

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

**Vol. 21**

**Compiled and translated by Oscar Sheynin**

**Berlin**

**2020**

## Contents

I am the author of contributions [ii], [vi], and [vii]

### Introduction by the compiler

I. R. J. Pulskamp, translator, Correspondence of Leibniz and Jacob Bernoulli [on probability], excerpts, no date

II. On the history of the statistical method in biology, 1980

III. A. L. Tzikalo, *A. M. Liapunov*, excerpts, 1988

IV. Alph. De Candolle, On a dominant language for science, 1873/1875

V. J. D. Sarna, B. Shapell, *Lincoln and the Jews*, excerpts, 2015

VI. Poisson's work in probability, excerpts, 1978

VII. Boscovich's work on probability, 1973

VIII. Jacob Bernoulli. On the law of large numbers, 1713

[oscar.sheynin@gmail.com](mailto:oscar.sheynin@gmail.com)

## Introduction by the compiler

### Notation

Notation **S, G, n** refers to downloadable file *n* placed on my website [www.sheynin.de](http://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

### General comments on some items

[i] So when had J. B. proven his main proposition (the law of large numbers)? The text above does not answer this question and we only know that its fragmentary proof appeared in a manuscript, in the author's *Meditations*, not later than 1690. This year, indeed, is the end of this manuscript whose excerpts had been since published (Jacob Bernoulli 1975).

The translation is decidedly bad. Pulskamp had slavishly copied the original: there should have been much more paragraphs, and the order of words in sentences is often clumsy (perhaps for the same reason). I have largely corrected these faults. The very translation is often either dubious or, especially in the last letters, unintelligible. Once more, I corrected at least some such deficiencies/mistakes.

Pulskamp was one of the two translators of the celebrated dissertation of Nicolas Bernoulli. Their translation is available in the Internet but it is almost useless. Suffice it to mention that Latin phrases 40 words long are rendered into English by equally long sentences and therefore make a complete nonsense.

[iv] The author (1806 – 1893) was an eminent botanist, foreign member of the academies of Sweden and Netherland, who influenced Asa Gray. His book (which I did not see) listed in the subtitle of this contribution was reprinted in 1885 and in 2018 (Amazon) and its German translation (Leipzig, 1911) was reprinted by Vero Verlag in 2015. The author denied Galton's idea about heredity being the main factor of talent. Instead, he put forward the influence of the surroundings which he understood in the wide sense.

The translation here reprinted is not good enough. The order of words is peculiar, many sentences are clumsy and certainly not thought out and instead of *easier more easy* etc. has appeared. I scarcely corrected her.

The author's prediction of the growth of population proved damnably wrong since the circumstances of life had drastically changed: families in the developed nations began to restrict the number of their children. This did not happen elsewhere, whereas in the world of Islam the main duty of women is still the production of children.

Artificial or rather constructed languages: in 1887, appeared the Esperanto. It had been used for about 80 years (and perhaps still a bit used) and among the Esperantists were Ho Chi Ming, President of

North Vietnam, and Josep Broz Tito, President of Yugoslavia. And there recently appeared the Interlingua – English Dictionary.

English is the first or second language for about a billion people and, in spite of the author's fear, did not separate into three different items. But specialists distinguish English of England, Welsh and Scottish, American, Canadian, Caribbean and ten other varieties of the same language. And I presume that the English of England proper has some dialects as well.

The Standard English resorts to dialectical and popular speech but is still rather a written than a spoken language. Human speech is a creation of unlettered people. Professors of New England or editors of *Times* influenced it but little. For that matter, in his § 24 Schlözer, writing in 1804 (see **S, G.** 86), counted 38 newspapers in London alone. And television is indeed influencing the language and certainly in the popular direction.

Usage creates errors and turns them into correct currency and language is always on the move. No one writes as Dickens did, and neither O'Henry did as Fenimore Cooper. Advances of science and technology are necessarily accompanied by the appearance of new words (or of new meanings of existing words).

And why should the Americans stick to the English of their Old Country? Pronounce *tomato* and *potato* in a different way? Write *plough* rather than *plow*? The author himself stated that men always choose the direct road!

My comments were partly based on authoritative sources and I conclude by stating that the author's pertinent opinion is hardly useful.

[v] This is a book about Abraham Lincoln (born 1809, assassinated 1865) President from 1860, intended to please the reader and at the same time to authenticate the contents. Photos are countless: portraits, texts of documents including handwritten notes by Lincoln, buildings, a map of Illinois of German origin (1845); paper of apparently best possible quality etc. Appended are chronologically arranged annotations of Lincoln's activity and an index of names and subjects.

Description is documented although not always definitely enough.

The content is much wider than implied by the title since much attention is devoted to Lincoln's attitude to slavery and in general the book portrays Lincoln's life and shows him as a humane and simple man. However, the separate episodes are sometimes difficult to date and, what is really frustrating, neither these episodes, nor the sketches which I mentioned above always stated whether Lincoln dealt with a Jew or a Gentile. In this sense the authors failed and I experienced much difficulty in isolating the former cases. Then, at least in one important case (two unnumbered pages of *Lincoln Jewish connections*) the simplicity of listing these connections is sacrificed to elegance.

In those days, everyone was religious, but apparently almost no Christian knew the true story of Judas. The *New Testament* states: *First*. God the Father commanded Jesus to die on the cross. *Second*, the Devil entered Judas who was unable to understand what he was doing. And, in addition, even up to nowadays no Christian theologian had publicly explained why did the Devil implement the command of

God the Father. It was and is so easy to blame Judas, the Jew ... And Jews never thought of opening the *New Testament*.

Explanation of some terms. *Elector*: a person elected by his party to vote in the election of a president and vice-president. *Abolitionism*: movement to abolish slavery. *Civil War*: 1861 – 1865 between abolitionists (*Union*) and other seven states (*Confederacy*) which allowed slavery (and depended on slave labour).

The appearance of Negroes led to unsolvable problems: when slavery was abolished, many of them moved northward and worsened the condition of white workers on the labour market. Most serious events followed ... The appearance of Turks as *Gastarbeiters* in Germany can be cited.

I begin with the authors' general statements (those whose pages are numbered by Roman numerals are due to Sarna). L and R denote the columns of the pages; C stands for information contained in the author's *Chronology*. Finally, **L** means Lincoln.

# I

**R. J. Pulskamp, translator**

## **Correspondence of Leibniz and Jacob Bernoulli**

from Leibnizens *Math. Schriften*.

Hrsg. C. I. Gerhardt. 1. Abt., Bd. 3. Hille

Internet

[Pulskamp numbered his Notes consecutively. I added my own Notes and numbered them otherwise, but also in an understandable way.]

The correspondents discussed the theory of probability. In 1666, Leibniz, as a boy, had written [published] his treatise *Arte combinatoria*. He knew about the works of Pascal and Huygens; in fact, he knew Huygens personally. In 1676, he had visited Johannes Hudde in Amsterdam who, in turn, had in 1665 also carried on a correspondence with Huygens regarding his (Huygens') treatise. He was in possession of De Witt's treatise about annuities. In addition, Caspar Neumann, deacon at St. Mary Magdalene in Breslau, Silesia, who examined the records which had been kept in Breslau concerning age, sex, year and month of deaths for many years, sent his observations to Leibniz. Leibniz provided the stimulus for the discussion in a letter written in April 1703. It is likely that the version of the *Ars Conjectandi* which had come down to us was completed at this time. R. P. [An unjustified statement.]

### **Postscript to Letter 11 from Leibniz to Bernoulli**

April 1703, Berlin. *Math. Schr.*, p. 71

I hear that the subject of estimating probabilities, which I consider important, has been not a little developed by you. I would like someone to treat mathematically the various kinds of games (in which there are beautiful examples of this subject). This task would be both pleasant and useful and it would not be unworthy of you or any other very serious mathematician. I have seen some of your stated theses and only a few of their discussion. However, I would like to have them all<sup>1</sup>.

In his letter dated 3 October 1703 Jacob replied Leibniz, described his work and the main proposition. He informed Leibniz that twelve years previously his theorem had been shown to his brother Johann. Jacob asked Leibniz for legal situations which would help in completing the work and also about the treatise of De Witt concerning annuities. Indeed, it will be seen that Jacob repeatedly insisted on obtaining a copy of De Witt's work to secure statistics. [In the sequel, I will not copy some of these annotations]

### **Extract from Letter 12 from Jacob Bernoulli to Leibniz**

3 October 1703, Basel. *Math. Schr.*, pp. 77 – 78

I would gladly like to know, most honourable Sir, from whom you know that I have been working on the subject of estimating probabilities. It is true that for many years past I have taken much pleasure in explorations of this sort since I scarcely think that anyone

else has thought more than I about these matters. I even had in mind to write a tract about this subject, but I have often put it off for years at a time because my natural laziness, which the weakness of my health as an accomplice has increased so much more, caused me to approach the writing very feebly.

I often wish I had a secretary who could fully divine my thoughts when they were gently hinted to him and could put them down in writing. Nevertheless, I have already completed the larger part of a book, but with an important part missing. There, I show how to apply the principles of the art of estimation to civil, moral and economic affairs. I will finish the book after I have solved a singular problem<sup>12.1</sup> which has not a small commendation of difficulty and a very large commendation of usefulness and which remained before my brother for twelve years, although he, when asked about the same problem some time ago by Marquis de l'Hopital, concealed the truth because of his eagerness to devalue my research<sup>12.2</sup>.

I will briefly tell you what the problem is. It is a known fact that the probability of any event depends on the number of possible outcomes with which it can or cannot happen. And so, it occurred to me to ask why, for example, we know with how much greater is the probability that a seven rather than an eight will fall when we roll a pair of dice and why indeed do we *not* know how much more probable it is for a young man of twenty years to survive an old man of sixty years than vice versa. This is the point: we know the number of possible ways in which a seven and in which an eight fall when rolling dice, but we do not know the number of possible ways which prevail in summoning a young man to die before the old man and which prevail in summoning an old man to die before a young man.

I began to inquire whether what is hidden from us by chance *a priori* can at least be known *a posteriori* from an occurrence observed many times in similar cases, i. e., from an experiment performed on many pairs of young and old men. For had I observed it to happen that a young man outlived the respective old man in one thousand cases and to happen otherwise only five hundred times, I could have safely enough concluded that it is twice as probable that a young man outlives an old man as it is that the latter outlives the former.

Moreover, although, and this is amazing, even the stupidest man knows by some instinct of nature *per se* and by no previous instruction, that the more observations there are, the less danger there is in straying from the mark. It requires not at all an ordinary research to demonstrate this fact accurately and geometrically. But this is not all that I want. In addition it must be inquired whether the probability of an accurate ratio increases steadily as the number of observations grows, so that finally the probability that I have found the true rather than a false ratio exceeds any given probability, or whether each problem, so to speak, has an asymptote, – that is, whether I shall finally reach some level of probability beyond which I cannot be more certain that I have detected the true ratio. For if the latter be true, we will be done with our attempt at finding out the number of possible outcomes through experiments. If the former is true, we will

investigate the ratio between the numbers of possible outcomes *a posteriori* with as much certainty as if it were known to us *a priori*.

And I have found that the former [the last mentioned] condition is indeed the case. Whence I can now determine how many trials must be set up so that it will be a hundred, a thousand, ten thousand etc. times more probable (and finally, so that it will be morally certain) that the ratio between the numbers of possible outcomes which I obtain in this way is legitimate and genuine. The following suffices for practice in civil life: to formulate our conjectures in any situation that may occur no less scientifically than in a game of chance. I think that all the wisdom of a politician lies in this alone. I do not know, most honourable Sir, whether anything of substance appears to you to be in these speculations. In any case, you will make me grateful if you could supply me with any legal situations which you think could be usefully applied to these matters. Recently, I found that a certain tract which had been unknown to me was cited in the printed Monthly Excerpts of Hanover: Pensionarius De Wit's *von Subtiler Aufrechnung bei valoris der Leib-renten*. Perhaps he has something doing here; I would very much wish to obtain his source from somewhere.

#### **Extract from Letter 13 to Jacob Bernoulli from Leibniz**

3 December 1703<sup>2</sup>. Berlin. *Math. Schr.*, pp. 83 – 84

The estimation of probabilities is extremely useful, although in several political and legal situations there is not much need for fine calculations as there is for the accurate recapitulation of all the circumstances<sup>13.1</sup>. I remember learning *not* from your brother but somewhere else that these matters had been dealt with by you. When we estimate empirically by means of experiments the probabilities of successes, you ask whether a perfect estimation can be finally obtained in this manner. You write that you have found this to be so.

There appears to me a difficulty in this conclusion: that results which depend upon an infinite number of cases cannot be determined by a finite number of experiments<sup>13.2</sup>. Indeed, nature has her own habits, born from the return of causes, but only *in general*. And so, who will say whether a subsequent experiment will not stray somewhat from the rule of all the preceding experiments because of the very mutability of things? New diseases continually inundate the human race, but if you had performed as many experiments as you please on the nature of deaths, you have not on that account set up the boundaries of the world so that it cannot change in the future.

When we investigate the path of a comet from any number of observations, we suppose that it is either a conic curve or another kind of simple curve<sup>13.3</sup>. Given any number of points, an infinite number of curves can be found passing through them. Thus, I show the following. I postulate (and this can be demonstrated) that given any number of points, some regular curve can be found passing through these points. Let it be given that this curve has been found and call it A. Now, let another point be taken lying between the points given but outside of this curve; let a curve pass through this new point and the points given originally according to the above postulate. This curve

must be different from the first curve but nevertheless it passes through the same given points through which the first curve passes. And since a point can be varied an infinite number of times, there will also be an infinite number of these and other possible curves. Moreover, observed outcomes can be compared with these points where the fixed underlying outcomes or their estimates can be compared with the model curve. It can be added that, although a perfect estimation cannot be had empirically, an empirical estimate would nonetheless be useful and sufficient in practice.

The person who composed the monthly Germanic excerpts of Hanover has been at my house. Pensionarius De Wit's article is flimsy when he uses that estimation known from the equal possibility of similar outcomes and hence shows that the problem of resurrections can be clearly solved by considering the fate of the Batavians<sup>3</sup>. And therefore he has written in Flemish so that he might appear to be on the same footing with the commoner<sup>13,4</sup>.

#### **Extract from Letter 14 from Jacob Bernoulli to Leibniz**

20 April 1704. Basel. *Math. Schr.*, pp. 87 – 89

Various questions about Certainty, Resurrections, Endowed Agreements, Conjectures<sup>4</sup> and other matters show me that the subject of estimating probabilities in legal affairs requires not only the recapitulation of circumstances but also the same computations and calculation which we are accustomed to use in reckoning the outcomes of games of chance. I will show how to do this clearly for each situation. Moreover, the difficulty you found with my empirical method in determining the ratio between the numbers of possible outcomes requires more examples, not those in which it is impossible by any means to agree upon the numbers themselves, but rather those in which the numbers can be learned *a priori*. In addition, I said that I could, in these examples, provide for you a demonstration (which my brother saw twelve years ago and approved of). That you may really understand more clearly what I think, I give you an example.

I place in an urn several hidden pebbles, black and white, and the number of white ones is twice the number of black ones. But you do not know this ratio and wish to determine it by experiment. And so, you draw one pebble out after another (replacing the pebble which you drew out in each single choice before you draw the next one so that the number of pebbles in the urn is not diminished). You note whether you have picked a white or a black one. Assume that you have two estimates of the two-to-one ratio which are quite close to one another, but different, one being larger, the other smaller, say, 201:100 and 199:100. Then I claim that I can determine scientifically the necessary number of observations so that with ten, a hundred, a thousand etc. times more probability the ratio of the number of drawings in which you choose a white pebble will fall within, rather than beyond those limits of the two-to-one ratio. And so, I claim that you can be morally certain that the ratio obtained by experiment will come as close as you please to the true two-to-one ratio.

But if now in place of the urn you substitute the human body of an old or a young man, the body which contains the tinder of sicknesses

within itself as the urn contains pebbles, you can determine in the same way through observations how much nearer to death the one is than the other. It does no good to say that the number of sicknesses to which each is exposed is infinite, for let us grant this: It is nevertheless known that there are levels in infinity and that the ratio of one infinity to another is still a finite number and can be expressed either precisely or sufficiently precisely for practical use. If sicknesses are multiplied with the passage of time then in any case new observations must be set up. It is certain that he, who thinks that the investigations of our ancient forefathers concerning the end of life be settled by the daily customary observations made in London, Paris or elsewhere, will grossly err from the truth.

In this situation, the example of investigating the path of a comet from several of its observed positions is almost apropos<sup>14.1</sup>. I would never use it to demonstrate a proposition, although in a limited way I can find an application since it cannot be denied that if five points have been observed, all of which are perceived to lie along a parabola, the notion of a parabola will be stronger than if only four points had been observed. Indeed, although there are an infinite number of curves which may pass through five points, there is nevertheless beyond this infinite number, rather an infinitely times more infinite number of curves which may pass through only the first four points and not through the fifth point, all of which are excluded by this fifth observation.

And yet, I admit that every conjecture which is deduced by observations of this sort would be quite flimsy and uncertain if it is not conceded that the curve sought is one of the class of simple curves. This, indeed, seems quite correct to me, since we see everywhere that nature follows the simplest paths.

I perceive from your description that the Belgian tract of Jean De Witt contains such things which prove my point very well. And so I ask as strongly as possible that you, most honourable Sir, send to me your copy of the book on any convenient occasion, since I have sought it in vain in Amsterdam. I shall return it faithfully on the next market day in Frankfurt together with the fourth and fifth part of my publications concerning infinite series whose latter part has been recently published and circulated<sup>5</sup>.

The next letters of Leibniz, possibly two or three are lost.

### **Extract from Letter 15 from Jacob Bernoulli to Leibniz**

2 August 1704. Basel. *Math. Schr.*, p. 91

I shall shortly receive two copies of the *Histoire de l'Académie des Sciences* [Paris] for the year 1701 from Father Varignon<sup>15.1</sup>. They must be sent to you and to my brother and I will arrange that the fourth and fifth parts of my theses on Infinite Series<sup>15.2</sup> be added for you. Conversely, I am expecting again from you the composition of the Pensioner De Witt on this market day. To that, if only you are able to add what you have formerly written concerning agreements. I would also like that you make available to me any example of the conditional legacy. Likewise, you may show by an example anything

you understand by way of annuities which are constituted on many lives. For not at all have I applied my mind according to the aforesaid to study in the judicial matters.

We can determine the ratio between the numbers of deaths although infinite by finite experiments not precisely but what amounts in practice sufficiently precise for constantly approaching nearer until the error becomes insensible with respect to which indeed it is common in Geometry itself. Thus, the ratio of the diameter to the circumference cannot be determined precisely except through the Cyclic numbers of Ludolphus<sup>6</sup> continued into infinitely, but it is nevertheless fixed by Archimedes within the ratios 7:22 and 71:223 sufficiently constricted to use. I exhibit a specimen of the art of conjecture in some games of chance, particularly regarding games of tennis which I treat in detail. But in the majority of card games I do not advance much and even less in games of draughts on account of the immense variety of combinations of many repeated throws of game stones they are able to undertake<sup>15.3</sup>.

#### **Extract from Letter 16 from Jacob Bernoulli to Leibniz**

15 October 1704. Basel. *Math. Schr.*, p. 92

Seeing what I understand from your last letters, my response to your previous letters has not been delivered to you, I send its copy with this. What concerns Mr. Hermann, he will respond to you himself. On this market day I vainly expected your Mr. De Wit treatise. Mr. Mencke perhaps will be able to be commissioned at the moment of the market day in Leipzig as intermediary of merchants.

#### **Extracts from Letter 17 to Jacob Bernoulli from Leibniz**

28 November<sup>7</sup> 1704. Berlin. *Math. Schr.*, pp. 93 – 94

The dissertation of Pensioner Witt, or rather the printed paper concerning life annuities, reasonably brief, certainly exists among my books and I wished to send it to you, but I have not yet found it. I shall nevertheless surrender that work when found. It is hidden somewhere at home. Besides, it contains nothing which can be very new to you.

My double dissertation on the Conditions was printed by the Academy of Leipzig<sup>17.1</sup>, if I remember well, in 1665. Two years after, corrected, it was refused, just as certain others of my small juridical reflections through that of Nuremberg where I had allowed that it may go in peregrinations to Altorf but the copies were lost and it is with difficulty that thereafter I have obtained one by chance through a friend. I intend some day to prepare a new edition.

In some insufficiently collected things (undoubtedly because of our [restricted] ability) there is no certainty with increased data just as with new years being added to the observations of death we approach nearer to the mean truth of the whole, but when continued truth is always approached in series of the Ludolphine kind. In games of pure reason (as in chess and ramparts<sup>8</sup>) or in those depending on chance as much as on reason as in cards which the Spaniards call *Hombre* or games which in our ordinary life we call *Verkehren*, it is not easy to calculate [the expectations of the players]. Whence we see the clever

players decide by considerations just as in military or medical matters. This is more sagacious and excel profound studies.

#### **Extracts from Letter 18 from Jacob Bernoulli to Leibniz**

28 February 1705. Basel. *Math. Schr.*, pp. 95 – 98

You will receive immediately from Mr. Varignon the *Histoire de l'Académie des Sciences de Paris* together with the fourth and fifth parts of the [of my] Propositions on series, unless you have already received them by chance (?). Once more I beg you to remember to pass to me the treatise of Mr. De Witt, if only it fell into your hands. No matter what it contains, it cannot be completely new to me. Just the same, if only you deem worthy to give me whatever you have published at some time, it will always be highly desirable to me. I have nothing of yours except *De arte combinatoria* and new *Hypothetical Physics*<sup>18.1</sup>.

Because it considers plausibility and in particular the evidence of an enlarged number of observations [The incomprehensible end of the phrase apparently mentions the pertinent future publication of the *Ars Conjectandi*.]

#### **Extracts from Letter 19 of Leibniz to Jacob Bernoulli**

April 1705. Hanover. *Math. Schr.*, pp. 98 – 103

I count for nothing those of my writings which you have, *De arte combinatoria* and *Hypothetical Physics*<sup>19.1</sup>. Indeed, it is almost of naivety composed during my first youth and published in 1666 and, I think, in 1670. In truth, they are diverse arguments of philosophy and mathesis [mathematics] that I have extracted from my journal.

I was still prevented to seek adequately the writing of Pensioner De Witt among my books. Nevertheless, I do not doubt that I will at last discover it whenever there will be time. Bur scarcely anything new will occur to you since it is based on the same foundation as anything else. Not only some learned men, but also Pascal in his *Arithmetic Triangle* and Huygens in the dissertation on chance had used that foundation. [Probably an error in translation is involved here. And (an additional remark): Graunt and Halley were never mentioned in the correspondence.] Just the same, the arithmetic mean may be supposed among equally uncertain [corrupted by error] events [observations, measurements]. Countrymen now enjoy the same foundation when they estimate the prices of farms, just as managers of monetary affairs when the principle of the mean establishes the revenues of prefectures for an appeared contractor.

On 16 August 1705 Jacob Bernoulli died. We close with a letter in which Leibniz errs in thinking that in a cast of two dice twelve is as likely as eleven. He also takes undeserved credit for engaging Bernoulli in the pursuit of the study of probability.

#### **Extract of a letter from Leibniz to Louis Bourquet**<sup>last letter 1</sup>

22 March 1714. Vienna. *Phil. Schr.*, Bd. 3, pp. 569 – 570

The art of conjecturing is founded on that which is more or less easy, or else more or less feasible, because the Latin *facilis* derived from *faciendo* wishes to say *feasible* word for word. For example,

with two dice it is as feasible to cast twelve points as to cast eleven for both can be had only in one way. But it is three times more feasible to cast seven because it can be made up by casting 6 and 1; 5 and 2; and 4 and 3, and one combination is here as feasible as the other<sup>last letter 2</sup>.

The Chevalier de Méré (author of the book of *Agréments*) was the first who gave occasion to these meditations that Pascal, Fermat and Huygens pursued. The Pensioner De Witt and Hudde have also worked there. The late Bernoulli has cultivated this matter on my exhortations. One regards the probabilities *a posteriori* by experience, and one must have recourse to the errors of reasons *a priori*. For example, it is equally probable that the infant who must be born is a boy or girl because the number of boys and of girls is found very nearly equal in this world. One is able to say that that which is done the most or the least is also the most or the least feasible in the present state of things, putting all the considerations together which must unite in the production of a fact.

### Notes by the Translator

1. Translation by Bing Sung. Translations from James Bernoulli. *Techn. Rept* No. 2. Harvard Univ., Dept. of statistics, 1966. [I glanced at this translation and call it a layman's and extremely unfortunate work. But perhaps Pulskamp selected some fairly good pieces or somewhat improved them. O. S.]
2. Sylla claims that this letter is dated 26 November 1703. [Pulskamp provided no reference. It is almost sure that he bore in mind Sylla (1998). O. S.]
3. Bing Sung: *Problem*: What is the possibility of resurrection? *Solution*: Look at the Proportion of Batavians who have been resurrected. (De Witt himself was a Batavian. (A Batavian is a Hollander. (Translator)
4. *De Assecurationibus, de Reditibus ad vitam, de Pactis datalibus. De Praesumptionibus*. Sylla renders this as *Insurance, Annuities, Dowry contracts and Presumptions*.
5. Translated by Bing Sung.
6. Rudolph von Ceulen computed  $\pi$  to 20 digits in a paper written in 1596 by extending the method of Archimedes. He ultimately computed  $\pi$  to 35 digits.
7. Sylla claims October.
8. The text reads *velut scaccorum et aggerum*. [This Note does not agree with the main text. *Ramparts* is the name of a game of old (now: a videogame). O. S.]

### Notes by O. S.

- 12.1. This phrase is self-contradictory.
- 12.2. In any case, Leibniz found out about the bad relations between the brothers Bernoulli from this letter. He became involved in their quarrels and was inclined to support Johann. Later Bernoulli (Bernoulli Jacob 1993, pp. 100 – 104) admonished Leibniz since he did not influence Johann to quit quarrelling.
- 13.1. Gauss (*Werke*, Bd. 12, pp. 201 – 204) stated that in applications of the theory of probability all circumstances ought to be allowed for. In the 19<sup>th</sup> century many of those authors who studied criminal statistics stated the same.
- 13.2. At the time and later, this statement justified God's *inscrutable ways*.
- 13.3. No other curve is possible. *Call it A*: an unnecessary addition which means that Leibniz had written somewhat carelessly. A *model curve* (end of letter) is likely a term invented by the Translator.
- 13.4. This is an arbitrary and wrong statement. De Witt compiled his document for the government and had to write in Flemish. It is now published (Jacob Bernoulli 1975, pp. 329 – 350) and its English translation comprises the second part of Hendriks (1852 – 1853).
- 14.1. See Note 13.3. Then, a curve which passes exactly through a number of given points reflects all the errors of observation and is therefore almost useless. And there exists a polynomial which approximates a given curve on a given interval

to any arbitrary  $\varepsilon$  (the Weierstrass theorem). Indeed, J. B. only thought about moral certainty (and Leibniz, at the end of his previous letter, agreed with this restriction).

**15.1.** I can only add that Varignon was taught in a Jesuit's college.

**15.2.** The five parts of the *Treatise on infinite series* were published in 1689 – 1704 and appeared in a German translation (Leipzig. 1909). The *Treatise* as a whole was appended in 1713 to J. B.'s *Ars Conjectandi*.

**15.3.** Tennis: apparently, one of its former varieties. Draughts: an obvious mistake (of the translator?): they do not involve any game stones.

**17.1.** Leibniz mentioned his dissertation *Ars combinatoria* for acquiring the highest university degree. *Conditions* is likely the translator's slip of pen instead of *combinations*. Academy of Leipzig is the present Sächsische Akademie der Wissenschaft. Altorf is now a commune in the north-east of France.

**18.1.** I cannot confirm the existence of that contribution.

**19.1.** I repeat note 18.1.

**Last letter 1.** Bourquet (1678 – 1742), naturalist and mathematician, see Jacob Bernoulli (1975, Index of names).

**Last letter 2.** Seven points occur not three, but six times (in addition, as 1 and 6, as 2 and 5 and as 3 and 4).

## Bibliography

**Bernoulli Jacob** (1975), *Werke*, Bd. 3. Basel.

--- (1993), *Briefwechsel*. Basel.

**Gini C.** (1946), Gedanken zum Theorem von Bernoulli. *Z. f. Schweiz. Volkswirtschaft u. Statistik*, 82. Jg., No. 5, pp. 401 – 413.

No other translations as compared with the present contribution but many comments.

**Hendriks F.** (1852 – 1853), Contributions to the history of insurance. *Assurance mag.*, vol. 2, pp. 121 – 150, 222 – 258; vol. 3, pp. 93 – 120.

**Kohli R.** (1975), Aus dem Briefwechsel zwischen Leibniz und J. B. In Bernoulli (1975, pp. 509 – 513).

No other translations as compared with the present contribution but many comments.

**Sylla Edith Dudley** (1998), The emergence of mathematical probability from the perspective of the Leibniz – J. B. correspondence. *Perspectives on science*, vol. 6, pp. 41 – 76.

The title of this paper is too promising.

**Wolf R.** (1858), Jacob Bernoulli aus Basel. *Biographien zur Kulturgeschichte der Schweiz*. 1. Cyclus. Zürich, pp. 133 – 166.

Describes the quarrels between the brothers Bernoulli.

## II

### On the history of the statistical method in biology

*Arch. hist. ex. sci.*, vol. 22, 1980, pp. 323 – 371

#### I. Introduction

As I see it, the statistical method in experimental science is primarily a method of reasoning based on the mathematical treatment of statistical data. I discuss the history of this method with special reference to the concept of randomness. The Biometric school, which was born at the close of the 19<sup>th</sup> century, aimed to develop methods to treat biological observations and to study statistical regularities in biology. It was thought that the amalgamation of this school with the “Continental” direction of demography and, at the same time, the penetration of the statistical method into other fields of science (for example, into meteorology) and its application in industry, created mathematical statistics. Now I rather believe that the English statisticians only stumbled across the discoveries of the Continental direction.

Anyway, there is a strong case for a separate study of the statistical method in biology up to the time of Galton, the immediate predecessor of the Biometric school, the more so since the development of genetics began only after the discovery, in 1900, of the previously unnoticed work of Mendel.

Elsewhere [124] I have discussed the formation and application (in particular, in biology and medicine) of the concepts of random and probable events, and I begin this article from the 18<sup>th</sup> century<sup>1</sup>. In accordance with the above-mentioned considerations my account includes and concludes with, the work of Darwin.

The main discussion takes up four sections, two of which (§§ 3 and 5) are devoted to the evolution of species while §§ 2 and 4 treat other problems. The fundamental nature of Darwin’s works, both within and beyond the bounds of the theory of evolution, has compelled me to discuss them separately, and exactly for this reason there are four main sections (with §§ 4 and 5 devoted to Darwin) rather than two.

Natural scientists and philosophers of the past have used at least three different interpretations of randomness [124, §§ 2.2, 8.1 and 9.1]:

1. (Randomness I). Chance is just ignorance of relevant causes.
2. (Randomness II). Chance is the result of a failure to attain a certain purpose (or lack of purpose whatsoever).
3. (Randomness III). Chance is an intersection of independent chains of determinate events.

Mathematically speaking, randomness was usually understood as  
(1) A random variable with a uniform (discrete or continuous) distribution.

(2) A mixture of relatively small random disturbances of the same order, i.e., “usually”, a random variable with a normal distribution.

(3) A “chaotic” variable with no law of distribution whatsoever.

I mentioned the first two cases [124, p. 140] but failed to notice the last one. All three will be met in the sequel. (My present discussion of randomness is [126a]).

Note, however, that unlike Poincaré (Ibidem, p. 100) biologists did not say that a chance event (and, “usually”, a uniform random variable) takes place when, in conditions of an unstable equilibrium, slight causes determine considerable effects.

I refer to Buffon only once. This outstanding natural scientist pronounced no statistical ideas in biology, and had to disguise his thoughts on the evolution of species [135].

I have not discussed either demography or moral statistics. At least the former is directly related to biology, but both of them should rather be separated from it. I added some information on medical statistics in my notes but, for the same reason, I have not treated this subject systematically.

My study seems important because it throws light on an unusual aspect of the history of biology. Also, my discussion (§ 4.7) of some problems solved by Darwin seems to be the first of its kind.

## **2. Various Statistical Problems before Darwin**

**2.1. Classification of life-forms in botany.** By the middle of the 18<sup>th</sup> century biologists had studied many thousands of species of all kinds of animals and plants and proposed various systems to classify them. The *méthodes naturelles* of classification [42, p. 29], which preserve “distances” between species, seemed most promising. Linné expressed a high opinion of these methods (Ibidem, p. 58), and Buffon and, especially, Adanson were their earnest partisans (p. 61). The latter classified plants [22, pp. cci – cccxiv] according to 65 (!) partial and therefore arbitrary systems, and strove to construct a natural method of classification in which similar species will be those considered alike in a maximal number of partial systems<sup>2</sup>.

Adanson [21, p. xi] formulated this aim in 1757, and added that it could be achieved only if *tous les objets* [species] were known. Aug. De Candolle, who possibly did not notice this remark, repeated it himself [42, p. 67], and noted that not all organs of an individual are equally important, so a natural classification must take this fact into account. Actually, De Candolle thus proposed to introduce expedient weights for (partial) distances between species in various systems of classification<sup>3</sup>.

De Candolle [42, p. 232] mentioned distances between species, thus dropping a further hint toward a construction of a mathematical model of the classification of species:

*La distance qui sépare chaque espèce, chaque genre, chaque tribu, chaque famille, peut être réellement calculée, sinon d'une manière absolue, au-moins d'après une méthode comparative.*

Adanson did not attain his goal. As it seems, he did not even leave any fragments of a natural classification of plants. Moreover, he would have been unable to cope with computational hindrances, let alone difficulties in the essence of the problem itself. However, Adanson [22] minutely discussed the history of classification of plants, and classified them himself according to partial systems (see above). He also compiled a questionnaire the answers to which were necessary for the completion of the partial classification of plants.

Incidentally, this is a good example of detecting imperfections just by a compilation of relevant statistical summaries.

Adanson sought to answer a question which pertains to multivariate statistics [127, p. 472]. He [22, p. cc] also compared botany with mathematics:

*Nous croions même lui [botany] trouver un rapport immédiat avec la Géométrie: elle a cela de commun avec elle, qu'elle ne distingue les Plantes que par leurs rapports de quantité, soit numérique, ou discrete, soit continue, qui nous donne l'étendue de leur surface ou leur grandeur, leur figure, leur solidité.*

Specialists in many branches of science could have offered similar remarks.

**2.2. Empirical laws.** The first empirical law applied in biology (in botany) was likely the so-called law of the sums of temperatures due to Réaumur [115, pp. 558 – 559; 48, p. 424]. According to it, leaves, flowers and fruits come out on plants of a given species after the sum of mean daily temperatures attains certain values (which also depend on a number of other factors).

The law of the sums of temperatures held its ground at least until 1875 [41]. In particular, Adanson [22, p. 87] applied it to compile a calendar for the appearance of leaves on plants of various species for the vicinity of Paris<sup>4</sup>.

Aug. De Candolle [43, t. 1, pp. 432 – 434] compared botanical observations which lasted a few years with results calculated according to this law. He (Ibidem, t. 2, p. 476) also criticized the law of the sums of temperatures, noting its obvious uncertainty. Thus, he wrote, the initial moment for recording the temperatures cannot be chosen arbitrarily, negative temperatures should not be taken into account, the notion of the mean daily temperature must be specified, etc.

Alph. De Candolle [39, t. 1; 41, pp. 13 – 20] continued the work of his father. Referring to a similarity between the action of temperature and the nature of kinetic energy, Quetelet [110, p. 242] proposed to replace the sums of temperatures by the sums of their squares. He calculated the date of the appearance of leaves on lilacs (presumably for the vicinity of Brussels) according to both laws, compared his figures with actual observations, and pronounced himself in favour of the new law.

But even he, to say nothing of the De Candolles or the savants before them (Réaumur, Cotte), proposed no definite procedure to calculate parameters of empirical formulas or to compare competing laws.

Quetelet applied his new law once more in 1849, and then in 1852. The first utterance on empirical laws and, indirectly, on the statistical method in general, seems to be due to Cuvier [54, p. 67] who noted that in case of need

*Nous devons suppléer au défaut de la théorie par le moyen de l'observation; elle nous sert à établir des lois empiriques, qui deviennent presque aussi certaines que les lois rationnelles, quand elles reposent sur des observations assez répétées<sup>5</sup>.*

A similar pronouncement is due to Aug. De Candolle [43, t. 2, p. 983]<sup>6</sup>:

*Quand on aura déterminé un grand nombre de fois, et dans des circonstances différentes, l'accroissement annuel des individus d'une même espèce, on pourra obtenir une moyenne de ces accroissemens, et alors la simple connaissance de la circonférence d'un arbre suffira pour connaître approximativement son âge, non pour les arbres jeunes où les irrégularités sont trop grandes, mais pour ceux qui passent un siècle, par exemple.*

De Candolle's work is a mine full of botanical data compiled by various scholars, himself included. Thus, he furnished information on the consumption of oxygen by flowers and leaves of plants in darkness (t. 2, p. 550), content of water and sugar in fruits (pp. 584 – 585), dates of germination of plants (pp. 646 – 647). In the third volume of his book De Candolle studied the influence of external forces (light, water, heat, atmospheric electricity) upon plants. Here he formulated a large number of empirical laws although hardly a single one of them was substantiated by observations<sup>7</sup>.

**2.3. Compilation of statistical data.** Camper [37, p. 153] published a comparative table of some body measurements of nine species of mammals. Thus he applied an embryo of the statistical method<sup>8</sup>.

In 1832 the British Association for the Advancement of Science established a statistical section. A permanent commission of this

section was founded in 1833 with Babbage, a fervent collector of all kinds of statistical data<sup>9</sup>, as its chairman.

Biology did not escape his attention [27, p. 376]:

*I took every opportunity of counting the number of the pulsations and of the breathings of various animals ... at another period ... [I generalized] the subject of inquiry, and [printed] a skeleton form [of a questionnaire] for the constants of the class mammalia (where?). It was reprinted by the British Association at Cambridge in 1833 [25] [actually in 1834], and also at Brussels in the Travaux du Congrès général de statistique ... 1853.*

In a more general work Babbage [26, p. 294] noted that he first published the questionnaire on mammals in 1826. He (pp. 295 – 299) proposed there 142 questions about mammals, for example:

*Number of young at a birth, number of pulsations per minute whilst the animal is in repose ..., temperature, average duration of life, proportion of males to females produced.*

Babbage also included questions concerning man, such as quantity of air consumed per hour, quantity of food necessary for daily support, average proportion of sickness amongst working classes. Lastly, Babbage listed questions on the geographical distribution of animals and plants<sup>10</sup>.

**2.4. Fisheries and cattle breeding: the statistical aspect.** Practical needs rather than Babbage's programme led to the initiation of statistical research in biology. The compilation and simple processing of statistical data relating to industry, commerce and agriculture dates back to the beginning of the 19<sup>th</sup> century, notably to the publication during 1821 – 1829 of the *Recherches* [61].

The first large-scale statistical study which concerned both economic activities of man and biology seems to be Baer's exploration of fisheries in Russia<sup>11</sup>. In particular, he and his associates (in the first place, N. Ya. Danilevsky) published nine volumes of *Untersuchungen* on fisheries [29].

Baer did not introduce any new statistical methods, and did not even use graphical devices to illustrate his subject. But, during a relatively short period, he compiled most important data concerning fisheries and put forward specific proposals, see for example volumes 2 and 5 of the *Untersuchungen* (1860 and 1863).

According to Valt's likely opinion [131, pp. 107 and 110] just these and similar researches [28] directed Baer towards theoretical problems in animal ecology<sup>12</sup>.

Cattle breeding became the scene of Pasteur's vast experiment [105]. He tested the effect of his vaccine against anthrax on many thousands of animals. The results of the experiment were indeed brilliant<sup>13</sup>, he did not have to worry about treating them

mathematically. However, Pasteur urged that scrupulous statistical studies be carried out. Thus, suspecting that earthworms might be responsible for the spread of epizootic, he [104, p. 262] put forward the following desiderata:

*Il serait à désirer qu'une statistique soignée mit en correspondance dans les divers pays les localités à charbon et celles à vers de terre.*

## **2.5. Geography of plants**

*2.5.1. Humboldt.* The compilation of biological or, rather, botanical statistical data was a necessary component of the geography of plants. This discipline dates back to the beginning of the 19<sup>th</sup> century. It was created by Humboldt<sup>14</sup> whose works even beyond biology I describe here in some detail.

Experimental science has to do with the knowledge of mean values, of the necessary. This was Humboldt's guiding principle, at least from 1815 onward (see below), and it seems that he was the very first natural scientist to say so<sup>15</sup>. [Next came the study of the deviations from mean values.]

Humboldt [81, Bd. 1, p. 18] disapproved of those who dogmatize *Statt den mittleren Zustand zu erforschen, um welchen, bei der scheinbaren Ungebundenheit der Natur, alle Phänomene innerhalb enger Grenzen oscilliren.*

He (p. 82) repeated his reasoning on mean values:

*Bei allem Beweglichen und Veränderlichen im Raume sind mittlere Zahlenwerthe der letzte Zweck, ja der Ausdruck physischer Gesetze; sie zeigen uns das Stetige in dem Wechsel und in der Flucht der Erscheinungen; so ist, z. B. der Fortschritt der neueren messenden und wägenden Physik vorzugsweise nach Erlangung und Berichtigung der mittleren Werthe gewisser Größen bezeichnet.*

Humboldt's achievements in meteorology and geography were indeed significant. Thus, he [78] introduced both isotherms (the lines of equal mean temperatures) and mean heights of continents [80].

Concerning isotherms, Humboldt [74, t. 3, chap. 6, p. 66] again advanced an argument in favour of isolating local disturbances:

*C'est le grand problème de la météorologie de déterminer les inflexions de ces lignes, et de reconnoître, au milieu des modifications produites par des causes locales, les lois constantes de la distribution de la chaleur<sup>16</sup>.*

Humboldt's picturesque writings inspired Alph. De Candolle [39, t. 1, p. v], who testified: *A l'âge de dix-sept ans, mes lectures favorites étaient les ouvrages de M. de Humboldt.* But it is more significant to record Humboldt's influence on Darwin [18, vol. 1, p. 305, 1.1845; vol. 2, p. 422, 1. 1881, 19, vol. 2, p. 26, 1. 1881; 18, vol. 1, p. 387, 1. 1844, Ibidem, p. 403, 1. 1854]<sup>16a</sup>:

(1) *My whole course of life is due to having read and reread as a youth his "Personal narrative"*<sup>17</sup>.

(2) [Humboldt is] *the greatest scientific traveller who ever lived ... He was wonderful, more for his near approach to omniscience than for originality.*

(3) *He [Humboldt] was more remarkable for his astounding knowledge than for originality*<sup>18</sup>.

Darwin's statement in items (2) and (3) can be specified. Indeed, Humboldt did not develop the kinetic theory of gases, or propose the evolutionary theory, or establish the periodic law in chemistry. But he attained outstanding achievements and, for that matter, in almost every branch of natural science, only because he guided himself by the statistical method, always compiling and processing relevant statistical data.

My assertion is completely valid in regard to the geography of plants. Humboldt studied the distribution of plants by air temperature (and other factors). He did not use or invent any special 'statistical methods, so from a theoretical point of view his studies were rather simple. Humboldt himself [77, p. 228] remarked:

*Il en est de la géographie des plantes comme de la météorologie; les résultats de ces sciences sont si simples, que de tout temps on a eu des aperçus généraux: mais ce n'est que par des recherches laborieuses et après avoir réuni un grand nombre d'observations précises, que l'on a pu parvenir à des résultats numériques, et à la connaissance des modifications partielles qu'éprouve la loi de la distribution des formes.*

And indeed Humboldt [76, p viii] called statistics a *defficilis* (labour consuming) science:

*Just as there is political arithmetic, or, in latino-barbare (!), statistics, an extremely difficult science, most part of it being moreover conjectural, there is also Arithmetica botanica*<sup>19</sup>.

Humboldt [79, p. 431] predicted the forthcoming birth of zoogeography. Its founder, or at least cofounder, was Wallace [133], who described the distribution of land mammals and birds in various regions of the earth and enumerated the species of each genus of animals.

2.5.2. *Alphonse De Candolle*. He [39, t. 1] unfolded important ideas pertinent to the geography of plants. Its main goal, as he (p. xii) maintained, was to

*Montrer ce qui, dans la distribution actuelle des végétaux, peut s'expliquer par les conditions actuelles des climats et ce qui dépend des conditions antérieures.*

De Candolle did not explain his terminology and, what is obvious, he would have failed to introduce any statistical methods for the

estimation of the influence of separate factors. But at least, like Humboldt before him (§ 2.5.1), he clearly understood the essence of what was to be done.

De Candolle (Ibidem, p. xvi) also formulated a sound opinion on the statistical method in general:

*Il y a dans chaque science, art ou objet d'étude, une partie statistique, soit numérique. On la retrouve en agriculture, en médecine, dans toutes les branches de l'administration, dans les sciences physiques, naturelles, et jusque dans les sciences morales. Elle occupe une très grande place dans la géographie botanique. Pour moi, j'en conviens, j'aime les chiffres autant que d'autres les détestent; mais ce qui me plaît, ce n'est pas d'accumuler des chiffres, c'est de montrer à quel degré il est nécessaire de choisir convenablement les valeurs, de les discuter, en d'autres termes, de les subordonner aux lois de la logique et du bon sens<sup>20</sup>.*

De Candolle accomplished a large amount of important practical work in the geography of plants. Also, in the second volume of his book [39], he raised problems about the evolution of the distribution of plants and the origin of cultured plants<sup>21</sup>.

De Candolle [39, t. 1, p. 458] advocated the use of sampling. Following the spirit of his times he did not discuss the accuracy of this method, and did not refer to Laplace [125, § 2.5.5], who had applied sampling to estimate the population of France.

2.5.3. *Geography of plants and statistics in general.* Geography of plants was specifically mentioned at the second *Congrès international de statistique* [46, p. xxiii] which resolved to include

*Dans le programme de la prochaine reunion ... sous le titre de Statistique physique, une nouvelle catégorie de questions à examiner relatives à la climatologie, ..., à la géographie végétale, spontanée et agricole, aux phénomènes périodiques de la vie des plantes et des animaux.*

A questionnaire on *Statistique physique* was indeed published in the proceedings of the next congress [47, pp. 390 – 397] with some of the questions concerning the geography of plants and zoogeography. The title of the questionnaire seems rather curious:

*Eléments que les sciences naturelles doivent fournir à la statistique (!) pour que celle-ci puisse représenter de la manière la plus complète les diverses manifestations de la vie sociale.*

I also note that the subject of two reports delivered at this congress (pp. 524 and 530) was the influence of meteorological conditions on crop yields. But the story ended then and there. The time was not yet ripe for statistics to take over any branch of biology. Moreover, it was rather naive to expect natural (or any other) science to supply statisticians with any data. Cf. § 5.8.1.

**2.6. Anthropometry.** The writings of Quetelet [111; 113] contain many dozens of pages devoted to various measurements of the human body, of pulse and respiration rate, to comparisons of weight and height with age etc. The term *anthropometry* seems to have been coined by Humboldt and introduced by Quetelet [112, p. 671]:

*Je ne crains pas de suivre l'exemple ... de Humboldt; il m'offrit, par un mot emprunté à la langue grecque, le titre de mon ouvrage [113].*

In the book itself Quetelet (p. 410) remarked:

*Il se présente ici une science tout à fait nouvelle; je m'estime heureux d'être un des premiers à la saluer, et de pouvoir applaudir aux succès des savants qui voudront l'approfondir.*

Quetelet likely came to think about his *homme moyen* in anthropometric terms, then generalized the new concept to include men's moral and intellectual qualities.

In his notes published in 1846 – 1848 Quetelet described his anthropometric measurements of men of various races and asserted that all races belong to a single species. However, his observations were not numerous at all, he made no attempt to evaluate the reliability of his conclusions<sup>22</sup> and did not adduce any general biological arguments [9, pp 263 – 280]. Moreover, contrary to Darwin Quetelet (chap. 2) paid no attention to variations between human races and did not compare variations in man and woman. According to Darwin (§ 5.2), males (and men in particular) enjoy larger variations in body measurements, and had Quetelet been interested in the theory of evolution he would have checked this opinion. As it is, Quetelet [113, p. 181] made only one comparison of variations in man and woman and did not comment at all, although his results contradicted the views held by Darwin of whose opinion he however was likely unaware.

### **3. The Evolution of Species: Period before Darwin**

#### **3.1. The Eighteenth Century**

*3.1.1. Linné.* The problem of mutability of species dates back to 1719 [23, p. 31]. About 1744 Linné had begun doubting their invariability [23, p. 32; 70, pp. 150 – 151]. However, intraspecific variations were then considered unimportant. This was the opinion of Linné himself [97, pp. 140 and 171] who also remarked [98, § 306, p. 342 and § 158, p. 232] on the need to study varieties:

(1) *The great usefulness of varieties ... has made the knowledge of them necessary in common life. Otherwise varieties belong not to botanists as such, but so far as they should take care that the species be not unnecessarily multiplied or confounded.*

(2) *The varieties of plants are accidental changes, generally owing to the climate, soil, exposure, heat ... and by a change of soil &c. [They] are generally reduced to their proper species.*

3.1.2 Adanson. He [23, p. 48; 22, p. cxv] held a similar opinion on the insignificance of variations:

(1) *Ces écarts ont aussi leurs loix & leurs bornes: en effet, plus on observe, plus on se convainc que ces monstruosités & variations ont une certaine latitude, nécessaire sans doute pour l'équilibre des choses<sup>23</sup> après quoi elles rentrent dans l'ordre harmonique préétabli par la sagesse du Créateur.*

(2) *On appelle Variété la différence qui se trouve entre les individus de même Espèce, différence accidentelle & peu durable.*

But then, Adanson [23a, p. 60] recognized accidental monstrosities capable of distorting the species concerned:

*Outre les variations accidentelles son [man's] corps est encore sujet à des monstruosités accidentelles qui se perpétuent d'un certain point pendant un certain nombre de générations et qui tendent à dénaturer son espèce si elles ne rentrent pas après un temps limité dans l'ordre naturel.*

3.1.3. Kant. He [82, p. 446] supposed that

*Bringt die Beschaffenheit des Bodens ... im gleichen der Nahrung nach und nach einen erblichen Unterschied oder Schlag unter Tiere einerlei Stammes und Rasse.*

He (p. 450) also attached great importance to latent intrinsic dispositions:

*Diese Fürsorge der Natur bringt bei der Wanderung und Verpflanzung der Tiere und Gewächse dem Scheine nach neue Arten hervor, welche nichts anders als Abartungen und Rassen von derselben Gattung sind.*

However, Kant (p 451) did not believe in randomness:

*Der Zufall oder allgemeine mechanische Gesetze können solche Zusammenpassungen nicht hervorbringen. Daher müssen wir dergleichen gelegentliche Auswickelungen als vorgebildet ansehen.*

3.1.4. Erasm Darwin. He [57, p. 238] maintained that birds acquire one or another beak in accordance with their food. He likely supposed that this adaptation was brought about by design. At least his other reasoning are instructive in this connection:

(1) *The strongest and most active animal should (!) propagate the species, which should (!) thence become improved (Ibidem).*

(2) *All warm-blooded animals have arisen from one living filament which the great first cause endued with animality, with the power of acquiring new parts etc. (Ibidem, p. 240).*

3.1.5. Maupertuis. Being a versatile scholar (a mathematician and astronomer in the first place), he earnestly busied himself with biology. He was

*The first to apply the laws of probability to the study of heredity ... Virtually every idea of the Mendelian mechanism of heredity and the*

*classical Darwinian reasoning from natural selection and geographical isolation is combined [in his works], together with De Vries' theory of mutations as the origin of species.*

This estimation is due to Glass [69, p. 60] who does not add that, for all his achievements, Maupertuis did not originate the evolution theory or initiate genetics. Just the same, it was not Lucretius who discovered, let alone studied, the Brownian motion!

Maupertuis [99, pp. 120 – 121] offered a stochastic explanation for the action of heredity:

*Dans la liqueur séminale de chaque individu, les parties propres à former des traits semblables à ceux de cet individu sont celles qui d'ordinaire sont en plus grand nombre, et qui ont le plus d'affinité.*

*Les parties analogues à celles du père & de la mère étant les plus nombreuses, & celles qui ont le plus d'affinité, seront celles qui s'uniront le plus ordinairement: & elles formeront d'ordinaire des animaux semblables à ceux dont ils seront sortis.*

However (p. 109), a child may also resemble one of his forefathers, while now and then occur large deviations, as for example (p. 121) a white child born of black parents<sup>24</sup>. Lastly (p. 123), although

*Le fonds de toutes ces variétés se trouve dans les liqueurs séminales mêmes, Maupertuis does not exclude l'influence que le climat & les aliments peuvent y avoir.*

Later Maupertuis [100, p. 11] supposed that species were produced by *un destin aveugle* and that only

*Un petit nombre [of individuals] se trouvoit construit de manière que les parties de l'animal pouvoient satisfaire à ses besoins.*

In 1751 Maupertuis [101, p. 148\*] recognized the possibility of the appearance of new species due to mutations (Randomness II) if and when

*Les parties élémentaires n'auroient pas retenu l'ordre qu'elles tenoient dans les animaux pères & mères: chaque degré d'erreur auroit fait une nouvelle espèce.*

This is similar to an earlier pronouncement [99, p. 110]:

*La Nature contient le fonds de toutes ces variétés: mais le hazard ou art [in case of domestic animals] les mettent en oeuvre.*

Still, Maupertuis [101, p. 146] did not believe in the omnipotence of randomness: *une attraction uniforme & aveugle* cannot produce eyes and ears. For the explanation of their origin

*Il faut avoir recours à quelque principe d'intelligence, à quelque chose de semblable à ce que nous appellons desir, aversion, mémoire.*

Maupertuis evidently recognized randomness only in the sense of a uniform (*uniforme & aveugle*) random variable. As noted above, his explanation of heredity was stochastic; exactly, but did he himself

think so? I am doubtful since the randomness involved was not uniform.

Maupertuis also considered a special problem on the probabilities of rare events. Suppose [102, p. 277] one person in 20,000 is six-fingered. Then the probability that a son, a grandson, and a great grandson of this person are all six-fingered is insignificant. If such cases do occur, then polydactyly must be hereditary.

*Je veux bien croire*, Maupertuis (p. 276) adds,

*Que ces doigts surnuméraires dans leur première origine ne sont que des variétés accidentelles ... mais ces variétés une fois confirmées par un nombre suffisant de générations ... fondent des espèces; & c'est peut-être ainsi que toutes les espèces se sont multipliées.*

Possibly because of such reasoning Maupertuis [100, p. xii] thought that

*Un nombre infini de probabilités est une démonstration complète, et pour l'esprit humain la plus forte de toutes les démonstrations.*

Maupertuis did not specify how many sons, grandsons etc. the six-fingered person should have. And in any case problems of this type originated not later than in 1713 [122, § 5]. See also § 4.7.2.

**3.2. Lamarck.** In the 18<sup>th</sup> century the variability of species was hardly connected with random variations of individuals, while Maupertuis' thought-provoking idea on the stochastic origin of heredity (§ 3.1.5) remained at best a hypothesis.

Lamarck was the first to proclaim evolution of species as a main principle of life. His ideas, not sufficiently substantiated by relevant facts, were forgotten right up to the end of the 19<sup>th</sup> century. Also, Lamarck noted the dissipation of order and regularity as a whole [85, §§ 920, 982 and 813]:

(1) *La nature ne forme rien, elle détruit toujours.*

(2) *Toute substance composée tend naturellement à se détruire.*

(3) *Tous les efforts de la nature ... sont perpétuellement dirigés vers ce seul but; savoir, d'opérer la destruction des composés quels qu'ils soient, et de rendre aux éléments qui les constituent, la liberté et leurs qualités naturelles, dont ils sont dépourvus dans leur état de combinaison.*

It is hardly worth mentioning that Lamarck knew nothing about thermodynamics. See also § 5.5.

**3.2.1. Randomness and necessity.** LAMARCK [90, p. 607] recognized Randomness I:

*All kinds of motions and changes occurring in [some] parts of the universe are governed by invariable laws of different orders ... Inviolable order and concord always reign [in the universe], ... all observed facts without exception are the result ... of motion and laws. ... The word chance signifies only ignorance of causes. ... What we*

*suppose to be disturbances does not take place in nature and is nothing but facts concerning isolated objects whose self-preservation is absolutely incompatible with the general order and laws governing all motions.*

And further (p. 632):

*The aims of meteorology would have been absolutely useless, unreliable and groundless if there could exist any part of nature which ... would not obey invariable laws, if that, which is called chance, might be a reality<sup>25</sup>.*

At the same time Lamarck [91, p 169] asserted:

*La nature a deux moyens puissans et généraux, qu'elle emploie continuellement ... Ces moyens sont. 1. L'attraction universelle ...; 2. L'action répulsive des fluides subtils, mis en expansion; action qui, sans être jamais nulle, varie sans cesse dans chaque lieu, dans chaque temps, et qui modifie diversément l'état de rapprochement des molécules des corps<sup>26</sup>.*

*De l'équilibre entre ces deux forces opposées ... naissent ... les causes des tous les faits que nous observons, et particulièrement de ceux qui concernent l'existence des corps vivans<sup>27</sup>.*

Lamarck (p. 173) calls the second force *très-irrégulière*. And it is this (*chaotic?*) force that leads to spontaneous generation; see § 3.2.3.

*Une ... cause accidentelle et par conséquent variable, a traversé ça et là l'exécution de ce plan [des opérations de la nature] (p. 133), i. e., the lay-out of the tree of animal life. He repeated this idea on p. 161 and even pointed out (pp. 454 – 457) specific corruptions probably occasioned in the general plan by random forces.*

Lamarck came out in favour of science founded on experiment. One of his principles read [95, p. 84]:

*Toute connaissance qui n'est pas le produit réel d'observation ou de conséquences tirées de l'observation, est tout-à-fait sans fondement, et véritablement illusoire.*

Nevertheless, his physical and chemical works proved almost useless because of his passion for general speculations [103, pp. 85 – 87]<sup>28</sup>. This very passion also permeates Lamarck's reasoning on the action of random forces<sup>29</sup>. For all that, his conclusions on the random disturbances of order in the tree of animal life make sense. Also, Lamarck's clear assertion on the importance of randomness in nature, though obscured by its (far-fetched) connection with repulsion, was about half a century ahead of its time.

3.2.2. *Evolution of Species and Randomness*. Lamarck [89, pt. 1, chap. 7] recognized the variability of species, and supposed that their evolution follows a definite pattern, viz.:

Change of external circumstances – appearance of requirements – emergence of new habits and efforts (of course, only in animals) –

greater exercise of relevant organs – their development – hereditary changes<sup>30</sup>.

Adducing numerous examples, he thus showed that the evolution of species was a universal phenomenon. Regrettably, he [128, p. 77] sometimes used *careless expressions* and introduced *unsubstantiated conjectures*. One such conjecture concerned the emergence of new habits and efforts in animals (see above)<sup>31</sup>. This fact perhaps provoked Darwin's flat denial of his work [18, vol. 2, p. 10, 1. 1859]:

*I do not know what to think about [an unspecified work of Lamarck], but it appeared to me extremely poor; I got not a fact or idea from it.*

At the same time Darwin [5, p. 8] acknowledged Lamarck's *Eminent service of arousing attention to the probability of all change in the organic ... world being the result of law, and not of miraculous interposition.*

As far as conditions of life may change at random Lamarck explained the evolution of species by accidental causes. Also, he [91, p. 198] thought that varieties might be produced by habits

*Contractées, soit accidentellement, soit autrement. Ainsi, l'homme ... offre lui-même des variétés remarquables dans son espèce, et parmi elles il s'en trouve qui paraissent dues aux [these very causes].*

Somewhat later Lamarck [92, p. 450] mentioned varieties engendered by random causes with no reference to habits at all. Considered in general, Lamarck's views seem not to be thoroughly reasoned out. On the one hand, he supposed that random causes played an important role in the formation of higher taxonomic categories (§ 3.2.1); on the other hand, he held that these causes only partly contributed to the emergence of varieties.

3.2.3. *Spontaneous generation of life*. Harvey and biologists before him believed in spontaneous generation of some creatures [124, p. 116]. Scholars of the 19<sup>th</sup> century were more careful, they recognizing at most the generation of simplest organisms. Moreover, spontaneous generation became divorced from the action of randomness. Thus Bastian [33, p. 244] argued:

*The phrase 'spontaneous generation' should be rejected. The phenomena hitherto referred to under this name are no more 'spontaneous' than are any others which take place in accordance with natural laws.*

Lamarck likely was the last biologist to admit the generation of comparatively developed forms. He [89, p. 62] thought that simple forms

*Sont des produits directs des moyens et des facultés de la nature.* This, he (p. 82) continued, is possibly the case with

*Les vers intestins, ... les moisissures, les champignons divers,*

*les lichens mêmes.*

And it is not any more possible to doubt spontaneous generation at least *à l'extrémité antérieure du règne végétal et du règne animal* [91, p. 179]. Lastly, the generation is due to irregular (i. e, random, see § 3.2.1) forces (p. 175).

Darwin [5, p. 118, see also p. 8] offered a reasonable explanation of Lamarck's attitude towards spontaneous generation:

*Why have not the more highly developed forms supplanted and exterminated the lower? Lamarck, who believed in an innate and inevitable tendency towards perfection in all organic beings, seems to have felt this difficulty so strongly, that he was led to suppose that new and simple forms are continually being produced by spontaneous generation.*

### **3.3. From Lamarck to Darwin**

*3.3.1. Estienne Geoffroy Saint-Hilaire.* Adhering to an old tradition and mentioning Pascal, Geoffroy Saint-Hilaire [64, p. 217; 65, p. 77, note] thought that variations in organisms are caused by external conditions<sup>32</sup>.

He also believed [62, t. 1, p 490, t 2, pp. 75 and 77] that deformities take place because of random causes acting upon the embryo. Probably he supposed that random causes manifest themselves only now and then. Indeed:

(1) Quoting Virey, Geoffroy [62, t. 2, p. 121] explained the origin of new species in the same way as the emergence of deformities:

*L'étude des monstres sera donc, pour le physiologiste et pour le philosophe, la recherche des procédés par lesquels la nature opère la génération des espèces.*

(2) Partly using published data [61] and partly drawing on information supplied by Fourier [61], Geoffroy [62, t. 2, p. 506] studied the birth-rate in Paris and remarked that the relative number of monsters among illegitimate children (whose mothers likely endured moral suffering and found themselves in strained circumstances) was rather small. However, he did not adduce any quantitative estimates.

He [63, pp. 290 – 291; 64 p. 227] attempted to prove that monster chicks hatch out of eggs kept in an incubator in wrong positions.

*Ayant opéré sur des masses, he wrote in the second instance, j'ai toujours obtenu le produit cherché. ... Nous nous trouvons aujourd'hui avoir fait trop de progrès dans la théorie des calculs de probabilités pour que cette argumentation [against my experimental results being representative] jouisse de quelque faveur.*

Mentioning probability theory, Geoffroy could have quoted Laplace, who, for example, proved that the difference between the ratios of male and female births in Paris and London was not accidental [125, § 2.5.3].

And it is somewhat strange that Geoffroy did not evaluate his final results or even publish his observations.

He [66, p. 645 and 646] declared his faith in the theory once more:

*Le calcul des probabilités a presque toujours été négligé des naturalistes; et cependant des applications nombreuses pourraient en être faites a presque toutes les branches des sciences naturelles. ... Si nous exceptons la minéralogie, le calcul des probabilités est, comme on l'a vu, la seule branche des sciences mathématiques qui soit applicable aux sciences naturelles.*

This opinion is rather interesting. For biology, the theory of probability (and mathematical statistics) had become the most important mathematical discipline, at least after Darwin and perhaps to this very day. Even certain branches of mineralogy use statistics.

3.3.2. *Goethe*. In 1790 he [71, § 5] distinguished *dreierlei Art* of metamorphoses in plants: *regelmäßig, unregelmäßig und zufällig*. Concerning accidental metamorphoses he (§ 8) said that they are *von außen, besonders durch Insekten gewirkt wird*. Lastly, Goethe (§ 30) noted that the moment when flowers appear on plants depends on external conditions<sup>33</sup>.

None of these considerations are either very interesting or new. Much more important is his later opinion [72, p. 120]:

*Die uns umgebenden Pflanzenformen seien nicht ursprünglich determiniert und festgestellt, ihnen sei vielmehr, bei einer eigensinnigen, generischen und spezifischen Hartnäckigkeit, eine glückliche Mobilität und Biagsamkeit verliehen, um in so viele Bedingungen, die über dem Erdkreis auf sie einwirken, sich zu fügen und darnach bilden und umbilden zu können.*

3.3.3. *Comte*. Other authors, especially Comte [45, No. 40, pp. 234 and 278; No. 42, pp. 444 and 446] put forward similar ideas on the importance of external conditions for the evolution of species. In the last instance Comte wrote:

*Chaque organisme déterminé est en relation nécessaire avec une système également déterminé de circonstances extérieures. ... il s'agit d'un équilibre mutuel entre deux puissances hétérogènes et indépendantes. Si l'on conçoit que tous les organismes possibles soient successivement placés ... dans tous les milieux imaginables, la plupart de ces organismes finiront ... par disparaître, pour ne laisser subsister que ceux qui pouvaient satisfaire aux lois générales de cet équilibre fondamental: c'est probablement d'après une suite d'éliminations analogiques que l'harmonie biologique a dû s'établir ... sur notre planète<sup>34</sup>.*

Comte seems to introduce intersections of two independent chains of random events (*équilibre entre deux puissances*). See a similar though less pronounced opinion of Goethe in § 3.3.2.

3.3.4. *Isidore Geoffroy Saint-Hilaire*. According to him [67, p. 422] variations between men

*S'expliquent, mais en partie, seulement, par l'influence du climat, du régime diététique et du genre de vie.*

Later he [68, pp. 430 – 436] formulated more general ideas:

*Les caractères des espèces ne sont ni absolument fixes, ni surtout indéfiniment variables. ... Ils sont fixes pour chaque espèce, tout qu'elle se perpétue au milieu des mêmes circonstances. Ils se modifient si les circonstances ambiantes viennent à changer* (p. 430).

Noting (Ibidem) that there exist *deux forces contraires: modificatrice, conservatrice*, Geoffroy (p. 436) concludes that species should be considered *relativement au monde actuel*.

The source which I have just referred to [68] is appropriately entitled *La théorie de la variabilité limitée de l'espèce*.

3.3.5. *Chambers*. He [44, p. 161] emphasized the importance of external conditions in the formation of species. He (p. 242) recognized statistical regularity in the behaviour of man and (pp. 261 – 263) described the causes of criminality, including the influence exerted by the way of life and social conditions. In connection with his reasoning on American languages Chambers (p. 218) referred to the problem due to Young [83] on the probability for the coincidence of words in two different languages. Regrettably he did not speak out in favour of stochastic considerations in biology.

3.3.6. *Roullier*. In some of his articles written in 1847 – 1856 he [119] adduced examples of hereditary changes in animals and plants brought about by external forces. Reflecting the prevailing belief of the day, he [117, p. 59] remarked that the *powerful influence exerted by external physical conditions is obvious and nowadays no one doubts it anymore*<sup>35</sup>.

3.3.7. *Cournot*. Like Comte (§ 3.3.3), he [51, p. 119] supposed that changes in external conditions brought about elimination of the less fit individuals. At the same time he thought that random causes were only secondary.

Thus, explaining the generally observed random differences between plants of the same species growing side by side, Cournot [51, pp. 126 – 127] concluded:

*Cet exemple peut donner l'idée de la part du hasard et de la multiplication indéfinie des combinaisons fortuites dans l'établissement de l'ordre final et des harmonies qui s'y remarquent. Mais il y a des limites à cette part du hasard, comme à la part des influences que la culture développe: le plus grand rôle dans la constitution de l'harmonie finale reste toujours à la force génératrice et plastique primitivement attaché au type originel, en vertu d'une harmonie préexistante.*

Two years after Darwin published his *Origin of Species* Cournot [52, p. 362] again pointed out that the differences between classes of animals (*les profondes divisions*) are due to a *plan supérieur*. He (p. 355) also compared *l'idée de type*<sup>36</sup> with a plate used for printing engravings:

*Une planche s'use par le tirage, et les épreuves du dernier tirage, qui ne diffèrent pas sensiblement les unes des autres, diffèrent sensiblement des épreuves du premier tirage.*

In other words, Cournot supposed that species are variable. Lastly, he (p. 416) offered a definition of a species calling it a

*Fond commun sur lequel brochent en quelque sorte les accidents du développement individuel, de la génération, de la transmission héréditaire, de manière à constituer des variétés individuelles, sporadiques, dont la science n'a point à s'occuper, et des variétés héréditaires ou des races plus ou moins anciennes, plus ou moins durables, mais dont l'ancienneté, la persistance et la durée ne sont point comparables à celles des espèces.*

Only in 1872 Cournot [53, p. 155] mentioned Darwin, raising two objections against his theory. First of all, he argued, biological phenomena are too diverse to be reduced to the struggle for food. Unconvincingly he adds:

*On dit d'un être raisonnable qu'il doit manger pour vivre et non pas vivre pour manger.*

Cournot's second and main objection concerns the absence or rarity of intermediate life-forms, an objection well known to Darwin himself [5, p. 157]. It is indeed strange that Cournot did not add any stochastic arguments.

Earlier in life he had published a treatise on probability theory with applications to natural science (except biology) [50] and, to say the least, he might have remarked on the need to substantiate Darwin's theory by certain stochastic considerations. I also note that Cournot was hardly satisfied with the old definitions of randomness.

#### **4. Various Statistical Problems: Darwin**

Here, I discuss Darwin's specific researches insofar as they relate to the statistical method. It seems opportune to remark right now that he explained the evolution of species by the accumulation of random variations (§§ 5.2 – 5.4) and that in biology the statistical method became established for good as a result of Darwin's work (§ 5.8).

Statistical tables and summaries with qualitative explanations of numerical results occur in a number of Darwin's writings [6, pp. 33, 48, 55, 90; 7, pp. 24 – 32; 8, vol. 1, pp. 183 and 285; 13, chapters 1 – 4; 15, pp. 49 – 55 and 88]<sup>37</sup>, but what I emphasize is that a good many of his works are permeated with statistics and that he used the

statistical method either manifestly or implicitly. In this connection Darwin's confession [18, vol. 1, p. 411, 1. 1855] is noteworthy: *I have no faith in anything short of actual measurement and the Rule of Three.*

**4.1. Facts and experiments.** The statistical method bears a direct relation to the methodology of experimental science [124, p. 121], and I shall therefore say a few words about Darwin the experimenter. As put on record by his son Francis [18, vol. 1, p. 126],

*The love of experiment was very strong in him, and I can remember the way he would say, "I shan't be easy till I have tried it."*

And Darwin himself [16, p. 141] testified:

*My industry has been nearly as great as it could have been in the observation and collection of facts.*

It was not without reason that Darwin (Ibidem, p. 49) remarked:

*On reading [the Zoonomia, a book written by his grandfather, Erasm] a second time ... I was much disappointed, the proportion of speculation being so large to the facts given.*

Indeed, in 1859 Erasm informed Charles [18, vol. 2, p. 28]:

*The à priori reasoning is so entirely satisfactory to me that if the facts won't fit in, why, so much the worse for the facts is my feeling.*

I note also that in 1867 Darwin [19, vol. 2, p. 6] had *come not to care at all for general beliefs without the special facts.*

Darwin's attitude towards facts as described above is confirmed in deed. Having studied the evolution of species for about twenty years, he [18, vol. 1, chap. 13]<sup>38</sup> published his main work [5] with extreme reluctance, considering it insufficiently corroborated by facts. It is remarkable that during his long later life Darwin did not essentially correct his theory; evidently he had been able to arrive at final conclusions (final, of course, only in the subjective sense) by drawing on rather incomplete information.

Darwin's experiments mainly had to do with the study of movements and pollination of plants and, also, with the life of earthworms (§ 4.7). As for the relevant statistical summaries and tables, I have referred to them in § 4.

According to Bell [35, p. 11], Darwin was not a first-class experimenter. For, my part, I might add that recording movements of plants Darwin [14, p. 435] sometimes restricted himself to a qualitative description, not even adding a slightest hint that a mathematical study of the respective curves might be advisable. But, to say the least, experiments were the life and soul of all of his specific researches.

I note a special measurement [5, p. 303]:

*I have estimated the areas [of regions in North America] by cutting out and weighing the paper [of H. D. Roger's beautiful map]. At the*

time, planimeters were not yet in general use. Also, what Darwin needed was approximate estimates of ratios of areas so that his method seems justified. But of course only if the *beautiful map* was compiled in an area-preserving projection.

**4.2. Records of observations.** Darwin always recorded his observations. Once he [11, p. 212] noted that he had *not picked measurements*<sup>39</sup> *out of a series, but [had] used all.* Elsewhere he [13, p. 150] pointed out:

*The erroneous numbers, however, are entered in the tables, that it may not be supposed that I have in any one instance tampered with the results.*

Yet another instructive example [16, p. 143] concerns an article which later proved to have been based on fabricated facts:

*A Belgian author ... stated that he had interbred rabbits in the closest manner without the least injurious effects ... I could not avoid feeling doubts, I hardly know why, except that there were no accidents of any kind, and my experience in breeding animals made me think this improbable.*

Fisher [59, § 18, p. 38] checked Darwin's experiments on pollination of plants (§ 4.7.3) and concluded that a corroboration of his inference was possible only because Darwin had published all relevant observations rather than mean results.

I note that the sacred duty of recording all observations, even rejected ones, came to be established in astronomy only by the middle of the 19<sup>th</sup> century [123, pp. 111 – 113].

**4.3. A null hypothesis in medical statistics.** Obviously bearing in mind an English translation of the *Lettres* [110], Darwin [18, vol. 1, p. 341, 1. 1850] noted:

*How true is a remark I saw the other day by Quetelet, in respect to evidence of curative processes, viz., that no one knows in disease what is the simple result of nothing being done, as a standard with which to compare homoeopathy, and all other such things.*

The original French version of the *Lettres* (pp. 348 – 349) contains the following passage:

*Pour juger avec connaissance de cause des avantages que peut présenter la thérapeutique, il faudrait commencer par rechercher ce que deviendrait l'homme affecté de telle maladie et abandonné aux seules forces de la nature.*

Disregarding the difference between homoeopathy and therapeutics in general, I remark that Darwin formulated the idea of a (still non-mathematical) null hypothesis more distinctly than Quetelet<sup>40</sup>.

**4.4. Sampling.** Considering the evolution of organisms Darwin [8, vol. 2, p. 401; 5, p. 29] put forward ideas which bear a direct relation to sampling:

*As man has domesticated so many animals and plants ... and as he certainly did not choose ... those species which would vary most, we may infer that all natural species, if exposed to analogous conditions, would, on an average, vary to the same degree.*

Darwin treats domestic animals and plants as a sample of organisms in general; evolution is a random variable and the mean evolution of the sample is approximately equal to the respective quantity over the patent population.

**4.5. Statistics related to man.** Darwin was always interested in statistical data related to biology and, in particular, to man<sup>41</sup>. Referring *to our highest authority on such questions, W. Farr*<sup>42</sup>, he [9, p. 214] adduced demographic data on mortality and birth-rate.

From Quetelet's works Darwin knew that variations of body measurements in man are of a statistical nature (§ 5.2) and his reference [9, p. 39] to an experimental study of the variations in the circulatory system of man seems quite appropriate:

*The chief arteries so frequently run in abnormal courses, that it has been found useful for surgical purposes to calculate from 1040 corpses how often each course prevails.*

**4.5.1. Collecting Statistical Data.** Statistical data occur in quite a few of Darwin's writings. In some cases he perhaps thought of using these data in due time; in other instances he may have collected data following his inclination towards a comprehensive description of biological phenomena.

**4.5.1.1. Susceptibility to diseases.** Darwin [9, p. 301; 18, vol. 2, p. 272, 1. 1864] compiled and circulated a questionnaire on the *relation between the colour of ... hair and liability to diseases of tropical countries*. He received no returns at all and even his questions remain unknown.

**4.5.1.2. Mental development of infants.** Darwin [19, vol. 2, p. 54, 1. 1881] spoke out in favour of a statistical study (a topical issue of his times) of the problem of mental development of coloured children and at the same time formulated his attitude towards statistics:

*I believe that isolated observations will add but little to our knowledge, whereas tabulated results from a very large number of observations, systematically made, would probably throw much light on the sequence and period of development of the several faculties. It would be desirable to test statistically ... the truth of the oft repeated statement that coloured children at first learn as quickly as white children, but that they afterwards fall off in progress.*

**4.5.1.3. Marriages between first cousins.** In connection with his botanical studies, Darwin [8, vol. 2, p. 104; 12, p. 465] became interested in the problem of the possibly harmful consequences of marriages between first cousins:

*Whether consanguineous marriages ... cause any injury will never be known with certainty until a census is taken with this object in view. My son, George Darwin, has done what is possible at present by a statistical investigation (regarding marriages between first cousins [58]). ... On the whole points to the evil [are] very small.*

Darwin [18, vol. 2, p. 309, 1. 1870] indicated his opinion on the desirability of such statistical studies once more, and noted [8, vol. 2, pp. 93 – 94] that

*There is good reason to believe ... that the evil effects of close interbreeding [of animals and plants in general] may be checked or quite prevented by the related individuals being separated for a few generations and exposed to different conditions of life.*

#### **4.6. Male and female births in animals.**

*Les lois que suivent à cet égard [in regard to the ratio of male and female births] les diverses espèces d'animaux me paraissent dignes de l'attention des naturalistes.*

Without knowing this opinion of Laplace [125, p. 158] Darwin became interested in the relative numbers of male ( $m$ ) and female ( $f$ ) births in animals in connection with problems in sexual selection. At first he [9, p. 328] thought that  $m > f$  with a large part of the males leaving no offspring at all. Rejecting this point of view (Ibidem), possibly because in man  $m \approx f$ , he (p. 332) started to compile relevant statistical data regarding race-horses and greyhounds. Moreover, Darwin (pp. 374 – 399) devoted a special study to the determination of the ratio sought for domestic animals, fish, birds, and even insects. In 1868 he [20, pp. 108 and 107] described this work in his correspondence:

(1) *I am having domesticated animals tabulated, & by patient enquiry, I hope to arrive at some degree of probability; certainty I fear is out of the question. With man alone we know positively that males are born in excess.*

(2) *The proportional numbers of male & female silk-moth has probably been observed [in France]. The whole subject is very intricate ... but I have often found that by patiently collecting facts, in relation to various classes [of animals?], a dim ray of light may be gained<sup>43</sup>. I am getting the results of breeding race horses, short horns & greyhounds, tabulated on a large scale.*

Darwin [18, vol. 1, p. 397] made known his interest in biological statistics even in 1846. Lastly, possibly in 1871 [19, vol. 2, pp. 41 – 42], he took up the same subject once more and, in particular, described the reasoning of his son George on the possible corruption of statistical data pertaining to the ratio  $m:f$  for illegitimate children.

#### **4.7. Specific problems**

4.7.1. *Extinction of higher taxonomic categories.* Darwin [2, pt. 1, pp. 40 – 41] attempted to explain the existence of discontinuities in the array of life forms. He offered a following example:

*If we take a man from any large family of 12 brothers and sisters in a state which does not increase, it will be chances against any one of them having progeny living ten thousand years hence; because at present day many are relatives, so that by tracing back the fathers would be reduced to small percentage: therefore the chances are excessively great against any two of the 12 having progeny after that distant period.*

Darwin (Ibidem, pp. 146 – 148) offered similar considerations concerning a constant or increasing population of a small settlement and concluded that

*There will be a period, though long distant, when of the present men not more than a few will have successors. [Of] two fine families one will [have] successors for centuries, the other will become extinct. ... Whole races act towards each other and are acted on, just like the two families.*

Darwin did not put forward any calculations, but he indirectly maintained that they are possible if, for a given group of men, the chance of their having a common progenitor is known.

Darwin [5, p. 116] formulated a similar assertion about species in general, again without furnishing any calculations and even without making any references to stochastic considerations:

*Of the species living at any one period, extremely few will transmit descendants to a remote futurity.*

Darwin had not mentioned a mathematical problem on the extinction of families which was formulated by statisticians and which belongs to the prehistory of branching processes [73, pp. 117 – 120].

#### 4.7.2. *Rare deviations.*

*If some rare deviation appears in parents and children, the mere doctrine of chances almost compels us to attribute its reappearance to inheritance,* Darwin [5, p. 26] maintained.

Elsewhere he [8, vol. 1, p. 449] formulated a specific problem:

*In a large population a particular affection occurs on an average in one out of a million, so that the à priori chance that an individual taken at random will be so affected is only one in a million. Let the population consist of sixty millions, composed ... of ten million families, each containing six members. On these data, Professor Stokes has calculated for me that the odds will be no less than 8333 millions to 1 that ... there will not be even a single family in which one parent and two children will be affected.*

The probability for at least one parent in a certain family to be affected is approximately equal to  $2 \cdot 10^{-6}$ . The corresponding

probability for two children out of four is  $C_4^2 10^{-12} = 6 \cdot 10^{-12}$  so that the probability that both these events occur independently is the product of these two numbers. Denote the number of families thus affected by  $\mu$  and the total number of them by  $n$  ( $n = 10^7$ ) and

$$P(\mu = 0) = \frac{(1.2 \cdot 10^{-10})^0}{0!} e^{-1.2 \cdot 10^{-10}} \approx 1,$$

$$P(\mu > 0) \approx P(\mu = 1) = \frac{1.2 \cdot 10^{-10}}{1!} e^{-1.2 \cdot 10^{-10}} \approx 1.2 \cdot 10^{-10},$$

$$\frac{P(\mu = 0)}{P(\mu > 0)} = \frac{1}{1.2 \cdot 10^{-10}} = 8.33 \cdot 10^9.$$

Darwin implies that because such families in England do exist ( $\mu > 0$ ), the rare deviation must be hereditary. He [8 vol. 1, pp. 469 and vol. 2, p. 56; 9, p. 189] made similar (and explicit) assertions several time more, but did not supply further numerical data.

4.7.3. *Pollination of plants.* Darwin studied the advantages of cross-fertilization as compared with spontaneous pollination. To this end he [12, p. 15] selected plants of *exactly the same age* which *were subjected ... to the same conditions and were descended from the same parents*. Despite these precautions, he found himself at a loss:

*As only a moderate number of ... plants were measured, it was of great importance for me to learn how far the averages were trustworthy. I therefore asked Mr. Galton ... to examine some of my tables of measurements.*

Having *examined the measurements ... with care, and by many statistical methods*, Galton confirmed Darwin's opinion that cross-fertilization was a preferable biological process.

What he did was to compare the two processes mainly in regard to the respective

(1) Sums of heights of the seedlings, or total number (and/or total weight) of seeds.

(2) Ordered heights of the seedlings (order statistics).

According to Galton's qualitative estimate, in both instances the results differed significantly. In the first case Galton could have calculated the corresponding variances and arrived at an expected difference between the observed quantities. As to the second instance, denote the heights of the seedlings by

$x_{(1)}, x_{(2)}, \dots, x_{(n)}$  and  $y_{(1)}, y_{(2)}, \dots, y_{(n)}$ ,

$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$  and  $y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$ .

Calculate the differences  $x_{(1)} - y_{(1)}, \dots$ . Galton was satisfied to notice that the signs of almost all of them were the same. A numerical evaluation of such data is now possible by the Spearman's rho, a coefficient of rank correlation. Darwin himself used only the first method of comparison whereas Fisher [59, § 18, p. 38], although he corroborated the conclusion about cross-fertilization, found out that the matter was not so obvious as Darwin (and Galton) had supposed.

Darwin himself used only the first method of comparison. Fisher [59, § 18, p. 38] corroborated his conclusions about cross-fertilization, and it turned out that it was not so obvious as Darwin (and Galton) had supposed.

*4.7.4. Different forms of flowers. If only two individuals of a dimorphic species happen to grow near together in an isolated spot, the chances are even that both will belong to the same form [13, p. 260]. If, however, the species is trimorphic, the chances are two to one in favour of their not belonging to the same form.*

This estimate is quite obvious. Darwin tacitly assumed that in both cases the different forms are equally numerous in individuals.

*4.7.5. The life of earthworms.* In studying the life of earthworms, Darwin and his son George together calculated the weight and mean displacement of earth and castings ejected from their burrows on unit area during a specified time [15]. This, of course, was a statistical research, and one which might well seem unrewarding at that. However, I am more interested in another part of Darwin's study, the one connected with earthworms dragging small objects into their burrows (Ibidem, pp. 52 – 55).

Darwin cut a few hundred triangles out of paper and strewed them about on the ground. After a while, earthworms carried away 315 of them. For 303 triangles Darwin managed to detect the specific part (1, 2 or 3) by which they were dragged. Here are the conclusions.

*If worms seized indifferently by chance any part [of the triangles], they would assuredly seize on the basal part ... far oftener than on either of the two other divisions. For the area of the basal to the apical part is as 5 to 1. ... The base offers two angles and the apex only one, so that the former would have twice as good a chance ... of being engulfed in a worm's month, as would the apex. ... Lastly, the proportion between the margins of the basal and apical parts is as 3 to 2 for the broad, and 2½ to 2 for the narrow triangles.*

The triangles were *elongated* (isosceles), some narrow, and some wide. Thus Darwin considered three possible cases of random dragging<sup>44</sup>, and reasonably rejected all of them in favour of a process that was not random.

Consider now the well-known Bertrand paradox. He [36, pp. 4 – 5] proved that in specific mathematical problems the expression *at random* (e. g., a random choice of a chord in a given circle) is not sufficiently definite. Darwin’s problem was of a physical rather than mathematical nature, and its ambiguity in regard to the choice of “randomness” is less obscure. Still, Bertrand could have referred to Darwin (and to Cournot [50, chap. 12] who criticized Laplace for an unreasonably restricted understanding of randomness in an astronomical problem).

### 5. Evolution of Species: Darwin

In § 5.1 I introduce a model of the evolution of species. Then, in §§ 5.2 – 5.5 I estimate its conformity with the ideas and statements of Darwin. Also in this section I consider (§ 5.6) a special problem, study (§ 5.7) Darwin’s understanding of randomness and discuss (§ 5.8) the importance of his work for the development of statistics.

**5.1. A stochastic model.** Consider an  $n$ -dimensional (possibly with  $n = \infty$ ) system of coordinates, (correlated!) body parameters of individuals belonging to a given species, (males and females separately) and introduce the respective Euclidean space  $A$  with the *usual* formula for the distance between points. There will somewhere exist a subspace  $B$  of optimal conditions for the life of that species, either over the habitat of the species or some part of it<sup>45</sup> and individuals situated far from  $B$  are apt to perish (or at least to leave less offspring). Individuals of the next generation will occupy other points of  $A$  and natural selection compels those individuals to approach  $B$ .

However, **1.**  $B$  moves in time with changes of external conditions. Abrupt changes can result in the disappearance of the entire species<sup>46</sup>. **2.** The direct path to  $B$  can be blocked en route by especially unfavourable local conditions. **3.** Individuals of given generation differ from their parents, but how?

Unknown probabilities enter everywhere and the model can only be qualitative<sup>47</sup>. For example, what is the probability of a given male mating a given female to produce healthy or weak offspring? The law of probability governs the random spread of the offspring around their parents. Many commentators of the 19<sup>th</sup> century only understood *uniform randomness*, cf. end of § 5.7 below, and (irrespective of that law) Darwin himself, see Note 48 in § 5.2 below, did not reflect the essential role of mutations in the *earlier* editions of the *Origin of Species* nor lived to see the rediscovery of the Mendelian laws.

Even neglecting the two last-mentioned circumstances, evolution is a random process, i. e. is extremely complicated, cf. Darwin’s illustration in note 68 below.

**5.2. Horizontal variations.** I introduce this term peculiar to a given generation as the projections of the deviations between individuals or points and the mean individual (point) on the coordinate axes. From the *Origin of Species* onward Darwin almost exclusively treated small variations<sup>48</sup> *due to unknown causes, ... without purpose, and in so far accidental* [19, vol. 1, p. 191, l. 1861].

Thus variations are accidental (Randomness I). Elsewhere [8, vol. 2, p. 416] Darwin discussed this important fact once more:

*Although each modification must have its own exciting cause, and though each is subjected to law, yet we can so rarely trace the precise relation between cause and effect, that we are tempted to speak of variations as if they arose spontaneously. We may even call them accidental, but this must be only in the sense in which we say that a fragment of rock dropped from a height owes its shape to accident.*

In most cases the change of conditions (and the consequent change of the positions of points *give the impulse to variability ... in a very indirect manner* [18, vol. 2, p. 517, l. 1881]<sup>49</sup>.

*By far the more important factor in the result*, Darwin [8, vol. 2, p. 415; 17, p. 336] concluded, *is the nature of the organisation which is acted on* rather than the surrounding conditions<sup>50</sup>. In particular, larger variations are peculiar to males rather than females [8, vol. 1, p. 457 and vol. 2, p. 404], and to men rather than women [9, pp. 344 and 852]<sup>51</sup>.

Also, DARWIN [5, p. 90] noted that a large variation in one individual would hardly perpetuate its kind *to the exclusion of the common form*. He explained that under domestication man usually breeds many individuals whose (slight) variations occur in the desired direction.

Darwin did not strictly define variations, but it seems that he meant deviations of some body measurement of an individual from the respective mean value for the species as a whole, i.e. *horizontal variations*. Thus, referring to Quetelet, he [10, p. 181] noted:

*It is known ... that the number of individuals [men] who exceed the average height by a given quantity is the same as the number of those who are shorter than the average by the same quantity.*

*.. So it is with the circumference of their chests; and we may presume that this is the usual law of variation in all the parts of every species under ordinary conditions of life.*

This passage shows that Darwin understood the statistical nature of variations. In a sense, he considered changes in mind and instinct of animals on a par with variations [2, pt. 1, p. 4], and once he [5, p. 232] even spoke about the natural selection of spontaneous variations of instincts.

The estimation of the relative influence of heredity and external conditions of life always interested Darwin. In 1876 he [18, vol. 2, p. 338] confessed:

*The greatest error which I have committed has been not allowing sufficient weight to the direct action of the environment ... independently of natural selection.*

**5.3. The emergence of varieties.** In accord with my general purpose I did not study Darwin's understanding of variability any further. In the same spirit, in this subsection I pay no attention to the discontinuous nature of heredity.

Darwin repeatedly claimed that the emergence of varieties was mainly peculiar to common species and genera:

(1) *Species which are most numerous in individuals have the best chance of producing favourable variations within any given period* [5, p. 104]<sup>52</sup>.

(2) *By far the most effective origin of well marked varieties and of species is the natural selection ... of those successive, slight & accidental (as we in our ignorance must call them) variations, which are ... advantageous to the individuals thus characterized ... hence [Darwin draws on statistical data] there would be a better chance of [new] varieties & species being thus formed amongst Common than amongst rare* [17, p. 136].

(3) *I have divided the N. Z. Flora as you [J. Lubbock] suggested. There are 329 [339] species in genera of 4 and upwards, and 323 in genera of 3 and less. The 339 species have 51 species presenting one or more varieties. The 323 species have only 37. [Darwin noted that 339:323=51:48.5. The last number in this proportion should be 48.6.] ... The case goes as I want it, but not strong enough; without it be general, for me to have much confidence in* [18, vol. 1, p. 462, 1.1857]<sup>53</sup>.

Darwin adduced similar data on the flora of other regions in another letter of the same year [19, vol. 1, pp. 98 – 100].

Thus he thought that the emergence of varieties is more probable in species which belong to larger genera. Having pronounced the same idea elsewhere, Darwin [5, p. 61] added (p. 62) that he tested *the truth of this anticipation* by suitably arranging the plants of twelve countries, and the coleopterous insects of two districts. And (p. 65), as the species of the larger genera *which present varieties, invariably present a larger number of varieties than do the species of the small genera. ... Larger genera tend to become larger.*

I comment on the stochastic nature of Darwin's reasoning in § 5.4.1. I also note that, referring to Alph. De Candolle and others, he [5, p 60] pointed out a determinate cause for the greater variability of

common species of plants, viz., their exposition to *diverse physical conditions*.

For the sake of completeness I add a few lines from Darwin's notebooks. Early in life he [2, pt. 2, p. 4] distinguished *two kinds of varieties, one approaching to nature of monster, hereditary, [the] other adaptation*. As to *the production of varieties* [2, pt. 1, p. 278], this takes place *per saltum*. However (p. 239) *any [large?] change suddenly acquired is with difficulty permanently transmitted*<sup>54</sup>. Also [2, pt. 2, p. 65] *any great change in species is reduced by atavism*.

These ideas seem to be self-contradictory. Obviously, Darwin significantly changed his mind during his later years. I conclude with yet another of Darwin's early opinions [1, p. 260] which he did not recall later: *There must be laws of variation – chance never would produce varieties*.

**5.4. Formation of species. A stochastic biological process.** Just how does the natural selection of favourable varieties work? Darwin [5, p. 59] supposed that *individual differences [are] the first steps towards slight varieties. ... The passage from one stage of difference to another [variety – subspecies – species] ... may be safely attributed to the cumulative action of natural selection ... and to the effects of the increased use or disuse of parts*<sup>55</sup>.

Because of the struggle for existence, Darwin [16, p. 120] also noted,

*Favourable variations would tend to be preserved and unfavourable ones to be destroyed. The result of this would be the formation of new species.*

In other words, the elimination of less fit individuals takes place, but then, exactly how does it all happen?

First, any slightest favourable variation in a given individual increases his chances of survival [3, p. 47; 4, pp. 119, 197 – 198, 242; 5, p. 444] and of leaving (a larger number of) descendants. With small values of the *vertical variations*<sup>56</sup> (between generations) favourable horizontal variations tend to remain preserved.

Second, variations consolidate after a few generations [4, p. 94; 2, pt. 3, p. 17; 8, vol. 2, p. 37]:

*A peculiarity generally becomes more firmly implanted after having passed through several generations; that is if one offspring out of twenty inherits a peculiarity from its parents, then its descendants will tend to transmit this peculiarity to a larger proportion than one in twenty; and so on in succeeding generations*<sup>57</sup>.

Third, sexual selection is extremely important. Indeed, both the direction and modulus of the vertical variation depend on each member of the respective parental pair with sexual selection influencing its formation. Other things being equal, the copulation of

the male (female) of a given species with female  $v_1$  (male  $u_1$ ) will be more probable, say, than with female  $v_2$  (male  $u_2$ ). And if even the weakest males of a given species ultimately find females, they usually copulate with the weakest females [9, p. 328].

Darwin (p. 325) concluded:

*It appears that the strongest and most vigorous males, or those provided with the best weapons, have prevailed ... and have led to the improvement of the natural breed or species. ... A slight degree of variability leading to some advantage, however slight, in reiterated deadly contests [between males] would suffice for the work of sexual selection.*

The divergence of sub-races and races thus achieved might be impeded, but, as a rule, it does not fade out<sup>58</sup>. Although in cross-breeds *there is a tendency in the young of each successive generation to produce the long-lost character* [5, p. 153], see also letter dated 1864 [19, vol. 2, p. 340], *the most distinct varieties of any one species of grass would have the best chance of ... supplanting the less distinct varieties* [5, p. 107]<sup>59</sup>. No wonder, then (p. 108), *Farmers find that they can raise most food by a rotation of plants belonging to the most different orders.*

Lastly [2, pt. 2, p. 30], *two varieties of many ages standing, will not readily breed together*. Darwin called this assertion *the most hypoth: part of theory*, but later [8, vol. 2, p. 82] pronounced a similar opinion concerning domestic animals, this time without any reservations.

*5.4.1. The stochastic nature of the evolution theory.* As I have shown above, Darwin used such usual words and expressions as *tendency, slight advantage, not readily* in a stochastic sense<sup>60</sup>. These and other household expressions [5, pp. 67, 159 and 164; 9, pp. 200 and 207] reveal the stochastic nature of Darwin's theory. But I especially emphasize that his whole doctrine hinges on the influence of slight constant causes, or very small differences between probabilities of two or several events, one and only one of which has to do with the reproduction of individuals in each generation. I have adduced relevant examples above and I can also refer to the *Origin of Species*, pp. 72, 75 and 81.

Darwin did not present any calculations since the lack of relevant data made them scarcely possible. Thus, for example, the estimation of the influence of sexual selection demands the knowledge of the probabilities of various possible copulations within a given species. I shall formulate Darwin's understanding of the action of slight causes in the following way. Let a vast series of trials be given. One and only one of a few possible independent events A, B, ..., C occurs in each trial. The probability of a certain event (say, A) is just higher than the probabilities of any other event. Then the total number of occurrences

of event A will be substantially larger than the respective numbers for events B, ..., C.

This (multinomial) scheme is rather simple; when only two events (A and B) are possible, it coincides with the case of Bernoulli (binomial) trials. It really seems that Darwin ought to have adduced a few methodological calculations to support his theory, using this latter case, if not the general pattern; see also § 5.8.2. Also, Darwin might have referred to the opinion of Laplace [96, p. xlviii], albeit not sufficiently clear and maybe even unconvincing [130, p. 11]:

*Il suit ... de ce théorème [of Bernoulli] que, dans une série d'événements indéfiniment prolongée, l'action des causes régulières et constantes doit l'emporter à la longue sur celle des causes irrégulières*<sup>61</sup>.

**5.5. The direction of evolution.** Darwin [8 vol. 1, p. 8] indirectly defined the direction of evolution (i. e., of the paths of a species to the subspace of most favourable conditions):

*We are almost compelled to look at the specialization or differentiation of parts or organs for different functions as the best or even sole standard of advancement. ... We may confidently believe that, on the whole, organization advances. Nevertheless a very simple form ... might remain for indefinite ages unaltered or unimproved*<sup>62</sup>.

Elsewhere Darwin [9, p. 253] briefly repeated his reasoning, this time mentioning Baer but adducing no reference. He [5, p. 143]<sup>63</sup> held that evolution ensues from a conflict of two opposite tendencies:

*There may truly be said to be a constant struggle going on between, on the one hand, the tendency to reversion to a less perfect state, as well as an innate tendency to new variations, and, on the other hand, the power of steady selection to keep the breed true. In the long run selection gains the day*<sup>64</sup>.

Darwin did not formulate his idea clearly enough: what he says means that (random) variations do not lead to the evolution of species; on the contrary, they counteract evolution! But in any case he thought that the evolution of species generally takes place in the direction from less probable to more probable structures. A comparison of evolution with the action of the second law of thermodynamics of course suggests itself. However, in respect to nature as a whole, the evolution of organisms is a transition from more probable to less probable structures (Fisher [58b, p. 40]).

I note also that Darwin [5, p. 321] did not believe in a reappearance of a species *once lost, even if the very same conditions of life, organic and inorganic, should recur*. Thus he rejected the possibility that any closed curve might occur in the process of evolution<sup>65</sup>. I understand that the corresponding problem has not been studied thoroughly enough.

**5.6. Evolution of Man and Society.** The evolution of species depends also on intraspecific competition and intraspecific struggle. Darwin [5, pp. 68 and 75] said a few words about these phenomena, and noted (p. 78, subtitle) *that struggle for life [is] most severe between individuals and varieties of the same species*. Of course Darwin also turned for suitable examples to the life of social animals (in particular, bees and ants) whose individual existence is altogether impossible.

One of his writings [9] is devoted to man. Here Darwin proves, from a biological point of view, that the moral sense of man will deepen:

(1) *The simplest reason would tell each individual that he ought to extend his social instincts and sympathies to all the members of the same nation. ... This point being once reached, there is only an artificial barrier to prevent his sympathies extending to the men of all nations and races* (p. 188).

(2) *We may expect that virtuous habits will grow stronger, becoming perhaps fixed by inheritance* (p 192).

I connect these passages with Laplace's naive belief that, owing to the laws of probability (§ 5.4.1), the world will be gradually reconstructed according to the *principes éternels de raison, de justice et d'humanité* [96, p. xlvi; 125, p. 182]<sup>66</sup>.

**5.7. Necessity and randomness.** Having gradually departed from creationism, the scholars of the 19<sup>th</sup> century had to assign a more or less significant role to randomness. Even creationists had no escape. Thus, Cuvier [54] likely thought that his revolutions were necessary (divine) acts, but their total number<sup>67</sup> must have caused some embarrassment.

Especially interesting philosophical thoughts are due to Darwin The regularity of the summary action of many random events is the cornerstone of his theory (§ 5.4.1)<sup>68</sup>. However, he himself proclaimed Randomness I [5, p. 128]<sup>69</sup>:

*I have hitherto sometimes spoken as if the variations ... were due to chance. This, of course, is a wholly incorrect expression, but it serves to acknowledge plainly our ignorance of the cause of each particular variation*<sup>70</sup>.

Trying his best to rescue Darwin from criticisms, Huxley [18, vol. 1, p. 553] emphasized his denial of randomness:

*The most singular of these, perhaps immortal fallacies [committed by Darwin's commentators] ... is that which charges Mr. Darwin with having attempted to reinstate the old pagan goddess, Chance. ... [His] whole theory crumbles to pieces if the uniformity and regularity of natural causation for illimitable past ages is denied.*

But then, Darwin [19, vol. 1, p. 395, 1. 1881] believed in Randomness II:

*Mr. Graham must have used “chance” in relation only to purpose in the origination of species. This is the only way I have used the word chance. ... On the other hand, if we consider the whole universe, the mind refuses to look at it as the outcome of chance, that is, without design or purpose.*

See also § 5.2 for a passage from Darwin’s letter dated 1861 (Ibidem, p. 191).

Lastly, Darwin also mentioned Randomness III [8, vol. 2, p. 236; 56, pt. 1, p 188]: [the action of selection]

*Absolutely depends on what we in our ignorance call spontaneous or accidental variability. Let an architect be compelled to build an edifice with uncut stones, fallen from a precipice. The shape of each fragment may be called accidental. Yet the shape of each has been determined by events and circumstances, all of which depend on natural laws; but there is no relation between these laws and the purpose for which each fragment is used by the builder. In the same manner the variations of each creature are determined by fixed and immutable laws, but these bear no relation to the living structure which is slowly built up through the power of selection.*

Being a deist, Darwin flatly rejected constant divine intervention:

(1) *Now the creationist believes these three [species of rhinoceroses] were created [separately one from another]; as well can I believe the planets revolve in their present courses not from one law of gravity but from distinct volition of Creator [3, p. 84].*

Darwin [2, pt. 1, p. 196; 4, p. 250; 19, vol. 1, p. 194, 1. 1861] used similar examples several times.

(2) *The idea of species having been created separately one from another makes the works of God a mere mockery and deception [5, p. 154].*

(3) *The view that each variation has been providentially arranged seems to me to make Natural Selection entirely superfluous, and indeed takes the whole case of the appearance of new species out of the range of science [19, vol. 1, p. 191, 1. 1861].*

As regards the origin of higher taxonomic categories, Darwin [18, vol. 2, pp. 145 – 146, 1. 1860; 19, vol. 1, pp. 321 – 322 (1. 1870) and 395 (1. 1881)] was unable to choose between chance and design:

*I cannot think that the world ... is the result of chance; and yet I cannot look at each separate thing as the result of Design.*

Darwin’s doubts were partly due to his agnosticism [20, p. 88, 1. 1879]:

*My judgement often fluctuates. ... In my most extreme fluctuations I have never been an atheist in the sense of denying the existence of a*

*God. I think that generally (and more and more as I grow older), but not always, that an Agnostic would be the more correct description of my state of mind.*

Still, Darwin [8, vol. 1, p. 12] formulated a definite opinion on the origin of life. He spoke out in favour of

*A few forms or of only one form having been originally created ... This more simple view accords well with Maupertuis' philosophical axiom of "least action."<sup>71</sup>*

### **5.8. Darwin's contribution to statistics**

*5.8.1. The Biometric School and the statistical method.* From 1867 onward Darwin came not to care for general beliefs (§ 4.1). Many of his contemporaries, Galton and Pearson included, felt themselves emancipated by his (and Wallace's) works and could have repeated this remark.

In 1886 Galton [109, p. 201, note] confessed:

*Owing to [my] hereditary bent of mind ... I was well prepared to assimilate the theories of Charles Darwin. ... [His publications] enlarged the horizon of my ideas. I drew from them the breath of a fuller scientific life, and I owe more of my later scientific impulses to the influences of Charles Darwin than I can easily express.*

Speaking in 1908, at the Darwin – Wallace celebration, Galton (Ibidem) expressed himself more definitely:

*The then new doctrine [due to Darwin and Wallace] ... burst the enthrallment of the intellect which the advocates of the argument from design had woven round us. It gave a sense of freedom<sup>71a</sup>.*

In 1923 Pearson [108, p. 23] expressed his own feelings in words of the same nature:

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

Pearson was a cofounder of the British Biometric school called into being by the need to introduce mathematical methods into biology. And this is how the founders of this school (W. F. R. Weldon, K. Pearson, C. B. Davenport) formulated their attitude towards Darwin and his theory [121, pp. 240 – 241 and 241]:

(1) *The starting point of Darwin's theory of evolution is precisely the existence of ... differences [between individuals of a race or species].*

(2) *May we not ask how it came about that the founder of our modern theory of descent made so little appeal to statistics?<sup>72</sup> The characteristic bent of C. Darwin's mind led him to establish the theory of descent without mathematical conceptions; even so Faraday's mind worked in the case of electro-magnetism<sup>73</sup>.*

*But as every idea of Faraday allows of mathematical definition, and demands mathematical analysis, so every idea of Darwin, variation,*

*natural selection, ... seems at once to fit itself to mathematical definition and to demand statistical analysis.*

It is opportune to recall Darwin's confession [16, p. 58] of his scant mathematical education:

*I have deeply regretted that I did not proceed far enough at least to understand something of the great leading principles of mathematics; for men thus endowed seem to have an extra sense.*

It is not easy to grasp the leading principles of any science. Still, a related question is warranted: suppose Darwin had known mathematics much better than in fact he did, what would have happened? First, he might have changed the bent of his mind and produced clear definitions<sup>74</sup>. Second, he might have used the terminology of probability, thus plainly pointing out the stochastic nature of the theory of evolution. Third, he might have corroborated or at least illustrated his considerations by stochastic calculations<sup>75</sup>. I return to such calculations in § 5.8.2. Here I add that the ideas of the Biometric School were not recognized all at once. Furthermore, at the turn of the 19<sup>th</sup> century the statistical method itself was not yet completely acknowledged in biology. Witness, for example, the opinion of a botanist [34, p. 45]:

*Undoubtedly the statistical method will possess its own merits until science will not acquire other, more accurate methods. However, statistical inferences cannot be considered on a par with laws of nature, because, for one thing, these inferences have to do only with specific phenomena. Statistics has no power to discover the true causes of the rules which it determines. As any other rules, these are liable to exceptions and [even] radical changes.*

The author is correct in stating that statistics does not provide a final corroboration of a theory. Physicists held similar views, though not up to the end of the 19<sup>th</sup> century. Schrödinger [121, § 3.2] described the situation in natural science, but he did not refer specifically to biology. He also maintained that Darwin *was the first scientific man aware of the vital role of statistics*<sup>76</sup> and that his theory is founded on the law of large numbers (§ 5.4.1).

5.8.2. *Stochastic reasoning and calculations.* From the time when discussions of Darwin's theory began, stochastic calculations really entered biology. Thus Danilevsky [56, pt. 1, p. 23], reckoning himself among *most resolute opponents of Darwin's doctrine*, adduced quite a few trial calculations (e. g., in pt. 2, pp. 92 – 93) the results of which in his opinion testified against Darwin<sup>77</sup>.

As years went by, and especially as Mendel's ideas were taken up, stochastic calculations became an inalienable feature of the evolution theory, and biology in general.

5.8.3. *Biologists before Darwin: Thoughts about Evolution*

I adduce their statements naming, in a generalized and sometimes formalized manner, their subject and explaining its essence. These statements mostly concern the evolution of species which began to interest biologists since mid-18<sup>th</sup> century and have to do with variations which perhaps gradually became a main object of biological research.

1. Adanson (1757, p. 61; 1763), Aug. P. De Candolle (1813/1819, p. 29). Natural classification of plants. *Organisms are points in a many-dimensional space*

2. Humboldt (1845, p. 82). Mean conditions (states) in nature. *Idea of random horizontal variation*

3. Maupertuis (1745/1756a, p. 120 – 121). *Vertical variations are random and small*

4. Maupertuis (1751, p. 146), Cournot (1851, p. 119). Randomness. *Its role in evolution is restricted*

5. Lamarck (1809/1873, pt. 1, chapter 7; 1817, p. 450), Maupertuis (1756b/1756a, p. 276). Change of external conditions. *Horizontal variations are random and can lead to changes in species*

6. Lamarck (an VIII (1800)/1906, p. 465; an X (1802)/1906, p. 511; 1809, t. 1, Chapter 7). Evolution of species. *Is a universal phenomenon*

7. E. Geoffroy Saint-Hilaire (1818 – 1822, 1822, p. 121). New species. *They originate due to random mutations*

8. Goethe (1831/1891, p. 120). Species of plants. *They change (bilden)*

9. Comte (1830 – 1842, t. 3/1893, No. 40, pp. 234 and 278, No. 42, pp. 444 and 446). Evolution of species. *Due to external changes*

10. I. Geoffroy Saint-Hilaire (1859, title of the pertinent section). Species and their evolution. *Species vary restrictively*

**5.9. A. R. Wallace.** Alfred Russell Wallace (1823 – 1913) was a humble co-founder of the theory of evolution. A biologist (and anthropologist), he also made important contributions in other fields of biology. I do not study his works (which are described by an immense literature) and only offer a few comments.

Wallace [132, p. 37] mentions tendencies and [134, p. 58] *probable outcomes (we should probably find, in accordance with the law of averages)*. He uses the same argument about slight causes [132, p. 38]:

*The scale on which nature works is so cast ... that any cause, however slight, and however liable to be veiled and counteracted by accidental circumstances, must in the end produce its full legitimate results.*

Wallace [132, pp. 37 and 35] used mathematical language to a somewhat greater extent than Darwin:

(1) *Though the doctrine of chances or averages can never be trusted to an a limited scale, yet, if applied to high numbers, the*

*results come nearer to what theory demands, and, as we approach to an infinity of examples, become strictly accurate.*

(2) *On the average the rule ... will invariably be found to hold true.*

Once Wallace even used the normal distribution to estimate the relative number of individuals of a given species with one or another variation in a certain body measurement.

Wallace is also meritorious for his other contributions, see for example [133].

## Notes

Referring to Darwin's correspondence, I use "I" instead of *letter dated*

1. Among scholars of modern times I [124, p. 116] have quoted Harvey, Kepler and Huygens. I (p. 117) also recall that Laplace was prepared to accept a tendency of changes in species. However, having read a few works written by Ray and Tournefort, I do not think that the 17<sup>th</sup> century was as a whole interesting from my point of view.

2. Cf. Darwin's opinion [5, p. 404]:

*If it [a trifling character] prevails throughout many and different species ... it assumes high value.*

3. Lamarck [84, pp. c – ci] discussed the same problem. If, for example, flowers of all kinds of plants possess calyxes, then, as he says, the *weight* of the calyx must be just 0.7.

4. The meteorologist Cotte [49, p. 96] praised Adanson's work calling it

*Peut-être le plus complet & le plus savant qui ait été fait sur la Botanique.*

5. In the same work Cuvier introduced an embryo of the notion of correlative empirical laws. Describing the correlation *des formes dans les êtres organisés*, he (p. 62) noted:

*Aucune de [parties of the body] ne peut changer sans que les autres ne changent aussi.*

The report of the Paris Academy of Sciences for 1827, drawn up by Cuvier [55, p. clxxxvii; 24, p. 91] includes an account of experiments carried out by G. De Busaraingues on male and female births in animals.

*Si l'on veut avoir plus de femelles, il faut employer des mâles jeunes et des femelles dans l'âge de la force, et nourrir celles-ci plus abondamment que ceux-là. Il faut faire l'inverse si l'on peut produire plus de mâles.* [The results of the experiment follow.] *Les oiseaux suivent la même loi que les moutons.*

6. It concerns a particular case, but De Candolle probably held similar thoughts about empirical laws in general.

7. Six laws described the influence of heat alone (pp. 1103 – 1109)! Here is the first one:

*La faculté de chaque plante et de chaque partie d'une plante pour résister aux extrêmes de la température, est en raison inverse de la quantité d'eau qu'elle contient.*

8. Camper [38] also wrote a sociological study of diseases in man.

9. He even communicated to Quetelet [111, t. 2, p. 226] how many drunkards did the London police force collect during each month of the year 1832.

10. The London (Royal) Statistical Society founded in 1834 [115a, p. 492] had declared, from the very beginning, its refusal to study compiled statistical data [Ibidem; 124, p. 121, note 103]. Probably Babbage was mainly responsible for this policy which was soon abandoned [124, p. 121; 73a].

- 11.** During his journey to Russia Humboldt [116, pp. 330 – 333] approximately estimated the catches in the Caspian Sea.
- 12.** Baer (1847) left an unpublished questionnaire on medical statistics which he compiled for the Military Medical Academy in Petersburg. (He was a member of a committee on medical statistics of this academy [114, p. 460]).
- 13.** Other types of vaccines as well as antibiotics are used nowadays against anthrax.
- 14.** Whereas Humboldt [76, p. xlvi] mentioned Tournefort, Linné and other scholars as his predecessors Alph. De Candolle [39, t. 1, p. vi] named Linné as Humboldt's only forerunner. Moreover, he credited Linné only with *des idées ordinairement très justes* [but] *parfois bizarres et erronnées*. De Candolle also contended that the geography of plants was developed by Humboldt, Aug. De Candolle and Brown. Cf. Darwin's opinion [19, vol. 2, p. 26, l. 188.1]: *I have always looked at [Humboldt] as, in fact, the founder of the geographical distribution of organisms*.
- 15.** But a poet had forestalled scholars! Humboldt [81, Bd. 1, p. 16] quotes Schiller: *Der Weise ... sucht das vertraute Gesetz in des Zufalls grausenden Wundern* etc. I [124, p. 97, note 2a] have referred to this passage but did not mention Humboldt.
- 16.** Humboldt did not directly busy himself with astronomy, but he [81, Bd. 3, p. 152] appreciated W. Herschel's studies:  
*W. Herschel den glücklichen Gedanken hatte gleichsam das Senkblei in die Tiefen des Himmels zu werfen und in seinen Stern-Achtungen die Sterne zu zählen, welche nach verschiedenen Abständen von der Milchstraße durch das Gesichtsfeld seines Telesopes gingen; wurde das Gesetz der mit der Nähe der Milchstraße zunehmenden Sternenmenge aufgefunden.* And (p. 175):  
*In der Vertheilung der Fixsterne an dem Himmelsgewölbe hat man erst angefangen gewisse Gesetze relativer Verdichtungen zu erkennen, seitdem W. Herschel [paved the way], etc.*
- Humboldt (p 401) also turned the attention of astronomers to an article of Schwabe on the regularities in the formation of sun-spots.
- 16a.** Sometimes I refer to  $m$  sources but quote only  $n$  ( $n < m$ ) passages. In these instances quotations are from the first  $n$  sources.
- 17.** A translation of *Relation historique*, this being a set of volumes included (as also, for example, the *Essai* [74]) in Humboldt's mammoth *Voyage aux régions équinoxiales*.
- 18.** Sec above (note 14) still another passage. A touch of irony or, alternatively, open-heartedness, is felt in this passage [16, p. 107]:  
*He [Sir John Herschel] never talked much, but every word which he uttered was worth listening to. I once met ... the illustrious Humboldt, who honoured me by expressing a wish to see me. ... I can remember nothing distinctly about our interview, except that Humboldt was very cheerful and talked much.*
- 19.** Except for the few words left in Latin, I have translated the last passage from Russian. Humboldt also engaged in political arithmetic. His *Voyage* [74] (see note 17) includes a politico-arithmetical account of Mexico in which he pays due attention to demography. Thus, though drawing on a small number of observations, Humboldt [74, t. 1, pp. 457 – 458] noted that, as in Europe, male births exceeded births of females.  
 Laplace [125, p. 158] noticed this fact. Lastly, I remark that Humboldt [75, plate 19] published a graph of a time series and used a bar graph. Despite his references [74, t. 1, p. 153 and elsewhere] to Playfair the work of the latter remained unnoticed until 1879 [60, p. 181].
- 20.** It is opportune to adduce Laplace's relevant opinion [96, p. cliii]:  
*La théorie des probabilités n'est, au fond, que le bon sens réduit au calcul.*  
 I suspect however that the same was then true about mathematics as a whole.
- 21.** Addressing himself to scholars of various specialities, De Candolle [39, t. 2,

p. 1342] recommended definite measures to promote the geography of plants. Thus he proposed to use the word *grade* (or, in English, *grad*) instead of *degré* (degree) for temperatures measured in the centigrade scale. He also noted that the Fahrenheit thermometer would be sooner or later abandoned, at least in scientific work. His ideas were in line with pressing demands voiced by statisticians of those times to introduce the metric system.

22. A grave shortcoming peculiar to all his work.

23. Thus Adanson felt that random events are necessary to achieve a state of equilibrium in nature. Regrettably, he did not elaborate. A similar meaning is sensed in an evasive pronouncement due to Kant [82, p. 447] and made almost at the same time:

*Eben in der Vermengung des Bösen mit dem Guten die großen Triebfedern liegen, welche die schlafenden Kräfte der Menschheit in Spiel setzen und sie nötigen, alle ihre Talente zu entwickeln und sich der Vollkommenheit ihrer Bestimmung zu nähern.*

His reasoning, directed against Maupertuis, was probably occasioned by the latter's good intention [101, p. 159\*] to accelerate the development of science and arts by encouraging children to follow the professional occupations of their fathers.

24. Elsewhere Maupertuis [101, pp. 146\* - 147\*] formulated similar assertions on the likeness between child and ancestor and on deformities.

25. Neither of the two passages is included in the part of Lamarck's manuscript published in the original French and English translation [93]. I had to translate both of them from Russian.

26. Although Boscovich considered both repulsion and attraction in his physical works, hardly anyone except Lamarck connected the former with randomness.

27. See also end of chap. 1 of same source.

28. I am not sure that Lamarck's meteorological forecasts (Ibidem) were impotent.

29. For example, he [91, p. 170] contended that without the action of a

*Force répulsive, la lumière, qui traverse sans cesse l'espace dans toute direction, ne serait point mise en mouvement.*

30. Lamarck first pronounced his ideas on the evolution of species in 1800 – 1802 [86, p. 465; 87, p. 511].

31. Packard [103, pp. 351 – 353] attempted to allay Lamarck's carelessness by remarking that the efforts of animals might have been *as much physiological, reflex, or instinctive as mental.*

32. He gave no reference. I quote his opinion from one of these sources [65, p. 69]:

*Le moment est enfin venu d'en constater l'existence, et de mettre en évidence qu'il est deux sortes de faits différentiels à étudier dans l'organisation. 1. Ceux qui appartiennent à l'essence des germes, et 2. Ceux qui proviennent de l'intervention du monde extérieur.*

33. Goethe (§ 102) was convinced that in principle a mathematical description of the forms of plants would be possible:

*Wir sind überzeugt, dass mit einiger Übung es nicht schwer sei, sich die mannichfaltigen Gestalten der Blumen und Früchte zu erklären; nur wird freilich dazu erfordert, dass man mit jenen oben festgestellten Begriffen der Ausdehnung und Zusammenziehung, der Zusammendrückung und Anastomose, wie mit Algebraischen Formeln bequem zu operiren.*

This problem is very complex. I do not know whether it is solved even now.

34. I note Comte's (Ibidem, No. 40, p. 329) negative attitude towards medical statistics:

*Cette prétendue application de ce qu'on appelle la statistique à la médecine, dont plusieurs savants attendent des merveilles, et qui pourtant ne saurait aboutir, par sa nature, qu'à une profonde dégénération directe de l'art médical, des lors réduit à*

*d'aveugles dénombrements.*

*Une telle méthode, s'il est permis de lui accorder ce nom, ne serait réellement autre chose que l'empirisme absolu, déguisé sous de frivoles apparences mathématiques.*

In particular (p. 330), the existence of variations leads to

*L'impossibilité manifeste de comparer judicieusement deux modes curatifs d'après les seuls tableaux statistiques de leurs effets, abstraction faite de toute saine théorie médicale.*

Comte's opinion was in conflict with contemporary ideas [126, § 5.4] but the middle of the 19<sup>th</sup> century saw a certain disappointment in the possibilities of statistics [125, p. 181].

**35.** Rouillier [118, p. 160] noticed (and explained) regularities in the distribution of spots on hides of domestic animals:

*Nous étions arrivé à deviner, sans voir le cheval, quels pieds sont blancs; nous nous mimes à parier trois contre un que nous saurions deviner, et ... nous fûmes en gain, c'est à dire quel plus des 3/4 des chevaux observés confirmaient la loi que nous avions remarquée.*

Rouillier detected this regularity looking out the window at animals during the few days of an illness. The Russian version of his article is entitled *To kill time*.

**36.** Cournot did not connect the type with any specific taxonomic category.

**37.** Darwin [14, p. 452] once noted in regard to two of his observations that their difference does not exceed the probable limit of error. He scarcely used this term in its exact sense. Elsewhere Darwin [5, p. 63] remarked that he had *endeavoured to test* [a certain conclusion] *numerically by averages*. This of course is not specific enough.

**38.** But isn't it just as possible that, irrespective of corroboration, Darwin was simply afraid of finding himself in the limelight? Cf. also the opinion of Tolstoy [129, chap. 29, p. 334], who decided:

*There exist an innumerable ... quantity of facts ... which one should investigate. Before investigating facts, it is necessary to have a theory on the basis of which facts are investigated, i. e. [a theory is necessary before] some facts or other are selected from among an innumerable quantity of them.*

**39.** These measurements included determinations of heights of ground features above sea level by barometric levelling. Darwin himself performed the field-work, if not the mathematical treatment of observations.

**40.** Quetelet [111, t. 2, p. 388] also supposed that it is

*Très-utile de le [a physician] voir aussi quand on est dans l'état de santé, afin qu'il pût bien étudier notre état normal et se procurer les éléments de comparaison nécessaires pour les cas d'anomalie et les indispositions.*

**41.** Darwin [18, vol. 2, p. 348, 1. 1872] wrote to De Candolle in connection with the latter's book [40]. He remarked that he took interest in several chapters, including one on statistics. De Candolle devoted this (very short) chapter to the relative stability of the number of accidents and crimes and mentioned Quetelet and Buckle. The book as a whole deserves special notice but in essence it belongs to the post-Darwinian = Galtonian period. Note that the year of its publication (1873) as given by De Candolle himself [40, preface] is obviously wrong.

**42.** William Farr (1807 – 1883) was the founder of the English national system of vital statistics [60, pp 188 – 190].

**43.** This opinion is hardly pessimistic. Elsewhere Darwin expressed even more confidence in statistics (§ 4.5.1.2) but then he [19, vol. 1, p. 168, 1. 1860] also understood the difficulties of weighing probabilities against each other.

**44.** Actually even four, because he additionally mentioned the possibility that all parts of the triangles could be equally attractive.

**45.** If a subdivision of the habitat is necessary (for example, if the species is geographically isolated) evolution should probably be considered for each region separately.

**46.** Darwin [9, p. 283] noted that the extinction of native populations under the pressure of civilization takes place because of rapid changes in the conditions of their life.

**47.** For the same reason Darwin's theory was just a hypothesis. But at least it led to the emergence of new fundamental problems and, together with the theory of heredity initiated by Mendel, determined the development of biology for many a decade.

**48.** This fact perhaps explains Darwin's pronouncement [5, p. 195]:

*In the earlier editions of this work I under-rated, as it now seems probable, the frequency and importance of modifications due to spontaneous variability.*

**49.** At the same time Darwin [9, p. 295] supposed that

*The immunity of civilised races [of man] and domesticated animals is probably due to their having been subjected to a greater extent to diversified or varying conditions, than the majority of wild animals.*

**50.** Possibly Darwin also had in mind the organisation of the species concerned; the measure of variation (or deviation) likely changes from one species to another.

**51.** Having been informed about a larger variability of females of some insects, Darwin [19, vol. 1, p. 181, l. 1861] confessed that this fact was new to him.

Anyway, to avoid a self-contradiction (see below), Darwin should have thought that men enjoy larger variability also in their mental faculties. But then his opinion [9, p. 858] is just wrong:

*We may ... infer, from the law of the deviation from averages, so well illustrated by Mr. Galton, in his work on 'Hereditary Genius' that if men are capable of a decided pre-eminence over women in many subjects, the average of mental power in man must be above that of women.*

Referring to his previous works, a modern author explains [61a] the prevalence of male births in man, the greater mortality of men, their larger variability and, for that matter, the very existence of the two sexes in animals by a single principle. The male sex, he contends, is responsible for the evolution of the species, for its adaptation to a changing world, while the female sex is more conservative and takes care of stability. Still, the greater variability of males in animals seems not to be proved.

**52.** Darwin (Ibidem, p. 46) repeated a similar statement in regard to domestic animals:

*As variations manifestly useful or pleasing to man appear only occasionally, the chance of their appearance will be much increased by a large number of individuals being kept.*

He (p. 86) also added that accidental destruction of individuals may be *ever so heavy*, provided the species remains numerous in individuals his (Darwin's) theory holds its ground.

**53.** Darwin could have estimated the significance of the observed divergence according to formulas due to Poisson [126, § 5.2.2].

**54.** Cf. Darwin's later statement [5, p. 90]:

*The preservation of any occasional deviation of structure, such as a monstrosity, would be a rare event.*

**55.** Elsewhere Darwin [10, p. 181] hypothetically remarked that the less useful parts of individuals of a species existing under unfavourable conditions diminish and ultimately disappear. He [5, p. 38] also mentioned the accumulation of (random) variations: fanciers, he argued, do not believe in the common origin of the various breeds of domestic animals (for example, pigeons). They

*Ignore all general arguments and refuse to sum up in their mind slight differences accumulated during many successive generations.*

See also note 68.

**56.** The variance functions of processes  $\xi_n(t)$  and  $\xi_v(t)$  slowly decrease as the time interval between the two respective generations increases.

**57.** But of course atavism is also possible [8, vol. 2, chap. 13].

**58.** This corresponds to the diversity of the paths of vertical variations.

**59.** Consider a related passage: *A breed, like a dialect of a language, can hardly be said to have a distinct origin* (Ibidem, p. 45). Darwin again and again [Ibidem, p. 434; 18, vol. 2, p. 152; 9, p. 137] compared the evolution of species with the development of languages:

*(1) Rudimentary organs may be compared with the letters in a word, still retained in the spelling, but become useless in the pronunciation. (2) Agassiz admits that the derivation of languages, and that of species or forms, stand on the same foundation, and that he must allow the latter if he allows the former, which I tell him is perfectly logical. Is not this marvellous?*

This passage is from a letter written by Asa Gray, but the last sentence is Darwin's comment. The opinion of Darwin, Gray and Agassiz seems no longer held (*Enc. Brit.*, vol. 14, 1965, p. 74, article *Linguistics*).

**60.** In some instances Darwin used mathematical language: *probable limit of errors* (§ 4); *doctrine of chances* (§ 4.7.2), *law of deviation from averages* and *law of variation* (§ 5.2) etc. However, such cases seem to be exceptions and, what is more, the terminology was not the best possible: Darwin understandably followed Quetelet rather than Laplace.

**62.** I [125, p. 180, note 3] have remarked that in 1864 M. Ye. Vashchenko-Zakharchenko connected the theory of Darwin (and Buckle's ideas on the development of society) with the same theorem due to Jacob Bernoulli.

**62.** Exactly the belief in the possibility of regress and, generally, in the capacity of simple forms for competition enabled Darwin [5, pp. 118 and 194] to reject the necessity of regular spontaneous generation regardless of Lamarck's views (§ 3.2.3).

**63.** Cf. Lamarck (see reference in note 62).

**64.** Darwin repeated this assertion on the next page of the *Or. spec.* (p. 144).

**65.** Assuming an unlimited time period, Poincaré later proved a theorem on the return of a dynamical system to a configuration arbitrarily near to its original one.

**66.** In 1873 Darwin [19, vol. 2, p. 43] called Galton's eugenic ideas utopian. Yet in his early life he [2, pt. 2, p. 220] remarked:

*Educate all classes, avoid the contamination of castes, improve the women (double influence) & mankind must improve.*

A modern author [58a] noted that mutual assistance became all the more necessary for man because children remain completely helpless for a rather long period of time. He put forward a number of relevant points and quoted P. A. Kropotkin, the anarchist and natural scientist.

Laplace's and Darwin's prediction proved damnably wrong.

**67.** His followers enumerated 27 revolutions (*Great Sov. Enc.*, vol. 11, 1973, p. 525). [This, the third edition of the *Encyclopaedia*, is available in an English translation.]

I recall that Kepler [124, pp. 121 – 122] had to construct a special theory to explain why God created just six planets.

**68.** It is remarkable that, at least as regards its complexity, Darwin [5, p. 77] compared the evolution of species with the accumulation of random influences or, as it seems, even with a random process:

*Throw up a handful of feathers, and all fall to the ground according to definite [but extremely complicated stochastic] laws; but how simple is the problem where each shall fall compared with problems in the evolution of species.*

See also note 55.

**69.** A similarity between Darwin and Kepler suggests itself. The latter did not recognize randomness either but he [124, § 8.1] had to admit it in fact.

**70.** I have quoted similar pronouncements in §§ 5.2 and 5.3, but there are still other passages of interest [8, vol. 2, p. 240; 9, p. 65]. I adduce one special remark due to Darwin [8, vol. 2, p. 57]:

*This power [inheritance] often appears to us in our ignorance to act capriciously, transmitting a character with inexplicable strength or feebleness.*

This point of view suggests a question: would have Darwin regarded a transmission of characters with one or another “biologically” acceptable law of distribution as a case of random heredity? I think that the answer is “No, not at all”.

**71.** I add a note on Baer. He

*Im Grunde teleologisch dachte [114, p. 367] and der Siegeszug des Darwinismus ... scheint ihm unbegründet zu sein und versetzte ihn in den Zustand eines verärgerten Staunens (Ibidem, p. 366).*

Most interesting is how Baer expressed himself in this connection [30, p. 6; 56, pt. 1, p. 194]:

*Dunkel regt sich hiebei in mir die Erinnerung, dass ich schon einmal von dem Bestreben das Zweckmäßige, ja Tieffinnige, durch Elimination des Anpassenden, durch zufällige Variabilität Erzeugten zu erreichen, gelesen oder gehört habe. In der Akademie von Lagado hat ein Philosoph, von dem richtigen Gedanken ausgehend, dass alle für die Menschen erreichbare Wahrheit doch nur durch Worte ausgedrückt werden könne, alle Wörter seiner Sprache in allen ihren grammatischen Formen auf Würfel geschrieben, und eine Maschine erfunden, welche diese auf allen Seiten beschriebenen Würfel nicht nur wendete, sondern auch in einander verschob.*

*Nach jeder Handhabung der Maschine wurden die sichtbaren Wörter abgelesen, und wenn drei oder vier mit einander einen Sinn gaben, wurde diese Wortfolge notirt, um auf diesem Wege zu aller möglichen Weisheit zu gelangen, die ja auch nur in Worten ausgedrückt werden kann.*

The good-for-nothing philosopher originally described in *Gulliver's Travels* obviously assumed that each word occurs in a given language equally often. This, of course, is not so. A similar supposition does not hold in biology either: even if a certain variation is a random variable with a uniform distribution, its hereditary transmission obeys much more complex laws. Baer wasn't a mathematician; he originally published his work [30] in a newspaper [114, p. 477], and, lastly, he did not recall his argument in further work [31; 32].

It seems that John Herschel [19, vol. 1, p. 190] referred to Swift in the same connection even before Baer. Swift borrowed his story from Raymond Lully (13<sup>th</sup> – 14<sup>th</sup> century).

**71a.** But did not Darwinism, in its turn, fetter later generations of biologists? I have no answer.

**72.** I would say: so little appeal to quantitative methods of treating statistical data.

**73.** A comparison of a biologist with a physicist is not quite proper.

**74.** See above the opinion of the founders of the Biometric school. I also refer to Ruse [120] who studied Darwin's notion *natural selection* with the express purpose of clarifying it.

**75.** I shall not add that Darwin would not have needed to consult Stokes (§ 4.7.2) and Galton (§ 4.7.3). His request for advice addressed to such scholars deserves high praise. But I am tempted to conclude that, rather than fulfilling the programme

which I outlined in the main text, Darwin would have rewritten the whole *Origin of Species* anew, postponing its publication for another twenty years.

76. I would say, of the accumulation of random influences, of statistical laws. And, anyway, Laplace has precedence over Darwin.

77. Danilevsky (pt. 2, pp. 465 – 478) also accused Darwin of involuntary bias in the appraisal of facts and suppositions, of introducing vague notions (see also my § 5.8.1), and of excessive confidence in the similarity between the evolution of domestic animals and life-forms in general. I note one of Danilevsky's statements (pt. 1, p. 185), which he, regrettably, did not render specific: *Randomness and necessity are not opposites at all.*

## References

### Ch. Darwin

1. *Darwin and the voyage of the Beagle*, pt. 3 (Notebooks). London, 1945, 149 – 268.
2. Notebooks on transmutation of species. *Bull. Brit. Museum (Nat. Hist.)*. Hist. ser., vol. 2, No. 2 – 4, 1960; vol. 3, No. 5, 1967. References in my text correspond to Darwin's original paging.
3. Sketch of 1842. In Darwin, Wallace. *Evolution by natural selection*. Cambridge, 1958, 41 – 88.
4. *Essay of 1844*. Ibidem, 89 – 254.
5. *Origin of species* (1859), London – New York, 1958. [Manchester, 1995.]
6. *The various contrivances by which ... orchids are fertilised by insects* (1862). London, 1885.
7. *The movements and habits of climbing plants* (1865). London, 1885.
8. *The variation of animals and plants under domestication*, vols. 1 – 2 (1868). London, 1885. [London, 1998.]
9. *The descent of man*, etc. (1871). London, 1901.
10. *On the males and complementary males of certain cirripedes*, etc. (1873). In *Coll. papers*, vol. 2. Chicago – London, 1977, 177 – 182.
11. *Geological observations*. London, 1876.
12. *The effects of cross and self-fertilisation in the vegetable kingdom* (1876). London, 1878. [London, 1989.]
13. *Different forms of flowers*, etc. London, 1877.
14. *The power of movements in plants* (1880). New York, 1966.
15. *The formation of vegetable mould*, etc. (1887). London, 1945. [London, 1989.]
16. *Autobiography*. London, 1958.
17. *Natural selection*. Cambridge, 1975.
18. *Life and letters*, vols. 1 – 2 (1887). New York, 1897, Volume 1 contains Darwin's *Autobiography* (see also [16]), F. Darwin's *Reminiscences of my father's everyday life* (pp. 87 – 136) and T. Huxley's *On the reception of the "Origin of species"* (pp. 533 – 558). [New York, 1969.]
19. *More letters*, vols. 1 – 2. London, 1903.
20. Further unpublished letters. *Annals of Sci.*, vol. 14, No. 2, 1958 (1960), 83 – 115.

### Other authors

21. **Adanson M.**, *L'histoire des coquillages* [du Sénégal]. Paris, 1757.
22. ---, *Familles des plantes*, pt. 1. Paris, 1763 Preface publ. separately: *Histoire de la botanique*. 2<sup>nd</sup> ed. préparée par l'auteur. Paris. Publié en 1864. Imprimé en 1867.
23. ---, Examen de la question, si les espèces changent parmi les plantes etc. *Hist. Acad. Roy. Sci. Paris avec Mém. Math. et Phys.*, 1769 (1772), pp. 31 – 48 of the *Memoirs*.
- 23a. ---, *Cours d'histoire naturelle*, t. 1. Paris, 1845.

24. **Babbage C.**, Letter on the proportionate number of births of the two sexes etc. *Edinb. J. Sci.*, new ser, No. 1, 1829, 85 – 104.
25. ---, On tables of the constants of nature and art (abstract) [115a, pp. 490 – 491].
26. ---, On tables, etc. *Annual report Smithsonian Instn.* 1856 (1857), 289 – 302.
27. ---, *Passages from the life of a philosopher.* London, 1864.
28. **Baer K.**, Untersuchungen ... ob die Menge der Fische im Peipus See abnimmt. *Zh. Ministerstva gos. imushchestva*, Bd. 43, No. 6, 1852, 248 – 302 (in Russian).
29. ---, *Issledovania o sostoianii ribolovstva v Rossii* (Untersuchungen über den Zustand der Fischerei in Russland), Bde. 1 – 9. Petersburg, 1860 – 1875 (in Russian).
30. ---, *Zum Streit über den Darwinismus.* Dorpat, 1873.
31. ---, *Über den Zweck in den Vorgängen der Natur.* In *Reden*, Tl. 2, *Studien aus dem Gebiete der Naturwissenschaften* (1876). Braunschweig, 1886, 49 – 105.
32. ---, Über Zielstrebigkeit in den organischen Körpern insbesondere. *Ibidem*, 171 – 234.
33. **Bastian H. C.**, *The beginnings of life*, vol. 1. London, 1872.
34. **Beketov A. N.**, *Nravstvennost i estestvoznamie* (Morals and natural science). Petersburg, 1892.
35. **Bell P. R.**, The movement of plants, etc., this being chapter 1 of **Bell et al.**, *Darwin's biological work.* Cambridge, 1959.
36. **Bertrand J.**, *Calcul des probabilités.* Paris, 1888. [Paris, 1907. Reprints, New York, 1970, 1972.]
37. **Camper P.**, *Leçons sur l'épizootie* (1769). *Oeuvr.*, t. 3. Paris, 1803, 7 – 146.
38. ---, *Les raisons physiques pourquoi l'homme est sujet à plus de maladies que les autres animaux* (1783). *Oeuvr.*, t. 2. Paris, 1803, 291 – 448.
39. **De Candolle Alph.**, *Géographie botanique raisonnée*, tt. 1 – 2. Paris, 1855.
40. ---, *Histoire des sciences*, etc. (1873). Genève – Bale, 1885.
41. ---, Sur la méthode des sommes de température, etc. *Arch. sci. phys. natur.*, t. 53, 1875, 257 – 280; t. 54, 5 – 47.
42. **De Candolle Aug. P.**, *Théorie élémentaire de la botanique* (1813). Paris, 1819.
43. ---, *Physiologie végétale*, tt. 1 – 3, this being his *Cours de botanique*, second partie. Paris, 1832.
44. **Chambers R.**, *Vestiges of the natural history of creation* (1844). London, 1887. Publ. anonymously.
45. **Comte A.**, *Cours de philosophie positive*, t. 3. Paris, 1893. First edition, tt. 1 – 6, 1830 – 1842.
46. **Congrès international de statistique.** Paris, 1855. *C. r.* Paris, 1856.
47. **Congrès international de statistique.** Vienna, 1857. *C. r.* Vienna, 1858.
48. **Cotte L.**, *Traité de météorologie.* Paris, 1774.
49. ---, Sur l'influence de la température relativement à la végétation. In author's *Mém. sur la météorologie*, t. 1. Paris, 1788, 68 – 99.
50. **Cournot A. A.**, *Exposition de la théorie des chances*, etc. Paris, 1843. [Paris, 1984.]
51. ---, *Essai sur les fondements de nos connaissances*, etc., t. 1. Paris, 1851. [Paris, 1975.]
52. ---, *Traité de l'enchaînement des idées fondamentales dans les sciences*, etc., t. 1 (1861). Roma, 1968. [Paris, 1982, 2019.]
53. ---, *Considérations sur la marche des idées*, etc., t. 2 (1872). Paris, 1934.
54. **Cuvier G.**, *Discours sur les révolutions de la surface du globe* (1812). Paris, 1861.
55. ---, Analyse des travaux de l'Académie ... pendant l'année 1827. Partie physique. *Mém. Acad. Roy. Sci. Paris*, t. 10, 1831, ciii – cxc.
56. **Danilevsky N. YA.**, *Darwinism*, vol. 1, pts. 1 – 2. Petersburg, 1885 (in Russian).

- 57. Darwin E.**, *Zoonomia*, vol. 2 (1796). London, 1801.
- 58. Darwin G. H.** Marriages between first cousins, etc. *J. [Roy] Stat. Soc.*, vol. 38, pt. 2, 1875, 153 – 182, 344 – 348.
- 58a. Efrominson V.**, Rodosdlovnaia altruizma (The genealogy of altruism). *Novi mir*, year 47, No. 10, 1971, 193 – 213.
- 58b. Fisher R. A.**, *Genetical theory of natural selection* (1930). New York, 1958.
- 59. ---**, *Design of experiments* (1935). Edinburgh – London, 1960. [Included with separate paging in author's *Statistical method, experimental design and scientific inference*. Oxford, 1990.]
- 60. Fitz Patrick P. J.**, Leading British statisticians of the 19<sup>th</sup> century. (*J. Amer. Stat. Assoc.*, 1960). *Stud. Hist. Stat., Prob.*, vol. 2. London, 1977, 180 – 212.
- 61. Fourier J. B. J.** (editor), *Recherches statistiques sur la ville de Paris et du département de la Seine*, tt. 1 – 4. Paris, 1821 – 1829.
- 61a. Geodakyan V. A.**, Amount of pollen as a means of transmission of ecological information, etc., *Zh. Obshchei biologii*, vol. 39, No. 5, 1978, 743 – 753 (in Russian, Engl. summary).
- 62. Geoffroy Saint-Hilaire E.**, *Philosophie anatomique*, tt. 1 – 2. Paris, 1818 – 1822.
- 63. ---**, Sur les déviations organiques provoquées et observées dans un établissement des incubations artificielles. *Mém. Mus. Hist. Natur. Paris*, t. 13, 1825, 289 – 296.
- 64. ---**, Mémoire ... sur les animaux ... vivant actuellement, et les espèces antédiluviennes et perdues. *Ibidem*, t. 17, 1818, 209 – 229.
- 65. ---**, Sur le degré d'influence du monde ambiant pour modifier les formes animales etc. *Mém. Acad. Sci. Paris*, t. 12, 1833, 63 – 92.
- 66. ---**, Naturaliste. *Enc. moderne*, t. 21, 1857, 642 – 647.
- 67. Geoffroy Saint-Hilaire I.**, *Essais de zoologie générale*. Paris, 1841.
- 68. ---**, *Histoire naturelle générale des règnes organiques*, t. 2. Paris, 1859.
- 69. Glass B.**, Maupertuis. In: *Forerunners of Darwin, 1745 – 1859*. **Glass et al.** (eds.). Baltimore, 1959, 51 – 83.
- 70. ---**, Heredity and variation in the 18<sup>th</sup> century. *Ibidem*, 144 – 172.
- 71. Goethe I. W.** Die Metamorphose der Pflanzen (1790). *Werke*, Abt. 2, Bd. 6/1. Weimar, 1891, 23 – 94.
- 72. ---**, Der Verfasser theilt die Geschichte seiner botanischen Studien mit (1831). *Ibidem*, 95 – 127.
- 73. Heyde C. C. Seneta E.**, *I. J. Bienaymé*. New York, 1977.
- 73a. Hiltz V. L.**, Allis exterendum, or, the origins of the Statistical Society of London. *Isis*, vol. 69, No. 246, 1978, 21 – 43.
- 74. Humboldt A.**, *Essai politique sur le royaume de la Nouvelle-Espagne* (1811). Paris, tt. 1 – 4, 1825 – 1827.
- 75. ---**, *Atlas géographique et physique du royaume de la Nouvelle-Espagne*. Paris, 1811.
- 76. ---**, *Nova genera et species plantarum*, t. 1. Paris, 1815. Russ. transl. of preface: 1936.
- 77. ---**, Sur les lois que l'on observe dans la distribution des formes végétales. *Annales chim. phys.*, t. 1, 1816, 225 – 239.
- 78. ---**, Des lignes isothermes et de la distribution de la chaleur sur le globe. *Mém. phys., chim. Soc. d'Arcueil*, t. 3, 1817, 462 – 602.
- 79. ---**, Sur les lois ..., second version. *Dict. Sci. Natur.*, t. 18, 1820, 422 – 436.
- 80. ---**, Versuch die mittlere Höhe der Continente zu bestimmen. (*Monats)ber. Preuss.* (= Berl.) *Akad. Wiss.*, Juli 1842, 233 – 244.
- 81. ---**, *Kosmos*, Bd. 1, 1845; Bd. 3, 1850. Stuttgart – Augsburg.
- 82. Kant I.** Von den verschiedenen Racen der Menschen (1775). *Werke*, Bd. 2. Berlin, 1912, 443 – 460.

- 83. Kendall M. G.**, T. Young on coincidences. (*Biometrika*, 1968) *Stud. Hist. Stat., Prob.*, vol. 1. London, 1970, 183 – 184.
- 84. Lamarck J. B.**, *Flora français*, t. 1 (1778). Paris, 1795.
- 85. ---**, *Recherches sur les causes des principaux faits physiques*, t. 2. Paris, 1794.
- 86. ---**, Discours d’ouverture ... an VIII [1800]. *Bull. scient. de la France et de la Belgique*, t. 40, 1906, 459 – 482.
- 87. ---**, Discours d’ouverture ... an X [1802]. *Ibidem*, 483 – 521.
- 88. ---**, Discours d’ouverture ... an XI [1803]. *Ibidem*, 523 – 543.
- 89. ---**, *Philosophie zoologique*, t. 2 (1809). Paris, 1873.
- 90. ---**, Aperçu analytique des connaissances humaines (MS, 1810 – 1814?). [94, 593 – 662]. Date of compilation ascertained in same source [94, p. 842].
- 91. ---**, *Histoire naturelle des animaux sans vertèbres*, t. 1. Paris, 1815.
- 92. ---**, Espèce. *Nouv. dict. hist. natur.*, t. 10, 1817, 441 – 451.
- 93. ---**, *Manuscripts at Harvard*. Cambridge (Mass), 1933.
- 94. ---**, *Izbrannye trudy* (Sel. Works), vol. 2. Moscow, 1959.
- 95. ---**, *Système analytique des connaissances positives de l’homme*. Paris, 1820.
- 96. Laplace P. S.**, *Essai philosophique*, etc. (1814). *Oeuvr. compl.*, t. 7, No 1. Paris, 1886, separate paging. [English translation, New York, 1995.]
- 97. Linné K.**, *Filosofia botaniki* (Botanic Philosophy) (1751). Petersburg, 1805 (in Russian). [*Philosophia botanice*. Stockholm, 1751.]
- 98. ---**, *Elements of botany*, this being an abridgement of *Botanic philosophy*. London, 1775.
- 99. Maupertuis P. L. M.**, *Venus physique* (1745). *Oeuvr.*, t. 2. Lyon, 1756, 1 – 133.
- 100. ---**, *Essai de cosmologie* (1750). *Oeuvr.*, t. 1. Lyon, 1756, i – xxviii, 1 – 78.
- 101. ---**, *Système de la nature* (1751). *Oeuvr.*, t. 2, 135 – 184.
- 102. ---**, *Lettres* (1756). *Ibidem*, 185 – 340.
- 103. Packard A. S.**, *Lamarck*. New York, 1901.
- 104. Pasteur L.**, avec la collaboration de **Chamberland** et **Roux**, Sur l’étiologie du charbon (1880). *Oeuvr.*, t. 6. Paris, 1933, 254 – 263.
- 105. ---**, Une statistique au sujet de la vaccination preventive contre le charbon etc. (1882). *Ibidem*, 414 – 417.
- 106. Pearson K.**, On the fundamental conceptions of biology. *Biometrika*, vol. 1, No. 3, 1902, 320 – 344.
- 107.** Omitted.
- 108. Pearson K.**, *Darwin*. London, 1923.
- 109. ---**, *Life, letters and labours of F. Galton*, vol. 2. Cambridge, 1924.
- 110. Quetelet A.**, *Lettres sur la théorie des probabilités*, etc. Bruxelles, 1846.
- 111. ---**, *Physique sociale*, tt. 1 – 2. Bruxelles, 1869.
- 112. ---**, Des lois concernant le développement de l’homme. *Bull. Acad. Roy. Sci., Lettr., Beaux Arts Belg.*, 39<sup>e</sup> année, t. 29, 1870, 669 – 680.
- 113. ---**, *Anthropométrie*. Bruxelles, 1871.
- 114. Raikov B. E.**, *K. E. von Baer*. Leipzig, 1968. Orig. publ. in Russian (1961).
- 115. Reamur R. A.**, Observations du thermomètre etc. *Hist. Acad. Roy. Sci. Paris avec Mém. math.-phys.* 1735 (1738), 545 – 576.
- 115a.** *Report Brit. assoc. advancement sci.* 1833. London, 1834.
- 116. Rose G.**, *Reise nach dem Ural* etc., Bd. 2. Berlin, 1842.
- 117. Rouillier R C.**, The Brazilian pig (1847). [119, pp. 57 – 62] (in Russian).
- 118. ---**, Repartition des taches blanches sur les animaux domestiques. *Bull. imp. naturalistes Mosc.*, t. 27, pt. 2, No. 4, 1854, 459 – 473. Simult. publ. in Russ. [119, pp 344 – 358].
- 119. ---**, *Izbrannye biologich. proizvedenia* (Sel. biol. works). Moscow, 1954.
- 120. Ruse M.**, Natural selection in the “Origin of species”. *Stud. Hist. Philos. Sci.*, vol. 1, No. 4, 1971, 310 – 351.

- 121. Sheynin O. B.**, Newton and the classical theory of probability. *Arch. Hist. Ex. Sci.*, vol. 7, No. 3, 1971, 217 – 243.
- 122. ---**, Finite random sums, etc. *Ibidem*, vol. 9, No. 4 – 5, 1973, 275 – 305.
- 123. ---**, Mathematical treatment of astronomical observations, etc. *Ibidem*, vol. 11, No. 2 – 3, 1973, 97 – 126.
- 124. ---**, On the prehistory of the theory of probability. *Ibidem*, vol. 12, No. 2 – 3, 1974, 97 – 141.
- 125. ---**, P. S. Laplace's work on probability. *Ibidem*, vol. 16, No. 2, 1976, 137 – 187.
- 126. ---**, S. D. Poisson's work in probability. *Ibidem*, vol. 18, No. 3, 1978, 245 – 300. Excerpts in this collection.
- 127. Sneath P. H. A.**, Mathematics and classification from Adanson to the present. In *Adanson*, vol. 2. Ed., **G. H. M. Lawrence**. Pittsburgh, 1964, 471 – 498.
- 128. Timiriazev K. A.**, *Istoricheskiy metod v biologii* (The historical method in biology), 1922. *Soch.*, vol. 6. Moscow, 1939, 13 – 237.
- 129. Tolstoy L. N.**, *Tak chto zhe nam delat?* (So what shall we do?), 1886. *Poln. sobr. soch.*, vol. 25. Moscow, 1937, 182 – 411 (in Russian).
- 130. Tutubalin V. N.**, *Granitsdy primenimosti* (The limits of usage (statistical methods and their possibilities). Moscow, 1977 (in Russian).
- 131. Valt Maie Kh**, Baer's ecological researches and the concept of the struggle for existence. In: *Peterburgskaia akademiya i Estonia* (The Petersburg academy and Estonia). Tallinn, 1978, 102 – 119. In Russian.
- 132. Wallace A. R.**, On the tendencies of varieties to depart indefinitely from the original type (1858). In author's *Contributions to the theory of natural selection*. London, 1870, 26 – 44.
- 133. ---**, *The geographical distribution of animals*, vols. 1 – 2 (1876). London, 1962.
- 134. ---**, *Island life*. London, 1880.
- 135. Wilkie J. S.**, The idea of evolution in the writings of Buffon. *Annals of science*, 1956, vol. 12, No. 1, 48 – 62; No. 3, 212 – 227; No. 4, 255 – 266.
- Andersson T.** (1929), Wilhelm Johannsen, 1857 – 1927. *Nordic Stat. J.*, vol. 1, 349 – 350.
- De Vries H.** (1905), Evidence of evolution. *Annual Rept Smithsonian Instn* for 1904, 389 – 396.
- Johannsen W.** (1912), Biology and statistics. *Nordic Stat. J.*, vol. 1, 1929, 351 – 361.
- Sheynin O.** (2014), Randomness and determinism etc. *Silesian Stat. Rev.*, No. 12 (18), 57 – 74.

### III

#### A. L. Tzikalo

A. M. Liapunov. Moscow, 1988 (Excerpts)

**P. 72.** For his final year in Kharkov, the faculty entrusted Liapunov the lectures in the theory of probability which acquired essential importance for the development of mechanics, statistical physics and various applications<sup>1</sup>. It and in particular its limit theorem attracted Liapunov long ago, even from his student years when he attended Chebyshev's course in that theory<sup>2</sup>.

Chebyshev outlined the proof of the limit theorem for sums of (independent) random variables but stipulated that its derivation was not rigorous in that the assumptions made were not accompanied by a determination of the boundaries of the ensuing error<sup>3</sup>.

Later Liapunov largely occupied himself with this topic and attempted to provide a more general and rigorous proof of that theorem. Markov communicated two of his memoirs to the Petersburg Academy of Sciences and they were published in 1900 and 1901<sup>4</sup>. Liapunov applied an entirely new method of characteristic functions<sup>5</sup>.

**P. 73.** This method opened up such wide possibilities that it by right occupies a central place in probability theory and it is opportune to note that Liapunov's research was only a short episode in his creative work.

It is impossible not to point out the rapid extension of the number of periodicals in which Liapunov published his work after 1893. Apart from the *Soobshcheniya* of the Kharkov mathematical society there were *Matematicheskii Zbornik*, *Proceedings* of the physical section of the *Obshchestvo liubitelei estestvoznania* (Society of Lovers of Natural Science), *antropologii i etnografii*, C. r. of the Paris Academy of Sciences, *J. pure et appl. math.* This fact certainly testifies to the attention paid by the leading national and foreign mathematicians to the work of Liapunov. Markov communicated his work to the Petersburg Academy, and E. Pickard, to the Paris Academy. He was member and later the President of that academy and a Fellow of the Royal Society.

**P. 74.** After Chebyshev's death, Liapunov published an essay dedicated to his memory, later reprinted as a separate booklet, in the *Soobshchenia*. He presented a vivid and precise portrait of the great scientist, expounded the main facts of his life and work, surveyed his most important work and paid attention to Chebyshev's inventions connected with analytic research in the theory of mechanisms.

All his life Liapunov reverentially regarded Chebyshev and the memory about him<sup>6</sup>. While still in Kharkov, Liapunov also translated three of his papers from French which were included in the first volume of Chebyshev's *Oeuvres* (1899 – 1907) and two more papers in volume 2; there also, appeared a paper which he translated from Russian.

**Pp. 80 – 81.** In 1901 – 1902, a commission elected by the Council of Kharkov University put forward a proposal for changing the

reactionary status of 1884 of Russian universities (Bagalei et al, 1906, pp. 260 – 262). Liapunov and V. A. Steklov<sup>7</sup> were among the twelve members of that commission, and here is an excerpt from the Note of that commission.

*The present purely bureaucratic order of universities is abnormal. ... We consider that the main evil is the spirit of indifference, which occurred as a necessary result of the Ministry's distrust of universities, and therefore a severe excessive regulation. Universities have become offices ... Trust of universities should in the first place lead to the substitution of bureaucratic supervision by freedom of scientific research and teaching.*

The Commission put forward and justified the following demand (Ibidem, pp. 263 – 264):

*The basis of the entire organization of a university should be the elective principle, and the Councils ought to elect professors and the teaching staff which can only be dismissed by a decision of the [appropriate] council. ... But nothing changed.*

**Pp. 84 – 85.** Here the authors describe Liapunov's answer to Nekrasov who had allegedly found mistakes in his work. I have described this and the related material, see Sheynin (2017, §§ 14.4 and 14.5)<sup>8</sup> and Liapunov (1901).

**Pp. 96 – 97.** The authors reprinted Liapunov's letter of 23 February 1905 to the President of the Petersburg Academy of Sciences, the Grand Duke Konstantin Konstantinovich, grandson of Nicholas I.

*Your Imperial Highness,*

*You were pleased to apply your circular letter to those academicians who had signed the Note [about the situation in Russian universities] and to censure them for their action. You consider their Note unlawful and capable of strengthening the students' unrest. At the same time the letter accuses the signatories of a careless discharge of their duty.*

*Such an attitude of your Imperial Highness to their action could not have failed to provoke a desire to express, directly and openly, how they see this business. For my part, I believe it my moral duty to submit the following for the kind consideration of Your Highness.*

*When signing that Note, we did not and do not think that we had violated the law. We suppose that each citizen has the right to express frankly his opinion about an issue closely related to the occupation to which he had devoted himself. So is there anyone except scientists and professors to whom enlightenment in our fatherland is closer?*

*Our conclusion quite naturally coincides with the resolution of the congress of the zemstvo figures since the scientific and pedagogic business suffers most of all from the bureaucratic system which dominates us. And we had openly stated only that which is desired by most of all the enlightened people in our fatherland and which had been recently proclaimed from the height of the throne<sup>9</sup>.*

*We are far from thinking that our action violated some legality. On the contrary, we thought it our direct duty to express what we see as a way out from the present difficult situation. We owed this moral duty to our fatherland to which we are beholden for our high standing and to our people for their means which provision our upkeep.*

*When signing that Note we were also far from thinking that we can strengthen the students' unrest. Being closely familiar with the mood of our young people and the causes which provoke the unrest, we were unable to think that way. On the contrary, as I believe, our Note is capable of somewhat abating the unrest which would have undoubtedly acquire an essentially sharper form had our professors and scientists remained dumb during the present moment of our historical life.*

*And, finally, with respect to the rebuke for carelessly discharging our direct duty. What ground has this heavy accusation? Is it really the sole act of signing that Note? There are moments when honest people should not and cannot keep silent, when even those who had only devoted themselves to science and had never before been interested in policy are unable to remain apathetic to public issues. And if people of science express themselves in such moments it does not mean that they forgot their direct duties. Indeed, the signatories include not a few generally known figures, professors who earned general respect by their activities and scientists who acquired world fame by their work.*

*These, your Highness, are the thoughts which I considered my duty to express.*

*Entirely devoted to your Imperial Highness, academician Liapunov.  
23 February 1905*

**P. 166.** Markov boundlessly trusted Liapunov and relied on his outstanding aptitude, unlimited carefulness and thoroughness in considering scientific issues. Their correspondence (Archive, Russian Academy of Sciences, Fond 257, Inventory 1, No. 20) testifies that he often turned to him for advice, for checking some calculations and proofs. Markov possibly felt that Liapunov more than he himself was capable of *positive* criticism and will be able to soften his own numerous critical comments and, in addition, to show authors proper directions or to take upon himself the filling of gaps and completion of a research. So it happened with the work of Sophia Kovalevskaya ...

Markov was only a year older than Liapunov but regarded him with a touching and almost fatherly sympathy. He actively attempted to propose Liapunov for an academician<sup>10</sup>. Indeed, Markov considered him most suitable for occupying the chair which was left vacant after the death of the great Chebyshev. Markov also took care after Liapunov's material needs. ...

**P. 171.** We see an essential distinction between the two outstanding researchers, Liapunov and Poincaré. Exactly because of his clearly expressed versatility and wide scope of scientific interests, Poincaré, unlike Liapunov, was not inclined to check and recheck his results, to fill up the gaps in a complicated work or to remove defects even after noticing them.

**P. 173.** An essential difference in their approaches and in the requirements to their own results and conclusions was revealed already during the first stage of the correspondence between Liapunov and Poincaré<sup>11</sup>. The latter thought that *in mechanics, it is impossible to require the same rigor as in pure analysis*, whereas Liapunov stated that any problem in physics or mechanics, one formulated quite

definitely in the mathematical sense, becomes *a problem of pure analysis and ought to be treated as such*. Steklov (Smirnov, 1992, p. 20) wrote:

*Innuendos and inaccuracies and sometimes non-rigorous proofs or even only hints at proofs often occur in the works of Poincaré whereas all Liapunov's considerations achieve perfection since he always speaks about and only about what can be proven with an irreproachable rigour.*

**P. 175.** [About Poincaré (read 1900)]. *To my greatest surprise, I did not find there anything essential. Its large part is devoted to expounding (and I ought to add, quite disorderly) the results known long ago. As to the issue which interests me, Poincaré only repeats in a very brief form that which he discussed in his in his old memoir of 1886.*

### Notes

1. The theory of probability had been very long ago applied to various branches of natural science (biology, medicine, meteorology, astronomy, see various chapters of Sheynin (2017).

2. Here and below, read *central limit theorem* (CLT) instead of *limit theorem*.

Liapunov not only attended Chebyshev's lectures in the theory of probability, he listened and wrote them down. This is a long story, see Sheynin (2017, beginning of § 13.2).

3. That was Chebyshev's remark about Poisson's proof of the CLT. And he did much more than just outlining the proof of that theorem (Sheynin 2017, § 13.1-4).

4. See appended Bibliography.

5. In 1810 – 1811 Laplace applied a particular case of a characteristic function (Sheynin 2017, p. 281).

6. By the end of the 19<sup>th</sup> century Russian mathematics began to lag behind. In Europe, mathematical analysis had been essentially developed by such scientists as Weierstrass whereas in Russia mathematicians remained mesmerized by Chebyshev. It was apparently Novikov (2002, p. 330) who let the cat out of the bag:

*In spite of his splendid analytical talent, Chebyshev was a pathological conservative.*

And here is Liapunov (1895/1946, pp. 19 – 20), see Sheynin (2017, p. 226) who should have known better: he called Riemann's ideas "extremely abstract"; his investigations, "pseudo-geometric" and sometimes, again, too abstract and having nothing in common with Lobachevsky's "deep geometric studies". Liapunov obliquely recalled Klein, but disregarded him. Klein had in 1871 presented a unified picture of the non-Euclidean geometry in which the findings of Lobachevsky and Riemann appeared as particular cases.

7. Steklov (1863/1864 – 1926), Liapunov's student, the future vice-president of the Russian Academy of Sciences.

A telling episode characterizes Russia's similar events of 1901 – 1902 (Slutsky 2010, pp. 280 – 283): In his autobiography of 1939 Slutsky wrote that he then actively participated in the students' unrest and was prohibited from entering (once more) any Russian higher academic institution. (He studied in the Munich Polytechnic School but did not like his new speciality and in 1905 was able to resume his former education in Russia.)

8. Nekrasov considered the CLT for the case of large deviations and justified this development by a more general interpretation of Chebyshev's study. Liapunov, however, refused to agree with Nekrasov (Sheynin 2017, p. 240). Cf. Note 6.

9. I only know about the later (October 17/30, 1905) Manifesto which ended unlimited autocracy and ushered a period of (lame) constitutional monarchy. In pursuing their own aims, the Bolsheviks disseminated a mock verse:

*The Tsar got a fright, // Signed a Manifesto,  
Freedom for the dead, // Arrest for the rest-0.*

10. In connection with this attempt, in 1901 he asked Liapunov about his ties with the West and Liapunov named seven scientists who had quoted him, Poincaré and Appell among them (Archive Russian Academy of Sciences, Fond 173, Inventory 1, Delo 11, p. 17).

11. See the note by the author (p. 174): Liapunov's letter to Steklov of 21 February 1903 (Smirnov 1992, pp. 335 – 337).

## Bibliography

**Bagalei D. I. et al** (1906), *Kratkiy ocherk istorii Kharkovskogo universiteta* (Short Survey of the History of Kharkov Univ.), 1805 – 1905. Kharkov.

**Chebyshev P. L.** (1853), *Sur l'intégration des différentielles irrationnelles ... Oeuvres*, t. 1.

--- (1857), *Sur l'intégration des différentielles ... Ibidem.*

--- (read 1861), *Sur une modification du parallélogramme articulé de Watt. Ibidem.*

--- (1874), *Sur le formes quadratiques. Oeuvres*, t. 2.

--- (1874), *Sur les valeurs limites des intégrales. Ibidem.*

--- (1881), *Sur les plus simples .parallélogramme qui fournissent un mouvement rectiligne ... Ibidem.*

--- (1899 – 1907), *Oeuvres*, tt. 1 – 2. Pétersbourg. New York, 1962. In Russian and French.

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

--- (1901), *Nouvelle forme du théorème limite de probabilités. Mém. Imp. Akad. Nauk*, Cl. phys.-math., t. 12, No. 5. Separate paging.

--- (1901), *An answer to Nekrasov. Zapiski Khark. Univ.*, t. 3, pp. 51 – 63. In Russian. **S, G**, 4.

--- (1901), *Sur une théorème du calcul des probabilités. C. r. Acad. sci. Paris*, t. 132, pp. 126 – 128.

--- (1901), *Une proposition générale du calcul des probabilités. Ibidem*, pp. 814 – 815.

**Novikov S. P.** (2002), *The second half of the 20<sup>th</sup> century ... Istoriko-Matematicheskie Issledovania*, t. 7(42), pp. 326 – 356. In Russian.

**Slutsky E. E.** (2010), *Collected Statistical Papers*. Berlin. **S, G**, 40.

**Smirnov V. I.** (1992), *Biografia A. M. Liapunova*. Moscow. In Russian.

**Poincaré H.** (read 1900), *Figures d'équilibre d'une masse fluide*.

## IV

### Alph. De Candolle

#### On a Dominant Language for Science

*Annual Rept, Board of Regents, Smithsonian Inst.*  
for 1874 (1875), pp. 239 – 248. Translated from author's  
*Histoire des sciences et des savants depuis deux siècles*. Genève, 1873 by Miss Miers

At the period of the Renaissance, Latin was the language employed by all the learned men of Europe. It had been carefully preserved by the Romish Church; and at that time not one of the modern languages presented a sufficiently rich literature to become its rival. But at a later period the Reformation disturbed the unity of the Romish influence. Italian, Spanish, French and English successively gained regular idioms and became rich in literary productions of every kind. And at last, eighty or one hundred years ago at most, the progress of science caused the inconvenience of the use of Latin to be felt.

It was a dead language, and in addition, it was wanting in clearness owing to its inversions, to its abbreviated words, and to the absence of articles. There existed at that time a general desire to describe the numerous discoveries that were being made, and to explain and discuss them without the necessity of seeking for words. The almost universal pressure of these causes was the reason for the adoption of modern languages in most sciences, natural history being the only exception.

For this, Latin is still employed, but only in descriptions, a special and technical part where the number of words is limited and the constructions very regular. Speaking truly, what naturalists have preserved is the Latin of Linnaeus, a language in which every word is precise in meaning, every sentence arranged logically clearly, and in a way employed by no Roman author. Linnaeus was not a linguist. He knew but little even of modern languages and it is evident that he struggled against many difficulties when he wrote in Latin. With a very limited vocabulary and a turn in mind which revolted equally from the periods of Cicero and the reticence of Tacitus, he knew how to create a language precise in its terms, appropriate to the description of forms and intelligible to students.

He never made use of a term without first defining it. To renounce this special language of the learned Swede would be to render descriptions less clear and less accessible to the *savants* of all nations [of any ...]. If we attempt to translate into the Latin of Linnaeus certain sentences in modern floras written in English or German, we quickly perceive a want of clearness. In English, the word *smooth* applied equally to *glaber* and *laevis*<sup>1</sup>. In German, the construction of sentences indicating generic or other characters is sometimes so obscure that in certain cases I have found it impossible to have them put into Latin by a German, a good botanist who was better acquainted than myself with both languages. It would be still worse if authors had not introduced many words purely Latin into their language. But, exclusive of paragraphs relative to characters, and

wherever successive phenomena or theories are in question, the superiority of modern languages is unquestionable. It is on this account that, even in natural history, Latin is every day less employed.

The loss, however, of the link formerly established between scientific men of all countries has made itself felt. From this has arisen a very chimerical proposal to form some artificial language which should be to all nations what writing is to the Chinese. It was to be based on ideas, not words. The problem has remained quite devoid of solution, and even were it possible, it would be so complicated an affair, so impracticable and inflexible, that it would quickly drop into disuse.

The wants and the circumstances of each epoch have brought about a preference for one or other of the principal European languages as a means of communication between enlightened men of all countries. French rendered this service during two centuries. At present, various causes have modified the use of this language in other countries, and the habit has been almost everywhere introduced that each nation should employ its own tongue. We have therefore entered upon a period of confusion. What is thought to be new in one country is not so to those who read books in other languages.

It is vain to study living languages more and more; you are always behindhand in the complete knowledge of what is being published in other countries. Few persons are acquainted with more than two languages, and if we try to pass beyond a certain limit in this respect, we rob ourselves of time for other things. There is a point at which the study of the means of knowledge hinders our learning. Polyglot discussions and conversations do not answer the intentions of those who attempt them. I am persuaded that the inconvenience of such a state of things will be more and more felt. I also believe, judging by the example of Greek as used by the Romans and French in modern times, that the need of a prevailing language is almost always recognized. It is returned to from necessity after each period of anarchy. To understand this we must consider the causes which make a language preferable and those which spread its employment in spite of any defects it may possess.

Thus, in the seventeenth and eighteenth centuries motives existed for the employment of French in preference to Latin throughout Europe. It was a language spoken by the greater part of the educated men of the period, a language tolerably simple and very clear. It had an advantage in its resemblance to Latin which was then widely known. An Englishman, a German was already half acquainted with French through his knowledge of Latin. A Spaniard, an Italian, was three parts [quarters] advanced in his study of the language. If a discussion were sustained in French, if books or translations written/made in this language, all the world understood.

In the present century, civilization has much extended north of France and population has increased there more than to the south. The use of the English tongue has been doubled by its extension into America. The sciences are more and more cultivated in Germany, in England, in the Scandinavian countries and Russia. The scientific centre of gravity has advanced from south toward the north.

Under the influence of these new conditions a language can only become predominant by presenting two characters. First, it must possess sufficient German and Latin words or forms to be within reach at once of the Germans and of the people who make use of Latin tongues. Secondly, it must be spoken by considerable majority of civilized people. In addition to these two essential conditions it would be well for the definitive success of a language that it should also possess the qualities of grammatical simplicity, of conciseness and clearness.

English is the only language which may, in fifty or a hundred years offer all these conditions united.

The language is still half German and half Latin. It possesses German words, German forms, and also French words, and a French method of constructing sentences. It is a transition between the principal languages used at present in science, as French was formerly between Latin and several of the modern languages.

The future extension of the Anglo-American tongue is evident. It will be rendered inevitable by the movement of the populations in the two hemispheres. Here is the proof which it is easy to give in a few words and a few figures.

At the present time the population stands thus (Almanach de Gotha, 1871<sup>2</sup>). English-speaking peoples in England, 31,000,000<sup>3</sup>; in the United States, 40,000,000, in Canada etc., 4,000,000; in Australia and New Zealand, 2,000,000; total, 77,000,000.

German-speaking peoples in Germany and a portion of Austria, 60,000,000; in Switzerland (German cantons) 2,000,000; total, 62,000,000.

French-speaking peoples in France, 36,500,000; in Belgium (French portion), 2,500,000; in Switzerland (French cantons), 500,000; in Algeria and colonies, 1,000,000; total 40,500,000.

Now, judging by the increase that had taken place in the present century, we may estimate the probable growth of population as follows<sup>3</sup>:

In England it doubles in fifty years, therefore, in a century (in 1970) it will be 124,000,000. In the United States, in Canada, in Australia, it doubles in twenty five. Therefore, it will be 736,000,000. Probable total of the English-speaking race in 1970, 860,000,000.

In Germany, the northern population doubles in fifty six to sixty years; that of the south, in one hundred and sixty seven years. Let us suppose one hundred years for the average. It will probably be, in 1970, for the countries of German speech, about 124,000,000.

In the French-speaking countries the population doubles in amount one hundred and forty years. In 1970 therefore it will probably amount to 69,500,000.

Thus the three principal languages spoken at the present time will be spoken a century hence with the following progression:

The English tongue will have increased from 77 to 860 millions.

The German tongue will have increased from 62 to 124 millions.

The French tongue will have increased from 40.5 to 69.5 millions.

The individuals speaking German will form a seventh part and those speaking French, a twelfth or thirteenth part of those of English

tongue and both together will not form a quarter of the individuals speaking English. The German or French countries will then stand toward those of English speech as Holland or Sweden do at present with regard to themselves [to Germany and France]. I am far from having exaggerated the growth of the Anglo-Australian-American populations. Judging by the surface of the countries they occupy, they will long continue to multiply in large proportion. The English language is besides more diffused than any other throughout Africa and Southern Asia. America and Australia are not, I confess, countries in which the culture of letter and science is so much advanced as in Europe, and it is probable that for as length of time, agriculture, commerce and industry will absorb all the most active energies. I acknowledge this. But it is no less a fact that so considerable a mass of intelligent and educated men will weigh decisively on the world in general.

These new peoples, English in origin, are mingled with a German element which in regard to intellectual inclinations, counterbalances the Irish. They have generally a great eagerness for learning and for the application of discoveries. They read much. Works written in English or translated into that tongue would, in a vast population, have a very large sale. This would be an encouragement for authors and translators that is offered by neither the French nor the German language. We know in Europe to what degree difficulties exist in the publication of books on serious subjects, but open an immense market to publishers and works on the most special subjects will have a sale.

When translations are read by ten times as many people as at present, it is evident that a greater number of books will be translated, and this will contribute in no small degree toward the preponderance of the English language. Many French people already buy English translations of German books, just as Italians buy translations in French. If English or American publishers would adopt the idea of having translations made into their language of the best works that appear in Russian, Swedish, Danish, Dutch etc. they would satisfy a public dispersed over the whole world and particularly the numerous Germans who understand English. Yet we are but at the beginning of the numerical preponderance of the English-speaking populations.

The nature of the language does not, at first sight, appear to have very great influence on its diffusion. French was preferred for two centuries and yet Italian was quite as clear, more elegant, more harmonious, had more affinity with Latin, and for a length of time had possessed a remarkable literature. The number, the activity of the French and the geographical position of their country were the causes of their preponderance. Yet the qualities of a language, especially those preferred by the moderns, are not without their influence. At the present time, briefness, clearness, grammatical simplicity are admired.

Nations, at least those of our Indo-European race, began by speaking in an obscure, complicated manner. In advancing, they have simplified and made their language more precise. Sanscrit and Basque, two very ancient languages, are exceedingly complicated. Greek and Latin are so in less degree. The languages derived from Latin are clothed in clearer and simpler forms. I do not know how

philosophers explain the phenomenon of the complication of language at an ancient period, but it is unquestionable. It is easier to understand the subsequent simplifications. When an easier and more convenient method of acting or speaking has been arrived at, it is naturally preferred. Besides, civilization encourages individual activity and this necessitated short words and short sentences. The progress of the sciences, the frequent contact of persons speaking different languages who find a difficulty in understanding each other, lead to a more and more imperious need for clearness. You must have received a classical education to avoid the perception of absurdity in the construction of an ode of Horace. Translate it literally to an uneducated workman, keeping each word in its place, and it will have to him the effect of building whose entrance-door is on the third story. It is no longer a possible language, even in poetry.

Modern languages have not all, to the same degree, the advantages now demanded of clearness, simplicity and briefness.

The French language has shorter words and less complicated verbs than the Italian. This in all probability has contributed to its success. The German has not undergone the modern revolution by which each sentence or portion of a sentence begins with the principal word. Words are also cut in two and the fragments dispersed. It has three genders whereas French and Italian have but two. The conjugations of many verbs are rather complicated. Nevertheless, modern tendencies weigh with the Germans, and it is evident that their language is becoming a little modified. Scientific authors especially exert themselves to attempt the direct modes of expression and the short phrases of other countries in the same way that they have abandoned the Gothic printed letters. Should they correspond with strangers, they often have the politeness to write in Latin characters. They willingly introduce in their publications terms taken from foreign languages, modifications sometime merely of form, occasionally fundamental. These attest the modern spirit and the enlightened judgement of the learned men so numerous in Germany. Unhappily, the modifications of form have no great importance, and the fundamental changes take place very slowly.

The more practical English language shortens sentences and words. It willingly takes possession of foreign words, as German does. But of *cabriolet* it makes *cab*; of *memorandum* it makes *mem* [memo]. It makes use only of indispensable and natural tenses – the present, the past, the future and the conditional. There are no arbitrary distinctions of genders; animated objects are masculine or feminine, the others are neuter. The ordinary construction is so sure to begin with the principal idea, that in conversation you may often dispense with the necessity of finishing your sentences. The chief fault of the English language, its inferiority in comparison with German or Italian, consists in an orthography absolutely irregular and so absurd that children take a whole year in learning to read<sup>4</sup>. The pronunciation is not well articulated, not well defined. I shall not go as far as Madame Sand in her amusing imprecations on this point. But there is truth in what she says. The vowels are not distinct enough. But, in spite of these faults, English, according to the same clever writer, is a well-expressed

language, quite as clear as any other, at least when English people choose to revise their MSS which they will not always do, they are in such a hurry!

English terms are adapted to modern wants. O you wish to hail a vessel, to cry *stop* to a train, to explain a machine, to demonstrate an experiment in physics, to speak in a few words to busy and practical people, it is the language *par excellence*. In comparison with Italian, with French, and above all with German, English has the effect, to those who speak several languages, of offering the shortest cut from one point to another.

I have observed this in families where two languages are equally well known, which often occurs in Switzerland. When the two languages are German and French, the latter almost always carries the day. Why? I asked of a German-Swiss established in Geneva.

*I can scarcely tell you; at home we speak German to exercise my son in the languages, but he always falls back into French of his comrades. French is shorter, more convenient.*

Before the events of 1870<sup>5</sup>, a great Alsatian manufacturer sent his son to study at Zürich. I was curious to know the reason why.

*We cannot induce our children to speak German, with which they are quite as familiar as with French. I have sent my son to a town where nothing but German is spoken that he may be forced to speak it.*

In such preferences you must not look for the causes in sentiment or fancy. When a man has choice of two roads, one straight and open, the other, crooked and difficult to find, he is sure to take, almost without reflection, the shorter and more convenient. I have also observed families where the two languages known in the same degree were English and French. In this case the English maintained supremacy, even in a French-speaking land.

It is handed down from one generation to another. It is employed by those who are in haste or who want to say something in as few words as possible. The tenacity of French or English families established in Germany in speaking their own language and the rapid disappearance of German in the German families established in French or English countries may be explained by the nature of the languages rather by the influence of fashion or education.

The general rule is this: In the conflict of two languages, everything else being equal, it is the most concise and the simplest that conquers. French beats Italian and German. English beats the other languages. In short, it need only be said that the simpler the language is, the easier it is to be learned, and the quicker can it be made available for profitable employment.

The English language has another advantage in family use: its literature is the most suitable to feminine tastes; and everyone knows how great is the influence of mothers on the language of children. Not only do they teach what is called *the mother tongue*, but often when well educated they feel pleasure in speaking a foreign language to their children. They do so gaily, gracefully. The young lad who finds his language-master heavy, his grammar tiresome, thinks very differently when his mother, his sister, or his sister's friend addresses herself to him in some foreign tongue. This will often be English, and

for the best of reasons: there is no language so rich in works (written in the spirit of true morality) upon subjects which are interesting to women – religion, education, fiction, biography, poetry etc.

The future preponderance of the language spoken by English, Australians and Americans thus appears to me assured. The force of circumstances leads to this result, and the nature of the language itself must accelerate the movement.

The nations who speak the English tongue are thus burdened with a responsibility which is well they should recognize at once. It is a moral responsibility toward the civilized world of the coming centuries. Their duty, as it is also their interest, is to maintain the present unity of the language at the same time admitting the necessary or convenient modifications which may arise under the influence of eminent writers, or be arranged by common consent. The danger to be feared is that the English language may, before another century has passed, be broken up into three languages which would be in the same relation to each other as are Italian, Spanish and Portuguese or as Swedish and Danish.

Some English authors have a mania for making new words. Dickens has invented several. Yet the English language already possesses many more words than the French and the history of literature shows that there is a greater need to suppress than to add to the vocabulary. No writer for three centuries past has employed nearly as many different words as Shakespeare, therefore there must have been many unnecessary ones. Probably every idea and every object had formerly a term of Saxon origin, and one of Latin or French origin, without counting Celtic or Danish words. The very logical operation of time has been to suppress the double or triple words. Why re-establish them? A people so economical in its use of words does not require more than one term for each thing<sup>6</sup>.

The Americans, on the other hand, make innovations of accent or orthography (they almost always spell *labour labor* and *harbour harbor*). The Australians will do the same if they do not take care. Why should not all possess the noble ambition of giving to the world one uniform concise language supported by an immense literature and spoken in the next century by eight hundred or one thousand million of civilized men? To other languages it could be as a vast mirror in which each would become reflected thanks to newspapers and translations, and all the friends of intellectual culture would have a convenient medium for the interchange of ideas. It would be rendering an immense service to future races and at the same time the authors and men of the English-speaking race would give a strong impulsion to their own ideas. The Americans, above all, are interested in this stability since their country is to be the most important of those of English tongue. How can they acquire a greater influence over Old England than by speaking her language with exactness?

The liberty of action permitted among people of English race adds to the danger of a division in the language. Happily however certain causes which broke up the Latin language do not exist for English nations. The Romans conquered nations whose idioms were maintained or re-appeared here and there in spite of administrative

unity. The Americans and Australians, on the contrary, have before them only savages who disappear without leaving any trace. The Romans were conquered and dismembered in their turn by the barbarians. Of their ancient civilization no evidence of unity remained unless it was in the Church which had itself felt the influence of the universal decline.

The Americans and Australians possess many flourishing schools. They have the literature of England as well as their own. If they choose, they can wield their influence by means of maintaining the unity of the language. Certain circumstances make it possible for them to do so; thus, the teachers and professors mostly come from the States of New England. If these influential men truly comprehend the destiny of their country, they will use every effort to transmit the language in its purity. They will follow classic authors and discard local innovations and expressions. In this question of language real patriotism (or, if you will, the patriotism of Americans really ambitious for their country) ought to be, to speak the English of Old England, to imitate the pronunciation of the English and to follow their whimsical orthography until changed by themselves. Should they obtain this of their countrymen, they would render to all nations and to their own an unquestionable benefit for futurity.

The example of England proves the influence of education upon the unity of a language. It is the habitual contact of educated people and the perusal of the same books which, little by little, is causing the disappearance of Scotch words and accent. A few years more, and the language will be uniform throughout Great Britain. The principal newspapers edited by able men also exercise a happy influence in preserving unity. Whole columns of the *Times* are written in the language of Macaulay and Bulwer and are read by millions of people. The result is an impression which maintains the public mind in a proper literary attitude.

In America the newspaper articles are not so well written but the schools are accessible to all classes and the universities count among their professors men especially accomplished in their use of the English tongue. If ever there should arise a doubt in the opinions of the two countries as to the advisability of modifying the orthography, or even making changes in the language it would be an excellent plan to organize a meeting of delegates from the principal universities of the Three Kingdoms [Great Britain], of America and Australia, to propose and discuss such changes. Doubtless they would have the good sense to make as few innovations as possible. And thanks to common consent the advice would possibly be followed. A few modifications in the orthography alone would render the English language easier to strangers and would contribute toward the maintenance of unity in pronunciation throughout Anglo-American countries.

**NOTES BY DR. JOHN EDWARD GRAY  
OF THE BRITISH MUSEUM**

It may be observed, in addition, that the people who use the English language in different parts of the world are a reading and book-buying

people and especially given to the study of quasi-scientific books as is proved by the fact of extensive sale which they command.

In support of this assertion I may quote the Baron Férussac's view of Wood's *Index Testaceologicus* in the *Bull. Sci. Nat.*, Paris, 1820, p. 375. He remarks:

*We observe with interest the number of subscribers that exist in England for an octavo volume on shells costing 186 francs. It is a curious fact which booksellers and authors will appreciate as it will afford them the means of seeing how a return is obtained for their outlay on such works in England compared with other countries. The number of subscribers is 280 of which 34 are females and 6 foreigners. Certainly all the rest of Europe could not produce as many, nor perhaps even the half of that number.*

How much more astonished would Férussac have been if informed that these were only the subscribers before publication and that 1,000 copies were sold! Since 1829 the sale of scientific books has much increased as is shown for example by the many editions of the work of Lyell and other naturalists, each edition being of 1,000 copies.

Most books in France and other continental countries can only be published when the government furnishes the cost, and they are chiefly published in an expensive form as a national display and are almost confined to their public libraries except for the sale of copies that are bought by English collectors.

In England, such works are generally published by individual enterprise and depend on the general public for their support and are published in a style to suit the different classes. Thus there are works of luxury for the rich, often published by individuals who confine themselves to the production of that class of books; very cheap works for the student and mechanic; and books of all intermediate grades produced by the regular publishers. The females of all grades are extensive readers of this class of books which I believe is chiefly the case with English speaking races.

Some of the scientific Swedes and Russians have published their papers in the English language or appended an abstract to them as [...]. The Danes and Dutch often publish *their* scientific papers in French as [...] who themselves read and write English [...].

De Candolle himself uses the French language with a very English construction but we believe that his work would have commanded the greatest number of readers if written in the English language which he reads and writes so fluently.

See also Galton's interesting article on the Causes which create scientific men (*Fortnightly Review*, March 1873, p. 346) which contains some interesting observations on de Candolle's work.

### Notes

1. The word *glaber*, in botany, means bald or not hairy [...] and *laevis*, smooth, not rough, but I know that they have been carelessly translated *smooth*, as de Candolle implies. J. T. G. (Gray).

2. No notice is taken of the English-speaking people in India and the East. J. E. G.

3. Almanach de Gotha 1870, p. 1039. Author.

4. Surprised, on one occasion, by the slowness with which intelligent English children learned reading [in a generalized sense!], I inquired the reason. Each letter

has several sounds or you may say that each sound is written in several ways. It is therefore necessary to learn reading word for word. It is an affair of memory.  
Author.

5. The Franco-Prussian war of 1870 – 1871. O. S.

6. A clever English writer has just published a volume on the institutions of the people called *Swiss* in English. He names them *Switzers*. Will there soon be *Deutchers*? Author

### **The figures mentioned**

**Bulwer-Lytton G. E. L.** (1803 -1873), writer, historian

**Lyell Ch.** (1797 – 1875), naturalist. His book on geology had great influence.

**Macaulay T. B.** (1800 – 1859), poet, historian, political figure

**Sand G.**, pen-name of female writer A. Dupin.

J. D. Sarma, B. Shapell

Lincoln and the Jews (excerpts)

New York, 2015

xiiL. Fully half a life L had Jewish friends and acquaintances, he repeatedly intervened on Jew's behalf, most famously, in 1862, when he overturned Ulysses G. Grant's order expelling "Jews as a class" from his war zone.

xiiiR. In 1809, when L was born, scarcely 3,000 Jews lived in the entire United States. By 1865 [more than 150 thousand]. [The authors never write, for example, 150 or 30 thousand.]

L was personally broadened by encounters with Jews and also worked to broaden America so that Jews might gain acceptance as equals nationwide. He himself appointed the first Jewish chaplain to the armed forces and placed many other Jews into positions of authority as well.

xiiiR. [He consistently attempted] to redefine America through phrases like "this nation under God" that embraced Jews and other non-Christians as insiders.

2L. Like his Puritan ancestors (and unlike Catholics of his time), he considered the Hebrew Bible an equal partner with the *New Testament*. He quoted and referenced the *Old Testament* about a third more times than he did the *New*. And in referencing the Deity some 420-plus times, he used the phrase *Saviour* but six. According to the *Coll. Works of Abraham Lincoln* [no exact reference], he mentioned Christ directly only once. He never referred directly to Jesus.

2R. In his lecture of 1858 L mentioned the *New Testament* scarcely twice, but referred to the *Old Testament* characters ad event fifty times.

7R. By the time of his death L had done more than any previous president to promote the Jews' advance in the American society.

8L. Thomas Jefferson [President in 1801 – 1809] writing to John Adams in 1813 decried Jews' wretched depravity of sentiment and manners. At the same time [he contradicted himself]. [Adams: President 1797 – 1801, see S, G, 116.]

12R, 13L. Julius Hammerslough, a German Jew, and a prominent businessman. In his obituary, the *New York Times* called him a warm friend of L.

13R. By the time of his death, L had acquired far more Jewish friends and acquaintances than any American president before him.

14L. Abraham Jonas, a member of the Jewish faith: for more than two decades they were friends and political allies.

29L. End of letter (1856) from L to Jonas: Your friend as ever.

30R. In 1861 L endorsed Henry Rice, a German Jew, for the position of a military storekeeper.

In 1858 **L** met Henry Greenebaum, another German Jew, see below. [Almost all the Jews mentioned by the authors were from Germany.]

32L. Much later Greenebaum called **L** the greatest man he ever met.

38L. In 1858, Abr. Jonas signed an appeal to come to the debate prior to the election to the Senate between Republican **L** and Douglas, Democratic party; it was published in local newspapers:

Hear the true principles of the Republican party expounded and the unsound doctrines of the Douglas Democracy exposed.

Jonas was the Republican chairman of the committee for the debate.

40R. Jonas promoted **L**'s candidacy for the nation's highest office.

47L. It carries a photo and information about Carl Schurz, a brigadier-general during the Civil War, **L**'s friend and advocate for Jews.

50L. Abram J. Dittenhoefer, Jewish lawyer. In 1864, he was Republican elector for **L**'s another re-election campaign.

52L. 1860: **L**'s letter to Jonas: You are one of my most valued friends.

54L. Jonas played an important backstage [successful] role to promote **L** in a Republican convention to nominate a candidate for presidency.

Jonas and Ditterhoefer attended in nonvoting capacities but Lewis Naphtali Dembitz, a Jewish delegate, actively participated in favour of **L**. He was a significant Republican political leader in Kentucky and a leading lawyer. He influenced his nephew, a future Supreme Court justice.

56L. Moritz Pinner, from Germany, campaigned together with Dittenhoefer and Jonas.

57R. They were a key component of **L**'s coalition.

58R. 1855, in a letter, **L** wrote: In Russia, despotism can be taken pure and without the base alloy of hypocrisy.

62L. During the presidential campaign of 1860, New York had the largest Jewish community in the USA. According to some estimates, Jews voted against **L** by a two-to-one margin. No explanation provided.

62R. Among **L**'s electors were Dittenhoeffer, Jonas and Kaufmann.

64L. Sigismund Kaufmann advocated the vote of German Americans for **L**.

65L Kaufmann led New York German Jews in supporting **L** for his (first term) presidency

65R. Kaufmann, a fiery Jewish refugee, was an elector prominent among German Americans.

66L. In 1860, Jonas warned **L** against a threat of assassination.

66R. Rabbi Morris Raphal (New Yorck) actively opposed abolitionism by referring to the *Old Testament*. Fierce debates followed since there were many other Jewish apologists for slavery. See also 71R.

Rabbi Isaac Mayer Wise, a leader of American Jewish community, very strongly expressed himself against the Republican party.

71LR. Many abolitionists were anti-Semitic. The scene was mixed.

71R. Many Jews and even Jonas' older brother supported slavery. [Strong commercial advantages followed for Gentile and Jew from cheap cotton and many of them were tied by relationships with the South.] Morris Raphal was an apologist of slavery.

72L. There existed different interpretations of the *Bible*.

72L. Chicago Jewish leader Abraham Kohn from Germany was a staunch Republican and L's partisan: I regard Kohn the best authority for his countrymen in Chicago (statement of L's aide).

73R. Daughter of Kohn about her father. He saw L like Moses freeing the slaves.

77R. Isaac Mayer Wise called L a Dagon [Dragon].

78L. Wise was repulsed by L's manner. Greenebaum stated (when?) that many most prominent Jews in Illinois supported L for presidency.

78R. President L repeated an ancient Jewish oath.

81L. In 1861, in Baltimore, 7000 Jews had divided opinions about the incoming president. The most orthodox Jews were against him.

81R. In his inaugural address L referred to Christianity as the basis for resolving the impending crisis. Protesting Jewish letters followed.

84L. The precise meaning of Christianity was somewhat complex. Catholics, Jews and Mormons were considered religious outsiders.

84R. L did not realize the offence (see 81R), but over time he learned better.

Isaac Mayer Wise was happy about Jonas becoming post master of Quincy.

85R, 87L, 89R. L accepted the regiment of Max Einstein.

87L. Max Einstein honourably served in Civil War as unofficial colonel.

90L. In 1865 L's wife remarked that he had wished to visit Jerusalem.

91L. In civil war, if possible, L appointed people on the basis of merit. Alfred Mordecai, Jr became second lieutenant whereas his father was possibly the highest ranking Jew in army (minded weapons and ammunition). In the 1850s, the wife of the Secretary of War Jefferson Davis, was greatly impressed by Mordecai.

91R, 92LR. During the Civil War there was a disproportionate number of Jewish quartermasters (perhaps 50 including the

quartermaster general) since many of them had a mercantile background and one of them was Moritz Pinner. The same happened in the army of the Confederacy.

Raphal evinced patriotism (apparently supported Union) and his son was an officer in the Union army, lost an arm.

97R. A disproportionate number of sutlers (sellers of provisions etc. for army in the field) were Jews since many of them had an indispensable mercantile background, cf. pp. 91 and 92.

Once more, see 71R, Morris Raphal is called an apologist of slavery.

100R, 101L. Union general appointed Henry Rice, a German Jewish merchant, as sutler and **L** approved appointment.

103L. A cavalry regiment commanded by Max Friedman elected a Jew as the chaplain. He was dismissed as not being a regularly ordained minister. Instead, the regiment elected another Jew who was also dismissed as not being a Christian.

103L, 104R. Isaac Mayer Wise opposed the restriction of chaplaincy to Christians.

105L. Colonel Max Friedman founded a regiment.

106RL. On **L**'s initiative chaplaincy was opened for those authorized by an ecclesiastic body.

107LR. Jewish chaplains began appearing.

112L. Leonard Myers, a Jew and a congressman, corresponded with **L** in 1862 – 1865.

115L. Grant later tried to distance himself from that order as drafted by a subordinate and not studied because of the press of warfare. See p. xiiL.

118R. By revoking [that order], ensuring that the chaplaincy was opened up to Jews and appointing numerous Jews to public and military positions of trust **L** dramatically improved the status of Jews in the USA.

Even Morris Raphal (cf. 71R) was impressed by **L**'s revocation of Grant's order (xiiL).

120R. Antisemitism was also rampant in the South during the war.

123 LR. Army telegrapher, Jew Edward Rosewater wrote to future wife: he and other telegraphers enjoyed interaction with **L**.

124LR. Some Jews opposed the Emancipation act of 1862.

124R. Issachar Zacharie, a chiropodist, was the closest **L**'s Jewish friend after Jonas. In the sequel, very much is stated about him and his high-ranking patients.

137L. On the eve of the Civil War Zacharie was **L**'s spy in New Orleans. He had to entice the Louisiana population, which tended to side with the secessionists, back into the Union, but was unsuccessful.

138R. Henry Wentworth Monk, 1827 – 1896, was a self-proclaimed prophet. See next pages.

139R – 140L. Outwardly, Monk appeared like Jesus, and **L** ushered him in from the crowd of those wishing to see him. Monk suggested:

Why not follow the emancipation of the Negro by emancipation of the Jew? We are blind to what goes on in Russia, Prussia and Turkey.

There can be no permanent peace in the world until the civilized nations ... atone for what they have done to the Jews ... by restoring them to their native home in Palestine and making Jerusalem the capital city of a reunited Christendom.

**L**: this is a noble dream. [After the War] you will see what leadership America ... will show to the world. (A contemporary letter by Monk's biographer.) Regrettably, this is not a real reference.

140L. **L**: I myself have a regard for the Jews [when stated?].

140R, 141L. 1863: Zacharie met highest-ranking member of Confederate cabinet, the Secretary of State, the Jew Judah P. Benjamin.

142L. He reported back to **L** about plans to end the War.

142R. Zacharie complained: **L** hesitated, his cabinet disapproved.

144L. Zacharie succeeds in freeing Jewish captives in the North (**L** pardons them).

Goodman L. Mordecai loyally served in Confederacy army. Later, as a blockade runner, he was pardoned by **L**.

147L. **L**, Gettysburg Address [1863]: This nation under God [shall have a new birth of freedom].

152LR. **L** reprieved Jewish deserter who attempted to see his dying mother.

159R. William Mayer (an Austrian Jew) became brigadier-general.

161L. Colonel William Mayer raised a regiment from New York. In 1863, **L** praised him for putting down New York draft riots and promoted him to brigadier-general.

163R. **L** ordered release of two Jewish prisoners since proof of their guilt was insufficient. In general, he pardoned many irrespective of faith.

167L, 168R. **L** sympathetically regarded humanitarian requests of Leonard Myers, Jewish congressman.

170L. Many interactions between **L** and Jews might seem trivial. Aggregated, however, they form a pattern. **L** insisted on treating Jews on the same basis as everybody else. This attitude was ever present.

174L. Five of Jonas' sons served in the Confederacy, at least two in the army.

174R. Jonas was one of **L**'s most ardent and able strategist and political campaigner.

178R. In 1864, Zacharie vigorously campaigned for the President's re-election as also electors Dittenhoefer and Isidor Bush (both Jews). They helped to carry their states for **L**. New York remained Republican and Missouri turned.

179L. In addition. Kaufmann is named as such an elector.

180LR. In 1864, **L** met with certain gentlemen of the Hebrew faith to discuss the Jewish vote in New York, possibly the first time that any president had ever formally discussed a like subject in the Executive Mansion.

181R. Zacharie campaigned.

183L, 185L. Leopold Blumenberg (a German Jew), opposed slavery, organized a volunteer regiment and became a commanding officer. **L** appointed him head of military police in a district. His

brother Rudolph engaged in illegal slave trade, imprisoned and pardoned by L for informing about it.

185L. In an era when anti-Semitism was commonplace, L openly sided with Zacharie (and another Jew) against the advice of his Secretary of War.

185R. In 1874, Ulysses S. Grant as president testified to the skills of Zacharie.

186L. Leopold Blumenberg became colonel and, following L's involvement, President Andrew Johnson promoted him to brigadier-General [when?].

187R. Reception 1865: Three prominent Jews were presented and cordially greeted by the Lincolns.

194LR. L's second inaugural address (1865) contained great many references to the *Old Testament*.

194R. L refused to include Christianity in the Constitution.

Until p. 202 there is much about L's conciliatory philosophy. Neither soldiers nor leaders of the South were persecuted.

206R. L's assassin: a deranged young actor John Wilkes Booth, stated that L meant nigger citizenship. He was described as being of Jewish "extraction" or "descent" who traced his ancestors back to Spain. His father was an unstable alcoholic, his brother, a favourite of a Jewish community and a son of a Jew as he himself stated.

217L. Jews prayed for the repose of L's soul.

218R. Dembitz claimed that L was sometimes known as Rabbi Abraham. Isaac Mayer Wise: L was fully a Jew. He supposed himself to be a descendant of Hebrew parentage. He said so in my presence. No date

221R. Some Jews rejoiced over his death.

225R. Jews played a very significant role in shaping [the] portrayal of L.

226R. In 1909, the centennial of L's birth, Victor David Brenker, a Eastern Europe Jew, designed his image on a coin, a 1-cent piece.

**Above and below, in Chronology (abbreviated C),  
the following Jews were mentioned more than once**

Blumenberg, Leopold, 183L, 185L, 1861C, 1865C

Dembitz, Lewis Naphtali, 54L, 218R, 1860C

Dittenhoeffer, Abram J., 30L, 54L, 56L, 57R, 62R, 178R, 1865C

Einstein, Max, 85R, 87L, 89R

Friedmann, Max, 103L, 105L

Geenebaum, Henry, 30R, 32L, 78L, 1855C

Hammerslough, Julius, 12R, 13L

Jonas, Abraham, 14L, 29L, 38L, 40L, 52L, 54L, 56L, 57R, 62R, 66L, 174LR; 1856C, 1862C

Kaufmann, Sigismund, 62R, 64L, 65R, 179L, 1861C

Kohn, Abraham, 72L, 73R

Myers, Leonard, 167L, 168R

Pinner, Moritz, 56L, 57R, 91R, 92L, 97R

Raphal, Morris J., 66R, 71R, 92R, 118R

Rice, Henry, 30R, 84R, 100R, 101L, 103L, 104R

Wise, Isaac Mayer, 66R, 77R, 78L, 218R

Zacharie, Isachar, 124R, 137L,140R, 141L, 142R,144R,  
178R,181R, 185LR, 1862C, 1865C

### **Chronology, pp. 228 – 233**

**1855.** L meets Greenebaum for the first time and meets Hammerslough.

**1856.** L and Jonas electors in 1856 presidential campaign

**1860.** Dembitz attends 1860 Republican National convention to support L

Kaufmann: Republican elector for 1860 presidential election

**1861.** L probably met Kaufmann for the first time.

**1862.** Three testimonials of L for Zacharie. Jonas: his loyal and sensible friend.

**1864.** L interviewed a soldier who had complained of anti-Semitism against general Butler.

L frees prisoner Abraham B. Samuels held by Butler.

**1865.** L Endorses appointment of Dittenhoefer

L writes to Secretary of War *About Jews* (Zacharie and Blumenberg).

L releases Markbreit, a POW tortured by Confederacy

## Poisson's Work in Probability

*Arch. hist. ex. sci*, vol. 18, No. 3, 1978, pp. 245 – 300

### 1. Introduction

The main works of Poisson (1781 – 1840) are devoted to mechanics (celestial mechanics included) and mathematics. In particular, he achieved much in the fields of definite integrals, equations in finite differences, partial differential equations, mathematical physics and probability. His work is not yet sufficiently studied and this is entirely true with regard to probability.

There are at least two biographies of Poisson [28], [45]. To one of them [28] a catalogue of his works originally compiled by Poisson himself [27] is appended; regrettably, the description of many items in this bibliography is rather incomplete.

I only provide some parts of my text since it was too difficult to reproduce Poisson's complicated formulas. On the other hand, his contribution [22], which I quoted many times, is now available in an English translation (S, G, 53). It also contains many formulas lacking here.

Now, I copy the titles of the left-out sections of the *Contents* of my original contribution.

3. Limit theorems (De Moivre – Laplace theorems – binomial trials with variable probabilities – central limit theorem).

5. Mathematical statistics (parameters of distribution – significance of discrepancies – theory of errors).

7. Poisson's memoirs. Twelve memoirs are briefly described, one memoir [3] is described adequately and one more is actually Poisson's *Programme* which is copied separately.

### 2. General Problems of Probability

**2.1. Probability, Mathematics and Logic** The theory of probability [22, p. 1] *a pris route son extension* in the 18<sup>th</sup> century, becoming

*Une des principales branches des mathématiques, soit par le nombre et l'utilité de ses applications, soit par le genre d'analyse auquel il a donné naissance.*

*Aucune autre partie des mathématiques n'est susceptible d'applications plus nombreuses et plus immédiatement utiles*, says Poisson (p. 36) elsewhere.

It is extremely difficult to estimate the comparative importance of various branches of mathematics, but as far as analytical methods (*genre d'analyse*) are concerned, it is sufficient to refer to Laplace who made essential advances in mathematical analysis and paved the way for such scholars as Fourier and Cauchy (as well as for Poisson himself).

*Le calcul des probabilités*, maintains Poisson (pp. 35 – 36), *a pour objet de déterminer dans chaque question d'éventualité ou de doute, le rapport du nombre des cas favorables à l'arrivée d'un événement ... au nombre de tous les cas possibles; de sorte que nous puissions*

*connaître ... la raison que nous avons de croire que cette chose soit vraie, ou que cet événement a eu ou aura lieu, et que nous puissions aussi ... comparer cette raison de croire, dans deux questions de nature toute différente ... Ses principes doivent être regardés comme un supplément nécessaire de la logique.*

The first assertion is characteristic of the heroic Laplacian period of probability, while the connection of probability with logic originated with Leibniz (and Lambert<sup>2</sup>); after Poisson De Morgan [41] even attempted to found initial probability theorems on logical calculus. The further development of the logical branch of probability in the 19<sup>th</sup> century was connected with Boole, Jevons and Venn whereas pertinent contemporary work is possibly connected with the general problem of correlation between mathematics and logic. Note that Poisson unreservedly trusted the so-called classical definition of probability.

**2.2. Humanism of the Theory of Probability.** Humanism is a distinctive feature of both Laplace's [57, § 4.2] and Poisson's theory of probability. Studying the lottery of France, *heureusement supprimée par une loi récente*, he [22, p. 68] proves its disadvantageousness for the public, notes that gamblers are apt to follow one or another (senseless) "policy", holds that games of chance are *la cause de beaucoup de malheurs et peut-être de crimes* (pp. 70 – 71) and that in any case they have not *créé ... de valeurs* (p. 72). Attempting to amend legal proceedings, Poisson (§ 6) may have striven to better the moral state of the nation. At any rate he (p. 21) held that the empirical data of the French criminal statistics is an important document *sur l'état moral de notre pays*.

**2.3. Concept of Randomness.** Poisson [22, p. 80] paid attention to the philosophical aspect of randomness:

*L'ensemble des causes qui concourent à la production d'un événement sans influencer sur la grandeur de sa chance, c'est-à-dire, sur le rapport du nombre de cas favorables à son arrivée au nombre total des cas possibles, est ce qu'on doit entendre par le hasard<sup>3</sup>. Thus (Ibidem) une chose est faite au hasard lorsqu'elle est exécutée sans rien changer aux chances respectives des divers événements qui peuvent arriver.*

For example, *agitations nombreuses* which precede a throw of a die are random because they do not influence the chance of various possible outcomes. This definition is hardly satisfactory because chance pertains to a "random" event (outcome) about which Poisson says nothing at all. Suppose, he (p. 165) adds, that with prior probability  $p$  cause  $C$  brings about an event  $P$ ; suppose also that causes  $B_i$ ,  $i = 1, 2, \dots, n$ , of the second order *en se combinant avec le hasard* (a reference to his previous page follows) are capable of producing  $P$  with probabilities  $r_i$ , even in the absence of  $C$ <sup>4</sup>. Then the posterior probability for the existence of  $C$ , given that  $P$  appears in every one of the  $n$  experiments, is

$$w = \frac{p}{p + (1 - p)r_1 r_2 \dots r_n}.$$

This explanation is hardly better than the original definition. Cournot [37, Chapter 4] accepted neither the former nor the latter. Following Aristotle as well as modern scholars [56, §§ 2.2 and 9.1], he reduced chance to the intersections of chains of determinate events.

Poisson did not formulate his problem definitely enough and it is difficult to interpret the derived formula; for example, difficult to offer its Bayesian justification.

Poisson also studied the problem of distinguishing between random and determinate events which De Moivre [53, § 2.2.2] considered as the main aim of probability and which Laplace (Ibidem, § 2.4.2; [57], epigraph), being a natural scientist, presented as a problem of statistical significance of observations. Suppose [22, p. 39] an urn contains an equal number of white and black balls. A sample *qui présente quelque chose de symétrique*, e. g. the occurrence of  $m$  white balls drawn with replacement one after another, implies that some determinate cause was at work. The same is true (p. 114) if the 26 letters of the alphabet are arranged in their *ordre naturel*<sup>5</sup>.

Also (p. 115), let the number of ordinary events be  $n$  and that of remarkable events be  $m$  ( $m < n$ ). Then the probabilities of the hypotheses that “a remarkable event has a definite cause” and “a remarkable event is random”<sup>6</sup> are as  $(1/m):[1/(m + n)]$ . This is of course an elementary consideration. Besides that, it is extremely difficult to distinguish beforehand between ordinary and remarkable events [53], p 229; 56, pp 113 – 114 and 125].

Poisson (p. 118) claimed that harmony in nature could not be explained by randomness:

*Quant aux phénomènes physiques, dont les causes nous sont encore inconnues, il est raisonnable de les attribuer à des causes analogues à celles que nous connaissons, et soumises aux mêmes lois. Leur nombre diminue au reste de jour en jour, par le progrès des sciences.*

Such opinions were widely spread at the end of the 19<sup>th</sup> century, when physicists seemed to become able to explain almost every phenomenon in nature. I also note that Poisson did not repeat Laplace’s profound conclusion [57, § 2.3.2] about the possibility of order being produced out of randomness (e. g., about the sameness of the composition of urns caused by random interchanges of balls between them).

**2.4. Subjective and Objective Probabilities.** Poisson [22, pp 30 and 31; 24, p. 60] distinguishes between the probability of an event and its chance: the former, as he understands, is subjective<sup>7</sup>, and the latter is objective. Thus probability measured by the ratio of the number of favourable cases ( $m$ ) to the total number of cases ( $n$ ) (see also § 2.1) could change with experience while chance is constant. Poisson first expounded his point of view in 1836, in a letter to Cournot who later included it in his book [37, pp. vi – vii; the passage below is from my translation of Cournot, **S, G.** 54]:

*Sir, With great pleasure will I read the work on the Doctrine of chances<sup>8</sup> which you propose to publish. What I am now completing will not hinder you at all, and I am leaving enough space for a more comprehensive book. I discern the same difference between the words*

*chance and probability as you, and strongly insist on it. As to your approach to the main problem, the probability of judgements, I will compare it with my own after reviewing and definitively accomplishing that part of my work. I only have to finish that and to copy the entire text before becoming able to begin printing. There are some problems whose solutions I will include in one of the last chapters provided that I complete them at least to my own satisfaction.*

*Finally, you will find in that work some metaphysical considerations and see that I am not at all denying that branch of human knowledge.*

It seems however that Poisson's distinction between probability and chance (though not between subjective and objective probabilities) did not essentially matter.

2.4.1. *A Problem on Subjective Probabilities.* An urn contains  $n$  balls, white and black. Required is the probability of extracting a white ball in the very first drawing [22, p. 47].

There are  $(n+1)$  equally possible cases concerning the number of white balls in the urn. The probability sought is thus

$$\frac{1}{n+1} \left[ \frac{n}{n} + \frac{n-1}{n} + \dots + 0 \right] = \frac{1}{2},$$

an expected result *puisque nous n'avons aucune raison de croire à l'arrivée d'une boule blanche plutôt qu'à celle d'une boule noire.*<sup>9</sup>

One may well question the validity of Poisson's reasoning. But then, granted the existence of a certain probability of extracting a white ball, this probability should be equal to  $1/2$  to provide minimal information about the contents of the urn.

Laplace [57, § 2.2; 58, § 3.3] pronounced similar statements but considered them tentative, subject to essential change while Poisson himself (p. 35) equated the case of  $p = 1/2$  with *la parfait perplexité de notre esprit entre deux choses contraires*<sup>10</sup>.

**2.5. Concept of a Random Variable and the Use of a Distribution.** Poisson [22, pp. 140 – 141] was the first to introduce the concept of a random variable:

*Supposons actuellement qu'au lieu de deux événements possibles (in Bernoulli trials) il y en ait un nombre donné  $\lambda$ , dont un seul devra arriver à chaque épreuve. Ce cas est celui où l'on considère une chose  $A$  d'une nature quelconque, susceptible d'un nombre  $\lambda$  de valeurs, connues ou inconnues, que je représenterai par  $a_1, a_2, \dots, a_\lambda$  et parmi lesquelles une seule devra avoir lieu à chaque épreuve ...*<sup>10a</sup>.

Denote by  $c_{ij}$  the chance that cause  $C_i$ , if only it takes place, leads to  $A$ 's value  $a_j$ ; also, denote by  $\gamma_i$ , the probability of cause  $C_i$ . Then (Ibidem)

$$\sum_{\gamma=1}^{\lambda} c_{i\gamma} = 1, \quad \gamma=1,2,\dots,v,$$

$$\alpha_j = \gamma_1 c_{1j} + \gamma_2 c_{2j} + \dots + \gamma_v c_{vj}, \quad j = 1, 2, \dots, \lambda$$

would be the *chance moyenne* of  $a_i$ .

Exactly these  $\alpha_j$  are, in modern terminology, the probabilities of  $a_i$ . Starting from the obvious condition that the sum of these probabilities is unity, we arrive at a unit sum of  $\gamma_i$ , a restriction not mentioned by Poisson.

Elsewhere he (p. 254) introduces a random variable  $A$  as

*Une chose quelconque, susceptible de plusieurs valeurs positives ou négatives, ... que nous supposerons des multiples d'une quantité donnée  $\omega$ .*

He then assumes  $\omega \rightarrow \infty$ , thus using a then standard method for transforming discrete into continuous.

To substantiate such transformation directly, Poisson [14, p. 637; 22, p. 274] also introduced functions which are now named after Dirac. [I omit the proof inserted in my original text.]

Poisson (pp. 155, 158 and 160) offers examples of random variables. In the first instance the values of the random variable are the observed values of a measured angle; in the second case they represent the duration of life of newborn babies, while in the last example they are the discrepancies between calculated and observed heights of water above the lowest local sea level.

Poisson's terminology is imperfect. He does not write *random variable A*, or *law of distribution of A* though he does use such expressions as *la loi de probabilité des diverses valeurs possibles de A* (p. 155) and *la chance moyenne (de cette valeur de A)* or *la valeur moyenne (de A)* (pp. 141 and 271). In one place Poisson (p. 291) identified the observed angle with the *véritable valeur* of  $A$ .

Besides this Poisson does not use random variables throughout. Thus, he repeatedly referred to quantity  $A$  on pp. 271 – 291, but returns to *événement E* when considering Bernoulli trials (p. 294). Neither does he directly use the notion of random variable for substantiating his law of large numbers (LLN) in § 4.3. Nevertheless, it is possible that subsequent scholars, notably Chebyshev, started from Poisson's *chose A*.

He first introduced a discrete random variable in 1829 [6, p. 3], connecting it with the cumulative distribution function (§ 2.6) rather than with any causes as described above but calling the expectation of this variable its *vraie valeur* (p. 19)<sup>11</sup>.

**2.6. Cumulative Distribution Functions.** Poisson [6, § 1] was the first to define that function for a discrete random variable

$$F_n(x) = P(x_n < x)$$

(subscript  $n$  denotes observation  $n$ ). Moreover, Poisson (Ibidem) defined the density as the derivative of  $F_n(x)$  and (§§ 3.1.1 and 3.3.3) used this second definition even for continuous random variables. Thus, his starting point was the cumulative function rather than the density. Poisson [24, pp. 63 and 80] introduced the cumulative distribution function for continuous random variables as well.

After Poisson's time Davidov [40; 51], possibly following him, introduced the cumulative distribution function once more, while

Liapunov [49, p. 132 of the 1954 edition] remarked that it might be applied. However, its essential use did not begin until the 20<sup>th</sup> century.

#### 4. Law of Large Numbers

**4.1. Enoncé.** Poisson's first pronouncements on the LLN are contained in his memoirs [17], [18], [20], [25, pp. 459 – 463]. In his main work [22, p. 7] he gives the following definition:

*Les choses de toutes natures sont soumises à une loi universelle qu'on peut appeler la loi des grands nombres. Elle consiste en ce que, si l'on observe des nombres très considérables d'événements d'une même nature, dépendants de causes constantes et de causes qui varient irrégulièrement sans que leur variation soit progressive dans aucun sens déterminé, on trouvera, entre ces nombres, des rapports à très peu près constants. Pour chaque nature de choses, ces rapports auront une valeur spéciale dont il s'écarteront de moins en moins, à mesure que la série des événements observés augmentera davantage, et qu'ils atteindraient rigoureusement s'il était possible de prolonger cette série à l'infini.*

In the *Table des matières* (p. i) this diffuse definition is called an *énoncé* verified by *exemples nombreux et variés*. The LLN is mentioned in this *Table* twice more (p. iii, annotation of §§ 52 – 54, and p. vi, annotation of § 104). In the former place (§ 4.2) Poisson mentions its deduction *déjà vérifiée dans le préambule* while in the latter place (§ 4.3) he speaks about completing the *démonstration, à priori*, of this law, *regardée jusque-la comme un fait d'expérience*. Indeed, Poisson offers numerous examples of the action of his law. Thus, it serves as the foundation for marine insurance (p. 8) and explains the existence of various stable quantities: of the mean sea level (p. 9); of **un intervalle moyen des molécules** (emphasized by Poisson himself but remained unnoticed), of receipts from indirect taxes and lotteries (p. 11); of the ratio of the number of convicted to that of those accused (p. 11)<sup>12</sup>.

**4.2. Main Propositions.** Poisson [22, pp. 138 – 142, §§ 52 and 53] describes his law, setting forth three principles.

(1) Denote probabilities (chances) for the occurrence of disjoint events E and F in  $\mu$  trials by  $p_1, p_2, \dots, p_\mu$ , and  $q_1, q_2, \dots, q_\mu$ . If E occurred  $m$  times and F occurred  $n$  times ( $m + n = \mu$ ) then the *moyenne de toutes ces chances* of these events are approximately equal to the relative frequencies  $m/\mu$  and  $n/\mu$ , respectively, are almost the same in each series of trials.

(2) Suppose that mutually disjoint causes  $C_1, C_2, \dots, C_v$ , the probabilities of whose action are  $\gamma_1, \gamma_2, \dots, \gamma_v$  are capable of producing an event E and that the chance of its occurrence given the action of cause  $C_i$ , is  $c_i$ . Then

$$\gamma = \gamma_1 c_1 + \gamma_2 c_2 + \dots + \gamma_v c_v$$

is the chance for the occurrence of E, almost the same in each series of trials.

In Poisson's opinion exactly this principle<sup>13</sup> proves principle (1). Principle (2) means that if  $P(C_i) = \gamma_i$ , and  $P(E/C_i) = c_i$ , then  $P(E) = \gamma$ . Nevertheless, Poisson did not prove here that the value of  $\gamma$  is stable.

(3) Suppose now that in each trial a certain magnitude  $A$  with chance  $\alpha_j$  assumes one or another of its possible values  $a_j$ . Then (p. 143) the essence of the LLN will be expressed by two approximate relations with the subscripts denoting some two series of observations

$$m_1/\mu_1 = m_2/\mu_2, s_1/\mu_1 = s_2/\mu_2.$$

For the first relation Poisson provides examples

(a) P. 145<sup>14</sup>. Each of  $\nu$  urns contains white and black balls in various ratios. A ball is extracted from an urn chosen at random and put back. With  $\mu$  such extractions and  $\mu$  much exceeding  $\nu$  the relative number of white balls thus extracted is almost independent of  $\mu$ .

(b) P. 148<sup>15</sup>: a coin chosen at random *parmi celles qui proviennent d'une même mode de fabrication* is tossed and returned in the general pile. Suppose that after a large number  $\mu$  of such tosses heads occurred  $n$  times. The relative frequency  $n/\mu$  of the occurrence of heads will be almost independent of  $\mu$ .

(c) P. 154. Substitute families for coins (and male and female births for heads and tails). The relative frequency of yearly male births in a given nation will be almost constant.

For the second relation Poisson's examples (pp. 155 – 160) are just those which he used to illustrate the concept of a random variable in § 2.5: substitute  $s/\mu$  by the arithmetic mean of observations, by mean duration of life and mean sea level.

**4.3. Proof.** Applying his method for the transition from discrete to continuous random variables (§ 2.5), Poisson [22, p. 277, § 104] notes that a formula from his § 3 proves principles (2) and (3) from § 4.2.

**4.4. Recognition.** For several reasons the LLN was not recognized all at once. Indeed,

(1) Poisson introduced his law rather clumsily: his alluring *énoncé* (§ 4.1) and general considerations (§ 4.2) likely became more widely known than the proof itself (§ 4.3).

(2) Bienaymé severely criticized it. Thus, in a non-mathematical work published in 1855 and reprinted in 1876, he [31, p. 204] declared that this law just does not exist<sup>16</sup>. To a certain extent his opinion was possibly caused by his own generalization of Bernoulli trials [43, § 4] which he published in 1839 but which he considered *il y a longtemps* before that date. In any case, Bienaymé's criticisms illustrate the fruitlessness of simply negating classical works.

(3) The third and last (indirect) reason seems to consist in the underestimation of the LLN by such scholars as Cournot, Quetelet and Lexis<sup>17</sup>.

Probably because of Bienaymé's criticisms Cournot [37] did not mention this law at all. Quetelet [52, pp. 313 – 315] correctly illustrated it by one of Poisson's own examples (example (b) from § 4.2) though in a somewhat restricted sense [32, p. 658]; still, he failed to use it in his work. His concept of the *homme moyen* and of a single index for the characteristic of criminal inclination of each man would have been unnecessary had he based himself on Poisson's law<sup>18</sup>.

Lexis, for his part, does not seem to have mentioned the LLN at all though he [48, p. 96] used Poisson's urn model (see example (a) from § 4.2) for describing mortality laws.

Criticisms of the LLN continued at least until 1888 [29, pp. xxxii and 94]. But it actually received recognition much earlier. In 1838, several times referring to Poisson, Bessel [60, esp. § 9] specifically mentioned a *Principe der grossen Zahlen*; in 1846 Buniakovsky [33, p. 35, footnote] mentioned this law in passing; in 1854 and again in 1857, Davidov [38, pp. 30, 36, 62; 39, p. 11] considered it to be a rather general law while in 1873 Laurent [46, p. 97] cautiously remarked:

*L'expression rigoureuse de cette loi n'est pas connue. Les effets de Poisson ne sont cependant pas restés infractueux*<sup>19</sup>.

The LLN was eventually recognized in the 20<sup>th</sup> century.

Remarking that Poisson was no statistician [59, p. 26], a qualification which seems too strong, Tschuprow (pp. 223 – 238) connected Poisson's researches with the Lexian problem of stability of statistical series thus emphasizing Poisson's role in the origination of the so-called Continental direction of statistics.

Exactly the discovery of diversity in "stabilities" and a mathematical substantiation of the methods for theoretical studies of stability constitute Poisson's immortal contribution to statistics, says Tschuprow (p. 235).

I ought to add that the LLN was accepted by Chebyshev [35, p. 259]:

*Cette proposition fondamentale de la théorie des probabilités, contenant comme cas particulier la loi de J. Bernoulli, est déduite par Mr. Poisson d'une formule qu'il obtient en calculant approximativement la valeur d'une intégrale définie, assez compliquée.*

Indeed (§ 4.3), Poisson verified his LLN by means of the central limit theorem which neither he, nor Laplace before him [58, § 4.2] proved rigorously<sup>20</sup>. on the other hand, Chebyshev succeeded in managing without it.

Later he [36] proved the LLN numbers, and, for that matter, a generalized version of same. But unfortunately these works by Chebyshev remained unknown in Europe; in any case they were not mentioned either by Bertrand or Poincaré.

**4.5. Modern Notion.** A sequence of random variables

$$\xi_1, \xi_2, \dots, \xi_n, \dots$$

is said to obey the LLN if there are sequences of numbers  $a_1, a_2, \dots, a_n, \dots$  and positive  $B_1, B_2, \dots, B_n, \dots$  such that for any  $\varepsilon > 0$

$$\lim P\left(\frac{1}{B_n} \left| \sum_{k=1}^n \xi_k - a_n \right| < \varepsilon\right) = 1, n \rightarrow \infty.$$

In particular, let those variables assume values 0 and 1 with probabilities  $\bar{p}, \bar{q}, \bar{r}, \dots$  and  $p, q, r, \dots$  respectively, thus illustrating

the occurrence of some event in the scheme of Poisson trials. In this case the relative frequency of the occurrence of this event will be equal to the mean value of  $\xi_k$ , or  $m/n$ . And, as proved by Poisson, with  $n \rightarrow \infty$  the absolute value of the difference between  $m/n$  and  $(p + q + r + \dots)/n$  will tend to become less than any  $\varepsilon$ .

#### 5.4. Applications of Mathematical Statistics in Medicine.

Principles governing the application of probability (more precisely: of mathematical statistics) in medicine and, in particular, the search of optimal methods of medical treatment were considered in the review [16]. The reviewers were of the opinion that medical statistics should be improved and their views (pp. 173 and 174) were quite justified:

(1) *En matière de statistique, c'est-à-dire dans les divers essais d'appréciation numérique des faits, le premier soin avant tout c'est de perdre de vue l'homme pris isolément pour ne le considérer que comme une fraction de l'espèce. ...*

*En médecine appliquée au contraire, le problème est toujours individuel ...*

(2) *La statistique mise en pratique, qui est toujours en définitive le mécanisme fonctionnant du calcul des probabilités, appelle nécessairement des masses infinies, un nombre illimité de faits non-seulement en vue d'approcher le plus près possible de la vérité, mais aussi afin d'arriver à faire disparaître, à éliminer, autant qu'il est possible, et à l'aide de procédés connus, les nombreuses sources d'erreurs si difficiles à éviter.*

It seems however, that neither the reviewers, nor Laplace [57, § 2.4] before them foresaw the great difficulties of mathematising medicine [16, p. 176]:

*.. La condition des sciences médicales, à cet égard (of lending themselves to mathematisation), n'est pas pire, n'est pas autre que la condition de tout les sciences physiques et naturelles, de la jurisprudence, des sciences morales et politiques, etc.*

Other reflections, interesting for the history of medicine, are offered, as e. g. (p. 169)

*.. Les nombreux établissemens créés pour le traitement des déviations de la taille ont révélé un grand nombre de maladies de ce genre qui seraient restées inaperçues sans cela.*

A few lines devoted to the same subject are contained in a footnote to the *Table des matières* (!) of Poisson's *Recherches* (p. vi):

*La médecine ne serait ni une science ni un art, si elle n'était pas fondée sur de nombreuses observations, et sur le tact et l'expérience propres du médecin, qui lui font juger de la similitude des cas et apprécier les circonstances exceptionnelles.*

#### 6. Application of Probability to Jurisprudence

Poisson's main contribution [22] is to a large extent devoted to stochastic investigations in jurisprudence, a fact reflected in the title of the book and emphasized on its opening pages (pp. 1 – 2):

*Parmi les applications de ce calcul [des probabilités], une des plus importantes est celle qui se rapporte à la probabilité des jugements, ou, en général, des décisions rendues à la pluralité des voix<sup>21</sup>.*

On p. 5 Poisson notes that

*Selon Condorcet, la chance d'être condamné injustement pourrait être équivalente à celle d'un danger que nous jugeons assez petite pour ne pas même chercher à nous y soustraire dans les habitudes de la vie.*

Poisson also reasonably argues that the probability of a convict's guilt should essentially surpass the probability of his innocence. He adds (p. 6) that, according to Laplace [44, p. 521], there should be

*Plus de danger pour la sûreté publique à l'acquittement d'un coupable, que de crainte de la condamnation d'un innocent.*

Poisson himself (pp. 388 – 389) also discusses these dangers (probabilities) and at least compares them but does not offer any appropriate criterion for decision making<sup>22</sup>. His main objective (p. 17) was to test the stability of the coefficient of conviction (§ 4.1)<sup>23</sup> and study the probability of legal errors with a view of ensuring the intercomparison of different legal procedures or criminal statistics of different nations.

Poisson's research occupies almost a hundred pages [22, chap. 5]. He begins by considering sentences passed by a single judge. Denote (p. 318) the probabilities of a correct sentence by  $u$ , of the defendant's guilt by  $k$ , of his conviction by  $\gamma$  and let  $p$  and  $q$  be the probabilities of the defendant's guilt given his conviction and of his innocence given his acquittal correspondingly. Then

$$\begin{aligned}\gamma &= ku + (1 - k)(1 - u) = 1/2 + 1/2(2k - 1)(2u - 1), \\ p &= ku/\gamma, \quad q = [(1 - k)/(1 - \gamma)]u, \quad u = p\gamma + q(1 - \gamma).\end{aligned}$$

These formulas are quite simple. But Poisson supposes that, as a rule,  $k > 1/2$  (so that if one assumes  $u > 1/2$ , which is a natural restriction,  $\gamma$  will also to be larger than  $1/2$ ). At first he (p. 4) accepts this condition for the *cour d'assises*, then assumes it even for courts of first instance. This condition can be granted only if  $k$  is regarded as a generalized mean index, but even so Poisson should have added that in each separate case courts ought to begin their proceedings with a presumption of the defendant's innocence.

Poisson (p. 4) even maintains that condition  $k = 0$  assumed by Laplace contradicts the rules,

*Qui sont la base de la théorie dont nous nous occupons* (he means the Bayesian approach), *exigent que l'on ait égard à toute présomption antérieure à l'observation lorsque l'on ne suppose pas, ou qu'on n'a pas démontré qu'il n'en existe aucune.*

For cases heard by the jury the probability of a conviction by  $(n - i)$  jurors out of the total number of them,  $n$ , is provided (p. 332) for the probability of a correct decision, common for all jurors. Then Poisson introduces  $t$  by the formula  $u = t/(1 + t)$ , assumes that  $n = m + 2i$ , so that  $m$  is the absolute majority of votes, and gets the probability of conviction

$$P_i = \frac{kt^m}{kt^m + (1 - k)}.$$

It depends on  $m$  but not on the total number  $n$  of jurors. This remark is due to Poisson. He does not add that the total probability of the defendant's guilt composed from separate  $p_i$ 's does depend on  $n$ . I also note that the case  $m < 0$  leads to a natural conclusion:  $p_i < k$ .

I do not describe Poisson's further deliberations sine they are hardly useful. Instead, I repeat an unusual idea from his *Introduction*: replace *guilt* and *innocence* (which are hardly ever known for sure) by *subject to be convicted*/*to be acquitted*.

I have put on record criticisms about applications of probability in jurisprudence by Poincaré, Poincaré and, possibly, Cauchy [54, p. 296]. I [57, § 2.10] also remarked that these applications have likely promoted the development of criminal statistics. I shall add now that Laplace [44, p. 523] definitely pointed out that his conclusions were based on the independence of judges (jurors) from one another. However, this qualification should be strengthened: formulas which do not allow for the interdependence of the opinions of judges (jurors) are hardly acceptable<sup>24</sup>.

### Appendix: Poisson's Programme [21]

Following my request, Professor F. Rosenfeld has found Poisson's *Programme* on p. 26 of *Programmes de l'enseignement de l'Ecole Polytechnique ... pour l'année scolaire 1836 – 1837*. Paris, Imprimerie royale, 1837. This *Programme* seems worthy of being quoted in full:

#### **Éléments du calcul des probabilités, et arithmétique sociale.**

Principes généraux du calcul des chances; probabilité simple, composée, partielle, totale. – Des épreuves répétées, théorème de Bernoulli. – Probabilité des événements à venir, déduite de l'observation d'événements antérieurs de même nature.

Espérance mathématique. – Application à divers cas, et particulièrement aux loteries.

Des tables de population et de mortalité. De la durée de la vie moyenne dans diverses contrées. Partage de la population suivant les âges et les sexes. De l'influence de la petite vérole, de l'inoculation et de la vaccine sur la population, et la durée de la vie moyenne.

Des bénéfices et des charges des établissements qui dépendent de la probabilité des événements. Des rentes viagères, des tontines, des caisses d'épargne, des assurances, des annuités, des fonds d'amortissement, des emprunts.

Des moyennes à prendre entre plusieurs résultats.

I conclude with a few remarks. First, in his lectures Poisson emphasized neither the LLN nor applications of probability to jurisprudence and mathematical treatment of observations. Second, the names of De Moivre, Bayes, Daniel Bernoulli or even Laplace are not mentioned. Third, in keeping with real life as also with the traditions of political arithmetic Poisson seems mainly to have included problems connected with mortality rather than with the ratio of male and female births or other problems of demographic statistics. Fourth and last, the term *arithmétique sociale*, possibly coined by Poisson himself, did not come into general usage perhaps because

Quetelet, in a broader context preferred to use another term, *physique sociale*.

*Acknowledgement.* R. Jaquel, D. Mackenzie, A. Moreau, F. Rosenfeld and E. Seneta have sent me copies of papers and/or reprints of their own contributions while I. Schneider provided a copy of the whole of Poisson's *Recherches* [22].

My subject occupies a special section in a relatively short account of the history of probability written by B. V. Gnedenko and myself for a later (in 1978) published monograph on mathematics in the 19<sup>th</sup> century.

### Notes

1. Referring to this work I usually indicate pages only.
2. Rather than mention any particular writing due to Lambert I refer here to the general spirit of his philosophical works.
3. See § 2.4 for Poisson's understanding of chance versus probability.
4. More precisely  $r_i$  are the corresponding chances for the occurrence of P. On p. 168 Poisson calls causes B accidental adding on p. 169 that *Ces causes variables et irrégulières, que l'on ne doit pourtant pas confondre avec le hasard, peuvent influencer sur la chance moyenne de arrivée de P.*
5. Similar examples are due to many scholars [57, § 3.3].
6. But what is a random event? The same old question once more!
7. Explaining his definition of (subjective) probability, Poisson (p. 33) notes in passing that with infinite  $m$  and  $n$  (e. g., if they are areas of certain figures) it could become irrational. Just as Laplace [57, § 2.4] he pays here no attention to the inadequacy of the classical definition of probability.
8. The complete lack of references to the LLN in Cournot's book (see also § 4.3) would hardly have pleased Poisson had he lived to see it.
9. I do not describe a second problem of the same kind (Ibidem).
10. Acting as a co-reviewer of a book on medical statistics and referring to Laplace, Poisson [16, pp. 176 – 177] remarks: *On le voit clairement, l'induction, l'analogie, des hypothèses fondées sur les faits et vérifiées, rectifiées sans cesse par de nouvelles observations, ... tels sont les principaux moyens de parvenir à la vérité.*
- 10a. The actual use of random variables dates back to the 17<sup>th</sup> century. Thus, if a gambler is entitled to one of three possible gains  $a$ ,  $b$  and  $c$  whose chances are  $p$ ,  $q$  and  $r$  respectively, his expected gain, calculated by Jacob Bernoulli [29], pt. 1, p. 9] and, in a particular case of  $p = q = r$ , by Huygens, is the expectation of a corresponding random variable, a notion which was not used in those times.  
From the mid-18<sup>th</sup> century onward random variables were effectively introduced into the theory of errors by Simpson and Lagrange, then by Laplace and Gauss. However, Poisson, in introducing a random variable formally and naming it, if only by a provisional term, made a major advance in the development of probability theory.
11. Poisson's terminology and notation underwent a certain evolution: in 1830 he [9, pp. 141 and 146] used the same letter A to denote an observed constant, also calling it *une chose quelconque* (or simply *une chose*).
12. The ratio or coefficient of conviction, as I shall call it. Regarding the stability of receipts from indirect taxes etc. see also the relevant opinion of Laplace [57, § 3.1]. Reasoning similar to that which leads to the concept of the mean interval between molecules, adds Poisson (p. 10); [57, § 3.2.2] referring to his *autres* works, is the base of the *calcul des forces moléculaires et du rayonnement calorifique dans l'intérieur des corps*.  
Poisson's works [7, pp. 369 and 370]; [12, pp. 270 – 272]; [11, pp. 4 and 5]; [14, p. 176]; [15, pp. 14, 65, 530]; [26, p. 7] contain direct or indirect pronouncements on molecular conditions of substance, local parameters of molecular interactions, etc., and in one instance [26, p. 12] he refers in this connection to the LLN. Poisson [10, p. 6] even uses the concept of the mean interval between molecules to introduce most important physical definitions: **1.** *Les molécules sont distribuées régulièrement, et, en général, inégalement resserrées en différents sens autour de chaque point; c'est le cas des corps cristallisés.* **2.** *Elles sont irrégulièrement distribuées; mais leur*

*intervalle moyen reste toujours égal en tous sens autour d'un même point, quelles que soient les pressions extérieures; ce cas est celui des fluides parfaits. 3. La même disposition a lieu dans les corps solides élastiques et non-cristallisés qui ne sont soumis à aucune force donnée; mais leurs molécules se resserent ou s'écartent inégalement en différents sens autour de chaque point, lorsque des forces de directions données agissent sur ces corps.*

**13.** On pp. 44 – 46 principle (2) is formulated for the particular case of  $\gamma_i = 1/v$  with no reference either to the LLN or to future series of trials.

**14.** This example first appeared in an earlier work [8, p. 280].

**15.** See also the previous memoir [18].

**16.** His first critical statement (in 1842) remained unpublished [31, p. 199].

**17.** Implied here is the fact that after Poisson, at least in Western Europe, probability was forgotten for many decades; see also my earlier publication [57, § 4.2]. This is what Mansion [50, p. 3] testified to:

*Il est peu de pays, croyons-nous, où le calcul des probabilités tiennne une place aussi considérable dans l'enseignement supérieur qu'en Belgique. ... En France ... il n'est enseigné qu'accidentellement, comme accessoire du cours de physique mathématique à la Sorbonne: à l'Ecole Polytechnique, on ne lui consacre que quelques leçons des cours d'analyse et d'astronomie. En Allemagne ... la théorie de la compensation des erreurs d'observation fait souvent l'objet d'une Vorlesung spéciale, mais rarement le calcul des probabilités y est exposé dans toute son étendue.*

Mansion goes on to attribute the high respect for probability in Belgium to the lasting influence of Quetelet, but Russia just did not exist!

**18.** Cf. Tschuprow [59], p. 227]:

*Poisson's generalized scheme irrevocably finishes off with the levelling tendencies of the simplified theory of statistical regularity advocated by Quetelet's disciples.*

**19.** Many years later Laurent [47, p. 22] put on record his adverse opinion of the criticisms due to Bienaymé: [the scheme of Bernoulli trials]

*a été généralisé par Poisson sous le nom de Loi des grands nombres. Bienaymé à écrit à plusieurs reprises contre la loi des grands nombres, mais ses critiques ... ne portent que sur un malentendu.*

**20.** For the relevant opinion of Chebyshev see my earlier publication [58, § 4. 2]. In 1853 this theorem was rigorously proved by Cauchy [42, p. 142], as at least stated Freudenthal.

**21.** Poisson discussed only one problem pertaining to voting procedures. As regards jurisprudence he was almost exclusively concerned with criminal trials.

**22.** Similar reasoning is due to Aristotle [56, p. 108]. It goes without saying that for him probability was a measure of subjective opinion only. However, excepting the case of statistical probability, the same seems likely regarding Laplace's and Poisson's stochastic considerations in jurisprudence. Considerations resembling those of Aristotle found their way into official documents, at least in regard to capital punishment. Thus according to the Kriegs-Reglement of Peter The Great (1716) published in Russian and German in vol. 5 of the *Complete code of laws of the Russian Empire from 1649* (Petersburg, 1830, 203 – 453; see p. 403):

*Viel besser ist, zehen schuldige zu befreien, als einen unschuldigen zum Tode zu condemniren.*

**23.** Poisson (pp. 16 and 377) remarked that this coefficient is comparatively stable for a given kind of crime and sex of the accused. Drawing on British criminal statistics for 1811 – 1832 Poisson (p. 23) noted that the rise in crime in Britain was followed by an increase in the coefficient of conviction, i. e. by a tightening, up of legal proceedings. A similar process took place in France (pp. 375 and 376): the relative number of those accused in the département de la Seine was four times as high as the national average and, correspondingly, the coefficient of conviction also came out to be somewhat higher.

Poisson (p. 376), remarked that the *répression des crimes* in that département was *plus nécessaire*, perhaps causing a *plus grande sévérité des jurés*.

This point of view seems to be at variance with Poisson's own statement (§ 2.2) to the effect that criminal statistics represents the moral state of the nation. Would it not be more correct to assume that the control of crime is primarily achieved by general social measures?

**24.** The general literature abounds with examples of such interdependence or of

causes affecting all jurors in the same way. Dickens (*Pickwick Club*, chap. 34) illustrated this problem with a touch of humour; Tolstoy (*Resurrection*, chap. 23 of pt. 1) narrated it with an outward calmness; while France (*Les dieux ont soif*, esp. chap. 16) used tragic irony to describe cases indeed horrible.

## References

### S.-D. Poisson

1. Observations relatives au nombre de naissances des deux sexes. *Annuaire le bureau des longitudes* pour 1825 (1824), 98 – 99.
2. Sur la probabilité des résultats moyens des observations. *Conn. des temps*, 1827 (1824), 273 – 302.
3. Sur l'avantage du banquier au jeu de trente-et-quarante. *Ann. math. pures et appl.*, t. 16, 1825 – 1826, 173 – 208.
4. Rapport sur les tontines. Prepared by Poisson, Lacroix, Fourier-rapporteur. (Fourier J. B. J., *Oeuvres*, t. 2. Paris, 1890, 617 – 633.)
5. Discours prononcé aux obsèques de ... Laplace. *Conn. des temps*, 1830 (1827), 19 – 22.
6. Suite du [2]. *Conn. des temps*, 1832 (1829), 3 – 22.
7. Sur l'équilibre et le mouvement des corps élastiques. *Mém. Acad. sci. Paris*, t. 8, 1829, 357 – 570.
8. Mémoire sur la proportion des naissances des filles et des garçons. *Ibidem*, t. 9, 1830, 239 – 308.
9. Formules des probabilités, relatives au résultat moyen des observations, qui peuvent être utiles dans l'artillerie. *Mémorial de l'artillerie*, No. 3, 1830, 141 – 156.
10. Sur l'équilibre des fluides. *Mém. Acad. sci. Paris*, t. 9, 1830, 1 – 88.
11. Sur les équations générales de l'équilibre et du mouvement des corps solides élastiques et des fluides. *J. Ecole polyt.*, t. 13, No. 20, 1831, 1 – 174.
12. *Nouvelle théorie de l'action capillaire*. Paris. 1831.
13. Discours prononcé aux funérailles de M. Legendre. *J. reine und angew. Math.*, Bd. 10, 1833, 360 – 363.
14. *Traité de mécanique*, t. 1. Paris. 1833 (seconde éd.).
15. *Théorie mathématique de la chaleur*. Paris, 1835.
16. Review of Civiale, *Recherches de statistique sur l'affection calculieuse*. *C. r. Acad. sci. Paris*, t. 1, 1835, 167 – 177. Co-reviewers: Dulong, Larrey, Double.
17. Recherches sur la probabilité des jugements, principalement en matière criminelle. *Ibidem*, 473 – 494.
18. Note sur la loi des grands nombres. *Ibidem*, t. 2, 1836, 377 – 382.
19. Note sur le calcul des probabilités. *Ibidem*, 395 – 400.
20. Formules relatives aux probabilités qui dépendent de très grands nombres. *Ibidem*, 603 – 613.
21. Programme du cours de calcul des probabilités à la Faculté des sciences (of the Ecole polyt.) pour 1836/1837. See above.
22. *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*. Paris, 1837, 2003. **S, G**, 53.
23. Solution d'un problème de probabilité. *J. math pures et appl.*, t. 2, 1837, 373 – 387.
24. Mémoire sur la probabilité du tir a la cible. *Mémorial de l'artillerie*, No. 4, 1837, 59 – 94.
25. Note sur la proportion des condamnations prononcés par les jurys. *C. r. Acad. Sci. Paris*, t. 5 1837, 355 – 357, 459 – 463.
26. Sur l'équilibre et le mouvement des corps cristallisés. *Mém. Acad. Sci. Paris*, t. 18, 1842, 3 – 152.
27. *Catalogue des ouvrages et mémoires scientifiques*. Paris, 1851.

### Other Authors

AHES = *Arch. Hist. Ex. Sci.*

- 28. Arago F.** (1850), Poisson. *Oeuvr. Compl.*, t. 2, 1854, pp. 593 – 671. Catalogue [27] appended on pp. 672 – 689. Author's Discours prononcé aux funérailles de Poisson appended on pp. 690 – 698.
- 29. Bernoulli Jacob** (1713), *Ars conjectandi*. Reprint: Bernoulli J. (1975, pp. 107 – 259). German transl.: *Wahrscheinlichkeitsrechnung*. Leipzig, 1899. Its reprint: Frankfurt/Main, 1999.
- 30. Bertrand J.** (1888), *Calcul des probabilités*. 2<sup>nd</sup> ed., 1907. Reprints: New York, 1970, 1972. Second ed. practically coincides with the first one.
- 31. Bienaymé I. J.** (1855), Sur un principe que M. Poisson avait cru découvrir et qu'il avait appelé Loi des grands nombres. *C. r. Acad. Sci. Morales et Politiques*, sér. 3, t. 11, pp. 379 – 389. Also: *J. Soc. Stat. Paris*, 1876, pp. 199 – 204.
- 32. Bortkiewicz L. von** (1894 – 1896), Kritische Betrachtungen zur theoretischen Statistik. *Jahrbüchern f. Nationalökonomie u. Statistik*, 3. Folge, Bde 8, 10, 11, pp. 641 – 680, 321 – 360, 701 – 705.
- 33. Buniakovsky V. Ya.** (1846), *Osnovania Matematicheskoi Teorii Veroiatnostei* (Principles of the Math. Theory of Probability). Petersburg.
- 34. Chauvenet, W.**, *Manual of spherical and practical astronomy*, vol. 2. Philadelphia, 1863.
- 35. Chebyshev P. L.** Démonstration élémentaire d'une proposition générale de la théorie des probabilités. *J. reine angew. Math.*, Bd. 33, 1846, 259 – 267.
- 36. Chebyshev P. L.**, Des valeurs moyennes. *J. math. pures et. appl.*, t. 12, 1867, pp. 177 – 184.
- 37. Cournot A. A.** (1843), *Exposition de la théorie des chances et des probabilités*. Paris, 1984. Editor B. Bru. **S, G**, 54.
- 38. Davidov A. Yu.** [1854], *Lektsii Matematicheskoi Teorii Veroiatnostei* (Lectures on Math. Theory of Probability). No place.
- 39. Davidov A. Yu.** (1857), Theory of mean quantities with its application to compilation of mortality tables. In *Rechii i Otchet, Proiznesennyye v Torzhestvennom Sobranii Moskovskogo Universiteta* (Orations and Report at the Grand Meeting of Moscow Univ.). Moscow, first paging.
- 40. Davidov A. Yu.** [1885], *Teoria Veroiatnostei, 1884 – 1885*.
- 41. De Morgan A.** (1847), *Formal Logic, or the Calculus of Inference, Necessary and Probable*. Second ed.: Chicago, 1926.
- 42. Freudenthal H.** (1971), Cauchy. *Dict. Scient. Biogr.*, vol. 3, pp. 131 – 148.
- 43. Heyde C. C., Seneta E.** Bienaymé. Proc. 40<sup>th</sup> session Intern. Stat. Inst. 1975. *Bull. Intern. Stat. Inst.*, t. 46, No. 2, 1977, 318 – 331.
- 44. Laplace P. S.** (1812), *Théorie analytique des probabilités*. *Oeuvr. Compl.*, t. 7, No. 1 – 2. Paris, 1886. Consists of two parts, an Introduction (1814, *Essai philosophique ...*, English translation: New York 1995) and supplements, see below. Theory of probability proper is treated in pt. 2. Sections 21 – 24 of Chapter 4: **S, G**, 58.
- 45. Lapparent, A. de.** Poisson. *Livre du centenaire de l'Ecole Polytechnique*, t. 1. Paris, 1895, 97 – 103.
- 46. Laurent H.** (1873), *Traité du calcul des probabilités*. Paris.
- 47. Laurent H.** *Statistique mathématique*. Paris, 1908.
- 48. Lexis W.** *Einleitung in die Theorie der Bevölkerungs-statistik*. Strassburg, 1875.
- 49. Liapunov A. M.** (1900), Sur une proposition de la théorie des probabilités. *Izvestia Imp. Acad. Sci. St. Pétersb.*, sér. 5, t. 13, pp. 359 – 386.
- 50. Mansion P.** (1904), Sur la portée objective du calcul des probabilités. *Mathesis*, sér. 3, t. 4. Suppl. 2. Separate paging.
- 51. Ondar Kh. O.** (1971, Russian), On the work of A. Yu. Davidov in the theory of probability and on his methodological views. *Istoriya i Metodologiya Estestven. Nauk*, vol. 11, pp. 98 – 109. **S, G**, 5.
- 52. Quetelet A.** (1846), *Lettres ... sur la théorie des probabilités*. Bruxelles.
- 53. Sheynin O. B.** (1971a), Newton and the classical theory of probability. AHES, vol. 7, pp. 217 – 243.
- 54. Sheynin O. B.** (1973a), Finite random sums. Historical essay. AHES, vol. 9, pp. 275 – 305.
- 55. Sheynin O. B.** (1974), On the prehistory of the theory of probability. AHES, vol. 12, pp. 97 – 141.

- 56. Sheynin O. B.** (1976), Laplace's work on probability. AHES, vol. 16, pp. 137 – 187.
- 57. Sheynin O. B.** (1977a), Laplace's theory of errors. AHES, vol. 17, pp. 1 – 61.
- 58. Tschuprow (Chuprov), A. A.,** *Ocherki po teorii statistiki* (Essays on the Theory of Statistics), 1909, 1910. Moscow, 1959.
- 60. Bessel F. W.** Untersuchungen über die Wahrscheinlichkeit der Beobachtungsfehler (1838). *Abhandlungen*, Bd. 2. Leipzig, 1876, 372 – 391.
- 61. Bienaymé, I. J.** [Calcul des probabilités]. *Procès-verb. Soc. Philom. Paris*, 1840, 23 – 26.
- 62. May, K.,** Probabilities of certain election results. *Amer. Math. Monthly*, vol. 55, No. 2, 1948, 203 – 209.

#### **Additional sources**

- Bru B.** (1981), Poisson, le calcul des probabilités et l'instruction publique. In Métivier et al (1981, pp. 51 – 94).
- (2013), Poisson et le calcul des probabilités. In Poisson (2013, pp. 333 – 355).
- Métivier M. et al** (1981), *Poisson et la science de son temps*. Paris.
- Poisson S.-D.** (1812), Review of Laplace (1812). *Nouv. Bull. des Sciences. Soc. Philomatique de Paris*, t. 3, pp. 160 – 163. **S, G**, 58.
- Poisson S.-D.** (2013), *Les mathématique au service de la science*. Paris. Editor, Yvette Kosmann – Shwarzbach.
- Sheynin O.** (2013), Poisson et la statistique. In Poisson (2013, pp. 357 – 366).

## VII

### R. J. Boscovich's Work on Probability

*Arch. hist. ex. sci.*, vol. 9, No. 4 – 5, 1973, pp. 306 – 324

#### Abstract

A description of Boscovich's (1711 – 1787) method of adjusting arc measurements is presented and its further history, including the incorporation of his method in the theory of linear programming, is traced in § 1. A special feature here is a description of a connection, discovered by C. G. J. Jacobi, between the method of least squares (MLSq) and one of the previous methods of adjusting observations.

Two manuscripts of Boscovich are considered in § 2. In one, he studied the stochastic behaviour of the sum of several random variables each having a particular discrete uniform distribution. Set in a context of the theory of errors, this undated manuscript, if written before 1756, is the first work where probability is applied to the theory of errors.

The second manuscript describes the practical side of the *lotto di Roma*. In § 3, Boscovich's reasoning on randomness in a physical context is noticed, and, as a sideline, a quotation from Maupertuis shows that, as regards determinism, he (as also Boscovich) is the precursor of Laplace.

For Boscovich's biography see [16], [25], [34], [35b] and Hald (1998, p. 97). On pp. 97 – 103 Hald provided a concise and meaningful modern description of his work.

#### 1. Mathematical Treatment of Observations

In 1750 – 1753, Boscovich, one of the last polymaths and, in particular, an astronomer, and another astronomer, C. Maire (1697 – 1767), conducted an arc measurement in Italy. After that, taking account of their own and other arc measurements, Boscovich deduced the parameters of the earth's spheroid.

From a formal point of view such deductions constitute an adjustment of indirect observations, as it is called in the classical theory of errors. The essence of this problem is to find the "real", "most plausible" values of  $m$  ( $m < n$ ) unknown quantities  $x, y, z, \dots$  from a redundant system of algebraic equations

$$a_i x + b_i y + c_i z + \dots + l_i = 0, \quad i = 1, 2, \dots, n \quad (1.1)$$

where, for the case of arc measurements,  $m = 2$ ,  $n$  is the number of these measurements,  $l_i$ , the free terms, are furnished by the (physically independent) measurements, and the coefficients are calculated according to the general theory of the figure of the earth<sup>1</sup>.

Strictly speaking, systems (1.1) are inconsistent, so that any "solution" is just a set  $(x, y, z, \dots)$  deduced by introducing some more or less reasonably chosen additional restriction imposed upon the residuals,  $\Delta l_i$ .

The MLSq introduced in the early 19<sup>th</sup> century,

$$\Delta l_1^2 + \Delta l_2^2 + \dots + \Delta l_n^2 = \min$$

is no exception. I [26], [29] described several previous methods and I emphasized that one of the principal ones was Boscovich's, which, moreover, still attracts the attention of mathematicians and geodesists [2], [5], [38].

Boscovich is the sole author of Book 5 of [24], written jointly by him and Maire and published originally in 1755 in Latin. A short summary of the Latin original published in 1757, also in Latin, has been reprinted recently with a Serbo – Croatian translation [9]. It was in this summary [9, pp. 90 – 91] and then in the commentary on B. Stay's poem, in 1760, that Boscovich first expounded his method reprinted in [24, Book 5], whence the following passage is taken (p. 501):

*On doit tirer une certaine ellipticité moyenne de tous les degrés [of the meridian] connus par les observations, comparés entre eux, en ayant égard au rapport que doivent avoir leurs différences, et aux règles de la probabilité touchant la correction qu'il convient de leur faire pour les réduire à ce rapport.*

*Le P(ère) Boscovich l'a fait dans un autre ouvrage au moyen d'une méthode très curieuse, et qui peut servir en plusieurs autres cas. Il en a exposé le résultat dans un extrait inséré dans les actés de l'Institut de Boulogne (Memoriae de Bononiensi Scientiarum et artium Instituto atque Accademia, t. 4, 1757, pp. 353 – 34, whose reprint is [9].) Il la développe dans ses Supplémens de la Philosophie en vers latins (in 1760), composée depuis peu par M. B. Stay. Nous insérerons ici cet article en entier.*

### 1.1. Boscovich's Method of Adjusting Arc Measurements.

Boscovich's equations of the type (1.1) were

$$D_i - D_0 - d_i x = 0, \quad i = 1, 2, \dots, n \quad (1.1.1)$$

where the unknown  $D_0$  was the length of one degree of the meridian at the equator, the second unknown,  $x$ , another function of the parameters of the earth's spheroid,  $D_i$ , the measured lengths of one degree of the meridian at latitudes  $\varphi_i$ , and  $d_i$ , the coefficients calculated according to the given  $\varphi_i$ 's.

The additional restriction imposed on the  $\Delta D_i$ 's were

$$\Delta D_1 + \Delta D_2 + \dots + \Delta D_n = 0, \quad (1.1.2)$$

$$|\Delta D_1| + |\Delta D_2| + \dots + |\Delta D_n| = \min. \quad (1.1.3)$$

Assuming as he did that  $m$  and  $M$  are the mean values of  $d_i$  and  $D_i$  respectively, Boscovich could have arrived at

$$(D_i - M) - x(d_i - m) = 0$$

and, according to the essence of his reasoning (see below), could have chosen for  $x$  the median of the quotients  $(D_i - M):(d_i - m)$ .

Actually he employed a geometric procedure, noticing that the

solution sought is given by a straight line passing through the “centre of gravity”, i. e., through the point  $C(m, M)$

$$y - M = k(x - m) \quad (1.1.4)$$

with  $k$  equal to the median of the quotients  $(D_i - M):(d_i - m)$ .

Thus the straight line (1.1.4) is such that the vertical distances  $\Delta D_i$  from points  $(d_i, D_i)$  to this line satisfy relations (1.1.2) and (1.1.3). The Boscovich method therefore means a choice of a straight line passing reasonably close to each point of observation  $(d_i, D_i)$ .

**1.2. Justification.** How did Boscovich arrive at his conditions (1.1.2) and (1.1.3)? In [9, p. 46] he calculated the mean difference of latitudes  $\Delta\phi$  of the end points of his arc measurement not by directly assuming the mean of the four of his observations but by calculating the mean of their various binary combinations.

Similarly, in [9, pp. 90 – 91], [24, pp. 483 – 484] he solves a system of equations of the type (1.1.1) by arranging them in binary groups (so that the number of equations in each group coincides with the number of unknowns), solving each group separately and calculating the mean value of each unknown over the whole set of groups. Such a procedure, equivalent to adopting (1.1.2) for each group of equations separately, as if allowing for the equal probability of errors of each sign, and, what is much more interesting, in a sense similar to the MLSq (§ 1.4), was commonly used in the 18<sup>th</sup> century (T. Mayer, Lambert).

Boscovich [24, p. 484] was not satisfied with this method:

*Le milieu même differe encore beaucoup de plusieurs d'entre ces ... déterminations*

which possibly meant (§ 1.3.4) that the distribution of errors in the arc measurements was such that the MLSq would not have been recommended.

As to the arrangement of measurements (of the differences in latitudes) in binary groups in case of direct observations, it is my understanding that, just like the coincidence of terms (*milieu, Mittel*) used in the adjustment of both direct and indirect observations, this method testifies to a desire of the 18<sup>th</sup> century scholars for a unified point of view regarding the treatment of observations.

Besides this, it possibly was the general feeling that at least in some instances preliminary arrangement in binary groups cancels out systematic errors, and thus provides a means for a qualitative estimation of the random errors by comparing different groups.

Being, as mentioned above, dissatisfied with this usual method of solving equations of the (1.1.1) type, Boscovich [24, p. 501] says:

*Mais pour prendre ce milieu, tel qu'il ne soit point simplement un milieu arithmétique, mais qu'il soit plié par une certaine loi aux règles des combinaisons fortuites et du calcul des probabilités; nous nous servons ici d'un problème que j'ai indiqué vers la fin d'une Dissertation insérée dans les actes de l'Institut de Boulogne ... et où je me suis contenté de donner le résultat de la solution. Voici le problème: étant donné un certain nombre de degrés (arcs), trouver la correction qu'il faut faire à chacun d'eux, en observant ces trois*

conditions, la première, que leurs différences soient proportionnelles aux différences des sinus versés d'une latitude double (= equations (1.1.1) must hold): la seconde, que la somme des corrections positives soit égale à la somme des négatives (= condition (1.1.2)): la troisième, que la somme de toutes les corrections, tant positives que négatives, soit la moindre possible, pour le cas où les deux premières conditions soient remplies (condition (1.1.3)).

The second condition is needed

*Par un même degré de probabilité, pour les déviations du pendule et les erreurs des Observateurs, dans l'augmentation et la diminution des degrés; la troisième est nécessaire pour se rapprocher autant qu'il se pourra des observations.*

Leading to the use of the median, the condition (1.1.3) was of course essentially different from *simplement un milieu arithmétique*. However, without using ideas and methods related to the theory of laws of distribution, which at those times was impossible, it was equally impossible to say just how the imposed conditions corresponded to the *règles des combinaisons fortuites*.

### 1.3. History of Boscovich's Method

1.3.1. *Prehistory*. I (§ 1.2) have remarked that condition (1.1.2) was commonly used in the 18<sup>th</sup> century. This, however, is not so important because a strict equality (1.1.2) could be eliminated together with one of the unknowns (that was just what Boscovich did, see § 1.1), but condition (1.1.3) was more essential. It is interesting that in an astronomical context Daniel Bernoulli formulated (1.1.3), although not in mathematical notation. On the other hand, as he himself stated, he was unable to provide a method for applying this condition.

1.3.2. *Laplace*. He repeatedly used Boscovich's method [20, pp. 506 – 516], [21, § 40 of vol. 2] in an analytical form and mentioned Boscovich [20, § 10]:

*Boscovich a donné pour cet objet (adjustment of arc measurements) une méthode ingénieuse ... mais, comme il l'a inutilement compliquée de la consideration des figures, je vais le présenter ici sous la forme analytique la plus simple.*

The *plus simple* is, however, rather doubtful. I shall describe it employing a somewhat different notation. Let the initial equations be

$$a_i - z - p_i y = x_i, \quad i = 1, 2, \dots, n,$$

where  $y$  and  $z$  are unknown and  $x_i$  are the residuals. The additional conditions are of course

$$\begin{aligned} x_1 + x_2 + \dots + x_n &= 0, \\ |x_1| + |x_2| + \dots + |x_n| &= \min. \end{aligned}$$

Summing all the equations and subtracting the sum from each of them, Laplace arrived at

$$b_i - q_i y = x_i \text{ or } h_i y - c_i, \quad h_i > 0 \quad (1.3.1)$$

with the quotients  $c_i/h_i$  arranged in decreasing order. Then

$$y = c_r/h_r$$

Here, the subscript  $r$  corresponds to the sum of the  $h_i$ 's up to and including  $h_{r-1}$ /including  $h_r$  which is smaller/larger than the half-sum of the  $h_i$ 's.

This solution is tantamount to the median. This becomes obvious after replacing (1.3.1) by

$$sy - g_i$$

with the  $g_i$ 's arranged in decreasing order.

Not surprisingly, Boscovich's method was used also by the American astronomer N. Bowditch [10], best known as the translator of Laplace's *Mécanique Céleste*. In the translation of its volume 2 (Boston, 1832, reprinted New York, 1966), in a footnote in § 40 (pp. 434 and 438 of the reprint), Bowditch noticed that the

*Method, proposed by Boscovich, and peculiarly well adapted to the present problem (of adjusting arc measurements), is not now so much used as it ought to be ...*

*We shall ... find ... that the method of the least squares, when applied to a system of observations, in which one of the extreme errors is very great, does not generally give so correct a result as the method proposed by Boscovich. ... The reason is, that in the former method, this extreme error [as any other] affects the result in proportion to the second power of the error; but in the other method, it is as the first power.*

I return to this matter in § 1.3.4.

1.3.3. *Gauss and Linear Programming*. Giving a concise description of the different methods of adjusting indirect observations, Gauss [14, § 186] describes also Boscovich's method:

*Laplace made use of another principle for the solution of linear equations whose number is greater than the number of the unknown quantities, which had been previously proposed by Boscovich, namely, that the sum of the errors themselves taken positively, be made a minimum. It can be easily shown that a system of values of unknown quantities, derived from this principle alone, must necessarily (except the special cases in which the problem remains to some extent indeterminate. C. F. G.) exactly satisfy as many equations out of the number proposed, as there are unknown quantities, so that the remaining equations come under consideration only in so far as they help to determine the choice (of the equations to be exactly satisfied; *ad optionem decidendam conferunt* is the Latin original). ...*

*Besides, Laplace qualifies in some measure this principle by adding a new condition: he requires, namely, that the sum of the differences [residuals], the signs remaining unchanged, be equal to zero. Hence it follows, that the number of equations exactly represented may be less by unity than the number of unknown quantities; [Hinc efficitur, ut multitudo acquationum exacte repraesentatarum unitate minor fiat quam multitudo quantitatum incognitarum] but what we have before*

*said will still hold good if there are at least only two unknown quantities [siquidem duae saltem incognitae affuerint].*

Laplace did not add any condition: (1.1.2) is original with Boscovich and I do not know whether Gauss had ever found out his mistake. Legendre informed Gauss about another mistake he had made, once more in a historical context and later Gauss partly exonerated himself (Sheynin 2017, pp. 138 – 139). In any case, Gauss is known to have been reluctant to mention other authors: he had neither time nor inclination for historical research. Not good enough for a usual mortal!

The alleged Laplace’s condition is a new equation, exactly represented (the right side of the equation is exactly equal to zero).

Thus it occurs that Boscovich’s method amounts to finding the “best” pair (in case of two unknowns) of equations and using this pair only and it also becomes clear that Gauss knew an important theorem of the future theory of linear programming.

*1.3.4. Median Versus Arithmetic Mean.* The apparent soundness of Boscovich’s computational results as testified by the use of his method by many scholars, notably by Laplace, calls for an explanation in terms of probability. In the case of two unknowns one of which is eliminated by condition (1.1.2) his method, as repeatedly stated above, is tantamount to the use of the median. The comparative merits of the arithmetic mean and the median had been studied beginning with Laplace, e. g., [22, p. 576], who introduced a special term for the use of the median (*méthode de situation*) and deduced a relation showing the comparative merits of both estimators in terms of the corresponding law of distribution.

Quite a few writings on the same subject are due to F. Y. Edgeworth. Kolmogorov [19] offered a simple criterion and he also asserted that the choice of the median was reasonable in case of unknown distributions. Nowadays the median is considered as a particular case of order statistics, and its possible use is studied in terms of the appropriate general theory.

Estienne [13] noticed an extremely simple fact: for distribution

$$\varphi(x, x_0) = C \exp(-h^2|x - x_0|) \quad (1.3.5)$$

the maximum probability of the realization of a given series of observations  $l_1, l_2, \dots, l_n$  is furnished by the use of the median. Indeed, this probability is proportional to

$$\exp(-h^2(|l_1 - x_0| + |l_2 - x_0| + \dots + |l_n - x_0|))$$

and its maximum corresponds to the choice of the median as the estimator of the unknown parameter  $x_0$ . Laplace used distribution (1.3.5) as early as in 1774.

The solution of a system of linear equations in  $m$  unknowns (the general case) is a linear combination of  $m$  quantities chosen out of the absolute terms  $l_1, l_2, \dots, l_n$  (§ 1.3.3) so that the statistical properties of the estimators of the unknowns are again dependent on the law of distribution of the errors of these terms. It could be possibly inferred

that the successful use of the Boscovich method implies that the distribution of errors in astronomical and geodetic observations were (and possibly are) in some sense close to the law (1.3.5).

**1.4. The Relation between the MLSq and that of Combining Equations.** As mentioned in § 1.2, the usual way of arriving at a sensible set of  $n$  unknowns was to arrange the initial equations into subsets of  $n$  equations each, solve each subset and average the values of each unknown, a procedure used with  $n = 2$ .

Jacobi [18] discovered a connection between this method and the MLSq. He derived a system of weights of the different subsets of equations which allowed the two methods to furnish identical results. However, as implicitly stated in § 1.2 (and repeated above), it was the usual practice to ignore these weights and take the usual (not weighted) arithmetic mean over all the subsets.

It seems that at least a sufficient condition for a proper weighing is that every subset of  $n$  equations has one and only one solution and that the matrix of the derived normal equations in the usual way is symmetric and positive definite. At present, this method is best found in Whittaker & Robinson [34a, p. 251].

Glensvik (1967) indicated that instead of solving those subsets it was possible to solve the appropriate parts of the normal equations and again to take the mean of the results. He also determined the weights of the thus derived unknowns<sup>2</sup>.

## 2. Boscovich's Manuscripts

Those manuscripts are not collected together in one place, and some may be irrevocably lost [15]. However [1], the University of California library possesses more than 180 of them (in his hand and in hands of copyists), his correspondence, including 420 of his letters and more than 1500 letters to him, his diary etc. Among the mathematical manuscripts at this library are two pertaining to the theory of probability which I describe in §§ 2.1 and 2.2.

From their general style it is evident that both these manuscripts are written for laymen, one of them being moreover directly addressed to a clergyman. and it remains an open question whether among Boscovich's correspondence not at the University of California and/or among the lost manuscripts there exist (existed) more profound writings on the same subject.

**2.1. De calculo Probabilitatum, etc.** [6]. In § 1 of the manuscript Boscovich formulates his problem: to find the probability of the error in the sum of errors of observations if the individual errors are equal to 1, 0 or  $-1$ . Although he does not say so, these possible values are meant to be equally probable. Then, though he speaks about probabilities, he actually calculates chances. Also in this section Boscovich explains the formula for calculating the number of combinations  $C_p^m$  and includes a table of this quantity (of binomial coefficients) up to and including  $p = m = 8$ .

In the next three sections Boscovich calculates the chances for the sum of errors to be 0, 1 and 2. For instance, a zero sum corresponds to the following cases:

(1) each observation has a zero error (one chance);

(2) one observation has an error equal to 1, another one, to  $-1$ , with zero errors of each of the rest observations ( $C_n^1 C_{n-1}^1 = n(n-1)$  chances);

(3) two observations have an error equal to 1, two other observations, to  $-1$  with zero errors of the rest observations ( $C_n^2 C_{n-2}^2$ ) etc.

In § 5 Boscovich calculates the chances for the sum of errors to equal 0, 1, 2, ..., 7, 8. In § 6 he explains in detail the law according to which these formulas are constituted and, in particular, notices that each series of the products of two combinations is continued until the upper index becomes equal to the lower. The last section, § 7, is a summary for  $n=1, 2, \dots, 7, 8$ .

No generalizations are considered: neither a large  $n$ , nor the general discrete uniform distribution with possible values of errors being,  $0, \pm 1, \pm 2, \dots, \pm N$ , nor, lastly, the continuous uniform distribution. Even the arithmetic mean of the errors, the estimator directly needed in practice, is not mentioned.

It terms of probability rather than chances the Boscovich problem could be formulated thus. Equally probable values of each of  $n$  random variables  $\xi_1, \xi_2, \dots, \xi_n$  are 0, 1, and  $-1$ . To find the law of distribution of their sum ( $\xi$ ). The generally known formula, whose equivalent Boscovich knew, is

$$P(\xi = a) = \Sigma [P(\xi_1 = x)P(\xi_2 = y) \dots P(\xi_n = w)], \quad (2.1.2)$$

with the summation extending over all the values of  $x, y, \dots, w$  complying with the conditions

$$x + y + \dots + w = a, \quad x, y, \dots, w = 1, \text{ or } 0, \text{ or } -1.$$

This is a formalized record of the problem, but it does not provide a means for actual computations.

The manuscript under consideration is undated and its evaluation is hardly possible. If it was written in 1750 – 1753, i. e. during his participation in the arc measurement, or at least before 1756, Boscovich should be regarded as the first to use a quantitative stochastic method in the theory of errors and, thus, as a precursor of T. Simpson. Otherwise, the manuscript is just not interesting, for its essence would have presented nothing new to Montmort or De Moivre even at the beginning of the century [31]. An indirect argument to the effect that the manuscript had been written before 1758 is proposed in § 3.

**2.2. Breve memoria sul lotto di Roma, etc.** [7]. The *Memoria* includes a courteous address to the Eminenza (Cardinal Lante) dated 1765 from which it follows that the Cardinal had already discussed the lotto with Boscovich, and, anyway, it is known that the latter was a frequent visitor of the Cardinal [16, p. 82].

The lottery discussed consisted of 90 tickets, five of which were drawn at a time. The gamblers enjoyed an option of guessing either one or some of, or even all, of the tickets drawn. The more tickets a

gambler endeavoured to guess, the more was his gain in case of success. The conditions of this classical lottery are explained in detail, and various related problems are described in [4].

The *Memoria* proper consists of three sections where

- (1) The rule for calculating the number of combinations is explained and complemented by examples related to the lotto;
- (2) The expectation of the receipts of the banker from each option is calculated, and the inference reached is that the lion's share of the receipts comes from the more venturesome, from those opting to guess more than one ticket.

Tacitly assumed is the equal probability of the choice of various tickets by the gamblers. Actually, however, some of the gamblers could have well pursued their own "systems", and, moreover, a large part of the gamblers in a given locality could have pursued one and the same system. Thus, for example, Laplace [22, *Essai philosophique*, Chap. 16] noticed that many gamblers erroneously thought that tickets which were not drawn for a long time will be more probably drawn. Alternatively, other gamblers favoured the opposite point of view!

Under other historical conditions other systems independent of previous drawings were used [37]. For example, the conditions of an illegal and primitive lottery which flourished in Harlem in this century<sup>3</sup> were that the better had to choose one or a few three-digit numbers. The odds were 1 to 1,000 and the payoff for winners 600 to 1.

*During Thanksgiving week, [37, p. 74], there was a superstition about the number 527. A great many played it and its combinations, 257, 275, ... If any of these numbers hit, most of the (petty) banks would go broke.*

In any case, a *system* pursued by a large part of gamblers constitutes a bias from an equally probable choice of tickets. Although Boscovich mentions a special *fondo sicuro*, he does not consider either the risk arising from a bias of this kind, or even from a possible guessing of four or five tickets. However, such considerations did not yet exist in those times. Besides, it is possible that at least in some cases bankers did not bear any risk at all or, alternatively, risked only a limited sum (if, for example, the total payoff did not exceed the total payment).

At the French lottery, about half a century later [11, § 59],

*L'expérience avait rassuré l'administration de la loterie sur l'influence que pouvaient exercer, à chaque tirage en particulier, les préjugés (= the "systems") ... et elle n'usait, plus du droit qu'elle s'était réservé, de fermer les numéros trop chargés.*

Cournot (§ 60) goes on to say that the *quine*, which is the option of guessing all the five tickets, had been finally suppressed:

*L'administration avait fini par supprimer cette chance, soit pour s'épargner toute inquiétude, soit parce que le quine se jouait trop rarement pour que le produit de la spéculation sur cette chance valût la peine d'en compliquer la comptabilité.*

Neither did Boscovich consider related theoretical problems such as the probability of drawing two or more consecutive numbers at once (Euler) or the probability of drawing each number at least once after a

given number of drawings (Laplace). And of course he did not forward any moral arguments against lotteries (although he possibly mentioned such arguments in oral discussions with the Cardinal) as did Laplace [23], [27]<sup>4</sup>.

Taken as a whole, the *Memoria*, only an elementary exposition of the practical side of a lottery, is of minor interest.

### 3. A Glimpse of a Stochastic Reasoning in Boscovich's Physics

His physics is essentially deterministic as is also his philosophy. For example, he writes [8, § 458]:

*In order that at any subsequent time it (the matter) may have the determinate state, which it actually has, it must be determined to that state, from the state just preceding. ... A preceding state cannot determine the one which follows it, except in so far as it itself has existed determinately. ... This preceding state ... derives it (its own determination) from one that precedes it.*

He then goes on to say that an infinite number of states with a zero determination each, taken together, have a zero determination, so that actually the determination must be given by a *Being* from the outside<sup>5</sup>.

Related passages are in Boscovich's neighbouring sections [8, §§ 547, 550, and 551]. In § 550 he speaks about *blind chance* in the same way as other scholars of the 18<sup>th</sup> century [27].

We ought to note that randomness was also present in Boscovich's physics. Thus, [8, § 481] the

*Connection that exists between the points (atoms) of matter forming the particle (of light, see his § 477, or, alternatively, particles in any bodies, § 478) and moving together with practically the same velocity will force the entire particle to move as a whole with the single motion that is induced by the sum of the inequalities pertaining to all its points; and this sum will still further approximate to equality. For, in circumstances that are fortuitous, distributed here and there at random, or concurring by chance, the greater the number taken, the more the sum of the irregular inequalities decreases.*

Such a decrease hardly occurs!

Boscovich did not formalize his reasoning, but it seems that in accord with the prevalent ideas of the 18<sup>th</sup> century [28] he had in mind an equally probable distribution of the velocities of points of matter, i. e. a discrete distribution with equally probable values of velocities being, say,

$$v, v \pm \Delta v, v \pm 2\Delta v, \dots, v \pm n\Delta v \quad (3.1)$$

(not the continuous case which, it seems, would have been contrary to his general philosophical outlook and certainly not the then unthought of Maxwell distribution, appropriate to a gas in equilibrium).

If so, (3.1) could be regarded as a generalization of the distribution which Boscovich studied in one of his manuscripts (§ 2.1), so that his reasoning just quoted is possibly a qualitative estimation of the behaviour of the sum of random variables. Written in the mid-18<sup>th</sup> century, it is not very important per se [31]. However, it is the first reasoning of this sort in a physical context and should be praised as such.

An interesting corollary follows: it seems possible that this reasoning, published in [8] in 1758, had been formulated after the writing of [6] where only the most elementary form of (3.1) is considered. The meaning of this possible fact is emphasized in § 2.1.

Other stochastic reasoning is in the same source [8, §§ 479 and 495]. In § 479 Boscovich writes:

*Imagine a sphere, that has for its semi-diameter the distance ... up to which the forces of each of the points (of matter) are fairly sensible. If the medium approximates sufficiently closely to homogeneity, and the sphere is divided into any two parts by a plane through the centre, the number of points of matter in each part will be nearly the same; and the sum of the forces will be very approximately the same, as the slight differences taken as a whole compensate one another in so great a multitude; for this is always the case in sufficiently numerous fortuitous combinations.*

As I see it, this passage also bears upon the distribution of sums of random variables, although neither here Boscovich formalized his reasoning. He did not arrive at a developed kinetic theory of heat. Whyte [35a] offered a possible explanation for the lack of such a theory in his writings (and, for that matter, in the 18<sup>th</sup> century in general):

*Though Boscovich was in principle concerned with all possible arrangements and modes of interaction of puncta ( atoms ) he concentrated his attention on those properties which appeared to him simplest. ... Though interested in the theory of probability, he did not consider applying statistical methods to random motions.*

But perhaps the main reason was Boscovich's general philosophical outlook.

*Acknowledgement.* Many authors [33], [36], [34], [12], [25] treated the subject of § 1, but my account of Boscovich's method proper and of its history is more detailed. In particular, neither Daniel Bernoulli nor Gauss had been mentioned before in connection with Boscovich.

In a concise form the subject of § 1.4 is described in [34a, § 128], a fact brought to my attention by Dr. C. Eisenhart. The subject of § 2 is completely original. Acknowledgement is here due to J. G. Michell, Assistant Librarian, University of California, for copies of Boscovich's manuscripts and to Olga V. Ioselevitch for an oral translation of [7]. Ivanovic [17], whose comments, however, are physical rather than mathematical, noticed the reasoning of Boscovich on randomness in a physical context (§ 3). I have used my previous writings. These include [26], which touches on the subject of § 1; [31], where a concise description of the subject of § 2.1 is given; [30], which has been deposited as a manuscript at the Institute for Scientific Information (Moscow) and is presently unpublished; [27], where, hidden in a footnote, are a few lines related to the subject of § 3 and, lastly, [32], where Daniel Bernoulli's anticipation of the main condition of Boscovich's adjustment procedure (§ 1) is considered in more detail.

## Notes

1. Actually,  $a_i = 1$ . I also notice that the case of only one unknown is called *adjustment of direct observations*. Systems (1.1) are linear since the parameters sought are approximately known and in any case since their approximate values can be calculated by solving any subsystem of (1.1) consisting of  $m$  equations.

2. Glaisher [14a, p. 614] reported that Cayley, to whom he communicated Jacobi's results, was already acquainted with them.

3. It seems that Laplace developed his negative attitude against lotteries following the reasoning of Condorcet [10a, p. 162]:

*Une loterie est un impôt volontaire, parce qu'il est payé par ceux qui veulent jouer à ce jeu.*

A similar pronouncement was due to Petty back in 1662 (Sheynin 2017, p. 38).

4 This problem is beyond the scope of the present paper, and I add only a few comments. First, a related passage occurs in Spinoza [32a, p. 31] who strove to prove that a chain of random events is impossible:

*Vielleicht wird jemand sagen, dass etwas Zufälliges ... eine zufällige (Ursache) habe.*

But then, he argues, this chain goes *ins Unendliche ... was offenbar falsch ist*. Spinoza supposes that this proves that *es aber keine zufälligen Dinge giebt*.

Second, the so called Laplacian determinism differs from that of Boscovich in that future events are also determined and in that no *Being* is supposed to exist. Third, this determinism is not his at all, and I think that it should be more correctly called ancient – Newtonian – Laplacian [27]. In particular, one of Laplace's immediate precursors in determinism was Maupertuis, whose opinion on this matter had not been noticed until now. This opinion, which I quote, is from his *Sur la divination* which constitutes *Lettre 18* of his *Lettres (Oeuvr., t. 2. Lyon, 1756, pp. 185 – 340)*. It occupies pp. 298 – 306 of the volume and the quotation is from p. 300:

*Ce n'est pas que tout étant lié dans la Nature, un esprit assez vaste ne put, par la petite partie qu'il aperçoit de l'état présent de l'Univers, découvrir tous les états qui l'ont précédé, et tous ceux qui doivent le suivre: mais nos esprits sont bien éloignés de ce degré d'étendue.*

Another Laplace's immediate precursors in determinism, at least as regards men's free will, was Kant [18a, p. 111]:

*Denn so zufällig wie auch immer die Entschliebung zum Heirathen sein mag, so findet man doch in eben demselben Lande, dass das Verhältnis der Ehen zu der Zahl der Lebenden ziemlich beständig sei, wenn man große Zahlen nimmt.*

5. Highly relevant is a recent article by T. A. Johnson *Numbers are called balm of Harlem* (*New York Times*, 1971, March 1, pp. 1 and 42).

## References

1. Archives of R. J. Boscovich (at the University of California). No place, no year, 5 pp. (unpubl.).
2. BEJAR J., Regresion en mediana y la programacion lineal. *Trabajos de estadistica*, t. 7, 1956, pp. 141 – 158.
3. BERNOULLI D., Recherches physiques et astronomiques sur le problème ... Quelle est la cause physique de l'inclinaison des plans des orbites des planets. 1735. In Latin, in author's *Werke*, Bd. 3. Basel, 1987, pp. 303 – 326.
4. BIERMANN K.-R., Problems of the Genoese lottery in the works of classics in probability. *Istoriko-matematich. issledovania*, vol. 10, 1957, pp. 649 – 670. In Russian.
5. BOLOTIN A. I., On a method of arriving at the minimum of a sum of absolute values of quantities linearly depending on a number of arguments. *Izv. vuzov. Geod. i aerofotos'emka*, 1965, No. 4, pp. 15 – 22. In Russian.
6. BOSCOVICH R. J., De calculo probabilitatum que respondent diversis valoribus summe errorum post plures observationes, quarum singule possint esse erronee certa quadam quantitate. Undated manuscript, 8 pp. From the Boscovich Archive, Dept of Rare Books and Special Collections, Univ. of California Library, Ms. No. 62.
7. BOSCOVICH R. J., Breve memoria sul lotto di Roma presentata a sua eminenza il signor cardinal Lante nella sua magnifica villa di Bagnaja. Manuscript in the hand of a copyist, 18 pp., including an address to the Eminenza, dated 1765. From the Boscovich Archive, etc., see [6], Ms. No. 65.
8. BOSCOVICH R. J., *Philosophiae naturalis theoria* (1758). Latin – English edition: Chicago – London, 1922.
9. BOSCOVICH R. J., De litteraria expeditione per pontificiam ditionem ad dimetiendos duos meridiani gradus et corrigendam mappam geographicam. In CUBRANIC N.,

- Geodetski rad R. Boskovicica*. Zagreb, 1961. The *Expositione* occupies the main part of the book.
10. BOWDITCH N. Observations of the comet of 1807. *Mem. Amer. Acad. Arts and Sci.*, vol. 3, pt. 1, 1809, pp. 1 – 17.
- 10a. CONDORCET M. J. A. N. CARITAT DE, Des impôts volontaires, et des impôts sur le luxe. (1788). *Oeuvr. compl.*, t. 14. Brunswick – Paris, 1804, pp. 162 – 190.
11. COURNOT A. A., *Exposition de la théorie des chances et des probabilités*. Paris, 1843, 1984. Editor B. Bru. **S, G**, 54.
12. EISENHART C., Boscovich and the combination of observations [35b, pp. 200 – 213]. A shorter version of same published in *Actes symp. intern. Boscovich 1961*. Beograd, 1962, pp. 19 – 25.
13. ESTIENNE J. E., Etude sur les erreurs d'observation. *Rev.artil.*, t. 36, 1890, pp. 235 – 259.
14. GAUSS C. F., *Theoria motus corporum coelestium* (1809). *Werke*, Bd. 7. Gotha, 1871 (the whole volume). English translation by C. H. DAVIS: Boston, 1857, 2009.
- 14a. GLAISHER J. W. L., On the method of least squares. *Monthly Notices Roy. Astron. Soc.*, vol. 40, No. 9, 1880, pp. 600 – 614.
15. HAHN R., The Boscovich archives at Berkeley. *Isis*, vol. 56, pt. 1, No. 183, 1965, pp. 70 – 78.
16. HILL E., Biographical essay of R. J. Boscovich. [35b, pp. 17 – 101].
17. IVANOVIC D. M., Molecules, stars and Boscovich's law. *Atti Conv. intern. celebr. 250 anniv. nascita Boscovic etc. 1962*. Milano, 1963, pp. 243 – 250.
18. JACOBI C. G. J., *Über die Bildung und die Eigenschaften der Determinanten* (1841). Ostwalds Klassiker No. 77, pp. 3 – 49. Leipzig, 1896, Hrsg. P. STÄCKEL.
- 18a. KANT I., Der einzig mögliche Beweisgrund zu einer Demonstration des Daseins Gottes. (1763). *Ges. Schriften*, Bd. 2. Berlin, 1912, pp. 63 – 163.
19. KOLMOGOROV A. N., La méthode de la médiane dans la théorie des erreurs. *Rec. math.*, t. 38, No. 3 – 4, 1931, pp. 47 – 50. In Russian; title of periodical and of contribution also in French.
20. LAPLACE P. S., Sur les degrés mesurés des méridiens et sur les longueurs observées du pendule, 1789 (1792). *Oeuvr. compl.*, t. 11. Paris, 1895, pp. 493 – 516. This memoir comprises §§ 8 – 14 of author's Sur quelques points du système du monde.
21. LAPLACE P. S., *Mécanique céleste*, t. 2 (1799). *Oeuvr. compl.*, t. 2. Paris, 1878.
22. LAPLACE P. S., *Théorie analytique des probabilités*. Supplément deuxième (1818). *Oeuvr. compl.*, t. 7, pt. 2, pp. 531 – 580. Paris, 1886.
23. LAPLACE P. S., Sur la suppression de la loterie. (1819). *Oeuvr. compl.*, t. 14. Paris, 1912, pp. 375 – 378. **S, G**, 58.
24. MAIRE [C. ], BOSCOVICH, [R. J.], *Voyage astronomique et géographique dans l'Etat de l'Eglise etc.* Paris, 1770.
25. MARKOVIC Z., *R. Boskovic*. tt. 1 – 2. Zagreb, 1968 – 1969; see t. 1, p. 360 – 365.
26. SHEYNIN O. B., On the history of adjustment procedures in the method of indirect observations. *Izv. vuzov. Geod. i aerofotos'emka*, No. 3, 1967, pp. 25—32. in Russian. **S, G**, 111.
27. SHEYNIN O. B., Newton and the classical theory of probability. *Arch. Hist. Ex. Sci.*, vol. 7, No. 3, 1971, pp. 217 – 243.
28. SHEYNIN O. B., J. H. Lambert's work on probability. *Arch. Hist. Ex. Sci.*, vol. 7, No. 3, 1971, pp. 244 – 256.
29. SHEYNIN O. B., On the mathematical treatment of observations by L. Euler. *Arch. Hist. Ex. Sci.*, vol. 9, No. 1, 1972, pp. 45 – 56.
30. SHEYNIN O. B., On two manuscripts of R. Boscovich pertaining to the theory of probability. Unavailable manuscript.
31. SHEYNIN O. B., Finite random sums. A historical essay. *Arch. Hist. Ex. Sci.*, vol. 9, No. 4 – 5, 1973, pp. 275 – 305.
32. SHEYNIN O. B., D. Bernoulli's work on probability. 1972. In Kendall M. G., Plackett R. L., *Studies in history of stat. and prob.*, vol. 2. London, 1977, pp. 105 – 132.
- 32a. SPINOZA B., Kurzgefasste Abhandlung von Gott, dem Menschen und dessen Glück. First publ. 1862. Berlin, 1874 (*Philos. Bibl.*, Bd. 18).
33. TODHUNTER I., *History of the mathematical theories of attraction and the figure of the earth*, vols. 1 – 2, 1873. A one-volume reprint: New York, 1962. See §§ 511 and 514.
34. VARICAK V., *Matematički rad Boskoviccev*. *Rad Jugoslav. Akad. znan. i umjetn.* t. 181, No. 47, 1910. Zagreb. See pp. 75 – 208 (the relevant part is on pp. 173 – 187).
- 34a. WHITTAKER E. T., ROBINSON G., *Calculus of observations*, 1924. London, 1949.
- 35a. WHYTE L. L., Boscovich's atomism [35b, pp. 102 – 126].
- 35b. WHYTE L. L., *R. J. Boscovich. Studies of his life and work*. London, 1961.
36. WOLF R., *Handbuch der Astronomie, ihrer Geschichte und Literatur*, Bd. 2, Dritter Halbbd. Zürich, 1892. See§ 425.
37. WOLFERT I., *Tucker's people*. New York, 1943.
38. ZUKHOVITSKY S. I., AVDEEVA L. I., *Linear and convex programming*. Moscow,

1967. In Russian.

HALD A., *History of mathematical statistics from 1750 to 1930*. New York, 1998.

GLEINSWIK P., Generalization of the Jacobi theorem. *Bull. Geod.*, t. 85, 1967, pp. 269 – 281.

SHEYNIN O., *Theory of probability. Historical Essay*. Berlin. **S, G**, 10.

## VIII

**Jacob Bernoulli**

### **On the Law of Large Numbers**

*Ars Conjectandi*. Basileae, 1713, pt. 4

#### **Contents**

##### **Preliminary explanation**

##### **Foreword**

1. The *Art of Conjecturing* and Its Contents
2. The *Art of Conjecturing*, Part 4
  - 2.1. Randomness and Necessity
  - 2.2. Stochastic Assumptions and Arguments
  - 2.3. Arnauld and Leibniz
  - 2.4. The Law of Large Numbers

##### **The Art of Conjecturing, Part 4**

- Chapter 1
- Chapter 2
- Chapter 3
- Chapter 4.
- Chapter 5

##### **Notes**

##### **References**

### **Preliminary explanation**

Since 1713 the *Ars Conjectandi* (AC) has appeared in a German translation whereas its Part 4 was translated into Russian and French (and, in an horrible unsatisfactory way, into English), see the references. The German translation, especially insofar as mathematical reasoning is concerned, is rather far from the original; the Russian text also somewhat deviates from Bernoulli; finally, the French translation is perhaps almost faultless in this sense, but the translator made several mathematical mistakes.

I do not read Latin and had to begin from the Russian text, but I invariably checked my work against the two other translations and the several English passages from the AC which Shafer (1978) had provided, as well as against the original with the help of a Latin dictionary. I am really thankful to Claus Wittich (Geneve) who kindly went over my own text and made valuable suggestions and corrections. I am confident that the final result is good enough and in any case better than any translation mentioned above, but any remaining shortcomings and/or mistakes are my own.

A few words about Markov are in order. He initiated, and then edited the 1913 Russian translation. The same year he put out the third, the jubilee edition, as he called it, of his treatise (see References) and supplied it with Bernoulli's portrait. Again in 1913, he initiated a special sitting of the Imperial [Petersburg] Academy of Sciences devoted to Bernoulli's work in probability and, along with two other mathematicians, delivered a report there, first published in 1914, reprinted in Bernoulli (1986) and available in an English translation (Ondar 1977/1981, pp. 158–163).

Later, in the posthumous edition of his treatise (1924), Markov improved Bernoulli's estimates (§ 2.4), as Pearson did at about the same time and, perhaps as an indirect result of his study of the AC, inserted there many interesting historical comments.

I had previously privately printed the same translation, see **S, G, 8**, but now I am not satisfied by it.

### **FOREWORD**

#### **1. The *Art of Conjecturing* and Its Contents**

Jacob Bernoulli (1654 – 1705) was a most eminent mathematician, mechanician and physicist. His AC (1713) was published posthumously with a Foreword by his nephew, Nicolaus Bernoulli (English translation: David (1962, pp. 133 – 135); French translation, Jacob Bernoulli (1987, pp. 11 – 12)). It is not amiss to add that N. B. (1709) published his dissertation on the application of the art of conjecturing to jurisprudence where he not only picked up some hints included in the manuscript of his late uncle, but borrowed whole passages both from it and even from the *Meditationes*, never meant for publication (Kohli 1975b, p. 541).

The *Meditationes* is Bernoulli's diary. It covers approximately the years 1684 – 1690 and is important first and foremost because it contains a fragmentary proof of the law of large numbers (LLN) to which Bernoulli indirectly referred at the end of Chapter 4 of Part 4 of the AC. Other points of interest in the *Meditationes* are that he (1975,

p. 47) noted that the probability (in this case, statistical probability) of a visitation of a plague in a given year was equal to the ratio of the number of these visitations during a long period of time to the number of years in that period. Then, Bernoulli (p. 46, marginal note) wrote out the imprint of a review published in 1666 of Graunt's book (1662) which he possibly had not seen; he had not referred to it either in the *Meditationes* itself or in the AC. And, lastly, at about the same time Bernoulli (p. 43) considered the probability that an older man can outlive a young one (cf. Item 4 in Chapter 2, Part 4 of the AC). All this, even apart from the proof of the LLN, goes to show that already then he thought about applying statistical probability.

Part 1 of the AC is a reprint of Huygens' tract (1657) complete with vast and valuable commentaries. Nevertheless, this form testifies that Bernoulli was unable to complete his contribution. Also in Part 1 Bernoulli (pp. 22 – 28 of the German translation), while considering a game of dice, compiled a table which enabled him to calculate the coefficients of  $x^m$  in the development of  $(x + x^2 + \dots + x^5 + x^6)^6$  for small values of  $n$ . Part 2 dealt with combinatorial analysis and it was there that the author introduced the *Bernoulli numbers*. Part 3 was devoted to application of the "previous" to drawing of lots and games of dice.

Parts 1 and 3 contain interesting problems: the study of random sums for the uniform and the binomial distributions; a similar investigation of the sum of a random number of terms for a particular discrete distribution; a derivation of the distribution of the first order statistic for the discrete uniform distribution; and the calculation of probabilities appearing in sampling without replacement. The author's analytical methods included combinatorial analysis and calculation of expectations of winning in each set of finite and infinite games and their subsequent summing.

Finally, Part 4 contained the LLN. There also we find a not quite formal "classical" definition of probability (a notion which he had not applied when formulating that law), a reasoning, in Chapter 2, on the aims of the art of conjecturing (determination, as precisely as possible, of probabilities for choosing the best solutions of problems, apparently in civil life) and elements of stochastic logic. Strangely enough, the title of Part 4 mentioned the completely lacking applications of the "previous doctrine" whereas his main theorem (the LLN) was not cited at all. This again testifies that Bernoulli had not completed his work. He did state, however (Chapter 4) that his LLN provided moral certainty which was sufficient for civil life and at the end of Chapter 2 he even maintained that judges must have firm instructions about what exactly constituted it.

Moral certainty had first appeared about 1400 (Franklin 2001, p. 69), but it was Descartes (1644, p. 323) who put it into circulation (above all apparently bearing in mind jurisprudence!). Huygens (Sheynin 1977, pp. 251 – 252) believed that proofs in physics were only probable and should be checked by appropriate corollaries and that common sense ought to determine the required degree of certainty of judgements in civil life. This latter statement seems much more reasonable than Bernoulli's rigid demand.

Bernoulli apparently considered the art of conjecturing as a mathematical discipline based on probability as a measure of certainty and on expectation which included (the not yet formally introduced) addition and multiplication theorems and crowned by the LLN.

## **2. The Art of Conjecturing, Part 4**

**2.1 Randomness and Necessity.** Apparently not wishing to encroach upon theology, Bernoulli (beginning of Chapter 1) refused to discuss the notion of randomness. Then, again in the same chapter, he offered a subjective explanation of the “contingent” but actually corrected himself at the beginning of Chapter 4 where he explained randomness by the action of numerous complicated causes. Finally, the last lines of his book contain a statement to the effect that some kind of necessity was present even in random things (but left too little room for it). He referred to Plato who had taught that after a countless number of centuries everything returned to its initial state. Bernoulli likely thought about the archaic notion of the Great Year whose end will cause the end of the world with the planets and stars returning to their positions at the moment of creation. Without justification, he widened the boundaries of applicability of his law and his example was, furthermore, too complicated. It is noteworthy that Kepler (1596) believed that the end of the world was unlikely. In this, the first edition of this book, his reasoning was difficult to understand but later he substantiated his conclusion by stating, in essence, like Oresme (1966, p. 247) did before him, that two [randomly chosen] numbers were “probably” incommensurable.

Bernoulli borrowed his example of finding a buried treasure from Aristotle (end of Chapter 1) but, unlike him, only indirectly connected it with randomness. The later understanding of randomness began with Maxwell and especially Poincaré, who linked it with (among other interpretations) with the case in which slight causes (digging the earth somewhere near) would lead to considerable effects (the treasure remained buried) and numerous complicated causes (here, he repeated Bernoulli). Poincaré also sensibly reasoned on the interrelations between randomness and necessity. On randomness see Sheynin (2014); new ideas took root late in the 20<sup>th</sup> century.

**2.2. Stochastic Assumptions and Arguments.** Bernoulli examined these in Chapters 2 and 3, but did not return to them anymore; he possibly thought of applying them in the unwritten pages of his book. The mathematical aspect of his considerations consisted in the use of the addition and the multiplication theorems for combining various arguments.

Unusual was the non-additivity of the deduced [probabilities] of the events under discussion. Here is one of his examples (Chapter 3, Item 7):

“Something” possesses  $2/3$  of certainty but its opposite has  $3/4$  of certainty. Both possibilities are probable and their probabilities are as 8:9. Koopman (1940) resumed, in our time, the study of non-additive probabilities whose sources can be found in the medieval doctrine of probabilism that considered the opinion of each theologian as probable. Franklin (2001, p. 74) traced the origin of probabilism to the year 1577, or, in any case (p. 83), to 1611. Nevertheless, similar

pronouncements on probabilities of opinion go back to John of Salisbury (the 12<sup>th</sup> century) and even to Cicero (Garber & Zabell 1979, p. 46).

I note a “general rule or axiom” concerning the application of arguments (pp. 234 and 236): out of two possibilities, the safer, the more reliable, etc. should be chosen.

On the subject of this subsection see Shafer (1978) and Halperin (1988).

Bernoulli derived many formulas which I had not copied. I believe that no one had or will ever apply them, but they are inserted in any full translations of the *Ars* and certainly in Bernoulli (1975).

**2.3. Arnauld and Leibniz.** Antoine Arnauld (1612 – 1694) was an extremely well known religious figure and philosopher, the main author of the influential treatise Arnauld & Nicole (1662). In Chapter 4 Bernoulli praised Arnauld and approved his reasoning on using posterior knowledge and at the end of Chapter 3 Bernoulli borrowed Arnauld’s example (1662, pp. 328 – 329) of the criminal notary. Other points of interest are Arnauld’s confidence in moral certainty and his discussion of the application of arguments. It might be reasonably assumed that Arnauld was Bernoulli’s “non-mathematical” predecessor.

In 1703, Bernoulli informed Leibniz about the progress in his work (Kohli 1975, p. 509). He had been compiling it for many years with repeated interruptions caused by his “innate laziness” and worsening of health; the book still lacked its “most important part”, the application of the art of conjecturing to civil life; nevertheless, he, Bernoulli, had already shown his brother [Johann] the solution of a “difficult problem, special in its own way” that justified the applications of the art of conjecturing.

Most important both in that letter and in the following correspondence of 1703 – 1705 (Ibidem, pp. 510 – 512) was the subject of statistical probabilities. Leibniz never agreed that observations could secure moral certainty, but his arguments were hardly convincing. Thus, he in essence repeated the statement of Arnauld & Nicole (1662/1992, pp. 304 and 317) that the finite (the mind; therefore, observations) could not always grasp the infinite (for example, God, but also, as Leibniz stated, any phenomenon depending on innumerable circumstances).

Leibniz’ views were possibly caused by his understanding of randomness as something “whose complete proof exceeds any human mind” (manuscript, 1686/1960, p. 288). His heuristic statement does not contradict a modern approach to randomness founded on complexity and he was also right in the sense that statistical determinations cannot definitively corroborate a hypothesis.

In his letter of 3 Dec. 1703 Leibniz (Gini 1946, p. 405) also maintained that the allowance for all the circumstances was more important than subtle calculations, and Bortkiewicz (1923, p. 12) put on record Keynes’ (1921) favourable attitude towards this point of view and indicated the appropriate opinion of Mill (1843/1886, p. 353), who had sharply contrasted the consideration of circumstances with “elaborate application” of probability and declared that the

“neglect of this obvious reflection” made probability “the real opprobrium of mathematics”. Bortkiewicz agreed that mathematicians had been sometimes guilty of such neglect, which, however, had nothing to do with the calculus of probability. In his Chapter 4, Bernoulli touched on medical statistics and, for my part, I note that its progress is accompanied by the discovery of new circumstances so that stochastic calculations ought to be made repeatedly. Thus, in the mid-19<sup>th</sup> century, amputation of a limb made under the newly introduced anaesthesia sometimes led to death from bronchitis (Sheynin 1982, p. 262) and the benefits of that procedure had to be critically considered. Circumstances and calculations should not be contrasted.

Bernoulli paid due attention to Leibniz’ criticism; more than a half of Chapter 4 of the AC in essence coincided with the respective passages from his letters to Leibniz (whom he did not mention by name).

In 1714, in a letter to one of his correspondents, Leibniz (Kohli 1975, p. 512) softened his doubts about the application of statistical probabilities and for some reason added that the late Jacob Bernoulli had “cultivated” the [theory of probability] in accordance with his, Leibniz’ “exhortations”.

On the correspondence between the two scholars see also Sylla (1998).

#### **2.4. The Law of Large Numbers**

*2.4.1. The Prehistory.* The LLN has its prehistory. It was thought, long before Bernoulli, that the number of successes in  $n$  “Bernoulli” trials with probability  $p$  was approximately equal to

$$\mu = np. \tag{1}$$

Cardano (Ore 1963, pp. 152 – 154 and 196), for example, applied this formula in calculations connected with games of dice. When compiling his mortality table, Halley (1694) assumed that “irregularities” in his data would have disappeared had he much more observations at his disposal. His idea can be interpreted as a statement on the increase in precision of formula (3), see below, with  $n$ ; it is likely, however, that these irregularities were occasioned by systematic corruptions.

A second approach to the LLN took shape in astronomy not later than during Kepler’s lifetime when the arithmetic mean became the universal estimator of the constant sought.

Similar but less justified statements concerning sums of magnitudes corrupted by random errors had also appeared. Thus, Kepler (Sheynin 1973, p. 120) remarked that the total weight of a large number of metal money of the same coinage did not depend on the inaccuracy in the weight of the separate coins (he should have mentioned the mean weight of a coin). Then, De Witt (Sheynin 1977, p. 214) stated that the then existing custom of buying annuities upon many ( $n$ ) young and apparently healthy lives secured profit “without hazard or risk”. The expectation of a gain  $Ex_i$  from each such transaction was obviously positive; if constant, the buyer could expect a total gain of  $nEx$ . There

also apparently existed a practice of an indirect participation of (petty?) punters in many games at once. In any case (Sheynin 1977, p. 236), both De Moivre and Montmort mentioned in passing that some persons bet on the outcomes of games. The LLN has then been known, but not to such punters, and that practice could have existed from much earlier times.

2.4.2. *Jakob Bernoulli*. Before going on to prove his LLN, Bernoulli (Chapter 4) explained that the theoretical “number of cases” was often unknown, but what was impossible to obtain beforehand, might at least be determined afterwards, i.e., by numerous observations. In essence, Bernoulli proved a proposition that, beginning with Poisson, is being called the LLN.

Let  $r$  and  $s$  be natural numbers,  $t = r + s$ ,  $n$ , a large natural number,  $v = nt$ , the number of [independent] trials (De Moivre (1712) was the first to mention independence) in each of which the studied event occurs with [probability]  $r/t$ ,  $\mu$  – the number of the occurrences of the event (of the successes). Then Bernoulli proved without applying mathematical analysis that

$$P\left(\left|\frac{\mu}{v} - \frac{r}{t}\right| \leq \frac{1}{1+c}\right) \geq 1 - \frac{1}{1+c} \quad (2)$$

and estimated the value of  $v$  necessary for achieving a given  $c > 0$ . In a weaker form Bernoulli’s finding meant that

$$\lim_{v \rightarrow \infty} P\left(\left|\frac{\mu}{v} - \frac{r}{t}\right| < \varepsilon\right) = 1 \quad (3)$$

where, as in formula (1),  $r/t$  was the theoretical, and  $\mu/v$ , the statistical probability.

Markov (1924, pp. 44 – 52) improved Bernoulli’s estimate mainly by specifying his intermediate inequalities, and Pearson (1925), by applying the Stirling formula, achieved a practically complete coincidence of the Bernoulli result with the estimate that makes use of the normal distribution as the limiting case of the binomial law; Markov did not use that formula apparently because Bernoulli had not known it, but then, on p. 55ff, he applied it without any connection with his previous reasoning.

In addition, Pearson (p. 202) considered Bernoulli’s estimate of the necessary number of trials in formula (2) “crude” and leading to the ruin of those who would apply it but had not found a single word appreciating the result achieved. On the contrary, he inadmissibly compared the Bernoulli law with the wrong Ptolemaic system of the world.

The very fact described by formulas (2) and (3) was, however, extremely important for the development of probability and statistics, and, anyway, should we deny the importance of existence theorems? For modern descriptions of Bernoulli’s LLN see Prokhorov (Bernoulli 1986) and Hald (1990, Chapter 16; 2003).

And so, the LLN established a correspondence between the two

probabilities. Bernoulli (Chapter 4) had indeed attempted to ascertain whether or not the statistical probability had its “asymptote”– whether there existed such a degree of certainty, which observations, no matter how numerous, would never be able to reach. Or, in my own words, whether there existed such positive numbers  $\beta$  and  $\delta < 1$ , that

$$\lim P\left(\left|\frac{\mu}{v} - \frac{r}{t}\right| < \beta\right) = 1 - \delta, v \rightarrow \infty.$$

He answered his question in the negative: no, such numbers did not exist and he thus established, within the boundaries of stochastic knowledge, a relation between deductive and inductive methods and combined statistics with the art of conjecturing.

Throughout Part 4, Bernoulli considered the derivation of the statistical probability of an event given its theoretical probability and this most clearly emerges in the formulation of his *Main Proposition* in Chapter 5. However, both in the last lines of that chapter and in Chapter 4 he mentioned the inverse problem actually alleging that he had solved it as well. I return to this point in § 2.4.3.

*2.4.3. Remarks on Later Events.* De Moivre (1756, p. 251) followed Bernoulli. Without any trace of hesitation, he claimed to have solved both the direct and the “converse” problems; he had expressed less clearly the same idea in 1738, in the previous edition of his book. De Moivre’s mistake largely exonerates Bernoulli, so that Keynes (1921, p. 402) wrongfully stressed that the latter “proves the direct theorem only”. It was Bayes who perceived that the two problems were different. He was the first to determine precisely the theoretical probability given the appropriate statistical data and for this reason I (Sheynin 2003) suggested that Bayes had completed the construction of the first version of probability theory. This, however, does not diminish the great merit of Bernoulli in spite of the much more precise results of De Moivre (for one of the problems) and Bayes.

I do not discuss Nicolaus Bernoulli’s version of the LLN, which he described in one of his letters of 1713 to Montmort (1713, pp. 280 – 285); see Youshkevich (1986) and Hald (1990, § 17.3; 2003). I myself (Sheynin 1970, p. 232; lacking in the original publication of 1968) noted that N.B. was the first to introduce, although indirectly, the normal distribution.

*2.4.4. Alleged Difficulties in Application.* Strangely enough, for a long time statisticians had not recognized the fundamental importance of the LLN. Haushofer (1872, pp. 107 – 108) declared that statistics, since it was based on induction [only partly], had no “intrinsic connections” with mathematics which is based on deduction [consequently, neither with probability]. A most noted German statistician, Knapp (1872, pp. 116 – 117), expressed a strange idea: the LLN was hardly useful since statisticians always made only one observation, as when counting the inhabitants of a city. And even later, Maciejewski (1911, p. 96) introduced a “statistical law of large numbers” in place of the Bernoulli proposition that had allegedly impeded the development of statistics. His own law qualitatively

asserted that statistical indicators exhibited ever lesser fluctuations as the number of observations increased.

All such statements definitely concerned the Poisson law as well (European statisticians then hardly knew about the Chebyshev form of the law) and Maciejewski's opinion likely represented the prevailing attitude of statisticians. Here, indeed, is what Bortkiewicz (1917, pp. 56 – 57) thought: the expression *law of large numbers* ought to be used only for denoting a “quite general” fact, unconnected with any definite stochastic pattern, of a higher or lower degree of stability of statistical indicators under constant or slightly changing conditions and given a large number of trials. Even Romanovsky (1961, p. 127) kept to a similar view and stressed the natural-scientific essence of the law and called it physical.

The text of Part 4 of the Art of Conjecturing follows.

**The Art of Conjecturing, Part 4 showing The Use and Application of the Previous Doctrine to Civil, Moral and Economic Affairs**

**Chapter 1. Some Preliminary Remarks**

**about Certainty, Probability, Necessity and Fortuity of Things**

Certainty of some thing is considered either objectively and in itself and means none other than its real existence at present or in the future; or subjectively, depending on us, and consists in the measure of our knowledge of this existence. Everything that exists or originates under the sun, – the past, the present, or the future, – always has in itself and objectively the highest extent of certainty. This is clear with regard to events of the present or the past; because, just by their existence or past existence, they cannot be non-existing or not having existed previously. Neither can you have doubts about [the events of] the future, which, likewise, on the strength of Divine foresight or predetermination, if not in accord with some inevitable necessity, cannot fail to occur in the future. Because, if that, which is destined to happen, is not certain to occur, it becomes impossible to understand how can the praise of the omniscience and omnipotence of the greatest Creator remain steadfast. But how can this certainty of the future be coordinated with fortuity or freedom [independence] of secondary causes? Let others argue about it; we, however, will not touch something alien to our aims.

Certainty of things, considered with respect to us, is not the same for all things, but varies diversely and occurs now greater, now lesser. Something, about which we know, either by revelation, intellect, perception, by experience, autopsia [direct observation; by one's own eyes] or otherwise, that we cannot in any way doubt its existence or realization in the future, has the complete and absolute certainty. To anything else our mind assigns a less perfect measure [of certainty], either higher or lower depending on whether there are more or less probabilities convincing us of its existence at present, in the past or the future.

As to *probability*, this is the degree of certainty, and it differs from the latter as a part from the whole. Namely, if the integral and absolute certainty, which we designate by letter  $\alpha$  or by unity 1, will be thought to consist, for example, of five probabilities, as though of five parts, three of which favour the existence or realization of some event, with

the other ones, however, being against it, we will say that this event has  $3/5\alpha$ , or  $3/5$ , of certainty.

Therefore, the event that has a greater part of certainty than the other ones is called *more probable*, although actually, according to the usual word usage, we only call probable that, whose probability markedly exceeds a half of certainty. I say *markedly* because a thing, whose probability is roughly equal to a half of certainty, is called *doubtful* or indefinite<sup>1.1</sup>. Thus, a thing having  $1/5$  of certainty is more probable than that which has  $1/10$ , although actually neither is probable.

*Possible* is that which has at least a low degree of certainty whereas the impossible has either no, or an infinitely small certainty. Thus, something is possible if it has  $1/20$  or  $1/30$  of certainty.

*Morally certain* is that whose probability is almost equal to complete certainty so that the difference is insensible. On the contrary, *morally impossible* is that which has only as much probability as the morally certain lacks for becoming totally certain. Thus, if morally certain is that which has  $999/1000$  of certainty, then something only having  $1/1000$  of certainty will be morally impossible.

*Necessary* is that, which cannot fail to exist at present, in the future or past, owing exactly to necessity, either *physical* (thus, fire will necessarily consume; a triangle will have three angles summing up to two right angles; a full moon, if in a node, will necessarily be accompanied by a [lunar] eclipse), – or *hypothetical*, according to which all that exists, or had existed, or is supposed to exist, cannot fail to exist (in this sense it is necessary that Petrus, about whom I know and accept that he is writing, is indeed writing), – or, finally, according to the necessity of a condition or agreement (thus, a gambler scoring a six with a die is necessarily reckoned the winner if the gamblers have agreed that winning is connected with throwing a six).

*Contingent* (both *free*, if it depends on the free will of a reasonable creature, and *fortuitous* and *casual*, if it depends on fortune or chance) is that which can either exist or not exist at present, in the past or future, – clearly because of remote rather than immediate forces. Indeed, neither does contingency always exclude necessity up to secondary causes. I shall explain this by illustrations.

It is absolutely doubtless that, given a certain position of a die, [its] velocity and distance from the board at the moment when it leaves the thrower's hand, it cannot fall otherwise than it actually does. Just the same, under a certain present composition of the air, and given the masses, positions, motions, directions, and velocities of the winds, vapours and clouds, as well as the mechanical laws governing the interactions of all that, the weather tomorrow cannot be different from that which it will actually be. So these phenomena take place owing to their immediate causes with no lesser necessity than the phenomena of the eclipses follow from the movement of the heavenly bodies. And still, usually only the eclipses are ranked among necessary phenomena whereas the fall of a die and the future weather are thought to be contingent.

The sole reason for this is that what is supposed to be known for determining future actions, and what indeed is such in nature, is not enough known. And, even had it been sufficiently known, geometric and physical knowledge is inadequately developed for subjecting such phenomena to calculation in the same way as eclipses can be calculated beforehand and predicted by means of known astronomical principles. And, for the same reason, before astronomy achieved such perfection, the eclipses themselves had to be reckoned as future chance events to not a lesser extent than the two other [mentioned] phenomena.

It follows that what seems to be contingent to one person at a certain moment, will be thought necessary to someone else (or even to the same person) at another time after the [appropriate] causes become known. And so, contingency mainly depends on our knowledge since we do not see any contradiction with the non-existence of the event at present or in the future, although here and now, owing to an immediate but unknown to us cause it is either necessarily realized, or ought to occur.

Not everything which brings us well-being or harm is called *happiness* or *misfortune* {*Fortuna prospera*, un *Bonheur*, ein Glück & *Fortuna adversa*, un *Malheur*, ein Unglück}, but only that which with a higher, or at least with the same probability would have possibly failed to occur. Therefore, happiness or misfortune are the greater, the lower was the probability of the well-being or harm that has actually occurred. Thus, exceptionally happy is the man who finds a buried treasure while digging the ground because this does not happen even once in a thousand cases. If twenty deserters, one of whom will be put to death by hanging as an example for the others, cast lots as to who of them remains living, those nineteen who drew the more favourable lot are not really called happy; but the twentieth who cast the horrible lot is most miserable. [In the same way,] your friend who came out unharmed from a battle in which [only] a small part of the combatants were killed should not be called happy, unless you will perhaps think it necessary to do so because of the special fortune of preserving life.

## **Chapter 2. On Arguments and Conjecture.**

### **On the Art of Conjecturing.**

#### **On the Grounds for Conjecturing.**

##### **Some General Pertinent Axioms**

Regarding that which is certainly known and beyond doubt, we say that we *know or understand* [it]; concerning all the rest, – we only *conjecture or opine*.

To make *conjectures* about something is the same as to measure its probability. Therefore, the art of *conjecturing or stochastics* {*ars conjectandi sive stochastice*}<sup>2,1</sup> is defined as the art of measuring the probability of things as exactly as possible, to be able always to choose what will be found the best, the more satisfactory, serene and reasonable for our judgements and actions. This alone supports all the wisdom of the philosopher and the prudence of the politician.

Probabilities are estimated both by the *number* and the *weight of the*

*arguments* which somehow prove or indicate that a certain thing is, was, or will be. As to *the weight*, I understand it to be the force of the proof.

*Arguments* themselves are either *intrinsic*, in every-day speech artificial, elicited in accordance with considerations of the cause, the effect, of the person, connection, indication or of other circumstances which seem to have some relation to the thing under proof; or *external* and not artificial, derived from people's authority and testimony. An example: Titius is found killed in the street. Maevius is charged with murder. The accusing arguments are: 1) He is known to have hated Titius (an argument from a *cause*, since this very hate could have incited to murder). 2) When questioned, he turned pale and answered timidly (this is an argument from the *effect* since it is possible that the pallor and fright were caused by his being conscious of the evil deed perpetrated). 3) Blood-stained cold steel is found in Maevius' house (this is an *indication*). 4) The same day that Titius was killed, Maevius had been walking the same road (this is *circumstance* of place and time). 5) Finally, Cajus maintains that the day before Titius was killed, he had quarrelled with Maevius (this is a *testimony*).

However, before getting down to our problem, – to indicating how should we apply these arguments for conjecturing to measure probabilities, – it is helpful to put forth some general rules or axioms which are dictated to any sensible man by usual common sense and which the more reasonable men always observe in everyday life.

1) *In such things in which it is possible to achieve complete certainty, there is no place for conjectures.* Futile would have been an astronomer, who, knowing that two or three [lunar] eclipses occur yearly, desires to forecast, on such grounds, whether or not there will be an eclipse during a full moon. Indeed, he could have found out the truth by reliable calculation. Just the same, if a thief says at his questioning that he sold the stolen thing to Sempronius, the judge who wants to conjecture about the probability of that statement by looking at the expression of the thief's face and listening to the tone of his voice, or by contemplating the quality of the stolen thing, or by some other circumstances, will act stupidly, because Sempronius, from whom everything can certainly and easily be elicited, is available.

2) *It is not sufficient to weigh one or another argument; it is necessary to investigate all such which can be brought to our knowledge and will seem suitable in some respect for proving the thing.* Suppose that three ships leave the harbour. After some time it is reported that one of them had suffered shipwreck and is lost. Conjectures are made: which of them? If only paying attention to the number of the ships, I shall conclude that each of them could have met with the misfortune in an equal manner. But since I remember that one of them was comparatively old and decrepit, badly rigged with masts and sails, and steered by a young and inexperienced helmsman, I believe that, in all probability, it was this ship that got lost rather than one of the others.

3) *We ought to consider not only the arguments which prove a thing, but also all those which can lead to a contrary conclusion, so that, after duly discussing the former and the latter, it will become*

*clear which of them has more weight.* It is asked, with respect to a friend very long absent from his fatherland, may we declare him dead<sup>2,2</sup>? The following arguments favour an answer in the affirmative: During the entire twenty years, in spite of all efforts, we have been unable to find out anything about him; the lives of travellers are exposed to very many dangers from which those remaining at home are exempted; therefore, perhaps his life came to an end in the waves; perhaps he was killed en route or in battle; perhaps he died of an illness or from some [other] cause in a place where no one knew him. Then, has he been living, he would have reached an age which only a few attain even in their homeland; and he would have written even from the furthest shores of India because he knew that an inheritance was expected for him at home. And so on in the same vein.

Nevertheless, we should not rest content with these arguments but rather oppose them by the following supporting the contrary. He is known to have been thoughtless; wrote letters reluctantly; did not value friends. Perhaps Barbarians held him captive so that he was unable to write, or perhaps he did write sometimes from India, but the letters got lost either because of the carelessness of those carrying them, or during shipwrecks. And, to cap it all, many people are known to have returned unharmed after having been absent even longer.

4) *For judging about universalities remote and universal arguments are sufficient; however, for forming conjectures about particular things, we ought also to join to them more close and special arguments if only these are available.* Thus, if it is asked, in general, how much more probable is it for a twenty-year-old youth to outlive an aged man of 60 rather than the other way round, we have nothing to take into consideration other than the distinction between the generations and ages. But if the question concerns two definite persons, the youth Petrus and the old man Paulus, we also ought to pay attention to their complexion, and to the care that each of them takes over his health. Because if Petrus is in poor health, indulges in passion, and lives intemperately, Paulus, although much older, may still hope, with every reason, to live longer.

5) *Under uncertain and dubious circumstances we ought to suspend our actions until more light is thrown. If, however, the necessity of action brooks no delay, we must always choose among two possibilities that one which seems more suitable, safe, reasonable, or at least more probable<sup>2,3</sup>, even if none of them is actually such.* Thus, if a fire has broken out and you can only save yourself by jumping from the top of the roof or from some lower floor, it is better to choose the latter as being less dangerous, although neither alternative is quite safe or free from the danger of injury.

6) *That which is in some cases helpful and never harmful ought to be preferred to that which is never either helpful or harmful. In our vernacular it is said Hilfft es nicht, so schadt es nicht* [Even if it does not help, it does not harm]. This proposition follows from the previous [considerations], because that which can be helpful is more satisfactory, reliable and desirable than that which under the same conditions cannot [be helpful].

7) *Human actions should not be assigned a value according to their outcomes* because sometimes the most reckless actions are accompanied by the best success, whereas, on the contrary, the most reasonable [may] lead to the worst results. In agreement with this, the Poet says: “May success be wanting, I wish, for him who would judge facts by their outcomes” [Ovidius, *Epistulae Heroidum* II, “Phyllis Demophoonti”, line 85]. Thus, someone who intends to throw at once three sixes with three dice, should be considered reckless even if winning by chance. On the contrary, we [ought to] note the false judgement of the crowd which considers a man the more prominent, the more fortunate he is, and for which even a successful and fruitful crime is mostly a virtue. Once more Owen (*Epigr[ammatum] lib[er] sing[ularis]*, 1607], § 216)<sup>2,4</sup> gracefully says:

*Although just now Ancus is believed to be a fool, it is argued that he is wise because the poorly conceived turned out successful [for him].* If something reasonably thought-out fails, even Cato will be judged a fool by the crowd.

8) *In our judgements, we ought to beware of attributing to things more than is due to them, ought not to consider something which is only more probable than the other as absolutely certain, nor to impose the same opinion on others.* [This is] because the trust attributed to things ought to be in a proper proportion to the degree of certainty possessed by each thing, and be less in the same ratio as its probability itself is. In vernacular, this is expressed as

*Man muss ein jedes in seinem Werth und Unwerth beruhen lassen* [Let each thing be determined by its value or worthlessness.]

9) *However, since complete certitude can only seldom be attained, necessity and custom desire that that, which is only morally certain, be considered as absolutely certain.* Therefore, it would be helpful if the authorities determine certain boundaries for moral certainty, – if, for example, it would be defined whether 99/100 of certainty be sufficient for resolving something, or whether 999/1000 be needed, so that a judge, unable to show preference to either side, will always have firm indications to conform with when pronouncing a sentence.

Anyone having knowledge of life can compile many more similar axioms, but, lacking an appropriate occasion, we can hardly remember all of them.

### **Chapter 3. On Arguments of Different Kinds and on How Their Weights Are Estimated for Calculating the Probabilities of Things**

He who considers various arguments by which our opinions and conjectures are formed will note a threefold distinction between them since *some of them necessarily exist and contingently provide evidence; others exist contingently and necessarily provide evidence; finally, the third ones both exist and provide evidence contingently.*

I explain these differences by examples. For a long time, my brother does not write me anything. I doubt whether to blame his laziness or his business pursuits, and fear that he may even have died. Here, there are threefold arguments for explaining the ceasing of the correspondence: *laziness, death, pursuits*. The first of these exists for sure (according to hypothetical necessity, since I know and accept that

my brother is lazy), but proves true [provides evidence] only contingently because laziness possibly would not have hindered him from writing. The second one contingently exists (because my brother could still be alive), but proves true without question because a dead man cannot write. The third one both exists and provides evidence contingently because my brother can have business pursuits or not, and if he has them, they need not be such that prevent him from writing.

Another example. I suppose that, according to the conditions of a game, a gambler wins if he throws seven points with two dice, and I wish to guess his hope of winning. Here, the argument for winning is the throwing of seven points. It necessarily indicates the winning (owing, indeed, to the agreement between the gamblers), but it only exists contingently, because, in addition to the seven points, another number of them can occur.

Excepting this difference between the arguments, another distinction can also be noted since some of them are pure, the other ones, *mixed*. I call an argument *pure* if in some cases it proves a thing in such a manner that on other occasions it does not prove anything positively. A *mixed* argument, however is such that in certain cases it thus proves a thing that on other occasions it proves the contrary in the same manner.

An example. Someone in a quarrelling crowd was cut with a sword; and, as trustworthy people who saw the incident from a distance testify, the perpetrator was dressed in a black cloak<sup>3.1</sup>. If Gracchus was among those quarrelling together with three others, all of them in black tunics, this tunic will be an argument in favour of Gracchus having committed the murder.

However, this argument will be mixed since in one case it proves his guilt, and, in three other cases, it demonstrates his innocence. Indeed, the murder was perpetrated either by him, or by one of the other three, with the latter instance being impossible without exonerating Gracchus. If, however, during the subsequent questioning Gracchus turned pale, the paleness of his face will be a pure argument because it demonstrates his guilt if occasioned by disturbed conscience. On the contrary, it would not prove his innocence had it been called forth by something else, since it is possible that he turned pale owing to another cause but still was himself the perpetrator of the murder [the murderer].

The above makes it clear that the force of proof peculiar to some argument depends on the multitude of cases in which it can exist or not exist, provide evidence or not or even provide evidence to the opposite of the thing. Therefore, the degree of certainty, or the probability engendered by this argument, can be deduced by considering these cases in accordance with the doctrine given in Part 1 [of this book] in exactly the same way as the fate of gamblers in games of chance is usually investigated.

And so, first, let an argument exist contingently and provide evidence necessarily. If some argument *both exists and indicates contingently*, ...

Then, if several arguments are collected for proving one and the same thing, the force provided by the totality of all the arguments is estimated in the following way ...

If, in addition to the arguments leading to the proof of a thing, there exist other pure arguments favouring the opposite, the arguments of both kinds ought to be weighed separately ...

It might happen that something has  $2/3$  of certainty whereas its opposite has  $3/4$  so that each of these contraries will be probable although the first of them is less probable than the opposite; namely, their ratio will be as  $2/3$  to  $3/4$  or as 8 to 9.

I cannot conceal here that I foresee many obstacles in special applications of these rules that can often lead to shameful mistakes if caution is not observed when distinguishing between the arguments. Indeed, sometimes such arguments can seem to differ which actually compose one and the same argument, and to the contrary: differing arguments can be accepted as a single argument. Sometimes an argument includes such premises which absolutely refute the opposite, etc. As an explanation, I only adduce one or two illustrations. In the example above concerning Gracchus, I assume that the trustworthy people who saw those quarrelling also noted that the perpetrator was red-haired and that Gracchus together with two of the others were distinguished by hair of that colour, but that no one of the latter was dressed in a black toga. In that case, if someone would have desired to compare the probabilities of Gracchus' guilt and innocence by the indications that Gracchus and three others were dressed in black, and also, that, again in addition to him, two others were notable for their red hair, and found that they are in a composite ratio of 1:3 and 1:2, or in the ratio of 1 to 6; and if he were to conclude that Gracchus is by far more likely to be innocent than to be the perpetrator of the murder, he would certainly have collated the matter in a most inept fashion. Actually, there are no two arguments here but only one and the same, resulting from two simultaneous circumstances, the colour of the dress and of the hair. Since both these circumstances are only conjoined in the case of Gracchus, they certainly demonstrate that no one else except him could have been the perpetrator.

Another example. It becomes doubtful whether a written document is fraudulently antedated. An argument to the opposite could be that the document was signed by the hand of a notary public, i.e., by an official and sworn person, with regard to whom it is unlikely that he might have permitted himself any fraud. Indeed, he would have been unable to do so without greatly endangering his honour and well-being; in addition, even from among 50 notaries hardly one would have dared to commit such a vile action. The following arguments could be in favour of an answer in the affirmative: This notary is very ill-famed; and could have expected greatest benefits from the fraud; and especially that he had testified to something having no probability, as for example that someone had lent 10,000 gold coins to another person, whereas, according to everyone's estimation, all his property then barely amounted to 100.

Here, if considering separately the argument from the character of the signatory, the probability that the document is authentic may be

valued as 49/50 of certainty. When, however, weighing the arguments favouring the opposite, it would be necessary to conclude that it is hardly possible that the document is not forged so that the fraud committed in the document is of course morally certain, that is, has 999/1000 of certainty. However, we should not conclude that the probabilities of authenticity and fraud are in the ratio of 49/50 to 999/1000, or almost of equality. Because, if I believe that the notary is dishonourable, I am therefore assuming that he does not belong to the 49 honest notaries detesting deception but that he is indeed the fiftieth who has no scruples of fulfilling his duties faithfully. This consideration completely destroys all the power of that argument, which in other cases could have been able to prove that a document is authentic.

#### **Chapter 4. On a Two-Fold Method of Investigating the Number of Cases.**

##### **What Ought To Be Thought about Something Established by Experience. A Special Problem Proposed in This Case, etc.**

It was shown in the previous chapter how, – given the number of cases in which arguments in favour of some thing can exist or fail to exist, can provide evidence or not, or even prove the opposite, – the force of what they prove, and the probabilities of things proportional to these forces, can be derived and estimated by calculation. We thus see that for correctly conjecturing about some thing, nothing else is required than both precisely calculating the number of cases and finding out how much more easily can some of them occur than the others. Here, however, we apparently meet with an obstacle since this only extremely seldom succeeds, and hardly ever anywhere except in games of chance which their first inventors, desiring to make them fair, took pains to establish in such a way that the number of cases involving winning or losing were determined with certainty and known and the cases themselves occurred with the same facility.

However, for most of other matters, depending either on the production of nature or the free will of people, this does not take place at all. Thus, for example, the number of cases is known in [a game of] dice. For each die there are manifestly as many cases as faces, and all of them are equally inclined [to turn up], since, owing to the similitude [congruence] of the faces and the uniform weight [density] of the die, there is no reason for one of them to turn up more easily than another<sup>4.1</sup>.

This would have happened if the forms of the faces were dissimilar or if one part of the die consisted of a heavier substance than the other one. In the same way, the number of cases for drawing a white or a black ticket from an urn is known, and known [also] is that [the drawings of] all of them are equally possible. Indeed, the number of tickets of both these kinds is evidently determined and known, and no reason is seen for one of them to appear more easily than any other.

But, who from among the mortals will be able to determine, for example, the number of diseases, that is, the same number of cases which at each age invade the innumerable parts of the human body and can bring about our death; and how much easier one disease (for example, the plague) can kill a man than another one (for example,

dropsy or, dropsy than fever), so that we would be able to conjecture about the future state of life or death? And who will count the innumerable cases of changes to which the air is subjected each day to form a conjecture about its state in a month, to say nothing about a year? Again, who knows the nature of the human mind or the admirable fabric of our body shrewdly enough for daring to determine the cases in which one or another participant can gain victory or be ruined in games completely or partly depending on acumen or agility of body?

Since this and the like depends on absolutely hidden causes, and, in addition owing to the innumerable variety of their combinations always escapes our diligence, it would be an obvious folly to wish to find something out in this manner. Here, however, another way for attaining the desired is really opening for us. And, what we are not given to derive a priori, we at least can obtain a posteriori, that is, can extract it from a repeated observation of the results of similar examples. Because it should be assumed that each phenomenon can occur and not occur in the same number of cases in which, under similar circumstances, it was previously observed to happen and not to happen. If, for example, it was formerly noted that, among the observed three hundred men of the same age and complexion as Titius now is and has, two hundred died after ten years with the others still remaining alive, we may conclude with sufficient confidence that Titius also has twice as many cases for paying his debt to nature during the next ten years than for crossing this border. Again, if someone will consider the atmosphere for many previous years and note how many times it was fine or rainy; or, will be very often present at a game of two participants and observe how many times either was the winner, he will thus discover the ratio of the number of cases in which the same event will probably happen or not also in the future under circumstances similar to those previously existing.

This empirical method of determining the number of cases by experiment is not new or unusual. Because the celebrated author of *L'art de penser*, a man of great intellect and acumen<sup>4.2</sup>, prescribes the like in Chapter 12 and in the next ones of the last part [of that book], and the same is also constantly observed in everyday life. Then, neither can anyone fail to note also that it is not enough to take one or another observation for such reasoning about an event, but that a large number of them are needed. Because, even the most stupid person, all by himself and without any preliminary instruction, is guided by some natural instinct (which is extremely miraculous) and feels sure that the more such observations are taken into account, the less is the danger of straying from the goal.

Although this is known by nature to everyone, its proof, derived from scientific principles, is not at all usual and we ought therefore to expound it here. However, I would have estimated it as a small merit had I only proved that of which no one is ignorant. Namely, it remains to investigate something that no one had perhaps until now run across even in his thoughts. It certainly remains to inquire whether, when the number of observations thus increases, the probability of attaining the real ratio between the number of cases, in which some event can occur

or not, continually augments so that it finally exceeds any given degree of certitude. Or [to the contrary], the problem has, so to say, an asymptote; i. e., that there exists such a degree of certainty which can never be exceeded no matter how the observations be multiplied, so that, for example, it is never possible to obtain more than a half, or than  $2/3$ , or  $3/4$  of certainty in deriving the real ratio of cases.

To make clear my desire by illustration, I suppose that without your knowledge three thousand white pebbles<sup>4.3</sup> and two thousand black ones are hidden in an urn, and that, to determine [the ratio of] their numbers by experiment, you draw one pebble after another (but each time return the drawn pebble before extracting the next one so that their number in the urn will not decrease), and note how many times a white pebble is drawn, and how many times a black one. It is required to know whether you are able to do it so many times that it will become ten, a hundred, a thousand, etc., times more probable (i. e., become at last morally certain) that the number of the white and the black pebbles which you extracted will be in the same ratio, of 3 to 2, as the number of pebbles themselves, or cases, than in any other different ratio. To tell the truth, if this failed to happen, it would be necessary to question our attempt at experimentally determining the number of cases. If, however, this is attained and we thus finally obtain moral certainty (in the next chapter I shall show that this is indeed so), then we determine the number of cases a posteriori almost as though it was known to us a priori. In social life, where the morally certain, according to Proposition 9 of Chapter 2, is assumed as quite certain, this is undoubtedly quite sufficient for scientifically directing our conjectures about any contingent thing in a no lesser way than in games of chance. Because, if we replace an urn for example by air or by a human body, which contain in themselves sources of various changes or diseases just as the urn contains pebbles, we will be able to determine by observation in exactly the same way how much easier can one or another event occur in these things.

To avoid false understanding, it ought to be noted that the ratio between the numbers of cases which we desire to determine experimentally is accepted not as precise and strict (because this point of view would have led to a contrary result and the probability of determining the real ratio would have been the lower the more observations we would have taken)<sup>4.4</sup>, but that this ratio be accepted with a certain latitude, that is, contained between two limits [boundaries] which can be taken as close as you like. Indeed, if in the example just provided concerning pebbles, we will assume two ratios,  $301/200$  and  $299/200$ , or  $3001/2000$  and  $2999/2000$ , etc., one of which is very near but greater, and the other one very near but smaller than  $3/2$ , it will be shown that, setting any probability, it can be made more probable that the ratio derived from many observations will be contained within these limits of  $3/2$  rather than outside.

This, then, is the problem that I decided to make here public after having known its solution for twenty years. Its novelty and the greatest utility joined with an equal difficulty can attach more weight and value to all the other chapters of this doctrine [of the *ars conjectandi*]. However, before exposing its solution I shall defend

myself in a few words from the objections to these propositions levelled by some scholars.

1. First, it was objected that the ratio of pebbles is one thing, whereas the ratio of diseases or changes in the air is something else. The number of the former is definite but the number of the latter is indefinite and vague. I answer this by saying that they both, in comparison to our knowledge, are equally indefinite and vague. However, we can imagine anything that is such in itself and in accordance with its nature, not better than a thing created and at the same time not created by the Author of nature because everything done by God is determined thereby.

2. Second, it is objected that the number of pebbles is finite and that of diseases etc. is infinite. Answer. Rather immense than infinite. But let us assume that it is indeed infinite. Even between two infinities a definite ratio is known to be possible and to be expressed by finite numbers either precisely or at least with any desired approximation. Thus, the ratio of each circumference to [its] diameter is definite. [True,] it is not precisely expressed otherwise than by an infinitely continued Ludolphus' cyclic number. However, Archimedes, Metius and Ludolphus himself<sup>4,5</sup> restricted that ratio within limit [boundaries] sufficiently close to each other for practice. Therefore, nothing hinders a ratio of two infinities approximately expressed by finite numbers to be determined by a finite number of experiments either.

3. Third, it is said that the number of diseases does not remain constant but that new diseases occur every day. Answer. We are unable to deny that diseases can multiply in the course of time; and he who desires to conclude from present-day observations about the times of our antediluvian forefathers will undoubtedly deviate enormously from the truth. But nothing follows from this except that sometimes we ought to resume observations just as it would be necessary to resume observations with the pebbles if it is assumed that their number in the urn is variable.

### Chapter 5. Solution of the Previous Problem

To explicate the long demonstration as briefly and clearly as possible, I will attempt to reduce everything to abstract mathematics, eliciting from it the following lemmas after which all the rest will only consist in their mere application.

**Lemma 1.** Suppose that a series of any quantity of numbers 0, 1, 2, 3, 4, etc., follow, beginning with zero, in the natural order and let the extreme and maximal of them be  $r + s$ , some intermediate, be  $r$ , and the two nearest to it on either side,  $r + 1$  and  $r - 1$ . If this series be continued until its extreme term becomes equal to some multiple of the number  $r + s$ , that is, until it is equal to  $nr + ns$ , the intermediate number  $r$  and those neighbouring it,  $r + 1$  and  $r - 1$ , will be augmented in the same ratio, so that  $nr$ ,  $nr + n$  and  $nr - n$  will appear instead, and the series itself

$$0, 1, 2, 3, 4, \dots, r - 1, r, r + 1, \dots, r + s$$

will change and become

0, 1, 2, 3, 4, ...,  $nr - n$ , ...,  $nr$ , ...,  $nr + n$ , ...,  $nr + ns$ .

With an increasing  $n$  both the number of the terms situated between the intermediate  $nr$  and one of the limiting terms,  $nr + n$  or  $nr - n$ , and the number of those terms which extend from these limits to the extreme terms  $nr + ns$  or 0 will thus increase. However (no matter how large will  $n$  be assumed), the number of terms following after the larger limit  $nr + n$  will never be more than  $s - 1$  times larger than, and the number of terms preceding the lesser limit  $nr - n$  will never be more than  $r - 1$  times larger than the number of them situated between the intermediate  $nr$  and one of the limits,  $nr + n$  or  $nr - n$ . Because, after subtraction, it is clear that between the greater limit and the extreme term  $nr + ns$  there are  $ns - n$  intermediate terms, and between the lesser limit and the other extreme term 0 there are  $nr - n$  intermediate terms, and  $n$  terms between the intermediate and each of the limits. However,  $(ns - n):n = (s - 1):1$  and  $(nr - n):n = (r - 1):1$ . It therefore follows, etc.

**Lemma 2.** A binomial  $r + s$  raised to any integral power is expressed by terms whose number exceeds by 1 the number of unities in the exponent.

Since a square [of a binomial] consists of three terms, a cube has 4, a fourth power has 5 terms, etc., as is known.

**Lemma 3.** For any power of this binomial (at least for an exponent equal to the binomial  $r + s = t$ , or to its multiple, for example, to  $nr + ns = nt$ ), a certain term  $M$  will be maximal if the number of terms preceding and following it are in the ratio of  $s$  to  $r$ ; or, which is the same, if the exponents of letters  $r$  and  $s$  in this term are in the ratio of the magnitudes  $r$  and  $s$  themselves. The term nearer to it from either side is larger than the more distant term on the same side; however, the same term  $M$  is in a lesser ratio to the nearer term than the nearer term to the more distant one if the numbers of intermediate terms are the same.

*Dem[onstration].* 1. Geometers know well enough that the binomial  $r + s$  raised to the power  $nt$ , that is,  $(r + s)^{nt}$ , is expressed by such a series:

$$r^m + \frac{nt}{1} r^{m-1} s + \frac{nt(nt-1)}{1 \cdot 2} r^{m-2} s^2 + \dots + \frac{nt}{1} r s^{m-1} + s^m.$$

[...] Since the number of all the terms except  $M$  is, according to Lemma 1,  $nt = nr + ns$ , and, as assumed, the numbers of the terms preceding and following  $M$  are as  $s$  to  $r$ , these numbers are  $ns$  and  $nr$  respectively. Therefore, in accordance with the law of the series, the term  $M$  will be

$$\frac{nt(nt-1)(nt-2)\dots(nr+1)}{1 \cdot 2 \cdot 3 \cdot \dots \cdot ns} r^{nr} s^{ns} \text{ or } \frac{nt(nt-1)(nt-2)\dots(ns+1)}{1 \cdot 2 \cdot 3 \cdot \dots \cdot nr} r^{nr} s^{ns}$$

call it (5.1), and in the same way the terms nearest to it on *the left* and *the right* are

$$\frac{nt(nt-1)(nt-2)\dots(nr+2)}{1\cdot 2\cdot 3\cdot \dots\cdot (ns-1)} r^{nr+1} s^{ns-1} \text{ and } \frac{nt(nt-1)(nt-2)\dots(ns+2)}{1\cdot 2\cdot 3\cdot \dots\cdot (nr-1)} r^{nr-1} s^{ns+1}$$

and in the same way the next ones on *the left* and *the right* are [...].

After a preliminary suitable cancellation of common multipliers from both the coefficients and the powers themselves, it becomes clear that the term  $M$  is to its nearest on the left as  $(nr+1)s$  to  $nrs$ ; this latter to the next one, as  $(nr+2)s$  to  $(ns-1)r$ , etc., and also that the term  $M$  is to its nearest on the right as  $(ns+1)r$  to  $nrs$ , this latter to the next one, as  $(ns+2)r$  to  $(nr-1)s$ , etc. But

$$(nr+1)s > nrs, \text{ and } (nr+2)s > nrs - r, \text{ etc.}$$

Also,

$$(ns+1)r > nrs \text{ and } (ns+2)r > nrs - s, \text{ etc.}$$

It follows that the term  $M$  is greater than either of the nearest terms on either side which [in turn] are greater than the more remote terms on the same side, etc. QED.

2. The ratio  $(nr+1)/ns$ , as is clear, is less than the ratio  $(nr+2)/(ns-1)$ . Therefore, after multiplying [them] by one and the same ratio  $s/r$ , the ratio

$$\frac{(nr+1)s}{nsr} < \frac{(nr+2)s}{(ns-1)r}.$$

Just the same, it is evident that the ratio

$$\frac{(ns+1)}{nr} < \frac{ns+2}{nr-1}.$$

Consequently, after multiplying [this inequality] by one and the same ratio  $r/s$ , also

$$\frac{(ns+1)r}{nrs} < \frac{(ns+2)r}{(nr-1)s}.$$

But the ratio

$$\frac{(nr+1)s}{nsr}$$

is equal to the ratio of the term  $M$  to its nearest term on the left and the ratio<sup>5.1</sup>

$$\frac{(nr+2)s}{(ns-1)r}$$

is the same as  $M$  has to the next one. And the ratio

$$\frac{(ns+1)r}{nrs}$$

is that of the term  $M$  to its nearest term on the right, and

$$\frac{(ns+2)r}{(nr-1)s}$$

is the ratio of that term to the next one. What was just shown may in the same way be also applied to all the other terms.

Therefore, the maximal term  $M$  is in a lesser ratio to the nearer term on either side than (if the intervals between the terms are the same) the nearer term is to the more distant one on the same side. QED.

**Lemma 4.** The number  $n$  in a binomial raised to the power  $nt$  can be taken so great that the ratio of the maximal term  $M$  to [any of the] two others,  $L$  and  $\Lambda$  distant from it by  $n$  terms on the left and on the right [respectively], would be greater than any given ratio.

*Dem[onstration].* Since in the previous Lemma the maximal term  $M$  was found to be equal to (5.1) the terms on the left and on the right,  $L$  and  $\Lambda$ , in accordance with the law of the [formation of the] series (adding  $n$  to the last multiplier in the numerators of the coefficients, and subtracting  $n$  from the last multiplier in their denominators, adding the same  $n$  to the power of one of the letters  $r$  and  $s$ , and subtracting it from the power of the other letter), will be

$$\frac{nt(nt-1)(nt-2)\dots(nr+n+1)}{1\cdot 2\cdot 3\cdot \dots\cdot (ns-n)} r^{nr+n} s^{ns-n}$$

$$\text{and } \frac{nt(nt-1)(nt-2)\dots(ns+n+1)}{1\cdot 2\cdot 3\cdot \dots\cdot (nr-n)} r^{nr-n} s^{ns+n}.$$

And after a suitable cancellation of common multipliers,

$$\frac{M}{L} = \frac{(nr+n)(nr+n-1)\dots(nr+1)s^n}{(ns-n+1)(ns-n+2)\dots nsr^n},$$

$$\frac{M}{\Lambda} = \frac{(ns+n)(ns+n-1)\dots(ns+1)r^n}{(nr-n+1)(nr-n+2)\dots nrs^n},$$

or

$$\frac{M}{L} = \frac{(nrs+ns)(nrs+ns-s)\dots(nrs+s)}{(nrs-nr+r)(nrs-nr+2r)\dots nrs},$$

$$\frac{M}{\Lambda} = \frac{(nrs + nr)(nrs + nr - r) \dots (nrs + r)}{(nrs - ns + s)(nrs - ns + 2s) \dots nrs}.$$

However, when  $n$  is assumed infinite, these ratios will [also] be infinitely large, because then the numbers 1, 2, 3 etc. will vanish as compared with  $n$ , and the numbers themselves  $nr \pm n \mp 1$ ,  $nr \pm n \mp 2$ , etc., and  $ns \pm n \mp 1$ ,  $ns \pm n \mp 2$ , etc. will have the same value as  $nr \pm n$  and  $ns \pm n$  [respectively], so that, after dividing [both parts of both last fractions] by  $n$ ,

$$\frac{M}{L} = \frac{(rs + s)(rs + s) \dots rs}{(rs - r)(rs - r) \dots rs}, \quad \frac{M}{\Lambda} = \frac{(rs + r)(rs + r) \dots rs}{(rs - s)(rs - s) \dots rs}.$$

It is clear that these ratios are composed of as many ratios  $[(rs + s)/(rs - r)]$  or  $[(rs + r)/(rs - s)]$  as there are multipliers whose number is  $n$ , that is, infinite since the difference between the first multipliers  $nr + n$  or  $ns + n$ , and the last ones,  $nr + 1$  or  $ns + 1$ , is  $n - 1$ . These ratios  $[M/L$  and  $M/\Lambda]$  will therefore be equal to  $[(rs + s)/(rs - r)]$  or  $[(rs + r)/(rs - s)]$  raised to an infinite power and therefore simply infinite. If you doubt this conclusion, imagine infinity [of ratios] in a continued proportion with their ratio being as  $rs + s$  to  $rs - r$  or  $rs + r$  to  $rs - s$ . The first ratio will be to the third as the square; to the fourth, as a cube; to the fifth, as the fourth [power], etc. Finally, the first ratio will be to the last one as infinite powers of the ratio  $[(rs + s)/(rs - r)]$  or  $[(rs + r)/(rs - s)]$ . It is known, however, that the ratio of the first [ratio] to the last one is infinitely large since the last one = 0 (see Coroll. to Prop[osition] 6 of our [Tractatus de] Seriebus Infinitis [etc.]<sup>5.2</sup>). It is therefore clear that infinite powers of the ratio  $[(rs + s)/(rs - r)]$  or  $[(rs + r)/(rs - s)]$  are infinite. It is thus shown that the ratio of the maximal term  $M$  to [any of the] two others,  $L$  and  $\Lambda$ , exceeds any assigned ratio. QED.

**Lemma 5.** Assuming the same as above, it is possible to imagine such a large number  $n$ , that the sum of all the terms from the intermediate and maximal  $M$  to both the [to any of the] terms  $L$  and  $\Lambda$  inclusive, is to the sum of all the other terms exterior to the limits  $L$  and  $\Lambda$ , in a ratio greater than any given ratio.

*Dem[onstration].* Let the terms between the maximal  $M$  and the limiting term  $L$  on the left be designated thus: the second one from the maximal<sup>5.3</sup>,  $F$ , the third one,  $G$ , the fourth one,  $H$ , etc.; and the second one beyond  $L$ ,  $P$ , the third one,  $Q$ , the fourth one,  $R$ , etc. Since according to the second part of Lemma 3

$$M/F < L/P, F/G < P/Q, G/H < Q/R, \text{ etc.}$$

and (Lemma 4), for an infinite  $n$ ,  $M/L$  is also infinite, and

$$M/L, F/P, G/Q, H/R, \text{ etc.} \quad (5.2)$$

are certainly infinite just as

$$\frac{F + G + H + \dots}{P + Q + R + \dots}$$

is. That is, the sum of the terms between the maximal term  $M$  and the limit  $L$  is infinitely greater than the sum of the same number of terms beyond and nearest to  $L$ . And since according to Lemma 1 the number of all the terms outside  $L$  is not more than  $s - 1$  times (i. e., not more than a finite number of times) greater than the number of terms between this limit and the maximal term  $M$ , and the terms themselves, in accordance with the first part of Lemma 3, become the smaller the further they are from the limit, the sum of all the terms between  $M$  and  $L$  (even without considering  $M$ ) will be infinitely greater than the sum of all the terms beyond  $L$ . In a similar way it is shown that the sum of all the terms between  $M$  and  $\Lambda$  is infinitely greater than the sum of all the terms beyond  $\Lambda$  (whose number, according to Lemma 1, is not more than  $r - 1$  times greater than the number of the former).

Therefore, finally, the sum of all the terms situated between the limits  $L$  and  $\Lambda$  (the maximal term may be excluded) will be infinitely greater than the sum of all the terms beyond these limits. Consequently, this statement persists all the more if the maximal term is included [in the first sum], QED.

**Explanatory Comment.** Those, who are not acquainted with inquiries involving infinity may object to Lemmas 4 and 5 in the following way: Although, if  $n$  is infinite, the multiples of the magnitudes expressing the ratios  $M/L$  and  $M/\Lambda$ , that is,  $nr \pm n \mp 1$ ,  $nr \pm n \mp 2$ , etc., and  $ns \pm n \mp 1$ ,  $ns \pm n \mp 2$ , etc. have the same value as  $nr \pm n$  and  $ns \pm n$  since numbers 1, 2, 3... vanish with respect to each multiplier, it can still happen that, taken together and multiplied one by another, they increase to infinity (because the number of multipliers is infinite) and will infinitely decrease, that is, make finite, the infinite powers of the ratios  $[(rs + s)/(rs - r)]$  or  $[(rs + r)/(rs - s)]$ .

I cannot obviate these scruples better than by showing now a method of deriving a finite number  $n$ , or a finite power of a binomial, for which the sum of the terms between the limits  $L$  and  $\Lambda$  has a larger ratio to the sum of the terms beyond them than any no matter how great given ratio, which I designate by letter  $c$ . Once this is shown, the objection will necessarily fall down.

To this end, I choose some ratio [larger than unity], less, however, than the ratio  $[(rs + s)/(rs - r)]$  (for the terms on the left), for example, the ratio  $[(rs + s)/rs]$  or  $(r + 1)/r$ , and multiply it by itself so many times ( $m$  times) that the product becomes equal or exceeds the ratio of  $c(s - 1)$  to 1; that is, until

$$[(r+1)^m/r^m] \geq c(s - 1).$$

When will this happen can be advantageously investigated by means of logarithms. Because, taking logarithms, we obtain

$$m\text{Log}(r + 1) - m\text{Log}r \geq \text{Log}[c(s - 1)]$$

and, after dividing, we find at once that

$$m \geq \frac{\text{Log}[c(s-1)]}{\text{Log}(r+1) - \text{Log}r}.$$

I continue in the following way. With regard to a series of fractions or multipliers

$$\frac{nrs + ns}{nrs - nr + r}, \frac{nrs + ns - s}{nrs - nr + 2r}, \dots, \frac{nrs + s}{nrs}$$

from which, according to Lemma 4, the ratio  $M/L$  is obtained by multiplying them one by another, it may be remarked that, although the separate fractions are less than the fraction  $[(rs + s)/(rs - r)]$ , they approach it the nearer the larger is the assumed  $n$ . Therefore, one of them will sooner or later become equal to the ratio  $[(rs + s)/rs] = [(r + 1)/r]$  itself. It should be therefore found out how great  $n$  ought to be taken for the fraction whose ordinal number is  $m$  to become equal to  $[(r + 1)/r]$  itself. But (as it is seen from the law of the formation of the series) the fraction of ordinal number  $m$  is

$$\frac{nrs + ns - ms + s}{nrs - nr + mr}.$$

Equating it to  $[(r + 1)/r]$ , we obtain

$$n = m + \frac{ms - s}{r + 1} \text{ so that } nt = mt + \frac{mst - st}{r + 1}.$$

I maintain that if this is the power to which the binomial  $(r + s)$  is raised, the maximal term  $M$  will be more than  $c(s - 1)$  times greater than the limit  $L$ . Indeed, for the assumed value of  $n$  the fraction of ordinal number  $m$  will be equal to  $[(r + 1)/r]$ , and the fraction  $[(r + 1)/r]$ , being multiplied by itself  $m$  times, that is [the fraction]  $(r + 1)^m/r^m$ , is (as constructed) equal or greater than  $c(s - 1)$ .

Therefore, this fraction [of ordinal number  $m$ ] multiplied by all the previous fractions will all the more exceed  $c(s - 1)$  since all these are greater than  $[(r + 1)/r]$ . Consequently, the product, being multiplied by all the following [fractions], will all the more exceed  $c(s - 1)$  because each of these is at least greater than unity. But the product of all the fractions expresses the ratio of the term  $M$  to term  $L$  and it is therefore absolutely clear that the term  $M$  exceeds the limit  $L$  over  $c(s - 1)$  times.

But, see (5.2), it follows that the second term after the maximal term  $M$  exceeds the second term after the limit  $L$  more than  $c(s - 1)$  times, that the third term [after  $M$ ] still more exceeds the third term [after  $L$ ], etc. Therefore, finally, the sum of all the terms between the maximal  $M$  and the limit  $L$  will exceed the sum of the same number of

maximal terms situated beyond this limit more than  $c(s - 1)$  times, and more than  $c$  times the same sum taken  $(s - 1)$  times.

Consequently, it is still more evident that it exceeds more than  $c$  times the sum of all the terms situated beyond the limit  $L$  whose number is not more than  $s - 1$  times greater [than the number of terms between  $M$  and  $L$ ].

I proceed in the same way with regard to the terms on the right. I take the ratio

$$\frac{s+1}{s} < \frac{rs+r}{rs-s},$$

assume that

$$\frac{(s+1)^m}{s^m} \geq c(r-1)$$

and determine

$$m \geq \frac{\text{Log}[c(r-1)]}{\text{Log}(s+1) - \text{Log}s}.$$

Then, among the series of fractions

$$\frac{nrs+nr}{nrs-ns+s}, \frac{nrs+nr-r}{nrs-ns+rs}, \dots, \frac{nrs+r}{nrs}$$

included in the ratio  $M/\Lambda$ , I assume that the fraction having ordinal number  $m$ , namely,

$$\frac{nrs+nr-mr+r}{nrs-ns+ms},$$

is equal to  $(s+1)/s$ . I derive therefrom

$$n = m + \frac{mr-r}{s+1} \text{ so that } nt = mt + \frac{mrt-rt}{s+1}.$$

After this, it will be shown just as before that the maximal term  $M$  of the binomial  $r + s$  raised to this power will be more than  $c(s - 1)$  times greater than the limit  $\Lambda$ , and also, consequently, that the sum of all the terms between the maximal  $M$  and the limit  $L$  will be more than  $c$  times greater than the sum of all the terms beyond this limit whose number is not more than  $r - 1$  times greater [than the number of terms between  $M$  and  $\Lambda$ ]. And so we finally conclude that, upon raising the binomial  $r + s$  to the power equal to the greater of two numbers,

$$mt + \frac{mst-st}{r+1}, \quad mt + \frac{nrt-rt}{s+1}.$$

the sum of all the terms included between the limits  $L$  and  $\Lambda$  will exceed more than  $c$  times the sum of all the other terms extending on either side beyond these limits. The finite power possessing the desired property is thus discovered, QED.

**The Main Proposition.** Now follows the proposition itself for whose sake all the previous was stated and whose demonstration ensues solely from the application of the preliminary lemmas to the present undertaking. To avoid tediousness, I name the cases in which some event can happen fecund or fertile; and sterile, those in which the same event does not occur. In the same way, I name the experiments fecund or fertile if some fertile case appears in them and infertile or sterile when we observe something sterile.

Let the number of fertile cases be to the number of sterile cases precisely or approximately as  $r$  to  $s$ ; or to the number of all the cases as  $r$  to  $r + s$ , or as  $r$  to  $t$  so that this ratio is contained between the limits  $(r + 1)/t$  and  $(r - 1)/t$ . It is required to show that it is possible to take such a number of experiments that it will be in any number of times (for example, in  $c$  times) more likely that the number of fertile observations will occur between these limits rather than beyond them, that is, that the ratio of the number of fertile observations to the number of all of them will be not greater than  $(r + 1)/t$  and not less than  $(r - 1)/t$ .

*Dem[onstration].* Suppose that the number of the available observations is  $nt$ . It is required to determine the expectation, or probability that all of them without exception will be fecund; that all of them will be such with one, with two, 3, 4, etc. being sterile. Since, according to the assumption, there are  $t$  cases in each observation,  $r$  of them fecund and  $s$  sterile, and because separate cases of one observation can be combined with separate cases of another one, and then again combined with separate cases of the third, the fourth, etc., it is easy to see that the Rule attached to the end of the notes of Proposition 12<sup>5.4</sup> of Part 1 [of this book] and its second corollary containing the general formula by whose means the expectation of the lack of sterile observations,  $r^m : t^m$ ; of the expectations of one, two, three etc. sterile observations

$$\frac{nt}{1} r^{m-1} s : t^m, \frac{nt(nt-1)}{1 \cdot 2} r^{m-2} s^2 : t^m, \frac{nt(nt-1)(nt-2)}{1 \cdot 2 \cdot 3} r^{m-3} s^3 : t^m, \dots$$

are here suitable.

Therefore (after rejecting the common term  $t^m$ ) it becomes clear that the degrees of probability, or the number of cases in which it can happen that all the experiments are fecund, or all excepting one sterile, excepting two, 3, 4, etc. sterile, are expressed, respectively, by

$$r^m, \frac{nt}{1} r^{m-1} s : \frac{nt(nt-1)}{1 \cdot 2} r^{m-2} s^2, \frac{nt(nt-1)(nt-2)}{1 \cdot 2 \cdot 3} r^{m-3} s^3, \dots$$

that is, by the terms themselves of the binomial raised to the power of  $nt$ , which were just studied in our lemmas. All the rest is now manifest. Namely, it follows from the nature of the series that the number of cases, which add  $nr$  fecund to  $ns$  sterile observations, is indeed [corresponds to] the maximal term  $M$  since, according to Lemma 3,  $ns$  terms precede, and  $nr$  terms succeed it. In the same way, the number of cases in which there occurred either  $nr + n$  or  $nr - n$  fecund observations with the others being sterile, are expressed by the terms  $L$  and  $\Lambda$ ,  $n$  terms apart on either side from the maximal term  $M$ . Consequently, the total number of cases in which there are not more than  $nr + n$ , and not less than  $nr - n$  fecund observations, is expressed by the sum of the terms situated between the limits and  $\Lambda$ . The total number of the other cases in which there occur either more or less fecund observations is expressed by the sum of the other terms beyond the limits  $L$  and [or]  $\Lambda$ . The power of the binomial may be taken so great that, according to Lemmas 4 and 5, the sum of the terms between the limits  $L$  and  $\Lambda$  inclusive is more than  $c$  times greater than the sum of all the other terms exceeding these limits. It is thus possible to take so many observations, that the number of cases in which the ratio of the number of fecund observations to the number of all of them does not exceed the limits

$$\frac{nr+n}{nt} \text{ and } \frac{nr-n}{nt} \text{ or } \frac{r+1}{t} \text{ and } \frac{r-1}{t},$$

is greater than  $c$  times the number of the other cases. That is, it will become greater than  $c$  times more probable that the ratio of the number of fecund observations to the number of all of them is contained between the limits  $(r + 1)/t$  and  $(r - 1)/t$  rather than beyond them. *Quod demonstrandum erat.*

When applying this to separate numerical examples, it is self-evident that the greater, in the same ratio, we assume the numbers  $r$ ,  $s$  and  $t$ , the narrower can be made the boundaries  $(r + 1)/t$  and  $(r - 1)/t$  of the ratio  $r/t$ . Therefore, if the ratio of the number of cases  $r/s$  that should be determined by observation is, for ex[ample], one and a half, I take for  $r$  and  $s$  not 3 and 2, but 30 and 20, or 300 and 200, etc. It is sufficient to assume  $r = 30$ ,  $s = 20$  and  $t = 50$  for the limits to become  $(r + 1)/t = 31/50$  and  $(t - 1)/t = 29/50$ .

Suppose in addition that  $c = 1000$ . Then, according to what was prescribed in the Explanatory Comment, it will occur that, for the terms on the left and on the right respectively<sup>5,5</sup>,

$$m > \frac{\text{Log}[c(s-1)]}{\text{Log}(r+1) - \text{Log}r} = \frac{42,787,536}{142,405} < 301,$$

$$nt = mt + \frac{mst - st}{r+1} < 24,728,$$

$$m > \frac{\text{Log}[c(s-1)]}{\text{Log}(s+1) - \text{Log}s} = \frac{44,623,980}{211,893} < 211,$$

$$nt = mt + \frac{mrt - rt}{s + 1} = 25,550.$$

From which, as it was demonstrated there, it will follow that, having made 25,550 experiments, it will be more than a thousand times more likely that the ratio of the number of obtained fertile observations to their total number is contained within the limits  $31/50$  and  $29/50$  rather than beyond them. And in the same way, assuming  $c = 10,000$  or  $100,000$  etc., we will find that the same is more than ten thousand times more probable if 31,258 experiments will be made; and more than a hundred thousand times if 36,966 experiments will be made; and so on until infinity, always adding 5708 other experiments to the 25,550 of them. This, finally, causes the apparently singular corollary: if observations of all events be continued for the entire infinity (with probability finally turning into complete certitude), it will be noticed that everything in the world is governed by precise ratios and a constant law of changes, so that even in things to the highest degree casual and fortuitous we would be compelled to admit as though some necessity and, I may say, fate<sup>5,6</sup>. I do not know whether Plato himself had this in mind in his doctrine on the restoration of all things according to which everything will revert after an innumerable number of centuries to its previous state.

### Notes

**1.1.** This remark conforms to information theory.

**2.1.** It was Bortkiewicz (1917, p. x) who noticed the new word in the *Ars Conjectandi* and put it into scientific circulation, although Prevost & Lhuillier (1799, p. 3) had preceded him. The *Oxford English Dictionary* included this word, which had already appeared in ancient Greece (Hagstroem 1940), with a reference to a source published in 1662.

**2.2.** Although an astrologer, Kepler (1610, § 115; p. 238 in 1941) simply refused to answer definitely the same question. Times had changed! Bernoulli resumed this discussion in his Chapter 3.

**2.3.** The application of stochastic reasoning to one single case conforms to modern ideas.

**2.4.** John Owen (1563 – 1622). Haussner (Bernoulli 1713, German transl., p. 311) saw five editions of his *Epigrams*.

**3.1.** A few lines below I write *black tunic*, and, at the end of the chapter, *black toga*. Bernoulli himself applied three different nouns.

**4.1.** This is the very old principle of indifference. It can be perceived, for example, in the use of the arithmetic mean in astronomy since Kepler's time.

**4.2.** Arnauld was the main author of *L'art de penser* (Arnauld & Nicole 1662).

**4.3.** Bernoulli wrote *stones*; the German translation mentioned *small stones* (Steinchen).

**4.4.** The maximal term of the binomial  $(r + s)^n$  is approximately equal to  $1/\sqrt{2\pi nrs}$  and therefore decreases with an increasing  $n$  as  $1/\sqrt{n}$ , see e.g. Feller (1950, § 3 of Chapter 6).

**4.5.** Adriaan Metius (1571 – 1635); Ludolph van Ceulen (1540 – 1610).

**5.1.** A misprint in this ratio was corrected without comment in all the translations.

**5.2.** Separate parts of Bernoulli's *Tractatus de Seriebus Infinitis* appeared in 1689 – 1704, and, for the first time as a single entity, in 1713 together with the *Ars Conjectandi*.

**5.3.** The “second” (repeated in the same sense in the Explanatory Comment below) is unusual: Bernoulli actually had in mind the term immediately neighbouring *M*. Cf.: February is the second month of the year, not the second after January. A similar remark is of course valid with respect to the “third” and the “fourth”.

**5.4.** Bernoulli wrongly referred to Proposition 13. Haussner (Bernoulli 1713, German transl., p. 262) corrected him without comment.

**5.5.** The excessive number of significant digits below was the result of a venerable but misleading habit.

**5.6.** Bernoulli obviously had in mind the archaic notion of the Great Year (“innumerable number of centuries”).

## References

### Jacob Bernoulli

(manuscript, ca. 1684 – 1690; partial publication 1975), *Meditationes* [Diary]. In Bernoulli (1975, pp. 21 – 90).

(1713, in Latin), *Ars conjectandi*. Reprint: J. Bernoulli (1975, pp. 107 – 259).

(1899), German translation of same by R. Haussner: *Wahrscheinlichkeitsrechnung*. Reprint: Frankfurt/Main, 1899.

(1913). Russian translation of Part 4 of same by Ya.V. Uspensky with Foreword by A. A. Markov. Reprint: Bernoulli (1986, pp. 23 – 59).

(1966), *Translations from James Bernoulli* by Bing Sung this being an English translation of Part 4 of the *Ars Conjectandi*. Dept of Statistics, Harvard Univ., Techn. Rept No. 2. The translation is very loose and therefore almost worthless.

(1975), *Werke*, Bd. 3. Basel. Ed., B. L. van der Waerden. In addition to the *Ars Conjectandi*, and *Meditationes* (partly) it includes reprints of several classical contributions and commentaries.

(1986), *O Zakone Bolshikh Chisel* (On the Law of Large Numbers). Includes reprint of J. Bernoulli (1913) with comments (O. B. Sheynin, A. V. Prokhorov, N. G. Gamkrelidze), three commentaries (Sheynin, *Bernoulli and the beginnings of the theory of probability*; Yu.V. Prokhorov, *The law of large numbers and estimation of the probabilities of large deviations*; A. P. Youshkevich, *Biography of Bernoulli*), all this preceded by A. N. Kolmogorov’s Foreword and Markov’s speech of 1913 at the session of the Petersburg Academy of Sciences observing the bicentenary of the law of large numbers. Editor, Yu.V. Prokhorov. Moscow.

(1987), *Jacques Bernoulli & l’ars conjectandi*, this being a Latin-French edition of Part 4 of the *Ars Conjectandi*. Translation by B. Lalande. Ed., N. Meusnier. Paris.

### Other Authors

**Arnould, A., Nicole, P.** (1662), *L’art de penser*. (Appeared anonymously.) Paris, 1992. English translation: Edinburgh – London, 1850.

**Bernoulli, Nicolaus** (1709), *De usu artis conjectandi in jure*. In Jacob Bernoulli (1975, pp. 284 – 326).

**Bortkiewicz, L.** (1917), *Die Iterationen*. Berlin.

--- (1923), *Wahrscheinlichkeit und statistische Forschung nach Keynes*. *Nord. Stat. Tidskr.*, t. 2, pp. 1 – 23. English translation: *Silesian Stat. Rev.*, No. 17/23, 2019, pp. 85 – 109.

**David, Florence Nightingale** (1962), *Games, Gods and Gambling*. London.

**De Moivre, A.** (1712, in Latin), *De mensura sortis, or, The measurement of chance*. *Intern. Stat. Rev.*, vol. 52, 1984, pp. 236 – 262 and commentary by A. Hald (pp. 229 – 236).

--- (1756), *Doctrine of Chances*. New York, 1967. Previous editions: 1718, 1738.

**Descartes, R.** (1644), *Les principes de la philosophie. Oeuvr.*, t. 9, pt. 2 (the whole issue). Paris, 1978, this being a reprint of the edition of 1647.

**Feller, W.** (1950), *Introduction to Probability Theory and Its Applications*, vol. 1. New York, 1957.

- Franklin, J.** (2001), *The Science of Conjecture*. Baltimore.
- Garber, D., Zabell, S.** (1979), On the emergence of probability. *Arch. Hist. Ex. Sci.*, vol. 21, pp. 33 – 53.
- Gini, C.** (1946), Gedanken zum Theorem von Bernoulli. *Schweiz. Z. Volkswirtschaft u. Statistik*, 82. Jg., pp. 401 – 413.
- Graunt, J.** (1662), *Natural and Political Observations Made upon the Bills of Mortality*. Baltimore, 1939. Editor, W.F. Willcox.
- Hagstroem, K.-G.** (1940), Stochastik, ein neues – und doch ein altes Wort. *Skand. Aktuarietidskr.*, t. 23, pp. 54 – 57.
- Hald, A.** (1990), *History of Probability and Statistics and Their Applications before 1750*. New York.
- (2003), *History of the Law of Large Numbers and Consistency*. Univ. Copenhagen, Dept. Applied Math. & Statistics, Preprint No. 2.
- Halley, E.** (1694), An Estimate of the Degree of Mortality of Mankind. Baltimore, 1942.
- Halperin, T.** (1988), The development of probability logic from Leibniz to MacColl. *Hist. and Phil. of Logic*, vol. 9, pp. 131 – 191.
- Haushofer, D. M.** (1872), *Lehr- und Handbuch der Statistik*. Wien.
- Huygens, C.** (1657), Le calcul dans les jeux de hasard. *Oeuvr. Compl.*, t. 14. La Haye, 1920, pp. 49 – 91.
- Kepler, J.** (1596, 1621, in Latin), *Weltgeheimnis*. Augsburg, 1923. English transl.: New York, 1981.
- (1610), *Tertius interveniens*. *Ges. Werke*, Bd. 4. München, 1941, pp. 149 – 258.
- Keynes, J. M.** (1921), *Treatise on Probability*. *Coll. Writings*, vol. 8. London, 1973.
- Knapp, G. F.** (1872), Quetelet als Theoretiker. *Jahrb. National-Ökon. u. Statistik*, Bd. 18, pp. 89 – 124. Reprinted in author's *Einführung in einige Hauptgebiete der Nationalökonomie*. München, 1925, pp. 17 – 53.
- Kohli, K.** (1975a), Aus dem Briefwechsel zwischen Leibniz und Jakob Bernoulli. In Jacob Bernoulli (1975, pp. 509 – 513).
- (1975b), Kommentar zur Dissertation von N. Bernoulli. *Ibidem*, pp. 541 – 553.
- Koopman, B. O.** (1940), The bases of probability. *Bull. Amer. Math. Soc.*, vol. 46, pp. 763 – 774.
- Leibniz, G. W.** (Manuscript 1686), *Allgemeine Untersuchungen über die Analyse der Begriffe und wahren Sätze*. In author's book *Fragmente zur Logik*. Berlin, 1960, pp. 241 – 303.
- Maciejewski, C.** (1911), *Nouveaux fondements de la théorie de la statistique*. Paris.
- Markov, A. A.** (1924), *Ischislenie Veroiatnostei* (Calculus of probability). Moscow. Fourth, posthumous edition. Previous editions: 1900, 1908, 1913. German translation of the second edition: Leipzig – Berlin, 1912.
- Mill, J. S.** (1843), *System of Logic*. London, 1886.
- Montmort, P. R.** (1713), *Essay d'analyse sur les jeux de hazard*. Published anonymously. New York, 1980. First edition, 1708.
- Ondar, Kh. O.**, Editor (1977, in Russian), *Correspondence between Markov and Chuprov*. New York, 1981.
- Ore, O.** (1963), *Cardano, the Gambling Scholar*. Princeton.
- Oresme, N.** (1966), *De proportionibus proportionum and Ad pauca respicientes*. Ed. E. Grant. Madison. Latin – English edition.
- Pearson, K.** (1925), James Bernoulli's theorem. *Biometrika*, vol. 17, pp. 201 – 210.
- Prevost, P., Lhuillier, S. A. J.** (1799), Sur l'art d'estimer la probabilité des causes par les effets. *Mém. Acad. Roy. Sci. et Belles-Lettres Berlin avec l'Histoire*, 1796, pp. 2 – 34 of the second paging.
- Romanovsky V. I.** (1961), *Matematicheskaja statistika*, Book 1. Editor, E. A. Sarymsakov. Tashkent.
- Shafer G.** (1978), Non-additive probabilities in the work of [Jacob] Bernoulli and Lambert. *Arch. Hist. Ex. Sci.*, vol. 19, pp. 309 – 370.
- Sheynin, O.** (1970), On the early history of the law of large numbers. In *Studies in the History of Statistics and Probability* [, vol. 1]. Editors, E. S.

- Pearson, M. G. Kendall. London, 1970, pp. 231 – 239. Reprinted from *Biometrika*, vol. 55, 1968, pp. 459 – 467.
- (1973), Mathematical treatment of astronomical observations. *Arch. Hist. Ex. Sci.*, vol. 11, pp. 97 – 126.
- (1977), Early history of the theory of probability. *Ibidem*, vol. 17, pp. 201 – 259.
- (1982), On the history of medical statistics. *Ibidem*, vol. 26, pp. 241 – 286.
- (2001), Social statistics and probability theory in the 19<sup>th</sup> century. *Historia Scientiarum*, vol. 11, pp. 86 – 111. Shorter version: *Jahrb. National-Ökon. u. Statistik*, Bd. 223, 2003, pp. 91 – 112.
- (2003), On the history of the Bayes theorem. *Math. Scientist*, vol. 28, pp. 37 – 42.
- (2006), Review of Sylla (2006). *Historia Scientiarum*, vol. 16, pp. 212 – 214.
- (2014), Randomness and determinism. *Silesian Stat. Rev.*, No 12/18, pp. 57 – 74.
- Sylla, E. D.** (1998), The emergence of mathematical probability from the perspective of the Leibniz – Jakob Bernoulli correspondence. *Perspectives of Sci.*, vol. 6, pp. 41 – 76.
- (2006), English translation of Bernoulli (1713). Baltimore.
- Youshkevich, A. P.** (1986, in Russian), N. Bernoulli and the publication of J. Bernoulli's *Ars Conjectandi*. *Russian Math. Surveys*, vol. 31, 1987, pp. 286 – 303.
- The title of the initial Russian journal (which is now being translated) was *Teoria veroiatnostei i ee primeneniya*.