevism were dead and dethroned Lenin. The pamphlet was signed *A. Tchouprov, Professeur d'Economie politique à l'Université de Moscou.*

It was Chuprov's father, Professor A.I. Chuprov, who had worked in Moscow University, but he died in 1908. Taking into account Chuprov's reserve of formulated opinions so characteristic of him and his position of observer rather than accuser of events in Russia, his authorship is also doubtful. Chuprov's colleagues and students, – and, above all, N.S. Chetverikov, the most beloved from among them, – were remaining in Russia and he had been maintaining personal and scientific contacts with them and was keeping an eye on Russian statistical publications, witness his review (Chuprov 1922b). At its beginning, he remarked that the development of statistics, that had begun in February 1917, was obviously continuing and this positive assessment contradicted the opinion current in the émigrés' circles.

At the same time, there is no reason to believe that Chuprov was prepared to return, as it would appear from an official anonymous document {translated in this collection}. There, he is described as being prepared to return with lack of money being the only obstacle for the move. We think that the author of that document was V.E. Den, Professor at the Polytechnical Institute, for whom its compilation was a possible step for securing himself against accusations of corresponding with a "defector".

There were no accusing political pamphlets, but an ardent desire to return to Soviet Russia was also absent. What existed was a life occupied by science and everyday disorder. In his fourth letter to Lutokhin Chuprov quite clearly formulated his attitude towards Russia: he obviously had not thought about returning home. But he maintained connections with his mother country. By publishing his paper (Chuprov 1924) devoted to the aims of theoretical statistics in the Soviet *Vestnik Statistiki*, he stirred up upon himself a wrathful storm. This apparently was the cause that sped up his premature death.

A few words about Dalmat Aleksandrovich Lutokhin (1886 – 1942). In 1903 he entered the Petersburg Polytechnical Institute. Was a student of the Paris school of social sciences, Berlin University, and Handelhochschule Berlin. In 1908 he went over to the Law faculty of Petrograd University and, in 1911, to the same faculty in Kharkov. Graduated from there in 1912; his candidate composition was entitled *S.Yu. Witte as Finance Minister*. After that, having secured the support of the Society for Financial Reforms and of V.I. Timiriazev, went to London for investigating issues of economic policy. After returning to Russia in 1914, he worked in various banks and in Petrograd municipality, and, from 1918, as expert in finance in the Directorate of Paper Industry. In January 1921 he was elected assistant chairman of the Russian Technological Society where he organized a number of sittings on the socio-economic after-effects of the war and the revolution. Taught economic geography at the Petrograd regional foodstuffs courses, actively participated in establishing the periodicals *Ekonomist*, *Ekonomicheskoe Vozrozhdenie*, *Utrenniki* and, from the end of 1921, directed the journal *Vestnik Literatury*.

In August 1922 the Petrograd *Cheka* (Special Committee for Combating Counter-Revolution and Sabotage) arrested Lutokhin and kept him in their cellar until February 1923. Then he was exiled from the Soviet Union. At first lived in Germany, then in Prague (more precisely, in Horni Mokropsy, a settlement near Prague). Sympathized with the Soviet Union and belonged to the "loyal opposition" with which Chuprov, as it follows from his letter, was not quite in agreement.

In 1927 Gorky helped him to return home. Lived in Leningrad, worked in various state bureaus and prepared his never published remembrances of emigration, – *У тех, кто бежал революции* (With Those Who Ran away from the Revolution) and *Пастыри зарубежные* (Our Émigré Preachers). Below, we incorporate an extract from the second source. After the assassination of Kirov {in 1934, possibly orchestrated by Stalin and in any case followed by a new wave of repressive measures against the *enemies of the people*} was exiled to Ufa. Died in the besieged Leningrad.

[Letter No. 1] Dresden, 8 March 1923

Highly respected Dalmat Aleksandrovich,

I shall be very glad to see you in Dresden. I myself do not intend to come to Berlin during the next weeks. And in general, as far as I can foresee, I shall remain in Strehlen {on the map of present-say Dresden, this is the name of one of its small parts}, so that you will almost surely find me here. But it would nevertheless be better to exchange letters beforehand because it still happens that I am away for some time. I advise you to weigh the Prague combinations {possibilities} very carefully. Life is extremely expensive there and housing is even more difficult than in Germany. It is impossible to live on a stipend {there} with a family, it would be necessary to earn extra money, and, as it is said, – to earn more than necessary for living in Germany without a stipend. In any case, I definitely advise you not to move at once with your family. At first, try it out on the spot leaving yours for the time being in Berlin. If you will attempt to secure a stipend, settle it with P.B. Struve[2] who will then accept your Master's examination. […] There is a Russian colony in Dresden, but I am standing aside and seeing no-one. As I heard, it would be hardly interesting for you to deliver a report here. There are two associations: one of them, very right-winged, and you will hardly go there since you are attracted to the same direction as Sor[okin][3] is; the other one is a Russian- German union of moderately inclined figures, they do organize reports, but are more interested in musical entertainment and dances. […] I have heard about the *Ekonomist* [4] but did not see even a single issue. In general, only a very small portion of Russian literature reaches me. All the best.

[**Letter No. 2**] Dresden, 27 April 1923

[…] A position of head of the Russian economic section in the *Ost-Europa Institut* in Breslau {present-day Wroclaw, in Poland} is being vacated. I heard about it today from Berlin and I am informing about it some of my students with whom I corresponded in this connection some time ago when such rumors had appeared. However, all those whom I had in mind are living outside Germany and it is rather probable that they would not want to move to Br[eslau]. What would be your attitude to such a combination? […]

[**Letter No. 3**] Dresden, 30 April 1923

[…] It is very consoling that even two comparatively suiting you combinations are beginning to show. Under these conditions Breslau is certainly not interesting. If M.V. Bern[atsky][5] will be able to secure for you an initial salary of 900 francs, I would advise you to move to Paris but at first without family […] And move your family after establishing yourself more firmly and familiarizing yourself with the situation in Paris. In the same way, even if Sor[okin] secures for you a stipend of one and a half thousand, which I consider not very probable, I would not advise you to move at once to Prague with your family. […]

[**Letter No. 4**] Dresden, 9 March 1924

[…] It is sad that you are still unable to balance your income and expenditure, but you can take comfort in that it would have now been more difficult had you remained in Berlin. The shift in the comparative cost of life is indeed striking! True, the prices that soared up beyond measure at the end of the previous year have begun to lower gradually, but on the other hand one item of expenses, and for that matter one of the most essential, especially for the likes of us, is still sharply raising: landlords are vigorously drawing the rent nearer to the prewar levels, and this brings to naught the reduction in the prices of foodstuffs, coal, etc. Until now, I am still dodging in my Strehlen rather successfully, but it is difficult for those less luckily settled down.

Returning from Italy at the end of November, I began preparations for a new tour; this time I am going North to the Scandinavian countries. I received an invitation to read a short course (six lectures) in Christiania {Oslo}, and, besides, to deliver there a report at the society of actuaries. I shall also read it en route in Copenhagen and perhaps in Stockholm on my way back. I am going at the end of April and will be absent for about a month, as I suppose. The tour as a whole interests me, and it is being paid for to such an extent that not only it itself but also the time spent in preparation will partly be compensated. In addition, I expect to get a little bit more if I submit the prepared material for publication; in this case, all my expended work will apparently be covered. Until now, I am writing everything in German, but hope to recast it also in Russian.[6] The construction of the loyal opposition about which you are writing {me} is in essence interesting but I am afraid that for a long time it is doomed to remain utopian: judging by what reaches us from Russia, the state customs there are not yet suitable. In general, even now life is there difficult for everyone excepting NEPmen {NEP = New Economic Policy, 1921 – ca. 1936. It meant some economic but certainly not political freedom} and specialists of the highest caliber. And breathing there is even more difficult for the likes of us. All the best.

Chuprov[7]
D.A. Lutokhin

Chuprov helped me to organize my report at the Russian colony in Dresden where he lived in 1923, and at the same time I visited him – and the Sistine Madonna. He was very well known among European statisticians, chairs in several foreign universities were offered him, but he was absorbed by scientific work and teaching would have demanded much of his time. “I can allow myself a glass of beer once in four months”, he told me, “but as soon as I begin reading lectures in Heidelberg, where I am invited, I would have to maintain acquaintance with the professorial staff”. And he preferred to lead a half-starved existence but totally devote himself to theoretical work. He was not a politician, but enjoyed great authority in public circles and attentively followed life in Russia. And the behavior of his colleagues embarrassed him. Chuprov told me that he had given his friend P.B. Struve a manuscript for *Russkaia Mysl* (Chuprov 1922a), but, after receiving its {appropriate} issue, became disconcerted and candidly informed P.B. that he will not continue publishing his papers anymore in such a milieu. And, retaining his good relations with Struve as with a comrade in scientific work, did not attend his jubilee. “Were you afraid to break a fast?”, asked him P.B. afterwards.

Chuprov was not an adherent of Bolshevism at all, and not a socialist either. He was a sincere liberal, to a large extent true to his father’s economic ideals, i.e., a liberal inclined towards the *Narodniki* {populists}[8]. But he loved Russia and did not become a reactionary and all positive for him in the revolution he indeed called positive. For example, Chuprov highly appreciated the Soviet organization of statistics and believed that in this respect Russia was now occupying the third place among cultured nations after the United States and England. Mentioning this in his lecture devoted to the correlation theory read in Prague in March 1925 [9], Chuprov expressed his desire that statisticians returning to Russia be on the same level as those who had remained and were working there.

He had published a paper in the Moscow periodical *Vestnik Statistiki* (see Chuprov 1922b) and his colleagues therefore defamed him. Upon moving to Prague, he intended to be nominated for the academic association, but the reporter who should have named his candidature, refused to do so after finding out about the paper in the Soviet journal. The “sanctions” did not stop there. The Czechs had not given him a professorial stipend and he had to accept for the time being a modest sinecure offered by Prokopovich.[10] This happened not long before his death. When travelling to the statistical congress in Rome, he fell ill en route and availed himself of an invitation to stay for a while in Géneva, by Gulkevich, the former Russian ambassador in Stockholm. There it was that he died.[11]

He had not thought of dying so soon. I remember our last conversation. He came to me in the Czech hamlet and was telling me about his father who lived to be 75 {a mistake, see Note 12}. Not long before his death, A.I. Chuprov[12] for the first time had to have his tooth extracted – a healthy tooth since his maxillary sinus became inflamed and it was impossible to penetrate it without sacrificing the tooth. I am afraid, as A.A. said, that I shall also have to undergo such an operation sometime, and he smiled widely screwing up his mouth and exposing his strong white teeth.

The physicians were at a loss for understanding what disease had suddenly weakened his heart to such an extent that he had to be kept in bed from which he never got up. It is indeed possible that the reception rendered him in Prague severely wounded Chuprov who was so delicate and sensitive to falsehood… Health was not sufficient for struggling against the "serried majority", a certain callousness was also needed but his heart occurred to be easily wounded …

Notes

1. He was elected on 29 November 1917 {see Struve et al in this collection}.

2. Petr Berngardovich Struve (1870 – 1944), an economist, philosopher and politician. In 1906 – 1917 taught at the Petersburg Polytechnical Institute. Master (1913) and Doctor (1917). In 1917 elected Full Academician of the Russian Academy of Sciences. Member of Denikin's "Special Conference" and of Wrangel's government. {A.I. Denikin, 1872 – 1947, and P.N. Baron Wrangel, 1878 – 1928 (see below), leaders of the White Movement opposing the Bolsheviks in the Civil War in Russia, 1918 – 1922.} In emigration since 1920 (Prague, Belgrade, Paris). In Prague, he was Professor at Russian Law Faculty (1922 – 1925). In 1908 – 1916 edited the periodical *Russkaia Mysl* (suppressed by the Bolsheviks in 1918). Struve resumed its publication in Sofia, then in Prague.

3. Pitirim Alkeksandrovich Sorokin (1889 – 1969), a sociologist. Graduated from the Law Faculty of Petrograd University (1914), Privat-Dozent from 1917. In 1914 – 1916 taught at the Psychoneurological Institute, in 1917 Kerensky's secretary for scientific issues. In 1920 – 1921 head of chair of sociology at Petrograd University. In 1922 exiled from the Soviet Union, lived in Berlin. {The Declaration on the establishment of the Soviet Union was signed 30 Dec. 1922 which means that Sorokin (and several other people, see below) were rather exiled from Soviet Russia.} Being invited by his friend Masaryk, the President of Czechoslovakia, to head the national Institute of Sociology, moved to Prague, then, in 1923, came to the USA. Professor at Harvard University.

Sorokin was friendly with Lutokhin, and, having found out about his exile, invited him to Prague. Their friendly relations had not however lasted long. The final break occurred after the latter published a sharp paper in the *Socialist-Revolutionary* journal *Volia Rossii* concerning Sorokin's article *Sociology of revolution*. Sorokin's letters to Lutokhin are published, see Two fates (Sorokin and Lutokhin. *Vazhsksaia Oblast*, No. 6, 1992, pp. 17 – 21.

4. *Ekonomist*, a periodical of the industrial-economic section of the Russian Technological Society published in Petrograd in 1921 – 1922. Its editorial staff included […] Lutokhin […] and others. It was suppressed after the appearance of Lenin's generally known article *On the importance of militant materialism.* Lutokhin himself (1923, p. 163) recalled:

> *When circulating the Ekonomist to various figures so as to secure comments, I thought it necessary to send a copy of the first two issues to Lenin.* […] *Lenin* […] *in 1922 burst out into an article against the Ekonomist and indicated that a young communist had brought him that periodical with enthusiastic comments, but that he allegedly either did not read it properly or did not understand it. Lenin had not mentioned that the journal was sent him by its editorial office. According to his opinion, the periodical deliberately or otherwise pursued serfdom tendencies and he advised the likes of the members of the Ekonomist staff to visit Western Europe so as to feel on their own hides all the delights of that "democracy" which they wish for Russia.* (*Already then* {?} *Lenin was unable to distinguish between serfdom and democracy…*) *Lenin wrote that article while on his sick-bed, probably durin a respite, and it was his political testament of sorts: keep an eye on "those advocating serfdom from among the intelligentsia". The GPU* {more correctly, OGPU, the predecessor of the notorious KGB} *had taken Lenin's article into account and the preparation of the lists of those, whom it was necessary to acquaint with the "delights of democracy in the West, had begun.*

5. Mikhail Vladimirovich Bernatsky (1876 – 1943), an economist and politician. Taught at St. Vladimir Kiev University from 1904, at Petersburg Polytechnical Institute as Dozent and then Professor. In 1917, head of section of labor at the Ministry for Commerce and Industry, Assistant Finance Minister, and, from September, Finance Minister. In 1918 – 1920, Finance Minister in Denikin's and Wrangel's governments. From 1920 lived in Paris. In 1924 most actively participated in the work of the economic section of the Russian Institute for Law and Economics at Paris University. In his letter of 17 May 1923 Bernatsky informed Lutokhin that it was practically impossible to find a position in Paris. […] (Inst. Russ. Lit., Fond 592, Item 73, No. 1).

6. Chuprov apparently had in mind his work (1925).
7. *Пастыри зарубежные* (Our Émigré Preachers), c. 61 – 62.
8. Lutokhin had in mind Chuprov's contributions devoted to village communities (1904; 1906a; 1906b).
9. Chuprov's lecture read for those attending the courses at *Prokopovich's Economic Room*.
10. Sergei Nikolaevich Prokopovich (1871 – 1955), an economist, statistician, politician. Doctor of philosophy (1913), head of the economic section of the Free Econiomic Society, chairman of section for insurance of working people at the Russian Technological Society, member of other learned societies. In 1917 chairman of the Main Economic Committee, assistant chairman of the Economic Council of the Provisional Government, Minister of Commerce and Industry, Minister of Foodstuffs. In 1922 exiled from the Soviet Union. Lived in emigration in Berlin, Prague, Géneva. In Berlin, and then in Prague established the *Prokopovich Economic Room.* In the 1920s – 1930s directed the periodicals *Ekonomichesky Vestnik* and *Russkiy Ekonomichesky Sbornik.*
11. Also see Hessen (1979, pp. 14 – 15):

> *Among those at the Russian legation was the former Professor at the Polytechnical Institute, and an eminent scholar A.A. Chuprov. He had been engaged in compiling financial-economic bulletins on the international situation for the legation. After an attentive study he had transferred his own small sum of money inherited from his father into cheap German marks which later lost all their value. The Imperial legation was then abolished and Chuprov moved to Dresden where he definitively buried himself in science and only violated his isolation for the sake of Gulkevich becoming ever closer to him. Together with Gulkevich he then moved to Switzerland when Nansen had taken the former on as his assistant for Russian refugees. And Chuprov, who had always seemed to be a picture of health, prematurely died of an insidious and for a long time undiscerned cardiac disease.*

{It is curious that Chuprov so mistakenly managed his inheritance. In 1920, the League of Nations appointed Nansen (an Honorary Member of the Petersburg Academy of Sciences and a Nobel peace prize winner, 1922) High Commissioner responsible for repatriation of prisoners of war.}
12. Aleksandr Ivanovich Chuprov (1842 – 1908), an economist, statistician and public figure. Professor at Moscow University in 1878 – 1899. Corresponding Member of the Petersburg Academy of Sciences (1887).

References

Chuprov, A.A. (1904), Community landownership. In *Нужды деревни* (The Needs of the Village), vol. 2. Psb, pp. 116 – 132.

--- (1906a), *Конституционно-демократическая партия и социализм* (The Constitutional-Democratic Party and Socialism). M.

--- (1906b), The land reform and the present leaseholder. *Poliarnaia Zvezda* (weekly. Psb), No. 9, Febr. 10, pp. 635 – 642. (R)

--- (1922a), The problem of indeterminism in the light of statistical physics. In Chuprov (1996, pp. 47 – 55).

--- (1922b), Westnik Statistiki, 1920 – 1922. *Nordisk Stat. Tidskrift*, Bd. 1, pp. 353 – 360. Russian transl.: *Voprosy Statistiki*, No. 1, 1999, pp. 11 – 13.

--- (1924), Ziele und Wege der stochastischen Grundlegung der statistischen Theorie. Ibidem, Bd. 3, pp. 433 – 493. Enlarged Russian versions: 1925 (two sources), reprint: Chuprov (1960, pp. 162 – 221).

--- (1925), *Grundbegriffe und Grundprobleme der Korrelationstheorie*. Leipzig – Berlin. Russian transl.: M., 1926 and 1960. Engl. transl.: London, 1939.

--- (1960), *Вопросы статистики* (Issues in Statistics). Collected articles in orig. Russian or translations into Russian. M.

--- (1996), *Материалы конференции к 70-летию со дня кончины* (Materials of Conference on the Occasion of 70 Years after His Death). Psb.

Hessen, I.V. (1979), *Годы изгнания. Жизненный отчет* (The Years of Expatriation. An Account of My Life). Paris.

Lenin, V.I. (1922), The importance of militant materialism. *Полное собрание сочинений* (Complete Works)

Lutokhin, D.A. (1923), Soviet censorship. (From my political reminiscences). *Arkhiv Russkoi Revolutsii* (Berlin), Bd. 12, p. 163. (R)

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

19. I.I. Eliseeva, A.I. Dmitriev. Chuprov

(Fragments).

Статистики русского зарубежья (Statisticians of the Russian Diaspora).Petersburg, 1998, pp. 77 – 89.

Aleksandr Aleksandrovich Chuprov was born 6 (18) February 1874 in Mosalsk, Kaluga province. Soon afterwards his family moved to Moscow. He grew up having been cared for by all the women around him, – by his mother Olga Egorovna, by his aunts, grandmother and the three sisters, – Olga, Maria and Elena. As the years went by, his father, a well-known professor at Moscow University, an economist and statistician Aleksandr Ivanovich Chuprov, was ever more influencing him. […]

G.K. Fediaevsky (1926), a graduate of the first class of 1907 of the Polytechnical Institute, published an extract from Chuprov's letter to him dated 29 Dec. 1924. Chuprov wrote there:

> *I myself passed 1924 just like the previous year: I sat at my desk in Dresden almost in complete solitude and worked indefatigably. There was however one interruption: in the spring I went on tour to Scandinavia. I read a small course in the theory of correlation in Christiania* {Oslo} *and Copenhagen. My trip was very fortunate. I was cordially welcomed and honored as a distinguished guest. And my lectures were also very well received. My audience was not very numerous but interesting for me: several professors, statisticians, actuaries and student-actuaries. No greenhorns there, all of them with scientific and practical experience. I was able to find the proper note, to offer them in my course just what they wanted. The remaining part of the year I was mostly engaged in trimming my course for publication, both in German, as read by me, and in Russian* […]
> *In January my comfortable life in Strehlen*[1] *will come to an end. After the German mark had stabilized, life became so expensive that it is impossible to earn enough by scientific and semiscientific writing. During the year my small emergency fund dried up so much that it is time for reconstructing life on some other basis, for returning to teaching. For the time being, I am moving to Prague to teach there at the Russian Law Faculty. Then time will show …*

Lectures read in Oslo formed the basis of Chuprov's book (1925) whose Russian edition (1926a) became possible owing to Chetverikov. Chuprov greatly influenced the development of statistics in Scandinavia; he published many {nine} articles in the *Nord. Statistisk Tidskr.* established in 1922 {and two papers in the *Skand. Aktuarietidskr*.}

Also in 1926 Churov published a fundamental paper where he actually defined the development of statistics in the 20th century. Its appearance was also engendered by his reports in Scandinavia. In the first part of the paper Chuprov stressed the connection between statistics and the theory of probability and formulated the problem of determining not those random numbers {magnitudes} "directly given by calculations" but rather the "real" quantities characterizing the studied phenomena.

Chuprov regarded probability theory as a philosophical and logical construction rather than as only a mathematical discipline. Consequently, for him mathematical probability, random variable, stochastic connections between variables were the initial notions of the stochastic theory of statistics.

The central proposition of the paper was a clear distinction between prior and posterior (empirical) laws of distribution and connections. "The similarity between the two systems is so great that there easily appears a temptation to restrict the necessary efforts to studying only one of them without spending time on constructing the other one", Chuprov (p. 184) stated and formulated the following problem (p. 189):

> *The stochastic theory of statistics should indeed help here by showing how could the numerical values of prior summary characteristics of the laws of distribution and connection be more or less precisely determined on the basis of these experimental data.*

Chuprov warned statisticians against pure empiricism and stressed the importance of the law of large numbers. And he had always been interested in the theory of sampling. He anticipated some of Neyman's (1903 – 1957) results, in particular, his well-known formula for the optimal distribution of a fixed size of sampling between strata (Seneta 1982).

Chuprov provided approaches to eliminating the "corrupting influence of chance" in the statistical data:

1) In cases of complete enumeration the influence of a concrete period ought to be eliminated; the data here should be regarded as "sampling in time", as a possible realization.

2) In cases of sampling from a really existing large general population (e.g., a sample of water from a river or sea).

3) In treating experiments (e.g., when testing a new medicine).

This attitude towards initial data led Chuprov to the development of the method of expectations for proving some limiting relations. It is in this vein that he studied the limiting behavior of the distribution of the coefficient of dispersion and calculated the central moments of arbitrary orders so as to determine the limiting distribution of the arithmetic mean of the values of a random variable obeying an arbitrary discrete distribution.

Chuprov held that, among the diverse problems solved by statistics, the discovery of connections between phenomena occupied one of the most important places with regard to its scientific and practical meaning, and he studied the nature of stochastic, and in particular of correlational connection. Already in his *Essays* (1909), and then in his book (1925a) issuing from the logic of the inductive method he explained the existence of "incomplete", "free" connections by the multiplicity of causes and effects. The differences between the complexity of the phenomena, between the "elementary causes" that determine their existence lead to the variablity of the connections between them. Suppose that there exist laws according to which a always follows A, b follows after B, c after C, … Assume also that there exist three phenomena

$$y_1 = a, \; y_2 = a + b, \; y_3 = a + b + c.$$

Obviously, y_1, y_2, and y_3 are influenced by one and the same cause A, but they differ owing to the influence of the additional causes B and C. A similar circumstance takes place with regard to the interconnections between the three phenomena. […]

The concept of stochastic connection is contrary to that of functional connection. When describing the types of relationship between variables, Chuprov was apparently the first to pay attention to the fact that a variable can be stochastically connected with a number of other variables separately, but at the same time be functionally dependent on these variables taken together. This remark reveals the possibility of investigating connections of such a type which are often encountered in economics. For example, the receipt from selling a commodity is equal to the product of the number of pieces sold and the price of a piece. This connection is functional but both the factors are connected stochastically and the same might be stated about the connection between receipt and the price or between it and the number of pieces sold.

Chuprov also developed the theory of stochastic series. […] Employing the method of expectations and invariably distinguishing between prior and posterior characteristics when developing the theory of stability of series, Chuprov concluded that the Lexian criterion of stability was not universally applicable and indicated the pertinent restrictions.

Chuprov (1925b) attempted to determine the real sex ratio by issuing from the life of human embryos in the womb and indicating that the available data testified with absolute certainty that male embryos were subject to higher mortality than their female counterparts. And since male births are more numerous, it followed that male conceptions were much more numerous than female conceptions (Chetverikov 1959).

Chuprov indicated that abortions, which were officially allowed in Leningrad since 1923 in accord with special procedures, provided an unique possibility for research. The registration of abortions regrettably concentrated on describing the women applying for the permission and the causes leading them to the abortion rather than on the age or sex of the embryos. He noted that such missing data would have provided the most reliable and direct statistical data on the sex of the embryos; indeed, neither the causes for abortion, nor its permission or refusal were connected with that sex so that the missing data could have been considered a representative sample.

After having created his school at the Petersburg – Petrograd Polytechnical Institute, Chuprov had been resisting to teach. Anziferov (1926) stated that

Under Russian conditions, the notion "*scientist*" *has almost always been* *connected with the concept of* "*instructor at an academic institution*". *The late Chuprov was also obliged to devote many years to teaching*. […] *But his natural inclinations always led him mostly to purely theoretical research* […]

Indeed, Chuprov had been engaged in an intensive intellectual work whose fruit can now be properly appreciated. His election, in 1923, to honorary membership of the Royal Statistical Society was an indication of acknowledgement of his merits.

Chuprov had no family of his own. His cousin, the writer A.V. Amfiteatrov (1862 – 1938) (1926) recalled that "Even in his early youth he exhibited something spinsterish, and looked as though being destined for a solitary path throughout life". […]

{By 1926} nine years had passed since Chuprov left Russia, but he was not considered an émigré. Obituaries were almost immediately published in one of the Leningrad papers {(Den 1926)} On May 23, 1926, the Executive Commission of Russian Statistical Coingresses held a commemorative sitting where Chetverikov reported on Chuprov's work and indicated that it was of special importance in the field of applying mathematics to social sciences. E.E. Slutsky described Chuprov's contributions and N.N. Vinogradova examined his pedagogic activities. The Commission expressed its condolences on the occasion of Chuprov's death, suggested that his {collected} works and a memorial volume be published and addressed its request to the Academy of Sciences and the Central Statistical Directorate repectively.

On May 30, 1926 a Grand Meeting of the Council of the Economic Faculty of the Leningrad Polytechnical Institute took place, again on the occasion of Chuprov's death (Izvestia 1928). At the time, those who called themselves *Chuprovites* and had the privilege of direct contacts with Chuprov were still alive. Their reports at the Meeting allow us to appreciate the scale of Chuprov's personality, his contribution to science and his special style of working with young scientists.

Chuprov's death united émigré Russian statisticians and economists. On April 27, 1926 a public sitting devoted to his memory was held by the Prokopovich Economic Room in Prague (Russk. Sbornik 1926).

Notes

1. {At the time, a settlement near Dresden; nowadays, a part of that city.}

References

Amfiteatrov, A.V. (1926), An extinct family. Newspaper *Vozrozhdenie*, 23 April. (R)

Anziferov, A.N. (1926), To the memory of Chuprov. Ibidem. (R)

Chetverikov, N.S. (1959), Chuprov as a demographer. In author's book *Статистические исследования* (Statistical Investigations). M., 1975, pp. 244 – 260.

Chuprov, A.A. (1909), *Очерки по теории статистики* (Essays on the Theory of Statistics). Later editions: 1910 and M., 1959.

--- (1924), Ziele und Wege der stochastischen Grundlegung der statistischen Theorie. NST, Bd. 3, pp. 433 – 493. Enlarged Russian version quoted in text: 1925.

--- (1925a) *Grundbegriffe und Grundprobleme der Korrelationstheorie*. Leipzig – Berlin. Russian edition: 1926 and 1960.

--- (1925b), Der behördlich genehmigte Abort in Leningrad. JNÖS, 3. Folge, Bd. 68, pp. 698 – 701.

F[ediaevsky], G.K. (1926), To the memory of Chuprov. Newspaper *Vozrozhdenie*, 31 May. (R)

Izvestia (1928), *Izvestia Ekonomich. Fak. Leningradsk. Politekhnich. Inst*., No. 1 (25), pp. 307 – 350. (R)

Russky Sbornik (1926), *Russk. Ekonomich. Sb*., No. 6. Prague, pp. 5 – 15 and 15 – 33. (R)

Seneta, E. (1982), Chuprov. *Enc. Stat. Sci*. Ed., S. Kotz. New York, vol. 1, p. 477.