Studies in the History of Statistics and Probability

vol. 23

Berlin
2020

Contents

Oscar Sheynin is the author of items i – v and vii

oscar.sheynin@gmail.com

Notation

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

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

I

G.-L. Leclerc de Buffon

Essai d'Arithmétique Morale, 1777

Essays on Moral Arithmetic
Translators J. D. Hey, T. M. Neugebauer, Carmen M. Pasca
www.Isf.lu/eng/Research/Working-Papers/2010

Introduction

In his classic work, Buffon discusses degrees of certainty, probability, the moral value of money, different evaluations of gains and losses; moreover, he proposes repeated experiments to determine the moral value of a game. We are first to translate this work completely.

Buffon (1707-1788) was acknowledged, and is mainly remembered, for his *Natural History*, of which during his lifetime he finalized 36 volumes. Another eight volumes were published posthumously.

His scientific work was generally based on the methods of empirical observation and experiment. Based on his evidence, he suggested that the organisms originated by spontaneous generation from smallest particles and that their development and diversity were due to climatic changes. Differing from the common view at his time, he believed that life first developed in the sea and that the stepwise development of the species occurred during long periods.

Preceding Darwin, he advanced the view of common descent of the species and discussed the relation between apes and man. By experimental evidence he showed that the earth was older than the 6,000 years calculated by theologians. This evidence led to a conflict with the church including the burning of his writings. He had to revoke officially his statement, but raised the point again in the *Essays*.

During the 1720s and 1730s, before dedicating himself to natural history, Buffon studied and contributed to mathematics. From 1727 (Weil 1961), he communicated, and maintained correspondence with Gabriel Cramer (1704-1752), professor of analysis and geometry at the University of Geneva. So, while it is likely that he started editing the *Essays* in the 1760s, his interest in the topic and many of his considered ideas date back to the early 1730s. The *Essay* contains 25 articles, the first two of them are introductory.

As he pointed out, the work is dedicated to the measurement of, and, more generally, to the valuation of uncertainty. Buffon defines levels of uncertainty or certainty by available evidence observed in nature. And generally, if the causes of a certain effect are unknown, but the effect is constantly repeated, he believes that it becomes physically certain. Conversely, if an effect has constantly failed to occur, he suggests that it is refuted.

Buffon also points out that a change in an assumed constant effect surprises us. Such effects have captured the attention of a great audience (Taleb 2007). He illustrates the concept of levels of certainty by the phenomenon of sunrise. He reflects on how an ignorant man sees the sunrise and sunset and learns to reinforce his belief (and decrease his doubt) by repeated experience, and reach certainty about the return of the sunrise (articles III-VI)[1]. He refers to physical certainty[2] as an "almost infinite" probability level to which he assigns a relationship as of one to $2^{2,189,999}$. The exponent is the number of days following the first day in 6,000 years which represent the accepted time of existence of man and 2 means that either the sun returns on the next day or not.

If one has only a limited number of observations of a constant effect so that physical certainty cannot be inferred. Buffon argues that uncertainty can be so much removed if only the number of experiences is sufficiently large, moral certainty about the effects can be achieved.

Moral certainty can be considered as a bounded rational judgment level, sufficiently high to draw conclusions on the certainty of constant effects. Given the limitations of time and resources on the number of experiences (observations), moral certainty is a substitute for physical certainty that enables to make such a judgment about the effects of nature even without understanding its causes. In Buffon's thinking, moral certainty implies a lower degree of certainty than physical certainty, and can be determined either by evidence, following a constant sequence, or by analogical reasoning based on testimonies of a constant sequence. The probability level that Buffon assigns to complete moral certainty is (1 – 1/10,000) (articles VII-IX). This level is based on the argument that a fifty-six year old man is fearless about dying during any given day, an event which according to his mortality tables, reported elsewhere (Buffon 1777A), occurred with the corresponding relative frequency.

Buffon refers also to different levels of probability for the case when the underlying effect is not constant. Discussing such chance effect in more detail, Buffon focuses on the 50-50 game, where an effect occurs as often as it fails. He suggests that, even in these games, the observation of a large number of results from a random device can give us an advantage in the game, if that device is biased (articles X-XI). Thus, he indicates that the underlying probabilities can also be learned from long series of observations, and can deviate from prior probabilities.

Generally, Buffon condemns participation in games of chance, judging them as generally harmful to overall well-being in society (article XII). In particular, he refers to the contemporary popular game of Pharaoh, apparently a forerunner of Poker. He argues that people behave dishonestly and irrationally with respect to such games. His judgment is based on: (1) a value function approach; and (2) a classification of individual income as either necessary or superfluous (articles XIII-XIV).

(1) His value function exhibits loss aversion, since it punishes losses more than it rewards gains. Gains are valued relatively to

income including the gains; losses are valued relatively to the prior income state, before subtracting the losses. Thus, since losses loom larger than gains, overall well-being is reduced even in fair games. Compared to the logarithmic utility function proposed by Daniel Bernoulli (1700-1782) which was published in 1738, losses are similarly valued but the Buffonian value function values gains less than losses. Loss aversion is quite accepted among modern researchers. For instance, Selten and Chmura (2008), in their application of the impulse balance theory to experimental data, assign a double weight to a loss as compared with gain. Such weighting was suggested by experimental evidence in Tversky and Kahneman (1992). We should mention here, even without further discussion, that modern loss aversion decisively differs from the Buffonian version in several respects. However, Buffon appears to have been the first to propose loss aversion within a utility approach. Unfortunately this approach was not embedded in a fully-fledged and fully-developed theory.

(2) He nevertheless suggests the existence of individual utility based on the requirements of personal needs according to one's position in society. He defines necessary income (necessary to sustain the social status of an individual) and superfluous income (over and above the necessary). This classification suggests that utility is structured in a way that necessary income represents a safety-first element, similar to Lopez's aspiration theory (1987). Because of this classification of income, losses lead to a greater loss of overall well-being (caused by loss aversion), if the loss in the game exceeds the necessary income and if the gain increases only the superfluous income, since the necessary income must be valued higher than the superfluous income. Finally, he assumes bounded utility in the sense of Cramer, i. e., that beyond a certain threshold the superfluous income gives no extra utility.

Buffon dedicates a large part of the *Essays* to a presentation and discussion of the Petersburg gamble (articles XV-XXII). His discussion is so broad that it includes almost all currently-known 'solutions' to the Petersburg paradox[3]. In a footnote he quotes at length the letter he wrote to Gabriel Cramer in 1730. Thus he proves that some of his ideas preceded those of Daniel Bernoulli (1738). He concludes that the Petersburg paradox – that is, the discrepancy between the intuitive value and the expectation – arises from two causes; first, the low probabilities of the exorbitantly high payoffs are estimated as zero[4], and, second, because the decreasing marginal utility of money the exorbitantly high payoffs leads to very low increase in values[5]. Buffon next raises the solvency problem, according to which the payoff in the gamble can only be finite, so that the value of the gamble must be based on a finite, rather short gamble length. A remarkable contribution to the Petersburg paradox is his determination of the game value by repeated experiment[6]. A child played out 2^{11} Petersburg gambles to yield an average payoff of about 5 Ecu per game. Issuing from this outcome, Buffon motivated the geometric payoff distribution that results by application of the law of large numbers. Table 1 shows the outcomes of the repeated experiment and the statistically expected outcomes. By reflecting on the theoretical distribution, Buffon concludes that the larger the number of repetitions the larger is the expected outcome. He partly anticipates the mathematical solution to the repeated Petersburg gamble as proposed by Feller (1945). In contrast to Bernoulli's utility function, his arguments stand the test of the Super-Petersburg paradox (Menger 1934), again by cancelling low probabilities and adopting the zero marginal value of the superfluous income.

Table 1. Buffon

Number of tails 1(1) 10
Buffon's observations 1060; 494; 282; 137; 56; 29; 25; 8; 6, – – –
Payoff 1, 2, 4, 8, …
Geometric approximation $2^{10}, 2^{9}, 2^{8}, \ldots, 2^{0}$

The total payoff was 10,057 Ecu, an average of 4.91 Ecu.

The contribution of the last two articles in the *Essays* preceding the concluding one (article XXV) is more to mathematics than to moral behaviour. Buffon presents his games of the franc-carreau and Buffon's needle (article XXXIII); the latter is the reason why his name still remains in the mathematical sciences. The contribution is important since it introduces geometrical probability. According to available information on the internet, Buffon conducted experiments by throwing a stick over his shoulder into a tiled room. In similar random experiments on Buffon's needle, the number π was approximated in the 19th century.

Finally, Buffon discusses the meaning of infinity (article XXIV) by highlighting its value as a mathematical tool which allows the generalisation of results and by pointing out that it is not a 'real' number[7].

Summing up, the *Essays* present a collection of articles of Buffon dedicated to judgment and behaviour, including experimental methods and mathematics. Regardless of a number of issues that one can take with his presentation from a modern-day perspective (as it lacks a fully-fledged and fully-developed approach, and is not always convincing), it appears remarkable to see how much thought, discussed today in the human sciences literature, was already expressed during the 18th century. Buffon's *Essays* is outstanding from the standpoint of today when compared to many other works of that age because of its many and remarkable contributions: the introduction of the valuation of a game by experimental method, thus highlighting also the importance of empirical evidence; the discussion based on philosophical arguments of levels of significant and negligible probabilities; the valuation of losses and its distinction from the valuation of gains; the distinction of necessary income from superfluous income; and, finally, the introduction of geometric probability.

Before concluding, one remaining question must evidently be addressed. Since this paper has been prepared for publication in

honour of Reinhard Selten, the reader may ask what is the relationship between Buffon, his *Essays*, and Reinhard Selten and his work?
We reply that although we do not know of any direct link, we see relationships and analogy in (1) spirit, (2) presentation, (3) methodology, (4) thought, and (5) the grandeur of scientific contribution.
(1) Although Reinhard Selten is evidently a modern researcher, in our view he shares the mind-set of the great savants of the age of the enlightenment, that is, the dedication to research driven only by the desire to find and communicate the truth about the nature of things. Owing to the conservativeness of academic reviewers who were not always open to Selten's unorthodox theories, and to his reluctance to making any changes in his work that can bias his vision of the truth, he would publish his famous papers unchanged in unknown academic journals rather than having the changed version published in a prestigious journal.

Even though the prestige of publishing in such journals would have helped his career at that time, he accepted the facts in a humorous way. We remember him saying that by publishing in an unknown journal "you can make a journal famous." He surely made several journals famous since his publications in quite unknown journals were frequently quoted.

This achievement is evidently more exceptional than a publication in a famous journal, but it is a rocky road for making any impact in the human sciences, even for an exceptionally brilliant mind. He also pointed to the fact that in earlier, as in the Buffon's times, researchers had to start writing an essay on scientific questions by first apologizing for daring to raise these questions at all (see article I).
(2) Selten's archer (Selten and Buchta 1994) who represents direction learning theory is as original and as illustrative as the great Buffonian metaphor of the blind man who learns by experience about the return of the sunrise. Nevertheless both presentations similarly describe the learning of information about unknown things and the adjustment in hindsight. The difference, indeed, is in the adjustment itself; while the blind man adjusts his belief, the archer adjusts his strategy. The archer learns hitting a target with an arrow similarly to a blind man, only by repeated experience through feedback information. If he is informed that the arrow missed the target to the left, on the next trial he is going to adjust the direction of the arrow to the right rather than to the left, and vice versa. This direction learning theory has shown to explain the behaviour of most experimental subjects in economics laboratory experiments in very different scenarios (see Selten 2004 for a review).

Impulse balance theory is based on this direction learning story; it allows point-predictions of behaviour by balancing potential positive and negative impulses in games of strategy (for a simple application of this concept, see Selten and Neugebauer 2006).
(3) Selten has used mainly mathematical and experimental approaches in his research. Similarly to Buffon, he was fascinated by mathematics from his youth, received a degree in mathematics before he turned to applied mathematics and experimental research.

During his career, his main interests have involved the measurement and valuation of games, both theoretically and experimentally, and in the construction of solution concepts based on (levels of) rationality.

(4) Selten sees limits to the benefit of mathematics in the understanding of human behaviour. In the language of Buffon (article II), we can imagine although we did not witness, that Selten could have said:

Maths involves the truth of definitions **[1]**. *This truth is only helpful if one understands the problem well, that is, if the analysis of the problem is based on the right definitions.*

The Reinhard Selten School believes in the merit of experimental research to uncover the fundamental definitions of human behaviour, and to support the building of a bounded rational system in the form of a toolbox if a fully-fledged theory and fully-developed approach is not available. Buffon also seems to favour a bounded rational system to human beings over pure rationality, since he states that rationality is cold and does not make man really happy. In particular, he accepts that less rarely people gamble from time to time because hope makes them happy. In the language of Selten, many people tend to avoid risk taking as much as they can, but on some occasions they enjoy and permit themselves to take risks: *Today I take a chance*.

In relation to Buffon's discussion of the Petersburg gamble, Selten accepts, at least from the behavioural viewpoint, a treatment of very low probabilities as zero. Discussing the outcomes of valuations of the Petersburg gamble by experimental subjects, Selten was not at all surprised that according to the data people behave as if they neglect very low probabilities in the Petersburg gamble; "of course they do that." **[2].**

(5) Originality and the greatness of ideas which lead to path-breaking approaches must be used to describe the work of both Buffon and Selten[8]. While we do not know about Buffon, Selten had a reason for being so original. He argued that he was slower than other researchers and therefore had to take greater steps. Both had a great impact on the development of research in games and experiments; Buffon as the founding father of experimental statistics, Selten as the founding father of experimental economics in Europe. Finally, both have influenced thinking, being thought provoking and created great interest in their subject.

Selten is very active with his current research. Recently, at the celebration of the twenty-fifth anniversary of the Bonn Laboratory of Experimental Economics, he said that he was the "money most-intensive researcher" at the economics department at Bonn. So, we expect many new researches, advancements of the field of bounded rationality and enlightening ideas of and by Selten.

The main text

I. I do not attempt to present general essays on morality here; that would demand more enlightenment than I can presume, and more art than I recognize. The first and most sensible part of morality is rather an application of the maxims of our divine religion than of a human science; and I cannot even dare to try matters where the law of God is our principle, and Faith our rationale. The respectful gratitude or rather the adoration, man has to his Creator; brotherly charity, or rather the love he has to his fellowman, are natural feelings and virtues written on an ingenuous mind; all that stems from this pure source bears the character of the truth; the light is so bright that the existence of error cannot obscure it, the evidence so great that it admits no argument or discussion, or doubt, and no other measures than conviction.

The measurement of uncertain things is my object here. I will try to give some rules to estimate likelihood ratios, degrees of probability, weights of testimonies, influence of risks, inconvenience of perils; and judge at the same time the real value of our fears and hopes.

II. There are truths of different kinds, certainties of different orders, probabilities of different degrees. The purely intellectual truths like those of Geometry all reduce themselves to truths of definition; it is a matter of resolving the most difficult problem only by understanding it well, and there are no other difficulties in calculation and in the other purely theoretical sciences than to untangle what we put in, and to untie the knots that the human mind has created in the study of the implications of the definitions and the assumptions that are used as the foundation and framework of these sciences. Doubtlessly, all their propositions can always be proven, because you can always go back from each of these propositions to the preceding equivalent propositions, and from these back to the definitions. It is for this reason that evidence[9] itself belongs to the mathematical sciences and only to them. Indeed, we must distinguish the evidence of reasoning from the evidence that comes through the senses, that is, discern the intellectual evidence from the physical intuition. The latter is only a clear apprehension of objects or images; while reasoning is a comparison of similar or identical ideas, or rather the immediate perception of their identity.

III. In the physical sciences, evidence is replaced by certainty; it is not measureable, since it has only one absolute property, the clear negation or affirmation of the matter it shows. But certainty is never positive and absolute, it requires several relations that we must compare and by which we can estimate the measure. Physical certainty, that is, the most certain of all certainties, is nevertheless only the almost infinite probability **[3]** which an effect or event that never failed to happen, will happen again; for example, because the sun has always risen, it is thenceforth physically certain that it will rise tomorrow. The reason for being is to have been, but a reason for ceasing to be is to have come into being. Consequently, we cannot say that it is equally certain that the sun will always rise, or at least we must assume a preceding eternity, equal to the subsequent perpetuity, otherwise it will end as it has begun. For we must judge the future by the evidence from the past; whenever something has always been, or if it always behaved the same way, we must be assured that it will be or will always behave in the very same way. By always I mean a very long time, and not an absolute eternity, the always of the future never being equal to the always of the past. The absolute of any kind, whatsoever, is not the responsibility of Nature or of the human mind. Men have considered as ordinary or natural all the events that have this kind of physical certainty. An always occurring effect ceases surprising us. In contrast, a phenomenon that has never occurred, or which has always occurred in the same way but ceases to occur or occurs in a different way will reasonably surprise us as so extraordinary that we consider it supernatural.

IV. However, those natural effects that do not surprise us, have everything necessary to surprise us. Which circumstances of causes, what collection of principles was necessary to produce a single insect, a single plant! What a prodigious combination of elements, motions and springs exist in the animal machine! The smallest works of Nature are subjects of the greatest admiration. The reason why we are not at all surprised by all these wonders is that we were born into this world of wonders, that we have always seen them, that our understanding and our eyes are equally accustomed to them; finally, because all were before and will still be after us.

Had we been born in a different world of a different physical shape and different senses, we will have other relationships with exterior objects, will see other wonders and not be surprised. The wonders of both worlds are based on the ignorance of the causes, and on the impossibility of knowing the reality of the things. We are only permitted to see the relation those things have with ourselves.

There are therefore two ways of looking at the natural effects; the first is to see them as they present themselves to us without paying attention to their causes, or rather without looking for their causes; the second is to examine the effects with the view to relate them to the principles and causes. These points of view are very different and offer differing reasons for surprise. One causes the sensation of surprise, and the other creates the feeling of admiration.

V. We will speak only about this first way to look at the effects of Nature. However incomprehensible and complicated they present themselves to us, we judge them as the most evident and simplest, and judge them only by their results. For example, if we cannot conceive or even imagine why things attract each other, we will certainly be satisfied by this actual attraction, and believe that they always had, and always had attracted/will attract each other. It is the same with other phenomena of all kinds. As unbelievable as they may appear to us, we will believe them when being sure that they have occurred very often, but doubt them if they have failed to happen as often as they occurred, and finally, will deny them if we believe for sure that they never occurred. In a word, according to having seen and recognized them or not.

But if the experience is the basis of our physical and moral knowledge, then analogy is the first instrument: when we observe that a thing occurs constantly in a particular way, we are assured by our experience that it will occur again in the same way. And when someone reports that a thing occurred in this or that way, and if these facts are similar to the other facts that we know by ourselves, then we believe them; on the contrary, if the fact has no analogy with ordinary effects, that is, with the things that are known to us, we must be in doubt; and if it is directly opposed to what we know, we do not hesitate to deny it.

VI. Experience and analogy may offer us different certainties, sometimes almost equal and sometimes of the same kind; for example, I am almost as certain that the city of Constantinople exists, although I have never seen it, as I am of the existence of the Moon, which I have seen so often. The testimonies in large number, when they concern things that have a full analogy with those that we know, can produce a certainty almost equal to the physical certainty. Physical certainty must be measured by an immense number of probabilities because it is produced by a constant sequence of observations, by what we call the experience of all the times. Moral certainty **[4]** must be measured by a smaller number of probabilities, since it presupposes only a number of analogies with what is known to us.

Assuming a man that had never seen anything, heard anything, we investigate how the belief and the doubt generate itself in his mind. We assume him that he is struck by the first appearance of the sun. He sees it shine from the top of the skies, then decline and finally disappear, so what can he conclude? Nothing, except that he saw the sun, saw it follow a certain route, and that he no longer sees it. But this star reappears and disappears again on the next day; this second sight is a first experience, that must produce in him the hope to see the sun again, and he begins to believe that it can return. Nevertheless he is very much in doubt, but the sun reappears; this third sight is a second experience which diminishes the doubt as much as it increases the probability of a third return. A third experience increases it until he no longer doubts that the sun returns a fourth time, and finally when he sees this shining star appear and disappear regularly ten, twenty, a hundred times, he will be certain that it always appears, disappears and moves the same way. The more similar observations he has, the greater will be his certainty to see the sun rise the next day.

Each observation, that is, each day, produces a probability, and the sum of these probabilities [!], since it is very high, produces physical certainty. Therefore, we will always be able to express this certainty by numbers from the time of our first experience, and it is the same for all the others effects of Nature. For example, if we want to reduce the duration of the world and of our experience to six thousand years, the sun will rise for us only 2 million 190 thousand times, and as to date back to the second day that it rose, the probabilities to rise the next day increase, as the sequence 1, 2, 4, 8, … or as 2^{n-1}.

I say for *us* [better: in our climatic belt] because for the climate of the polar region it is not altogether so.

We[10] will have (for a sequence of natural numbers up to 2,190,000) $2^{2,189,999}$. This is already a prodigious number of which we cannot form an idea, and it is by this reason that we must look at the physical certainty as being composed by an immensity of probabilities. Indeed, when moving the day of creation back by only two thousand years, this immensity of probabilities becomes even 2^{2000} greater[11] **[5]**.

VII. But it is not so easy to estimate the value of analogy, or consequently, to measure moral certainty. In truth, it is the degree of probability that provides analogic reasoning its power; but in itself an analogy is only the aggregate of relations with known things. Nevertheless, according to that aggregate or those relations in general if more or less strong, the consequence of analogic reasoning will be more or less certain, without ever absolutely certain. For example, if a witness who, I suppose, possesses common sense, tells me that a child has just been born in this city, I will believe him without hesitating, since that fact is nothing than very ordinary, but on the contrary has an infinity of relations with the known things, that is, with the births of all other children. I will believe him, however, without being absolutely certain about it. If however the same man tells me that this child was born with two heads, I will believe it again, but more weakly, since a child with two heads has less relations with known things; if moreover he added that the newborn baby in addition has six arms and eight legs, I should have good reason to hardly believe it. But however weak my belief was, I could not refuse it entirety; this monster, although very special, nevertheless is composed only of parts that have some relations with known things. Only their assembly and very extraordinary number will provide me a strong inclination to distrust him. The power of analogic reasoning will therefore be always proportional to the analogy itself, that is, to the number of the relations with known things, and it is not a matter of ensuring good analogic reasoning, but of adjusting oneself well to all the circumstances, of comparing them with analogous circumstances, of aggregating the number of these, of taking a model of comparison to which we will relate this found value. Then we will exactly have the probability, that is, the degree of power of the analogic reasoning.

VIII. There is therefore a prodigious distance between the physical certainty and the certainty of the kind that we can deduce from most of the analogies. The former is an immense sum of probabilities that forces us to believe; the other is only a lower or higher probability, and often so low that it leaves us perplexed. The doubt is always inversely proportional to the probability; that is, it is always the greater the lower the probability. In the order of the certainties produced by the analogy, we must place the moral certainty, which seems even to take the centre between doubt and physical certainty; and this centre is not a point, but a very extensive line, whose boundaries it is quite difficult to determine. We can feel that it is a certain number of probabilities that equals the moral certainty, but what number is it? And can we hope to determine it as precisely as that by which we have just represented the physical certainty?

After having reflected on it, I have thought that of all the possible moral probabilities, the one that most affects man in general is the fear of death, and I felt from that time that any fear or any hope, whose probability is equal to that which produces the fear of death, can morally be taken as the unit for relating the measure of all other fears. I relate to it even the probability of hopes, since there is no difference between hope and fear, other than from positive to negative; and the probabilities of both must be measured in the same way. I seek therefore for what is actually the probability that a man who is doing well, and consequently has no fear of death, dies nevertheless in the twenty-four hours.

Consulting the mortality tables **[6]**, I see that there are only 10,089 to bet against one, that a 56 year old man will live more than a day. Now, since any man of that age, when reason is fully mature and the experience has all its force, nevertheless has no fear of death in the 24 hours, although there are only 10,089 against one that he will die in this short interval of time. I conclude that any equal or lower probability must be regarded as zero, since any fear or any hope below 1/10,000 must not affect or even occupy for a single moment our heart or mind[12].

IX. We can conclude that physical certainty relates to moral certainty as $2^{2,189,999}$**:**1000 and that whenever an event, whose causes we absolutely ignore, occurs in the same way thirteen or fourteen times in a row, we are morally certain that it will occur again even a fifteenth time, for $2^{13} = 8192$, and $2^{14} = 16,384$, and consequently when this event has occurred thirteen times, there is 8,192 to bet against 1 that it will occur a fourteenth time; and when it has occurred fourteen times, there is 16,384 to bet against 1 that it will occur even a fifteenth time, which is a higher probability than the one of 10,000 against 1, which is higher than the probability that makes moral certainty.

Someone will be perhaps able to tell me, that although we do not have the fear or the worry of sudden death, the probability of sudden death must be zero, and that its influence on our conduct is morally zero. A man whose mind is beautiful, when he loves someone, will he not reproach himself to delay a day the measures that must assure the happiness of the loved person? If a friend entrusts us with a considerable deposit, will we not put the same day a written comment on this deposit? Therefore, in these cases we act as if the probability of sudden death is something, and we have reason to act thus. Therefore we must not always regard the probability of sudden death as zero.

This kind of objection vanishes after considering that we often do more for others than for ourselves! When we comment in writing at the very moment when receive a deposit, we are only honest to the owner of the deposit for his tranquillity, and not at all because of the fear of our death in the next 24 hours. It is the same attentiveness that makes happy someone or us, but not the feeling of the fear of an approaching death that guides us, it is our own satisfaction that drives us, we seek to enjoy all possible at the earliest.

A reasoning that might appear more justified is that all men are inclined to flatter; that hope seems to arise from a lower degree of probability than fear; and that consequently we are not entitled to substitute the measure of the one with the measure of the other: fear and hope are feelings but not concrete things. It is possible, and even more likely, that these feelings do not measure the precise degree of probability, so must we assign them an equal measure or no measure at all?

I reply that we measure not feelings, but rather the reasons that must originate them, so that any wise man must estimate the value of these feelings of fear or hope only by degrees of probability. Indeed, even when Nature, for the happiness of man, had given him more inclination towards hope than fear, it is not less true that probability is the true measure of both [inclinations]. It is not even by applying this measure that we can figure out our false hopes, or to reassure on the unfoundedness of fears.

Before finishing this article, I must observe that we must beware of mistaking what I called effects of unknown cause. Since I understand only the effects of which the causes, although unknown, must be supposed constant, such as the natural effects, any new discovery in physics, noted by thirteen or fourteen confirmed experiences, already has a degree of certainty equal to that of moral certainty, and this degree doubles with each new experience. By multiplying, we ever more approach physical certainty. But we must not conclude that the effects of chance follow the same law. It is true that in a sense these effects are among those for which we ignore the immediate causes, but we know that in general these causes are quite far from the supposed constancy. On the contrary, they are necessarily variable and volatile as much as possible. Thus even by the notion of chance itself, it is evident that there is no connection, no dependence, between its effects. Consequently the past cannot have any influence on the future, and we will be very much and even completely mistaken by wanting to infer from previous events any reason for or against future events. If a card, for example, has won three times in a row, it is not less probable that it will win a fourth time. As long as the laws of the game ensure the equality of chances, we can bet on equal terms on winning or losing, no matter how many times we won or lost previously. Presuming or believing the opposite, as some players do, is to go against the principle of chance itself, or forget that by the conventions of the game it is always played the same way.

X. In the effects for which we see the causes, a single test is sufficient to cause physical certainty. For example, I see that in a clock the weight makes the wheels turn, and the wheels make the pendulum go; I am certain from that time without need of reiterated experiences, that the pendulum will always perform the same way as long as the weight makes the wheels turn. This is a necessary consequence of an arrangement that we made ourselves while constructing the machine, but when we see a new phenomenon, an effect in the still unknown Nature, and since we are ignorant about its causes, which can be constant or variable, permanent or sporadic, natural or accidental, we do not have other means to obtain certainty than by repeating the experience as often as necessary. Nothing here depends on us and we only know as much as we experiment; we are only assured by the effect itself and by its repetition. As soon as it has occurred thirteen or fourteen times in the same way, we already have a degree of probability equal to moral certainty that it will occur even a fifteenth time, and from this point we can soon cover an immense interval, and conclude by analogy that this effect depends on the general laws of Nature; that it is consequently as old as all the other effects and that there is physical certainty that it always will occur as always.

In the risks that we arranged, balanced and calculated ourselves, we must not say that we are ignorant of the causes of the effects: we are ignorant of the true immediate cause of each effect in particular; but we clearly see the first and general cause of all the effects. I do not know, for example, and I even cannot imagine in any way, what is the difference of the movements of the hand, to pass or not to pass ten with three dice, which nevertheless is the immediate cause of the event, but I certainly see by the number and the make of the dice, which are here the main and general causes, that the chances are absolutely equal, so you are indifferent to bet on passing or not passing ten. I see moreover that these same events, if they happen, have no connection, since to every throw of the dice the risk always is the same, and nevertheless always new; that the past throw cannot have any influence on the throw to come; that we can always equally bet for or against. Finally that the longer we play, the greater the number of effects for and against, the nearer they will approach equality. After making sure that the outcome of every experience exactly opposes the experience of natural effects, I wish to say, the certainty exists in the inconstancy instead of the constancy of the causes. And each test leads to doubling the probability of the replication of the effect, that is, of the certainty that the cause is constant. In the effects of risk, on the contrary, each test increases the certainty of the inconstancy of the cause by showing us ever more that it is absolutely volatile and totally indifferent to produce the one or the other of these effects.

When a gamble is in its nature perfectly fair, the player has no reason to choose this or that side; since finally it necessarily follows from the supposed fairness of this game that there are no good reasons to prefer the one or the other side. And one of them can only be determined by the wrong reason, the logic of the gamblers seems completely wrong. Even the good minds that allow themselves to play, fall into absurdities after which they soon blush as reasonable men.

XI. Besides, all this supposes that the risks are balanced and rendered equal, as in the game passing ten with three dice. These dice are as perfect as possible, that is, they are exactly cubic, the material is homogenous, the numbers are painted and not marked hollow, so that the weight on one face is not more than on any other. But it is impossible to make anything perfect, and no dice are made so strictly precise, and we are often able to recognize by observation on which side the imperfection of the devices of chance tips the risk. Thus, it is only necessary to observe attentively for a long time the sequence of events, to count them exactly, to compare their relative numbers. If one of them essentially exceeds the other, we will be able to conclude, with great reason, that the imperfection destroys the perfect equality of the risk, and provides a stronger inclination to one side than to the other. For example, I suppose that before playing passing ten, one of the gamblers was subtle enough, or rather, rogue enough to throw a thousand times the three dice in advance, and to recognize that in these trials there were six hundred that passed ten. He will then have a great advantage over his opponent by betting on passing.

This difference that stems from the imperfection of the devices can therefore be recognized by observation, and it is for this reason that the gamblers often change dice and cards when their luck is against them. Thus, however obscure the destinies may be, however opaque the future may appear to us, by reiterated experiences we may nevertheless become, in some cases, enlightened about future events, as if we were beings or rather superior natures who immediately deduce the effects from their causes. And among the very things that seem to be pure risk, as games and lotteries, we can again recognize the inclination of the risk. For example, in a lottery drawn every fortnight, by studying the published winning numbers, and noting those that most often won during a year, two, three consecutive years, we may deduct, with reason, that the same numbers will win again more often than the others. Indeed, however vary the motion and the position of the device of chance, it is impossible to render it perfectly enough to maintain absolute equality of chances. There exists a certain routine to proceed, to place, to mix the tickets, which even in the midst of confusion produces a certain order, so that certain tickets must come out more often than the others. It is the same with the arrangement of cards to play; they have a kind of sequence so that we can grasp some terms by force of observations; because while assembling them with the hand one follows a certain routine, and the gambler himself when shuffling them follows his routine. Everything is done in a certain way more often than in another, and an attentive observer of results collected in large number will always bet with great advantage, that a certain card, for example, will follow such other card **[7]**. I say that this observer will be greatly advantageous, because the prior risks must be absolutely equal. A least inequality, a least degree of higher probability greatly influences the game, which is in itself only a multiplied and always repeated bet. If such a difference only amounts to one hundredth, it is evident that in a hundred throws the observer will gain his stake, that is, the sum that he risks every time; so that in a long run a gambler equipped with these dishonest observations, cannot fail to ruin all his opponents. But we offer a powerful antidote against the epidemic evil of the passion of play, and at the same time some preventives against the illusion of this dangerous art.

XII. It is generally known that the game is a greedy passion, where the habit is ruinous, but this truth has perhaps never been shown unless by sad experience about which gamblers did not reflect enough to correct it consciously. A gambler, whose wealth is exposed every day to the vagaries of chance, exhausts himself little by little and finally finds himself necessarily ruined. He attributes his losses only to the same risk that he blames for unfairness, he equally regrets what he lost and what he did not win; greed and false hope gave him claims on the goods of others. So humbled to find himself in necessity and no longer having the means to satisfy his greed, he is despaired and blames his ill-fated star. He does not even imagine that this blind power, the fortune of the game, actually marches by an indifferent and uncertain pace, but that at each step **[8]** it nevertheless tends to a goal, and pulls to a certain end. And that is the ruin of those who try it. He does not see that with time the apparent indifference it has for good or evil necessarily produces evil, that a long random sequence is a fatal chain whose extension causes misfortune. He does not feel that regardless of the hard tax of the cards and of the even harder tribute paid to the roguery of some opponents, he has passed his life making ruinous conventions; that finally the game by its very nature is a vicious contract in its principle, a harmful contract to each contractor in particular, acting contrary to the good of any society.

This is not at all a speech of vague morals, the above describes precise truths of metaphysics which I subject to calculation or rather to the strength of reason. Truths that I pretend to show mathematically to all those who have their minds clear enough and the imagination strong enough to combine without geometry and to calculate without algebra.

I will not speak about the games invented by artifice and worked out by avarice, where the chance loses a part of its rights, where the fortune can never balance, because it is invincibly entailed and always obligated to lean to one side. I want to say that all those games where the chances are unequally divided offer at once an assured and dishonest gain, and leave the other only with a sure and shameful loss. In Pharaoh, the banker is only a roguish solicitor and the punter a fool who is not mocked by [tacit] agreement.

It is in games in general, in the most equal game, and consequently the most honest that I find a vicious essence. Even in the word *game* I include all the conventions, all the bets where one puts at risk a part of his goods to obtain a similar part of the goods of others; and I say that in general the game is a misunderstood pact, disadvantageous to both parties. Its effect is to make the loss always greater than the gain; and to capture the good and transfer it to the evil. The demonstration of that is as easy as evident.

XIII. Take two men of equal fortune, each having 100,000 pounds of goods, and suppose that they stake 50,000 pounds, that is, half of their goods in one or more throws of the dice. It is certain that whoever wins increases his goods only by a third, and that whoever loses diminishes his by half. The loss is a sixth part larger than the gain since there is this difference between the half and the third **[9]**. The agreement to play is detrimental to both, and consequently essentially vicious.

This reasoning is not false, it is true and exact. Although one gambler only lost precisely what the other won; this numerical equality of the sum does not prevent the true inequality of the loss and gain. The equality is only apparent, but the inequality very real. The agreement both men make when betting half of their goods is tantamount to the effect of another agreement which no one had decided to make. That would be the agreement for each to throw a twelfth part of his goods into the sea. Before they risk half of their goods, it can be shown that the loss is necessarily a sixth greater than the gain. This sixth must be considered as a real loss, that can indifferently befall one or the other, and consequently it has to be divided equally.

And what will happen if the two men decided to gamble all their goods? One will only double his fortune, and the other, annihilate his. What proportion is here between the loss and the gain? The same as between all and nothing; the gain of the one is only equal to a rather modest sum, and the loss of the other is numerically infinite, and morally so great that the work of a lifetime will perhaps not suffice to regain his goods.

When one gambles all his goods the loss is therefore infinitely greater than any gain; it is greater by a sixth part when gambling half of goods, and by a twentieth part when gambling a quarter of goods [25/125 = 1/5, 25/100 = 1/4, 1/4 – 1/5 = 1/20]. In a word, however small the portion of fortune that one risks in the game, there is always more to lose than to gain; the agreement to gamble is thus a vicious contract that tends to ruin both parties. This is a new but very useful truth, that I wish to be known to all those who, for greed or laziness, spend their life gambling.

It has been often asked why people are more sensitive to loss than to gain, but a fully satisfactory answer is impossible unless not doubting the truth just presented. Now however the response is easy: people are more sensitive to loss than to gain, since actually whereas numerical equality is supposed, the loss is nevertheless always and necessarily greater than the gain. The feeling is generally an implicit reasoning only less clear, but often brighter, and always surer than the direct product of reasoning. One feels that the gain does not give us as much pleasure as the loss causes us pain. This feeling is only the implicit result of the reasoning just presented.

XIV. Money must not be estimated by its numerical quantity: if the metal, that is merely the sign of wealth, was wealth itself, if the happiness or the benefits which result from wealth were proportional to the quantity of money, men will have reason to estimate it numerically, by its quantity, but it is barely necessary for the benefits derived from money to be in just proportion with its quantity. A rich man with an income of 100,000 Ecus is not ten times happier than the man who has only 10,000. More: as soon as a certain boundary is passed, money has almost no real value, and cannot increase the wellbeing of its possessor. A man who discovered a mountain of gold will not be richer than that who found only one cubic fathom **[10]**[13].

Money has two values both arbitrary and conventional. One measures the benefits to the individual, the other determines the well-being of the society. The first has only been estimated very vaguely; the second is suitable for justly estimating by comparing the quantity of money with the proceeds of the land and the labour of men.

To offer successfully some precise rules about the value of money, I examine special cases in which the mind easily grasps the necessary combinations. By induction, they lead us to estimating the value of money in the general case, for the poor, the rich, even for the more or less wise.

For the man who in his budget, whatever it is, has only the necessary, money has an infinite value; for those who abounds in superfluous, money has almost no value anymore. But what is necessary, and what is superfluous? By *necessary*, I understand the income which a person ought to spend to live as always. He can be comfortable and even pleasurable, but the habit soon creates needs. And so, in the *superfluous* I do not include any of the usual pleasures.

Superfluous is the income that can bring us new pleasures. The loss of necessary is felt infinitely, so when you risk a considerable part of this necessary, the risk cannot be offset by any however great hope. On the contrary the loss of the superfluous has limited effects; and if even superfluous income is still more sensitive to loss than to gain, it is because in fact the loss is generally always greater than the gain. This happens because ordinary feelings are based on common concepts or on simple induction, but the delicate feelings depend on exquisite and elevating ideas, and are in fact only the results of several combinations often too subtle to be clearly noticed, and almost always too complicated to be proved.

XV. Those mathematicians who have evaluated games of chance, and whose research in this field deserves praise, had considered money only as a quantity susceptible to growth and diminution, without other value than number. They had estimated the relations between gain and loss by the numerical quantity of money. They had calculated the risk and hope relative to this very numerical quantity.

We consider the value of money from a different point of view, and our principles solve some embarrassing cases by ordinary calculation, solve, for example, the game of heads and tails, where two men (Peter and Paul) play against each other under these conditions **[11]**: Peter throws a coin until tails shows up. If that occurs at the first throw, Paul gives him one Ecu; if only on the second throw, Paul gives him two Ecus; then four, eight, …, Ecus, always doubling that number. Peter can only win, and his gain will at least be an Ecu, perhaps two, four, eight, sixteen etc. Ecus, and finally an infinity of Ecus.

So how much Peter must give Paul to compensate him, or, what amounts to the same, what sum is equivalent to the hope of Peter.

This problem was proposed to me for the first time by the blessed Mr. Cramer, the famous professor of mathematics at Geneva, during my trip to this city in 1730. He told me that it was previously proposed by Nicolas Bernoulli to de Montmort (1708/1713, pp. 402 – 407).

I thought about this problem for some time without finding the knot, I could not see that it was possible to agree mathematical calculations with common sense without introducing some moral considerations. I expressed my ideas to Cramer, and he told me that I was right, and that he had also solved this question by a similar approach. He showed me then his solution almost identical to the one printed later in 1738 in the *Mémoires* of the Academy of Petersburg **[12]** by Daniel Bernoulli on the measure of chances. I saw that most of the ideas of Dan. Bernoulli agreed with mine, which gave me great pleasure since I always have, in addition to his great talents in geometry, considered and acknowledged Dan. Bernoulli as one of the best minds of this century. I found also that the idea of Cramer is indeed justified, and worthy of a man who has given us proofs of his skill in all mathematical sciences, and to whose memory I do justice [he died in 1752] with so much more pleasure than there was to the company and friendship of this scholar whom I owe part of the first knowledge that I acquired in this field.

Montmort solves this problem by ordinary rules, and he says, that the sum equivalent to the hope of that person who can only win, is equal to the sum of the sequence 1/2, 1/2, 1/2, …, etc. Ecu continued to infinity, and that consequently this equivalent sum is an infinite sum of money. The reason on which this calculation is based is that there is one half of probability that Peter will have one Ecu; one quarter of probability that he will have two; one eighth of probability that four; etc. to infinity; consequently it is necessary that Peter gives Paul as an equivalent half of an infinity of Ecus.

This is mathematically true and cannot be disputed. Montmort and the other geometricians thought that this problem was well resolved, but this solution is so far from the true: instead of giving an infinite sum, or even a very large sum, which already is quite different, no man of common sense will give twenty or even ten Ecu to replace that person who can only win.

XVI. The reason for this extraordinary contradiction between common sense and calculation has two causes. First, probability must be considered as zero as soon as it is very low, that is, below 1/10,000; second, the relation between the quantity of money and its resulting benefits should be accounted for. A mathematician estimates money by its quantity, but the moral man must estimate it otherwise. For example, if we propose to a man with a mediocre fortune to put 100,000 pounds in a lottery, because there is only 100,000 to bet against one that he will win 100,000 times 100,000 pounds; it is certain that the probability to obtain the promised, it is certain, I say, mathematically speaking, that his hope will be worth his stake; but this man will make a very big mistake to risk this sum, and even a larger since the probability of winning is very low. Although the money to win increases in proportion, 100,000 times wining 100,000 pounds will not double the benefits that he will have with 50,000 times those same 100,000 pounds, or ten times as much benefit as with 10,000 times 100,000 pounds. For the moral man, the value of money is not proportional to its quantity, but rather to the benefits that money can buy. It is obvious that this man must risk only in proportion to the hope of these benefits, which he must not calculate by the numerical quantity of those sums. The quantity of money, beyond certain limits, cannot further the increase in his happiness. He will not be happier etc.

XVII. To understand the stated connection and the truth of all that I have advanced, we examine the proposed question more closely than the geometricians did. Indeed, ordinary calculation cannot resolve it because of the morale which is causing difficulties with mathematics. Let us see if other rules enable us to reach a solution which does not violate common sense, and at the same time is in accordance with experience. This research will not be useless, and we furnish the means to estimate exactly the price of money and the value of hope in all cases. The first thing I remark is that mathematical calculation gives as equivalent to the hope of Peter an infinite sum of money but that sum cannot morally have more than 30 terms, since the sum will already amount to 520,870,912 Ecus, that is, as much money as exists perhaps in the whole kingdom of France.

An infinite sum of money does not exist, and all hopes based on the next terms to infinity do not exist either. The moral impossibility destroys mathematical possibility[14].

XVIII. But how to estimate it, how to find this value for different quantities? Can we give precise and general rules for this estimation? It seems that everyone must evaluate his state, and estimate the quantity of money proportional to this state and to the usage he can make from it. However, this approach is still vague and too special to serve as a principle. I believe that more general and safer methods can be found and the first method that presents itself is to compare mathematical calculation with experience. Indeed, in many cases, repeated experience can explain the effect of chance as surely as if deduced immediately from the causes.

I have therefore made 2048 experiences, I played 2048 times this game by letting a child throw the coin in the air. Those trials produced 10,057 Ecus in all. The equivalent to the hope of the person who can only win is almost five Ecus for every trial **[18]**. In this experiment, there were 1060 trials that produced only one Ecu, 494 trials that produced two Ecus, 232 trials, four Ecus, 137 trials, eight Ecus, 56, sixteen Ecus, 29, 32 Ecus, 25, 64 Ecus, eight trials, some (?) 128, and finally six trials, 256. I regard this result generally as good, because it is based on a large number of experiences, and agrees with another mathematical and indisputable reasoning by which one finds almost the same equivalent of five Ecus.

Here is this reasoning. If we perform 2048 trials, there must naturally be 1024 trials that produce only one Ecu each, 5012 trials, two, 256, four, 128, 64, sixteen, 32, 32, sixteen, 64, sixty-four, eight, 128, four, 256, two, 512, and one, 1024, and finally one, which we cannot estimate. We can neglect it without appreciable error, because it can be supposed, without violating more than slightly the equality of chance, that there were 1025 instead of 1024 trials that produced only one Ecu. Besides, the equivalent of this trial cannot be more than fifteen Ecus, since all the terms beyond the 30th give such great sums that they do not exist, and that consequently the greatest supposed equivalent is fifteen Ecus. Adding together all these Ecus which I naturally must expect by the indifference of risk, I have 11,265 times five Ecus for 2048 trials, very roughly five Ecus and one half as the equivalent which agrees with the experience to within 1/11 [10,057/2048 = 4.91; 5.50 – 4.91 = 0.59; 0.59/5.50 = 1/10.7.]

I feel, although, that criticism is possible, that this type of calculation that gives five Ecus and one half for an equivalent for 2048 trials, will give a greater equivalent, if a much larger number of trials is added; because, for example, only 1024 trials produce very roughly an equivalent of five Ecus; only 512 trials, no more than very roughly four Ecus and one half; only 256, no more than four Ecus, and thus always diminishes. But the reason for concern is in the outcome which we cannot estimate but which constitutes a considerable part of the totality. It is even much more considerable if playing still less, and consequently a large number of trials is necessary, like 1024 or 2048 which are so large that this outcome can be supposed negligible, or even a zero. Following the same argument, we find that playing

1,048,576 [= 1024^2] trials, the equivalent is almost ten Ecus, but we must consider everything in the morale: it is impossible to perform so many trials.

And between playing only one and the largest morally possible number of trials, this reasoning yields an average equivalent of five Ecus. Thus I still say that the equivalent to the hope of the person who can only win is five Ecus instead of an infinite sum.

XIX. Let us see if according to this estimation it would not be possible to derive the money value that corresponds to the benefits resulting from it. The progression of probabilities is

$$1/2,\ 1/4,\ 1/8,\ \ldots,\ 1/2^{\infty}. \qquad (1)$$

The progression of the sums of money

$$1,\ 2,\ 4,\ \ldots,\ 2^{\infty-1}. \qquad (2)$$

The sum of all these probabilities, multiplied by all the money at stake is $\infty/2$ which is the equivalent of the hope of the man who can only win. But we saw that in reality this sum is only five Ecus. It is therefore necessary to look for such a sequence, whose terms multiplied by the appropriate probabilities provide five Ecus. This sequence is a geometric progression, namely

$$1,\ 9/5,\ 81/25,\ \ldots$$

It represents the geometric quantity of money provided by experience, and consequently its moral and real value. Here is therefore a general estimation sufficiently close to the value of money for all possible cases, and independent of any assumption. For example, by comparing the two sequences, it is seen that two thousand pounds do not double the benefit of one thousand pounds, but increases it less by 1/5, i. e., by 1800 pounds. [Similar examples follow.]

A miser is like a mathematician, both esteem money by its numerical quantity, but a sensible man considers neither the mass nor the number, he sees only the possible ensuing benefits. He reasons better than the miser, and discerns better than the mathematician. For the miser and mathematician an Ecu that the poor has set aside to pay a necessary tax and an Ecu that completes the bags of the financier have just the same value. The mathematician will consider them equal, the other will grab both Ecus with equal pleasure, but the sensible man will count the Ecu of the poor for a Louis, and the Ecu of the financier, for a Liard.

XX. Another consideration which supports this estimation of the moral value of money is that a probability must be regarded as zero as soon as it is only 1/10,000, as low as the unfelt fear of death in 24 hours. We may even say that, since the intensity of this fear is considerably stronger than the intensity of all the other feelings of fear or hope, we must consider a fear or a hope, that has only a probability of 1/10,000. as almost zero.

The weakest man can draw the lots without any emotion, if the ticket of death were mixed with 10,000 tickets of life; and the strong man must draw without fear, if this ticket is mixed with a thousand. Thus in any case in which the probability is under a thousandth, we must look at it as almost zero.

Now, in our study, already beginning with the tenth term of sequence (1), the probability is 1/1024. Therefore, morally thinking, we must neglect all the following terms, and limit all our hopes at this tenth term; so that five Ecu are left as the desired equivalent. This confirms the accuracy of our estimation. By rearranging and cancelling all the calculations where the probability becomes lower than 1/1000, there will no longer remain any contradiction between mathematical calculation and common sense. All difficulties of that kind disappear. The man pervaded by this truth will not anymore abandon himself to vain hopes or false fears; he will not gladly give his Ecu to obtain a thousand, unless he does clearly see that the probability of success is higher than one thousandth. Finally, he will leave his frivolous hope of making a great fortune with small means.

XXI. So far I have only reasoned and calculated for the truly wise man who is determined only by the weight of reason; but must we not give some attention to the great number of men whom illusion or passion deceive, and who are often quite comfortable with being deceived? Is there yet nothing to lose if things are always presented just as they are? Is hope, however low the probability, not a good for all men, and the only good for those unfortunate? After having calculated for the wise man, we therefore also calculate for the much less rare man who often enjoys his mistakes more than his reason. Calculate regardless of cases in which, in spite of lacking means, a glimmer of hope is a supreme good; regardless of these circumstances under which a restless heart can only rest on the objects of its illusion, and only enjoys its desires. Are there not thousands and thousands of occasions in which even wisdom must throw forward a volume of hope in the absence of real evidence? For example, the desire to do good, recognized in those who hold the reins of government, albeit without exercise, spreads on all the people a sum of happiness one cannot estimate. Hope, albeit vain, is therefore a real good, its enjoyment is appreciated by anticipation of all other goods. I am compelled to admit that full wisdom does not imply full happiness of man, that unfortunately the mere reason had at any time only a small number of cold listeners, and never created enthusiasts. A man stuffed with goods will not yet be happy without hope for more. With time, superfluous income becomes really necessary and the mere difference between the wise and the non-wise is that the latter, at the very moment he reaches an overabundance of goods, converts this lovely superfluous income into sad necessary income, and raises his state in accordance with his new fortune. But the wise man will use this overabundance only to spread the benefits and to obtain some new pleasures. He spares consumption of the superfluous income and multiplies its enjoyment.

XXII. The display of hope is the business of all money swindlers. The smart art of preparing a lottery consists in presenting large sums with very low probabilities, which is soon swollen by the spring of greed. These swindlers still enlarge this ideal product, dividing it and offering it for very little money, which everyone can loose. A hope that, though much weaker, seems to be a part of the grandeur of the total sum. People do not know that, when the probability is below one thousandth, hope becomes zero however large is the promised sum.

Anything, however great it may be, is necessarily reduced to nothing as soon as it is multiplied by nothing, as it is here: the large sum of money, just as any number, multiplied by a zero probability, is always zero. People are also unaware that independently from that reduction of the probabilities to nothing, hope suffers a successive and proportional decline of the moral value of money, which is always less than its numerical value. When hope seems to double, it only increases by 9/5, when seems to quadruple, increases only by $(9/5)^2$ etc.

We see how much moral hope differs in each case from the numerical hope and the wise man must therefore reject as false all the propositions, though proven by calculation, where a very large quantity of money seems to compensate a very low probability. If he wants to risk with less disadvantage, he must never allocate his funds to a large venture, it is necessary to divide them. To risk 100.000 francs on a single vessel or 25,000 on each of four vessels, is not the same thing, because you will have 100,000 for the moral hope in the latter case, but only 81,000 [= $25{,}000(9/5)^2$] in the former. It is by this same reason that the most surely profitable businesses are those where the mass of the debt is divided between many creditors. The owner of the mass will suffer only light setbacks but will not be ruined.

In the moral sense, playing for high stakes means playing a bad game. A punter in the game of pharaoh who takes into his head to push all his cards until fifteen will lose about a quarter on the moral product of his hope. Indeed, his numerical hope is to pull 16, but his moral hope is only 13104/125 [104/125 = $(9/5)^3$].

It is the same for countless other examples that one could give; and everywhere it shows that the wise man must put at risk the least possible, and that the prudent man who, through his position or his business, is forced to risk large funds, must divide them, and subtract from his speculations the hopes for which the probability is very low, albeit the sum to obtain is proportionally also large.

XXIII. In the science of probabilities analysis is the only instrument that has been used until now to determine and fix the ratios of chance; Geometry appeared hardly appropriate for such a delicate matter. Nevertheless, looking closely, you will easily recognize that this advantage of Analysis over Geometry is quite accidental, and that chance according to whether it is modified and conditioned is in the domain of geometry as well as in that of analysis. To be assured of this, it is enough to see that games and problems of conjecture ordinarily revolve only around the ratios of discrete quantities. The human mind, rather familiar with numbers then measurements of size, has always preferred them. To put therefore Geometry in possession of its rights in the science of chance is only a matter of inventing some games that revolve on size and on its ratios or to analyse the small number of those of this nature that already exist. The free tile [franc-carreau] game can serve as an example, and here are its very simple terms.

In a room parquetted or paved with equal tiles of an unspecified shape **[13]** one throws an Ecu in the air. A gambler bets that after its fall this Ecu will be located on a single tile; the second bets that on two tiles, that is, covers one of the joints which separates them; a third gambler bets, that on two joints, and a fourth, that it will be located on three, four or six joints. Required are the chances of each of these gamblers.

To begin with, I seek the chances of the first and second gamblers. I inscribe in one of the tiles a similar figure, distant from the tile borders by the length of half the diameter of the Ecu. The ratio of their chances is as that of the areas of the circumscribed ring and the inscribed cell. That can be easily shown, because as long as the centre of the Ecu is in the inscribed cell, it can be located only on a single tile: by construction, this inner cell is everywhere distant from the contour of the outer cell by the distance equal to the radius of the Ecu. In contrast, as soon as the centre of the Ecu falls outside the inscribed cell, the Ecu will necessarily be located on two or several cells, since then its radius is greater than the distance between the contours of both cells. And yet, all points where this centre of the Ecu may fall are represented in the first case by the area of the ring which is the remainder of the cell. Therefore the ratio of the chances of the gamblers is the ratio of those areas. Thus to equalise their chances it is necessary that the areas of the inscribed cell and the ring be equal to half of the total surface of the cell.

I enjoyed calculation, and found that to play a fair game on two square cells, ratio of the side of the outer cell to the diameter of the Ecu must equal $1:[1-(1/\sqrt{2})]$, almost three and half times greater than the diameter of the coin.

To play on triangular equilateral tiles, […], On lozenge tiles, […]. Finally, on hexagon tiles, […]

I have not studied other figures, because the above are the only ones by which a space can be filled without leaving some intervals for other figures **[14]**; and I do not think it is necessary to tell that if the joints of the tiles have some width, they give advantage to the gambler who bets on the joint, and that to render the game even more equal by giving to the square tile a little more than three and half times, to the triangular six times, to the lozenges four times, and to the hexagons two times the diameter of the coin.

Now I seek the chances of the third gambler who bets that the Ecu will be located on two joints; and I inscribe in one of the cells a similar figure as I have already done. Next I extend the inscribed sides of this figure until they meet those of the cell. The ratio of the chances of this third gambler to those of his opponents is as the sum of the spaces enclosed between the extension of these lines and the borders of the tile to the remainder of the surface of the tile. To show this properly suffices it to be well understood.

I have also calculated this case, and found that to play a fair game on square tiles …

I omit the solution of several other cases, for example the cases in which one of the gamblers bets that the Ecu will fall only on one joint or on two, on three, etc. They are not more difficult than the preceding; and besides, they are rarely met. But suppose that instead of throwing a round piece, as an Ecu, we throw a piece of another shape, as a squared Spanish pistole, or a needle, a stick, etc. This problem demands a little more geometry, although in general it is always possible to solve it by comparing areas.

I suppose that the floor is simply divided by parallel joints. We throw a stick in the air, and one gambler bets that the stick will not cross any of the parallels on the floor, and the other, that the stick will cross some of these parallels. What are their chances? This game can be played on a checkerboard with a sewing needle or a headless pin.

These examples suffice to give an idea of the games that can be imagined on the relations of extents. Several other problems of this type are also interesting and even useful. For example, how risky it is to cross a river on a more or less narrow plank; what must be the fear of a lightning bolt or a shell burst **[15]**.

XXIV. From the very beginning, we find Infinity in Geometry, and since the earliest times Geometricians caught sight of it; the squaring of the circle and the treatise on *Numero Arenae* [Sand Reckoner] by Archimedes prove that this great man gave thought to infinity, and we must even share some of his thoughts; we have extended his ideas, though handled in different ways, and finally we have found the art of applying calculus to his ideas. But the basis for the metaphysics of the infinite had not changed at all, and only recently some Geometricians gave us views on infinity that are different from those of the Ancients, but so far from the nature of the things and truth, that they were neglected in the works of those great mathematicians. Hence arose all the opposition, all the contradictions that we suffer in calculus; hence arose the disputes between Geometricians on how to calculate, and on the principles from which it derives. We were surprised by the of miracles which these calculations produced, and confusion followed. It was believed that infinity produced all these wonders; it was imagined that the knowledge of infinity had been refused in all the centuries and reserved for ours. Finally, infinity was built on systems that only served to obscure thought.

Let us say therefore a few words on the nature of infinity, which, while enlightening seems to have blinded men. We have clear ideas about magnitude, we see that things in general can be augmented or diminished, and the idea that a thing becomes larger or smaller is as familiar to us as about the thing itself. It is possible to augment or diminish whatever is thus presented to us or only imagined, nothing stops, nothing destroys this possibility. We can always conceive a half of the smallest thing and a double of the largest thing; we can even imagine that it can become a hundred a thousand, a hundred thousand times smaller or larger. It is in this possibility of growth without boundaries that consists the true concept we must have on infinity.

We derived this concept by issuing from the concept of the finite. A finite thing has ends, boundaries; an infinite thing is the same finite thing from which we remove these ends and boundaries. The idea of infinity is thus only a concept of deprivation, and has nothing of a real object. Here is not the place to show that space, time, duration, are no infinite realities (?); it will suffice to prove that there exists no number infinite or infinitely small, or larger or smaller than an infinite number, etc.

[Natural] numbers are only an assembly of units of the same kind; the unit is not at all a number, it designates a single thing in general. But the first number 2 denotes not only two things, but two similar things, two things of the same kind. It is the same for all the other numbers: yet these numbers are only representations, and never exist independently of the things that they represent; the characters that they designate do not give them any reality at all. They require a subject or rather an assembly of subjects to represent, to make their existence possible. I understand their intelligible existence since they can only have real values. An assembly of units or of subjects can never be other than finite, we will always be able to assign the parts of which it is composed. Consequently numbers cannot be infinite whatever the growth one gives them.

But, we may ask, is not the last term of the natural sequence 1, 2, 3, etc. infinite? Are there no last terms of other even more infinite sequences than the last term of the natural sequence? It seems that in general the numbers have to become infinite in the end, but can they still grow? I reply, that this growth to which they are susceptible evidently proves that they cannot at all be infinite; I say further that there is no last term in these sequences; a supposition of a last term already destroys the quintessence of these sequences which consists in the succession of the terms that can be followed by other terms, and these other terms again by others, all of the same nature as the preceding. This is to say, all finite, all composed of units; thus the supposition that a sequence has a last term, which is an infinite number, contradicts the definition of the number and the general law of sequences.

Most of our errors in metaphysics come from the reality that we give to ideas of deprivation. We know the finite, we recognize its real properties, we examine it, and when considering it after this examination we do not recognize it anymore, and believe to create a new being, whereas we only destroyed a part of what we had formerly known.

We must therefore consider infinity in small, in large, only as a deprivation, an entrenchment of the concept of the finite, which can be used like an assumption. In some cases, it can help to simplify the concepts, and generalize Sciences. Thus, all art is reduced to capitalizing on this assumption, to attempts of applying it to the subjects under consideration. All the merit is therefore in the application, in the possible use.

XXV. All our knowledge is based on relations and comparisons, everything in the Universe is therefore relation; and subject to measurement. Even all our ideas are relative have nothing absolute. As we have explained, there are different degrees of probabilities and certainty. And even evidence is more or less clear, more or less intense, according to the different aspects, to the relations under which it is presented. Truth transmitted and compared by different minds appears under more or less strong relations since, as formerly, the result of affirming or negating a proposition by all men in general seems to give weight to the truths that are best shown and most independent from any convention.

The properties of matter that appear to us evidently distinct from each other have no relation between them; extent cannot be compared to gravity, inscrutability to time, movement to surface, etc. These differing properties have in common only the underlying subject, and that ensures their being. Each of those properties considered separately asks therefore for a measure of its kind, that is, a measure different from all the others.

Notes by the translators

1. Reinforcement learning is an algorithm that tries to maximize payoffs under uncertainty by successively increasing the weight of better experiences (Gigerenzer & Selten 2001). In contrast, Buffon believes that learning focuses only on judgment and not on strategy.

2. So, we should warn readers at the outset that Buffon does not use the word *certainty* in the current absolute sense. We think of a certain event as one that is bound to happen, that is, an event that has probability 1 of happening. However, and rather schizophrenically, we also say that one event is more certain than another. This latter sense is that of Buffon. He refers to certainties of different orders and hence uses the word certainty as a synonym for probability or likelihood.

3. Neugebauer (2010) offered a concise survey of the literature and the concepts of solving the Petersburg paradox accompanied by experimental-economics evidence.

4. The treatment of very low probabilities remains a very controversial issue. As Selten (1998, p. 51) points out:

In general, it is very difficult to judge how small a very small probability should be. Usually there will be no good theoretical reasons to specify a probability as 10^{-5} *or rather than* 10^{-10}. *… The value judgment … that small differences between small probabilities should be taken very seriously and that a wrong description of something extremely improbable as having zero probability is an unforgivable sin … and is unacceptable*.

5. One may criticise Buffon for not referring to the correspondence of Nicholas Bernoulli (1687–1759). In 1728 he had referred to moral certainty in a letter written to Gabriel Cramer and it is likely that Buffon was aware of this fact. Nevertheless, the merit of presenting this explanation to a greater audience is an important contribution since the statement of Nicholas had not been widely heard.

6. This experiment seems to be the first conducted statistical experiment ever reported. It was replicated later by various researchers. Though the experimental approach is used to elicit the money value of the game, it is not an economic experiment, since the economic behaviour of subjects is not the purpose of study.

According to the definition by Sauermann and Selten (1967, p. 8, our translation) this condition is crucial:

It is advisable to attribute to experimental economic research only such experiments in which the economic behaviour of experimental subjects is observed. So-called simulation experiments, which only consist in the computation of numerical examples for theoretical models on electronic computers, do not belong, in this sense, to experimental economic research.

An economic experiment on valuation of lotteries would have to elicit the willingness to pay from experimental subjects. A frequently used elicitation procedure in this context is the BDM approach as used in Selten et al. (1999). For a more general introduction to economic experiments see Hey (1991).

7. Following the 25 described articles, Buffon dedicates another ten articles to "arithmetic and geometric measures". These articles look to us less relevant to the study of human sciences and are therefore omitted in the translation.

8. In fact, we do not intend to give here a fair and thorough appraisal of Reinhard Selten's contributions. Many important articles have been reprinted in Selten (1988, 1999). Buffon

9. Buffon uses the word *evidence*, which we also understand as something obvious or evident.

10. Buffon obviously means days rather than years.

11. See below the results of mortality tables. [They are not included in the translation.]

12. I communicated this idea to Daniel Bernoulli, one of the greatest geometricians of our century, and most experienced in all of the science of probabilities. Here is his response that he gave me in his letter, dated from Basel dated March 19, 1762.

I strongly approve, Sir, your way to estimate the limits of moral probabilities; you consult the nature of man by his actions, and you suppose indeed, that no one worries in the morning to die that day; hence, he will die according to you, with probability one in ten thousand; you conclude that one ten-thousandth of probability should not make any impression in the mind of man, and consequently this one ten-thousandth has to be regarded as an absolute nothing. This is doubtless the reasoning of a Mathematician-Philosopher; but this ingenious principle seems to lead to a lower quantity, because the absence of fear is certainly not in those who are already ill. I do not fight your principle, but it seems rather to lead to 100,000/1 rather than to 10,000/1

To make myself better understood, suppose that in a lottery where there is only a single prize and ten thousand blanks, a man takes only one ticket, I say that the probability to obtain the prize is only as one against ten thousand, his hope is zero since there is no more probability, that is, no reason to hope for the prize, than there are fears of death within twenty-four hours. This fear does not affect him in any way, so the hope for the prize must not affect him more, and even much less, since the intensity of the fear of death is much greater than the intensity of any other fear or of any other hope. If, despite the evidence of this demonstration, this man insists on wanting to hope, and if a similar lottery is played every day, and he persists in buying a new ticket every day, always hoping to win the prize, one could, to disabuse him, bet with him at equal odds that he will die before winning the prize.

It is the same in all the games, the bets, the perils, the risks; in all cases. In a word, if the probability is lower than 1/10,000, for us it must be, and it is in fact, absolutely zero; and by the same reason in all cases in which this probability [the chances are higher] *than 10,000/1, it is for us the moral certainty most complete.*

I confessed to Mr. Bernoulli that as the one ten-thousandth is taken according to the mortality tables that represent the *average man*, that is, men in general, well or sick, healthy or ill, strong or weak, there is perhaps a little more than ten thousand to bet against one, that a man, healthy and strong will not die in the 24 hours; but it is hardly necessary to increase this probability to one hundred thousand. Moreover, this difference, although very large, changes nothing in the main implications that I draw from my principle. Buffon

13. Old measure; one cubic fathom ≈ 6.12 m^3.

14. Here is what I left then in writing to Cramer, and of what I kept the original copy.

Montmort is satisfied to reply to Nic. Bernoulli that the equivalent is equal to the sum of the sequence 1/2, 1/2, 1/2, 1/2, etc. Ecus continued to infinity, i.e., ∞/2, and I do not believe that his mathematical calculation can be disputed. But far from giving an infinite equivalent, there is no man of common sense who will want to give twenty, or even ten pounds.

The reason of this contradiction seems to consist in the relation between money and its resulting advantage. A mathematician in his calculation estimates the money only by its quantity, i.e., by its numerical value, but the moral man must estimate it otherwise, only by the advantages or pleasure that it can obtain. It is certain that he must behave according to this view, and to estimate money only in proportion to the resulting advantages, and not relatively to the quantity which beyond certain limits cannot at all increase his happiness. For example, he will hardly be happier with thousand million than he would be with hundred, or with hundred thousand million more than with thousand million. He will be greatly mistaken to risk his money beyond certain limits. If, for example, 10,000 Ecus were all his goods, he will make an infinite mistake to risk them. I believe therefore that his mistake will be infinite, when these 10,000 pounds were a part of his necessary income, that is, when they are absolutely necessary for him to live, as he was raised and always lived. If that sum is as part of his superfluous income, his mistake diminishes, and the less they constitute his superfluous income the more diminishes his mistake; but it will never be zero, unless he considers this part of his superfluous income indifferent, or that he only considers the expected sum necessary to give him in proportion as much of pleasure as this very sum is larger than the one that he risks. We cannot at all give rules, since there are people for whom hope itself is a greater pleasure than they can experience by enjoying their stake.

To reason therefore more certainly, we must establish some principles; I say, for example, that the necessary income is equal to the sum that one is obliged to spend to continue living as always. The necessary income of a King is, for example, ten millions of revenues (because otherwise he will be a poor king). For a nobleman, 10,000 pounds of revenue (otherwise a decent man will be a poor nobleman); for a peasant, 500 pounds, because otherwise he will be miserable, he cannot spend less to live and nourish his family.

I suppose that the necessary income cannot give us new pleasures, or to speak more exactly, I will not account for any of the pleasures or advantages that we always had, and according to that, I define the superfluous income as what can provide other pleasures or new advantages; I also say that the loss of the necessary income is felt infinitely, it cannot be compensated by any hope. On the contrary the feeling of loss of the superfluous income is limited, and consequently it can be compensated.

I believe that gamblers feel this truth, because the loss, even if small, always gives more pain than an equal gain gives pleasure, and is that which leads to wounded pride, since I suppose that the game is indeed of pure chance. I also say that the quantity of the money within the necessary income is proportional to what it returns to us, but that within the superfluous income it begins diminishing, and the more the greater the superfluous income becomes.

I leave you, Sir, to judge these ideas, etc. Geneva, October 3 1730.

15. For this reason, one of our most skilful geometricians, the blessed M. Fontaine, introduced a declaration of Peter's goods, since indeed he can give as an equivalent only the totality of goods he owns. See this solution in his mathematical works, Paris, 1764.

General comment

The text had many unnecessary details and repetitions as well as insufficient explanations and the style is just so-so. No mistake here since the translators had slavishly copied it (see below). But about 2006 I saw the text of the same contribution as reprinted in 1954 devoid of such faults. The editors apparently improved the text which was quite proper since the original text certainly remained intact.

The translators of the text above were diletanti, absolutely ignorant of proper work, ignorant of mathematics (did not even know the term *exponent*). The enumeration of notes is very complicated and I am not sure that my changes are altogether correct. They preserved very long sentences with one or more semicolons inside, their text was sometimes mysterious and they had not made a slightest effort to improve their style. Here is a mild example from the beginning of § VII: *But it is not so easy to do the estimation here* … Follow my Comments on Buffon himself as seen in his contribution.

Notes by O. S.

[1] Definitions ought to have content and not contradict one another, but truths are alien to them.

[2] A low probability of a fire or of a loss of a ship in usual weather had been very early accounted for by insurance.

[3] Buffon thought that probability can exceed 1 and even be infinite,
[4] Descartes introduced moral certainty (apparently, in the first place, for jurisprudence) and Jacob Bernoulli (*Ars conjectandi*, end of chapter 2 of pt. 4) advocated its official introduction, again, for jurisprudence. Cournot (1843, § 47) introduced physical (as though instead of moral) certainty. Buffon did not refer to Descartes and, in turn, apparently no one followed him.

However, even Huygens (1669) indicated that certainty has many levels and proposed $(1 - 10^{-11})$, also without justification. Buffon (see below) selected the *human* insignificant probability by issuing from moral considerations.

[5] The 2000 was not explained. In addition, the exponent should have been multiplied by 365.

[6] Buffon published his mortality table in 1777, in t. 4 of the *Supplements* to *Nat. Hist*. I used another table (1989, p. 102). There, the probable duration of human life at age 56 was 13 years and 5 months, or about 4833 days. I added 15 days due to uncertainty of month and got 4848 days. For the uniform distribution which Buffon apparently applied, duration of life is 9696 rather than the stated 10,089.

[7] These considerations can seem unnecessary but Fienberg (1971) described a similar deviation from *uniform randomness*.

[8] Why *at each step*?

[9] Here, is apparently Buffon's reasoning. 50/150 = 1/3, 100/100 = 1 (a doubling). In the former case the increment is relative to the increased capital, in the latter, relative to the initial capital. So the gain increased the means of the first gambler by a half, but the loss of the second gambler halved his means. Below, Buffon provided a similar example. Denote now both gain and loss by α, $0 < \alpha < 1$ and assume that the capital of both gamblers is 1. Then always $\alpha/(1 + \alpha) < \alpha/1 - \alpha)$.

[10] For the ordinary reader, these enormous quantities of gold, these great numbers make little sense.

[11] It was called Petersburg game since Daniel published his memoir (1738) in that city. Nicholas Bernoulli invented it although for a die rather than for a coin, see Montmort (1708/1713, p. 402). In a letter of 1728 to the inventor (appended by Daniel to his memoir) Cramer replaced the die by a coin and in another letter of 1732 he proposed the term *moral expectation*. It was in connection with that game that Daniel introduced his celebrated moral expectation.

Condorcet (1784/1998, p. 394) reasonably argued that such a game represented only one trial and its study should be based on many games. Freudenthal (1951) independently voiced the same ideas and suggested that the roles of the gamblers should be each time decided by lot. On the early history of the Petersburg game see Spieß (1975), on its history in general, Jorland (1987), Dutka (1988), who studied it, as Buffon did, see below, by statistical simulation. One of the latest commentators was Aaronson (1978).

For a long time moral expectation remained fashionable and Laplace (1812, p. 189) called the classical expectation *mathematical*. His specification is still (undeservedly) used although, as it seems, only in the French and Russian literature.

[12] The *Memoirs* of the Petersburg Academy had not been published in the 18th century, and *Commentarii* should have been mentioned instead. *After Daniel Bernoulli* meant in t. 5 of that edition, but no Cramer's memoir is contained there. After all, Buffon possibly meant *after the text* of Daniel. Indeed, Daniel (1738) had appended the text of Cramer's letter of 1728 to Nicholas Bernoulli and acknowledged Cramer's priority in introducing moral expectation (although more primitive than his own).

[13] Suffice it to say *cells*. A *ring* (below) is the surface (the area) between the contours of any two figures. Denote the areas of the cells S and s, then the areas of the rings will be $S - s$. By the condition of the problem $S - s = s$. Then, Buffon undoubtedly meant filling only by congruent regular polygons or lozenges.

[14] Paragraph XXIV is only significant in that it shows that Buffon was not really versed in mathematics. He did not know that that science studies systems of notions which do not necessarily exist in nature. The last paragraph is barely interesting.

References

Bernoulli, Daniel, 1738. *Specimen theoriae novae de mensura sortis.*
English translation 1954 by L. Sommer with footnotes by Karl Menger, Exposition of a new theory on the measurement of risk, *Econometrica* 22(1), 23-36.
Bernoulli, Nicholas, 1728. Letter 9 to Gabriel Cramer. In B. L. van der Waerden, ed., *Werke*, Bd. 3. Jakob Bernoulli . Basel, 562-563.
Buffon, G. Le Clerc, 1777A. *Probabilités de la durée de la vie*. In: *Oeuvr. Compl.*, t. 3. Nouveau Edition, 1829, pp. 95-268.
--- 1777. *Essais d'Arithmetique Morale*. Ibidem, Nouveau Edition, pp. 338-405.
Feller, W., 1945. Note on the law of large numbers and 'fair' games. *Annals of Math. Stat.* 16, 301–304.
Gigerenzer, G., Selten, R., 2001. *Bounded Rationality: the Adaptive Toolbox*. Cambridge & London.
Hey, J. D., 1991. *Experiments in Economics*. Blackwell.
Lopez, Lola L., 1987. Between hope and fear: The psychology of risk. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*.
Menger, K., 1934. Das Unsicherheitsmoment in der Wertlehre. *Z. der Nationalökonomie* 51, pp. 459-485.
Neugebauer, T., 2010. *Moral impossibility in the Peterburg gamble: a literature survey and experimental evidence*. LSF working papers, Univ. Luxembourg.
Sauermann, H., & Selten, R., 1967. Zur Entwicklung der Experimentellen Wirtschaftsforschung. In Sauermann, H. (Ed.) *Beiträge zur Experimentellen Wirtschaftsforschung*. Tübingen, 1-8.
Selten, R., 1988. *Models of Strategic Rationality. Theory and Decision.* Library, Series C: Game Theory, Mathematical Programming and Operations Research, Dordrecht-Boston-London.
Selten, R., Buchta, J., 1994. *Experimental sealed-bid first price auctions with directly observed bid functions*. Discussion Paper B-270, Univ. of Bonn. In: Budescu, D., Erev, I., Zwick, R. (Eds.), 1999. *Games and Human Behaviour: Essays in Honour of Amnon Rapoport*. Lawrenz Associates Mahwah, NJ.
Selten, R., 1998, Axiomatic characterization of the quadratic scoring rule. *Experimental Economics* 1(1), 43-61.
Selten, R., Sadrieh, A., Abbink, K., 1999. Money does not induce risk neutral behaviour, but binary lotteries do even worse. *Theory and Decision* 46, 211-249.
Selten, R., 1999. *Game Theory and Economic Behaviour: Sel. Essays*. 2. Vol. [vol. 2?].Cheltenham-Northhampton.
---, 2004. Learning direction theory and impulse balance equilibrium. In Cassar, Alessandra and Friedman, D. (eds.) *Economics Lab: An Intensive Course in Experimental Economics.*
Selten, R., Neugebauer, T., 2006. Individual behaviour of first-price auctions: The Importance of feedback information in experimental markets. *Games and Economic Behaviour* 54: 183-204.
Selten, R. Chmura, Th., 2008. Stationary concepts for experimental 2x2 games. *Amer. Econ. Rev*. 98(3), 938-966.
Taleb, N., 2007. *The Black Swan, The Impact of the Highly Improbable*. Random House.
Tversky, A., Kahneman, D.: (1992). Advances in prospect theory: cumulative representation of uncertainty. *J. Risk and Uncertainty* 5: 297-323.
Weil, F., 1961. La correspondance Buffon-Cramer. *Rev. d'Hist. des Sciences et de leurs Appl.* 14 no 2, pp. 97-136.

Aaranson J. (1978), Sur le jeu de Saint-Pétersburg. *C. r. Acad. Sci. Paris*, t. A286, No. 3, 839 – 842.
Bernouli Jacob (1975), *Werke*, Bd. 3. Hrsg. B. L. van der Waerden. Basel.
Borel E. (1943, 1946, 1993). *Les probabilités et la vie*. Paris.
Buffon J. L. L. (1989), *Histoire naturel de l'homme et des animaux*. Paris. Includes reference to data of 1772. Apparently composed of two volumes or their parts.
Condorcet M. J. A. N. de Caritat (1784), *Sur le calcul des probabilités*. In

author’s *Arithmétique sociale*. Paris. Editors B. Bru et al, 385 – 436.
Cournot A. A. (1843), *Exposition de la théorie des chances et des probabilités*. Paris, 1984. Editor B. Bru. **S, G,** 54.
Czuber E. (1884), *Geometrische Wahrscheinlichkeit*. Leipzig.
Dutka J. (1988), On the St. Petersburg paradox. *Arch. Hist. Ex. Sci*., vol. 39, 13 – 29.
Fienberg S. E. (1951), Randomisation and social affairs: the 1970 draft lottery. *Science*, vol. 171, 255 – 261.
Fontaine de Bertin A. (1764), *Mémoires* … Paris.
Freudenthal H. (1951), Das Petersburger Problem in Hinblick auf Grenzwertsätze der Wahrscheinlichkeitsrechnung. *Math. Nachr.*, Bd. 4, 184 – 192.
Huygens C. (1669), Correspondance. Published in *Oeuvr. compl*., t. 6. La Haye, 1895, 531 – 532. **S, G,** 85.
Jorland G. (1987), The Saint-Petersburg paradox, 1713 – 1937. In *Probabilistic revolution*, vols. 1 – 2. Cambridge (Mass.), vol. 1. Editor L. Krüger et al., 157 – 190.
Laplace P. S. (1812), *Théorie analytique des probabilités*. *Oeuvr. compl*., t. 7. Paris, 1886.
Montmort P. R. (1708, 1713), *Essay d’analyse sur les jeux de hazard*. New York, 1980.
Spieß O. (1975), Zur Vorgescichte des Petersburger Problems. In Bernoulli Jakob (1975, pp. 557 – 567).

II

Galton, Pearson, Fisher … Gone with the wind?

The students at University College London accused Galton and Pearson (not yet Fisher) of racism and demanded that Galton's name be removed from buildings and the Lecture Theatre and the same concerned the Pearson Building, see R. Langkjaer-Bain, Cause for concern: Galton's troubled legacy, *Significance*, vol. 16, No. 3, 2019, pp. 16 – 22, a journal of the Royal Statistical Society London.

His paper is balanced; I find there a cautious comment of Deborah Ashby, the RSS President: there should be no *blanket condemnation of the whole man* (of Galton); the author's remarks: *Those who judge Galton may be judged as harshly in future; Such* (racist) *views were not unusual for the period* (he also mentions Fisher and a few other eminent scholars), but Debora Ashby more accurately stated that *almost the entire population* held such views. Finally, here is Ernst Haeckel (1914), *The History of Creation*, vol. 2, p. 429 (sixth edition):

The Caucasian … man has from the time immemorial been placed at the head of all races of man.

But there, in *Significance*, is also a blunt statement of Professor Joe Cain, head of the UCL department of science and technology studies. He *told the University he will not lecture in the theatre as long as Galton's name is on it*. The RSS had set up a commission to decide the *troubled* issue of renaming buildings *this summer*, and the students had certainly won. I very much doubt that a black professor, the head of that commission, had objectively decided that issue. The RSS had succumbed to the hoi polloi!

I defend the accused since defence is still needed. Aristotle (*Politics*) declared that slavery was necessary and useful; until 1861, when serfdom was abolished in Russia, many if not most Russian public and literary figures were serf-holders. And what about the immense tribe of anti-Semites? Martin Luther, the leader of German Reformation, was its prime member, his wicked German book *On the Jews and Their Lies* appeared in 1543. So down with Aristotle? With Luther?

Just suppose now that Professor Cain (let alone some of those students) is an anti-Semite. Would he hesitate to denounce Galton? No, with a capital n. Anti-Semitism is now almost *comme il faut*! Forgotten is Disraeli's prophetic saying: *The Lord deals with the nations as the nations deal with the Jews*! Yes, prophetic since militant Muslims, who replaced the millions of Europe's exterminated Jews, are already all but governing its great chunks.

So here the students cum professor can see their real aim, but they are blind as so many bats. Racism has reversed its sign. Many black people and Arabs (and certainly the militant Muslims) are the racists today. And anti-Semitism is flourishing.

A Jew called Jesus appropriately asked (Matthew 7:2-5): *Why do you look at the speck of sawdust in your brother's eye and pay no attention to the plank in your own eye*?

Will those students rest content with their ill-begotten victory? Hardly. I recall the experience of Soviet statistics. In 1909, Lenin called Pearson *a conscientious and honest enemy of materialism* and *one of the most consistent and lucid Machians* (pp. 190 and 274 in the 1961 edition of his *Materialism and Empiriocriticism*). And for a very long time Soviet statisticians refused to study Pearson's contributions. Some were simply afraid, others honestly obeyed the obvious command, but at the same time many of them had thus been happily concealing their ignorance.

So I think that many British students (not only from UCL) will now study Pearson against their wishes (*malgré lui*), as little as possible, and the result will be almost pitiful.

The decision carried be the RSS is lengthy, see

https://www.ucl.ac.uk/provost/sites/provost/files/recommendations-ucl-eugenics-inquiry-more-group-
university-college-london-february-2020.pdf

https://www.ucl.ac.uk/provost/sites/provost/files/ucl history of eugenics inquiry report.pdf

I summarise. Ten members, a subgroup, of the special committee which was set up to resolve the entire issue at hand, included eight academics, five of them professors (yes, Cain was certainly among them), studied the entire problem of eugenics and its role in society and understandably the names of Galton and Pearson were its victims. I am sufficiently qualified to add: none of them was a historian of statistics. Many lengthy quotations from Galton, and some from Pearson, are included complete with a conclusion: there exists

A clear link between Galton's eugenics, imperialism and national socialism [and Nazi Germany].

The need to study the history of eugenics is stressed. But history of science is a delicate subject and can hardly be promoted by present-day specialists. In addition, lecturers will likely remain on the safe side and be biased (guess: in what sense?).

I note that even pseudo-sciences had played a positive role: the history of alchemy is the prehistory of chemistry; western astrology tended to apply mathematical methods and attempted to separate usual and unusual events (cf.: to separate randomness from determinism) whereas Kepler the astrologer introduced qualitative correlation between heaven ad earth. And here is one of the conclusions of that subgroup:

Eugenics and the idea of heredity were the basis for his [Galton's] *idea of 'regression to the mean*'.

Above, I have expressed my opinion about the blindness of the RSS attitude and I hope that a time comes when Galton and Pearson will be appraised in the proper historical context. Their Weltanschauung was unavoidable, and had Professor Cain been their contemporary, he would be their companion.

Francis Galton (1822 – 1911), polymath, father of the Biometric school, main promoter of fingerprints, discoverer of anticyclones. Invented the word *eugenics*, studied and controversially advocated it, Was a pioneer in studying heredity. Awarded the Darwin medal by the Royal Society.

Karl Pearson (1857 – 1936), Fellow of Royal Society. Main originator of the Biometric school and main predecessor of Fisher in mathematical statistics. Promoted eugenics, denied (at least in words) *negative* eugenics. His apparently reasonable recommendations were either unrealisable or would have led to most deplorable results.

Ronald Fisher (1890 – 1962), Fellow of Royal Society, main originator of modern mathematical statistics, geneticist and eugenicist. Studied eugenics in Pearson's spirit and even more resolutely.

Je vous approuve en tous points, la "politiquement correct" en histoire des sciences est une absurdité et une régression aux temps de l'inquisition ou des procès staliniens. B. Bru: Professor, Paris

Fisher's memorial window to be removed from Gonwille and Caius College, Cambridge. Dr. R. W. Farebrother, Liverpool

And now I enlarge on our situation since the episode described above is but a tiny illustration of our general situation. In 1918 – 1922 O. Spengler published a book in German (two volumes) whose title, in English, is *Decline of the West*. No, not Decline but Downfall!

It was seen long ago by the widespread feeling of guilt: we, the white race, had enslaved the blacks. So now we are aiding and abetting black racism … But there hardly exists a single nation which, over the centuries, did not perpetrate crimes against some other people. And Catholicism committed horrible deeds against millions of innocents.

Judges and courts are known to guide themselves by dated formalities, therefore to exonerate criminals and award large moneys to other obvious criminals and victims of their own stupidity. Unbelievably, some prisoners are living quite comfortably. The famous Breivik, who shot about 80 youngsters in cold blood, ought to be happy with his life in prison in a spacious cell with television and computer and his wife often staying with him. And he was awarded very considerable compensation for … having received lukewarm coffee. Poor Breivik!

All this is downfall pure and simple! It is occurring since the natural, necessary requirements of a state are disregarded en masse, often openly. And the birthrate of the white race is low enough to ensure its virtual disappearance …

I recall some authors whose names I did not memorize: the appearance of human beings was a mistake of the Creator or evolution …

III

On A. N. Kolmogorov's letters to V. P. Efroimzon

Introduction

1. Vladimir Pavlovich Efroimson (E.), 1908 – 1989, was a geneticist, Doctor of biological sciences and co-founder of national genetics. In 1929 he was expelled from a university for defending Chetverikov, a most prominent geneticist, and I have not seen anywhere that he had ever graduated from a university.

In 1932, he did three years for participating in the Free philosophical society. I did not establish it, but anything *free* was an anathema! In 1949 – 1955 he did time once more for allegedly slandering the soldiers of the Red Army in the aftermath of the war. And then E. experienced great difficulties when applying for a post. The usual true cause of his new difficulties (to put it mildly) in the post-war period was his ardent denunciation of Lysenko, Stalin's battering ram for subduing the entire science. In the above, I used the entry on E. in vol. 10, 2001, of the (*Kratkaia*?) *elektronnaia evreiskaia enz.* (Short (?) Electronic Jewish Enc.). It seems to be translated in the Internet.

E.'s study of Lysenko and Lysenkoism was published in pieces in all four yearly issues of *Voprosy Istorii Estestvoznania i Tekhniki* in 1989. There, E. (No. 3, p. 102 note) added, regrettably without substantiation, that genetics was rooted out in Nazi Germany.

Fisher (1948) also attacked Lysenko, and in addition I quote the opinion of Kolman (1982, pp. 213 – 214):

I was disgusted since his opponent, the official Vavilov school at the Academy of Agricultural Sciences, had for a long time prevented him from practically proving his innovatory ideas, slighted him since he, a provincial agronomist-breeder lacking higher education, invaded the sanctum sanctorum of those pontiffs of science. And I was delighted by the enthusiasm with which he developed his concepts.

At the beginning he sincerely believed in being right and ardently upheld his ideas, but, after gaining authority and having felt power, he did not mind anymore to apply administrative, forceful methods of struggling with his convinced enemies. Who knows whether he himself had not repeatedly participated in hounding them to death or that he did not "doctor" his experiments if they had not confirmed his theory.

In December 1985, during a premiere of a film documentary about Vavilov, the leading Soviet geneticist and a member of the Royal Society, E. spoke out without permission, not mincing his words:

Vavilov did not die [in the labour camp], *he croaked like a stray dog from hunger and cold.*

Oh, yes! Vavilov was guilty since he impulsively promised that very soon genetics will achieve grand practical results.

E. also stated that the Soviet Union was a *land of slaves governed by nomenklatura thugs* (shpana).

E.'s main works are (all in Russian): *Genetics of Genius* (Genetika i genialnost), 1998. Apart from its constitutive writing it contains *Pedagogic Genetics* (Pedagogicheskaia genetika), 1998 which was written in 1974 – 1977, and a paper *Origin of altruism* first published in 1971 in an adapted form. Apparently not included was G*enetika etiki i estetiki* (*Genetics of ethics and aesthetic*) (1995). I also mention the included booklet *Predposylki genialnosti* (*Preconditions of Genius,* 1998) and a manuscript on the history of Jews (not included). Not included either was *Vvedenie v medizinskuyu genetiku* (Intro. into Med. Genetics). The proof of E.'s early Russian contribution *Genetics of Silkworm* was scattered perhaps of some infringement on dialectical Marxism.

And here is a quotation from *Preconditions* … (part 1, end of chapter 4) which hints at Israel, cf. Note 7:

Even a small country of, say, five million inhabitants, but having developed and realized 10% of its potential geniuses and talented men, will after 50 years leave behind a country of a hundred times more inhabitants which left barriers for the development and realization of its potential geniuses.

I glanced at *Genetics and Genius* in the Internet and E.'s statements described in my Notes are from that source that lacked paging.

2. Kolmogorov's letters likely contain something barely known about his work with school students and his interest in psychoses. Incidentally, I think that he mentioned Stalin's psychosis or psychoses in one of the places blackened by the Archive. And now a reservation: Pontriagin (1980), that anti-Semite supreme alongside Vinogradov and Shafarevich, justly stated that Kolmogorov's recommendations concerning school students in general were sky-high above reality.

In a mildly form the chair of mathematics at the Plekhanov Institute in Moscow where I worked remarked that the graduates of the Kolmogorov boarding school were not attuned to applied mathematics.

Yes, geniality is fraught with inconvenience for ordinary people, and I myself experienced Kolmogorov's impatience on a tiny scale (Gnedenko and Sheynin 1978). I noted the appearance of the Dirac delta-function in Laplace, but Kolmogorov, the main editor of the source, struck out my discovery since it made no sense in the language of generalized functions (although was still noteworthy!).

Concerning § 15 I note that Kolmogorov was a Russian (but certainly not a Soviet) patriot. As a hardly needed illustration I recall a chair of mathematics in Berlin, a Russian, telling me that Kolmogorov once swam in cold water and commented afterwards: *We are Russians, not Germans* or words to that effect. See also Letter 2, end of Item 3, and Note 15. It seems that at heart Kolmogorov was devoted to socialism *with a human face*.

The very fact of Kolmogorov's correspondence with E. is noteworthy. He also talked with E. over the telephone (§ 15, P. S.) and mentioned their future meeting (did it occur?). It remains unknown which of his contributions had E. sent Kolmogorov. But anyway, it was likely a draft.

3. But where are E.'s letters? Here is an edited text of my tiny publication (*Math. Intelligencer*, vol. 39, No. 4, 2017, p. 46):

Where are Kolmogorov's posthumous papers?

In a worthwhile tradition, the Archive of the Russian Academy of Sciences (RAN) collects and keeps the posthumous papers of its late members. Kolmogorov died in 1987, so I asked the Academy for permission to look at his papers. I found that RAN did not have them. Staff at their Archive advised me to inquire at the Archive of Moscow University where Kolmogorov had been a staff professor. I had twice inquired there but received no answer and asked the Presidium of RAN. An anonymous representative from the Class of Mathematical Sciences answered in writing that nothing was known about Kolmogorov's papers. Period! They obviously did not dare say anything more.

A colleague told me that Albert Shyraev, professor at Moscow University, perhaps keeps those papers. Twice I wrote to him but received no answer.

Shyraev (albertsh@mi.ras.ru)! He hurriedly published a paper (1989) which described Kolmogorov's merits in mathematics complete with a list of his publications. After its extremely superficial examination I found two omissions; in addition, translations of his works were not mentioned. But the main point is that Shyraev is unscrupulous. Novikov (1997) explained how irresponsibly he managed to promote that crazy Fomenko and in his § 3 washed his hands of the business: *Allow me to keep silent about Shyraev's role*.

Another episode is insignificant as compared with the above but just as disgusting. In 2001, the yearly journal *Istoriko-Matematicheskie Issledovania* published a paper by Yu. V. Chaikovsky who, without even a trace of justification, invented the Jacob Bernoulli – Cardano law of large numbers. I was member of the editorial board, did not know anything beforehand and resigned. The Editor, S. S. Demidov, *explained*: Shyraev recommended the manuscript. Such an obliging person is really needed, and he is now President of the International Academy of History of Science. That scientific body had however degenerated and is hardly needed at all.

Quite recently I searched for Kolmogorov's papers anew and it really seems that the situation had not changed. Furthermore, I found out that the financial circumstances of the Archive of RAN are horrible so that even their inestimable treasures are in danger.

I also found out that the *Archive* is keeping E's papers in a special *fond*. Twice asked them to prepare for me a copy of his letters to Kolmogorov, but received no answer, perhaps because of those circumstances.

Kolmogorov wrote both You, Yours and you, yours and I left it at that. A few sentences were grammatically wrong and I corrected them. Then, some words were also grammatically wrong and I italicized them in translation which sometimes seems curious.

Both letters from Kolmogorov are kept by the Archive of RAN, Fond 2024, Inventory 1, Delo 354, pp. 1 – 11

Letter 1, 10 Dec. 1977

Dear colleague, I am sending you my remarks to ensure at once the possibility for you to decide to what extent we are fellow-travellers and not to form exaggerated assumptions. I found much interesting in your manuscript and hope to find *notlittle* (ne malo) at our meeting.

1. About the genius of great men

At the beginning of the previous century France needed one single emperor. It is very difficult to estimate how many candidates potentially fit for filling that post were among the officers promoted by the revolution. Similarly, it is unclear what measure of "genius" in military leadership, management and personal courage needed Joan of Arc for accomplishing her mission. The information which you found in the *Larousse* dictionary[1] about her constitution is very interesting. It is naturally connected with her disposition to have hallucinations. Then follow the milieu and the sense of her exceptional mission. The war [apparently, the German – Soviet war of 1941 – 1945] showed us that in an appropriate situation courage reaching utmost limits is not so exceptional.

2. General and specific endowment

I was extremely surprised and, I would say, saddened by what you had stated about that subject on pp. 45 – 47. A "titanic purposefulness" without a proper point of application seems to me *some what* (chem to) abnormal and quite *un desirable* (ne zhelatelno). Then, to discern in proper time and cultivate special gifts is not at all simple and is [even] central for the system of upbringing.

3. To take care of the children of talented men and geniuses is naturally the duty of their parents[1].

If a talent is really inherited, it usually does not vanish. More important is the problem about the inborn and acquired components of talent. How strong is the former if talent had not revealed itself in previous generations because of its "polygene" feature. This problem interests us when planning a system of upbringing and education.

4. Geniuses, talent and psychoses

When assuming that the manic depressive psychosis[1] and schizophrenia are the two main psychoses it would be natural to turn our attention on both. As it seems, this is indeed happening in the literature. I know well enough the data on Moscow mathematicians. Quite pronounced manic depressive psychosis **text blackened by the Archive**. Schizophrenics among us are much oftener. You certainly know better, but the situation with musicians is possibly the same.

5. Gout

For me, this subject was new and as far as I understand, your great work of four years was inserted just there [devoted …]. I do not dare criticize your final conclusions, but I note that the method of statistical comparisons with a constant frequency of 0.4% seems to me unfounded[1].

The frequency of gout in various social strata and times is apparently sharply different so that comparisons should be made between homogeneous groups. I have not statistically studied it, but after simply following our classical literature it apparently becomes possible to establish that the gout was most widely spread among the Russian nobility of the nineteenth century[1]. And the frequency of gout among talented men and geniuses, who belonged to that nobility, ought to be compared to that particular frequency.

It is possible to approach such comparisons by selecting a random sample of those families which had not revealed special talents and studying the archives of their remembrances and letters.

6. Uric acid

Is it possible to check directly the correlation between its concentration and mental activity? Otherwise the hypothesis remains not too convincing[1].

7. Four mechanisms

One of them is constructed on a single example of Joan of Arc. The second one on three examples (Lincoln, Anderson and Prof. Nikolsky[1]). I would not name them in a general reasoning. But the existence of the "schizoid" type of talented men and geniuses seems doubtless.

Are not the chemical and hormonal explanations too categorical? The example of enormous quantities of drunk coffee (Napoleon and others) does not seem fortunate. It will be then too easy to become the emperor of France.

I got the impression that you regret the impossibility of stimulating talent and genius just by inserting uric acid. And *there fore* (iz za) you recommend much more complicated methods of stimulation.

8. Kolmogorov left out this number

9. Selection of talent

In the narrow field of mathematicians I may be considered a specialist. The boarding school which I head provides *notbad* (ne plokhie) results. We have to select at age fifteen. Had we better possibilities of establishing summer camps for teenagers 13 or 14 years old and of selecting from them after becoming closely acquainted with each , we would have preferred this lesser age. But still 13 years seems doubtful. And I will certainly advise not to bother with those of 12 years.

<u>Specific</u> abilities which we need are apparently formed later. Most members of mathematical study groups for those of 12 years later scatter. Girls occupy there the first place but already at 15 years of age most of them lose interest in mathematics.

This certainly does not mean that ability is not needed for studying mathematics. Suitable training should begin earlier. But this general quick-wittedness and ingenuity can be successfully developed even by fishing, birdwatching, playing games etc. Gifted "wild" boys, if being interested in mathematics at age 14, can already at 19 publish their own scientific work. **Text blackened**.

In music serious training of receptivity and technique should certainly begin earlier, and still earlier for circus performers and sportsmen since sport became professional[1].

10. Speeding up development

Freedom is *undoubtedly* (nesomneno) better, therefore we certainly should not prohibit external school-leaving examinations. I think however that both parents and teenagers should be warned that that method is dubious. And in any case no preparatory summer schools for external examinations ought to be established.

I have a rather considerable personal experience of work with child prodigies and in particular of their considerable frustration which happens sometimes. I tell them that, in the gymnasium, the greatest mathematician of our century, Hilbert, as he himself said, did not hurry too much to study mathematics since being sure that in due time he will become an excellent mathematician[1].

11. Tests

The ban on tests is now lifted. Our Academy of Pedagogic Sciences applies them and, in appropriate instances, recommends tests. They certainly reign supreme in the Anglo-American world. Last year I attended an International Congress on Mathematical Education in Karlsruhe, Germany. And I can confirm that in our field, in Germany and France, tests play a most modest role. They strongly criticize the English system of selection of 10 – 12 year old children for entering "grammar schools" which lead to universities.

In France and Germany I feel myself at home, but I little know America [the USA] on the occasion of having *nocommand* (iz za nevladenia) of English[10]. But still I suspect that you exaggerate the value of the MERIT programme[10]. In America [in the USA] everything at once assumes an immense scale and is skilfully advertised. But did this programme become the main method of promoting gifted youths? I will ascertain this as far as it concerns student-mathematicians by asking my American colleagues. For the present, I would be grateful for indicating the materials which are at your disposal.

12. Early childhood

Here, I quite sympathise with you. Allowing for all the conditional character of the IQ, a publication of the data based on that indicator would have been useful[11]. I heard that considerable measures to ensure the mothers a possibility of remaining home with children during the first years of their life are, or are being implemented in Hungary. If you speak out on this subjects you need to inquire and secure precise information.

13. Cutting down the size of school classes

It seems that *some thing* (chto to) is done in that direction once more in Hungary. In France, until recently, classes were separated in two groups, as it is done here with respect to foreign languages[12]. Gabi had recently abolished this system but, instead, curtailed the size of classes to 30 school students. Regrettably, 40 students are already planned here for many years ahead. It would be very important to achieve changes here.

14. Editing the proposals

For justifying your proposals (§ 16 [where is it?]) very little is needed from genetics and age-specific psychology. It would be reasonable to restrict the appropriate document by the necessary only.

15. Patriotic motifs and the criticism of capitalism

In a paper and in the respective [future] address such passages are extremely unfortunate. "Plutocracy" which "was compelled" to allow the democratic forces to enter its milieu etc. are fantastic [expressions] and will favourably impress no one.

With deep respect [signature follows]

PS. I found the book of Volotskoy. It occurred that I meant exactly it when speaking with you over the telephone. It begins by a short introduction by P. M. Zinoviev which is however less substantial than I thought before seeing it. But the last, the twelfth chapter seems interesting. There, in accord with Kretschmer[13], the cyclic and the schizoid characters are described by fluctuations between the two appropriate poles with a third epileptoidnic polarity. The latter did not apparently find a wide response.

After glancing at your example of cyclic geniuses and talented men I began to think that schizoids are also found there. Is it so? I have recently reread Oscar Wilde and was astonished that the gout was apparently extremely usual in the circles of the English society which he described.

And my suspicion that here we have to do with what is called nonsense correlation had essentially strengthened.

Letter 2. 19 Jan. 1978

Highly respected colleague,

I begin with what offended you in my first letter. I should have apparently avoided any irony when speaking about the uselessness of some passages in your *proiect* (proэkt of an appeal to high instances. But my aim was quite serious: to caution you against a mistake. An excessive ideological zeal in such documents impresses our leading circles in quite the opposite direction: it provokes mistrust.

But somewhat later about the outlook of some or other appeals to the top people. At first I would like to appear as your *assistent* (pomoshnik) in in the search for truth.

1. I see no need to charge my collaborators with verifying your card indices etc. Let us issue from assuming that they were compiled conscientiously. But any statistician would have turned your attention to the danger of what is called “nonsense correlation”. In any publication you should prevent beforehand such objections rather than “give way to despair” because of my fault finding. I did not suggest to begin in earnest the story of the prevalence of gout among the Russian nobility. However, such studies, for example, of various layers of the English society had been probably accomplished long ago.

But you shouldn’t appeal to the public without discussing such issues of the methodology of the statistical approach to the business at hand.

I became interested in the data about the professors of the Michigan University. But here also a suspicion of nonsense correlation concerning age appears at once. It is curious that the “prevalence of investigative interest” offers exactly the least correlation with the uric acid. But I suspect that the American researchers themselves had foreseen such an objection and somehow warded it off.

Generally speaking, I note however that you had convinced me in that the role of the ill-starred uric acid is similar to the part of caffeine. It was new to me that this issue was widely illuminated in the literature before you. The appearance of an essay in Russian about the correlation of the prevalence of gout with endowments and the stimulating action of the uric acid, for example in the journal *Priroda,* would certainly be very desirable.

2. Any essay about the correlation of endowments with psychoses should touch not only manic depressive psychosis but schizophrenia as well. You write that there are many examples among artists and men of letters. Do not **text blackened by the Archive** belong to them? I shall not yet enumerate real and gifted schizophrenics and schizoids in the younger generation, but there are many of them.

Among mathematicians of the 20^{th} century beyond our country I name L. E. J. Brouwer, the founder of mathematical intuitionalism and a topologist. Poincaré ad Hilbert would have deserved attention. It is curious that Brouwer and **text blackened** … open the long list of the representatives of mathematical logic. This is already a detail which definitely seems not to be random. In general, a vast literature apparently exists about talented men and geniuses among schizophenics. Do you belong there **text blackened** …

For realization, the talents of schizophrenics naturally need prolonged remissions. The course of the illness is cyclic. As far as I know, such courses do not give grounds for confusing it with the manic depressive psychosis to which a cyclic course is predominantly ascribed.

I note in passing that Ivanovs, in the *Dostoevsky Clan* [Volotskoy (1934)] were undoubtedly schizophrenics and schizoids. Volotskoy's opinion that schizoid *sidebyside* (na riadu) with prevalent "epileptoid" features were a specific feature of Dostoevsky himself as well, does not seem to me that nonsensical. I only met with the concept of "epileptoids" in Volotskoy's book and in Zinoviev's introduction to it. I do not know whether it was widely recognized. But in any case a fine essay in Russian on correlations of endowments with psychoses would have been desirable.

3. I intend to ask Neyman, the head of American statisticians, about the results of the programme of revealing talented men by means of tests and about subsequent support for them. Regrettably, correspondence is slow. I do not wait an answer (?) in the near future.

I am grateful to you for sending me a copy [copies] from the journal *Amerika*[14]. They had not convinced me that the notorious programme worth sixty million had occupied such a central place like you imagine in the activities directed at promoting talented men in the USA. It is curious that you require dozens of billions.

For the time being this is all that I can offer as assistance to your inquiries. In your last letter you strengthen still more your horrible forecasts (in six or seven years the USSR will lose its rank of a superpower). I think that this is fantastic. I know well enough the deficiencies of our system and even the danger of their aggravation (see the latest decision about the school[15]) but happily our competitors and in particular our main competitor, the USA, have their own deficiencies. When we meet I can tell you about my estimates of the future and the observations on which they are founded. Now, however, I formulate something about my mood which will hardly change because of our meeting.

A. I have written about considering definitely unfortunate any mention of studies of psychoses, uric acid etc. along with proposals for pre-school and school upbringing.

B. Proposals about pre-school upbringing with justification of the importance of its individual character ought to come, in the first place, from psychologists. Perhaps it will be possible to cooperate, for example, with Zenkov[16] and his collaborators. But argumentation by means of genetics can also play a certain part. It will be really essential to base ourselves on the achievements of the socialist countries of Mid-Europe [Central Europe] which had apparently overcome us.

C. It is beneficial to propagandise tests. But I do not see any definite programme of promoting talented men founded exactly on tests. Ideal are certainly studies open for all when selection is accomplished all by itself: the lazier themselves will scatter. And in many directions we are not so far from such an ideal. But if a competition with appraisal is indeed unavoidable, tests will be useful, although only in a secondary role.

Thus, higher institutions entrance examinations: a test for elimination is certainly unfit (negodny). A serious written worksheet followed by an interview is needed. Therefore I imagine that your entire concept of an all-embracing "testing" and education of specialists "testologists" is mistaken. For example, tests of mathematics will certainly be better when compiled by able mathematicians somewhat acquainted with that task.

E. [D is missed.] Proposals about curtailing the size of classes and about special work with school students of the higher forms are certainly very important. But I think that they should be put forward independently; references to genetics can rather hinder. Details also during personal meeting.

For me, talks with you will be interesting, but I thought it beneficial to disappoint you beforehand by establishing definite bounds for the matter about which we agree. In the issues in which I myself have at least *some thing* (kakoi to) like a small weight (school students of higher forms) I would not see much benefit from my support of your all-embracing proposals. I would rather lessen my capability of doing *some thing* (koe chto) useful.

Yours (signature)

The Archive appended Kolmogorov's postal address (in a building for the staff of Moscow University).

Notes

1. *Larousse*: a multi-volume encyclopaedic dictionary. E. (=Efroimson) and apparently Kolmogorov thought about its ten-volume edition of 1960 – 1964 with an additional volume in 1968. E.: Joan of Arc's behaviour was determined by her Morris syndrome (in inborn disturbance of the gender development). In his *Genetics and Genius* E. considers in detail various syndromes and stimulation by gout, see Notes 3 and 5.

2. Item 3 was not separated from the context. And here is the difference between talent and genius (E., without mentioning any source): a genius creates what he is obliged, a talented man creates what he can.

3. The manic depressive psychosis: in its manic phase the mental process is accelerated. E. apparently also mentioned schizoids and epileptoids. The former are submerged in their inner world with a preponderance of abstract thinking, the latter: explosiveness untidiness.

4. The method of statistical comparison is likely the rank correlation.

5. Here is N. I. Nekrasov, *Who Is Happy in Russia*: A house-serf boasts that he earned his gout by drinking much expensive wine, so that in this respect he is a nobleman. Gout was thought to be caused by gluttony, heavy drinking and various excesses.

E. described in detail how the victims of gout almost became the vehicle for the history of society. In particular, he mentioned Boris Godunov as an outstanding statesman who had nothing to do with the assassination of Prince Dmitry. Historians knew it but were afraid to oppose widespread calumny.

6. Uric acid is structurally very similar to caffeine and the stimulation of the brain by gout can elevate its activity to the level of talent and geniality (E., partly supporting himself by a source of 1955).

7. Anderson, likely the physicist and Nobel-prize winner Carl David Anderson (1905 – 1991). E. thought that Lincoln, Anderson (and de Gaulle) had the Marfan syndrome (a form of gigantism). Kolmogorov (title of § 7) mentioned four mechanisms, E. mentioned four conditions determined by the society which are necessary but not sufficient for the appearance of geniuses. In the first place, as he thought, they emerge after the breakdown of caste, class and other restraints. Here is Novikov (1997, p. 72), about the crazy A. T. Fomenko who curtailed the chronology of civilization by one and a half thousand years:

As it appears, the 75 year old Nikolsky was mightily attracted by the new theory and communicated his manuscript for publication.

Nikolsky was an academician and an eminent mathematician and Novikov certainly had not hinted at any psychosis. The moral atmosphere which reigned in the Soviet Academy is shown by Fomenko's carrier: he was elevated to the very top and managed the science of the land.

8. Professional sport did not officially exist in the Soviet Union, but leading sportsmen's way of life had been *professional* except for payment.

9. Geniuses are exceptions and ought to be studied individually. Gauss behaved quite differently.

10. At least Kolmogorov easily read English literature apparently including fiction.

11. Quoting another author, E. stated that the IQ was a most important instrument which prevented science from degenerating into a system of castes. It is certainly difficult to compile a test, and a tested person could have better or worse answered another question. In his *Pedagogic Genetics* E. said that the IQ test was banned since in the first place the needed people were those devoted to the authorities rather than the clever ones.

12. A class was separated if its students studied different languages, say German (the prevalent language before 1945) and French or English. I have not found Gabi who was apparently a French (?) high official in the educational field.

13. Ernst Kretschmer (1888 – 1964), a German psychiatrist and psychologist.

14. The journal *Amerika* had been published monthly in the USA in 1946 – 1948 and 1956 – 1994. Its circulation in the Soviet Union, perhaps except 1946, was restricted to the utmost and sometimes banned. Stalin, who had been suspecting his own shadow, could not have decided otherwise. Something was possibly done to slower correspondence with the *capitalist surrounding*, see just above.

Also a bit above Kolmogorov discussed the national American scholarship programme MERIT. It was initiated in 1955 and is managed by a privately founded corporation. I found a description of its work but still have no answer to Kolmogorov's question.

15. The decision of the highest Party and government organs of 22 Dec. 1977. It required a strengthening of the ideological direction but neither mathematics, natural sciences, or foreign languages were mentioned there and had to suffer.

The Soviet Union had existed only a bit longer than E. thought so that Kolmogorov was wrong. The perestroika was doomed to fail since the entire mighty nomenklatura was rotten and most if not every constituent republic demanded real independence from Russia proper. Brezhnnev's claim that there appeared a new entity, a Soviet man, was proved damnably wrong.

16. Psychologist Leonid Vladimirovich Zenkov. He died 27 Nov. 1977, but Kolmogorov obviously had not yet known it.

References

Fisher R. A. (1948), What sort of man is Lysenko. *Coll. Works*, vol. 5. Adelaide, 1974, pp. 61 – 64.

Gnedenko B. V. & Sheynin O. (1978, in Russian), Theory of probability. In *Math. in the 19th Century*, [vol. 1]. Editors, A. N. Kolmogorov, A. P. Youshkevich. Basel, 1992, 2001, pp. 211 – 288.

Kolman E. (1982), *We Should Not Have Lived That Way*. New York. In Russian.

Novikov S. P. (1997), Mathematics and history. *Priroda*, No. 2, pp. 70 – 74. **S, G,** 78.

Pontriagin L. S. (1980), On mathematics and the quality of its teaching. *Kommunist,* No. 14, pp. 99 – 112. In Russian.

Shyraev A. N. (1989), A. N. Kolmogorov: in memorial. *Teoria veroiatnostei i ee primenenia*, vol. 34, pp. 5 – 118. That journal is translated as *Russian Math. Surveys*.

Volotskoy M. V. (1934), *Khronika roda Dostoevskogo* (Chronicle of the Dostoevsky Family).

IV

Bertrand's work on probability

Arch. hist. ex. sci., vol. 48, 1994, pp. 155 – 199

1. Introduction

1.1. General information. Joseph Louis Francois Bertrand (1822 - 1900) contributed to several branches of mathematics. In 1855 he translated into French Gauss's writings on the theory of errors and method of least squares (MLSq)[1]. A few of his notes on the theory of probability and combination of observations appeared in 1875 - 1884. Then, during 1887-1888, he published 25 more notes on the same subject. Bertrand's *Calcul des probabilités* appeared in 1888 (§ 1.2) and his last note on probability was dated 1892.

In 1856 Bertrand was appointed professor at the *Ecole Polytechnique* and, in 1862 he became professor at the Collège de France as well. A member of the Paris Academy of Sciences, he was its permanent secretary from 1874 until his death. Lévy (1900, p. 72) indicates that Bertrand taught probability *à diverses reprises* both at tbe Collège *dans son enseignement moins éleve* and at the Ecole. Darboux (1902, p. XLII) testifies that in 1878 Bertrand abandoned his teaching at the Collège, but in 1886, had to resume his activities there[2]. This fact likely explains his sudden interest in probability as manifested by his publications of 1887-1888.

For the first time ever, I describe in full Bertrand's work on probability and error theory. Beginning with § 2, my account follows his treatise [2] and 1 usually refer to it just by page numbers of its first edition. 1 also turn attention to Bertrand's notes[3]. Their large number, and the appearance of most of them during just two years, compelled me to refer to them without indicating their date. Consequently, my system of mentioning Bertrand is unusual.

Here are my general remarks. Except for Items 2 and 3 they concern all of Bertrand's writings.

1. Bertrand mentioned some of his predecessors (De Moivre, Laplace, Bienaymé), but did not refer to other scholars, notably to Chebyshev.

2. His treatise contains mistakes and misprints[4]. The conditions of many problems are stated carelessly and drawings are completely lacking. Verbal explanation s, sometimes given instead of formulas, are irritating.

3. The treatise is badly organized.

4. Bertrand uses the term *valeur probable* on a par with *espérance mathématique*.

5. His literary style is extremely attractive.

I left out some of the topics discussed by Bertrand, namely the description of classical least squares and the bivariate normal law which he introduced largely for the sake of discussing target shooting. 1 have somewhat changed his inconsistent notation and introduced Gauss symbol $[aa]$ for $a_1^2 + a_2^2 + \ldots + a_n^2$.

1.2. Bertrand 's Treatise [2]. According to the *American National Union Catalog Pre-1956 Imprints* (vol. 50, p. 591) the treatise was published in 1888 and again in 1889. The *Catalog* also mentions the second edition of 1907, *conforme à la* 1. I saw no other reference to the edition (or printing) of 1888 and, moreover, there seems to be some ambiguity with its dating. Thus, Rouché (1888b) published a review of the treatise (a long one, intended for non-specialists) stating that the book [had] appeared in 1889. On pp. 561 and 577 of his review, below its main text, the publisher inserted the relevant date of the periodical, Décembre 1888.

Again, the *C. r. Acad. sci. Paris*, t. 107 (1888), twice mentioned the edition of 1889. First, on p. 671, it said that (on 29 October) Bertrand *présente à l'Académie l'Ouvrage qu'il vient de publier sous le titre de Calcul des probabilités*. Then, on p. 705, the *Bulletin bibliographique* that listed the titles of the *Ouvrages recus* on the same date (29 October) began with Bertrand's *Calcul* (Paris, Gauthier-Villars et Fils, 1889) [!].

I conclude that the treatise was indeed first published in 1888, but that at least some of its copies were dated wrongly, either on purpose or otherwise. In the *Préface* to his treatise, on p. V, Bertrand states that his book is a *résumé de Lecons faites au Collège de France* and that he attempted to discuss *les résultats les plus utiles et les plus célèbres* basing them *sur les démonstrations les plus simples*, cf. § 18. The *Préface* (pp. V-VI) also contains a passage which reminds me Laplace's *Essai*:

La plupart des réflexions suggerées par l'étude approfondie des questions souvent controversées ont été proposées dans un Travail dégagé de toute intervention des signes algébriques, imprimé déjà depuis plusieurs années.

1.3. Some other of Bertrand's contributions. Here I describe other contributions as far as they touch on the theory of probability. I also indicate cases where he failed to mention it.

1. In the *Thermodynamique* [1] Bertrand (p. XI) centred his *explications principales autour de trois noms …, Sadi Carnot, Mayer et Clausius*. He did not refer to Boltzmann or mention probability. Poincaré, in 1892, maintained a similar attitude (Sheynin 1991, pp. 141-142).

2. In *D'Alembert* [3] Bertrand (pp. 49-55) did not really express his subject's studies in probability. However (pp. 49-50), he correctly stated that D'Alembert had refused to consider the calculus of probability as *une branche légitime des mathématiques*. Bertrand (p. 51) also suggested that D'Alembert was *toujours prêt à déclarer impénetrable ce qui lui semble obscur* and that (p. 55) the vagueness of his contributions stemmed from want of pedagogical experience.

3. Bertrand [4] devoted a few pages to Pascal's work in probability. Pascal offered *un principe* for solving the problem of points (p. 316); *Les problèmes sur le hasard … avaient à ses yeux le premier rang* (p. 317); and, in general (p. 315),

Sans Pascal, la science n'aurait pas eu le livre de Jacques Bernoulli et l'admirable théorème qui le termine.

The first two statements were left unsubstantiated whereas the last one was somewhat too strong. Finally, Bertrand (p. 313) denied Condorcet's low opinion (1847, p. 608) of Pascal's merits in probability.

4. Bertrand published a large number of long reviews in the *Journal des savants* (Anonymous, 1902). One of these, in 1896, was devoted to Huygens's *Oeuvres complètes,* tt. 2-6 (not to tt. 2-4, as stated in the source just mentioned). Some discussion of Huygens's work in probability was quite in order. However, Bertrand did not even mention this topic.

In 1887 Bertrand [37] reviewed Laplace's *Théor. anal. prob.* He considered Bayesian approach, moral expectation, application of probability to jurisprudence, and the theory of errors and largely repeated himself in his treatise. Lacking was a justified assessment of Laplace's classic.

In another review Bertrand [38] offered some comments pertaining to the theory of errors. Following Gauss (§ 14, Item 3) he (p. 211) recommended combining measurements made by several geodesists to obtain a large total number of measurements and ensure a plausible estimation of the general precision.

2. The probability of random events. Bertrand (p. 2) introduced the classical definition of probability and went on to discuss pertinent problems. He (p. 4) indicated that to choose

Au hasard, entre un nombre infini de cas possibles, n'est pas une indication suffisante[5].

Thus, the probability of selecting at random a real number x greater than 50 from those between 0 and 100 was not necessarily 1/2. At random might mean that x^2 rather than x itself should have possessed a uniform distribution so that $P(x^2 < 2500) = 1/4$. I describe now some other problems. In the first two of them Bertrand continues to discuss uniform randomness.

1. *On trace au hasard une corde dans un cercle* [of radius r]. *Quelle est la probabilité* [p] *pour qu'elle soit plus petite que le côté du triangle equilateral inscrit* (p. 4)?

Bertrand considered three cases: **a)** One end point of the chord was fixed and all of its directions were equally probable (distributed uniformly); $p = 1/3$. **b)** The chord's direction was fixed, and all of its distances from the centre of the circle were equally probable; $p = 1/2$. **c)** Nothing was fixed, but the location of the middle of the chord in any point of the circle was equally probable; $p = 1/4$[6].

2. *On fixe au hasard deux points sur la surface d'une sphére; quelle est la probabilité pour que leur distance soit inférieure à 10*' (p. 6)?

Without mentioning anyone, Bertrand gave differing solutions due to Laplace and Cournot (1843, § 148). He returned to this problem on pp. 170-171, this time mentioning Michell, who, in 1767, had attempted to calculate the probability that two stars out of a certain number of them were close to each other.

Suppose that the distance between two stars is α or less. Then, if the first star is at point A, the second one is located at any point of the segment whose vertex is A. The ratio of the segment's surface area to that of the hemisphere is $\alpha^2/2 = 1/236,362$. Forgetting his previous solutions Bertrand now calculated the probability P that only two stars out of 230 (of those of the first three magnitudes) were at a distance α or less from each other. He noted that

$$P = 1 - (1 - 2/\alpha^2)^n, \; n = 230\text{x}229/2 = 26{,}335$$

and $P = 0.103403$, actually, $P = 0.105441$, or, I should say, 0.105.

Confirming an earlier statement (p. 7), Bertrand (p. 170) maintained that

L'ingénieux argument de Mitchell [!] *ne peut pas cependant fournir d'évaluation numerique*[7].

He reasonably remarked that other peculiarities of the sidereal system rather than the small distance between stars could have been evaluated (instead).

I (1984, § 5) have described the history of Michell's problem[8]. Here, I only repeat that both Newcomb and Fisher studied it by means of the Poisson distribution.

3. An urn contains m balls. One of them is extracted and returned back. Determine the probability P that after n such drawings at least k specified balls will be extracted at least once (p. 14).
The answer is

$$P = m^{-n} - \{m^n - k(m-1)^n + [k(k-1)/2](m-2)^n] - \ldots \pm (m-k)^n\}$$

De Moivre (1725, p. 44; 1756, p. 315) used a similar formula (of inclusion and exclusion) for the probability of compatible intersecting events to calculate *the Value of the Longest of any Number of Lives*.

4. A ballot (p. 18). Candidates A and B scored m and n votes respectively, $m > n$. If all the possible voting records were equally probable, what is the probability P that, during the balloting, A was always ahead of B?
Following André (1887), Bertrand gave a simple proof that

$$P = (m - n)/(m + n). \qquad (1)$$

He published formula (1) even before André. Suppose that $s = m + n$ and that $P_{m,s}$ combinations of the votes are favourable. Then [10]

$$P_{m+1,s+1} = P_{m,s} + P_{m+1,s}.$$

Bertrand did not explain how he solved this equation in finite partial differences[9].

Barbier (1887) generalized formula (1). He maintained that for natural values of k the probability that during the balloting A always scored more than k times the number of votes registered for B was

$$P = (m - kn)/(m + n), \; k < m/n$$

Barbier did not supply the proof. According to Takacs (1967, p. 2), this was not published until 1924[10].

In all, the first two chapters of Bertrand's treatise containcd 40 problems. He did not solve all of them but always provided the answer. As he himself (p. 44) noted, a number of problems were due to De Moivre.

3. Espérance mathématique

This is the title of the third chapter of Bertrand's treatise. Poisson (1837, pp. 140-141) was the first to introduce a special, although temporary term, *une chose quelconque*, for a variable taking a number of values with corresponding probabilities. Bertrand (p. 61), without mentioning Poisson or his term, considered *grandeurs* (discrete random variables) and formulated two theorems concerning their expectations. Thus,

La valeur probable d'un produit, <u>quand les facteurs sont indépendants</u>, est le produit des valeurs probables des facteurs.

He (e. g., on p. 80) used the term *la valeur probable* on a par with *l'espérance mathématique*[11]. Although he translated Gauss's contributions on the theory of errors in French (1855), he did not follow the Master who had preferred the term *valor medius*, see Sheynin (1991, pp. 139-140).

Bertrand offered pertinent problems.

1. An um contains a large number m of balls marked 1, 2, …, m. One of them is extracted and returned back and n such drawings are registered. How large is Pierre's expected gain if he receives a franc for each turning point in the number sequence thus obtained (p. 53)?

Following an old tradition that goes back to Montmort and De Moivre, Bertrand introduced an unnecessary character[12]. A much more general problem is due to Bienaymé who discussed it in 1874 and again in 1875. Heyde & Seneta (1977, § 5.11) described his work and later developments including Bertrand's note [7], or rather two separate notes the second of which was a slightly improved version of the first one. Bertrand noticed that a number chosen out of three [natural] numbers *absolument inconnus* will be maximal with probability 1/3. Consequently, he derived at once that Pierre's expectation was $2/3n$, and he repeated this reasoning in his treatise. Bertrand reasonably omitted from his book a previous Statement [7] that, if ten consecutive terms of a random number sequence were not turning points, the probability for the next one to be such a point was 10/11.

Heyde & Seneta (pp. 125-126) credit Bertrand with using the method of indicator variables (variables, taking values 0 and 1). He did not however introduce them directly. Furthermore, Chebyshev (1867, p. 183) preceded Bertrand.

2. The Buffon needle (p. 54). A needle falls on a set of parallel straight lines. The distance between adjacent lines is a, the length of the needle is s, $s < a$. Pierre (again Pierre!) receives a franc if the needle intersects a line. Determine his expected gain.

This is a classical problem. In particular, Laplace (1812, Chapter 5) noted, while solving it, that the Buffon needle provided a means for empirically determining π. Cf. Bertrand's recommendation about the ratio $(E\xi - E\xi)^2 : (E|\xi - E\xi|)^2$in note 19.

Bertrand solved this problem and stated that the expectation depended on the length, but not the shape of the needle whereas the corresponding probability depended on both these parameters. This, he added, also mentioning the previous problem (Problem 1),

Montre la différence entre le calcul de la probabilité et celui de l'espérance mathématique.

Strange indeed!

Bertrand referred to Barbier, obviously, to Barbier (1860). Barbier gave much thought to generalizing this problem and, on p. 275, indicated that he preferred to derive expected values in accord with Bertrand's advice. Many subsequent authors took up the Buffon problem. One of them was Poincaré (1912, Chapter 8).

3. The Petersburg Problem (p. 62). Gambler A throws a coin. If heads appears at once, he receives a franc from B. If however heads shows up only at the k-th throw ($k = 2, 3, \ldots$), A gets 2^{k-1} francs. Determine A's expected gain.

This problem or game was devised by Nicolaus Bernoulli and became generally known after Daniel Bernoulli, in 1738, solved it in a memoir published in Petersburg by introducing and using the notion of moral expectation.

A's expected gain is infinite. Nevertheless (p. 63), *Qui voudrait … risquer 100fr à un tel jeu*? Bertrand denied the use of moral expectation. However, it is now applied in economics. Thus, von Neumann & Morgenstern, in 1953, gave an axiomatic foundation to the neo-Bernoullian theory of subjective value (Jorland 1987, p. 179), also see Shafer (1988). Dutka (1988) offered a good survey of the history of the Petersburg Problem, but did not mention the contributions of Freudenthal (1951) or Aaronson (1978). I note only that the problem (the game) has become a legitimate object of stochastic study. Thus, Freudenthal considered a series of such games in each of which the gamblers took turn by lot.

Bertrand first discussed the Petersburg Problem in a previous note [13] without however making any progress. Nevertheless, he formulated there a proposition about the duration of a fair play between two gamblers having an equal number of coupons. The probability that the play will last $\underline{n}$ games, as he indicated, was proportional to $n^{-3/2}$. This fact does not seem to be well known, but it can be proved by considering the limiting case ($n \to \infty$) of formula from Item 5 of § 5.

4. The binomial distribution

4.1. Theory. Bertrand (pp. 69-80) proved the De Moivre - Laplace theorem mistakenly calling it after Jacob Bernoulli[13]. In the process, he expended four pages (72-76) for proving the Stirling theorem. He correctly indicated that it *était connu de Moivre* whereas Stirling had determined the value of its constant[14]. As an example, Bertrand (p. 76) compared the values of 20! with its approximation according to the Stirling theorem. He retained 14 or 15 digits in his calculations although what he actually needed was only the ratio of the two numbers, 1.00417.

Bertrand did not believe in the Poisson law of large numbers. He

(p. XXXII) maintained that this *découverte … se distingue bien peu des lois connues du hazard* and also stated that Poisson himself was *à peu près seul* who had attached to it *une grande importance*[15]. Bertrand was mistaken on both counts. In particular, I (1978, § 4.4) have described how the Poisson law gradually won recognition. In a passage there quoted, Chebyshev (in 1846) called it *Cette proposition fondamentale*. Again, while denouncing Quetelet's *homme moyen,* Bertrand (pp. XLII-XLIII) did not notice that this notion would have been much more acceptable had Quetelet based it on the Poisson law[16] rather than on Jacob Bernoulli's theorem.

Bertrand (pp. 80-82) used a generating function for determining (in later notation) $E\xi^2$ and $E|\xi|$ or, rather, $E(\xi - E\xi)^2$ and $E|\xi - E\xi|$ for the number of occurrences of an event in n Bernoulli trials[17]. In the first case he considered

$$(p + q)^n = p^n + A_1 p^{n-1} q + \ldots + A_{qn} p^{pn} q^{qn} + \ldots + q^n. \qquad (1)$$

[…]. He obtained

$$(E\xi - E\xi)^2 = npq. \qquad (2)$$

He obviously assumed that $E\xi = qn$[18].

According to modern definition, the generating function for the binomial distribution is $(q + ps)^n$, or $(p + qs)^n$. In an earlier note Bertrand [8] in a similar way determined $E(\xi/\underline{n - p})^2 = pq/n$ and stated that, consequently, as $n \to \infty$, $P(|\xi/n - p| > \varepsilon) \to 0$. What was lacking, here and elsewhere (e. g., when discussing target shooting on pp. 244-245 of his treatise), was the Bienaymé-Chebyshev inequality. Cf. note 13.

Formula (2) can be obtained in a quite elementary way by replacing successes and failures in the Bernoulli trials by indicator variables and directly calculating the expectation sought.

Bertrand's second case was more interesting. He noticed that […]

$$E\,|\,\xi - E\xi\,| = \sqrt{2npq/\pi}$$

and that both $E(\xi - E\xi)^2$ and $E|\xi - E\xi|$ coincided with their respective values calculated directly by means of the appropriate normal distribution[19].

4.2. Problems. Bertrand offered a few pertinent problems.

1 (p. 89). Determine the probability that in n = 20,000 Bernoulli trials with p = 0.45 the number of successes will be more than $n/2$.

In an earlier note Bertrand [18] solved the same problem in a roundabout way by first calculating, after Gauss (1816, § 5), the even moments of the normal distribution, $E(\xi/a)^{2s}$ and determined their minimal value.

2 (p. 94). An urn contains λp white and λq black balls. Determine the probability that after n drawings without replacement $(np - k)$ white balls will be extracted.

In an earlier note Bertrand [17] gave only the final result for $\lambda, n \to \infty$

$$P = \frac{1}{2\sqrt{\pi npq}}\sqrt{\frac{\lambda}{\lambda - .n}}\exp[-\frac{k^2\lambda}{2pqn(\lambda - n)}]$$

and noticed its similarity with the De Moivre-Laplace limit theorem. He (p. 1202) appropriately indicated that the variable probability of drawing a white ball was

En quelque sorte un régulateur de la proportion normale prévue par le théorème de Bernoulli.

In his treatise, Bertrand provided a proof by using the hypergeometric distribution. De Moivre (1712, pp. 247-248; 1756, pp. 86-89) was the first to solve a problem involving drawings without replacement by introducing this distribution[20].

3 (pp. 106-107). How many fair games between two gamblers will be necessary for one of them to lose more than 100,000 stakes with probability 0.999?

According to the De Moivre – Laplace theorem it follows that […] approximately $n = 0.62 \times 10^{14}$. […]

4.3. Le hasard, à tout jeu, corrige ses caprices. In § 4.2 (Problem 2) I quoted Bertrand's pronouncement on the regulating influence of chance. Both Daniel Bernoulli and Laplace left much more interesting statements on this subject (Sheynin 1976, pp. 150-151)[21]. After Bertrand Poincaré 1896, p. 150ff) explained important regularities in nature, see also Sheynin (1991, § 8.1)), by the action of randomness. For his part, Bertrand had uttered a few pertinent remarks and I chose one of them [2, p. XX] as the title of this subsection. Another such statement is on p. L:

Le hasard est sans vertu; impuissant dans les grandes affaires, il ne trouble que les petites[22]*. Mais, pour conduire les faits de nature à une fin assurée et précise*[23]*, il est, au milieu des agitations et des variétés infinies, le meilleur et le plus simple des mécanismes.*

At the same time he justly denied that a run of unfavourable (say) events will be compensated in the near future. The roulette, he (p. XXII) noted, *n'a ni conscience ni mémoire*.

5. The gambler's ruin

The problem of the duration of play, or of the gambler's ruin, has important applications and is a venerable chapter of the theory of probability. Many commentators, Thatcher (1957), Takacz (1969), Kohli (1975), Feller (n. d., Chapter 14), Hald (1990, Chapters 20 and 23) discussed the history of this subject, but did not mention Bertrand. Here are his problems.

1 (pp. 111-113). An expected infinite duration of play. Gamblers A and B have $2m$ francs each. The stake is 1 franc and their game is fair. Determine the expected duration of play, $\varphi(2m)$.

Suppose that at first the gamblers risk only m francs each. Then the loser should play for the second half of his money and is either ruined altogether or recovers his loss. Consequently,

$$\varphi(2m) = 2\varphi(m) + 1/2\varphi(2m),\ \varphi(2m) = 4\varphi(m).$$

If now B is infinitely rich and A has only m francs, the expected duration of play will be

$\varphi(m) + 1/2\varphi(2m) + 1/4\varphi(4m) + 1/8\varphi(8m) + \ldots =\cdot$

$\varphi(m) + 2\varphi(m) + 4\varphi(m) + 8\varphi(m) +\ldots = \infty.$

On pp. 132-133 Bertrand, without mentioning these considerations, generalized his problem by tacitly assuming that the play was unfair. He issued from De Moivre's formula (also see Item 3 below)

$P_m = (\lambda^m - 1):(\lambda^{m+n} - 1).$

Here, A had m counters and B had n of them, and $\lambda = q/p$ and p and q are the probabilities of their winning each game respectively[24].

Assuming now that each gambler had m francs, Bertrand got […]

$\varphi(2m) + 2\varphi(m) + 2\lambda^m\varphi(2m):(1 + \lambda^m)^2.$

Here Bertrand dropped his investigation *qui n'intéresse pas le Calcul des probabilités*. I suspect however that it was the difficulty of subsequent work that thwarted his attempts.

2 (pp. 116-117). Ruin in a fair game with unequal probabilities of winning and losing. Gamblers A and B have m and n counters, their stakes are a and b, and they win each game with probabilities p and q, respectively. Determine the probability of ruining (P)[25].

[I only formulate the next problems.][26 – 29]

3 (pp. 117-119).The same problem, but the game is not fair.

4 (pp. 119-122). Determine the probability that gambler A having a finite number of counters will be ruined in an indefinitely long play *contre tout adversaire qui se présente.*

Bertrand noted that in one of the possible cases the outcome remained indefinite.

5 (pp. 122-123). Ruin after a specified number of games. Gambler A has m counters, the probability of his winning any given game is p and he plays with an infinitely rich partner. Determine the probability that he will be ruined in exactly n games, $n > m$. Supposing that n and m were of the same parity, Bertrand wrote down the probability that A loses $(n + m)/2$ games and wins $(n - m)/2$ times

$$P(n) = \frac{m}{n} C_n^{(n-m)/2} p^{(n-m)/2} q^{(n+m/2}.$$

6 (pp. 126-127). Ruin in a fair game with unequal probabilities of winning and losing.

7 (pp. 128-131).The same problem, but the game is not fair anymore.

6. The probability of causes

This is the title of one of the chapters of Bertrand's treatise. Of course, the main problem here was to determine the posterior probability of a hypothesis in a Bayesian setting. At the same time he denied the Bayesian approach (end of section)! Laplace (1814, p. CXLVIII) had mentioned the English scholar, even if late in the day, whereas Bertrand did not.

Here are his main problems.
1 (pp. 148-149). An urn contains white and black balls. Draws with replacement produced m balls of the first kind and n of the second one. Determine the most probable composition of the urn.
Denote the probability of extracting a white ball by x. Then the probability of the sample drawn is proportional to

$$y = x^m(1 - x)^n$$

and the most probable composition corresponds to

$$\hat{x} = m/(m + n). \quad (1)$$

Suppose that x is a random variable taking values $(0, 1, \dots, m + n)$ and having density

$$\psi(z) = Cz^m(1 - x)^n. \quad (2)$$

Then, to repeat my earlier remark about Laplace's pertinent work (Sheynin 1976, p. 155),

$$Ex = (m + 1):(m + n + 2) \quad (3)$$

whereas $\hat{x}$ is only an asymptotically unbiased estimator of Ex[30] (and at the same time the statistical probability of the appearance of a white ball).
2 (pp. 149-151). Bertrand continued: *Chaque hypothèse sur la valeur de x a une probabilité. Nous devons en chercher la loi.*
Assuming that x had a uniform prior distribution, he determined the probability of

$$x = p = m/(m + n) + \varepsilon,\ 1 - x = n/(m + n) - \varepsilon,$$

$$P = C\exp[-\varepsilon^2(m + n)/2pq].$$

Bertrand then indicated that *le seule différence* between this formula and the local De Moivre - Laplace theorem was that this time not the prior probability, but number m was known exactly. This is not the whole truth. In its own standard notation, the integral theorem can be written down as

$$P(-\alpha \le \frac{\mu/n - p}{\sqrt{pq/n}} \le \alpha) = \sqrt{\frac{2}{\pi}}\int_0^\alpha \exp(-z^2/2)dz$$

(μ is the number of the occurrences of the event studied in n trials) but in the Bayesian setting (known μ and unknown p)[31]

$$P(-\alpha \le \frac{\mu/n - p}{\sqrt{\mu(n-\mu)/n^3}} \le \alpha) = \text{the same} \quad (*)$$

In these cases, respectively,

$\sigma_\mu = \sqrt{pqn}$ and $\sigma_p = \sqrt{\mu(n-\mu)/n^3}$

and both formulas describe the behaviour of a normed and centred variable, $|\xi - E\xi|:\sigma_\xi$.

2a (p. 180). An urn contains white and black balls. After n drawings with replacement m white balls have appeared. Determine the probability P that the urn has $(m/n + z)$ white balls.

[In two cases Bertrand provided different answers, both times without justification.]

3 (pp. 152-153). N white and black balls are placed in an urn by lot so that balls of each colour have equal probabilities of being chosen. Determine the most probable composition of the urn if extractions with replacement produced m white balls and n black ones.

Suppose that the urn contains $(N/2 - z)$ white balls and denote $z/n = y$. The probability of this hypothesis is, for a large N, proportional to $\exp(-2Ny^2)$ and the probability of the sample is $(1/2 - y)^m(1/2 + y)^n$.

La probabilité de la cause, c'est-à-dire de la valeur y, is then proportional to the product of these two probabilities and takes its maximal value at point

$$y = (n - m)/[2(N + m + n)].$$

This problem is elementary, but it discusses, in a natural way, unequal prior probabilities[32].

4 (pp. 158-160). A coin is tossed 4040 times producing 2048 heads and 1992 tails (Buffon). Determine the probability that the coin is irregular.

What Bertrand wanted to know was the probability P that the probability of heads, p, exceeded 1/2. As he (p. 157) noted,

On a cherché quelquefois, non la probabilité de chaque hypothèse [of each value of p, since this is too difficult], *mais la probabilité* [of $p > 1/2$].

Let $p = 1/2 + z$, $0 < z \le 1/2$. Then

$$y = (1/2 + z)^{2048} (1/2 - z)^{1992} \approx \exp(-8080z^2 + 112z)$$

so that P is proportional to

$$\int_0^{1/2} \exp(-8080z^2 + 112z)dz.$$

The probability of $z \ge 0$ is proportional to the same integral with term $-112z$ instead, the ratio of the two integrals is 4.263 and $P = 0.81$.

Bertrand tacitly assumed that the prior probability was distributed uniformly. He added that Poisson (1837, p. 229) [by means of his own form of the De Moivre - Laplace integral theorem] had obtained the same result. I myself have also gotten the same answer by using the Bayes theorem (*). In addition, this proposition provides $\sigma_p = 0.0079$

which of course corresponds to $Ep = 2048/4040 = 0.507$ and is consistent with the value of P (above).

In an earlier note Bertrand [24] reasoned on Buffon's experiment, unjustly called Poisson's calculations *longs et difficiles*, but did not yet estimate the probability P.

5 (pp. 166-169). In 10,000 games of roulette red appeared 5,300 rather than 5,000 times. Determine the probability that the roulette (the mechanical device itself) was faulty.

Or, compare the hypotheses $P_1(0.529 \le p \le 0.531)$and· $P_2(0.499 \le p \le 0.501)$. Bertrand noted that $P_1 + P_2 < 1$, but that their ratio nevertheless provided the answer sought. Actually, however, he calculated the ratio

$$p_1^{5300}(1-p_1)^{4700} \div p_2^{5300}(1-p_2)^{4700}$$

for $p_1 = 0.500$ and $p_2 = 0.530$. He indicated that this ratio will be equal to P_1/P_2 only if the prior probabilities of those two hypotheses were the same. Since these were unknown, he abandoned his problem.

6 (pp. 173-174). Determine the probability that the sun will arise tomorrow. Suppose that an event occurred m times in s trials. The density law of the random variable that describes this result is given by expression (2) with $n = s - m$ and the probability of the next occurrence of the event in trial $(s + 1)$, or, I would say, the expectation of the event, Ex, is equal to fraction (3) with the same n.

Accordingly, Bertrand calculated the probability of a sunrise, given that $m = s = 2{,}191{,}500$ and $n = 0$, maintaining, however, that the problem was meaningless. Later authors (Polya 1954, p. 135; Feller (n. d., § 5.2) shared this opinion. For my part, I (1976, p. 162) have remarked that this problem is due to Hume, see also Zabell (1989). Price (Bayes 1764, pp. 150-151) had offered it along with a few similar ones; then Bufffon and Laplace turned their attention to it.

Polya's argument was hardly convincing. I repeat that Price posed his problem only to explain how to proceed when prior knowledge was completely lacking. It would be unwise to suggest that he (or Buffon, or Laplace) did not understand the actual situation.

7 (pp. 276-278). The next problem is related to the Bayesian approach. Heads appeared $m = 500{,}391$ times in $n = 10^6$ tosses of a coin. Is it correct to conclude that $p = 0.500391$?

Bertrand began by maintaining that three tosses were not enough for estimating the probability of heads. He added that even for large values of n the precision of the statistical probability was not *mieux justifiée*, and that not a single digit of the calculated *p mérite confiance*.

Consider two hypotheses, namely: the probability sought is $p_1 = 0.500391$; or, it is $p_2 = 1 - p_1$. BertrandD assumed that the prior probabilities of p_1 and p_2 were equal and reasonably took $p \approx q \approx 1/2$ when calculating, according to the local De Moivre-Laplace theorem,

$$\frac{P_1}{P_2} = \frac{P(p = p_1)}{P(p = p_2)} = \exp(-\frac{\Delta p^2}{2npq}) \approx 3.40.$$

What should have the reader inferred? What should he have thought about the strangest statement about the statistical probability of heads? Of course, even the most probable value of a random variable with a continuous distribution had an infinitely low probability: but I still do not understand why Bertrand did not believe in statistical probability, which should be considered almost on a par with its theoretical counterpart (Sheynin 2017, pp. 68 – 69).

An easier way of calculation was to determine directly $P_1/P_2 = (p_1/p_2)^{m-n}$.

Bertrand did not really believe in the Bayesian approach. He (pp. 160-161) considered just one throw of a coin. Proceeding as in Problem 4, he got the probabilities of *heads* and *tails,* 3/4 and 1/4.

Une telle conséquence suffirait pour condamner le principe [of Bayesian inference], he declared thus strengthening his earlier negative remark [24, p. 637]. A modern author (Roberts 1978, p. 12) has acknowledged the difficulty of considering one trial. Again, Cournot (1843, § 95), obviously following Laplace, stated that the Bayesian inference becomes ever more objective as the number of trials increases[33].

7. Order statistics

Bertrand proved several theorems on the mean values of order statistics. Following him I start from the normal density

$$\varphi(x) = \frac{k}{\sqrt{\pi}} \exp(-k^2 x^2). \qquad (1)$$

For a random variable Δ having this $\varphi(x)$

$$E\Delta^2 = 1/2k^2,\ E|\Delta| = 1/k\sqrt{\pi}.$$

Here are Bertrand's theorems (problems):

1 (pp. 198-199). Observations are randomly divided in pairs. Determine the mean (expected) value of the modulo larger error [ξ, not mentioned by Bertrand] in a pair. Answer:

$$E\xi = \frac{\sqrt{2}}{k\sqrt{\pi}}.$$

2 (pp. 199-200). Estimate the modulo smaller error m. Answer:

$$Em = \frac{2-\sqrt{2}}{k\sqrt{\pi}}.$$

The reader is left to guess whether the sum $E\xi + Em$ agreed with the unknown given sum.

3 (pp. 200-201). Observations are divided as above. Determine the mean value of the square of the modulo larger error in a pair.
The solution is similar to that of Problem 1:

$$E\xi^2 = \frac{1}{2k^2}(1+2\pi).$$

4 (pp. 201-202). Observations are divided in groups of three. Determine the mean (expected) value of the square of the modulo largest error in a group. Answer:

$$E\xi^2 = \frac{1}{2k^2} + \frac{2\sqrt{3}}{\pi}.$$

5 (pp. 216-217). Determine the mean (expected) value of the modulo least error in a series of *n* observations.

Bertrand was mistaken. Bortkiewicz (1922, pp. 198-201) described his work on order statistics, but did not follow up the approximations in Problem 5 and did not therefore correct Bertrand.

8. Jurisprudence

Bertrand (p. 319) denied the work of Condorcet, and continued:

Laplace a rejeté les résultats de Condorcet, Poisson n'a a pas accepté ceux de Laplace; ni l'un ni l autre n'a pu soumettre au calcul ce qui y échappe essentiellement: les chances d'erreur d'un esprit … devant des faits mal connus et des droits imparfaitement définis.

Cournot was guilty as well: he allegedly supposed that judges decided their cases independently from one another. This accusation [2, pp. 325-326] was simply wrong[34].

Bertrand did not offer any positive recommendations. He ended this chapter by repeating Mill (1886, p. 353) to the effect that probability, misapplied to jurisprudence, had become the r*eal opprobrium of mathematics*. He first quoted Mill in an earlier note [24, p. 638]. In both cases he translated *opprobrium* as *scandale*; *deshonneur* (disgrace) is more precise. Elsewhere, however, Bertrand (p. XLIII) expressed a somewhat different opinion; Mill's accusation *est injuste*, he maintained, since the three French scholars were content to give only approximate results. Does this mean that he had a change of heart?

Bertrand's chapter on jurisprudence was also published separately [35] and the two texts are practically identical.

Cournot (1843, § 121) concisely and clearly formulated the benefits of applying statistics to jurisprudence. His insight is all the more interesting since he opposed such scholars as Poinsot and even Cauchy (Sheynin 1973, p. 296, note)[35] and since pertinent stochastic studies have recently been resumed (Heyde & Seneta 1977, p. 34). Also see Zabell's historical review (1988).

9. The laws of statistics

Bertrand (pp. 307-311) remarked that population statistics had to do with variable probabilities so that the binomial distribution could not always be applied to it. Another complication was the dependence between events, but in this connection he mentioned only meteorology[36]. He (pp. 310-314) illustrated his ideas by urn schemes.

1. There are n urns with white and black balls in each of them. The probabilities of extracting a white ball are $p_1, p_2, ..., p_n$, respectively. Out of each urn drawings with replacement are made.

2. There is only one urn with the probability of the same event $p = (p_1 + p_2 + ... + p_n)/n$ and sn drawings with replacement are made out of it. The number of white balls extracted in those cases will be

$$\underline{sp}_1 + z_1;\ \underline{sp}_2 + z_2;\ \ldots;\ \underline{sp}_n + z_n$$
$$spn + z.$$

The expected numbers of these balls will be the same, so that

$$E(z_1 + z_2 + \ldots + z_n) = Ez.$$

However, the two cases are not identical since

$$Ez^2 > E(\sum_{i=1}^{n} z_i^2).$$

Consequently, it is not allowed to lump together urns of different content.

Pages XXIX-XXX of Bertrand s treatise and his earlier note [33, p. 1312] provide a clue to his attitude: he mentioned Dormoy and the *coefficient de divergence*[37]. Dormoy and Lexis originated the· so-called *dispersion theory*, which studied the stability of statistical series (of the underlying probabilities) in demography by comparing the oscillations between their terms with those of a series of Bernoulli trials[38]. Later statisticians became interested in a second goal, viz., in revealing possible dependence between terms of statistical series.

For a long time Dormoy's contributions remained unnoticed, and modern statisticians connect the theory only with Lexis whose main pertinent works appeared in the 1870's. Bertrand's failure to mention Lexis is regrettable. And Bortkiewicz (1930), who compared the merits of the two men, concluded, on p. 53:

Ging Dormoy das Verständnis für Anwendungen der Wahrscheinlichkeitsrechnung auf Erfahrungstatsachen in so starkem Masse ab, dass es der historischen Gerechtigkeit nicht entspricht, ihn, so oft von Dispersionstheorie die Rede ist, in eine Reihe mit Lexis zu stellen.

Bertrand next considered the Gompertz and the Makeham laws of mortality. This is a special topic and I leave it aside[39]. On the whole this chapter hardly matched its title.

10. The normal law and the method of least squares

Bertrand considered two derivations of the normal law. At first he (pp. 29-31) described the demonstration which is usually named after Maxwell but dates back to Adrain, and justly stated that it essentially demanded the (lacking) independence of the deviations (or errors). Ellis made the same remark in 1850. Bertrand also took up Gauss's substantiation of 1809. First, he (pp. 180 -181) was dissatisfied with the Master's *postulatum* (as he called it on p. 176) of the arithmetic mean since the mean of the observations did not correspond to that of their functions, e. g., of their logarithms. This statement was worthless at least since Gauss had considered only observations.

Second, Bertrand (p. 177) objected to the principle of maximum likelihood[40]. However, Gauss himself, while offering his mature justification of the MLSq (1823, § 6), introduced an integral measure of precision and abandoned his old line of thought. Furthermore, he explained this new attitude in his letters to Encke (in 1831) and Bessel (in 1839). His correspondence with Bessel was published in 1880.

Third, Bertrand (p. 177) maintained that the density law of observational errors should be assumed not as $\varphi(\Delta)$, as Gauss had written it, but as $\varphi(X, \Delta)$ where X was the quantity measured. His explanation was not really clear, but he evidently meant that an error of reading, which is a component of Δ, depended on the difference between X and the nearest number that could be directly read on the scale of the measuring device. Some argument!

Fourth, Bertrand (p. 180) concluded:

Le théorème [obviously, the normal distribution] *semble confirmé; il est mis en défaut: la formule … est seulement approchée, elle devrait être rigoureusement exacte*.

He referred to his wrong formula in § 6. Note also that Gauss's derivation of the normal law was valid for a finite number of observations whereas the formula just mentioned held for large values of n.

Bertrand' s previous note [23] contained the last two arguments. Here, he more correctly (although, once again quite unnecessarily) made use of another formula from § 6.

The note ended with a Statement not to be found in the treatise:

Je ne crois pas me montrer téméraire, en supposant que cette grave objection [the last one] *apercue par Gauss qui n'en a rien dit, est la cause de ses efforts plusieurs fois renouvelés pour substituer une théorie nouvelle à celle qu'il avait d'abord établie*[41].

Elsewhere Bertrand (p. XXXIV) indicated that Gauss

En proposant en 1809 une hypothèse sur la théorie des erreurs … ne prétendait nullement établir la vérité mais la chercher.

Perhaps partly true.

Bertrand [20; 2, pp. 181-183] proved that the principle of maximum likelihood, given any density function $\varphi(\Delta)$, did not lead to either the geometric mean of observational errors $x_1, x_2, \ldots, x_n$, or to $\sqrt[n]{x_1^2 + x_2^2 + \ldots + x_n^2}$. He believed that this was an additional objection to Gauss's derivation, but I fail to understand why.

No wonder that Bertrand was in favour of Gauss's later substantiation of the MLSq. He stated that Gauss had solved the problem *rigoureusement* (p. 248) and that *La théorie nouvelle semble preférable* (p. 268). At the same time it seemed strange to him (*ll doit sembler étrange*) that the density law of errors

Soit sans influence sur les conclusions d'une théorie dans laquelle elle joue un si grand rôle (p. 267).

Accordingly, he noticed that for small errors an expansion of any even law in powers of x yielded

$$a + bx^2 = a\exp[(b/a)x^2].$$

This reasoning is also contained in Bertrand's note [30].

In general, frequency laws are unknown, otherwise Bertrand's objection above is valid.

11. Precision and weight

Bertrand (p. 208) introduced the notions of precision and weight. Assuming that observational errors obeyed the normal law (7.1) and following Gauss (1823, § 7), he called k the precision of an observation and stated that k^2 was its weight. On pp. 254-257 Bertrand extended the definition of weight to any density law. Suppose that X is an unbiased linear estimator of the real value of an observed quantity whose measurements are $x_1, x_2, \ldots, x_n$:

$$X = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n,\ \lambda_1 + \lambda_2 + \ldots + \lambda_n = 1$$

and choose such coefficients λ that provide the minimal variance of X. Then, as Bertrand stated, λ_i was the weight of x_i. He remarked however that *dans le cas général* precision cannot be defined *avec la même rigueur*. Why not? According to Gauss (1823, § 7) even in this case precision is the square root of weight. True, he did not essentially use this notion, nor did Bertrand.

Bertrand (pp. 208-211) went further and offered what I call definitions (2a) and 2b):

2a) If, for observational errors δ and Δ

$$P[x \le \delta \le x + dx] = P[\alpha x \le \Delta \le \alpha(x + dx)]$$

then the first observation is x times more precise than the second.

2b) Given, two series of observations, I and II, such that m observations from I provide the same inference as n observations from II . Select an observation (A and B) from each series. Then the weights of A and B are as n**:**m[42].

Now, in each case the definitions contradict the earlier ones since Bertrand, while illustrating them, took into account only the exponential term of the normal law. For example, when comparing two series of observations having different weights, he equated the values of the respective likelihood functions disregarding their numerical coefficients.

Bertrand tried to isolate such laws of error for which his definitions 2a) and 2b) made sense, obtaining function

$$\psi(x) = C\exp(-\,ax^{2n})$$

At the same time he declared that his demonstration was *sans intérét pour le Calcul des Probabilités*. I do not understand him. He began by stating that, if the precisions of two series of observations were in the ratio of h, then $\psi_1(hx)/\psi(x) = C$, a condition which was not satisfied for normal laws with different values of k. The earlier exposition of this subject [15] was not better.

12. Estimating the precision of observations

Throughout his treatise Bertrand adopted the normal law (7.1) as the density of observational errors[43]. He paid much attention to estimating *k*, the natural measure of observational precision (cf. § 11). Indeed, he (p. 190) calculated the first four absolute moments of function (7.1) and, following Gauss (1816, § 5), equated them with the corresponding empirical sums of absolute terms. Thus, for the first two moments, if $e_1, e_2, \ldots, e_n$ are the errors of observation[44],

$$(|e_1| + |e_2| + \ldots + |e_n|)/n = S_1/n = 1/k\sqrt{\pi}, \tag{1}$$

$$(e_1^2 + e_2^2 + \ldots + e_n^2)/n = S_2/n = 1/2k^2. \tag{2}$$

He (pp. 191-194) then determined

$$E(\frac{S_1}{n} - \frac{1}{k\sqrt{\pi}})^2 = \frac{1}{2nk^2}(1 - \frac{2}{\pi}) = d(\frac{1}{k\sqrt{\pi}})^2, \tag{3}$$

$$E(\frac{S_2}{n} - \frac{1}{2k^2})^2 = \frac{1}{2nk^4} = d(\frac{1}{2k^2})^2 \tag{4}$$

so that dk^2 was smaller in the second case[45].

Bertrand (pp. 194 -195) also attempted to find such λ's which led to minimal values of

$$E(\frac{1}{k} - \lambda\frac{S_1}{n})^2 \text{ and } E(\frac{1}{k^2} - \lambda\frac{S_2}{n})^2. \tag{5a,b}$$

He obtained

$$\lambda = \frac{2\sqrt{\pi}}{2 + (\pi - 2)/n} \text{ and } \lambda = \frac{2n}{n+2} \tag{6a,b}$$

whereas formulas (3) and (4) corresponded to $\lambda = \sqrt{\pi}$ and $\lambda = 2$ respectively. He did not indicate that, in later terminology, his new formulas were biased.

Expression (2) is essentially due to Laplace (1816)[46], only his *e*'s were residuals so that Laplace's estimator was biased.

Bertrand (pp. 195-198) also offered an alternative method of estimating precision. The most probable value of *k*, as he noted, corresponded to

$$Ck^2\exp(-k^2[ee]) = \max,$$

a condition which led him to formula (2). He (p. 196) maintained, however, that the *valeur probable* [the expected value] … *doit en général être préférée*[47]. Accordingly, he calculated ·

$$Ek = C\int_0^\infty k^{n-1}\exp(-k^2[ee])dk. \tag{7}$$

Using an inaccurate approximation

$$\Gamma(x) = x^{x-1}\sqrt{2\pi x}\exp(-x),$$

he assumed a large n, obtained $Ek = \sqrt{2S_2/n}$ and stated that this expression was in accord with formula (2).

Without providing any calculations Bertrand [22] earlier published
1. The optimal values of λ for expressions (5a, b) with the first of these values differing from number (6a);
2. The most probable value of k;
3. Ek, again differing from the corresponding number derived from formula (7), and three formulas for $1/2k^2$, namely (2); its version corresponding to expressions (5b) and (6b); and another modification which he introduced without explaining it.

13. The sample variance (Gauss)

Gauss (1823. §§ 37-38) extended the use of the Laplacian formula (12.2) to any distribution and replaced it by the modern expression

$$m(=\frac{1}{k\sqrt{2}}) = \sqrt{\frac{[vv]}{\pi-\rho}}. \qquad (1)$$

Here, v's were the residuals, π was the number of observations, and ρ, the number of unknowns. In the "Selbstanzeige" to pt. 2 of his "Theoria combinationis" he stated that the change in the denominator was necessary both in actual fact and in keeping with the *Würde der Wissenschaft*[48]. His formula provided an unbiased estimator of the variance of observational errors.

Bertrand (pp. 203-206) also proved formula (1), though only for the normal distribution. Then, without mentioning this substantiation, he (pp. 298-302) borrowed a proof in the general case from Guyou (1888). Just the same, he wrote down the variance of the arithmetic mean first for the normal distribution (p. 218), then for the general case (p. 250).

Finally, Bertrand (pp. 251-252) estimated the error of Gauss's formula (1). He stated this problem as a derivation of the *valeur probable de la constante* m^2.

Denote the errors by e_i, $i = 1, 2, \ldots, n$. Then, as he showed,

$$E(m^2 - [ee]/n)^2 = (h_4 - m^4)/n. \qquad (2)$$

Here h_4 was the fourth moment of the unspecified density law of the errors. He could have added that Gauss (1823, § 40) effectively derived this formula though only for the normal distribution. However, Gauss considered $[vv]$ rather than $[ee]$, hence his result was

$$\mathrm{var}\, m^2 = 2m^4/(\pi - \rho). \qquad (3)$$

According to modern notions (Cramér; Kolmogorov et al, see below) formulas (2) and (3) should be corrected by replacing m^4 by the unknown quantity $(Ei^2)^2$. Indeed, $Em^2 = E_i^2$ but m^2 is not Ei^2.

Cramér (1946, § 27.4) obtained a formula similar to, but more precise than relation (2): it included terms proportional to $1/n^2$ and $1/n^3$. Gauss's main result (1823, § 40) was not the indirectly proved formula (3) but the derivation of bounds for varm^2. Helmert in 1904 and, in 1947, independently, three authors including Kolmogorov improved his finding. Liapunov, in a note published posthumously in 1975, showed that m^2 was a consistent estimator of the variance. Finally, Student, in 1908, determined the distribution of $m^2 = [ee]/n$ for the normal distribution. He had also stated that, again for that distribution, the arithmetic mean and m were independent, but only Fisher proved this proposition. Cf. § 17.

14. The Gauss formula criticized

Bertrand unmethodically objected to formula (13.1). Below, I collected and arranged his remarks.

1. Systematic errors make estimation impossible. Offering important historical examples, Bertrand (pp. XL-XLI and 304-305) explained that the existence of systematic errors[49] had misled astronomers into trusting their numerical results far too much. Thus, Laplace thought that he estimated the mass of Jupiter (m) to within $\Delta m/m = 1/100$ and declared that the corresponding odds were 999,308 to 1[50]. *Quelle ostentation de consciencieux savoir*, exclaims Bertrand (p. XXXIX).

Indeed, natural scientists should always suspect the presence of systematic errors. Summing up Bertrand (p. 304) stated:

Le calcul de la précision d'un système d'observations et l'évaluation qu'on en déduit pour la confiance méritée par le résultat ont compromis plus d'une fois la méthode des moindres carrés.

Gauss (1823, § 2) remarked that the MLSq was not intended to deal with systematic errors. What was perhaps lacking in practical work was a sound discussion of observations, an attempt to estimate the influence of systematic errors. Bertrand himself (p. 238), when describing target shooting, refused to decide what was more important, the general deviation of the hit-points from the centre of the target or the measure of their scatter. He might have added that both parameters should characterize the results of the shooting. Nothing better was ever proposed either in this case or for estimating the plausibility of observations in general.

2. Observations actually have unequal weight:

On suppose, a priori, toutes les mesurese également précises; il est impossible, dans le plupart des cas, de croire a cette égalité: c'est faute de connaître aucune raison de préférence qu'on accepte l'équivalence des résultats. Mais, connues ou inconnues, ces raisons, si elles existent, doivent exercer une influence sur l'erreur réelement commise [2, p. 304].

Bertrand did have a point. What he failed to mention was that astronomers took utmost care to secure equal weight both in observation and station adjustment. This fact can be confirmed by the practical work of Gauss and Bessel, by Gauss's correspondence and by other sources. Schreiber (1879, p. 141), drawing on Gauss's *Protokolle* of his field work, testified:

Er auf jeder Station so lange gemessen hat bis er meinte, dass jeder Winkel sein Recht bekommen habe. Er hat dann … [his observations] *als gleichgewichtig und von einander unabhängig in die Systemausgleichung eingeführt.*

3 (pp. 221-222, 274-277, 295, 303-306). Residuals are not sensitive enough; prior information is more important. Of course prior information is essential. Exactly for this reason astronomers, especially following Gauss and Bessel, always attempted to estimate (and minimize) the influence of all possible errors, both instrumental and external[51].

The scatter of observations is measured by their deviations from the arithmetic mean (by the residuals). Yes, small residuals can be misleading, but Bertrand offered no advice. Now, two baselines measured at the ends of a chain of triangulation, as well as the discrepancies of the triangles themselves, provide a powerful "external" check of precision. Astronomical observations furnish a similar control. Therefore, ten (say) triangles of a triangulation lead to 10-12 plausible values of v_i and formula (13.1) becomes reliable enough.

Gauss gave thought to this circumstance. In his correspondence he stressed that a small discrepancy in a triangle did not yet testify to the worthiness of the three observations. And in a letter to Gerling of 17.4.1844 (which Bertrand could not have seen) he combined observations made at several stations and estimated their average precision from a large number of measurements thus obtained. Cf. Bertrand's own later remark in § 1.3 (Item 4). Elsewhere he [27, p. 887] even declared himself in opposition to Gauss and Bessel since these scholars believed in estimating precision by the Gauss formula.

4. The Gauss formula can be improved by introducing appropriate parameters [32][52].

Five angles, $s_1, s_2, …, s_5$ are measured at a station[53] between straight lines $CA_1, …, CA_4$ and angles $A_1CA_4 > A_2CA_4 > A_3CA_4$. Before the adjustment the angles are considered unknown. Three of them should be adjusted, the other two (or three) can then be calculated.

Bertrand wrote out the discrepancies w_1 and w_2 in terms of s_i and obtained the empirical variance of the observations [54]

$$m^2 = [vv]/(5-3) = (3w_1^2 + 3w_2^2 - 2w_1w_2)/16. \qquad (1)$$

In general, as he noted, his example led to [Bertrand wrote out the general formula]. Finally, he [32] found out that the optimal solution indeed corresponded to formula (1). Note that this follows from Gauss's finding and even from Bertrand's own result.

Bertrand then arrived at a lesser value for this variance, but he disregarded the bias of this new estimate of the variance and (p. 1262) wrongly concluded that Gauss's principle, which still remained an *admirable résultat algébrique* [!], was not *la meilleure*.

Previously Bertrand [31] did not yet calculate the variance but he stated there that Gauss

A exprimé par des formules précises des appréciations simplement plausibles et proposées comme telles; on a transformé en théorèmes devenus classiques des conditions très légèment motives.

He did not explain himself sufficiently. As I understand it, he was dissatisfied with equating an empirical expression with its mean value. However, he himself adopted the same condition[55].

5 (p. 222). *L'application du Calcul des probabilités à l'étude des erreurs d'observation reposé une fiction*.

Here, as also on p. 212 and earlier [25, p. 701], Bertrand meant that (in later terminology) the observational error was not a random variable. He explained that the observations were corrupted by systematic errors (Item 1) and blunders. The main point is however that to this very day the treatment of observations is still founded on stochastic and statistical considerations. True, largely because of the circumstance stated the theory of errors remains a very special chapter of mathematical statistics. For the same reason the Gauss formula is much better suited for estimating the precision of observations after adjusting them in accord with all existing geometric conditions than for indicating when the astronomer can stop his work on a station (than for applying the sequential analysis). I especially refer to Item 2 above.

15. Rejection of outlying observations

Bertrand (p. 211) believed that outlying observations were *presque avec certitude* worse than the others. Accordingly, he (p. 213) proposed a definite test. Reject any observation, he maintained, which differed from the mean $\overline{x}$ more than by λ, $|\overline{x} - x_i| > \lambda$ and

$$p = \frac{2k}{\sqrt{\pi}} \int_0^{\lambda} \exp(-k^2 x^2) dx \qquad (1)$$

with an arbitrarily chosen p. He also noted that the number of observations retained will be m or about np where n was the number of observations. Bertrand first described his test in an earlier note [25][56].

Astronomers started putting observations to stochastic tests in the mid-19th century. Apart from the three-sigma rule the most widely known criterion was due to Chauvenet (1863, vol. 2, pp. 558 - 566) whose arguments led him to choose, in formula (l), $p = 1/2n$[57].

16. Station adjustment

Bertrand used his standard example (see § 14, Item 4) to describe station adjustment; or rather to show how to estimate the precision of observations made at a triangulation station. He rather unsuccessfully indicated that dependence between sums of observed angles should be accounted for, but his description did not picture the real world and remained unnecessary. And his conclusion about the estimation of the precision of the unknowns [of the appropriate estimators] even before observation was known long ago[58].

17. Approaching an important theorem in mathematical statistics

While attempting to prove that the empirical variance was not a proper measure of the precision of observations (§ 14), Bertrand twice

considered distributions connected with the squared sum of observational errors.
1. Given, observational errors $x_1, x_2, \ldots, x_n$ and denote

$$\Delta^2 = \Sigma(x_i - x_j)^2, i, j = 1, 2, \ldots, n, i < j.$$

The essence of Bertrand's conclusion was that we ought to allow for the difference between Δ^2 and $E\Delta^2$. However, Gauss (1823, §§ 37 – 38) remarked that such differences must necessarily be neglected. Bertrand nevertheless inferred that he refuted

Cette maxime de Laplace, *Le Calcul des probabilité n'est autre chose que le bon sens réduit en formules*[59].

But then, in his deliberations Bertrand came close to proving the mutual independence of the arithmetic mean and the variance for a normal sample[60].

2. Bertrand considered the adjustment of the angles of a triangle, but at the end of the 19th century that subject hardly interested anyone[61].

18. Conclusions

As indirectly stated by Bertrand himself (§ 1.2), he did not try to develop the theory of probability. But how did his contemporaries evaluate his work?

Poincaré (1894, p. 159) stressed Bertrand's critical attitude to the theory and indicated that it was compelled to accumulate *les hypothèses tacites … souvent arbitraires*, and that Bertrand *les avez dénoncées impitoyablement*[62].

Lévy (1900, p. 71) declared that Bertrand's *Ouvrage est et restera un chef-d'oeuvre* while Darboux (1902, pp. XLII) remarked that his *Thermodynamique* (1887), *Calcul des probabilités*, and *Lecons sur la théorie mathématique de l'éctricité* wbich appeared in 1890 *peuvent être considérés comme le couronnement de ses recherches sur les applications des mathématiques à la philosophie naturelle*[63].

At tbe same time, Darboux (pp. XLII-XLIII) politely noticed that these books were not *traites complets*; that Bertrand *a eue d'écarter les parties de la science qui sont encore en travail*; and that, at least as far as his *Calcul* was concerned, he had to pursue methodological aims:

Laplace a mis en Oeuvre les théories mathématiques les plus élevées. Bertrand les écarte résolument pour se mettre a la portée du plus grand nombre de lecteurs.

In 1916, Darboux (p. XXXIV), following Lévy, called Bertrand's treatise a *chef-d' oeuvre*. Nevertheless, as before, he added a grain of salt: Bertrand *s'était borné à critiquer et à démolir*.

Here now is my own opinion. Bertrand's *Calcul* [2] is not a masterpiece. He was mistaken in denying the Bayesian inference (end of § 6) and even statistical probability (§ 6, Item 7); he considered the estimation of the precision of observations (§ 14) without possessing astronomical know-how, and superficially criticized the Gauss formula; his discussion of the application of probability to jurisprudence (§ 8) and his recommendations about rejecting outlying observations (§ 15) were contradictory.

Add to this that Bertrand's treatise is badly organized and written carelessly (obviously, in great haste); that scholars such as Bayes, De Morgan, Chebyshev [!], Helmert, Meyer, Liagre, or Edgeworth are not mentioned and even Cournot is all but absent; and that he did not append any bibliography[64]. Again, Bertrand was apparently carried away by his desire to find fault with everything (and, for that matter, by his enviable literary style).

However, I do not agree with Le Cam (1986, p. 81) who declared that

Bertrand and Poincaré wrote treatises on the calculus of probability, a subject neither of the two appeared to know.

I (1991, pp. 164-165) took issue with him as far as Poincaré was concerned. In addition, as indicated throughout my paper just mentioned, Bertrand was Poincaré's main source of inspiration[65]. I also note that in principle a scholar can substantially develop science even without mastering his subject. Darwin, who did not know probability, originated the stochastic theory (more correctly, hypothesis) of evolution.

What else? Le Cam did not say that Bertrand's *Calcu*l [2] was the only contemporary source covering the entire theory of probability. Meyer's important contribution (1874), in spite of its having been translated into German, apparently was not nearly as popular and in any case its range was not so wide. Moreover, the mere fact that the permanent secretary of the Paris Academy of Sciences and a leading French mathematician took much interest in probability, even if deriding it at the same time, made this discipline fashionable. Bertrand initiated the revival of probability in France and I am sure that in this field he stimulated not only Poincaré (above) but also Bachelier and Borel.

I conclude by listing some of Bertrand's specific achievements in probability. His works contain important materials such as the ballot problem (§ 2), an interesting use of a generating function (§ 4.1); the determination of the normal approximation of the hypergeometric distribution (§ 4.2); some theorems on order statistics (§ 7); a critical . note on Gauss's second justification of the MLSq (§ 10); an approach to the theorem on the independence of the mean and the mean square error for a normal sample (§ 17); and, his best known discovery, an elegant proof that the term *at random* is not precise (§ 2). Finally, it was Bertrand who translated into French Gauss's classical work on the treatment of observations.

Acknowledgements. This paper represents a part of a research programme on the history of the theory of errors performed at the Mathematical Institute of the University of Cologne with the support of the Axel-Springer Stiftung. I have discussed Gauss's later attitude to his first substantiation of the MLSq (cf. § 10) and the practical aspect of estimating the precision of observations (§ 14) in my earlier article (1994). Professor J. Pfanzagl has read a preliminary version of this paper and indicated some mistakes and ambiguities.

Notes

A few Notes are omitted since I have shortened my initial text.

1. The long period of time that passed from the appearance of this translation until

Bertrand began his own work on probability testifies that he had no early interest·in the treatment of observations. Without going into detail I (1991, p. 140. note 6) repeat that Bertrand himself [6] later stated that Gauss, who died in 1855, had no time for commenting on the proofs of the translation. I note in addition that Gauss is known to refuse to publish his manuscripts in French.

Gauss's German *Abhandlungen* (1809), as well as their Russian counterpart published in 1957, additionally include §§ 172 - 174 and 187-189 of the *Theoria motus* and Gauss's "Selbstanzeige". The following passage from Bertrand's "Avertissement" to Gauss (1855) shows that he was then a novice. Gauss's contributions, Bertrand wrote,

N'exige ni commentaires ni annotations. Les questions de priorité sur lesquelles se sont engagées des discussions assez vives, y sont traitées brievément mais de la manière la plus nette et la plus loyale. J'ai donc dû me borner au rôle de traducteur: c'était le seul qui fût utile, et le seul d'ailleurs que Gauss m'eût autorisé à prendre.

2. Darboux writes:

Il avait abandonné, en 1878, son cours … et croyait bien avoir renoncé pour toujours aux mathématiques [!].

3. Only three of these [9; 21; 34] are not mentioned. The first one constituted the essence of just one page, and gave its name to the introduction to Bertrand's treatise [2]. The two last-mentioned notes are hardly interesting.

4. I mention only some of them. However, I must say that Bertrand, in the *Préface*, managed to misname the title of Laplace's classic (1812).

5. Elsewhere, while discussing another problem (see Item 2 below), Bertrand (p. 171) mentioned the *nom vague de hasard*. His criticism of the classical definition of statistics (due to De Moivre!) was extremely superficial. It was begging the question and, much worse: it was not a definition but a formula for calculation.

6. Darwin, in 1881 (Sheynin 1980, § 4.7.5), preceded Bertrand. He studied how earthworms dragged small objects into their burrows, whether or not they seized indifferently by chance any part of their find, and considered several possible versions of chance.

Any number of solutions was possible. Thus (Czuber 1924, p. 117): fix one end point of the chord on the circumference of the given circle and suppose that the other end is situated with equal probability anywhere on the circumference; or, choose both end points of the chord in the same way as the second one was just chosen; or, draw the chord through any two points each of them situated with equal probability in any point of the circle.

Much more interesting were the findings of De Montessus (1903) who proved that the probability sought was a function of a continuous argument, of the distance between the centre of the circumference and a point chosen randomly on any of its diameters, either inside the circle or otherwise. A discussion followed; see pp. 343-348 and 464-466 of the same source. See also Sheynin (2003).

7. Bertrand (pp. XVIII-XIX) formulated the same opinion about this problem and the calculation of the probability of the next sunrise (§ 6). Cournot (1843, §§ 232-239 and § 8 of his Résumé) discussed such events whose *probabilité philosophiques*, as he called them, were difficult to express in numbers. He connected this topic with the difficulty of separating remarkable outcomes from ordinary ones. Bru (Cournot 1843, comment on p. 355) referred in this connection to several authors including Laplace. Fries (1842/1974, Preface) introduced philosophical probabilities a bit previously.

8. Professor R. L. Plackett, in a private letter, has since remarked that I did not mention Gower (1982). Also see McCormach (1968).

9. Its origin is evident. The notation however does not correspond to the conditions of the problem since numbers $(m + 1)$ and $(s + 1)$ do not belong to it. The use of such equations in probability goes back to Lagrange.

10. Takacs (1967, p. 3; 1969, p. 895) also stated that De Moivre had proved formula (1) while determining the probability that a gambler, playing with an infinitely rich adversary, will be ruined exactly after n games. Actually, however, Takacs supplemented De Moivre's result (e. g., 1756, pp. 204-210) who had calculated the probability of ruin in not more than n games. Bertrand (§ 5) applied formula (1) for solving the former problem and thus provided an illustration for Feller's subsequent remark (n. d., p. 69):

A great many important problems may be formulated as variants of some generalized ballot problem.
11. At least once Bertrand [8] denoted an expected value by symbol E writing E … (rather than Eξ …) but Meyer (1874, Chapter 3) preceded him in this sense.
12. I (1977b, p. 236) suggested that such bystanders were petty businessmen laying bets on games played by others. Now I note in addition that scholars mentioned these outsiders because they were unable to discuss expected values of *random variables*.
13. Elsewhere he (p. 101) mistakenly connected the Bernoulli theorem with both (rather than with only the second of the) expressions

$$E[(\mu - np)/n]^2 = E(\mu/n - p)^2 = pq/n, \lim P(|\mu/n - p| < \varepsilon) = 1, n \to \infty.$$

Quite unnecessarily, he derived these formulas while describing a fair game by means of a variable taking values $(p; -q)$ with probabilities $(q; p)$ rather than by discussing the usual Bernoulli scheme. Also see note 19.

Up to the end of the 19th century the De Moivre - Laplace theorem was indeed attributed to Jacob Bernoulli (Pearson 1924). De Moivre (1733, p. 243) claimed that

It is now [it was then] *a dozen years or more* since he had discovered his theorem but (p. 244) *had desisted from proceeding farther* [from formulating it decisively] *till my* [until his] *worthy and learned Friend Mr. James Stirling*
disclosed to him the value of the constant in the Stirling theorem, as it is now called. Without explaining the difference between Jacob Bernoulli's theorem and his own, De Moivre (p. 254) called his predecessor an

Acute and judicious Writer who had assigned the Limits within which, by the repetition of Experiments, the Probability of an Event may approach indefinitely to a Probability given.

Todhunter (1865, pp. 190 - 193) described De Moivre's work (including his previous efforts). While correctly stating that De Moivre had considered the particular case of $p = q = 1/2$ he did not add that his hero (De Moivre 1733, p. 250, Corollary 10) had noted that the theorem held for the general case, or that the very title of De Moivre's memoir had indirectly reflected this fact. Todhunter did not even say that any mathematician would have immediately generalized De Moivre's main account. In essence my remarks are known, see Sheynin (1970, pp. 207 and 209), Hald (1990, p. 487). Nevertheless, until this very day De Moivre's achievement are sometimes understood in the narrow sense (Schneider1988, p. 118).
14. This was the second time (end of § 2) that Bertrand mentioned De Moivre. He (§ 5) also borrowed the latter's method of solving a problem on the gambler's ruin.
15. On p. 94 Bertrand remarked that the law, as introduced by Poisson, *manque non seulement de rigueur mais de précision*.
16. Bertrand (p. LXIII) only spitefully and mistakenly ridiculed Quetelet (Sheynin 1986, p. 297) and repeated Cournot's consideration (1843, § 123): a human being with average height, average weight etc. was impossible. Professor W. Kruskal, in a private letter of 1987, remarked that Cournot's reasoning can be strengthened since $E\xi^3 \neq (E\xi)^3$. Then, Frechet, in 1949, replaced the *homme moyen* by the *homme typique*. In any case, Quetelet's notion is useful in economics.
17. On pp. 244-245 Bertrand used a generating function in a similar way. Laplace (1812, pp. 428-429) applied a generating function to determine $E\xi$ for the binomial distribution. Chebyshev (1880, p. 172) did the same to deduce $E\xi$ and $E(\xi - E\xi)^2$.
18. Earlier Bertrand (see below) supposed that the probability of "success" was p rather than q. In both cases he introduced two contrary events without distinguishing between them.
19. On pp. 101-103 he offered a demonstration *plus simple encore* of the Bernoulli theorem, an independent proof that, for the binomial distribution, $E|\xi - E\xi|/n \to 0$. He (p. XXII) also noted that *avec de la patience* it was possible to obtain empirically the number $\pi/2$ since it was roughly equal to $(E\xi - E\xi)^2 : (E|\xi - E\xi|)^2$. As I understand him, he (pp. XXVI - XXVII) suggested that if this ratio for a series of the differences between male and female births deviated significantly from $\pi/2$, statisticians should reveal the reason why. Finally, he (p. XXIV) described the results of his study of a ten-place logarithmic table. He tested the distribution of its seventh digit calculating this same ratio for a series of numbers, but did not explain his method adequately.

20. This problem is due to Huygens (Sheynin 1977b, pp. 241 and 245) who approached it differently.·
21. However, according to modern notion, their models described an increase in entropy and disorder.
22. Later, Poincaré (1912, p. 1) developed this idea, see Sheynin (1991, p. 161).
23. Bertrand thus reveals his philosophical credo.
24. Jacob Bernoulli was the first to put forward formula (2) (in another form) but De Moivre preceded him in publishing it. See Todhunter (1865, pp. 62-63).
25. Bertrand expressly regarded ruin as complete lack of counters rather than as having less than a (or b) of them. Markov (1912, pp. 142-146) took pain to investigate this last-mentioned case. Referring to his earlier and hardly noticed contribution (1903), he also stated, on p. 146, that, in particular, he had given *notwendigen Ergänzungen* to Bertrand's *unstrengen Betrachtungen* on the expected duration of play (see Item 1). I do not describe Markov's considerations. I believe that Bertrand's problem was not really important since nothing changed if quantities a and b were eliminated altogether by replacing pb by p_1 and qa by q_1.
26. Omitted.
27. Omitted.
28. Omitted.
29. Omitted.
30. On the history of formula (3), see Zabell (1989).
31. Bayes himself (1765) was obviously dissatisfied with De Moivre's theorem: he studied the case of a large number of trials without considering the process of $n \to \infty$. Nevertheless, I (1969, pp. 43-44 and 51-54; 1971, pp. 329-330) believe that formula (*) represents the Bayes theorem. I have now noticed that Meyer (1874, p. 200) was of the same opinion, but did not substantiate it. Both Bayes and Laplace (1812, Chapter 6) discussed only uniform prior distributions. It was evidently von Mises (1919, § 9) who first dwelt on the general case. Formula (*) and the previous formula describe the two versions of the law of large numbers (Sheynin 2017, pp. 68-69).
32. On p. 148 Bertrand solved a similar problem involving discrete prior probabilities.
33. In § 93 Cournot mentioned the case of one trial and stated that the evidence thus obtained was insufficient.
33a. Omitted.
34. Bertrand should have mentioned Laplace and Poisson, who had obviously understood that the study of their subject in full was hardly possible. At the very least, Laplace (1816, p. 532) expressly presumed independence. Cournot (1843, § 206) stated that his deductions were founded on the same assumption and discussed the possibility of revealing dependence between the decisions of judges. He (§ 12) even formulated a test of independence for a series of observations. In this connection it is worthwhile to recall the later dispersion theory, § 9.

Poisson's (1837) main aim was the minimization of the unavoidable miscarriage of justice by determining the optimal majority vote of the jurors. See also Sheynin (2017, pp. 131-132).
35. Poincaré, in a rare source (Sheynin 1991, p. 167), denied the application of probability *aux sciences morales* and declared (1912, p. 20) that in law courts people act like the *moutons de Panurge*.
36. He could have noticed for example that for different individuals the chances of dying in a given year depended in the same way on the general conditions of life.
37. In his note, he gave no formulas and offered two examples without sufficiently explaining them.
38. Bertrand (p. XXV) posed a relevant question about the variations between the differences of male and female births (but did not elaborate): *sont-elles assimilables aux résultats capricieux du hazard*?
39. Earlier Bertrand [29] treated this material somewhat differently and in addition linked it with Simpson's. Czuber (1921, pp. 276 -278) discussed this subject in more detail.
40. Poincaré (1896, p. 174) stated that Bertrand had recommended to choose not the maximal, but the expected value of the likelihood function as the estimator sought. Czuber (1924, p. 301) repeated this remark but neither he nor Poincaré gave the

exact reference. I think that they had in mind Bertrand's opinion on a related subject; see § 12.
41. Bertrand largely repeated this statement in his treatise (p. 247).
42. I have edited Bertrand's vague description.
43. He (pp. 257-258) also mentioned the Cauchy distribution (correctly attributing it to Poisson) only to reject it as incompatible with observations. Earlier he [12] additionally referred to Bienaymé as his like-minded predecessor. In both sources Bertrand indicated that the Cauchy distribution could describe the use of a weathercock for determining the cardinal points of a compass. His exposition is very loose and leaves room for objection since the error of the weathercock is not literally infinite.
44. On p. 202 he remarked that he had previously considered errors rather than residuals.
45. For errors Δ_i distributed according to the normal law (7.1),

$$E|\Delta| = 1/k\sqrt{\pi},\ E\Delta^2 = 1/2k^2.$$

Suppose now that $z = |\Delta| - E|\Delta|$. Then

$$Ez^2 = (\pi - 2)/2\pi k^2$$

which is an easily verified expression, and

$$E|\Delta| = (2/k\sqrt{\pi})[\theta(1/\pi) - 1 + \exp(-1/\pi)],$$

$$\theta(t) = (2/\sqrt{\pi})\int_0^t \exp(-x^2)dx.$$

Writing these formulas without proof, Bertrand [36] noted that

$$Ez^2/(E|z|)^2 \approx E\Delta^2/(E|\Delta|)^2.$$

He was not prepared to regard this fact as a coincidence (as it obviously was) and left the issue undecided.
46. He thus did not always restrict himself with his main estimator of precision, the absolute expectation of error. Note that Laplace (1818) introduced the normal law as the actual distribution of errors, not as its limit, thus once more extending his usual approach. See Sheynin (1977a, § 7.1).
47. This is the remark to which Poincaré (note 40) likely referred.
48. Here and below, the residuals are of course those that appear after adjusting the observations by least squares.
49. He usually called them constant. Gauss (1823, § 1) recognized two kinds of errors, random (*irregulares seu fortuiti*) and systematic (*constantes seu regulares*). I see no reason why Bertrand should have hesitated to use the more general term.
Laplace (see below) added for measurements made the same way.
50. Laplace (1814, p. LX; 1816, p. 518) gave the odds as 10^6**:**1. He indirectly excluded systematic errors by adding that future observations will be made the same way.
51. I agree with the first half of Bertrand's final remark (p. 306):
La probabilité pourquoi le hasard, et non la perfection des mesures, les rendre compatibles après de petites corrections peut être considérée comme une impossibilité. On peut, en conséquence, quand la somme des carrés des erreurs est petite, accepter sans crainte le résultat, mais il est téméraire d'évaluer en chiffres la confiance qu'il doit inspirer.
52. Much less is contained in Bertrand's treatise (pp. 281-282).
53. Bertrand used the same example several times [32; 2, pp. 264-265; 272-273, 81-282]. He did not however explain it sufficiently. The description of the imagined figure is my own, and it was not easy to achieve it. Note that an astronomer would have hardly left his station without measuring all of his angles (in this case, angle A_2CA_3 was forgotten).

54. Here and elsewhere, Bertrand could have hardly managed without determining [vv]. I do not therefore agree with his statement [2, p. 288] that *les meilleures valeurs des inconnues* can be obtained by least squares *sans s'astreindre au calcul préalable des erreurs commises sur les grandeurs directement déterminées*. In addition, these errors (residuals) are necessary for the final estimation of precision.
55. Gauss (1823, § 38) remarked that such expressions (he discussed appropriate magnitudes) could be either greater or less than their mean value, but that the difference will decrease with the increase in the number of observations.
56. It is difficult to say what Bertrand (p. XXXVIII) meant by stating that, to oppose the evidence of

Une formule démontrée, c'est a peu pres comme si, pour refuser à un homme le droit de vivre, on alléguait devant lui un acte de décès authentique.

Contrary to his opinion, I would interpret these words as an argument against the rejection of outliers! Even more: Bertrand himself (p. 305) indicated that

Les observations sont des témoins; si elles sont, avant l'épreuve, jugées dignes de confiance, leur déclaration, quelle qu'elle soit, doit être recueillie et conservée.

I have only one explanation of his *avant l'épreuve*: before comparing one observation with another. Barnett & Lewis (1978, p. 360) had nevertheless done away with new statistical tests:

The major problem in outlier study remains the one that faced the very earliest workers […]: *what is an outlier and how should we deal with it*?
57. Professor L. N. Bolshev (1922 – 1978), in an unpublished note which he gave me, estimated the level of significance of Chauvenet's test. According to Chauvenet, the expected number of variables satisfying condition $|\xi_i| > x$ is

$$n[1 - \varphi(x/\sigma)],\ \varphi(x/\sigma) = \frac{1}{\sqrt{2\pi}}\int_0^{x/\sigma} \exp(-z^2/2)dz$$

where σ is presumed to be known. Consequently, the maximal observation $|\xi_i|$ is rejected if

$$n[1 - \varphi(\max|\xi_i|/\sigma)] > 1/2.$$

Now

$$P(\max|\xi_i| > x) \le nP(|\xi_1| > x);$$

$$P(\max|\xi_i| > x) \ge npP(|\xi_1| > x) - [n(n-1)/2]P(|\xi_1|, |\xi_2| > x) \ge$$

$$nP(|\xi_1| > x) - n^2[P(|\xi_1| > x)]^2,$$

so that

$$1/4 \le P(\max |\xi_i| > x) \le 1/2.$$

58. Bertrand called this fact paradoxical. Obviously, he forgot that Gauss (1823, § 21) had explained this fact long ago. True, in adjusting triangulation rough values of its elements were needed to compile some of the necessary initial conditions (e. g., the base condition).
59. This is not an exact quotation. See Laplace (1814, p. CLIII).
60. This is what Heyde & Seneta (1977, p. 67, note) put on record although with respect to Bertrand's later work (see Item 2). In any case, both Laplace (1816), see Sheynin (1977a, p. 36), and Helmert in 1876, preceded Bertrand.
61. Omitted.
62. Elsewhere I (1991, p. 138, note 2) quoted Poincaré in full.
63. No wonder! Hilbert, in 1900, assigned the theory of probability to physics. However, I am unable to agree with Darboux (1902, pp. XLIII-XLIV) on another point: he stated that the theory of errors was *objet des études de toute sa* [Bertrand's] *vie*.
64. That Laplace was guilty in the same sense is no excuse.

65. Poincaré could have studied the work of other previous scholars. That he failed in this respect is quite another story.

References

J. Bertrand

1. *Thermodynamique*. Paris, 1887.
2. *Calcul des probabilités*. Paris, 1888 (see § 1.2). Second edition, 1907. It only differs from the first one in that the page numbers do not coincide. Reprints: New York, 1970 and 1972. 1 did not see the second of these; the first one is a reproduction of the first edition, a fact which is not stated in the publisher's evasive annotation.
3. *D'Alembert*. Paris, 1889.
4. *Pascal*. Paris, 1891.
5. *Eloges académiques*, nouv. sér. Paris, 1902.

Except Items [9] and [35] all articles are from the *C. r. Acad. sci. Paris*.

6. Sur la méthode des moindres carrés; t. 40, 1855, 1190-1192.
7. [Note relative au théorème de M. Bienaymé], t. 81, 458 and 491-492.
8. Sur la théorie des épreuves répetées; t. 94, 1882, 185-186.
9. Les lois du hasard. *Rev. deux mondes*, 15 Avr. 1884, p. 758.
10. Solution d'un problème; t. 105, 187, p. 369.
11. Without title, on duration of play. Ibid., 437-439.
12. Note sur une loi de probabilité des erreurs. Ibid., 779-780.
13. Sur un paradoxe au problème de Saint-Petersbourg. Ibid., 831-833.
14. Théorème relatif aux erreurs d'observation. Ibid., 1043-1044.
15. Sur ce qu'on nomme le poids et la précision d'une observation. Ibid., 1099-1102.
16. Sur la loi des erreurs d'observation. Ibid., 1147-1148.
17. Sur les épreuves répétée, Ibid., 1201-1203.
18. Sur l'association des électeurs par le sort, t. 106, 1888, 17-19.
19. Démonstration du théorème précedent. Ibid., 49-51.
20. Sur la loi de probabilité des erreurs d'observation. Ibid., 153-156.
21. Probabilité du tir à la cible. Ibid., 232-234, 387-391, 521-522.
22. Sur la détermination de la précision d'un système de mesures. Ibid., 440-443.
23. Sur la rigueur d'une démonstration de Gauss. Ibid., 563-565. ·
24. Sur l'indéternination d'un problème résolu par Poisson. Ibid., 636-638.
25. Sur la combinaison des mesures d'une même grandeur. Ibid., 701-704.
26. Sur la valeur probable des erreurs les plus petites dans une série d'observations. Ibid., 786-788.
27. Sur l'évaluation a posteriori de la confiance méritée par la moyenne d'une série de mesures. Ibid., 887-891.
28. Sur l'erreur à craindre dans l'évaluation des trois angles d'un triangle. Ibid., 967-970.
29. Sur les lois de mortalité de Gompertz et de Makeham. Ibid., 1042-1043.
30. Sur la méthode de moindres carrés. Ibid., 1115-1117.
31. Sur la précision d'un système de mesures. Ibid., 1195-1198.
32. Sur les conséquences de l'égalite acceptée entre la valeur vraie d'un polynôme et sa valeur moyenne. Ibid., 1259-1263.
33. Note sur l'introduction des probabilités moyennes dans l'interprétation des résultats de la statistique. Ibid., 1311-1313.
34. Note sur le tir à la cible, t. 107, 1888, 205-207.
35. Sur l'application du calcul des probabilités à la théorie des jugements. *Mém. Soc. philom. Paris à l'occasion centenaire de sa fondation 1788-1888*, 69-75.
36. Note sur un théorème du calcul des probabilités; t. 114, 1892, 701-703.

Reviews

37. Review of P. S. Laplace, *Théorie analytique des probabilités. Oeuvr. compl.*, t. 7, Paris, 1886, *J. des savants*, Nov. 1887, 686-705.
38. Review of C. Lallemand, *Nivellements de haute précision*. Paris, 1889. Ibid., Avr. 1895, 205-213.

Other authors

Abbreviations: AHES = *Arch. hist. ex. sci.*
C. r. = *C. r. Acad. sci. Paris.*

Aaronson J. (1978), Sur le jeu de Saint-Petersbourg. *C. r.*, t. A 286, 839-842.
André D. (1887), Solution directe du problème résolu par·M. Bertrand. *C. r.*, t. 105, 436-437.
Anonymous (1902), Liste de travaux de J. Bertrand [5, pp. 387-399].
Barnett V., Lewis T. (1978), *Outliers in statistical data*. Chichester, 1984.
Barbier E. (1860), Le problèmc de l'aiguille et le jeu du joint couvert. *J. math. pures et appl.*, t. 5, 273-286.
--- (1887) Généralisation du problème résolu par·J. Bertrand. *C. r.*, t. 105, p. 435.
Bayes, T. (1764-1765), An essay towards solving a problem in the doctrine of chances, pts. 1-2. First part reprinted in *Biometrika*, vol. 45 (1970) and in *Studies hist. stat. and prob.* vol. 1. Eds, E. S. Pearson & M. G. Kendall, 134-153. London, 1970.
Second part: *Phil. Trans. Roy. Soc.* vol. 54, 296-325.
Bortkiewicz L. von (1922), Die Variationsbreite beim Gaußschen Fehlergesetz. *Nord. stat. tidskr.* t. 1, 11-38 and 193-220.
--- (1930), Lexis und Dormoy. *Nord. Stat. J.* vol. 2, 37-54.
Chauvenet W. (1863), *Manual of spherical and practical astronomy*, vols. 1-2. Philadelphia-London. Reprint of the edition of 1891: New York, 1960.
Chebyshev (Tchebicheff) P. L. (1867), Des valeurs moyennes. *J. math. pures appl.*, t. 12, 177-184.
--- (1880 [1936]), Theory of probability. [*Teoria veroiatnostei.*] Lecturcs rcad in 1879-1880. Published by A. N. Krylov from A. M. Liapunov's notes. Moscow-Lcningrad. Great many misprints. Krylov obviously did not read the proofs (if existed).
Condorcet M. J. A. N. Caritat De (1847), Eloge de Pascal. *Oeuvres*, t. 3, 567-634. Reprint: Stuttgart - Bad Cannstatt, 1968. The date of the contribution is not stated.
Cournot A. A. (1843), *Exposition de la théorie des chances et des probabilités*. Paris, 1984. Ed., B. Bru. **S, G,** 54.
Cramér H. {1946), *Mathematical methods of statistics*. Princeton.
Czuber E. (1921-1924), *Wahrscheinlichkeitsrechnung*, Bde. 1-2. Leipzig-Berlin.
Darboux G. (1902), Eloge historique de J. L. F. Bertrand [5, pp. Vll-LI]. Lu 1901.
--- (1916), Eloge historiquc d'Henri Poincaré. Lu 1913. In: Poincaré, H., *Oeuvres*, t. 2. Paris, VII-LXXI.
De Moivre, A. (1712), De mensura sortis etc. Engl. transl.: *Intern. stat. rev.*, vol. 52, No. 3, 1984, 229-262. With commentary by A. Hald.
--- (1725), *Annuities on lives*. London. Last edition: De Moivre (1756, pp. 261-328).
--- (1733), A method of approximating the sum of the binomial $(a + b)^n$ etc. In De Moivre (1756, pp. 243-254).
--- (1756), *The Doctrine of chances*. London. Previous editions: 1718, 1738. Reprint: New York, 1967.
De Montessus R. (1903), Un paradoxe du calcul des probabilités. *Nouv. annales math.*, t. 3, 21-31.
Dutka J. (1988), On the St. Petersburg paradox. AHES, vol. 39, No. l, 13-39.
Feller W. (n. d., after 1976), *An introduction to probability theory and its applications*, vol. l. Third ed. 1968, revised printing. New York.
Fries J. F. (1842). *Versuch einer Kritik der Principien der Wahrscheinlichkeits-rechnung*. Braunschweig. *Sämtl. Schriften*, Bd. 14. Halen, 1974.
Freudenthal H. (1951), Das Petersburger Problem in Hinblick auf Grenzwertsätze der Wahrscheinlichkeitsrechnung. *Math. Nachr.*, Bd. 4, 184-192.
Gauss C. F. (1809), *Theoria motus*. Gcrman transl.: Aus der Theorie der Bewegung etc. In author's *Abh. zur Methode der kleinsten Quadrate*. Hrsg. A. Börsch, P. Simon. Berlin, 1887, 92-117. Vaduz, 1998. English translation: Boston, 1865, 2009.
--- (1816), Bestimmung der Genauigkeit der Beobachtungen. *Werke*, Bd. 4. Göttingen, 1880, 109-117.
--- (1823), Theoria combinationis … Gcrman transl.: Theorie der den kleinsten Fehlern unterworfenen Combination der Beobachtungen. In author's *Abh.*, 1887, 1-53. English translation by G. W. Stewart. Philadelphia, 1995.
--- (1855), *Méthode des moindres carrés*. Paris. Traduits … et publiés avec l'autorisation de l'auteur par J. Bertrand.

Gower B. (1982), Astronomy and probability: Forbes versus Michcll etc. *Annals of sci.*, vol. 39, 145-160.
Guyou E. (1888), Note relative à l'expression de l'erreur probable d'une système d'obscrvations. *C. r.* t. 106, 1282-1285.
Hald A. (1990), *A history of probability and statistics and their applications before 1750*. New York.
Helmert F. R. (1875), Über die Formeln für den Durchschnittsfehler. *Astron. Nachr.*, Bd. 85, 353-366.
--- (1876), Genauigkeit der Formel von Peters etc. Ibidem, Bd. 88, 113-132.
Heyde C. C., Seneta E. (1977), *I. J. Bienaymé*. New York.
Jorland G. (1987), The St. Petersburg paradox. In: *Probabilistic revolution*, vol. l. Cambridge (Mass.)-London, pp. 157-190.
Kohli K. (1975), Spieldauer. In: Bernoulli J. *Werke*, Bd. 3. Basel, pp. 403-455.
Kruskal, W. (1946), Helmert's distribution. *Amer. math. monthly*, vol. 53, 435-438.
Laplace P. S. (1812), *Théorie analytique des probabilités*, livre 2. *Oeuvr compl.*, t. 7, No. 2, Paris, 1886, pp. l 81-496.
--- (1814), *Essai philosopliique* etc. Ibid. No. 1. Separate paging. English translation New York, 1995.
--- (1816), Supplément premier to Laplace (1812). Ibid. No. 2, pp. 497-530.
--- (1818), Deuxième supplément to Laplace (1812). Ibid. pp. 531-580.
Le Cam L (1986), The central limit theorem around 1935. *Stat. science*, vol. 1, No. l, 78-96.
Lévy M. (1900), Funérailles de J. Bertrand. *Bull. sci. math.*, t. 24, 69-75.
McCormach R. (1968), J. Michcll and H. Cavendish: weighing the stars. *Brit. J. Hist. Sci.*, vol. 4, 126-155.
Markov A. A. (1903), On the .gambler's ruin. *Izvestia fis.-mat. obshchestva Kazan Univ.*, ser. 2, vol. 13, No. 1, 38-45. In Russian.
--- (1912), *Wahrscheinlichkeitsrechnung*. Leipzig-Berlin. Translation of Russian edition of 1908. Other Russian editions: 1900, 1913, 1924.
Meyer A. (1874), Calcul des probabilités. *Mém. Roy. Sci. Liège*, sér. 2, t. 4, separate paging. German transl. by E. Czuber, *Vorlesungen über Wahrscheinlichkeits-rechnung*. Leipzig, 1879.
Mill J. C. (1886), *System of logic.* London. First ed. 1843.
Mises R. von. (1919), *Fundamentalsätze der Wahrscheinlichkeitsrechnung*. (*Math. Z.*, Bd. 4.) *Sci. papers*, vol. 2. Providence, RI, 1964, 35-56. ·
Pearson K. (1924), Historical note on the origin of the normal curve. *Biometrika,* vol. 16, 402-404.
Poincaré H. (1894), Bcrtrand. Discours prononcé au jubilé de Bertrand. In author's *Savants et écrivains*. Paris, 1910, pp. 157-161.
--- (1896, 1912, 1987), *Calcul des probabilités*. Paris.
Polya G. (1954), *Mathematics and plausible reasoning*. Princeton.
Poisson S. D. (1837, 2003), *Recherches sur la probabilité des jugements* etc. Paris. **S, G,** 53.
Roberts H. V. (1978), Bayesian inference. In: *Intern. Enc. Stat.* vol. 1, pp. 9-16. New York-London. Ed. W. H. Kruskal, Judith M. Tanur.
Rouché E. (1888a), Sur un problème relatif à la durée du jeu. *C. r.*, t. 106, 47-49.
--- (l8&8b), Review of Bertrand [2]. *Nouv. annales math.*, sér. 3, t. 7, 553-588.
Schneider I., Hrsg. (1988), *Die Entwicklung der Wahrscheinlichkeitstheorie von den Anfangen bis 1933*. Darmstadt.
Schreiber [O.] (1879), Richtungsbeobachtungen und Winkelbeobachtungen. *Z. Vermessungswesen*, Bd. 8, 97-149.
Seneta E. (1983), Modern probabilistic concepts in the work of E. Abbe and A. DeMoivrc. *Math. scientist*, vol. 8, 75-80.
Shafer G. (1988), The St. Petersburg paradox. *Enc. stat. sci.*, vol. 8, 865-870. · Eds, S. Kotz, N. L. Johnson. New York.
Sheynin O. B. (1969), On the work of Bayes in probability theory. In Russian. Actually unavailable.
--- (1970), On the history of the De Moivre-Laplace limit theorems. *Istoria i metodologia estestv. nauk*, No. 9, 199-211. In Russian. **S, G,** 103.
--- (1971), On the history of some statistical laws of distribution. *Biometrika*, vol. 58. *Stud. hist. stat. and prob.*, vol. 2. Eds, Sir Maurice Kendall, R. L. Plackett. London, 1977, pp. 328-330.

--- (1973), Finite random sums etc. AHES, vol. 9, 275 – 305.
--- (1976), Laplace's work on probability. AHES, vol. 16, 137 – 187.
--- (1977a), Laplace's theory of errors. AHES, vol. 17, 1 – 61.
--- (l977b), Early history of the theory of probability. AHES, vol. 17, 201-259.
--- (1978), Poisson's work in probability. AHES vol. 18, 245-300.
--- (1980), On the history of the statistical method in biology. AHES, vol. 22, 323-371.
--- (1984), On the history of the statistical method in astronomy. AHES, vol. 29, 151-199.
--- (1986), Quetelet as a statistician. AHES, vol. 36, 281-325.
--- (1991), Poincaré's work on probability. AHES, vol. 42, 137-171.
--- (1994), Gauss and geodetic observations. AHES, vol. 46, 253-283.
--- (2003), Geometric probability etc. *Hist. Scientiarum*, vol. 13, 42 – 53.
--- (2017), Theory of probability. *Historical essay*. Berlin. **S, G,** 10.
Takacs L. (1967), *Combinatorial methods in theory of stochastic process*. New York.
--- (1969), On the classical ruin problems. *J. Amer. Stat. Assoc*., vol. 64, 889-906.
--- (1982), Ballot problems. *Enc. stat. sciences*, vol. 1, 183-188.
Thatcher A. R. (1957), A note on the early solutions of the problem of the duration of play. **(***Biometrika*, vol. 44). *Stud. hist. stat. and prob*., vol. 1, 127-130.
Todhunter I. (1865), *A history of the mathematical theory of probability*. NewYork, 1949, 1965.
Zabell S. L. (1988), The probabilistic analysis of testimony. *J. Stat. Planning and Inference*, vol. 20, 327-354.
--- (1989), Rule of succession. *Erkenntnis*, Bd. 31, 283-321.

V

On the history of the principle of least squares

Arch. hist. ex. sci., vol. 46, 1993, pp. 39 – 54

1. Introduction

1.1. The principle and the method of least squares. The classical method of least squares (MLSq) as used in treating observations served to determine the *true values* of the constants sought and to estimate the plausibility of the obtained results. The principle of least squares (PrLSq), as I call it, was only a certain (least-squares) condition imposed on the residuals (Δw_i) of an inconsistent redundant system of equations

$$a_i x + b_i y + c_i z + \ldots + w_i = 0, \ i = 1, 2, \ldots, n \qquad (1.1)$$

with the number (m) of the unknowns less than n. This principle led to a reasonable solution of such systems without however any estimate of plausibility. In 1809 Gauss justified this principle but not yet the MLSq. I use the colloquial expression least squares when the difference just explained is not important.

1.2. Adjustment of direct and indirect observations. Suppose that system (1.1) has only one unknown. This case is called *direct* (otherwise *indirect*) *observations*. The connection between these cases was understood: the unknowns in both of them were called *milieu* or *Mittel*. Furthermore, when dealing with several unknowns, practitioners usually also thought about direct observations. Thus, Boscovich (p. 501 of his joint contribution of 1770), while substantiating one of the conditions of his main method of solving systems (1.1), mentioned a property of *usual* random errors, i. e., to the case of direct observations. The sum of the residuals, as he reasoned, should vanish[1] since there existed

Un même dégré de probabilité pour la deviations du pendule et les erreurs des Observateurs, dans l'augmentation et la diminution des dégres [in the lengths of the degrees of the meridian arcs].

1.3. The aim of this paper. Gauss had used the PrLSq since 1794 or 1795 and Legendre, in 1805, was the first to publish the MLSq. I am describing related findings before and including 1805 and pay attention to the treatment of direct observations. I begin by studying an important fact in the early history of the arithmetic mean.

I discuss various related topics and in § 7 I defend Euler and Gauss from the horrendous attack by Stigler. Incidentally, the indifference of the statistical community to that attack testifies that it is seriously ill. And nowadays it would be really unpleasant to disown the long-standing hero …

2. Kepler and the arithmetic mean. That Kepler several times chose the arithmetic mean as an appropriate estimator is not surprising. But in one case Kepler (1609/1992, p. 200) adjusted four observations and without any explanation selected as the final value the "medium ex aequo et bono" (in fairness and justice). But that

Latin expression occurred in Cicero and carried an implication *Rather than according to the letter of the law*. It follows that Kepler, who likely read Cicero, called the ordinary arithmetic mean *the letter of the law*, i.e., the universal estimator [of the parameter of location]. In Kepler‘s time or somewhat earlier the arithmetic mean became regarded as law. In more detail see Sheynin (2017, p. 32).

3. The centre of gravity

3.1. Cotes. According to Cotes's celebrated recommendation posthumously published in 1722, the centre of gravity is the estimator of the true position of an unknown point given its observations, see English translation by Gowing (1983). The choice of ordinary or weighted means as estimators of the *true values* of the sought constants is a simple corollary of this recommendation. Such means were used even before Cotes (§ 2) but his authority obviously supported the common feeling[1a]. Laplace lui-même (1814, p. CL) testified that

La régle de Cotes fut suivie par tous les calculateurs,

see however a bit below.

Perhaps independently from Cotes Lambert (1765, §§ 20 – 24) used centres of gravity for fitting straight lines or curves to sets of observational points. This was an example of the connection between the two versions of adjustment calculations (§ 1.2).

Laplace (1812/1886, p. 352) provided a similar example by stating that Mayer (§ 4.3) had used Cotes's rule. He (pp. 351 – 353) also corrected his previous statement (above): no one followed Cotes until Euler and Mayer, but that after 1750

Les meilleurs astronomes ont suivi cette méthode, et le succès des Tables qu'ils ont construites à son moyen en a constaté l'avantage.

Laplace had not noticed that between Cotes and Euler (1749) there was hardly anyone who could have applied or not applied that rule. I also note that *adjustment* at least connotes mechanical considerations and that the German *ausgleichen* was derived from mechanics (Gerling 1843, p. 18):

Diese Analogie [between adjusting observations and calculating the position of the centre of gravity] *hat Veranlassung zu der Redenart gegeben: Die Beobachtungen in's Gleichgewicht setzen, welche gleichbedeutend ist mit dem Ausdruck ausgleichen*.

3.2. Simpson. Stigler (1984, p. 619) discovered the following problem in Simpson's unpublished papers:

From four given points to draw four lines meeting in one and the same point, that ye sum of all their squares shall be a Minimum.

Call it Problem No. 1. Simpson did not bother to write out its solution, neither did he explain its origin. Denote the coordinates of the given points by (x_i, y_i), $i = 1, 2, \ldots, n$, $n = 3$ or 4, and the point sought by (x, y). Then the sum of

$$(x_i - x)^2 + (y_i - y)^2 = \min \qquad (3.1)$$

and x and y are the appropriate means.

No stochastic considerations are involved here, but (x, y) defines the centre of gravity of the given points. The problem is indirectly connected with the treatment of observations and even with least squares. A similar argument is in § 4.1.

Fermat posed this problem with condition (3.1) but with a square root of its left side. His problem has a long and interesting history. For example, Gini, in 1914, examined it while discussing Quetelet's *homme moyen*. See also Farebrother (1990) and especially Monjaret (1991, pp. 48 – 57).

4. Combining equations

Among several standard problems which led to systems (1.1) was the case of adjusting a chain of triangulation laid out between two baselines. This problem was not special but at least up to the 19th century it stymied mathematicians and astronomers, see end of § 7.1. In any case the early methods of combining equations were not connected with triangulation.

Here, I describe a problem in land surveying (§ 4.1), its relation to least squares (§ 4.2) and, in §§ 4.3 and 4.4, discuss special methods of solving systems (1.1) also somehow related to least squares. And in § 4.5 I dwell on the Boscovich method of solving these systems, interesting in another sense.

The problem of § 4.1 is logically connected with the pairing of equations (§ 4.3), but it is only seen when using least squares and was not noticed, if at all, up to 1805. This conclusion is also applicable to the methods of § 4.4.

The solution of some systems of equations by two astronomers could differ, which possibly led Gauss to remark in a letter to Schumacher of 24.6.1850:

Tob. Mayer nicht nach einem systematischen Princip, sondern nur nach hausbackenen Combinationen gerechnet hat.

He referred to Mayer's manuscripts, but likely the published method was the same. In an earlier letter of the same year Gauss himself recommended the same although only for solving a down to earth problem.

4.1. A problem in land surveying. Suppose that the lines drawn from the given points in Problem No. 1 (§ 3.2) do not intersect in one and the same point but rather form a *quadrangle* (or *triangle*, or *polygon*) of *errors*. It is required to determine a plausible position of the *intersected* point without allowing for condition (3.1). This is a standard problem encountered in land surveying for at least a few centuries. Call it Problem No. 2. Denote the equations of the lines by

$$a_i x + b_i y + c_i = 0. \qquad (4.1)$$

The point (x_0, y_0) such that

$$\sum \frac{1}{a_i^2 + b_i^2}(a_i x_0 + b_i y_0 + c_i)^2 = \min$$

will obviously correspond to the least-squares solution of system (4.1).

4.2. Subsystems of equations and least squares. Jacobi and Binet, independently, proved that least-squares solutions of systems (1.1) with $m = 2$ can be represented[2] as (Whittaker & Robinson 1958, p. 251)

$$x_0 = (A_{12}C_{12} + A_{13}C_{13} + \dots + A_{k-1,k}C_{k-1,k})/D,$$

$$y_0 = (B_{12}C_{12} + B_{13}C_{13} + \dots + B_{k-1,k}C_{k-1,k})/D,$$

$$D = C_{12}^2 + C_{13}^2 + \dots + C_{k-1,k}^2$$

and all the magnitudes are the appropriate determinants corresponding to equations i and j ($i, j = 1, 2, \dots, n$, $i < j$) and k is the number of combinations of n things two at a time.

In the general case the least-squares solution is a weighted mean of the solution of m equations with weights C_{ij}^2. For $m = 2$, see Problem 2 in § 4.1, the solutions (A_{ij}/C_{ij}; B_{ij}/C_{ij}) are of course the coordinates of the points of intersection of lines i and j. In both cases it is presumed that the partial solutions exist.

The solutions of Problems 1 (§ 3.2) and 2 do not coincide. Nevertheless, to quote Jacobi (1841, § 1), the connection between solving equations by least squares and by combining them into subsystems provides *eine tiefere Einsicht* into the nature of both methods.

The case of two unknowns was extremely important since the figure of the Earth was generally accepted to be an ellipse of rotation, see also §§ 4.3 and 4.5.

Jacobi also remarked that his theorem was of no practical value, but I hesitate to agree with him. I also note that Gleinsvik (1967) proved two related theorems.

1. Instead of directly solving the subsystems of system (1.1) calculate and solve the corresponding subsystems of normal equations. Then (an almost obvious conclusion) the least-squares solution of system (1.1) is (still) a weighted mean of all the thus obtained solutions.

2. The same inference is valid even if each of the subsystems consists of s equations, $m < s < n$.

4.3. A primitive use of subsystems. In the 18th century one of the methods of solving such redundant systems as (4.1) consisted in arranging the equations in pairs, solving each pair separately and calculating the ordinary mean value of each of the two unknowns over the entire set of pairs. Boscovich (Maire & Boscovich 1770, pp. 483 – 484) had used this procedure before devising his main method of adjusting observations. Mayer (1750), as he himself noted, had to abandon that procedure since the necessary work would have been too tiresome. But least squares had not at once superseded cases involving a moderate number of subsystems (Muncke 1827, p. 872).

Pairing of equations apparently had a second aim: elimination of some systematic errors and a qualitative estimation of the residual random scatter. Thus, drawing on Tycho's biographer Dreyer, Plackett (1958, pp. 122 – 124) noted that Tycho had combined his observations of the right ascension of a certain star in appropriate groups of two[3]. He added three single observations to the twelve thus obtained and calculated the general mean assigning equal weight to each of the fifteen values[4].

The next step was natural. Boscovich (Cubranic 1961, p. 46) calculated the mean difference of latitudes of the end points of his meridian arc measurement not in the usual way, but as a mean over all possible binary combinations of these differences. A few years later he similarly paired his equations.

4.4. Linear combinations of equations. Mayer (1750) whom I mentioned in §§ 3.4 and 4.3, had to solve 27 equations in three unknowns, so he skilfully combined these equations in three disjoint groups, and calculated three sums of nine equations each. Laplace (1788) applied a subtler method. He had 24 equations in four unknowns, combined them into four groups and calculated the sums of all equations; of some 12 of them minus the sum of the other 12 equations; of the sum of some 12 equations; of the sum of the rest 12 equations. [I am unaware of an adequate explanation of his trick.]

4.5. The method of Boscovich. In 1770, he had to consider systems

$$x + b_i y + w_i = 0,\ i = 1, 2, \ldots, n \qquad (4.2)$$

and introduced conditions

$$w_1 + w_2 + \ldots + w_n = 0,$$
$$|w_1| + |w_2| + \ldots + |w_n| = \min \qquad (4.3)$$

the first of which can be readily disposed of by summing up all he equations so that I only discuss condition (4.3).

In the general case of systems (1.1) this condition leads to exactly m (which is the number of the unknowns) zero residuals as noticed by Gauss (1809, § 186)[5]. This means that he knew an important theorem in linear programming (Sheynin 1973a, p. 311). It follows that at least theoretically the Boscovich method consists in solving all the subsystems of m equations each and choosing the solution which corresponded to condition (4.3). Instead of applying all these subsystems (with equal or unequal weights) only one of them is used.

Laplace is known to have solved systems (4.2) by the Boscovich method. In 1832, in his translation of Laplace's *Mécanique céleste*, Bowditch remarked that this method was *peculiarly well adapted* to the adjustment of meridian arc measurements, that it was superior to least squares and ought to be used oftener. This last remark invites a comparison between the arithmetic mean and the median.

In 1760, Simpson met Boscovich and contributed to the latter's method, see § 3.2.

Biot (1811, pp. 198 – 206) described the method of Mayer (§ 4.4), highly praised it as *également applicable* to researches in physics and chemistry and even called it *la seule employee par les astronomes* (p. 203). But on p. 202 Biot concluded that only least squares ensured *la plus grande chance possible d'exactitude*. He referred to Legendre and Laplace but certainly not to Gauss. He did not mention Boscovich at all.

Evidently, Biot overestimated the merits of least squares. About twenty years later Gergonne 1821, p. 186 and 191) denied this accepted view[6].

5. Euler

5.1. The principle of least squares. While commenting on Daniel Bernoulli's memoir of 1778, Euler (1778) recommended the use of an arithmetic mean with posterior weights (x) as the estimator of the *true value* of a measured constant. Suppose that there are n observations

$$\Pi + a,\ \Pi + b,\ \Pi + c,\ \dots,$$
$$a + b + c + \dots = 0.$$

Then, according to Euler,

$$x = \frac{\alpha a + \beta b + \gamma c + ...}{\alpha + \beta + \gamma + ...}, \qquad (5.1)$$

the weights were

$$\alpha = r^2 - (x - a)^2,\ \beta = r^2 - (x - b)^2,\ \gamma = r^2 - (x - c)^2,\ \dots$$

and r was the difference between x and the observation *which is to be all but rejected* [= semi-range].

Euler noted that x (5.1) can be determined from condition

$$\alpha^2 + \beta^2 + \gamma^2 + \dots = \max \qquad (5.2)$$

and that alternatively x was the real root of least absolute value of the equation

$$nx^3 - nr^2x + 3Bx - C = 0, \qquad (5.3)$$

$$B = a^2 + b^2 + c^2 + \dots,\ C = a^3 + b^3 + c^3 + \dots$$

and x was the correction to the arithmetic mean.

Now (Sheynin 1972, p. 50), condition (5.2) is heuristically tantamount to the PrLSq. Indeed, neglect the fourth powers of the errors of observation, i. e., $(x - a)^4$, … and this condition becomes

$$(x - a)^2 + (x - b)^2 + (x - c)^2 + ... = \min. \qquad (5.4)$$

And so x, the correction of the arithmetic mean, effectively vanishes. Euler did not mention it but his numerical examples corroborate my statement.

5.2. The arithmetic mean persists. Von Zach applied Euler's method without indicating that it did not really differ from choosing the arithmetic mean. He (p. 414) introduced six observations whose arithmetic mean was 1.06211 (at least the last two digits were superfluous) and equated r to the range rather to the semi-range of his observations[7]: $r = 0.05485$. Issuing from equation (5.3), he (p. 491) got

$x^3 - 218.206x - 3{,}668{,}999 = 0,\ x = -\ 0.0017.$

I myself calculated somewhat different numbers but the end result did not change. I think that each astronomer worth his salt, had he only noticed that condition (5.2) led to principle (5.4), would imagine that the latter can be applied to the general case of adjusting systems (1.1), see § 1.2[8].

6. Huber and Adrain

6.1. Huber. Daniel Huber (1768 – 1829) was a Swiss astronomer and mathematician. Several authors (Merian 1830; Wolf 1858; Cantor 1881; Fellmann, 1972 and 1992) indicated that he had preceded Gauss and Legendre in discovering the MLSq (I would say the PrLSq). Fellmann, in both cases, declares that this statement *gilt als erwiesen*.

Huber did not publish his discovery and even its date remains unknown. From what follows it may be inferred that it was before 1802 and perhaps even before 1790. Huber began his career as an astronomer, and in 1790, after writing several astronomical papers he was invited to work at the Danzig observatory. However, he declined to leave Switzerland and soon afterwards became professor of mathematics at Basle. From about 1802 he doubled as librarian. In a small way Huber continued his astronomical activities, observations included, at least until 1805 (Wolf 1858, p. 447). In 1815, he became engaged in geodetic work in the field.

The crucial passage concerning least squares is this (Merian 1830, p. 148):

Es ging ihm wie manchen isoliert lebenden Gelehrten in kleinen Städten, dass er manchen guten Gedanken oft lange mit sich herumtrug. … So hatte er z. B. schon in früheren Zeiten, durch eignes Nachdenken, die späterhin durch Gauss und Legendre bekannt gewordene Methode der kleinsten Quadrate … aufgefunden.

By implication, the *earlier times* were those before 1802, as Merian said somewhat before the passage just quoted. However, Dutka (1990) referred to a forgotten paper (Spieß 1939) who had quoted Huber himself. Huber mentioned Legendre's *Maßstab of least squares*.

6.2. Adrain. In 1809, the American mathematician Robert Adrain (1775 – 1843) justified the PrLSq and the normal distribution at about the same time as Gauss and applied it for solving several problems (Dutka 1990). It was Hogan (1977) who ascertained the date 1809 in spite of the formal date 1808 stated in Adrain's journal.

For a very long time Adrain's discovery remained barely noticed and the mathematical involved was unbelievably low. Several authors even before Dutka studied his paper. My own contribution (1965) merits attention only since I discussed his efforts to justify adjustments in land surveying.

7. Defence of Euler and Gauss

It is absolutely necessary to refute Stigler's (1986) venomous fabrications which, for the first (and hopefully the last) time abused the memory of two greatest scientists and at the same time, owing to the blindness of the statistical community, promoted Stigler to the heights of the scientific ladder.

7.1. Euler. Stigler (1986) attacked Euler (1749) no less than nine times, on pp. 5, 55 and 159 and on six occasions on pp. 27 – 34 although the discussion in each case centred on one and the same point. On p. 28 he wrongly and perniciously declared that Euler had

Distrusted the combination of equations, taking the mathematician's point of view that errors actually increase with aggregation.

Again (p. 55), Euler believed that

The combination of observations made under different conditions would be detrimental.

Observations made by different astronomers, perhaps using different techniques etc. are indeed difficult to combine, even psychologically, and different conditions are impossible in metrology, but Stigler obviously did not mean such cases. Euler's failure (to use the proper expression) was occasioned by reasons other than faulty statistical work and Stigler himself had mentioned them.

[After I had refuted him in this paper, Stigler (1999, p. 318), who never acknowledges his mistakes, declared that in 1778, by recommending a return to the arithmetic mean, Euler was proceeding *in the grand tradition of mathematical statistics*. Actually, however, because of Euler's particular requirement, he should have chosen the median!]

Without discussing them, I note now that Euler solved a system of equations by an elementary form of the not yet developed method of minimax. Stigler stated that that method was not good enough but he did not mention that Euler, like Kepler previously and Laplace afterwards, had applied it to ascertain whether or not a theory stood an observational test. Witness indeed Kepler's celebrated remark that the Ptolemaic system of the world did not fit Tychonian observations and take into account Laplace's doubts about the exact ellipticity of the figure of the Earth. He used the minimax method in quite a few of his contributions but I only refer to my own description of his reasoning (Sheynin 1977, pp. 48 – 49). And could it not be that Euler did not proceed further since he felt that his research was not comprehensive?

Stigler did not name the mathematicians who had believed that errors increase with propagation, so I help him. Laplace and Legendre feared that, at least in triangulation this can indeed occur. At the turn of the 18th century the French *Commission des Poids et Mesures* (Laplace, Legendre, Delambre, Méchain and several *savants étrangers*) *refusa de compenser les angles* of a certain chain of triangulation between baselines. Instead, the Commission decided to divide the chain in two parts and to adjust separately the triangles of each (and apparently to adjust somehow the length of the side of triangulation common to both parts). Laplace (1812/1886, pp. 590 – 591) described this episode, see also Méchain & Delambre 1810, pp. 415 – 433). Bru (1988, pp. 227 – 228) who reminded his readers about this fact also noted on pp. 225 – 226) that Maupertuis, Bouguer and Condamine had experienced great difficulties when deciding how to adjust triangulation if at all and that La Caille, in

1744, was the first who dared to perform arbitrarily that procedure but only elementarily.

7.2. Gauss, the main culprit. Stigler (p. 143): Only Laplace saved Gauss' justification of the PrLSq from passing *relatively unnoticed* and joining *an accumulating pile of essentially ad hoc constructions*.

More (p. 146), since a single glaring stupidity was not enough: there is no indication that, before Legendre's publication Gauss had noticed the *great general potential* of least squares; that (pp. 145 and 146) Gauss had

Solicited reluctant testimony from friends that he [had] *told them of the method before 1805.* […] *Although Gauss may well have been telling the truth about his prior use of the method, he was unsuccessful in whatever attempts he made to communicate it before 1805*.

Even before 1805 Gauss used the PrLSq to calculate the orbit of a lost dwarf planet (Ceres) by means of only a few available observations. His success stirred up curiosity and astronomers became eager to acquaint themselves with Gauss' method of calculation, see for example Volk (1957, p. 208). It was impossible to forget the yet unpublished principle. And since Gauss used it systematically it is lunacy to suggest that he failed to realize its general potential.

After *Theoria motus* was published the PrLSq became immortal. Owing, in particular, to its classical elegance his own mature substantiation of the PrLSq (1823) remained barely noticed. As to the *ad hoc construction*, it can almost refer to Legendre. Gauss was *unsuccessful in communicating* etc. is nonsense. For example, Bessel (1832, p. 27) became familiar with Gauss' invention before 1805 *durch eine mündliche Mittheilung*. That Gauss *solicited* … was a damned lie whereas the phrase *may well have been telling the truth* … was appropriate with respect to a suspected rapist. See all this in detail in my *Antistigler* (**S, G,** 31).

Acknowledgement. This paper represents a part of a research programme on the history of the theory of errors performed at the Mathematical Institute of the University of Cologne with the support of the Axel-Springer-Stiftung. Dr. C. Eisenhart informed me about Whittaker & Robinson (1958), see § 4.2. The reference to Delambre (Note 1a) is due to Prof. F. Schmeidler. Dr. F. Nagel and Dr. E. A. Fellmann sent me some important materials and the latter allowed me to quote from his then yet unpublished paper. I benefited from suggestions made by Prof. J. Pfanzagl who had looked through a preliminary version of this paper.

Notes

1. I return to this condition more than once and it is opportune to add that it is called after Cauchy (1853, p. 40). He introduced it without any references. His work was devoted to interpolation and commentators do not usually connect it with my subject. Noteworthy exceptions are Idelson (1947, § 21) and Linnik (1961; § 5 of chapter 15 of Russian edition of 1962). The work of Cauchy is beyond my chronological boundaries and I only note that his *method*, as testified by Idelson, was applied at least up to the 1940s, to solve systems (1.1). Delambre (1814, p. 309) preceded Cauchy while solving large systems of equations.

1a. Cf. Delambre (1827, p. 455):

Ce moyen a été employé de nos jours par plusieurs géomètres, qui ont pu le trouver d`eux-mêmes, il n'en est pas moins juste d'en faire honneur à Cotes, qui paraît en avoir en la première idée.

1b. Stigler does not refer to Laplace's *Oeuvres complètes* but it seems that he had mentioned the same passages as I did.

2. Actually they considered the general case of $2 \le m < n$.

3. Evidently, Tycho began by designing his observations.

4. Perhaps he had sufficient reasons to assign double weight to single observations.

5. He had not proved his statement. Waterhouse (1990) offered a possible reconstruction of the proof. It resembles, even if heuristically, the treatment of systems (4.3) as described by Boscovich.

6. I (1976, p. 180, Note 1) stated that similarly even Poisson had not mentioned Gauss either in his memoirs on the theory of errors of 1824 and 1829 or in 1833, in his obituary of Legendre. This former attitude of French mathematicians towards Gauss is well known. Concerning the optimality of the MLSq it is now known that it it is not universal (Petrov 1954). Thus, cf. above, the median can be a better estimator of the location parameter than the arithmetic mean.

7. Apparently a precaution since the number of observations was too small.

8. I (1973b, p. 123) remarked that Gauss, in 1809, had possibly issued from Daniel Bernoulli and Euler (both 1778) but he hardly saw their contributions.

9. My review (1988) was not detailed enough. I repeat and enlarge on my remarks only insofar as they are relevant here.

References

AHES = this *Archive*

Bessel F. W. (read 1832); Über den gegenwärtigen Standpunkt der Astronomie: In author's *Popüläre Vorlesungen* … Hamburg, 1848, pp. 1 – 33.

Biot J.-B. (1811), *Traité élémentaire d'astronomie physique*, t. 2. Paris. Second edition.

Boscovich R. J. (1757), De litteraria expeditione per pontificam ditionem … *Memoriae de Bononiensi scientiarum et artium Instituto atque academia*, t. 4, pp. 353 – 394, this being a summary of the original Latin edition of Maire & Boscovich (1770). See Cubranic (1961).

Bru B. (1988), Laplace et la critique probabiliste des mesures géodésiques. In *La figure de la Terre*. Eds, H. Lacombe, P. Costabel, pp. 223 – 244. Paris.

Cantor G. (1881), Huber Daniel. In *Allg. deutsche Biogr.*, Bd. 13, pp. 228 – 229. Reprint: Berlin, 1969.

Cauchy A. L. (1853), Sur l'évaluation d'inconnues déterminées par un grand nombre d'équations approximatives du premier degré. *Oeuvr. Compl.*, sér. 1, t. 2. Paris, 1900, pp. 36 – 46.

Cubranic N. (1961), *Geodetski rad R. Boskovicá*. Zagreb. The main part of the book is occupied by a reprint of Boscovich (1757) and its Serbo-Croatian translation.

Delambre J. B. J. (1814), *Astronomie théorique et pratique*, t. 2. Paris.

--- (1827), *Histoire de l'astronomie aux dix-huitième siècle*. Paris.

Dutka J. (1990), R. Adrain and the method of least squares. AHES, vol. 41, pp. 171 – 184.

Euler L. (1749), Recherches sur la question des inégalités du mouvement de Saturn et de Jupiter. *Opera omnia*, ser. 2, t. 25, pp. 45 – 157. Turici.

--- (1778), Observationes … Translation: Observations on … Bernoulli. See Kendall (1961).

Farebrother R. W. (1990), Further details of contacts between Boscovich and Simpson. *Biometrika*, vol. 77, pp. 397 – 400.

Fellman E. A. (1972), Daniel Huber. *Neue deutsche Biogr.*, Bd. 9, pp. 689 – 690.

--- (1992), Same. *Schweizer Lexikon*, Bd. 3, p. 493.

Gauss C. F. (1809), Theoria motus. German transl. (1887) of relevant part. *Abh. zur Methode der kleinsten Quadrate*. Hrsg. A. Börsch, P. Simon. Berlin, pp. 92 - 117. Vaduz, 1998.

--- (1975 – 1997), *Werke, Ergänzungsreihe*, Bde 1 – 5. Hildesheim. Correspondence with Bessel (Bd. 1, 1880/1975); with Gerling (Bd. 3, 1927/1975);

with Olbers (Bd. 4, 1909/1976); with Schumacher (Bd. 5, 1836/1975).
Gergonne J. D. (1821), Dissertation sur la recherche du milieu le plus probable … *Annales math. pures et appl.*, t. 12, No. 6, pp. 181 – 204. Published anonymously.
Gerling C. L. (1843), *Ausgleichungsrechnung* … Hamburg.
Gleinsvik P. (1967), Generalization of the theorem of Jacobi. *Bull. géod.*, t. 85, pp. 269 -281.
Gowing R. (1983), *Roger Cotes: natural philosopher*. Cambridge.
Hogan E. R. (1977), R. Adrain: American mathematician. *Hist. Math.*, vol. 4, pp. 157 – 172.
Idelson N. I. (1947), *Sposob naimenshikh kvadratov* (Method of least squares etc.). Moscow. **S, G,** 58 (Chapter 1).
Jacobi C. G. J. (1841), *Über die Bildung und die Eigenschaften der Determinanten*. *Ostwald Klassiker* No. 77. Leipzig, pp. 3 – 49.
Kendall M. G. (1961), Daniel Bernoulli on maximal likelihood. Incorporates translations of D. B. and Euler (1778). *Biometrika*, vol. 48, pp. 1 – 18. Reprint: *Stud. hist. stat. prob.,* Eds, E. S. Pearson, M. G. Kendall. London, 1970, pp. 155 – 172. Commentaries superficial.
Kepler J. (1929), *Neue Astronomie* (1609, Latin). München – Berlin. *New Astronomy*. Cambridge, 1992, 2015.
Lambert J. H. (1765), Theorie der Zuverlässigkeit der Beobachtungen und Versuche. In author's *Beyträge zum Gebrauche der Mathematik* …, Tl. 1. Berlin, pp. 424 – 488.
Laplace P. S. (1788), Théorie de Jupiter et de Saturne. *Oeuvr. Compl*., t. 11, Paris, 1895, pp. 95 – 239.
--- (1812), *Théorie analytique des probabilités*. Ibidem, t. 7, No. 1 – 2. Paris, 1886.
--- (1814), *Essai philosophique* … Included with separate paging as Introduction to later editions of the *Théor. anal. prob*. English translation: New York, 1995.
Linnik Yu. V. (1961), *Method of least squares* … Oxford. Initially in Russian, 1958. Not suited for geodesists.
Maire C., Boscovich R. (1770), *Voyage astronomique et géographique* … Paris. Initially in 1755, in Latin.
Mayer T. (1750), Abhandlung über die Umwalzung des Monds um seine Axe … *Kosmogr. Nachr. und Sammlungen* für 1748, pp. 52 – 183.
Méchain, Delambre (1810), *Base du système métrique*, t. 3. Paris.
Merian P. (1830), Daniel Huber. In *Verh. der allg. schweiz. Ges. für die ges. Naturwiss. in ihrer 16. Jahresversammlung zu St. Gallen 1830*. St. Gallen, pp. 145 – 152.
Monjardet B. (1991), Eléments pour une histoire de la médiane métrique. In *Moyenne, milieu, centre. Histoire et usages*. Eds, J. Feldman et al. Paris, pp. 45 – 62.
Muncke (1827), Erde. In *Gehlers phys. Wörterb*. Bd. 3. Leipzig, pp. 825 – 1141.
Petrov V. V. (1954), The method of least squares and its extreme properties. *Uspekhi matematich. nauk*, vol. 1, pp. 41 – 62. In Russian.
Plackett R. L. (1958), The principle of the arithmetic mean. *Biometrika*, vol. 45, 130 – 135. Reprint. *Stud. hist. stat. prob*., pp. 121 – 126.
--- (1972), Discovery of the method of least squares. *Biometrika*, vol. 59, pp. 239 – 251. Reprint: *Stud. hist. prob. stat*., vol. 2. London, 1977. Eds, Sir Maurice Kendall, R. L. Plackett, pp. 279 – 291.
Sheynin O. B. (1965), On Adrain's wok in the theory of errors. *Istoriko-matematich. issledovania*, vol. 16, pp. 325 – 336. **S, G,** 1.
--- (1972), On mathematical treatment of observations by Euler. AHES, vol. 9, pp. 45 – 56.
--- (1973a), Boscovich's work on probability. AHES, vol. 9, pp. 306 – 324.
--- (1973b), Mathematical treatment of astronomical observations. AHES, vol. 11, pp. 97 – 126.
….--- (1976), Laplace's work on probability. AHES, vol. 16, pp. 137 – 187.
--- (1977), Laplace's theory of errors. AHES, vol. 217, pp. 1 – 61.
--- (1979), Gauss and the theory of errors. AHES, vol. 20, pp. 21 – 72.
--- (1988), Review of Stigler (1986). *Centaurus*, vol. 31, pp. 173 – 174.
--- (1992), Al-Biruni and the mathematical treatment of observations. *Arabic sciences and phil*., vol. 2, pp. 299 – 306.
--- (n. d.), *Antistigler*. **S, G,** 31.

….--- (2017), *Theory of probability. Historical essay*. Berlin. **S, G,** 10.
Spieß W. (1939), Kann man für D. Huber Ansprüche als Erfinder der Methode der kleinsten Quadrate geltend machen? *Schweiz. Z. Vermessungswesen u. Kulturtechnik,* Bd. 37, pp. 11 – 17, 21 – 23.
Stigler S. M. (1976), The anonymous Professor Gergonne. *Hist. Math.*, vol. 3, pp. 73 – 74.
--- (1984), Boscovich, Simpson and a 1760 manuscript note. *Biometrika*, vol. 71, pp. 615 – 620.
--- (1986), *The* (!!) *history of statistics*. Cambridge, Mass.
--- (1999), *Statistics on the table*! Cambridge, Mass.
Volk O. (1957), Astronomie und Geodäsie bei Gauss. In *Gauss, Gedenkband*. Hrsg. H. Reichardt. Leipzig, pp. 205 – 229.
Waterhouse W. C. (1990), Gauss's first argument for least squares. AHES, vol. 41, pp. 41 – 52.
Whittaker E. T., Robinson G. (1958), *Calculus of observations*. London.
Wolf R. (1858), Daniel Huber von Basel. In author's *Biographien zur Kulturgeschichte der Schweiz*, 1. Cyclus. Zürich, pp. 441 – 462.
Von Zach F. X. (1805), Versuch einer auf Erfahrung gegründeten Bestimmung terrestrischer Refractionen. *Monatl. Correspondenz*, Bd. 11, No. 5, pp. 389 – 415; No. 6, pp. 485 – 504. Published anonymously. Von Zach was Editor.

VI

William Farr

Vital statistics

Humphreys N. A., Editor, Vital statistics. *Memorial volume* (extracts)
Sanitary Inst. Gr. Brit., 1885, pp. 166 – 205

1. *Life and death in England*. How the people of England live is one of the most important questions that can be considered; and how, of what causes, and at what ages, they die is scarcely of less account; for it is the complement of the primary question teaching men how to live a longer, healthier and happier life.

The vital units to be specially dealt with are persons living and persons dying in the ten years 1861 – 70 only distinguishing them into units representing males and females of different ages and occupations, losing life year after year by various causes in about 627 districts extending from the borders of Scotland to the English Channel and from the Irish Sea to the German Ocean [North Sea]. The deaths in the several classes have to be compared with the population enumerated at three decennial censuses, in corresponding groups.

The long series of Tables offers a retrospect extending over the ten years, with which it is compared. The primary object is to determine what the death-toll [equivalent to rate of mortality – author] is at the several ages, and what the causes of the loss of life are, under different circumstances. The importance of this determination will become apparent by enumerating some of the relations the mortality bears to other orders of facts. There is a relation betwixt death and sickness; and to every death from every cause there is an average number of attacks of sickness, and a specific number of persons incapacitated for work. Death is the extinction of pain. There is a relation betwixt death, health and energy of body and mind. There is a relation betwixt death, birth and marriage. There is a relation betwixt death and national primacy: numbers turn the tide in the struggle of populations, and the most mortal die out[1].

There is a relation betwixt the forms of death and moral excellence or infamy; men destroy themselves directly or their fellows under the most varied mental conditions; they may die by indulgence in excesses, by idleness, or by improvidence. Death is met especially in primeval races not only in conflicts with each other, but in conflicts with other races of animals – directly with great carnivorous quadrupeds or creeping poisonous serpents, and indirectly with four-footed animals, winged birds, and multitudinous insects, blighting or consuming food. Death is also wrought by low but organised parasites in the body. It is still more frequently the result of elementary molecules (zymads[2]) which, though of no recognised form, evidently thrive, propagate, die in the bodies of men, disintegrating or devitilizing their issues.

There is finally a relation betwixt death and the mean lifetime of man. If a life passing through a given time is represented by a line, death is the point of termination as birth is the point of origin. And a generation of men born together is represented by an indefinite number of such lines of life. The natural lifetime of man is a century; that age under ordinary conditions is, as the Etruscans[3] remarked, attained by at least *one* in every considerable generation, and they made it their seculum[4]; as in that time are passed through all the phases of childhood, youth, manhood, maturity, and monumental age[1]. The mean lifetime in the healthiest districts of England, and in the healthiest ranks, is 49 years[5]. And we have no evidence that under the most favourable conditions it exceeds 50 years.

Actually individual life varies in duration from a second to a century. And the relation to be shown here is between the dying by different causes and the living at every stage of the march of a generation through time. The mean lifetime of a generation may be the same, and yet the several lifetimes of the individuals of which it is composed may vary infinitely. Under the actual laws of mortality, great numbers die in infancy, few in adolescence, more in manhood, and, after infancy, the greatest numbers by the English Table[6] at the age of *seventy-three*, the numbers born living fallen in the proportion of ten born alive to two then surviving.

It is evident that an entire revolution in the life of the human race would follow if every person born lived the average lifetime of fifty years, or if half the deaths happened in infancy and the other half at the end of 100 years or at any very advanced age. What we observe actually is that in certain conditions the mean lifetime sinks to half its standard length, and that this is the result of the high mortality in the first five years, of the reduced mortality in adolescence, and of the increasing mortality in the manhood up to the ultimate term of life. Few old people surviving and few dying therefore after four score years, especially in such unfavourable conditions as exist in Liverpool.

Under the existing state of things, of the constituent lives of every generation a certain number dies at every age of causes to be investigated under two heads: direct and organic, including diseases and injuries, and remote and indirect, namely, the causes of those diseases and injuries. Before entering upon the investigation two preliminary questions have to be discussed. (Supplement to 35th Annual report, pp. 3 – 4.)

Arranging the districts of England in the order of their mortality, it is found that the annual mortality in the various groups ranges from the rate of 15 to 39 per 1000; the birth-rate from 20 to 40 per 1000, and it is seen that, in the next Table, as the death-rate increases, the birth-rate increases, so that in all the districts with a mortality under 25 per 1000 the natural increase of population is very constant.. The mortality increases with the density of the population, and thus every additional death is met by an additional birth.

2. *Density of population, death-rate, birth-rate, excess of births over deaths, and increase of population per 1000 persons living, in seven groups of districts arranged in the order of mortality.* [Farr himself describes the collected data (below), and I do not copy this table.]

In the first stage of the scale, that is, in the 54 healthy districts, the death-rate is 16.7, the birth-rate, 30.1. In the second stage, 19.2 and 32.2; in the third stage, 22.0 and 35.6, in the fourth stage, 25.1 and 38.1.

The natural increase of population in each of these four stages ranges from 13.0 to 13.6, or is severally 13.4, 13.0, 13.6, 13.0. When the mortality reaches the *fifth* stage the death-rate is 27.8 and the birth-rate 39.1. After that point, while the death-rate increases to 32.5 in Manchester and 38.6 in Liverpool, the birth-rate reduces to 37.3 and 37.6, and there is a decrease of indigenous population, which if it should go on might end in a decrease of population in geometric progression.

Should the deaths in the districts where the mortality is 22.0 per 1000 be reduced by sanitary measures to the same level as in the districts where the mortality is 19.2, the births might be reduced in the same or a greater degree, namely, from 35.6 to 33.2; and should the death-rate be brought down to 16.7, the birth-rate might be reduced, as in the healthiest districts to 30.1; the deaths falling 5.3, the births actually fell 5.5 per 1000 as shown in the table. The fall of the birth rate is observed in the existing circumstances of this country. It maintains an uniform increase in districts under different laws of mortality, but it is not a necessary consequence of a reduced death-rate, and if, in the opinion of the parties concerned, their prospects are good, they marry and procreate children at the same rate as before. In that case the population increases faster, whereas in a depressed condition the births fall off until the population becomes stationary or declines.

There is no inevitable connection between the gradual reduction of the mortality of the whole kingdom to the rate of 17 per 1000 and the more rapid increase of population. Because the birth-rate may of itself fall to the level of that now prevailing in the healthiest districts and leave the increase of population as it was. Statesmen are not then, by alarming cries of increase of population in a faster geometrical progression, to be deterred from the noblest work in which they can engage. For it is certain that population as it improves in England will not increase faster than the requirements of industry in all its forms at home or the new openings of colonial enterprise abroad. (Supplement to 35th Annual report, pp. xii – xiv.)

3. *Probable decrease of mortality*. There are many obstacles to the sanitary progress of a nation, and it is evident that at present they can only be overcome in part, but there is no ground for despair. There has been progress. The mean lifetime of sovereigns and peers is prolonged, it was in past ages much shorter than the lifetime of the unhealthy labourers in the cities of today. The mortality of the city of London was at the rate of 80 per 1000, in the latter half of the seventeenth century, 50 in the 18th century and 24 at present. The mortality in the liberties of the city of London within and without the walls[7] was in the four plague years 1593, 1625, 1636, 1665, at the rate of 24, 31, 13 and 43 per cent.

In the city alone 90,472 persons died of plague in the four epidemics, and 55,604 of other diseases. The enumerated population of the city was 130,178 in 1631. In the cholera epidemic year of 1849 the mortality from all causes in the metropolis was only 3 per cent. And in the last two epidemics there was a further decline. Thus it is as certain that the high mortality can be reduced by hygienic appliances down to a certain limit as it is that human life can be sacrificed.

The analysis of the causes of the mortality renders it still further certain that the actual mortality of the country can be reduced. Many of the destroyers are visible, and can be controlled by individuals, by companies, and by corporate bodies, such as explosions in coal mines, drowning in crazy ships, railway collisions, poisonings, impurities of water, pernicious dirt, floating dusts, zymotic contagions, crowding in lodgings, mismanagements of children, neglects of the sick, and abandonments of the helpless or of the aged poor.

Furthermore, including the London district of Hampstead, there are fifty-four large tracts of England and Wales which actually experience a mortality at the rate of only seventeen per 1000, less by *five* than the average mortality per 1000 of the whole country, less by *ten* than than in nine districts and less by *twenty-two* than the mortality reigning for ten years in Liverpool. Now the healthy districts have a salubrious soil, and supply the inhabitants with waters generally free from organic impurities. The people are by no means wealthy, the great mass of them are labourers and workpeople on low wages, whose families get few luxuries and very rarely taste animal food. Their cottages are clean but are sometimes crowded, and impurities abound, the sanitary shortcomings are palpable[8].

It will not therefore be pitching the standard of health too high to assert that any excess of mortality in English districts over 17 *annual deaths* to every 1000 living is an excess not due to the mortality incident to human nature, but to foreign causes to be repelled, and by hygienic expedients conquered.

It is right to state that the real is greater than the apparent mortality of these districts. They are increasing and contain an undue proportion of population at the younger healthiest ages, so that a correction for this makes the mortality 20 instead of 17. That is the rate of their stationary mortality if the population were stationary, if births equalled deaths and there were no migration.

The mean annual deaths at the rate of 22.4 in the ten years 1861 – 70 were 479,450 in England. And had the rate of mortality been 17 the annual deaths would not have exceeded 363,617, so the overplus due to the operation of causes existing but less destructive in the healthier districts was 115,838. The hope of saving any number of these 115,838 lives annually by hygienic measures is enough to fire the ambition of every good man who believes in human progress. (Supplement to 35th Annual report, pp. viii – ix.)

4. *Possibilities and difficulties of extending human life*. The laws of life are of the highest possible interest, even if the knowledge of those laws gave man no more power over the course of human existence than the meteorologist wields over the storms of the atmosphere, or

the astronomer over the revolutions of the heavens. But all human laws proceed on the belief that the lives of individuals and of the communities can, within certain limits, be regulated for good or for evil. And as latterly this has been questioned, it becomes necessary to discuss the problem: can lifetime be prolonged by knowledge of the causes that cut it short, or by any means within a nation's power?

To live long is a natural aspiration, and in the early years of the marvellous science of chemistry the alchemists sought with as much ardour as they sought the philosopher's stone for an *elixir vitae* to confer on man perpetual prime. They promised him, by its discovery, immortality upon earth. The possibility of this seems to have been an ancient belief, for in one of the oldest legends man had been told that he should not die, that he should live for ever. And it had in it some grounds, or it could never have led the first [Roger] Bacon, Descartes, Franklin and Condorcet to intimate that human life might be prolonged indefinitely.

The forces, as well as the constituents of the body are in truth indestructible, but they are fugitive, and they are perpetually passing out of the men of existing generations into other forms. The flame of consciousness shines in one life only for a while. But the alchemists were right when they saw virtues in minerals and trees to prolong as well as to shorten life, to check disease and to set the body free. For if mercury, arsenic, antimony, iron, potash, soda, magnesia, phosphorus, chlorine, iodine, sulphur, in their various salts and acids; if strychnia, quinine, opium, chlorophorm, aether, ipecacuanha, camphor, and alcohol will kill, they are also cure in the hands of the skilful. Surgery too has its great triumphs. Therapeutics is not a delusion, the Healer is a reality. But no drug can do more than prolong life for a time, the man raised from the grave dies in the end.

Life can be lengthened by regimen, by dietetics, which Celsus says engaged in his day the most eminent professors (?) of medicine in Rome, because it is the most potent and philosophical, dealing in regimen of mind and body and medicinally controlling aliment, air, sleep and exercise. The influence of the external world of air, water, soil and climate on health and length of life was placed beyond doubt by the great treatise of Hippocrates. And Moses had before indicated the exclusion of the sick by zymotic diseases from the Congregation. In these latter days science has gone further and shown under which conditions the lifetime is long or short. And the science of life, yet only in its infancy, will make further progress and solve many problems hitherto held to be insoluble, when hygiene is cultivated in all the medical schools. The genius of agriculture, of engineering, of industry and commerce is growing every year and handling new power in new machines, is supplying new means of existence and banishing fatal impurities.

Descent is easy and onward motion over a level road is not difficult, but every step upwards to a higher state encounters obstacles. And so it is in the improvements of the human race. Of this a few examples are instructive. Smallpox is a fatal disease and after it had been learnt that a milder type could be induced artificially, fatal to few of the inoculated, the practice was introduced in London and was publicly performed in the years 1746 – 63 on 3434 persons at the smallpox hospital. Only 60 of whom it is said died of the disease[9]. The mortality varied in different places but it was nowhere considerable. What appeared so well fitted to justify Lady Mary Wortley Montague's exultation when she learned in Turkey that "ingrafting" rendered smallpox harmless? In 1718 she wrote:

I am patriot enough to take pains to bring this useful invention into fashion in England.

But it was found after it was brought that the deaths from smallpox in London compared with the deaths from all other causes and also the absolute mortality increased considerably when inoculation became common. Large numbers of children and adults remained unprotected and inoculation kept the *varioilads* alive in an artificial nursery. Inoculation is now made illegal.

Again, hospitals were opened to receive people attacked by this dreadful disease and to afford them the advantages of watchful attendance and skilful advice. This was carried out in London, but the mortality of the patients in the hospital was double the mortality by the disease outside[10]. Here was another apparent failure. But vaccination was a great advance on inoculation. The danger of the operation was quite inconsiderable, and cowpox, unlike smallpox, never scattered abroad the seeds of disease. In 1771 – 80 smallpox in London was the cause of 100 in every 1000 deaths, in 1831 – 5, of 27, in 1861 – 70, of 11, and in the absolute mortality by this disease there was a large reduction.

In the last two decennials 1851 – 70, the mortality per 100,000 by smallpox remained stationary in London at 28. In all England the mortality per 100,000 by smallpox declined from 22 to 16, or to the extent of 6. But population growing denser the mortality by scarlet fever rose from 88 to 97, thus increasing 9, or one and a half times as much as the mortality by smallpox decreased. The mortality by measles, diphtheria, and whooping-cough also increased. Vaccination diminished the chances of taking smallpox and though it did not afford absolute security[11], it reduced the danger of its attacks. But, density of population increasing, other zymotic principles appeared to find in its absence freer scope for their destructive operations In quenching the flames at one point the good work is begun but it is not ended. Can zymotic diseases of all kinds never be quenched?

Out of pity for poor children Foundling hospitals were erected but the babies nearly all perished, and a greater number than ever were abandoned[12]. Had these hospitals succeeded the race of child-abandoning men must have been multiplied.

Another example is offered by the drainage of towns. In London the fatal refuse which had been retained in the houses was conveyed by water into the drains and into the Thames. And this was an advance on the previous state of things, but the sewers were charged with impurities. They put houses by their effluvia in communication with each other, and poured zymotic elements into the waters which were distributed by companies to the houses of both the wealthy and the indigent. And even at the present hour the sewage is pumped into the Thames which it pollutes and obstructs, instead of being distributed over the land to which it belongs. The same difficulty in disposing of sewage is encountered in all English towns.

In the early ages the English population was scattered in slight dwellings over woods, meads and undrained marshland where they suffered from agues, rheumatisms, and famine fevers. As the people multiplied they assembled in cities and partook of a few of the advantages of civilization. But the increase of density brought new dangers, and, as the proximity of houses exposed towns to conflagrations, it laid their inhabitants open to devastating maladies and to destructive pestilences. The people flocked in numbers to London in the reigns of Henry VIII, of Elizabeth, and of James [1508 – 1557, 1558 – 1603, 1685 – 1688], and the sweating sickness and fevers and the oriental plague decimated the population[13].

The Restoration [1649 – 1660] brought country families to the metropolis, and the plague made its ever memorable swoop. The manufactures, the mines, and the great works hat create subsistence for thousands, collect workmen in towns as ill-provided with sanitary appliances as ill-organised camps. And thus Lancashire, Yorkshire, Durham, South Wales are still in a high degree insalubrious. Until the Legislature led by Lord Shaftesbury, intervened, the lives of young children and mothers were barbarously sacrificed in the factories. Here is seen again the success with which evil poisons the healing springs of industry.

The low wages of large numbers of artisans in towns deprives them of the means of healthy life. Latterly wages have risen and they had the command of those means to a larger extent, but unfortunately the consumption of spirits and other stimulants absorbed their wages to the no small detriment of health. To sweep up the dusty and close workshops they are apt to be made draughty, so difficult is it to improve the health of artisans.

In the last twenty years the towns of England have increased from 580 to 938, their population from nine to fourteen millions, and the health of the whole population of the country has remained stationary.

Breeders reject weakly animals from their stock, and thus achieve success. By the care now taken of the humblest member of the human race the weakly, it is said, survive. They marry and propagate, and thus, as some contend, the proportion of inferior organizations is raised. The imbecile, the drunkard, the lunatic, the criminal, the idle, and all tainted natures were once allowed to perish in fields, asylums, or gaols, if they were not directly put to death, but these classes and their offspring now figure in large numbers in the population[14]. (Supplement to 35th Annual Report, pp. v – viii).

5. *Relative mortality of males and females at seven age periods in eight groups of districts. 1861 – 1870*. The following table [not reproduced] affords valuable evidence of the varying incidence of the effect of density population and the insanitary conditions upon males and females living at seven age periods, in various groups of districts in which the annual rate of mortality during the 10 years, 1861 – 1870, ranged from 15 to 39 per 1000. The rates prevailing at each age period in each sex are compared with the rates that ruled in the 53 districts to

show the relative excess at each age period, and of each sex, in the more healthy districts. Out of 2111 male children living under 5 years of age 100 die annually in the healthy districts 284 in the Manchester district and 349 in the district of Liverpool. (Supplement to 35th Annual Report, p. clxii.)

6. *Mortality of children* (aged 0 - 5 years), 1861 – 1870. The first thing to observe is, that the fatality children encounter is primarily due to the changes in themselves. Thus, 1,000,000 children just born are alive but some of them have been born prematurely. They are feeble, they are unfinished. The molecules and fibres of brain, muscle, bone are loosely strung together, the heart and the blood on which life depends have undergone a complete revolution. The lungs are only just called into play. The baby is helpless, for his food and all his wants he depends on others. It is not surprising then that a certain number of infants should die. But in England the actual deaths in the first year of age are 149,493 including premature births, deaths by debility and atrophy. Diseases of the nervous system, 30,637 and of the respiratory organs, 21,995. To convulsions, diarrhoea, pneumonia, bronchitis, their deaths are chiefly ascribed. Little is positively known and this implies little more than that the brain and spinal marrow, nerves, muscles, lungs, bowels fail to execute their functions with the exact rhythm of life.

The first two are said by pathologists to be often rather symptoms of diseases unknown than diseases in themselves. The total dying by miasmatic diseases is 31,266, but it is quite possible that several of the children dying of convulsions die in the early stages of some unrevealed zymotic disease whose symptoms have not had time for development. Convulsion is a frequent precursor in children of measles, whooping-cough, scarlet fever. Indeed, Dr. C. B. Radcliffe well remarks

In the fevers of infancy and early childhood, especially in the exanthematous forms of these disorders, convulsions not infrequently takes the place occupied by rigor in the fevers of youth and riper years[15].

Many of the cases of pneumonia may also in like manner be hooping-coughs and other latent zymotic diseases. In the second year of life pneumonia, bronchitis, and convulsions are still the prevalent and most fatal diseases. Many also die of measles, whooping-cough, scarlatina and diarrhoea. Scarlet fever[16] asserts its supremacy in the second, third, fourth and fifth years of life. Whooping-cough is at its maximum in the first year, measles in the second, scarlatina in the third and fourth. Thus these diseases take up their attacks on life in succession and follow it onwards.

The deaths from all causes under the age of five years are 263,182. The number ascribed to infanticide is very few but the death by suffocation (overlaying) etc. are more numerous, and so are the deaths directly referred to the *want of breast-milk*. The total to deaths by burns, injuries, drownings and all other kinds of violence, are 5175.

By a physiological law 511,745 boys are born in England to 488,255 girls, and by another law 141,387 boys and 121,795 girls die in the first five years of life. At the end of five years the original disparity in the numbers of the two sexes is so much reduced that at the age of five years the boys only slightly exceed the girls in number. The greater mortality of boys is due to difference of organisation, for the external conditions are substantially the same in which boys and girls are placed.

Great as is the influence of organisation itself, the difference of external circumstances and sanitary condition exercise a very real influence on life, disease and death in childhood. Thus, even in the healthy districts of the country , out o 1,000,000 born, 175,410 children die in the first five years of life, but in Liverpool district, which serves to represent the most unfavourable sanitary conditions, out of the same number born, 460,370, nearly half the number born, die in the five years following their birth. This is 284,960 in excess of the deaths in the healthy districts.

The above table [not reproduced] shows how many children die from the several groups of causes in the healthy districts; in all England; and in the Liverpool district. [Nineteen diseases were entered separately and among them cholera, cancer, scrofula and tabes mesenterica and hydrocephalus had occurred for the first time.]

There is a greater increase in Liverpool rom smallpox and measles than from scarlet fever; and diphtheria was more fatal in the healthy districts than in all England. Diarrhoea and cholera were greatly aggravated in the other districts of England, so were hooping-cough and fever under which were registered typhus, typhoid, infantile remittent and relapsing fever. The diseases of the lungs are more fatal to children in Liverpool than diseases of the brain.

The children of Norway fare better than the children of sunny Italy, to which it may well be still an *officina gentium*. Out of 10 children born alive the deaths in the first five years of life are in Norway 17, Denmark and Sweden 20, England 26, Belgium 27, France 29, Prussia 32, Holland 33, Austria and Spain, 36, Russia 38, Italy 39. Russia is almost as fatal to her children as Italy.

In a paper read before the [London] Statistical Society the methods of determining the rates of mortality were described, and I collected information ass to the treatment and management of children in Scotland, Norway, Sweden, France and Austria. The subject was taken up in England by the Obstetrical Society who published an able report based on returns, on the birth and treatment of English children[17]. I have not yet received papers from Russia or Italy.

The mortality of infants evidently depends to some extent on the midwifery of a country, on the way the children are fed by the mothers, on the water, and on the cleanliness observed as well as the other sanitary conditions. (Supplement to 35th Annual Report, pp. xxviii – xxx.)

[Appended was the beginning of Farr‘s undated letter to the Registrar General about mortality in 1861 – 1870. Its text essentially coincided with what had been included in the beginning of § 1.]

Notes

1. Difficult to understand.
2. Organisms responsible for infectious diseases.
3. Etruscans lived in ancient Italy from 900 BC to first century AD.
4. Their standard roughly equal to the potential lifetime of a person or of a renewal of a population.
5. See *Census Report* of 1851, vol. 1, p. xv. W. F.
6. There are many tables but which one was thus called?
7. Districts of London enjoying privileges granted to that city.
8. A contradiction.
9. Duvillard cites Dr. Jurin who prepared a table showing that out of 447 inoculated with effect 9 died. De Monro inoculated 5554 persons of which 72 died. Dr. Gregory set down the mortality at 3 in a 1000. [Also Duvillard?]. By natural smallpox the mortality per 1000 attacked ranged from 150 to 300. W. F.

I can refer to Duvillard (1806), *Analyse et tableau de l'influence de la petite vérole sur la mortalité à chaque âge et de cette qu'on préservatif* … Paris.

10. The mortality in the Smallpox Hospital was at the rate of 25% in 1748 – 1763 for 1634 of 6456 patients died. For later returns see Letter to Registrar General in Appendix to 34th *Annual Report*. W. F.
11. At first, the technique of vaccination was unknown. For example, how long had the vaccine remained usable?
12. In Paris, a Hospice des enfants-trouvés was established in 1688. In the mid-19th century the mortality of foundlings was extremely high and it was found out that they were first of all baptised and only then fed and left to sleep. This fact is certainly gushed up, but many years ago, I found this fact.
13. And what about the periods between those reigns?
14. The eugenics (Galton, Pearson) took up this touchy subject.
15. In Reynolds, *System of Medicine*, vol. 2, p. 593, article On diseases of spinal chord.
16. Farr applies the terms *scarlet fever* and *scarlatina*. I have not found any difference between them.
17. Mortality of children in the principal states of Europe. *J.* [*London*] *Stat. Soc.*, vol. 29, pp. 1 – 35. W. F.
18. Republished I the Appendix to the Registrar General's 34th *Annual Report*, pp. 225 – 229. W. F.

Farr (1807 – 1883) was an epidemiologist and a cofounder of medical statistics. His 35th Annual Report was published after 1870 although the pertinent data were applied in the paper.

VII

Inverse law of large numbers

Istoriko-Matematich. Issledovania, issue 14 (49), 2011, pp. 212 – 219

1. General Information

In 1713, there appeared the posthumous work of Jacob Bernoulli *Ars Conjectandi* (AC). Its fourth part contained his law of large numbers (LLN), a term due to Poisson, and I refer only to that part. In 1913, V. Ya. Uspensky translated it into Russian and the translation appeared the same year complete with Markov's Foreword. The second edition of that translation (Bernoulli 1986) was also supplemented by notes and commentaries by several authors including Prokhorov (1986).

In chapter 1, not quite formally, and without mentioning equal probabilities of cases, J. B. introduced the classical definition of probability (in spite of popular belief, due to De Moivre rather than Laplace). Nevertheless, in chapter 5 Bernoulli formulated his *Main Proposition*, i. e., the LLN, in terms of favourable and unfavourable cases, but once more mentioned probability in the last lines of the AC (see below in § 2).

He certainly had no time to complete his work. Thus, the title of part 4 included the lacking applications of the art of conjecturing *To civil, moral and economic affairs*. It is also possible that J. B. did not want to interrupt himself from his main aim, although this possibility does not reject my main conclusion.

The LLN can be described as the description of the stochastic convergence of the statistical probability $\hat{p}$ to the constant theoretical probability p, i. e., as, first, the proof that when the number of (independent) trials increases unboundedly, the limit of $\hat{p}$ will be p. And, second, the estimation of the rapidity of that convergence.

This estimation proved unfortunate. Markov (1900/1924, pp. 44 – 52) essentially improved it, whereas Pearson (1925) achieved even better results by applying the Stirling formula which J. B. had not known. Markov (pp. 104 – 115) repeated his study, this time applying that formula, but did not refer to his former study. This was an example of his disregard of readers. Other estimations followed but we may only mention Prokhorov (1986).

Pearson (1925, p. 202) passed over in silence the existence of the limit (see above) and made an unforgivable historical mistake by comparing J. B.'s theorem with the mistaken Ptolemaic system of the world. His son, Egon, edited his father's posthumous lectures of 1921 – 1993 (1978), and, on p. 230 stated that he had omitted the pertinent lecture which essentially repeated the paper of 1925. On the very first page of that book Pearson stated that

A most fundamental principle of statistics has been attributed to Bernoulli instead of its real discoverer, De Moivre.

2. The inverse theorem

Todhunter (1865, p. 73) noted that J. B. *himself proposed to employ it* [the LLN] *inversely*. And Pearson (1925, p. 205) remarked that, after proving his law, J. B. *turns around* and states that p must become ever nearer to $\hat{p}$, that is, will be estimated ever better. Actually J. B. only remarked in the last lines of his AC, that

If observations of all events be continued for the entire infinity (*with* [statistical!] *probability finally turning into complete certitude) it will be noticed that everything in the world is governed by precise laws and a constant law of changes …*

Markov (1914/1986, pp. 10 – 11) confirmed that J. B. had indeed thought bout the inverse problem but as an example he only cited by far not the most interesting urn problem.

Even in 1685 or 1686 J. B. (1975, pp. 46 – 47) considered posterior probabilities (yes, forestalling its definition of 1713, § 1), the possibility of a man to outlive another one. In a marginal note on p 46 he wrote out the bibliographic information of a review of Graunt's classical work of 1662. So J. B. knew about Graunt although possibly did not read his work.

Then, in 1713, in chapter 4, J. B. indeed considered the inverse problem. The entire chapter was in essence devoted to its qualitative description, partly in his final response to Lreibniz' objections which he voiced in his correspondence of 1703 – 1705 with Bernoulli.

De Moivre (Sheynin 2007a, p. 315) also thought that his limit theorem could be turned around and only Bayes and Price (Ibidem, p. 318 and Note 1) understood that the inverse problem ought to be studied separately (p. 316), and study they did.

3. A special case of the inverse problem: the non-existent theoretical probability

In some of his examples in chapter 4 Bernoulli considered the estimation of non-existing probabilities[1]. For mathematics, similar cases present no problems and statistics borrowed the expression of the theory of errors, *real value* of a measured constant, even such which did not exist in nature (Sheynin 2007b). Here are two examples.

Gauss (1816, §§ 3 and 4) looked for the real values of measures of precision, and Fisher (1922, pp. 309 – 310), after introducing fundamental notions about the properties of statistics, mentioned, on the very next page, real values of measures of precision.

It turned out that statisticians had not grasped the LLN, in the first place owing to the uncertainty connected with the notion of probability. Herschel (1817/1912, p. 579) illustrated a special case:

It may be presumed that any star promiscuously chosen … out of such a number [exceeding 14 thousand] *is not likely to differ much from a certain mean size of them all*.

He certainly had no data, whereas we now know that by their size stars monstrously differ from one another, probabilities are here irrelevant, *ex nihilo nihil*. They do not belong to a single totality which explains such differences.

4. Statisticians about the law of large numbers

Here, we may ignore the Poisson and Chebyshev forms of the LLN[2]. Even Quetelet who indeed rendered essential service to statistics, practically had not mentioned it. His notion of the average man and inclinations to marriage and crime were constant for an age group. He (1846, p. 216) mentioned that Poisson form only in connection with mean stature.

In the 1880s the continental direction of statistics was born (in Germany). Its initiator was Lexis. He (1877, pp. 15 – 18) acknowledged that equally possible cases can be imagined if the statistical probability indeed tended to a certain value and furthermore if the circumstances connected with the studied event sufficiently resembled the conditions of games of chance. I would say: he meant that the fulfilment of the necessary conditions for the LLN should be somehow checked.

But later that same Lexis (1886, p. 437) concluded that, because of the equally probable cases the theory of probability is subjectively justified. The same cases haunted him much later (1913, p. 2091).

That picture was not really pretty, but here is something worse. Maciejevski (1911, pp. 94 – 98) introduced a *statistical* LLN instead of the Bernoulli proposition that allegedly impeded the development of statistics. His own law qualitatively asserted that statistical indicators exhibited ever lesser fluctuations as the number of observations increased.

Bortkiewicz (1917, pp. 56 – 57), the chronologically second main author of the Continental direction, thought that the LLN ought to be understood only in the sense which it acquired in statistics, i. e., for denoting *a quite general and independent from any definite stochastic pattern* stability of the statistical indicators if the number of observations is large and their circumstances are only weakly variable.

And Romanovsky, in one of his earliest works (1912, p. 22), stated:

In the very beginning of the calculus of probability there must be a law on which all [its] *applications to reality rests. In all justice, this law can be called the LLN. It is independent from both the Bernoulli and Poisson theorem and is their basis. It reads: If a trial on which an … event having probability p can occur is repeated n times with n being sufficiently large, then this event must happen about np times.*

On p. 18 Romanovsky noted that the Bernoulli theorem forfeits its sense if sense is lacking in the notion of probability of an event in an isolated trial. Markov (1911/1981, p. 150) stated the same but Rmanovsky had thus undermined his own definition of probability. Later Romanovsky (1924, p. 15n) agreed with Bortkiewicz (see above) and, without referring to his former statement, decided that the LLN is tantamount *to many theorems on the probability in which a large number of trials is essential*. Finally, he (1961, p. 127) stressed the general scientific essence of the LLN and called it physical.

In a modern reference book (Prokhorov 1999) a few articles are devoted to the LLN. The first of them begins, on p. 60, with recognition of the LLN as *a general principle*. This means a recognition of its physical sense if not physical essence.

Mises statistically defined probability and thus essentially completed, in the logical sense, the thoughts of Bernoulli, De Moivre and Bayes. While describing his theory, Tutubalin (1977, p. 15) noted:

Kolmogorov's axiomatic is the one widely accepted. However, the concepts of practical applications largely follows the Mises idea.

Yes, workers on natural science can, and even are compelled to refer to Mises, how otherwise can they begin any study? But then, they should have been content with Bayes.

The doubts led Lexis (§ 4) to a formulation of a test for checking the equality of the probabilities of an event in different series of observations, to a study of *the stability of statistical series*. For a few decades this study remained the main subject of the theoretical work of Continental statisticians (and their attitude towards non-existing probabilities had appropriately changed). True, Chuprov (1918 – 1919) almost completely refuted the practical significance of the Lexian test, but as a result of those studies the statistical thought had essentially livened and some incidental results were achieved.

Let the studied event whose probability is supposed constant and equal p occur a_i times in series i. Then its variance can be calculated either by the Gauss non-parametric formula or by the formula only valid for the binomial distribution

$$\sigma = pqn,\ q = 1 - p \qquad (1)$$

where n is the number of trials.

According to the ratio of these estimates Lexis (1879, § 6) separated the stability of series of independent trials in two classes. It is needless to describe his thoughts in detail, see above, and I only notice that the application of formula (1) seems mistaken. It assumes that p is known whereas Lexis tested this assumption by issuing from statistical data.

To put it otherwise, it was necessary to issue from the formula for the variance in the Bayes problem and it is difficult to understand why Chuprov did not notice it. This formula can be seen in the German translation of the Bayes formula (Sheynin 2007a, § 2) or in Czuber (1903/1908, p. 186) whom Chuprov (1909/1959, p. 159) cited although in connection with other matters.

The inverse LLN, ass I called it, became for a long time an unnoticed item of the classical study of Bayes. The lack of comprehension of the difference between the versions of the LLN (understandable at those times) led Jacob Bernoulli and De Moivre to mistaken statements. A similar remark is valid for Lexis.

Notes

1. Thus, in one of his examples Bernoulli (chapter 4) noted that it is impossible to say how easier a disease (for example, plague) can kill a man than another one (for example, dropsy, or dropsy than fever).

2. The Poisson form of the LLN is interesting in another sense. He admitted that theoretical probabilities (in the plural) can be unknown. As an example, he (1837, p. 10) asserted that there existed a mean interval between the molecules of a body.

Bibliography

Шейнин О. Б., Sheynin O. (2007a, Russian), On the history of the Bayes theorem. *Istoriko-Matematicheskie Issledvania*, issue 12 (47), pp. 312 – 320.

--- (2007b), The true value of a measured constant and the theory of errors. *Historia Scientiarum*, vol. 17, pp. 38 – 48.

Bernoulli J. (manuscript, 1975), Meditationes. In author's *Werke*, Bd. 3. Basel. Partial publication.

--- (1986), *O zakone bolshikh chisel* (On the law of large numbers). Moscow. Contents: Author's *Ars Conjectandi*, pt. 4 and related material,

Bortkiewicz L. (1917), *Die Iterationen*. Berlin.

Chuprov A. A. (1909), *Ocherki po teorii statistiki* (Essays on the theory of statisics). Moscow, 1959.

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

Czuber E. (1903), *Wahrscheinlichkeitsrechnung und ihre Anwendungen* …, Bd. 1. 1908. New York, 1968.

Fisher R. A. (1922), On the mathematical foundations of theoretical statistics. *Phil. Trans. Roy. Soc*., vol. A222, pp. 309 – 368.

Gauss C. F. (1816), Bestimmung der Genauigkeit der Beobachtungen. In author's *Abhandlungen zur Methode der kleinsten Quadrate* (1887). Hrsg A. Börsch, P. Simon. Vaduz, 1998, pp. 129 – 138.

Herschel W. (1817), Astronomical observations and experiments … *Scient. Papers*, vol. 2, pp. 575 – 591. London, 1912.

Lexis W. (1879), *Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft*. Freiburg i/B.

--- (1886), Über die Wahrscheinlichkeitsrechnung und deren Anwendungen auf die Statistik. *Jahrbücher f. Nationalökonomie u. Statistik*, Bd. 13 (47), pp. 433 – 450.

--- (1913), Review of A. A. Kaufman (1913), *Theorie und Methoden der Statistik*. Tübingen. *Schmollers Jahrbuch f. Gesetzgebung, Verwaltung u. Volkswirtschaft in Deutschen Reiche*, Bd. 37, pp. 2089 – 2092.

Maciejewski C. (1911), *Nouvaux fondements de la théorie de la statistique*. Paris.

Markov A. A. (1911, Russian), On the basic principles of the calculus of pobability and on the LLN. In Ondar (1977/1981, pp. 149 – 153).

--- (1900), *Ischislenie veroiatnostei* (Calculus of probability). Later editions: 1908, 1913, 1924 (posthumous, Moscow). German translation, 1912.

--- (1914, Russian), Bicentennial of the LLN. In Bernoulli (1986, pp. 9 – 16).

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

Pearson K. (1925), James Bernoulli's theorem. *Biometrika*, vol. 17, pp. 201 – 210.

--- (1978), *History of Statistics in the 17th and 18th Centurie*s … Lectures 1921 – 1933. London, Editor, E. S. Pearson.

Poisson S.-D. (1837), *Recherches sur la probabilité des jugements* … Paris, 2003. **S, G,** 53.

Prokhorov Yu. V. (1986, Russian), The LLN and the estimation of the probabilities of large deviations. In Bernoulli (1986, pp. 116 – 150).

--- , Editor (1999), *Veroiatnost i matematicheskaia statistika. Enziklopedia* (Probability and math. statistics. Encyclopaedia). Moscow.

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

Romanovsky V. I. (1912), *Zakon bolshikh chisel i teorema Yakoba Bernoulli* (On the LLN and the Jacob Bernoulli theorem). Warsaw.

--- (1924, Russian), Theory of probability and statistics according to some newest works of Western scholars. *Vestnik Statistiki*, No. 4 – 6, pp. 1 – 38; No. 7 – 9, pp. 5 – 34

--- (1961), *Matematicheskaia statistika,* vol. 1. Tashkent.

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

Tutubalin V. N. (1977), *Granitsi* … (Boundaries of applicability. Methods of probability and statistics and their capability). Moscow. **S, G,** 108.

VIII

Vladimir Niksa

Die ökonometrische Bewegung und die Statistik

Stat. Vierteljahresschrift, Bd. 1, 1948, pp. 20 – 31

1. Allgemeine Bemerkungen

Am 29. Dezember 1930 wurde in Cleveland in den Vereinigte Staaten die *Econometric Society* gegründet. Auf Einladung von Prof. Irving Fisher, Ragnar Frisch[1] und Charles F. Roos versammelte sich eine Gruppe von Wirtschaftswissenschaftlern, Statistikern und Mathematikern mit der Ziele, eine neue internationale Vereinigung zur Förderung der Zusammenarbeit von Wirtschaftstheorie, Statistik und Mathematik ins Leben zu rufen[2]. Seit Jänner 1933 erscheint die Zeitschrift *Econometrica*, das Organ der Ökonometrischen Gesellschaft, wo Abhandlungen ökonometrischen Charakters veröffentlich werden. Die ökonometrische Bewegung zählt heute 800 Wissenschaftler, die auf der ganzen Erde verteilt sind. Deren Kern befindet sich jedoch in den Vereinigte Staaten von Amerika.

Die Ökonometrie befasst sich mit dem Studium der quantifizierbaren Seite der Wirtschaftstheorie, wobei mathematisch-statistische Denkprinzipien[1*] und Forschungsarten als Grundlage dienen. Sie ist weder mit der allgemeinen Wirtschaftstheorie, noch mit Mathematik, noch mit der Statistik identisch, sondern bildet eine gesunde Synthese dieser drei Standpunkte. Die Ökonometrie unterscheidet sich von der älteren mathematischen Wirtschaftslehre wesentlich: die mathematische Ökonomie hatte weder die geeigneten Anfangsprämissen, noch verifizierte sie ihre Lehrsätze, sondern war eine Abart der deduktiven Betrachtungsweise, die im Vergleich mit der ökonomischen Wirklichkeit notwendigerweise den Zusammenbruch erlitt.

Da erhält die Mathematik keinen Primät, wie bei den älteren Meistern, wo die mathematische Ökonomie manchmal in die ökonomische Mathematik überging, sondern ist nur eine Hilfsgerät. Die Klassische mathematische Ökonomie bekommt statistische Formen, wodurch sie dem faktischen Leben von allen bisherigen Methoden in den Wirtschaftswissenschaften am nächsten steht. Diese höchste Ausmaß an Objektivität verdankt die Ökonometrie den statistischen Forschungsarten.

Vierzehn komplette Jahrgänge der *Econometrica*, mit über 5500 Seiten dienten als Basis für die Bewertung und Würdigung der ökonometrischen Bewegung. Es ist kaum einen Problem der quantitativen Ökonomie zu begegnen, welches nicht ökonometrisch bearbeitet wurde. Dabei wurden die älteren Probleme ins richtige Licht gestellt, zahlreiche neue aufgeworfen und gelöst, insbesondere die Forschungsarten quantitativ erweitert und quantitativ verfeinert.

Um das *Wirtschaftsgesetz* und seine logische Natur drehen sich sämtliche methodologischen und philosophischen Fragen der Wirtschaftstheorie. Dabei ist das Methodenproblem sehr oft

Selbstzweck geworden, statt als Werkzeug für das Eindringen on die komplizierte Gestalt des Wirtschaftslebens zu dienen. Sämtliche *Methodenschulen* in der ökonomischen Theorie mussten Schiffbruch erleiden, weil zwischen den durch ihre Methode gewonnenen Gesetzen, Idealtypen, Entwicklungstendenzen und der Wirklichkeit eine unüberbrückbare Kluft lag.

Es ist klar, dass weder die naturwissenschaftliche Betrachtung der Wirtschaftslehre, noch ein relativistischer Historismus weder ein dialektischer Materialismus noch ein psychologischer Subjektivismus imstande sind, als das einzige Erkenntnismittel der Ökonomie zu bestreben. Meiner Auffassung nach wurde der faktische Stoff der Ökonomie zu wenig beachtet. Weder Induktion noch Deduktion können die Massenerscheinungen variablen Charakters – mit denen das ganze Gebiet des Wirtschaftslebens durchsetzt ist – restlos erklären.

Diese quantitativen Erscheinungen wurden seitens der mathematischen Richtung schon früh bemerkt, die auf Grund des Vorhandenseins der quantitativen Vorgänge Mathematik auf die Ökonomie applizierte, im Glauben, reine exakte Ökonomie aufzustellen. Diese quantitative Wirtschaftslehre wies gewisse Vorteile, sowie gewisse Nachteile gegenüber der amathematischen, qualitativen Ökonomie auf. Ihr Hauptverdienst liegt darin, einen Wink für die richtige Beurteilung der Methodenfrage gegeben zu haben. Ihre Funktionalität und ihr deduktiv-spekulativer Charakter in reiner mathematischer Form gibt jedoch einen zu strengen Maßstab. Ihre Exaktheit grenzt oft an die Übertriebenheit der deduktiv aufgebauten Methoden des Klassizismus und Psychologismus.

Jede Methode darf nicht so sehr vorausgesetzte Ziele al den Stoff der zu untersuchenden Erscheinungen berücksichtigen. Da die ökonomischen Massenphänomena keine typischen, sondern variable Erscheinungen darstellen, kommt eben die Statistik als beste und sicherste Forschungsart in Betracht. Indem die Statistik in ihren Untersuchungen stark nach der Mathematik greift – die ältere quantitative Schule baute überhaupt auf der Mathematik – wird das mathematische Rüstung in jeder quantitativen Methode der Ökonomie unerlässlich. Auf dem Grundgedanken einer mathematisch-statistischen Betrachtung zwecks Vertiefung der Wirtschaftstheorie beruht die ökonometrische Bewegung, die wahrhaftig die beste Methode für die Erforschung der Wirtschaftsvorgänge und Aufbau einer Wirtschaftstheorie darbietet.

So bedeutet Ökonometrie keinesfalls eine Neuauflage der älteren mathematischen Schule. Die Fehler der Altmeister sind erkannt, Mathematik ist kein Prius [Latin: Not the most important], sondern nur ein – wenn auch wertvolles – so doch nur ein Hilfsgerät. In der Verfassung der Ökonometrischen Gesellschaft ist das Ziel – Förderung der Wirtschaftstheorie in Beziehung zur Statistik und Mathematik – klar festgelegt.

Erkenntnistheoretische Abhandlungen von Ragnar Frisch[3], Irving Fisher[4], J. A. Schumpeter[5] und Haavelmo[6] sind neben der Satzung der Gesellschaft die Quelle für die Beurteilung der Ziele und Aufgaben der Ökonometrie.

Ökonometrie ist weder mit der Wirtschaftsstatistik, noch mit der Anwendung der Mathematik auf die Wirtschaftslehre noch mit der allgemeinen ökonomischen Theorie identisch. Sie bedeutet keine Sekte, keine Schule, sie beabsichtigt durchaus nicht einen Methodenstreit zu eröffnen oder die Bedeutung und Notwendigkeit der qualitativen Ökonomie zu vermindern oder sogar zu leugnen. Aber das Vorhandensein des Quantitativen in der Wirtschaftslehre berechtigt die Anwendung der Statistik und Mathematik[7].

Man könnte einwenden, dass die Statistik schon von einer Richtung besonders hochgehoben wurde, sich aber nicht besonders behaupten konnte (historischen Schule[2*]). Einerseits bekommt das Statistik im Rahmen der Ökonometrie eine ganz andere Rolle als in der historische Schule zugewiesen, anderseits ist die ontologisch-staatenkundliche, amathematische Statistik im Sinne Roschers wesensverschieden von der Statistik im modernen Sinne. Statistik hat von dieser Epoche bis zum heutigen Stand eine ungeheure Entfaltung erlebt. Wahrscheinlichkeitstheoretisches Denken, Kurvenanalyse, Indexzifferntheorien, Korrelations-, Trendberechnung und andere Errungenschaften der modernen Statistik existierten damals noch nicht. Was der Historismus als *Statistik* benützte, wird heutzutage als *Statistik* abgelehnt.

Zwischen der Entwicklung der Statistik und der der Wirtschaftslehre bestehen viele Berührungspunkte. Zur Zeit der Universitätsstatistik konnte man kaum eine Disziplin von der anderen unterscheiden, *Staatsmerkwürdigkeiten* waren ein mixtum compositum aus Statistik, Ökonomie, Soziologie, Finanzehre und anderen Staatswissenschaften[3*]. Die politische Arithmetik, im Unterschied zur Achenwall'schen Richtung ein mathematisch ausgeprägter Zweig der historischen Statistik, war eng mit den ökonomischen Problemen verknüpft.

In der Evolution der beiden Wissenschaften lassen sich viele Parallelen erkennen. Die allgemeinen, herrschenden Ideen einer Epoche fanden in den beiden Wissenschaften ein adäquates Echo. Als ökonomischer Ausdruck der Staatsidee im aufgeklärten Absolutismus kommt Merkantilismus und Kameralistik [accountancy]. In der Statistik herrschte die Göttinger Schule. Der Siegeszug der Naturwissenschaften kulminierte auf dem Gebiete der Wirtschaftstheorie im Physiokratismus und der Klassik, die statistische Ausdrucksform ist die Sozialphysik Quetelets und die Physiologie der Gesellschaft Knies'. Dem ökonomischen Historismus entspricht der Umwandlungsprozess der staatenkundlich orientierten in die soziologisch gerichtete Statistik Georg v. Mayrs. Die religiös-ethische Richtung ist in beiden Disziplinen vertreten.

Die Spaltung der modernen Statistik in eine logische und mathematische Richtung ist nur die Fortsetzung der alten Antithese Achenwall – Petty, die sich auch regional mit den Einflusszonen der Universitätsstatistik und der politischen Arithmetik deckt. Die Dogmengeschichte lehrt übrigens, wie die kontinentalen europäischen Vertreter der mathematischen Ökonomie (Cournot, Dupuit, Gossen, Walras) es viel schwieriger hatten als ihre englischen Kollegen (Jevons, Edgeworth, Marshall). Obwohl der Kontinent einen Cournot, Walras und Pareto hervorbrachte, zeigt die geographische Verteilung der Vertreter der Ökonometrie, dass die anglo-amerikanische Welt weit voran ist, während Mitteleuropa – die Wiege der Göttinger Schule – der mathematischen Ökonomie schwer zugänglich bleibt und als eine Burg der *logisch* orientierten Statistik ihren mathematischen Denkformen widerstrebt.

2. Relativität der statischen Erkennismöglichkeiten

Mit Rücksicht auf die Zusammensetzung der quantitativen Erscheinungen auf dem Gebiete der Wirtschaftswissenschaften, auf die losere Art der Kausalzusammenhänge, auf das Wirken des Zufalls und der Wahrscheinlichkeit, auf die Stochastizität der Verbindungen, lässt sich auch unter Mitwirkung der Statistik keine präzise und funktionelle, sondern nur eine relative, wahrscheinliche Grundstruktur, die Wesensform, ermitteln. Obwohl diese Erkenntnisse, von Regelmäßigkeiten nur einen relativen Wert besitzen, sind sie absolut genommen der höchste Grad einer positiven Eindringung in die komplizierte Welt der Sozialerscheinungen, Statistik ist das weitestgehende Mittel in der Erkenntnis der ökonomischen Massenerscheinungen,.

Sämtliche deduktiv arbeitende Schulen begingen einen schweren Fehler, indem sie künstliche Abstraktionen, leicht verständliche Schemen, zu *Wirtschaftsgesetzen* erhoben haben. Für uns sind das Pareto'sche Gesetz der Einkommensverteilung[8], die statistisch bewiesenen regelmäßigen Zusammenhänge zwischen Wettergestaltung und Erntezyklen (Moore[9], Beveridge[10]), oder die allgemeinen Gesetzmäßigkeiten einer säkularen Entwicklung in der Preisbewegung der industriellen und landwirtschaftlichen Güter[10] viel wertvoller als sämtliche unbewiesene deduktive Spekulationen.

Die gesamte Gebäude der Statistik beruht auf der Wahrscheinlichkeit: die Zusammensetzung einer Masse, die Festlegung einer Mindestgröße der statistischen Masse für die Zulassung des Gesetzes der großen Zahl[en], die Bestimmung von statistischen Maßzahlen, die statistische Ursachenforschung, sowie auch die dadurch ermittelten Tatbestände und Regelmäßigkeiten[11].

3. Geschichte der ökonometrischen Bewegung

3.1. Bis zur Gründung der ökonometrischen Gesellschaft. Die Geschichte der Ökonometrie ist noch nicht geschrieben worden. Dieses Kapitel verfolgt auch keinesfalls die Absicht, eine solche zu geben, sondern nur die Ideen, die zur Bildung der gegenwärtigen ökonometrischen Bewegung führten, zusammenbefasst darzustellen.

Dem Wesen nach liegen ihre Wurzeln in der ökonomischen Theorie, Statistik und Mathematik. Jedoch sind in einem historischen Rückblick nur diejenigen Ideen von Bedeutung, die sich mit der quantitativen Studie der Wirtschaftstheorie befassten, vor allem die mathematische Ökonomie und die mathematische Statistik.

Zuerst begann die quantitative Wirtschaftstheorie mit der Darstellung einzelner ökonomischer Vorgänge in den Formen der rudimentären Mathematik oder in inadäquater Übertragung der mathematischen Gedanken auf das ökonomische Gebiet. Cournot, Dupuit, Gossen, Jevons, Walras, Edgeworth, Pareto, Moore und Schumpeter schufen die Fundamente der quantitativen

Wirtschaftslehre, an die sich eine lange Reihe neuerer Vertreter mit neuen Ideen und Anschauungen anschließt. Die quantitative Wirtschaftstheorie begann mit der Statik, verbreitete sich später auf die Dynamik. Die ursprünglich stark deduktive mathematische Ökonomie hat ihre Vervollkommnung in der mathematisch-statistischen Ökonometrie gefunden.

1. Fisher hat die Bibliographie der mathematischen Wirtschaftslehre in vier Abschnitte geteilt:

Von Ceva bis Cournot (1711 – 1837)

Von Cournot bis Jevons (1838 – 1870)

Von Jevons bis Marshall (1871 – 1889)

Von Marshall bis ersten amerikanischen Ausgabe Cournots (1890 – 1897).

Wie schon gegen Ende des vorigen Jahrhunderts die mathematische Wirtschaftslehre stark entwickelt war, ist aus der Zahl der Werke ersichtlich.

Erste Periode dauerte 127 Jahre mit 27 Werke, d. i. auf 1 Werk 4.30 Jahre. **Zweite Periode** dauerte 33 Jahre, d. i. auf 1 Werk 9 Monate. **Dritte Periode** dauerte 19 Jahre mit 114 Werke, d. i. auf 1 Werk 2 Monate. **Fierte Periode** dauerte 8 Jahre mit 142 Werke, d. i. auf 1 Werk 3 Wochen.

Es ist unendlich schwer, die komplette Liste der Männer, die die quantitative Ökonomie vorwärts gebracht haben, insbesondere nach der Jahrhundertwende bis zur Gründung der *Econometric Society*, zu erfassen. Es gibt eine große Anzahl von Wissenschaftlern, die ihr Lebenswerk noch keinesfalls beendet haben, die fortwährend neue Probleme aufwerfen, neue Erkenntniswege bahnen, neue Forschungsarten schaffen. Seit dem Jahre 1897 ist die mathematisch-ökonomische Literatur stark angewachsen, die Zahl der Schriften ist allzu groß geworden, so dass man nur einige wirkliche Pioniere behandeln kann. Deswegen sollen seit der Jahrhundertwende mehr die Hauptrichtlinien und deren Spitzenträger erfasst werden, statt eine vollständige Bibliographie anzuführen, weil auch die Fälle des Stoffes jede diesbezügliche Vollkommenheit hindert.

Die Darstellung der Geschichte der Ökonometrie fällt besonders schwer. Soll man starke Persönlichkeiten oder allgemeine Ideen mehr berücksichtigen? I. Fisher berücksichtigte mehr, teilweise einseitig, starke Bahnbrecher, seine Perioden sind auch ungleichmäßig. Unsere Wendepunkte sind Cournot, Jevons – Walras, I. Fisher – Pearson, die Eröffnung des Harvard Institutes und die Gründung der Econometric Society, wobei das Erscheinen von Cournots *Recherches* (1838) und die Bildung der Ökonometrischen Gesellschaft (1930) die zwei wichtigsten Ereignisse bedeuten.

Mit wenigen Ausnahmen blieb Mitteleuropa für die quantitative Ökonomie verschlossen, vor allem das deutschsprachige Gebiet (Ausnahmen: Schneider, O. Anderson[4*], Weinberger). Es ist kein Zufall, dass Schumpeter, Marschak, Wald, Stachle, Haberler und Tintner – heutzutage führende Wissenschaftler auf dem Gebiet der quantitativen Ökonomie und der Statistik, Fellows der Ökonometrischen Gesellschaft – das mitteleuropäische Milieu verließen. Ökonometrie fand eine viel bessere Aufnahme in den

Niederlanden Jan Tinberger), in Polen (Zawadski, Wisniewski, Lange, Kalecki), besonders auch in den skandinavischen Ländern (Ragnar Frisch, einer der bedeutendsten Gründer der Ökonometrie, Haavelmo, Koopmans, Zeuthen, Wold, Ohlin, Myrdal.

3.2. Die Gründung der Ökonometrischen Gesellschaft. Jeder Anfang ist schwer. Die Vorgeschichte der Ökonometrischen Gesellschaft ist mit dem Namen Irving Fishers eng verbunden, der schon in Jahre 1912 vergeblich versuchte, in Rahmen der *American Association for the Advancement of Science* eine Arbeitsgemeinschaft, zusammengesetzt aus Wirtschaftswissenschaftlern, Statistikern und Mathematikern, ins Leben zu rufen, um die quantitative Wirtschaftswissenschaft zu fördern. Auch im Jahre 1920 scheiterten seine darauf abzielenden Anstrengungen.

Im Jahre 1930 diskutierten Prof. Ragnar Frisch und Dr. Charles F. Roos mit Prof. Irving Fisher die Möglichkeit der Gründung einer internationalen Gesellschaft. Fisher war zuerst – mit Rücksicht auf seine ergebnislos gebliebene Pioniertätigkeit in den Jahren 1912 und 1920 – ziemlich reserviert, jedoch gewährte er Frisch und Roos volle Unterstützung, die darauf eine rege Tätigkeit entwickelten, um den Weg für die Ökonometrische Gesellschaft zu bahnen.

Am 29. Dezember 1930 zu Cleveland (Ohio) kam die Gründung der *Econometric Society* zustande. Zum ersten Präsidenten wurde der Professor der Yale University Irving Fisher einstimmig erwählt. Die Mitglieder des ersten Rates waren Luigi Amoroso (Rom), Ladislaus von Bortkiewicz Berlin, gestorben August 1931), A. L. Bowley (London School of Economics), Francois Divisia (Ecole Nationale des Pontes et Chaussées, Paris), Ragnar Frisch (Universität Oslo), Charles Roos (Smithsonian Instn, Washington), Joseph Schumpeter (Bonn) und Wl. Zawadski (Universität Wilna).

Für die weitere Entwicklung der Ökonometrie ist von allergrößten Bedeutung die Finanzierung einer neuen Zeitschrift durch Alfred Cowles, als Organ der Gesellschaft. Cowles, Sekretär und Schatzmeister der *Econometric Society*, gründete auch die *Cowles Commission for Research in Economics*, die mit der University of Chicago affiliert ist und in den nächsten Beziehungen zur Ökonometrie steht. Seine Weitsichtigkeit, Großzügigkeit und Generosität ermöglichten es wesentlich, dass die Ökonometrie eine so erfolgreiche Entwicklung nehmen konnte.

Mit der wissenschaftlichen Leitung der neuen Zeitschrift, die den Namen *Econometrica* erhielt, wurde Prof. Frisch betraut. *Econometrica* ist das Organ der Ökonometrischen Gesellschaft, sie erscheint seit Jänner 1933 vierteljährlich. *Econometrica* bringt wissenschaftliche Abhandlungen ökonometrischen Charakters, Berichte von den Versammlungen der *Econometric Society* und andere Mitteilungen, die sich auf die Ökonometrie beziehen. Keine Abhandlung wird seitens der *Econometrica* wegen ihres zu mathematischen Inhaltes abgewiesen, anderseits erscheinen in *Econometrica* auch zahlreiche Aufsätze nichtmathematischen Charakters.

Die Zeitschrift ist jedoch kein Magazin für die ökonometrische Literatur, sondern nur ein Art *Clearing House*, wo nur die besten, die erlesensten, die bedeutendsten Forschungen auf dem Gebiete der

Ökonometrie Platz finden[12]. Jede Meinung darf man offen vertreten, jedoch ist der Verfasser für seine Theorie und Anschauung persönlich verantwortlich. In den Zeilen der *Econometrica* herrscht volle Freiheit des Gedankens, Diskussion und Kritik sind willkommen.

Gegenwärtig befindet sich *Econometrica* im fünfzehnten Jahrgang. Bis jetzt sind vierzehn Jahrgänge erschienen, sogar auch in den Kriegsjahren in etwas verminderten Umfang (1942 – 1944). Herausgeber ist Ragnar Frisch.

4. Einiges über die Leistungen der Ökonometrie

4.1. Die Grundlagen der Ökonometrie, Das Wesen sowie den erkenntnistheoretischen Ausgangspunkt der Ökonometrie haben Ragnar Frisch, Joseph Schumpeter und Irving Fisher dargestellt[13, 14, 15].

Die stochastische Natur der Ökonometrie hat Trygve Haavelmo[16] geklärt. Er gab die theoretische Grundlage für die Analyse der Beziehungen zwischen ökonomischen Variablen, die er auf der Wahrscheinlichkeitstheorie und Statistik fußen lässt. Hagstroem bringt die Unvollkommenheiten der klassischen Theorien vor, betont die Notwendigkeit der Einführung der Wahrscheinlichkeitselemente in die Ökonomie und einer nachträglichen statistischen Verifikation[17]. Das Problem der Differenzen- und Differenzialgleichungen in der Ökonometrie haben besonders ausführlich Frisch-Holme, James & Belz[18, 18a] bearbeitet.

Die stochastische Natur des dynamischen Wirtschaftssystems fand ihre Erklärung bei Hurwicz[19], welcher das Problem der Differenzialgleichungen richtigstellte. Die Frage der Abschätzung der linearen stochastischen Differenzgleichungen auf statistischer Grundlage wurde von Mann & Wald[20] in das richtige Licht gestellt. Lange[21] und Samuelson[22] entwickelten Multiplier-Theorien. Von Bolza[23] stammt die mathematische und physikalische Formulierung der Invarianten, die für die wirtschaftlichen Variablen anwendbar sind, die ihren Wert trotz Transformation der Koordinaten behalten, wobei er sich auf die Arbeiten von Joule und Einstein stützt. Selbstverständlich ist die Parallele der Physik nicht bis zur letzten Konsequenz durchführbar (Le Corbeiller, Mayer, Knight[24, 25, 26]).

Die Flucht in die reine Physik und Mathematikbedeutet keinesfalls eine Rettung für diejenigen ökonomischen Probleme, die sich mit Hilfe der logischen und statistischen Methoden nicht einwandfrei formulieren lassen.

Einen guten Überblick der Entwicklung der statischen und dynamischen Auffassung der Wirtschaftstheorie gab Jan Tinbergen[27] über die Probleme der Zeit und des wirtschaftlichen Gleichgewichts handelte Akermann[28]. Von Divisia stammt ein Vorschlag, wonach man in der Ökonomie die Konstanz (stock) und Bewegung (flux) gleich de Prinzipien der Bevölkerungsstatistik einführen sollte. Dauer der konstanten und fluiden Elemente wäre zu bestimmen, sowie auch der erneuernden Faktoren, die der Natalität und Mortalität der Demographie entsprechen[29].

4.2. Beiträge zur Geschichte der Ökonometrie. In *Econometrica* wurden auch die Probleme, die sich auf die Geschichte der quantitativen Wirtschaftslehre beziehen, behandelt. So gaben Fisher[30] und Schumpeter[31] in kurzen Umrissen die allgemeine Entwicklung der älteren mathematischen Wirtschaftslehre und Ökonometrie. Die Klassiker der Ökonometrie wurden besonders beachtet. Gustavo die Veccio beleuchtete Francesco Fuoco, einen Representanten der älteren italienischen mathematischen Schule, der mit dem Popularisator des Smithianismus auf dem Kontinent J. B. Say kämpfte[32]. Erich Schneider widmete eine größere Abhandlung Johann Heinrich Thunen[33].

Cournot, der bedeutendste Bahnbrecher der quantitativen Wirtschaftslehre, fand gebührende Würdigung; René Roy[34] beleuchtete die Bedeutung Cournots für die mathematische Ökonomie sowie auch seine ökonomische Lehre[35], Irving Fisher[36] gab einen Rückblick seit der Übersetzung Cournots ins Englische und die Entwicklung der Ökonometrie, wobei er die Verdienste Cournots sowie die gegen ihn vorgebrachte Einwendungen sine ira et studio [absolutely impartially] darstellte.

Nichol[37] schildert die Tragödien im Leben Cournots, der, obwohl ein bedeutender Wirtschafts-wissenschaftler, Mathematiker, Wahrscheinlichkeits-theoretiker und Philosoph, bei seinen Zeitgenossen kein Verständnis fand[5*]. Hicks[38] gab eine eine umfangreiche Biographie Léon Walras, Marget[39] kommentierte die Neuauflage seiner Werke, *Econometrica* veröffentlichte die Walrassche Korrespondenz mit Jevons und Cournot, mit einer Bemerkung von Antonelli[40]. Die Ökonometrische Gesellschaft wendete sich mit einer besonderen Adresse[41] an die Universität zu Lausanne, wo der große Wissenschaftler lehrte.

Winifred & Stanley H. Jevons[42] veröffentlichten eine umfangreiche Studie über das Leben und die Werke ihres Vorfahren, A. L. Bowley[43] gedachte Edgeworth, auch eine kleine Notiz über Marshall[44] gehört hierher. Amaroso[45] publizierte die bisher beste Studie über Pareto, Neisser[46] analysierte seine Theorie der Produktion, Millikan und Travaglini[47, 48] beleuchteten Pareto als Soziologen. Aus der Feder Akermanns[49] stammt eine Studie über Knut Wicksell. Das Lebenswerk Colsons wurde von seinem Landsman Roy[50] dargestellt. Die Bedeutung des tragisch geschiedenen Henry Schulz würdigen Hotelling und Douglas[51, 52]. Obwohl nicht unmittelbar Biographien, dienen doch zur Beleuchtung und Lehre von wichtigen Verfassern einige Aufsätze: von Garver[53], der das Steuerproblem Edgeworths analysierte, sowie auch die kritischen Bemerkungen Hotellings[54].

Ragnar Frisch[55] widmete sich dem Steuertheorem Dupuits, Duncan[56] hat Marshalls Paradox in Beziehung auf die Richtung der Verschiebung in der Nachfrage dargestellt. Geiringer[57] gab eine neue Erklärung für die anormale Dispersion Lexis. Die Lehre Keynes' wurden von Samuelson, Hicks und Kaldor[58, 59, 60] kritisiert. In einer analytischen Studie Kaldors[61] über die Entwicklung der Theorien des Kapitals, in der Abhandlung Marschaks[62] über Identität und Stabilität in der Wirtschaftstheorie, sowie in den kritischen Bemerkungen Tinbergens[63] zu den Konjunkturtheorien gibt es viel wertvolles geschichtliches Material, ebenso wie in den jährlichen Übersichten aus Wirtschaftstheorie, Konjunkturforschung und Statistik, die oft weit in die Vergangenheit greifen.

References

Lacking place of journal publications is *Econometrica.* Notes 1* … are my own

1. Ragnar Frisch. Editorial, vol. 1, pp. 2ff, 1933.
2. Constitution of the Econometric Society, vol. 1, No. 1, pp. 106ff, Article 1, 1933.
3. Ragnar Frisch. Editorial, vol. 2, No. 1, p. 99, 1933.
4. Irving Fisher, Mathematical methods in the social sciences, vol. 9, No. 3 – 4, pp. 185ff, 1941.
5. J. A. Schumpeter, The common sense of econometrics, vol. 1, No. 1, pp. 5ff, 1941.
6. Trygve Haavelmo, The probability approach in econometric, vol. 12, Supplement, pp. 1ff., July 1944.
7. J. A. Schumpeter, see [5].
8. V. Pareto, *Course d'économie politique*. Lausanne, 1897. *Manuele d'economia politica*, p. 372. Milano, 1906. A. L. Bowley, The nature and purpose of the measurement of social phenomena, 2nd edition, pp. 106ff, 198ff. London, 1923. Bresciani-Turroni, Pareto's law and the index of inequality of incomes, vol. 7, pp. 133ff, 1939. W. Winkler, Einkommen. *Handwörterbuch der Staatswissenschaft*[*en*?], 4th edition, Bd. 3, pp. 255ff, 396ff.
9. H. L. Moore, *Economic cycles. Their law and causes*, pp. 119ff. New York, 1914. Sir William Beveridge, Wheat prices and rainfall in Western Europe. *J. Roy. Stat. Soc*., vol. 85, pp. 434ff, May 1922.
10. N. D. Kondratieff, Die Preisdynamic der industriellen und landwirtschaftlichen Waren. *Arch. f. Sozialwissenschaft u. Sozialpolitik*, Bd. 60, pp. 1ff.
11. L. v. Bortkiewicz, Realismus und Formalismus in der mathematischen statistik. *Allgem. Stat. Archiv*, Bd. 9, pp. 225ff, 1915. [That was his answer to a criticism of his alleged law of small numbers. He also formulated his answer in a posthumous paper translated in *Annals of math. stat*., vol. 2, pp. 1 – 22, 1931.] A. L. Bowley, *Elements of statistics*, 5th edition, p. 225. E. Czuber, *Die philosophischen Grundlagen der Wahrscheinlichkeitsrechnung*. Wien, 1935. F. Y. Edgeworth, On the application of the calculus of probabilities to statistics. *Bull. Inst. Intern. Stat*., 13 May 1909, pp. 506ff. [Or, the author's *Writings in prob., stat. and economics*, vol. 2, pp. 1 – 22, 1996. Editor Ch. R. McCann Jr, Cheltenham, UK, Brookfield, USA.] E. Kamke, *Einführung in die Wahrscheinlichkeitstheorie*. Leipzig, 1932. J. v. Kries, *Prinzipien der Wahrscheinlichkeitsrechnung.* Tübingen, 1927, pp. 114ff. Initially published in 1886. R. v. Mises, *Probability, statistics and truth*. New York, 1981. Initially published in German, in 1928. W. Winkler, *Grundriss der Statistik*, pp. 37ff. Berlin, 1931.
12. R. Frisch, Editorial., pp.1ff, 1933.
13. Ragnar Frisch, Editorial, vol. 1, No. 1, pp. 1ff, 1933. Note on the term econometrics, vol. 4, No: 1, p. 95. The responsibility of the econometrician. vol. 14, No. 1, pp. 1ff, 1946.
14. A. J. Schumpeter, The common sense of the econometrics, vol. 1, No. 1, pp. 1ff, 1933.
15. Irving Fisher, Mathematical method in the social science. vol. 9, No. 3 – 4, pp. 185ff.
16. Trygve Haavelmo, The probability approach in econometrics. E, Bd. 12, pp. 1ff, Supplement, July 1944. [Abbreviation E not explained, ought to be *Econometrica*, vol., but in later sources, only *Econometrica.*]
17. K. G. Hagstroem, Pure economics as a stochastic theory, vol. 6, No. 1, pp. 40ff, 1938.
18. Ragnar Frisch-Holme, The characteristic solution of a mixed difference and differential equation occurring in economic dynamics, 3rd edition, No. 2, pp. 225ff, 1935. [A book?]

18a. R. W. James & M. H. Belz, The significance of the characteristic solutions of mixed difference and differential equations, 6th edition, No. 4, pp. 326ff, 1938,

19. Leonid Hurwicz, Stochastic models of economic fluctuations, vol. 12, No. 2, pp. 148ff, 1944.

20. H. B. Mann & A. Wald, On the statistical treatment of linear stochastic difference equations, 11th edition, No. 3 – 4, pp. 173ff, 1943. [Second edition? A book?]
21. Oscar Lange, The theory of multiplier, E 11, No. 3 – 4, pp. 227ff, 1943.
22. A. Paul Samuelson, Fundamental multiplier identity, E 11, No. 3 – 4, pp. 221ff, 1943.
23. Hans Bolza, The conception of invariants in dynamic economics, E 4, No. 1, pp. 80ff, 1936.
24. Le Ph. Corbeiller, Les systèmes autoentretenus et le oscillations de relaxation, E, vol. 1, No. 1, pp. 328ff, 1933.
25. Joseph Mayer, Pseudo-scientific method in economics, E, vol. 1, No 4, pp. 418ff, 1933. Rejoinder, Там же, pp. 430ff.
26. H. Frank Knight, Reply, E, vol. 1, No. 4, pp. 428ff, 1933.
27. J. Tinbergen, Annual survey of significant development in general economic theory, E, vol. 2, No. 1, pp. 13ff, 1934.
28. Johann Akermann, Annual survey of economic theory: the setting of central problem. E, vol. 4, No. 2, pp. 97ff, 1936.
29. Francois Divisia, L'utilité d'une théorie générale des ensembles renouvelées, E, vol. 1, No. 3, pp. 325ff, 1933.
30. Irving Fisher, see [15].
31. J. Schumpeter, see [14].
32. Gustavo del Veccio, Francesco Fuoco, opponent of J. B. Say on the use of algebra in political economy, E, vol. 1, No. 2, pp. 220ff, 1933.
33. Erich Schneider, Johann Heinrich von Thünen, E, vol. 2, No. 1, pp. 1ff, 1934.
34. René Roy, Cournot et l'école mathématique, E, vol. 1, No. 1, pp. 13ff, 1933.
35. Он же, L'oeuvre économique d'Augustin Cournot, E, vol. 7, No. 2, pp. 131ff, 1939.
36. Irving Fisher, Cournot forty years, E, vol. 6, No. 3, pp. 198ff, 1938.
37. A. J. Nicol, Tragedies in the life of Cournot, E, vol. 6, No, 3, pp. 193ff, 1938.
38. J. R. Hicks, Léon Walras, E, vol. 2, No. 4, pp. 338ff, 1934.
39. W. Arthur Marget, Note on the new edition of the works of Léon Walras, E, vol. 5, No. 1, pp. 103ff, 1937.
40. Léon Walras et sa correspondance avec Augustin Cournot et Stanley W. Jevons, avec une note d'Etienne Antonelli, E, vol. 3, No. 1, pp. 119ff, 1935.
41. Address of the Econometric Society to the University of Lausanne, E, vol. 2, No. 4, p. 337, 1934.
42. Jevons H. Winefrid & Jevons H. Stanley, Wiliam Stanley Jevons, E, vol. 2, No. 3, pp. 225ff, 1934.
43. A. L. Bowley, Francis Ysidro Edgeworth, E, vol. 2, No. 2, pp. 113ff, 1934.
44. Alfred Marshall: the mathematician as seen by himself, E, vol. 1, No. 2, pp. 221ff, 1933.
45. Luigi Amoroso, Wilfredo Pareto, E, vol. 6, No. 1, pp. 1ff, 1938.
46. H. Neisser, Note on Pareto's theory of production, E, vol. 8, No. 3, pp. 253ff, 1940.
47. Max Milligan, Pareto's sociology, E, vol. 4, No. 4, pp. 324ff, 1936.
48. Volrico Travaglini, comments on Milligan's review of Pareto's sociology, E, vol. 5, No. 3, pp. 301ff, 1937.
49. Johann Akermann, Knut Wiksell, a pioneer of econometrics, E, vol. 1, No. 2, pp. 113ff, 1933.
50. René Roy, Clement Colson, E, vol. 8, No. 3, pp. 193ff, 1940.
51. Harold Hottelling, The work of Henry Schulz, E, vol. 7, No. 2, pp. 97ff, 1939.
52. H. Paul Douglas, Henry Schulz as colleague, E, vol. 7, No. 2, pp. 104ff, 1939.
53. Raymond Garver, The Edgeworth taxation phenomenon, E, vol. 7, No. 4, pp. 402ff, 1933.
54. Harald Hottelling, Note on the Edgeworth taxation phenomenon. Prof. Garver's additional condition on demand functions, E, vol. 1, No, 4, pp. 408ff, 1933.
55. Ragnar Frisch, The Dupuis taxation theorem, E, vol. 7, No 2, pp. 154ff, 1939.
56. A. A. Duncan, Marshalls paradox and the direction of shift in demand, E, vol. 6, No. 4, pp. 375ff, 1938.
57. Hilda Geiringer, New explanation of nonnormal dispersion in the Lexis theory, vol. 10, No. 1, pp. 53ff, 1942.
58. A. Paul Samuelson, Lord Keynes and the general theory, E, vol. 14, No. 3,

pp. 187ff, 1946.
59.J. R. Hicks, Keynes and the classics, a suggested interpretation, E, vol. 5, No. 2, pp. 147ff, 1937.
60. Nicholas Kaldor, On the theory of capital. Rejoinder to Prof. Knight, E, vol. 6, No. 2, pp. 163ff, 1938.
61. Он же, Annual survey of economic theory. The recent controversy on the theory of capital, E, vol. 5, No. 3, pp. 201ff, 1937.
62. Jakob Marschak, Identity and stability in economics: a survey, E, vol. 10, No. 1, pp. 61ff, 1942.
63. Jan Tinbergen, Critical remarks on some business cycle theories, E, vol. 10, No. 2, pp. 129ff, 1942.

1*. My likely inadequate explanation: principle of responsible dealings.
2*. The English historical school in economics which sought a return of inductive methods.
3*. This is a mistake. Schlözer, for example, warned that a Merkwürdigkeit was only taken into account if it promoted or hindered the goal of the state.
4*. Anderson published his main works in German and became the first statistician in West Germany. In general, the nationality of authors is not really important anymore.
5*. I have translated Cournot (**S, G,** 54) and harshly criticized him.

Commentary

This important paper ought to be criticized. Niksa mentioned many authors by second names only. Most of them were included in his Bibliography, but it is difficult to identify the other ones.

Understandably, he did not know that P. H. Laurent had published a *Traité* (1902) which was much occupied by the science *nouvelle created by Walras and his disciples about which he read at Sorbonne* (undated letter of Poincaré kept in his Dossier at the Paris Academy of sciences and addressed to the President [presumably, of that academy]).

But it is extremely important that the author entirely neglected Soviet/Russian scientists. **His paper is thus defective**.

I am unable to elaborate, but it is impossible to believe that such scientists like Kolmogorov, Bernstein (from Odessa) or Smirnov did not promote, even if only indirectly, the advent of econometrics. Then, Slutsky was a forerunner of that science (Zarkovic 1956/1977, p. 484), see also Kolmogorov (1948/2002).

For many years Chuprov had been attempting (not really successively) to bring together the Biometric school and the Continental direction of statistics. He also all but destroyed the Lexian theory of stability of statistical figures and thus posed the problem of solving it (Sheynin 1990/2011, pp. 141 – 143).

This led official Soviet statisticians to accuse him, and Bortkievicz, and Süssmilch (!) of attempts to perpetuate the stability of capitalism … (Там же, pp. 159 – 160). After Chuprov's death there appeared only one obituary written by his former close friend and published in a Leningrad newspaper. True, several obituaries written by Soviet statisticians appeared abroad.

Kolmogorov A. N. (1948/2002), Obituary. E. E. Slutsky. *Math. scientist*, vol. 27, pp. 67 – 74.
Laurent P. H. (1902), *Petit traité d'économie politique mathématique*. Paris.
Sheynin O. (1990/2011), *A. A. Chuprov. Life, work, correspondence.* Göttingen.

Zarkovic S. S. (1956/1977), Note on the history of sampling methods in Russia. In *Studies in history of statistics and probability*, vol. 2. Editors, Sir Maurice Kendall, R. L. Plackett. London, pp. 482 – 485.

IX

Rose-Luise Winkler

Ein unveröffentlichtes Manuskript[1] von Boris M. Hessen: „Materialien und Dokumente zur Geschichte der Physik“ (Druckfahnen – 1936 (?), ca. 700 Seiten, russ.)[1]

Sitzungsberichte der Leibniz-Sozietät 92 (2007), 133–152
der Wissenschaften zu Berlin. Kurzvortrag
vor der Klasse für Sozial- und Geisteswissenschaften am 14.12. 2006,
gewidmet Boris Hessen anlässlich seines 70. Todestages am 20. Dezember 2006

Der sowjetrussische, aus der Ukraine stammende Wissenschaftler Boris Michajlovic Hessen (russ. Gessen), Physiker, Philosoph, Soziologe und Wissenschaftshistoriker, gehört zu jener Generation marxistisch orientierter Wissenschaftler, die auf tragische Weise und vor ihrer Zeit zu Tode kamen. Wie wir heute wissen, wurde B. M. Hessen am 20. Dezember 1936 aufgrund konstruierter Anschuldigungen vom Obersten Militärgericht der UdSSR zum Tode verurteilt und am gleichen Tag hingerichtet. Er war 43 Jahre alt.

Sein Schicksal steht stellvertretend für eine Vielzahl von Wissenschaftlern, die den Stalin'schen Repressionen zum Opfer fielen. Das genaue Datum seiner Verurteilung und seines Todes wurde erst spät bekannt: mit der Veröffentlichung von Angaben aus der Ermittlungsakte Gessen aus dem Zentralen Archiv des KGB durch Gennadij E. Gorelik im Jahr 1992[2]. Bis zu diesem Zeitpunkt wurde angenommen, er sei 1938 verstorben. Man kann sich kaum eines makabren Gefühls erwehren, wenn auch heute noch seine Lebensdaten in wichtigen Veröffentlichungen falsch angegeben werden[3]. Das Datum seines Todes wurde in früheren Veröffentlichungen offenbar fälschlich angegeben. Das läßt sich aus der 2003 erschienenen, sehr umfänglichen, Dokumentation über die Kommission zur Geschichte des Wissens an der AdW [Akad. Wiss.] der UdSSR, schließen[4].

Diesen Angaben zufolge wurde Hessen nicht später als am 1. September 1936 verhaftet und kam in der Verbannung um[5]. Ob es sich hier um eine bewusste Verfälschung der Ausgangsdaten oder nur um Unkenntnis bzw. mangelnde Sorgfalt handelt[6], bedarf konkreter Nachweise, die mir zum gegenwärtigen Zeitpunkt nicht zur Verfügung stehen. Das von G. E. Gorelik angeführte „Stenogramm“ einer Versammlung am Physikalischen Institut der Akademie der Wissenschaften in Moskau (FIAN) von 1937 macht deutlich, den Teilnehmern war nicht bewusst, dass Hessen zu diesem Zeitpunkt schon nicht mehr am Leben war. Auch an der Physikalischen Fakultät der MGU, haben 2 Vollversammlungen der Studenten und der Aspiranten stattgefunden, konnte sich Eugen L. Fejnberg[7] im Sommer 2005 erinnern. Er hatte als Student in der Zeit von 1930 bis 1935 Vorlesungen bei Hessen gehört. Wessen man Hessen bezichtigte, wurde laut Fejnberg nicht offen gelegt. Angeblich habe er ein „verräterisches Lehrprogramm der Physik“[8] (вредительская программа физики) erstellt. Der Physiker Grigorij S. Landsberg hätte das Programm verteidigt. „Er selbst habe es erstellt und nicht Hessen“.[9] An den Zeitpunkt dieser Versammlungen konnte sich Fejnberg nicht mehr genau erinnern. (1936?[10]) „Man sprach von 10 Jahren Verbannung für derartige Vergehen. Hessen habe einen kleinen Zirkel über philosophische Fragen der Naturwissenschaften für Studenten durchgeführt. Er hätte sich daran beteiligen können, habe sich jedoch für einen anderen Zirkel entschieden.“

Ob es sich bei den angeführten Auseinandersetzungen um den Inhalt des nicht mehr zur Veröffentlichung gelangten Manuskripts, die oben genannten Druckfahnen handelte, dem ein bestimmtes von Hessen bereits in seinem berühmten Vortrag von 1931 begründetes Programm, zugrunde liegt? Noch lassen sich zu dieser Frage keine Antworten geben. Die wichtigsten Zeitzeugen, die darüber Auskunft geben könnten, leben schon lange nicht mehr.

Die schon genannte Quellendokumentation belegt, Hessen wurde durch den Beschluss der Vollversammlung der Akademie der Wissenschaften vom 29. April 1938 aus der Akademie ausgeschlossen.[11] Ein entsprechender Beschluss vom 5. März 1957 rehabilitierte ihn.[12] Nach Angaben der russischen Gesellschaft „Memorial“ befinden sich die sterblichen Überreste von B. M. Hessen wie auch die des mit ihm verurteilten Arkadij O. Apirin auf dem Moskauer Friedhof Donskoe.[13] Als Volksfeind und Verräter gebrandmarkt, wurden seine Arbeiten für lange Zeit aus den Bibliotheken in seiner Heimat entfernt und in wissenschaftlichen Veröffentlichungen verschwiegen. Für seine Rehabilitation hat sich vor allem Igor E. Tamm[14] eingesetzt, mit dem er zusammen in Edinburgh studierte und seit der Zeit seiner Kindheit befreundet war.[15] Igor E. Tamm verlor in jenen schicksalsschweren Jahren 1936/1937 mit Boris Hessen nicht nur einen seiner engsten Freunde, sondern auch seinen Bruder und mehrere ihm nahestehende Verwandte und Schüler.[16] Hinsichtlich der falschen Angabe des Geburtsdatums von Hessen scheint ein Irrtum[17] vorzuliegen, da infolge seiner Verhaftung die meisten Personalunterlagen über seine Person und Tätigkeit aus den einschlägigen Archiven entfernt wurden oder nicht mehr zugänglich waren und sind. Einige Nachweise sind erhalten, die von mir aufgefunden werden konnten: aus den Beständen von der Kommunistischen Akademie (Komakademija) ein persönlicher Arbeitsplan von 1924 (handschriftlich) und zwei Lebensläufe von Boris Hessen, ein handschriftlicher von 1924[18] und ein maschinenschriftlicher aus der Handschriftenabteilung der Staatlichen Russischen Bibliothek von 1930[19]. Aus beiden Lebensläufen geht sein Geburtsjahr (1893)[20] eindeutig hervor. Hinzuziehen kann man auch den Immatrikulationsnachweis von 1913–1914 von der Universität Edinburgh, in dem er seinen Namen (Hessen in lateinischer Schrift) und sein Alter mit 20 angibt. Dieser Nachweis ist ebenfalls handschriftlich.[21] Von I.E. Tamm ist ein Nachweis über sein Studium in Edinburgh von 1913–1914 abgedruckt (Non-Graduation Certificate, Faculty of Arts), mit der Unterschrift von E.T. Whittaker.[22] Es ist anzunehmen, dass Hessen gleichfalls ein solches Zertifikat erhielt.

Den bisherigen kargen biographischen Darstellungen zu Boris Hessen (über seine Familie konnten trotz mehrfacher Versuche meinerseits bisher keine zuverlässigen Angaben ermittelt werden) liegen hauptsächlich die in den beiden genannten Lebensläufen von ihm angegebenen Daten, Tätigkeitsfelder und Veröffentlichungen zugrunde sowie die schon genannten Erinnerungen über Igor E. Tamm. Einige spärliche Angaben können aus den Archiven der Russischen Akademie ergänzend herangezogen werden. Eine auf der Auswertung von Archivunterlagen der Moskauer Universität beruhende Darstellung gibt Leonid V. Levšin in dem Buch „Die Dekane der Physikalischen Fakultät an der Moskauer Universität", im Jahr 2002 zur bevorstehenden 250-Jahrfeier der MGU (2005) herausgegeben.[23] Seine lückenlose zeitliche Erfassung der Lehre in der Physik umfasst einen historischen Zeitraum von etwa 1756 bis heute, ab 1805 bis 1930 sind die Dekane der physikalisch-mathematischen otdelenije und später der Fakultät aufgeführt, ab 1930 besteht erstmals eine eigenständige physikalische otdelenije, 1933 in eine Fakultät umgebildet, deren erster Dekan Boris Michajlovič Hessen war.[24] Von Februar 1931 bis November 1934 war Hessen demzufolge Dekan der ersten eigenständigen Physikalischen Fakultät der Moskauer Universität, sowie ab 1930 bis zu seiner Verhaftung Direktor des Physikalischen Instituts an der MGU. 1934 wurde im Zusammenhang mit der Übersiedlung der Akademie von Leningrad nach Moskau das FIAN gegründet, dessen Direktor Sergej I. Vavilov, Boris Hessen als stellvertretenden Direktor an das Institut holte. Sein Nachfolger im Amt als Dekan wurde der Physiker Semen E. Chajkin. Hessen war Mitglied mehrerer wissenschaftlicher Gremien (der naturwissenschaftlichen Sektion der Komakademie, der AdW [Akad. Wiss.] der UdSSR, der GUS [Gemeinschaft unabhäng. Staaten, СНГ]) und nachfolgender Redaktionskollegien von Zeitschriften: „Naturwissenschaften und Marxismus" (Естествознание и марксизм), „Ergebnisse der Physik" (Успехи физики), „Physikalische Zeitschrift der Sowjetunion (von 1932–1936 in Charkow in deutscher Sprache vom Obersten Volkswirtschaftsrat der UdSSR herausgegeben), der Reihe „Biographien herausragender Persönlichkeiten" (Биографии замечательных людей) und der Großen Sowjet-Enzyklopädie (1. Auflage) sowie der Reihe Übersetzungen „Klassiker der Naturwissenschaften". Seit 1928 publizierte Hessen zu Fragen der theoretischen Physik, der Methodologie, Philosophie und Geschichte der Naturwissenschaften und zu Fragen der Lehre und Ausbildung in der theoretischen Physik und in den Naturwissenschaften. Am bekanntesten wurde sein Vortrag „Социально-экономические корни механики Ньютона"[25], den er 1931 auf dem 2. Internationalen Kongress für Geschichte der Wissenschaft und Technik in London hielt.[26] Dieser zählt heute zu den klassischen Arbeiten in der Wissenschaftsforschung und -soziologie. Der Vortrag wurde in 6 europäische Sprachen und ins Japanische übersetzt und mehrfach aufgelegt.[27] Der Mehrzahl der Übersetzungen liegt die englische Fassung von 1931 zugrunde, die bedauerlicherweise eine Vielzahl von Übersetzungsschwächen vor allem in der Wiedergabe von Fachbegriffen sowie irrtümliche Angaben über

Personennamen und Orte enthält'[28] Der bereits im Titel gewählte Begriff „sozialökonomische Ursprünge" (Grundlagen, Wurzeln) wurde beispielsweise in der deutschen Übersetzung in Analogie zur englischen Fassung sinnentstellend in sozial und ökonomisch zerlegt.[29] Er wurde von Hessen im Sinne des auf Marx zurückgehenden formationstheoretischen Herangehens verstanden, abgeleitet von dem Begriff der „ökonomischen Gesellschafts-formation". In der englischen Übersetzung dieses Begriffs (von K. Marx und F. Engels selbst) finden sich die Ausdrücke „economic formation of society",“economic social formation".[30] Ein Begriff sozialökonomische Formation oder auch das Adjektiv sozial-ökonomisch dagegen findet sich nicht, dieser geht wahrscheinlich auf Hessen zurück und ist später gleichbedeutend mit der Beschreibung von Fragen der gesellschaftlichen Determination von sozialen Erscheinungen (wie Wissenschaft, Kunst, Kultur, Produktion u.a.) in der marxistisch orientierten Soziologie und Wissenschaftsforschung verwendet worden.

Es spricht für die Produktivität des Ansatzes von Boris Hessen, wenn dieser trotz der Übersetzungsschwächen eine so weitgefächerte Diskussion initiierte, wie sie in der Hessen-Rezeption seit 1931 zum Ausdruck kommt. Die Wirkung seines Beitrages ist vergleichbar der des von Thomas S. Kuhn eingeführten Paradigma-Begriffs in der Wissenschaftsforschung in den 1960–70er Jahren. Der Begriff der sozialökonomischen Determination ist in die Folgezeit einer der wichtigsten Grundbegriffe für soziologische Analysen geworden, da er Aussagen zum Verhältnis von Gesellschaftsformation und Wissenschaft in empirisch erfassbare und interpretierbare Sachverhalte übersetzt. Boris Hessen hat damit eines der Kardinalprobleme der wissenschafts-soziologischen Forschung formuliert und an einem prägnanten Objekt Fragen dazu aufgeworfen. (Man könnte auch fragen, war Newton eine Ausnahmeerscheinung? Oder wie ist die Einsteinsche Relativitätstheorie in dieser Hinsicht heute zu verorten?) In welcher Beziehung stehen der vom Marxismus geprägte formationstheoretische Ansatz und die Auffassung von der Moderne von heute? Diese Fragestellung wird in der Gegenwart kaum thematisiert bzw. bewusst gemieden. Sie ist weder von geringem theoretischen Interesse[31] oder mangelnder Bedeutung noch gelöst, sondern stellt eher ein Entwicklungsproblem der Gesellschafts-wissenschaften von heute dar.

Auch die der russischen Veröffentlichung beigefügten historischen Quellen und Literaturangaben, die in der englischen Fassung nicht enthalten sind, sind noch heute von Interesse. So beispielsweise die im Deutschen nicht bekannte Satire „Burleskes Urteil – gefällt vom Hohen Gericht des Parnasse aufgrund der Klage von Magistern, Medizinern und Professoren der Universität Stagire im Land der Chimeren: Bewahrung der Lehre von Aristoteles" von Nicolas Boileau (zuerst 1671).[32] Hessen stellt in dem (wiederaufgefundenen) Manuskript die Geschichte um diese Satire ausführlich dar, eine kleine soziologische Lektion in Fragen um das Verhältnis von Wissenschaft und Macht im universitären Machtkampf zwischen Zentralgewalt und Provinz. Sie wurde von einem russischen Physiker aus dem

Französischen übersetzt und in einer Geschichte der Physik publiziert.[33] In ihr wird das Verhalten der scholastischen Naturphilosophen gegenüber der experimentellen, empirischen Erforschung der Natur gegeißelt. Es wird deutlich, dass wissenschaftliches Erkennen nicht den Beschlüssen von staatlichen und anderen Gremien unterliegen kann, erst recht keinen Urteilen von Gerichten. Fragen der Feststellung von Wahrheit oder Falschheit im Prozess der wissenschaftlichen Erkenntnis können nicht per Gerichtsbeschluss (über Lehrverbote) geregelt werden. Analoge Situationen, wie sie in dieser Satire dargestellt werden, sind wohl in der Geschichte der Wissenschaft in allen Ländern keineswegs selten. Zu Hessens Zeit traf dies auch für die Wissenschaft in der UdSSR zumindest in Teilbereichen zu: Empirische soziologische Forschungen unterlagen zunehmend Restriktionen, in der Physik gab es die Auseinandersetzung um die Einsteinsche Relativitätstheorie, in der Hessen selbst öffentlich bezichtigt wurde, dem Einfluss bürgerlicher Einstellungen zu unterliegen.

Hessen, der in seiner Tätigkeit als Wissenschaftler soziologische Methoden (beispielsweise Zeitbudgetanalysen) anwandte, um die Probleme der wissenschaftlichen Arbeit von Physikern zu diskutieren, setzte sich auch intensiv mit Fragen der sozialen Organisation von Wissenschaft auseinander.

Er war Mitglied der gesellschaftswissenschaftlichen Klasse der Akademie (seit 1933) und stand als Physiker im Zentrum der physikalischen Arbeiten an der Akademie und an der Universität. Die in der Rezeptionsgeschichte der Arbeiten von B. M. Hessen bestehende Diskrepanz, eine unausgewogene Darstellung seines physikalischen, philosophischen, soziologisch orientierten und wissenschaftshistorischen Schaffens wird durch das vorliegende unveröffentlichte Manuskript offenbarer. Über seine physikalischen Arbeiten ist wenig bekannt, und schriftlich kaum etwas überliefert. Es kann als ein Studienmaterial für angehende Physiker, Naturforscher und an der Geschichte der Physik interessierte Philosophen, Soziologen und Wissenschaftshistoriker angesehen werden.

Hessen ist bemüht, die Rolle der historischen Untersuchung für das Verständnis der physikalischen Kategorien deutlich zu machen. Er legt großen Wert auf die Kenntnis der Originalquellen. So präsentiert er eine Zusammenstellung von Originalquellen zur Geschichte der Physik, die er seinem Konzept entsprechend in drei große Themen gliedert: Thema I. Die sozialökonomischen Voraussetzungen (Hervorhebung R.-L.W.) der klassischen Physik, Thema II. Die Entstehung und Entwicklung der Hauptprinzipien der klassischen Mechanik und die Auseinandersetzungen im 17. Jahrhundert darum, Thema III. Das Problem der Bewegung in der Physik Newtons. Der Kampf von Materialismus und Idealismus um diese Frage im 17. Jahrhundert. Im Thema I und III erfahren einzelne Abschnitte seines Vortrages von 1931 eine vertiefende Darstellung, was sich anhand eines Vergleichs mit der Feingliederung des Vortrages leicht feststellen lässt. So stimmen vielfach die Zwischenüberschriften im Wortlaut überein.

Im erhalten gebliebenen Vorwort begründet Hessen sein Herangehen ausführlich:

„Der vorliegende Band von Dokumenten und Materialien hat zum Ziel, den Leser mit den Originalquellen zur Geschichte der Physik bekannt zu machen. Von analogen Textsammlungen, die es in der westeuropäischen Literatur gibt und die zumeist eine Zusammenstellung von Auszügen aus den klassischen Arbeiten in chronologischer Reihenfolge darstellen, unterscheidet er sich vor allem durch die Auswahl und die Darbietung des Quellenmaterials. Das physikalische Material wird auf dem Hintergrund der sozialökonomischen Verhältnisse. (Hervorhebung R.-L.W.) der entsprechenden Epoche vorgestellt. Daraus erklärt sich der im Vergleich zur üblichen Geschichte der Physik große Anteil an ökonomischem und technischem Material.“(S. 6) ... Der vorliegende Band stellt sich nicht die Aufgabe, eine systematische Darstellung der Geschichte der Physik zu geben, sondern widmet sich einer Reihe von Themen, häufig auch aus weit voneinander entfernten Perioden. Das bietet uns die Möglichkeit, einzelne Momente in der Geschichte der Wissenschaftsentwicklung, ihre sozialökonomischen Voraussetzungen und Peripetien in der ideologischen Auseinandersetzung vollständiger und umfassender zu beleuchten. (S. 7) ...“

Für den Band wurde eine Reihe von vorhandenen Übersetzungen benutzt und anhand der originalsprachlichen Fassungen überprüft, wie Hessen im Vorwort vermerkt. Für einen großen Teil des Materials wurden erstmals Übersetzungen in Russisch angefertigt. Jedem Kapitel ist eine kurze Einführung vorangestellt, in der die Auswahl der Quellen begründet und inhaltliche Orientierungen für die Darstellung erfolgen. Diese ist vielfach mit ausführlichen historischen Kommentaren von Hessen versehen. Das Kapitel II stellt eine Art Chrestomathie zur Geschichte der Physik des 17. Jahrhunderts dar. Viele der hier aufgeführten Erstübersetzungen sind für den russischen Leser teilweise auch heute noch nicht verfügbar (zum Beispiel die Artikel von A. E. Haas „Antike Dynamik“, von Johann Bernoulli „Über die Dynamik Newtons und Descartes“, oder von Rodger I. Boskovič „Über die Prinzipien des Aufbaus der Mechanik“, oder die (erst 1993 von Ju. A. Danilov, in ВИЕТ 1 (1993): 30–45, veröffentlichten) Boyle-Lectures von R. Bentley und sein Briefwechsel mit Is. Newton).[34] Auch die herangezogenen sozialwissenschaftlichen und ökonomischen Quellen waren vielfach jüngsten Datums. Ob Hessen Kenntnis von unveröffentlichten Arbeiten von K. Marx und F. Engels hatte, lässt sich aus den Materialien nicht erschließen.

So erfolgte beispielsweise die Veröffentlichung der „Deutschen Ideologie“, auf die sich Hessen in seinem Vortrag von 1931 stützt, 1927 in deutscher Sprache vom Marx-Engels-Lenin-Institut (Moskau) und 1933 in russischer Sprache.[35] Dieses Werk gehört zum Grundbestand des sozialwissenschaftlichen Wissens, zur Zeit Hessens waren diese Quellen neuartig, und wissenschaftlich kaum erschlossen. Für den deutschen Leser sei zusätzlich auf die russischen Quellen verwiesen, die im allgemeinen in analogen Arbeiten zur Geschichte der Physik keinen Niederschlag finden.

Das von Hessen konzipierte Buch stellt somit eine nicht nur für die 1930e Jahre neuartige Sichtweise einer Physikgeschichte dar, es kann darüber hinaus auch als Pionierleistung für die in dieser Zeit entstehende Wissenschaftsforschung gelten. Das Manuskript dürfte Ende 1935 oder Anfang 1936 an den Verlag gegangen sein. So teilte Hessen in einem Brief vom 26. Juni 1935 an J. G. Growther diesem mit, dass er eine dritte wesentlich überarbeitete und erweiterte Auflage seines Vortrages vorbereite und fragte, ob dieser eine englische Ausgabe übernehmen würde.[36] V. S. Kirsanov vermutet das Jahr 1936, da seiner Meinung nach bei einem früheren Datum das Buch 1936 schon in den Verkauf hätte gehen müssen.[37]

Die Frage, wer die Übersetzungen der originalen Quellen anfertigte, lässt sich nicht mehr beantworten, da Angaben darüber in den Druckfahnen nicht enthalten sind. Das im Nachlass von A. P. Juškevič aufgefundene Exemplar der Druckfahnen ist den Angaben von Kirsanov zufolge dem bekannten, herausragenden Übersetzer Vladimir Solomonovič Gochman (1880–1956), dem Schwiegervater von Juškevič, zugehörig. Eine Autorschaft von Gochman selbst lässt sich aber ausschließen. Zu den von Kirsanov[38] besprochenen unveröffentlichten Manuskripten aus den 1930er Jahren gehört auch eine neu Übersetzung von Newtons „Principia", für die Hessen als Herausgeber vorgesehen war. Diese war in einer 7-bändigen Gesamtausgabe Newtons 1934 von S. I. Vavilov konzipiert und vom Verlag bestätigt.

Im folgenden geben wir den Inhalt nach dem Manuskript der Druckfahnen wieder. Ein Gesamttitel ebenso wie ein Inhaltsverzeichnis lag dem Manuskript nicht bei. Auch ist es unvollständig. Fehlende Abschnitte sind vermerkt. Die Seitenangaben entsprechen den Seitenzahlen der Druckfahnen.

Anlagen:

1. handschriftlicher Lebenslauf von Boris Hessen (8. Juli 1924) aus dem Bestand Lebensläufe der Komakademie: F. 364. Opis 3a. Nr. 17. Bl.3.
2. persönlicher Arbeitsplan zur Naturwissenschaft 1924/25, handschriftlich. Ebenda: F. 364.Opis 3a. Nr. 17. Bl. 4,

3. Wiedergabe beider Dokumente in russisch und deutsch.

Lebenslauf B. Gessen

Ich wurde 1893 geboren. 1913 habe ich die 8. Klasse eines Gymnasiums beendet. Von 1913–1914 habe ich an der Universität Edinburgh /Schottland/studiert. / Faculty of Science department of Pure Science. Ich belegte dort die folgenden Vorlesungen und Übungen und legte die Prüfungen ab: Einführung in die Analysis und den ersten Abschnitt zur Differentialrechnung bei Prof. Whittaker und analytische Geometrie bei Dr. Carse.u. Wärme bei Prof. Barkla und ein physikalisches Praktikum bei Dr. Carse. Anorganische Chemie und ein chemisches Praktikum bei Prof. Walke. ... Dr. Dobbin. Da es während des imperialistischen Krieges nicht möglich war nach England zu gelangen, studierte ich zwei Jahre 1914-1916 an der ökonomischen Fakultät des Petrograder Polytechnikums. Dort habe ich zur Statistik bei A. A. Čuprov und Mares gearbeitet und beschäftigte mich auch mit mathematischer Statistik. Ebenfalls war ich Hörer an der mathem.-physik. Fakultät der Petrograder Universität, an der ich als Jude nicht angenommen wurde. In diesen zwei Jahren hörte und arbeitete ich zu Fragen der Differential- und Integralrechnung, Prof. Uspenskij und Selivanov, Anwendung der Analoga in der (Geometrie?) – Adamov, Höhere Algebra – Ju. Sokockij, Theorie der Bestimmung von Integralen, Sokockij, Integration von Differentialgleichungen – Steklov. Zu diesen Fächern konnte ich natürlich kein Examen ablegen. Außerdem beschäftigte ich mich selbständig mit Philosophie und ein wenig mit Geschichte der Mathematik.

Seit Beginn der Revolution war ich in der Parteiarbeit und in der propagandistischen Arbeit tätig: 1917 bis zum Oktober als Sekretär der Organisation der Internationalisten in Jelisavetgrad, nach dem Oktoberumschwung als Sekretär des Rates der Arbeiterdeputierten, 1919 ab August – Mitglied des Kollegiums der Abteilung für Volksbildung. Von 1919–1921 zunächst Instrukteur in der politischen Arbeit und in den Abteilungen ... und der Abteilung zur Ausbildung von Personal. Seit 1921 bis heute bin ich an der Sverdlov-Universität tätig und lehre Politische Ökonomie, ich bin Leiter für den Lehrzyklus Ökonomie und die Lektorenausbildung. Ich beherrsche deutsch, französisch, englisch und lateinisch. 8.VII.1924 Unterschrift (B. Gessen)

An die Leitung der Hochschule für Rote Professur B. Gessen

Persönlicher Arbeitsplan zur Naturwissenschaft 1924/25

Die Grundlage für meinen Arbeitsplan sind die Besonderheiten meiner naturwissenschaftlichen Ausbildung, die in meinem curriculum vitae dargelegt sind: bei einer vergleichsweise hinreichenden mathematischen Vorbildung fehlt mir eine systematische physikalische Ausbildung. Außerdem macht die siebenjährige Pause (1917–1924) in meiner Tätigkeit eine Wiederholung und Auffrischung meiner Kenntnisse notwendig.

In Mathematik: Gründliche Wiederholung der Differentialgeometrie und der gewöhnlichen Differentialgleichungen nach Czuber – Vorlesungen über Differential- und Integral Rechnung, nach Steklov –

gewöhnliche Differentialgleichungen und Elemente der Variationsrechnung. Vorlesungen, die an der St. Petersburger Universität 1912/13 gehalten wurden.

In Physik: Gründliche Wiederholung eines Kurses zur experimentellen Physik nach Edser. Properties of Matter. Edser Heat. Eichenwald Elektrizität. Theoretische Physik im Umfang von Haas. Einführung in die theoretische Physik. Bd. 1, neueste Auflage. Helmholtz. Dynamik der diskreten Massenpunkte.
Laborarbeiten: einige Aufgaben zur experimentellen Physik (ein allgemeines Praktikum habe ich in England absolviert) und praktische Übungen zu elektrischen Schwingungen.
November 1924 B. Gessen
Nachweis: Foto. Archiv der MGU. F. 46. Opis 1. Ed. 52. L.1

Anmerkungen

ИКП (Abk.) Институт Красной Профессуры
Стеклов, Владмир Андреевич (1864-1926), russ. Mathematiker, (1912), Vizepräsident der Akademie der Wissenschaften der UdSSR 1919-1926, Organisator und Direktor des Physikalisch-mathematischen Instituts 1921-1926, OM [ordinary member of the Acad. Sci.]
Чупров, Александр Александрович (1874-1926), russ. Mathematiker und Statistiker. Lehrte 1902-1917 an der Ökonomischen Fakultät des Polytechnischen Instituts in Petersburg.
Czuber, Emanuel (1851-1925)

Notes

1. Die Druckfahnen wurden im Dezember 2004 von Vladimir S. Kirsanov im Nachlass von A.P. Juškevič aufgefunden und mir freundlicherweise nach der Übertragung auf eine CD-ROM im Frühjahr 2005 zur Verfügung gestellt. Dem Beitrag liegt eine gekürzte, überarbeitete Fassung meines Vortrages „Boris Hessen and the Origins of Sociology of Science in Soviet Union (Russia)“ auf dem XXII. Internationalen Kongreß für Wissenschaftsgeschichte vom 24.–30. Juli 2005 in Peking zugrunde.
2. Г.Е. Горелик. Москва, физика, 1937, ВИЕТ, 1992. № 1: 15–32; dt.: G. E. Gorelik. Physiker unter Stalin. Kap. 5. Braunschweig 1995: 98–133.
3. So ist in dem Handbuch: „Академия наук. Персональный состав“ 2. Auflage (Bd. 1 u. 2., Verlag der Wissenschaft Moskau Bd. 2. 1999: 176) der nachfolgende Eintrag vermerkt:

Гессен Борис Михайлович. Родился 28 августа 1883 г., Елизаветград Херсонской губ., Умер 9 августа 1938 г., Москва (?), Философ. Член-корреспондент по отделению общественных наук (философия), с 1 февраля 1933 г. Analog finden sich diese Daten in der CD-ROM „The Russian Academy of Sciences 1724–1999“ , in russ. und engl. erschienen (Gessen Boris Mikhailovich, Born on 28.08. 1883, Died on 09.08.1938. Philosophy. Corresponding Member of the division of Social Sciences; since 01.02.1933) sowie in Internet-veröffentlichungen.
4. Siehe: Комиссия по истории знаний – 1921–1932 гг. Из истории организации историко-научных исследований в Академии наук. Составители: В. М. Орел, Г. И. Смагина. Изд.-во Наука. Петербург 2003: 580, 681.

Das Jahr 1936 ist hier richtig, aber das Geburtsjahr falsch angegeben (1883). Ebenda. Bedauerlicherweise werden heute oft Daten ohne vorherige Prüfung übernommen. Im Fall Hessen spiegelt sich dies auch in der internationalen Literatur wider.
5. Das Jahr 1936 ist hier richtig, aber das Geburtsjahr falsch angegeben (1883). Ebenda.

6. Bedauerlicherweise werden heute oft Daten ohne vorherige Prüfung übernommen. Im Fall Hessen spiegelt sich dies auch in der internationalen Literatur wider [повтор Прим. 4].
7. Evgenij L'vovic Fejnberg (1912 – 10.12. 2005), KM [corr. member] (1966) und OM [ordinary member] (1997) an der Russischen Akademie der Wissenschaften. Anfang August 2005 konnte ich ein Gespräch mit Eugen L'vovič in Moskau in seiner Wohnung führen.
8. Vgl. dazu auch: Е.Л. Фейнберг. Вавилов и Вавиловский ФИАН//Эпоха и личность. Физика. Очерки и воспоминания. Физматлит Москва 2003: 241.
9. Там же.
10. Das könnte zutreffen, denn die Verhaftung Hessens konnte nicht unbemerkt bleiben.
11. Vgl. Комиссия по истории знаний – 1921 – 1932 гг. Из истории организации историко-научных исследований в Академии наук. A.a.O.
12. Тм жне. Dem ging die Rehabilitation durch das Oberste Militärgericht der UdSSR vorher:
Am 21. April 1956 wurde die Verhaftung von Hessen annulliert mangels eines zureichenden Grundes.
13. Internetveröffentlichung von Memorial. (http://www.memo.ru), vgl. auch Boris Hessen. Wikipedia (free encyclopedia).
14. I. E. Tamm (1895–1971), Nobelpreis für Physik 1958.
15. Anlässlich des 100. Geburtstages von I. E. Tamm veröffentlichte die Zeitschrift „Priroda" ein Sonderheft mit Erinnerungen von Zeitgenossen, Schülern und einer Reihe von originalen Dokumenten von ihm, in denen auch das Schicksal von Boris Hessen Erwähnung findet. Vgl. Специальный выпуск. Природа № 7 (959) Июль 1995. К 100-летию Игоря Евгеньевича Тамма. См. также: Е. Л. Фейнберг.Эпоха и личность.A.a.O.: 60.
16. Ср. Воспоминания о И. Е. Тамме. Изд. второе, дополненное -. Отв. Ред. Е. Л. Фейнберг. „Наука". Москва 1986: 285–288, 298. Reminiscences about I. E. Tamm. Ed. By E. L. Feinberg. Nauka Publisher Moscow 1987.
17. Möglicherweise ein Druckfehler. Dieses Datum ist auch im Handbuch « Научные работники Москвы.» Часть IV. Ленинград 1930: 63, № 137 enthalten. Dort heißt es:
Гессен Б. М. доц. Каф. Истории и философии естествознания при МГУ, н. сотр. Комакадемии; физика, методология точн. естествознания, обоснование статистич. механики и теории относительности. ~ Пл. Свердлова, 2-й дом Советов, кв. 21, тел. 2-80-77 (16 VIII 83 Елизаветград).
18. Ср.: Автобиография Б. М. Гессена. 8. VII. 1924. Архив АН СССР. Фонд 364 (Komakademie/IKP). Опись 3а. № 17. Л. 3. (im Anhang wiedergegeben)
19. Автобиография Б. М. Гессена. Отдел рукописей ВГБИЛ, Фонд 384 (В. И. Невский), папка 6, ед. Хр. 15.
20. Auf diesen beiden Dokumenten beruhen meine Angaben zu Hessens Geburtsjahr: Vgl. meine Kurzbiographie B. M. Hessen // Portraits of Russian and Soviet sociologists. Special Issue. Berlin–Moskau 1990: 126–130, (dt. u. russ. 1987–1988: 208–210, 168–170). Eine gekürzte Wiedergabe ist enthalten in: Социологи России и СНГ XIX –XX вв. Biобиблиографический справочник. Эдиториал УРСС. Москва 1999: 64.
21. University of Edinburgh Matriculations (1913–1914), 45, Nr. 873. Die Universität verlangt für eine Kopie dieses Nachweises eine Summe von 12 Pfund. Die Kopie befindet sich in meinem Privatarchiv.
22. Ср. И. Е. Тамм в дневниках и письмах//К 100-летию Игоря Евгеньевича Тамма. A.a.O: 137.
23. Л. В. Лёвшин. Деканы физического факультета Московского университета. Москва, 2002. (Leonid V. Levšin, Leiter der Otdelenije für experimentelle und theoretische Physik an der MGU, seit 1969). Das Archiv der MGU war mir persönlich nicht zugänglich und in den Jahren 2004–2005 wegen Umzug in den Neubau der Bibliothek geschlossen.
24. Т ам же: 18, 198–203.
25. 1933 und 1934 als Einzelveröffentlichung erschienen. Wiederabgedruckt 1992 und 1998 in: Из истории социологии науки: советский период 1917–1935, У истоков формирования социологии науки. Россия и Советский союз. Первая треть XX. века. Тюмень. Hrsg. R.-L. Winkler.

26. The Social and Economic Roots of Newton‘s Principia. // Science at the Crossroads. Papers presented to the International Congress of the History of Science and Technology held in London from June 29th to July 3rd, 1931 by the Delegates of the URSS, Russian Foreign-Languages Press, Kniga, London 1931: 149–212.
27. Englisch (Sidney 1946, Lexington Mass. 1968, London 1971, New York 1971) schwedisch (Stockholm 1972), deutsch (Frankfurt a.Main 1974), spanisch (Havanna 1985, Montevideo 1988, Pentalfa, Oviedo 1999, Barcelona 2001), französisch (Paris 1978 u. 2006), italienisch (Bari: De Donato 1977), japanisch (S.R. Mikulinskij spricht von zwei Auflagen: С. Р. Микулинский. Очерки развития историко-научной мысли. Москва 1989).
28. Auf verschiedene aus der Übersetzung herrührende Fragen wurde in der Literatur verschiedentlich hingewiesen. Der volle Umfang der Übersetzungsschwächen wurde jedoch erst mit dem Textvergleich zur russischen Ausgabe und mit der Prüfung der dem Vortrag von Hessen zugrundeliegenden Quellen deutlich. Bemühungen von DDR-Wissenschaftlern um eine neue Übersetzung und Editierung blieben bis Ende der 1980er Jahre ohne Erfolg.
29. Noch sinnentstellender ist die in der deutschen Übersetzung von J. G. Crowthers Arbeit von 1935 wiedergebene Formulierung „Die sozialen und volkswirtschaftlichen Ursachen von Newtons Principia". Vgl. Einführung. Grosse englische Forscher: aus dem Leben und Schaffen englischer Wissenschaftler des 19. Jahrhunderts. Berlin 1948:
30. Für die Hilfe bei der Auffindung der englischen Termini bei K. Marx und F. Engels zum Formationsbegriff bedanke ich mich hier bei Frau Regina Roth vom Akademienvorhaben MEGA (?) an der BBAW [Berlin-Brandenburgische Akad. Wiss.].
31. Eine Neuübertragung des Vortrages ins Deutsche wurde von der Autorin vorgenommen (Drucklegung in Vorbereitung).
32. 1671 forderten Theologen und Mediziner der Pariser Universität eine Regierungsentscheidung zur Verurteilung der Lehren R. Descartes. In einer beißenden Satire machte N. Boileau diese Forderungen der gelehrten Scholastiker lächerlich.
33. Sie stammt, wie ich jetzt dem unveröffentlichten Manuskript entnehmen konnte, von N. A. Ljubimov. Vgl. Н .А. Любимов. *История физики*. Bd. 1–3. St. Petersburg 1896. Bd. 3: 508–511. (nach der Ausgabe: *Œuvre*. De Boileau-Despreux. Paris 1798: 391).
34. Vgl.: В. С. Кирсанов. Уничтоженные книги: эхо сталинского террора в советской истории науки. ВИЕТ. 2005. №.4: 122.
35. Die Erstveröffentlichung erfolgte 1927 durch D. B. Rjasanov. *Zur Problematik der Erstveröffentlichung der „Deutschen Ideologie"* vgl.: Erfolgreiche Kooperation. Das Frankfurter Institut für Sozialforschung und das Moskauer Marx-Engels-Institut (1924–1928). Beiträge zur Marx-Engels-Forschung. Neue Folge. Sonderband 2. Argument Verlag Berlin Hamburg 2000.
36. C. A. J. Chilvers hat diesen Briefwechsel 2003 veröffentlicht: „The dilemmas of seditious men: the Crowther-Hessen correspondence in the 1930s//BJHS [Brit. J. Hist. Sci.] 36 (4) December 2003: 432.
37. См. Note 34.
38. Там же: 119–124.
39. Sir Edmund Whittaker (1873-1956), fourteenth Prof. of Mathematics in 1912 in the Faculty of Science. He established the first mathematical laboratory for numerical computation.
R. M. Birse. Science at the University of Edinburgh 1583-1993. An Illustrated History to mark the Centenary of the Faculty of Science and Engineering 1893-1993. The Faculty of Science and Engineering. The University of Edinburgh 1994: 97.
40. Charles Barkla, the eleventh Prof. of Natural Philosophy, chair of Physics at King`s College London since 1909, Nobel Prize for Physics 1917. Ebenda.

Commentary

Some places are difficult to understand and those who do not read Russian are seriously disadvantaged. The author made no attempt to present orderly her valuable paper. Thus, Notes 34 and 37 were identical and books are not distinguished from other sources.