

Studies in the history of statistics and probability, vol. 25

Oscar Sheynin

Berlin. 2021

Contents

- i. O. Sheynin, *History of the theory of errors*, 2021
- ii. A. DeCandolle, On a dominant language for science, 1873
- iii. J. M. Keynes, Review of Czuber, 1911
- iv. J. Cournos, London under the Bolsheviks, 1919
- v. H. Cramér, Mises' work in probability and statistics, 1953
- vi. R. Frisch et al, Constitution of the Econometric Society, 1933
- vii. R. Frisch, Editorial, 1933
- viii. Jos, A, Schumpeter, Common sense of econometrics, 1933
- ix. T. Andersson, Statistics or chaos, 1929

I

History of the theory of errors

Second edition, essentially revised

First edition: Deutsche Hochschulschriften 1118
Hänsel-Hohenhausen, Egelsbach, 1996

Contents

- 0. Introduction
- 0.1. Goals of the theory of errors
- 0.2 Interrelations with statistics
- 0.3. Astronomy and geodesy
- 0.4. Why and when did the theory of errors emerge
- 0.5. The scope of this book
- 0.6. Some terminology and notation
- 1. The early history
 - 1.1. Bounds and estimators
 - 1.2. Regular observations
 - 1.3. Optimal circumstances of observation
 - 1.4. Ptolemy
 - 1.5. Some explanation
 - 1.6. Al-Biruni
 - 1.7. Galileo
 - 1.8. Tycho Brahe
 - 1.9. Kepler
- Notes
- 2. The eighteenth century
 - 2.1. The arithmetic mean
 - 2.2. Mayer
 - 2.3. Lambert
 - 2.4. Boscovich
 - 2.5. Simpson
 - 2.6. Lagrange
 - 2.7. Daniel Bernoulli
 - 2.8. Euler
- Notes
- 3. Laplace
- 4. The nineteenth century before 1809
 - 4.1. Solution of redundant systems of equations
 - 4.2. A venerable problem in land surveying
 - 4.3. Huber
 - 4.4. Legendre
 - 4.5. Adrain
 - 4.6. Gauss
- Notes
- 5. Gauss
 - 5.1. *Theoria motus*
 - 5.2. Determination of the precision of observations

- 5.3. Theoria combinationis
- 5.4. Practical considerations. The error theory surveyed

Notes

6. From Gauss to Helmert and beyond

- 6.1. Some new work
- 6.2. Physics, chemistry, meteorology
- 6.3. The normal law
- 6.4. The normal law modified
- 6.5. The theory of errors and statistics

Notes

7. Helmert

- 7.1. Rejection of outlying observations
- 7.2. Revealing systematic errors
- 7.3. Revealing systematic errors: the Abbe criterion
- 7.4. The sums of natural powers of errors
- 7.5. The chi-squared distribution
- 7.6. The Peters formula
- 7.7. The Gauss formula
- 7.8. Anticipating the Student – Fisher theorem
- 7.9. The precision of the mean square error
- 7.10. Is unbiasedness necessary?

Notes

8. Stable laws (Lévy)

- 8.1. Random errors
- 8.2. The precision of observations
- 8.3. The mean square error
- 8.4. A new concept of precision
- 8.5. Stable laws

Notes

9. The determinate theory of errors

- 9.1. The eighteenth century
- 9.2. Laplace
- 9.3. Gauss and Bessel
- 9.4. Helmert

Note

Bibliography

Acknowledgement

With sincere gratitude I recall the late Professors Youshkevitch, who was always favourably (too favourably) disposed towards me, and Truesdell, the Editor of the *Archive for History of Exact Sciences*, who had to busy himself with my English and compelled me to pay due attention to style. In 1991, after moving to Germany, I became able to continue my work largely because of Professor Johann Pfanzagl's warm support. In particular, he secured a grant for me (which regrettably dried up very long ago) from Axel-Springer Verlag. In my papers, I acknowledged the help of many colleagues including the late Doctors Chirikov (an able mathematician whose bad health thwarted his scientific career) and Eisenhart as well as the advice offered by R. Farebrother and R. L. Plackett.

0. Introduction

The booklet of 1996 was one of a very few attempts to describe the history of the theory of errors. Since then, I dwelt on much of its materials in my later papers and in my monograph (2017). Now, I have essentially revised my account of 1996. In particular, I included much (but by far not all the possible) information from my monograph. However, now that I reached the venerable age of 95+, the obvious difficulty of reprinting formulas compelled me to omit much information from my initial booklet.

One more point. Pearson (1892, p. 15) declared that the unity of science consists only in its method rather than subject. I think that *science* ought to be understood as a *given branch of science*. And the method can only be *theory*, theory of statistics, theoretical (rather than the narrower mathematical) statistics. I wholeheartedly agree. The first inference is that statistics need not have its own subject. Then, medical statistics, for example, is the application of the statistical method to medicine, and the theory of errors, its application to the treatment of observations (measurements). It follows, that the theory of errors is not alien to statistics. See however § 5.4.

O.1, O.2. The aims of the theory of errors.

Its relation with statistics

The theory of errors is a discipline that attempts to determine the best plausible results of measurements made in experimental science. Consider an example. Stations A and B are given and the position of C is fixed by intersection. Several questions should be answered.

1. How does the form of the triangle ABC influence the precision of determining point C?

2. How and at what time of the day should the angles be measured to minimize the influence of unavoidable random and systematic errors which depend on local meteorological conditions?

3. How to estimate several measurements of (each) angle?

4. How precise are the calculated elements of the triangle and the coordinates of point C?

5. Suppose now that point C is intersected from three stations, A, B and D, to eliminate blunders and increase precision. Lines AC, BC and DC will not generally meet at any single point, so where is the most plausible position of point C and how precise is it?

The answer to the first question is given by the sine theorem and the errors of the sides can be calculated since the prior approximate values of angles and their errors are known. The answer to the second question is ensured by an appropriate programme of observations. For example, if about a half of the measurements of each angle is done in the morning, and a half, in the evening.

Other questions require stochastic considerations. The theory of errors thus has two parts: the stochastic branch which treats the results of measurement and the determinate branch which examines the entire process of measurement to ensure best possible results. The same is the aim of the exploratory data analysis and the experimental design which belong to theoretical (not mathematical) statistics. Nowadays they swallowed the determinate branch of the theory of errors, it does not exist anymore.

Random errors are a particular case of random variables and the stochastic branch of the theory of errors heavily relies on probability theory whose development from the mid-18th century and perhaps up to the 1920's was largely determined by the need to justify and advance mathematical treatment of observations. Thus, Poincaré (posth. publ. 1921, p. 343) confessed that for the first edition of 1896 of his treatise *La théorie des erreurs était naturellement mon principal but*.

And Lévy (1925, p. VII) indicated that, without that theory his main contribution on stable laws *n'aurait pas de raison d'être*. He was mistaken: for the theory of errors stable laws are not needed.

In turn, I emphasize that mathematical statistics borrowed variance and the principles of maximal likelihood and minimal variance from the theory of errors.

To continue with some questions. In the case of one unknown (of *direct observations*) it is required to choose its final value, given its observations $x_1, x_2, \dots, x_n$ and to estimate its plausibility.

In general, we need to adjust *indirect observations*, to deduce final values for the unknown constants $x, y, z, \dots$ from a redundant system of equations

$$a_i x + b_i y + c_i z + \dots + l_i = 0, i = 1, 2, \dots, n \quad (1)$$

with given coefficients and measured free terms; and to estimate the plausibility of these values and/or their functions. The linearity of equations (1) is not restrictive: the approximate values of $x, y, z, \dots$ are either known or can be calculated from any *square* subsystem of (1).

For *physically* independent free terms equations (1) are inconsistent and any set $(\hat{x}, \hat{y}, \hat{z}, \dots)$ which leads to reasonable residual free terms v_i has to be admitted as a solution. The MLSq is no exception, it requires that the sum of the squares of v_i is minimal among all possible sets of $(\hat{x}, \hat{y}, \hat{z}, \dots)$.

0.3. Astronomy and geodesy

An expedient treatment of astronomical observations allowed Kepler to establish the actual system of the world (§ 1.9). (Not necessarily meridian) arc measurements allow geodesists to determine the general figure of the Earth. Newton proved that the Earth is a flattened ellipsoid of rotation, and arc measurements were also needed to confirm/refute his theory. To this end, two measurements are necessary, but many more are needed to compensate hopefully the local deviations from the figure of the Earth and increase precision. Not the semi-axes of the ellipsoid, a and b ($a > b$) but a and the flattening $(a - b):a$ were usually determined.

Suppose now that points A and B are situated on the same meridian and that O denotes the centre of the Earth. Then angle AOB is the astronomically measured latitudinal difference between A and B, and the length of AB is indirectly measured by a chain of triangulation and a base (much better, by bases at the ends of the triangulation).

By the end of the 18th century arc measurements became also necessary for the introduction of the metric system of measures ($1 \text{ m} = 10^{-7}$ of a quarter of a meridian) and cartography. Pendulum observations became an important additional means for establishing the flattening of the ellipsoid. Nowadays they, along with other instruments, are also used for studying the Earth's gravitational field.

0.4. Why and when did the theory of errors emerge?

It was needed to understand the essence and effects of random and systematic errors; to formulate the main aims and methods of a yet non-existing theory; and to describe these aims and methods in a separate source. Consequently, I hold that the error theory emerged in the 18th century and now I attempt to isolate several stages of its development.

The first stage. Scientists enjoyed full power over their observations and only a (small) part of the data might have been used while the rest of them remained unknown to the scientific community. Ptolemy embodied this attitude whereas Tycho Brahe apparently heralded the coming of a new period. So did Graunt, but in population statistics. It was Tycho to whom we owe the origin of experimentally checked ideas on observations and their preliminary treatment much

more than to Ptolemy and it was Galileo who formulated the elements of the theory of errors,

The second stage. All the observations were, or should have been generally known, but they were treated either subjectively or at best without proper stochastic or statistical interpretation (Boscovich).

The third stage. The treatment of observations was and is accompanied by statements on the stochastic and/or statistical properties of the final results. Anyway, these properties are known now. This stage had begun somewhat before Boscovich published his findings. In the second half of the 18th century frequency laws were introduced and studied (Simpson, Lambert); the arithmetic mean justified; the principle of maximum likelihood put forward (Lambert); the precision of observations estimated (Lambert, badly); the theory of errors and its name coined (Lambert); indirect observations adjusted; and the principle of least squares heuristically anticipated (Euler).

The fourth stage. The classical theory was developed, mostly by Gauss. Important supplementary work followed, methodological improvements achieved and generalisation and improvement of Gauss's ideas attempted.

It was Gauss who shaped the treatment of observations into a practical tool. Unlike Laplace, he did not presume a very large number of observations. Then, here is an important feature of his scientific work (Subbotin 1956, p. 297): Lagrange and Laplace had

Restricted their attention to the purely mathematical aspect [of determining the orbits of celestial objects] whereas Gauss thoroughly worked out his solution from the point of view of computations and took into account all the conditions of the work of astronomers and [even] their habits.

As to the determinate error theory, its main stages were

The period before the 18th century. The existence of errors and general measures protecting against them and/or minimizing their influence became known.

The 18th century and Laplace. Differential formulas for estimating the precision of geodetic networks and the influence of observational and instrumental errors were applied and the two kinds of errors (random and systematic) isolated.

Gauss and Bessel. The hunt for errors began in earnest. The instrument and the methods of observation were considered faulty in every possible way unless and until thoroughly checked. Means for eliminating or minimizing each conceivable error were devised.

After Gauss. Helmert. The precision of geodetic networks was studied more extensively. For a general adjustment separate parts of large and only gradually constructed networks began to be replaced by arcs of geodesics.

0.5. The scope of this book

Drawing on many sources and my own papers (see Bibliography) I trace the history of the theory of errors from Ptolemy to Kepler and Galileo, to the period of arc measurements, to Lambert and Simpson, to Laplace, Gauss and Bessel, and finally to the beginning of the 20th century.

From previous sources I mention, first of all, Helmert (1872) and Idelson (1947). Both are dated but still useful. And, beginning from Idelson mathematical statistics started to invade actively the theory of errors. I also recommend Farebrother (1999) written on the modern level, and Hald (1990; 1998).

I loathe mentioning Stigler (1986). He was the first and hopefully the last who dared to slander the memories of Euler and mostly Gauss. **No one** defended those giants which means that the scientific community (perhaps not only statisticians) is seriously ill. Even until now, and especially in his fatherland, the USA, Stigler is considered a great scientist. Shame indeed! Hald (1998, p. XVI) soiled himself by calling Stigler's book *epochal*.

I (1990a, 1990c) refuted his astonishing declarations which he even repeated in 1999 slightly less impudently.

0.6. Some terminology and notation

Without loss of generality I use equations (1) only in two or three unknowns and I call these systems *linear* without any adjectives, redundant algebraic. I also apply two terms, random variable and normal law irrespective of their introduction. I also use notation $\bar{x}$ for arithmetic means of x_i and $E\xi$, the expectation of random variable ξ . I apply the exceptionally apt and elegant Gaussian (1811, § 13) notation of the type

$$[ab] = a_1b_1 + a_2b_2 + \dots + a_nb_n,$$

cf. symbol $[vv]$ in the condition of least squares. Laplace obstinately refused to apply this notation, and later French scientists did not dare apply it either. In the same source Gauss introduced further symbols

$$\begin{aligned} [bb,1] &= [bc] - [ab][ac]:[aa], \\ [cd,2] &= [cd,1] - [bc,1][bd,1]:[bb,1] \text{ etc.} \end{aligned}$$

The commas later disappeared but otherwise such symbols are convenient in solving normal equations by successive elimination of the unknowns (Gauss).

Other points. Abbreviation: CLT = central limit theorem; MLSq = method of least squares; W-i = The *Werke* of Gauss, Bd. i. Then,

S, G, i stands for downloadable file i from my website www.sheynin.de which is being diligently copied by Google, Oscar sheynin, Home. I apply this notation in cases of rare sources or those which I translated into English.

1. The early history

The attitude of ancient astronomers towards the treatment of observations should be explained. Below, I attempted to achieve that goal.

1.1. Bounds and estimators

Archimedes (1925, pp. 68 – 69) was one of the first to state that neither human faculties, nor instruments ensure sufficient plausibility of observation. However, he continued, since this subject was *often*

treated, it was unnecessary to elaborate. So we do not know anything more from him.

Ancient astronomers realised that their observations were imperfect and attempted to establish bounds for the measured magnitudes. Thus, Toomer (1974, p. 139) remarked that this

Became a well-known technique ... practised for instance by Aristarchus, Archimedes and Eratosthenes.

As an example, I cite Aristarchus (1959, p. 403) who maintained that

The diameter of the sun has to the diameter of the Earth a ratio greater than ... 19:3 but less than ... 43:6 [6.33 and 7.17].

Suppose now that the observations of an unknown constant x are $x_1, x_2, \dots, x_n$ with $x_1 \leq x_2, \dots, \leq x_n$. The difference $(x_n - x_1)$ is called the range (of observations) and the end points can be the bounds of x . But the range tends to increase with the increase of n , and the bounds should be estimated by indirect evidence or theoretical considerations.

Again, bounds hardly helped when an observed magnitude served as an initial parameter or argument in difficult calculations. Even worse when several such magnitudes had to be used^{1.1}.

Thus, bounds did not eliminate the need to assign some point estimator for x . Ancient astronomers hardly applied any universal estimator such as the arithmetic mean. They likely chose some number taking into account previous knowledge and their own subjective feelings as well as convenience of subsequent calculations (Neugebauer 1950, p. 252). Perhaps only once Ptolemy (I 12, p. 63; H 68)^{1.2} explained his choice: he chose the midrange, $(x_1 + x_n):2$. The greatest possible error became minimal, cf. the method of minimax (§ 1.9).

The value of twice the obliquity of the ecliptic was known to be greater than $47^\circ 40'$ and less than $47^\circ 45'$, he stated, so that

*We derive very much the same ratio as Eratosthenes, which Hipparchus also used. For the arc ... is approximately $(11:83)360^\circ$ [$=47^\circ 42' 39''$]. He then remarked that *One takes the point halfway between the two extrema* [the midrange]^{1.3}.*

1.2. Regular observations

Another noteworthy feature of ancient astronomy was the understanding of the need to observe regularly. Ptolemy (III 1, p. 132; H 194) testified that Hipparchus had regularly observed the length of the tropical year. Not much is known about this scholar whose work substantially helped Ptolemy to develop his classical system of the world and whom Ptolemy (IX 2, p. 421; H 210) called *a great lover of truth*. This telling remark apparently means that Hipparchus was not afraid of revealing discrepant results, cf. Toomer (1974, p. 140).

It is of course unknown to what extent did Hipparchus realise that regular observations provide a means for diminishing the influence of some (of random) errors and eliminating other (systematic) errors.

1.3. Optimal circumstances of observations

The third and last feature of ancient astronomy with which Neugebauer (1950, p. 250) credits even Babylonian astronomers of the Seleucid period, was the use of optimal circumstances of

observation. For example, at certain times a given error in registering the moment of an astronomical phenomenon has a much lesser influence on the final result than at other times. Or (Ptolemy V 14, p. 252; H 417), the equality of two certain angles can be established easier than their magnitudes. Here, the optimal approach was connected with the possibility of altogether excluding measurements inevitably corrupted by considerable error.

Elsewhere Neugebauer (1948, p. 101) remarked that observations in ancient times *were more qualitative than quantitative*. I am not satisfied with this expression, *qualitative observation*, which Aaboe & De Solla Price (1964) even included into the title of their contribution: almost all observations are quantitative. But I ought to add that in general ancient science, unlike its current counterpart, was qualitative and that the choice of optimal circumstances of observation was one of the aims of the disappeared determinate branch of the theory of errors. This choice can be easily determined beforehand by elements of the differential calculus, but in antiquity trial and error was the only means for achieving success. Ancient astronomers were apparently able to succeed.

No wonder that they often selected one or a few results and used the other observations *only* for a rough check. Thus Ptolemy (III 1, p. 137; H 203) *abandoned* observations *conducted rather crudely* and Al-Biruni (1967, pp. 46 – 51) rejected four indirect observations of the latitude of a certain town in favour of its single and simple direct measurement.

1.4. Ptolemy

Ptolemy was a theoretician rather than an observer. Nevertheless, he adhered to at least two of the three features of ancient astronomy: he observed regularly and conducted his observations under optimal circumstances. With regard to the former see Ptolemy (III, 1, pp. 132 and 136; H 194 and 201) and here are three worthy testimonies.

1) Kepler (1609/1992, p. 324):

We have hardly anything from Ptolemy that we could not with good reason call into question prior to its becoming of use to us in arriving at the requisite degree of accuracy.

2) Laplace (1796/1884, p. 413):

Hipparchus among all ancient astronomers deserves the gratitude of astronomy for the large number and precision of his observations, for important conclusions which he had been able to make by comparing them with each other and with earlier observations, and for the witty methods by which he guided himself in his research.

Ptolemy, to whom we are mostly indebted for acquainting us with his work, had invariably based himself on Hipparchus' observations and theories. He justly appraised his predecessor...

And on the next page:

His Tables of the Sun, in spite of their imperfection, are a durable monument to his genius and Ptolemy respected them so much that he subordinated his own observations to them.

3) Newcomb (1878, p. 20):

All of Ptolemy's Almagest seems to me to breathe an air of perfect sincerity.

And it is dead certain that Ptolemy knew all there was to be known about optimal circumstances. Here is just one of his pronouncements (IX 2, p. 423; H 213):

Planetary observations that are most likely to be reliable are those in which there is observed actual contact or very close approach to a star or the moon and especially those made by means of the astrolabe instruments.

I do not think that he did not follow his own indirect recommendations, or still less, that he (R. R. Newton 1977, p. 379) was

The most successful fraud in the history of science who invented all his observations. True, much can be said about Ptolemy's obscure ways, see Kepler's opinion above. Modern astronomers loyal to Ptolemy agree that he rejected, adjusted or incorporated *a great array of materials as he saw fit* (Gingerich 1983, p. 151) and that he was an opportunist ready *to simplify and to fudge* (Wilson 1984, p. 43).

Ptolemy's attitude towards rounding off is also difficult to understand. Commentators seem to agree that (Neugebauer 1948, p. 113)

The ancients [including Ptolemy] were little concerned about the influence of rounding off and accumulated errors. Often the errors are of the same order of magnitude as the effect under consideration.

Perhaps (Lourier 1934, p. 37) approximate calculations were indeed *Attributed ... to the realm of lower, applied science ... unworthy of inclusion into scientific mathematics.*

That Ptolemy had at least a general knowledge of systematic and random errors of observation is obvious: it could not have been otherwise. And here are some of his own statements (VIII 6, p. 416; H 203 and 1956, III, 2, p. 231):

The difference between

The observers themselves and the atmosphere in the regions of observation can produce variation in and doubt about the time of the [phenomenon], as has become clear, to me at least, from my own experience and from the disagreements in this kind of observations.

Practically all other horoscopic instruments on which the majority of the more careful practitioners rely are frequently capable of error, the solar instruments by the occasional shifting of their position and of their gnomons, and the water clocks by stoppages and irregularities in the flow of water from different causes and by mere chance.

1.5. Some explanation

Now I venture to formulate several principles which possibly governed the scientific behaviour of ancient astronomers.

1. They regarded their observations as private property. They rejected some data without informing anyone else about it, let alone about the thus discarded results.

2. Astronomers made known some other observations for use by colleagues, sometimes without references: everyone knew what others does (done). Many authors beginning at least with Al-Biruni

(Shevchenko 1988, p. 175) believe(d) that Ptolemy had copied Hipparchus' work. Possibly he did, but in good faith.

3. Understanding that, in spite of the precautions taken, most of their observations were corrupted by considerable errors, astronomers doctored the obtained results. They felt that an observation was not a definite number or point, but almost any number lying within the estimated bounds¹⁻⁴. From a modern point of view, in case of *bad* distributions, it is possible to leave whichever of such observations and reject all the others.

Ptolemy's cartographic activities indirectly corroborate this conclusion. It seems (Berggren 1991) that he was mainly concerned with semblance of truth [with general correctness] rather than with mathematical consistency. A related fact is known even in the Middle Ages (D. J. Price 1955, p. 6):

Many medieval maps may well have been made from general knowledge of the countryside without any sort of measurement or estimation of the land by the "surveyor".

Theoretical considerations partly replaced measurements in the Chinese meridian arc of 723 – 726 (Beer et al 1961, p. 26). At the end of their paper the authors suggested that it was considered more suitable to present harmonious results.

1.6. Al-Biruni

He was the only Arab scholar to surpass Ptolemy and be a worthy forerunner of Galileo and Kepler. It is natural that he adhered to the sound traditions of his predecessors.

1. He (1967, p. 203) used bounds:

As to the latitude of Bagdad, different observers have found that it is neither less than 33°20' or greater than 33°30' and the approved one is 33°25' because it is also the mean between these two.

Apparently the midrange carried some weight, but the *also* implied that some other estimator was used. Here are two other quotes. On p. 168: *As to the halving of the interval between the two times, it is a rule of procedure adopted by calculators for minimizing the errors of observation so that the time calculated will be between the upper and the lower bounds.*

And on p. 237: Al-Biruni will *rely* on a certain amount

Because it is close to the average between the smaller amount ... and the larger amount ... and because the indirect method ... produces an amount which is not far from that amount.

2. Al-Biruni (1967, p. 32) recommended to observe latitudes *continually* to forecast landslides etc. Given the precision of such observations, his advice is useless. But in the same source Al-Biruni mentions his regular observations made for astronomical purposes and on p. 65 testifies that Al-Battani *declared that he had repeated his observations over many years.*

3. I adduce only one passage (Al-Biruni 1967, p. 58) on the use of optimal circumstances of observation:

The use of sines engenders errors which become appreciable if they are added to errors caused by the use of small instruments and errors made by human observers.

Two more points: Al-Biruni's estimators and his thoughts about elimination of a certain systematic error.

4. At least once Al-Biruni (1967, p. 83) chose a comfortable and common-sense but hardly universally accepted estimator:

Now all the testimonies that we have adduced point out collectively that the [obliquity of the ecliptic] is 23° plus one third and one quarter of a degree. The slight excess or defect in some [of the observations] is due to the instrument.

I corroborate my conclusion made just above by Al-Biruni's (1983, pp. 60 – 62) manner of determining the densities of metals. Sometimes he assumed the mode of the measurements but in other cases he chose either the midrange or an unspecified value between the extreme observations.

5. Describing the determination of the longitudinal difference $\Delta\lambda$ between two cities Al-Biruni (1967, p. 155) recommended that

Observers of an eclipse [should] obtain all its times [phases] so that every one of these in one of the two towns can be related to the corresponding time in the other. Also, from every pair of opposite times that of the middle of the eclipse must be obtained.

The differences between the $\Delta\lambda_i$ will characterise the random component of the error of $\Delta\lambda$.

1.7. Galileo

Maistov (1974, pp. 30 – 34) was the first to dwell on Galileo's (1632, pp. 280 – 318)^{1,5} thoughts on the treatment of observations. They were prompted by discussions on whether the new star of 1572 was situated *below* the moon, between it and the stars or among the stars.

Astronomers attempted to measure the diurnal parallax of the new star although even its annual parallax was too insignificant. But Galileo formulated important propositions about the properties of [usual random] errors. He stated that the modulo lesser errors were more probable and that positive and negative errors were equally likely. And Galileo also argued that in adjusting observations the sum of the absolute values of the corrections should be the least possible.

The proper discussion of Galileo's work is in Hald (1990, pp. 149 – 160).

1.8. Tycho Brahe

Wesley (1978, p. 52) maintained that [in astronomy; otherwise Graunt should be also named] Tycho was

The first to see that it is ... necessary to take long series of observations so that random, instrumental and human error can be averaged out^{1,6}.

It is possible that Tycho had programmes of observation capable of averaging out, to a certain extent, all types of errors. However, ancient astronomers made regular observations long before him and likely more or less understood the difference between random and systematic errors.

Wesley (p. 51) also states that Tycho combined measurements made by many instruments, but how did he treat his observations when at least one instrument had to be temporarily taken out of service? A systematic shift in the mean result would have occurred.

Anyway, Tycho enabled Kepler to construct a new (the new) system of the world. The development of observational know-how and programmes rather than furthering of more abstract ideas in treating observations was the essence of Tycho's life. Tycho also worked out a curious system of the world with the Earth at its centre, the Sun moving round it (as in the Bible!) and the other planets rotating round the Sun. Kepler rejected this system as well.

Tycho's attempts undoubtedly included some adjustment of observations but they remain hardly studied. A special point here is Tycho's derivation of the right ascension of a certain star, see Plackett (1958, pp. 122 – 123). To eliminate the systematic effect of parallax and refraction Tycho combined 24 of his observations into (12) pairs and calculated the general mean of these pairs and of 3 (obviously already corrected) separate observations. He assigned equal weight to each of the 15 values thus obtained. His mean (with the degrees omitted) was $28^{\circ}9'$ whereas the weighted mean with the weight of the separate observations of $1/2$ would have been $28^{\circ}2'$. The difference is insignificant but an explanation (such as my own) was necessary but lacking.

The pairing considerably eliminated systematic influences. Thus, the first separate observation was 44° but the components of the first pair differed from it by $3^{\circ}32'$ and $-4^{\circ}21'$ ^{1.7}. Apparently Tycho calculated the intermediate means to estimate the influence of the residual random or largely random errors.

1.9. Kepler

Unlike Ptolemy or Al-Biruni, Kepler (1606/2006, p. 163) left us his pronouncement on randomness:

But what is chance? Nothing but an idol, and the most detestable of the idols; nothing but a contempt for God, sovereign and almighty, as well as for the highly perfect world come out of his hands.

At the same time, he had to leave room for randomness in his system of the world: understandably, the eccentricities of the planetary orbits did not obey any law of nature.

In discussing and treating observations, Kepler did not distinguish between the two main kinds of error, but his arguments, whether explicit or indirect (below) belonged to random errors. First, he (1992, pp. 210/63, 215/71 and several other places) stated that errors were inevitable. Once he (p. 286/113) used the term *uncertainty* on a par with *latitudo* of the observations which apparently meant *range*. But the range tends to increase with the number of observations, and it is not a trustworthy measure of precision. However, nothing better was then known.

Second, Kepler (1992, p. 520/254) hinted at the equal probability of errors of both signs:

If ... we take a mean ... as if saying that in the two observations ... there were some small observational errors in opposite senses ...

This is also an approach to justify the choice of the arithmetic mean although only in the simplest case of two observations. An indirect and more general remark to the same effect is contained in Kepler's letter of 1627 to the Senate of Ulm (Caspar & von Dyck 1930, p. 248):

The total weight [more correctly, a coin's average weight] of many coins of the same mintage hardly depends on the precision of the weights of the individual coins. That was a rudiment of the law of large numbers.

In one instance Kepler (1609/1992, p. 200/63) collected four observations of the right ascension of Mars. With the degrees omitted they were

$$x_1 = 23'39'', x_2 = 27'37'', x_3 = 23'18'', x_4 = 29'48''.$$

He chose, as his final value,

$$x = (x_1 + 2x_2 + x_3)/4$$

which is a weighted arithmetic mean with weights 1, 1, 2 and 0. He called it *medium ex aequo et bono*. Kepler did not explain his choice; his formula was reconstructed by J. J. Filliben (Eisenhart 1976, p. 356). It indeed fitted the Latin expression (*In fairness and justice*), but it is much more important that that expression goes back to Cicero^{1.8} and implies *Rather than according to the letter of the law*. Thus, already in Kepler's time the arithmetic law was apparently the letter of the law. Without enjoying any special status in antiquity, it came to the fore, perhaps in the late 16th century and is still with us.

Kepler (1609/1992, pp. 521 – 524/255 – 256) also adjusted indirect observations. His procedure was not sufficiently general: with another data he would have possibly been compelled to devise another approach. Here is his remark (p. 523/256) that exemplifies his ad hoc attempts:

Since the first and third position of Mars ... agree rather closely, some less thoughtful person will think that it [the final position] should be established using these, the others being somehow reconciled. And I myself tried to do this for rather a long time. But since the second and fourth could not be reconciled [they had to be accounted for as well].

It is certain that in adjusting numerous observations, Kepler did not sin against common sense and kept the corrections *within the limits of observational precision* (1609/1992, p. 334/143). His most celebrated pronouncement (p. 286/113) which heralded the refutation of the Ptolemaic system of the world corroborates his statement quoted just above:

The divine benevolence had vouched us Tycho Brahe, a most diligent observer from whose observations the 8' error [which is inadmissible] in this Ptolemaic computation [in fitting this to the Ptolemaic system] is shown.

I believe that Kepler thus made use of the principle of minimax, of seeking such a solution of equations for which the modulo maximal residual is minimal. His equations were not lineal, not even algebraic, but they could have been linearised. Now, if even the minimal value of a residual is not reasonably small, the underlying theory has to be rejected (or the observations defective, which was here impossible).

Kepler had to proceed by trial and error (only Laplace devised the proper algorithm).

The minimax principle does not belong to the stochastic theory of errors (its place is in the statistical decision theory)^{1.9}, It is however interesting to note that this principle corresponds to generalised least squares, and in case of direct observations it leads to the midrange. Daniel Bernoulli (1778, § 10) found this estimator *less wrong* [apparently less deviating from the arithmetic mean than he thought before]. He did not elaborate.^{1.10}

Notes

1.1. In our day the mean square error of a function of several arguments can easily be calculated, given their errors rather than bounds. Even so, some possible circumstances, as dependence between the arguments or their different stochastic behaviour, should be considered.

1.2. References such as this one are to the English translation of Ptolemy's *Almagest* (1984).

1.3. According to the Mishnah text of the Talmud (second century; Rabinovich 1974, p. 352), the volume of the *standard* hen's egg was supposed to be the mean of the two volumes for the largest and the smallest egg (the midrange).

1.4. This attitude would have been in keeping with the qualitative gist of ancient science which provided a general impression of nature rather than its quantitative description. Thus, ancient geographers isolated climatic belts but did not know anything about measuring air temperature. And ancient physicians (Hippocrates) believed that fat men are apt to die earlier, but they did not register age at death.

Both examples also illustrate the origin of the first stage of the statistical method (Sheynin 1982, p. 242) during which qualitative inferences have been made in the absence of statistical data.

1.5. Earlier, in 1964, Maistrov described them in an ad hoc Russian paper, but did not mention that already Buniakovsky (1846, Chapter on history of probability) remarked that Galileo had considered errors of observation.

1.6. This subdivision is hardly sufficient since instrumental and human errors can be either systematic or random.

1.7. It is these figures that explain why the mean of a pair was equivalent to a separate corrected observation.

1.8. Cicero, *Rede für A. Cäcina*, § 65. Published together with translations of other of his pieces, each with its own paging, thus constituting a book without any general title-page or date. Stuttgart.

1.9. Lambert (1765a, § 420) mentioned the minimax principle, but confessed that he was unable to use it *auf eine allgemeine Art und ohne viele Umwege*. Obviously, he was unaware that Euler (1749, §§ 122 – 123) had achieved some success in applying it.

1.10. In meteorology, the extreme observations are perhaps more apt to be extraneous than elsewhere and the midrange is consequently more dubious. In any case, Jacob Bernoulli in his *Diary* partly published in Latin (1975, p. 47) argued that, when estimating the mean height of the barometer, the arithmetic mean should be preferred to the midrange. However, even in the 19th century, in spite of further objections meteorologists still made use of this statistic (Sheynin 1984b, p. 74).

2. The eighteenth century

During the second half of that century extremely important achievements in treating observations were made and the theory of errors emerged. I do not dwell here on Laplace's early memoirs, see next chapter.

2.1. The arithmetic mean

Recall (§ 1.9) that during Kepler's lifetime or somewhat earlier the arithmetic mean apparently became generally accepted as the

estimator of the *true value* of the measured constants. Cotes (posth. publ. 1722, p. 22; see also Gowing 1983) seems to be the first who left a direct statement about the mean and maintained that it provided the *most probable* [estimate] of the true position of an object. He did not explain what he meant by *most probable* or exemplify his rule, but his authority apparently supported the common feeling. Laplace (1812/1886, p. 351 – 353) indicated that astronomers began to apply the Cotes rule after Euler (1749). But even before that Picard (1693/1729, pp. 330, 335, 343) called the arithmetic mean the *véritable* value.

Delambre 1827, p. 455) cautiously remarked:

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

And now Condamine (1751, p. 223):

En prenant ... un milieu entre un grand nombre d’observations, on court peu de risque de se tromper. Et quand même il y a en auroit dans ce grand nombre quelques-unes de sensiblement défectueuses, le moyen résultat seroit à peine alteré: puisque l’excès ou le défaut de celles-si se partageant entr’elles et toutes les autres, changeroit peu le résultat.

This is apparently a good example of the feelings at that time.

2.2. Mayer

Mayer (1750) was apparently the first to solve linear systems by arranging their equations into subsystems (groups) and calculating intermediate partial solutions. He had 27 equations in three unknowns and combined them into three disjoint groups of nine equations each. Then he derived his unknowns by solving the new system. Each group was obtained by summing up the pertinent equations. If, for example, a group consisted of the first nine equations, then, as Mayer assumed, the sum of the residual free terms vanished

$$v_1 + v_2 + \dots + v_9 = 0.$$

He remarked that arranging his 27 equations three at a time etc. was too difficult, but suppose he had 31 or 25 equations, how would he proceed?

Mayer was mostly interested in only one of his unknowns, call it x . All equations of his first group had the largest positive coefficients a_i of x ; in the second group the a_i had the largest negative values. Thus $[aa]$ was reasonably large, and the weight of the least-squares estimator of x was adequate.

Mayer also estimated the increase in precision of x with the increase in the number of equations and decided that a nine-fold increase of that number led to the same increase of precision. For his time, his mistake was understandable.

In a letter to Schumacher of 24.6.1850 Gauss (Peters 1865, Bd. 6, p. 90) remarked that

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

He referred to Mayer's manuscripts, but it is likely that Mayer's trick was almost the same in both cases. And Gauss himself, in a letter to Schumacher of 22 Febr. (Ibidem, pp. 66 – 67) described a similar procedure. True, he thought of calibrating an aneroid rather than quantifying a regularity of nature^{2,1}

2.3. Lambert

Lambert was a versatile scholar. He partly devoted some of his contributions (1760; 1765a, 1765b) to the treatment of observations. In 1760, he classified errors according to their origin (§ 282), attempted to prove that extreme observations should be rejected (§§ 287 – 291) and estimated the precision of observations (§ 294). Then he introduced an unspecified continuous frequency curve (§ 296) and formulated the principle of maximal likelihood (§ 303).

Below, I describe all three of his works; items 1 – 3 belong to the first of them.

1. Rejection of outliers. Lambert proved that the observation corrupted by the modulo greatest error should be rejected, but the necessary information is lacking and his dubious advice is moreover worthless.

2. The precision of observations. Denote the arithmetic mean of all n observations by u , and by v if an extreme observation is rejected. Lambert maintained without proof that u hardly deviated more than by $|u - v|$ from the true value sought. If the rejected observation was a blunder, Lambert's estimate will be too conservative and in any case it is not connected with the number of observations. It would be better to use the range of several such differences calculated for $n, n - 1, n - 2, \dots$ for $n = 10 - 12$.

3. The principle of maximum likelihood. His § 295 is difficult to understand. He drew an unspecified unimodal density curve whose behaviour was in keeping with the properties of *usual* observational errors. He called the ordinates of this curve PN, QM, ... the *true numbers of occurrences* of the corresponding errors CP, CQ, ... where C was the mode of the curve. Then Lambert supposed that those errors occurred $n, m, \dots$ times and demanded that

$$PN^n QM^m \dots = \max$$

and explained the geometric method of deriving the maximum by subtangents of the curve. The exponents were not needed. Note that Lambert introduced a continuous curve but applied notions peculiar to the discrete case.

4. The arithmetic mean. Lambert (1765a, § 320) called that mean *certainly the most secure* [estimator] if only errors of both signs were equally possible. He (1765b, § 3) added that that mean tended to the true value sought. This, the limit property of consistency, as it is called in statistics, holds for linear estimators of the parameter of location rather than of the true value. Lambert (1765a, § 441) also remarked that the use of the mean is based on its maximal probability. This is only true if the mean coincides with the mode of the frequency curve. Lambert (1765a, §§ 443 – 445) partly based all these assertions

on the comparison of the arithmetic mean with the midrange and his reasoning was not really quantitative (or not clear enough).

5. A frequency curve. As in 1760, Lambert (1765a) again classified errors according to their origin (§ 311). He then experimentally checked the properties of errors of an elementary graphic procedure (§§ 435 – 436) and derived the frequency curve of errors which occurs in pointing a geodetic instrument (§§ 429 – 430). He decided that that curve was a circumference since there was no reason for any alternative.

6. Adjustment of indirect observations. Lambert (1765b, for example § 24) fitted straight lines to sets of observational points. He arranged these points into two groups, calculated their centres of gravity and drew those lines through these centres. It is not difficult to show that Lambert's method is similar to Mayer's (§ 2.2) and the later method of Boscovich (§ 2.4).

Lambert (1765b, § 22) stated that the straight line should be constructed twice, the second time only using $n - 1$ empirical points. The difference between the lines will measure the plausibility of the obtained results^{2,2}.

7. The *Theorie der Fehler*. This term is due to Lambert (*Vorberichte* to Bd. 1 of his *Beiträge* (1765a or 1765b). Then (1765a, § 321) he defined its goals as discovering the relation between errors, their consequences (*Folgen*), the circumstances of measurement and the trustworthiness of the instruments. Neither Laplace nor Gauss ever applied this term, but Bessel (1820, p. 166; 1838b, § 9) either borrowed it or even introduced it independently. Soon afterwards Fischer (1845) used it as the title of his Abschnitt 1 and Liagre in 1853 and Airy in 1861 included the new term in the titles of their books in French and English respectively^{2,3}.

2.4. Boscovich

1. Meridian arc measurements. In 1750 – 1753 Boscovich, with another astronomer, Maire, carried out a meridian arc measurement in Italy. He then deduced the parameters of the Earth's ellipsoid of rotation by adjusting the data of several scientists.

When dealing with direct observations, Boscovich (Cubranic 1961, p. 46) sometimes calculated the arithmetic mean by a peculiar procedure. He observed the zenith distances of the same two stars at both ends of his meridian arc (Rome and Rimini) taking two measurements in Rome, but only one at Rimini. Then, for each star, he calculated two values of the latitudinal difference between the two stations, $\Delta\varphi$. He thus got its four values and tacitly assigned equal weight to each. The observations at Rimini received double weight and it is difficult to say why.

Next Boscovich computed the halfsums of $\Delta\varphi_i + \Delta\varphi_j$, $i, j = 1, 2, 3, 4$, $i \neq j$ and calculated the arithmetic mean of these values which was his final $\Delta\varphi$. Perhaps he thus qualitatively estimated the scatter of observations and hoped to isolate defective measurements^{2,4}.

Boscovich' initial method of adjusting indirect observations (Maire & Boscovich 1770, pp. 483 – 484; Cubranic 1961, pp. 90 – 91) was in this respect similar. At first he considered linear systems in two unknowns (the parameters of the figure of the Earth) by arranging his

equations into binary groups, solved each separately and calculated the mean value of each unknown over the whole set of groups. Thus, for equations i and j

$$v_i + v_j = 0.$$

A similar condition corresponds to the choice of the arithmetic mean of two direct observations. See the proper method of dealing with such groups in § 4.1.

Boscovich compared the results deduced from all ten binary combinations of five meridian arc measurements and decided that another method of adjustment was needed. The mean [the *milieu*], as he argued (1770, p. 501), should not be a simple arithmetic mean, it ought to be

Tied by a certain law to the rules of fortuitous combinations and the calculus of probabilities.

Accordingly, the conditions imposed on the residual free terms should be, as he stated,

$$\begin{aligned} v_1 + v_2 + \dots + v_n &= 0, \\ |v_1| + |v_2| + \dots + |v_n| &= \min. \end{aligned}$$

The first condition, as Boscovich argued, followed from the equal probability of errors of both signs and the second

Was necessary to approximate the observations as closely as possible.

Nevertheless, the ideas and methods related to frequency curves of the observational errors were still lacking and Boscovich was unable to say just how his conditions were *tied* to stochastic rules^{2,5}.

The first of the two conditions above can be readily met by summing up all the initial equations and eliminating one of the unknowns.

While translating Laplace's *Mécanique celeste* into English, Bowditch (Laplace 1832, § 40, Note) remarked that

The method of 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 rather say that the robustness of the Boscovich method is occasioned by its connection with the median.

2. Random sums. In an undated manuscript Boscovich studied the probability (more precisely, the chances) of the error of the sum of n observational errors which took the values $-1, 0, 1$ with equal probabilities of $1/3$. He concluded his purely combinatorial calculations by considering eight errors but he had not mentioned either a large n or a more general discrete uniform distribution or even the arithmetic mean. Add to this that Boscovich explained the formula for calculating the number of combinations and it becomes evident that he compiled his manuscript for laymen. Then, he tacitly assumed

mutually independent observations. Finally, he had predecessors, notably Galileo (1718) and even Leibniz (manuscript of 1676 published 1956). Both studied throws of dice.

Elsewhere Boscovich (1758, § 481, also see § 479) reasoned about particles of matter moving together with practically the same velocity and rather obscurely stated that the sum (not the mean) of n irregular inequalities between the velocities tended to disappear with $n \rightarrow \infty$. This was a wrong conclusion possibly explained by his previously described calculations^{2,6}. His manuscript is interesting if it predated Simpson (§ 2.5).

2.5. Simpson

In modern times long series of observations had appeared (Tycho and later astronomers, notably Bradley). While introducing his discovery of nutation Bradley (1750, p. 17) stated:

This points out to us the great advantages of cultivating [astronomy] as well as every other branch of natural knowledge, by a regular series of observations and experiments^{2,7}.

Nevertheless, some natural scientists, although hardly astronomers, held that a single experiment (not observation) can be more valuable than an entire series of them (Boyle, posth. publ. 1772, p. 376). It seems that this reasonable viewpoint was sometimes stretched. Thus, Simpson (1756, p. 82) aimed to refute

Some persons, of considerable note, have been of opinion and even publicly maintained, that one single observation, taken with due care, was as much to be relied on, as the mean of a great number.

Simpson was the first to publish a contribution on our subject, and a really important one at that.

He assumed (before Lambert, § 2.3) that errors obey a density law (and thus were random variables!). Only Poisson (1837, pp. 140 – 141) formally introduced a special and obviously temporary name for such variables (*chose a*). Then, Vasiliev (1885, pp. 127 – 131) discussed *random magnitudes* (a term still applied in Russian). On p. 133 he remarked that *random errors have all the properties of random magnitudes*.

At first Simpson supposed that the errors were distributed according to a discrete uniform law but then he went over to the discrete triangular law.

The uniform law. The values of errors

$$-v, -(v+1), \dots, -1, 0, 1, \dots, (v-1), v \quad \#$$

are equally probable. Let ε_i be the error of observation i and the number of some chances be N . Then

$$N = N(\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_n = m)$$

is the coefficient of r^m in the development of

$$(r^{-v} + \dots + r^0 + \dots + r^v)^n = r^{-vn}(1-r)^{-n}(1-r^{2v+1})^n$$

whence

$$N = C_{n+q-1}^q - C_n^1 C_{n+q-w-1}^{q-w} + C_n^2 C_{n+q-2w-1}^{q-2w} - \dots = \\ C_{n+q-1}^{n-1} - C_n^1 C_{n+q-w-1}^{n-1} + C_n^2 C_{n+q-2w-1}^{n-1}$$

Here $q = nv + m$, $w = 2v + 1$. The series continues until the binomial coefficients are still defined. The corresponding probability is of course equal to $N/(2v+1)^n$.

The triangular law. The probabilities of errors are now proportional to

$$1, 2, \dots (v-1), v, (v-1), \dots, 2, 1$$

and Simpson noted that

$$N = N(\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t = m)$$

is the coefficient of r^m in the development

$$[r^{-v} + 2r^{-v+1} + \dots + (v+1)r^0 + \dots + 2r^{v-1} + r^v]^t = \\ r^{-vt}(1-r)^{-2t}(1-r^{v+1})^{2t}$$

so that

$$N = C_{p-1}^{n-1} - C_n^1 C_{p-w}^{n-1} + C_n^2 C_{p-2w}^{n-1} - \dots$$

Here, $n = 2t$, $p = tv + m + n$ and $w = v + 1$,

$$P[\sum(\varepsilon_i/t) = m/t] \text{ and } P[(\sum \varepsilon_i/t) \leq m/t].$$

In the second instance he referred to the *method of increments* (to the summation of generalised powers such as)

$$x(x-1)(x-2) \dots (x-k+1).$$

Simpson thus effectively applied generating functions of the type

$$f(r) = \delta_{-v}r^{-v} + \delta_{-v+1}r^{-v+1} + \dots + \delta_0r^0 + \dots + \delta_{v-1}r^{v-1} + \delta_vr^v,$$

assuming that

$$\delta_k = 1 \text{ and } \delta_{-v} = \delta_v = 1, \delta_{-v+1} = \delta_{v-1} = 2 \text{ etc.}$$

The derivation of many formulas by generating functions is due to De Moivre (1711, p. 240, without proof; 1730, pp. 191 – 197; 1718/1756, pp. 39 – 41) who had not studied observational errors. Simpson himself (1740, Problem 22) studied throws of dice.

Simpson's second case is nearer to reality and he (1757) soon considered its continuous version. He included a graph of the error of

the mean but it did not have the distinctive form of the normal distribution.

The relations between De Moivre and Simpson became horrible and Karl Pearson (1978, pp. 145 and 184) justly called Simpson

A most disreputable character and an unblushing liar and a thorough knave at heart.

2.6. Lagrange

He was a pure mathematician but published a long memoir on the probability of sums and means of errors (1776). He considered a number of discrete and continuous distributions which included, for example, the cosine law. Without mentioning Simpson^{2.8} he included the cases of the uniform and triangular distributions. Lagrange (§ 18) was the first to apply the term *La courbe de la facilité des erreurs*.

Karl Pearson (1978, pp. 587 – 612) described Lagrange's memoir in detail. On p. 599 he noted that Lagrange, in his Problem 6, had evaluated a term of a multinomial *by what amounts to the Stirling theorem* and reached the normal multivariate surface^{2.9}.

2.7. Daniel Bernoulli

Johann III Bernoulli (1785) published a passage from a manuscript of Daniel Bernoulli (1769/1997) which he had received in 1769 but which was written much earlier, as D. B. had told him. There, Daniel assumed the density law of observational errors as a semi-ellipse or semi-circumference. Its radius r was ascertained by assigning a reasonable maximal error of observation and its location parameter assumed equal to the weighted arithmetic mean with posterior weights

$$p_i = r^2 - (\bar{x} - x_i)^2.$$

Here, included are the observations and the usual arithmetic mean. Successive approximations were possible.

Short (1763) first applied a weighted mean. It required a subjective choice of weights and only provided a correction to the ordinary mean which tended to vanish for even densities.

In his published memoir D. B. (1778) denied the arithmetic mean since it meant that (§ 5)

The most skilful shot will have no advantage over a blind man.

Small errors, he continued, are more probable than large ones^{2.10} and the density curve can be a semi-ellipse a semi-circumference or, as he finally decided, an arc of a parabolic curve. He did not know that the variance of the result will then change.

Bernoulli then proposed the maximum likelihood estimator of the location parameter. He (§ 9) supported his idea by indicating that, when one out of several possible and incompatible events had occurred, it should be thought that it was the most probable event.

He listed a few reasonable restrictions for the density curve but added that it should cut the abscissa axis almost perpendicularly and selected a semi-circumference with radius equal to the greatest possible for the given observer error. Then he (§ 11) wrote out the proper likelihood function but preferred to work with its square; see my remark above.

For three observations^{2.11} his likelihood equation was of the fifth degree. Bernoulli numerically solved it in a few instances. It is easy to present his equation in a form which allows to determine the maximum likelihood estimator by successive approximations. Strangely enough, the appropriate formulas are lacking in Bernoulli's memoir and even stranger since the results contradicted his own preliminary statement about the skilful shot. The astronomers of his time would not agree with weights increasing towards the tail of a distribution, so did not D. B. avoid criticism by the mentioned lack of formulas? And it was not known that such weights are expedient in the case of some distributions.

In his second memoir on the theory of errors Bernoulli (1780) studied errors of pendulum observations. Suppose that out of $2N$ oscillations of a pendulum per day, about 86,400, $(N + \mu)$ are too slow and have a period of $(1 + \alpha)$ sec whereas $(N - \mu)$ are too fast and have period $(1 - \alpha)$ sec. The total error of the pendulum per day will then be $-2\mu\alpha$ or 2 sec for $\mu = 100$ and $\alpha = 0.01$ sec.

Earlier Bernoulli (1770 – 1771) considered a similar problem in population statistics. He estimated the relative numbers of male and female births and (before Laplace and not referring to De Moivre) arrived at the De Moivre – Laplace limit theorem^{2.12}. Suppose that $2N = 20,000$ are born, m of them boys and that the probabilities of births of both sexes are the same or different. For a large number of vibrations Bernoulli derived the normal distribution of the accumulated error of the pendulum.

Finally, Bernoulli isolated two kinds of errors, systematic (*chronicarum*), almost constant, and random (*momentanearum*) whose influence is proportional to the square root the appropriate time interval. He had not considered the essence of random errors. Neither did he generalise his adopted pattern of the oscillations of the pendulum so that it would correspond to uneven probabilities of male and female births or mention the possible dependence between consecutive swings of the pendulum. But he was the first to apply the normal law in the theory of errors, to use the probable error^{2.13} and to isolate the two kinds of error^{2.14}.

2.8. Euler

Euler (1778) commented on Daniel Bernoulli's memoir of the same year. He was almost blind and misunderstood D. B. and apparently followed Bernoulli's general considerations about the skilful shot. Then, Euler (§ 6) denounced the principle of maximum likelihood. He argued that the adjustment should not depend on the possible rejection of an outlier and remarked, in § 7, that there was no need

To have recourse to the principle of the maximum since the undoubted precepts of the theory of probability are quite sufficient to resolve all questions of that kind.

Actually, his requirement led to the choice of the median. In the positive part of his commentary Euler recommended instead of the arithmetic mean an estimate which led almost to the ordinary mean. He derived a condition which was heuristically similar to the principle of least squares.

A small deviation from the mentioned condition did exist. It was occasioned by inevitable deviations of the observations from the proposed or tacitly assumed symmetrical law, and Bernoulli himself noted it when he adjusted sets of three observations.

Notes

2.1. Even in the 19th century calculations involving tricks of some kind were not at all rare. Thus, Bessel (1826, p. 229) studied the calibration of thermometers and remarked that it was too difficult to apply least squares. It would mean the solution of 26 initial equations with the same number of unknowns. Such work can be carried out by several calculators but his problem hardly warranted such efforts.

2.2. See Laura Tilling (1975, pp. 201 – 206) for a general discussion of Lambert's use of graphs in natural-scientific work.

2.3. In a letter to me of 1971 the late Egon Pearson explained why his father (1978) did not describe Lambert in his lectures:

Curiously, I find no reference to Lambert in these lectures. It was not because his writings were in German of which my father was an excellent scholar. I suppose that he selected the names of the personalities he would study from a limited number of sources, e. g., Todhunter, and that these did not include Lambert's name. Of course, K. P. was over 70 by the time his history lectures passed the year 1750, and no doubt his exploration was limiting itself to the four Frenchmen, Condorcet, D'Alembert, La Grange and Laplace.

Todhunter (1865) did refer to Lambert but had not described his work.

2.4. More interesting examples of such computations, at least in the 19th century, were those in which each observation was made by *its own* instrument(s). Thus, when measuring the relative acceleration of gravity at a station, Sabine (1821) obtained four numbers which showed the applied clock and pendulum.

2.5. Not needed.

2.6. Cf. Kepler's error (§ 1.9) concerning the total weight of a large number of coins. Even much later Helmert (1905, p. 604) warned his readers against the mistaken belief that the sum of random errors tends to vanish.

2.7. I also quote Descartes (1637, p. 63):

Je remarquais, touchant les experiences, qu'elles sont d'autant plus nécessaires qu'on est plus avancé en connaissance.

2.8. A possible but inadequate reason for this omission was the heated dispute over priority between De Moivre and Simpson: Lagrange apparently did not want to be even indirectly involved in it. De Moivre was a scholar of a much higher calibre (which Simpson clearly recognised) and 43 years the senior of them. At least on several important occasions Simpson did not refer to De Moivre. Being accused by the latter of *making a show of new rules and works of mine* (De Moivre 1743, p. xii), Simpson (1775, p. 144)

Appealed to all mankind, whether in his treatment of me he has discovered an air of self-sufficiency, ill-nature and inveteracy, unbecoming a gentleman.

For more details see Sheynin (1973a, p. 279).

2.9. In this memoir Lagrange managed to use generating functions in the continuous case and thus to pave the way for characteristic functions, to provide the *first dictionary of Laplace transforms* (Seal 1949, p. 72; Sheynin 1973a, § 2).

2.10. Karl Pearson (1978, p. 268) aptly remarked that small errors are more frequent and therefore *have their due weight in the arithmetic mean* (which therefore need not be abandoned).

2.11. The values of observations should be chosen in accord with the appropriate density law, but for three observations this remark is hardly important.

2.12. Daniel only derived the local theorem since he applied summation of limit probabilities instead of their integration.

2.13. In 1669, Lodewijk and Christian Huygens, in their correspondence with each other introduced the probable duration of life (Huygens 1669b, pp. 531 – 532 and 537).

2.14. It is evident that Ptolemy had some understanding of this subject, cf. § 1.4. I also quote a letter by D. T. Whiteside of 1972 to me:

Newton in fact but not in explicit statement had a precise understanding of the difference between random and structurally inbuilt errors. He was certainly, himself, absorbed by the second type of inbuilt error, and his many theoretical models of different types of physical, optical and astronomical phenomena were all consciously contrived so that these structural errors should be minimized. At the same time, he did, in his astronomical practice, also make suitable adjustment for random errors.

3. Laplace

3.1. Introduction

1. *General information.* Laplace was a natural scientist (an astronomer and physicist). He impressively contributed to mathematics but did not consider himself a mathematician. Thus, after applying integrals of functions of complex variables, he (1810a, p. 304) remarked that *géomètres* will hopefully get interested in this subject. In the theory of probability Laplace did not offer even a heuristic definition of a random variable or consider densities (or characteristic functions which he himself introduced) as mathematical objects in their own right. For example, when applying several formulas for deriving the density of a sum of random variables (Sheynin 1973a, § 3), he did not write out any of them, he rather solved isolated problems and never thought about solving similar problems in a similar way.

Scientists before Laplace treated probability theory as a branch of pure mathematics, but he resolutely and properly directed it to applied mathematics. Only this change enabled him to make his discoveries in natural science. But, as a result, the level of abstraction of his probability theory was not high enough and, after him, it had to be constructed anew.

2. *Ignorance of causes.* Several times in his earlier memoirs Laplace (1774; 1776, p. 148; 1781) derived laws of error acting out of ignorance. His definition of probability (which goes back to Jacob Bernoulli and definitively to De Moivre) was no better: it is circular and not a definition but a formula for calculation. True, in practice, we additionally have only the Mises approach fraught with theoretical hindrances. And Laplace (1776, pp. 144 – 145) even stated that ignorance of causes led to the origin of probability theory. Actually, that theory describes regularities inherent in mass phenomena.

Some philosophers of science tend to deny ignorance of causes as a valid starting point, to reject *the principle of indifference* (Keynes 1921, p. 44), or, as it was called previously, *principle of non-sufficient reason*. At the same time, Laplace (1798, p. 135; 1802 or 1803, p. xi) did not contradict Newton^{3.1} and was prepared to change his assumptions in the light of new observations:

On doit considerer que, les lois les plus simples devant toujours être préférées, jusqu'à ce que les observations nous forcent de les abandonner.

Telle est la faiblesse de l'esprit humaine, qu'il a souvent besoin de s'aider d'hypothèses pour lier les faits entre eux. En bornant les hypothèses à cet usage, en évitant de leur attribuer une réalité qu'elles n'ont point, et en rectifiant sans cesse par de nouvelles

observations, on parvient enfin aux véritables causes, ou du moins aux lois des phénomènes. L'histoire de la Philosophie nous offre plus d'un exemple des avantages que peuvent ainsi procurer les hypothèses, et des erreurs auxquelles on s'expose en les réalisant.

3. *Remarks on notation.* Laplace changed it from one piece of work to another and putting it in order proved too difficult. However, in several instances I introduced some notation of my own. Note also that Laplace was careless about inequalities. Thus, he did not distinguish their strict and non-strict versions.

3.2 Laplace's theory of errors

I refer readers to my later contribution (2017, § 7.2) but retain my appropriate notes.

3.3. Laplace's theory of probability

Here, I refer to the same contribution, § 7.1.

4. The nineteenth century before 1809

4.1. Solution of redundant systems of equations

The earliest method of dealing with them was to form all the subsystems of two equations and calculate the mean value of each unknown over all the subsystems. Jacobi and Binet independently proved that the least-square solution of the system was a weighted mean of those partial solutions. Actually, their proof remained valid for any number k of unknowns of a redundant system. For $k > 2$ (Mayer, § 2.2) such work was however too difficult.

4.2. A venerable problem in land surveying

Graphical intersections of a point from several stations form a polygon of errors on a sheet of a surveyor's table. A plausible position of the point naturally meant that that point ensured the minimal sum of squared distances from the lines of the intersections. However, in the 18th century (and earlier) surveyors likely selected any reasonable point in the polygon of errors.

4.3. Huber

He was a Swiss astronomer and mathematician. By 1790 he compiled several astronomical papers and soon afterwards became professor of mathematics in Basle. Merian (1830, p. 148) stated that somewhat before 1802 he discovered the principle of least squares but that

Es ging ihm wie manchem isoliert lebenden Gelehrten in kleinen Städten, dass er manchen guten Gedanken oft lange mit sich herumtrug. ... So hatte er z. B. schon in früher Zeiten durch eignes Nachdenken, die ... Methode der kleinsten Quadrate ... aufgefunden.

Later commentators from Wolf (1858) onward agreed with this testimony. However, Dutka (1990) discovered a forgotten paper (Spieß 1939) whose author had concluded that Huber did not really invent the principle of least squares. He drew on Huber's unpublished calculations and adjustments of *eine grosse vermessungstechnische Arbeit* and remarked (p. 12) that *die M. d. kl. Q. wenn gekannt, so doch nicht gefunden habe*.

Spieß (Ibidem) quoted Huber's work:

Jeder beobachtete Winkel gibt eine solche Gleichung ... und kann ... durch das Legendre'sche Maßstab der kl. Quadrate der wahrscheinlichste Wert von ... bestimmt werden.

4.4. Legendre

He (1805, pp. 72 – 73; translation by Hald 1998, p. 119) introduced the principle of least squares for solving redundant systems of linear equations:

Among all the principles that can be proposed I think there is no one more general, more exact and easier to apply than that which we have applied in the preceding researches and which consists of making the sum of the squares of the errors [of the residual free terms] a minimum. In this way, there is established a sort of equilibrium among the errors, which prevents the extremes to prevail and is well suited to make us know the system most near the truth.

Legendre also indicated that the absolute values of the extremes (of the residuals) should be confined within the shortest possible interval. This, however, was achieved by the minimax method.

Instead of describing the known conflict between Legendre and Gauss I suggest that Legendre could have remarked:

Let Gauss state whatever he wishes, but the scientific community will agree that I am the inventor of that principle.

4.5. Adrain

In 1809 rather than in 1808, as proved by Hogan (1977), Adrain justified the normal law, derived the principle of least squares and applied it to solve several practical problems. He also claimed that lack of space compelled him to postpone two other studies (until 1818), reprinted in Stigler (1980, vol. 1).

The normal law. First case. Errors x and y are the errors of the measurement of lines a and b along the coordinate axes and

$$x/a = y/b, x + y = c.$$

The first equation characterised systematic errors, the second was arbitrary.

Then Adrain supposes that for points on the same circumference

$$x^2 + y^2 = \text{Const},$$

again arbitrarily. In both cases Adrain tacitly assumed independence of errors. Kac (1939) and Linnik (1952) weakened this assumption.

Gauss also applied the idea that errors possess a density law which can be derived by issuing from their properties. And in spite of its deficiency John Herschel, Maxwell and other scientists repeated Adrain's second justification of the normal law.

Adrain also wrote out the joint distribution of two (independent) errors and noted that the appropriate contour lines were ellipses, ellipses of errors^{4,2}.

The principle of the arithmetic mean. It follows immediately. For the case of several unknowns we obtain the principle of least squares^{4,3}.

A problem in land surveying A traverse is laid out with measured sides and bearings taken by compass. Its adjustment was rather suited for systematic influences, but, interesting enough, Adrain adjusted the directly measured magnitudes rather than their functions.

In 1818 (see above)^{4,5} Adrain applied the principle of least squares to study the figure of the Earth but now he had a predecessor: Biot (1811, pp. 167 – 169) applied the same principle to study pendulum observations. Adrain revealed two mistakes in the calculations of Laplace (1798 or 1799, § 42 of Livre 3) or rather in his adjustments by the Boscovich method.

Adrain (1818a) obtained flattening $1/319$ of the Earth's ellipsoid, although, insignificantly, he calculated $(a - b)/b$ instead of $(a - b)/a$.

Finally, Adrain (1818b) indirectly determined the greater semi-axis a and calculated the radius of a spherical Earth

$$r = (2a + b)/3 = 3959.36 \text{ miles.}$$

With his flattening this means that $a = 3963.49$ miles. Assume that $1 \text{ m} = 39.370113$ inches, then $a = 6378.613 \text{ km}$ whereas for the Krasovsky ellipsoid of 1940 $a = 6378.245 \text{ km}$. Adrain's flattening was too small. For flattening $1/298.3$ (Krasovsky) $a = 6379.094 \text{ km}$, not bad either.

Abbe (1871) was likely the first European author to notice Adrain (1809). Adrain's papers of 1818 in which he cited his first contribution became more or less known. Thus, Olbers wrote to Gauss on 24.2.1819 (Schilling 1900, p. 711):

Auch ein Amerikaner schreibt sich, soviel ich an seiner ... Abhandlung schließen kann, die Erfindung der Methode der kleinsten Quadrate zu. Er ... beruft sich auf seine, doch erst 1808 herausgekommene Algebra (?).

Gauss did not comment. Strasser (1957) had not mentioned Adrain (1818a, b) which possibly meant that those papers were forgotten.

4.6. Gauss

He (1809a; 1809b, § 186) applied the principle of least squares from 1794 or 1795. In the second instance he called it

Unser Princip, dessen wir uns seit dem Jahre 1795 bedient haben and in both cases he mentioned Legendre, although not in his main memoir of 1823.

Legendre (letter to Gauss of 31.5.1809, see W-9, p. 380) was badly offended. He rightfully stated that priority is only established by publication. Gauss did not answer and Legendre (1820, pp. 79 – 80) charged him with appropriating that principle. Above (end of § 4.4) I suggested another possible Legendre's response.

I quote two commentators, May (1972, p. 309) and Biermann (1966, p. 18):

Gauss cared a great deal for priority. ... To him this meant being first to discover, not first to publish. And he was satisfied to establish his dates by private records, correspondence, cryptic remarks in publications. ... Whether he intended it so or not, in this way he maintained the advantage of secrecy without losing his priority in the eyes of later generations.

What is forbidden for usual authors, ought to be allowed for Gauss and in any case we must respect his initial considerations.

Laplace (end of § 8.8) objectively described the studied discovery but did not mention that Legendre had not justified it. Then Laplace offered his own version of the theory of errors only suitable for a large number of observations and other conditions of the CLT (whose proof he never achieved). To their detriment, other French scientists including Poisson followed him and did not even mention Gauss.

Gauss (30 Jan. 1812; W-10/1, p. 373) answered Laplace. He had applied that principle long before 1805 but had no desire to publish a fragment. Gauss' preliminary reports had appeared in the *Göttingische gelehrte Anzeigen* although Legendre hardly saw the report (1809a). It was reprinted (W-6, pp. 59 – 60) and here is an excerpt

The author has been applying the main principles which are here considered for fourteen years now and long ago had communicated to his astronomical friends. They lead to the same method which Legendre ... published a few years ago.

I (Sheynin 1999a, 1999c) described the possible cases in which Gauss could have applied the MLSq before 1805 and named many of his colleagues and friends to whom he had communicated his discovery. Unexpectedly, von Zach who allegedly refused to testify to Gauss' priority, had not until 1805 known the formulation of the principle of least squares. Furthermore, he (1813, p. 98n) indirectly agreed with the latter's statements by repeating them without any qualification:

The celebrated Dr. Gauss was in possession of that method since 1795 and advantageously applied it when determining the elements of the four new planets as it can be seen in [Theoria motus].

Regrettably, not seen there.

This passage is more important than Zach's editorial acceptance of Gauss' priority noticed by Dutka (1996, p. 357).

Gauss' priority is not universally accepted. Thus, Marsden (1995, p. 185) who nevertheless did not mention the opposite opinion of Brendel (1924) or Galle (1924, p. 9). In any case, Gerardy (1977) who drew on archival sources discovered that in 1802 – 1807 Gauss had participated in land surveying (partly for his own satisfaction) and concluded on p. 19, note 16, that Gauss started to apply the method not later than in 1803. Sorrowfully Gerardy concentrated on simple calculations and his testimony was not quite definite.

Gauss could have also applied his invention for trial calculations or short cuts. Then, mistakes in the data and weighing of observations possibly occurred as well.

Now, communication of the discovery, see also Plackett (1972). Among those informed before 1805 were Bessel (1832, p. 27) and Wolfgang Bolyai, the father of the co-founder of the non-Euclidean geometry Janos or Johann Bolyai, and Olbers about whom it was known long ago. In 1812, Olbers promised Gauss to state publicly that Gauss informed him a few years before 1805. He fulfilled his promise in 1816 since before that he had not published anything suitable for

such a mention, see the appropriate volume of the *Catalogue of scientific literature* of the Royal Society.

Notes

I heavily drew on a later source and my Notes do not reflect this fact, but I leave the Notes without any changes.

4.1. Concerning the 18th century see § 2.8.

4.2. While dwelling on the history of correlation theory Pearson (1920) and Walker (1928) considered later authors, Plana and Bravais who had studied the error of the situation of a point on a plane and in space but concluded that those later authors did not promote the emergence of the theory.

Both Pearson (pp. 185 – 187) and Walker (p. 469) attribute a similar study to Gauss (1823b; 1828) who only worked with the normal distribution in 1809 and 1816. And even then, contrary to their statement, Gauss did not introduce product terms.

4.3. Coolidge (1926) maintained that Adrain's library had included Legendre's memoir, but from when?

4.4. Adrain considered a closed transit, but it does not matter, whether closed or not.

4.5. Indeed, Hogan (1977) plausibly dated Adrain's paper.

5. Gauss

See Gauss before 1809 in § 4.6.

5.1. *Theoria motus* (1809)

This book appeared in Latin since the publisher requested a translation of its original German text. Here is the comment of Olbers (letter of 27.6.1809 to Gauss; Schilling 1900, p. 436):

Sie hatten wohl Recht, wenn Sie mir sagten, dass durch die successive Ausbildung Ihre Methode, wie sie jetzt ist, der anfänglichen Form derselben kaum mehr ähnlich ist. Auch die lateinische Umarbeitung scheint mir, so viel ich mich noch von der damals nur flüchtigen Durchsicht des deutschen Textes erinnere, noch vieles mehr vervollkommet zu haben.

The German text did not survive.

The treatment of observations occupies only a small part of the book. Thus

1) The Boscovich method. Suppose that n equations in m unknowns ($n > m$) are adjusted by that method. Then, as Gauss (1809b, § 186) remarks, exactly m residual free terms disappear and (§ 174) disapproves. He thus mastered an important theorem in linear programming. Only m equations out of the initial n are taken into consideration and the practitioner should choose the proper m equations.

Somewhat below, also in § 186, Gauss takes into account the second Boscovich condition but mistakenly attributes it to Laplace and (§§ 188 – 189) apparently agrees that the Boscovich method ensures a first approximation.

2) The normal distribution (§§ 175 – 177). Gauss considered the treatment of direct observations and formulated the principle of the arithmetic mean:

Wie ein Axiom pflegt man nämlich die Hypothese zu behandeln ... dass alsdann das arithmetische Mittel ... wenn auch nicht mit absoluter Strenge, so doch wenigstens sehr nahe den wahrscheinlichsten Wert gebe^{5.3}.

Then follows the well-known derivation of the normal distribution for a unimodal and *im allgemeinen* even density of the observational errors^{5,4}

$$\varphi(x) = h/\sqrt{\pi} \exp(-h^2 x^2), h > 0$$

where h is the *gradus praecisionis* (§ 178)^{5,6} of an observation.

To justify the applied principle of [maximal likelihood], Gauss (§ 176) proves the “fundamental principle” of inverse probability for the case of equal probabilities of the various hypotheses. However, the principle of the arithmetic mean (above) implies this restriction (Whittaker & Robinson 1924/1949, p. 219).

Possibly he was not satisfied with his derivation from the very beginning. His wording of the principle of the arithmetic mean and of the properties of the density of observational errors contain qualification remarks whereas the obtained principle of least squares became an axiom. Again, it is difficult to believe that Gauss was pleased with the appearance of a universal law of error. Later he (1821/1887, pp. 193 and 194; 1823a/1887, p. 196) remarks that his derivation depended on a hypothetically assumed distribution. And here is Bertrand’s opinion (1888a, p. XXXIV): Gauss does not claim to establish the “vérité”, he attempts to search for it. Bertrand (pp. 180 – 181) also remarks that the mean of the values of some function does not coincide with the mean value of its arguments, which, in his opinion, testifies against the principle of arithmetic mean. Gauss, however, considered direct measurements. Note also that Gauss (his letter to Encke of 1831; W-8, pp. 145 – 146) “not without interest” acquainted himself with the attempt of his correspondent to justify the arithmetic mean by deterministic analytical axioms. Many authors made similar efforts and Zoch (1935) concludes that, although they were unsuccessful, the postulate of the arithmetic mean can nevertheless be established without stochastic considerations. His finding is unrelated to the theory of errors, but those investigations apparently served as the point of departure for the theory of invariant statistical hypotheses and estimators (Lehmann 1959, chapter 6).

3) The principle of least squares follows immediately (§ 179), and it *muss überall ... als Axiom gelten*.

Helmert (§ 7.6) and Merriman (1877, p. 165) effectively remarked that Gauss, after he derived the density of errors, did not distinguish errors and residual free terms. Helmert indicates this fact but pays scant attention to it. Apparently he knew nothing about stable laws.

4) The precision of the arithmetic mean^{5,7}. In § 181 Gauss proves that, certainly in case of the application of least squares, the precision of the arithmetical mean is $[aa]$ times higher than the precision of a single observation. Here, the arithmetic mean is calculated for the case in which an unknown constant is determined by equations of the type

$$a_i x = m_i.$$

Gauss (1845/1873, p. 143) left a lesser known statement about the arithmetic mean. He remarks that the random variations corrupting observations mostly compensate one another so that the mean becomes ever more reliable as the number of observations increases. This is “generally absolutely right”, and often led to “splendid results” in natural sciences. However, Gauss continues, an important condition, often overlooked and difficult to check, is that the disordered variations ought to be entirely independent from each other.

5) The precision of a sum. In § 183, note, Gauss wrote out the appropriate formula without justifying it. He obviously considered normally distributed summands. The note was found in Gauss’s own copy of the *Theoria motus* and published only in his *Werke*.

. Suppose that

$$x = a + b + c + \dots,$$

then

$$h_x = 1 \div [(1/h_a^2) + (1/h_b^2) + (1/h_c^2) + \dots]^{1/2}.$$

Gauss does not explain his note; the terms above are probably normally distributed since he only introduced h for that law. However, he apparently derived this formula in the general case.

It proves that Gauss knew that the sum of several normal laws is normal, and, second, that its error increases as the square root of the number of its summands.

In his Second *mémoire* of 1829 Fourier provides formulas for the errors of functions of random errors. There also he mentions the error of the height of the Cheops pyramid. It had 203 steps, the square root of 203 is ca. 14, and he decides that the error is 14 times larger than the error of the measurement of a step. Fourier joined the army as a scientific advisor during the French campaign in Egypt and Syria in 1798 – 1801 and it is then that he busied himself with the height of that pyramid. Recall that in 1780 Daniel Bernoulli studied the precision of pendulum observations. There, he derived their diurnal error and forestalled the Fourier decision.

6) The precision of the [estimators of the] unknowns (§ 182; 1811, § 13). Suppose that these estimators are determined by solving a system of normal equations in accordance with the Gauss method of successive eliminations. Then, assuming that the precision of a direct measurement is unity, the precision of the estimator of the last unknown is equal to the root of its coefficient in the last reduced equation^{5,8}.

5.2. Determining the precision of observations (1816)

1) The precision of the measure of precision h . Suppose that the errors of m [independent] observations are $\alpha, \beta, \gamma, \dots$. Then the most probable value of that magnitude is determined by the condition

$$h^m \exp[-h^2(\alpha^2 + \beta^2 + \gamma^2 + \dots)] = \max$$

and is therefore equal to

$$h_0 = \{m/[2(\alpha^2 + \beta^2 + \gamma^2 + \dots)]\}^{1/2} = 1/\sigma\sqrt{2}.$$

In the last expression, which is my own, σ is the mean square error of an observation. Gauss also indicates that for the normal distribution

$$P(|x| \leq \rho\sqrt{h}) = 1/2, \text{ and } r = \rho/h$$

is the probable error which Bessel (1816, pp. 141 – 142) formally introduced.

Let

$$S_n = |\alpha|^n + |\beta|^n + |\gamma|^n + \dots, K_n = \int_{-\infty}^{\infty} x^n \varphi(x) dx.$$

Then, for large values of m ,

$$P(-\lambda \leq S_n - mK_n \leq \lambda) = \theta \{\lambda/[2m(K_{2n} - K_n^2)]^{1/2}\},$$

where mK_n is the most probable [the mean] value of S_n , Gauss treated absolute moments^{5,9}.

Helmert (1876b) and then Lipschitz (1890) proved that formula, but Cramér (1946, § 28.2) noted that it is a particular case of the CLT.

Finally, Gauss derives a formula for the absolute moments of the normal law

$$mK_n = S_{n0} = m\Gamma[(n-1)/2]/h^{n\sqrt{2}}/\pi, \Gamma(x) = \Gamma(x+1),$$

so that h (and therefore r) can be estimated by the mean value of S_n . Comparing the probable intervals of r for different n , Gauss concludes that $n = 2$ secures its best estimator.

In one of his letters of 1825 Gauss (W-8, p. 143) objects to the probable error as “depending on a hypothesis” [on the law of distribution]. Still, again in his correspondence (Sheynin 1994a, p. 261), and even in a paper (1828b), he applies it quite a few times. Natural scientists, for example Mendeleev and Newcomb, followed suit and Bomford (1971, pp. 610 – 611) “reluctantly” changed from probable to mean square error in that edition of his book. However, L. O. Struve (1887, last, unnumbered, page) proposed to abandon the probable error.

2) Denote $1/h\sqrt{2} = \alpha$ and let $n = 2$. Then

$$[m(K_4 - K_2^2)]^{1/2} = \alpha^2\sqrt{(2m)}$$

the sum of squares S_2 is distributed normally $N[m\alpha^2; \alpha^2\sqrt{(2m)}]$. This is the asymptotic chi-squared distribution, cf. Cramér (1946, § 20.2).

5.3. Theoria combiuationis (1823b)

I consider the main part of this memoir. There Gauss provides his definitive justification of the MLSq by the principle of maximum weight [of minimal variance], and I add a few words about its supplement (1828a).

Yaroshenko (1893) based the MLSq on the Bienaymé – Chebyshev inequality. It can be argued that he had not really achieved anything new, but that inequality adds a new dimension to the Gaussian approach.

1) Random errors and the density of observational errors.

Gauss (§ 1) distinguishes between random (*irregulares seu fortuiti*) and systematic (*constantes seu regulares*) errors. The former unyield to calculation, they are caused by imperfection of human organs of sense or instruments or brought about by external reasons (§§ 1 – 3). The notion of random variable was not yet formally introduced.

2) The measure of precision. Gauss (§ 6) introduces a measure of precision [the variance]

$$m^2 = \int_{-\infty}^{\infty} x^2 \varphi(x) dx$$

and calls it the mean error to be feared, *des mittleren zu befürchtenden Fehler, errorum medium metuendum* (1821/1887, p. 194; 1823b, § 7).

In his letters to Encke of 23 Aug. 1831 (W-8, pp. 145 – 146), to Bessel of 28.2.1839 (pp. 146 – 147), to Schumacher of 25 Nov. 1844 (reported by Helmert in his Introduction to Gauss 1887) and in § 7 of the *Theor. Motus*) Gauss stresses that an integral measure of precision is preferable to a local measure^{5,11}. He (1823b, § 6) also indicates that the quadratic function is the simplest [from integral measures], and in 1821 he (1887, p. 192) dwells on his choice in more detail: it is also connected with

Some other, extremely important advantages which no other function possesses. However, it is possible to select any other even degree.

Possible in spite of the advantages of the variance? Bienaymé (1853/1867, pp. 167 – 169) proves that a formula of the type below, is not valid for any other even exponent; see a clear exposition of this proof in Idelson (1947, pp. 269 – 271).

Therefore, Bienaymé continues, the choice of the variance is unavoidable. I doubt, however, that, as he believed (p. 169), Gauss was here mistaken. The sample variance is distribution-free.

3) An inequality of the Bienaymé – Chebyshev type. Gauss (§ 9) examines the probability

$$\mu = P(|\xi| \leq \lambda m)$$

for a [unimodal] density of observational errors ξ with variance m^2 and proves (§ 10) that

$$\lambda \leq \mu \sqrt{3} \text{ for } \mu \leq 2/3 \text{ and } \lambda \leq (\text{for } [2/3 \sqrt{1 - \mu}]) \mu \leq 1.$$

Cramér (1946, § 15.7 and Example 4 to Chapters 15 – 20) easier proves this “remarkable” theorem, as Gauss called it, whereas Seal (1967/1970, p. 210) indicates, that Gauss wished to abandon the universality of the normal distribution since it occurred that, anyway, $P(|\xi| \leq 2m) \geq 0.89$. But may we forget his own, although indirect arguments and doubts?

4) Independence. In § 18 Gauss offers his not quite formal definition of independent functions of observations: they should not contain common observations. Only in § 19 he adds that those functions are linear; indeed, otherwise his statement contradicts the Student – Fisher theorem for the normal distribution on the independence of the sample variance and the arithmetic mean.

Therefore, if some observation is common for two functions of observational results, the errors of these functions are not independent from one another and the mean value of their product does not therefore vanish. And in one of his examples, Gauss calculates the variance of a linear form of independent random variables.

Gauss (1809b, § 175; 1823b, § 15) mentions independence even earlier but without explanation, and, later he (1826/1887, p. 200; 1828, § 3) describes the mutual dependence of magnitudes known from observation by the existence of functional connections between them. This means, for example, that the adjusted angles of a triangle, since their sum is equal to 180° plus the spheroidal excess, are dependent on one another.

In mathematical statistics the definition of independence is different. An orthogonal transformation of independent and normally distributed magnitudes leads to their as though “adjusted” values, to their linear forms of a certain type, which are nevertheless independent (the Fisher lemma, Cramér 1946, § 29.2). Here is K. Pearson’s appropriate statement (1920/1970, p. 187) which I do not however understand: for Gauss

The observed variables are independent, for us [they] are associated or correlated. For him the non-observed variables are correlated owing to their known geometrical relations with observed variables; for us, [they] may be supposed to be uncorrelated causes, and to be connected by unknown functional relations with the correlated variables.

True, in his studies of heredity Galton was only interested in mistakes!

According to Kregel (2011), the modern notion of independence of events is due to Bohlmann whom Kolmogorov in 1933 does not mention. Kolmogorov introduces independence of events and random variables.

5) The principle of maximum weight for [unbiased] estimators^{5.13}. Gauss describes this subject ponderously. For that matter, Helmert (1872) and Idelson (1947) are in general much better understood. Suppose that, without loss of generality, the initial equations are

$$a_i x + b_i y = G_i = g_i + \varepsilon_i, i = 1, 2, \dots, n$$

where ε_i is the error of the free term g_i . The estimators of the unknowns might be represented by linear forms, for example by $x = [\alpha G]$ with unknown coefficients α_i so that

$$m_x^2 = [\alpha\alpha]m^2$$

where m^2 is the variance of an observation.

It is easy to prove that $[a\alpha] = 1$, $[b\alpha] = 0$ and the condition of maximal weight is

$$W = [\alpha\alpha] - 2Q_{11}[a\alpha] - 2Q_{12}[b\alpha] = \max$$

where Q_{11} and Q_{12} are the Lagrange multipliers. Similar considerations, and, in particular, a similar estimation of precision is also possible for the other unknowns. The weights of the estimators of the unknowns are calculated by the Lagrange multipliers of the type of Q_{ii} which, like the other multipliers Q_{ij} , are determined from the same normal equations with partly unit and partly zero free terms. Thus, $[\alpha\alpha] = Q_{11}$.

It follows that such formulas can be used even before observation: the general layout of the geodetic network and the crude values of its angles are obtained during reconnaissance and ensures the calculation of the Q_{ij} . And (what Gauss does not know) these multipliers are connected with covariations; thus, $Q_{12} = E(xy)$.

6) The estimator of the sample [variance]. Gauss (§§ 37 – 38) proves that, for n observations and k unknowns, the unbiased sample variance and its estimator are, respectively,

$$m^2 = E[vv]/(n - k), m_o^2 = [vv]/(n - k)$$

where v_i are the residual free terms of the initial equations. Instead of the mean value, the sum of squares $[vv]$ itself has to be applied. Coupled with the principle of maximal weight (of least variance), these formulas provide effective estimators, as they are now called. Gauss (1823a/1887, p. 199) remarks that the acceptance of his formula instead of the previous expression, whose denominator is equal to n , is demanded by the “dignity of science”.

Gauss stresses that the derived estimator is unbiased; however, the practically applied estimator is not m^2 , but the biased m . Furthermore, unbiased estimators do not exist in every case and some bias is allowed (Spratt 1978, p. 194). Finally, I note that Czuber (1891, p. 460) testifies that Helmert thought that $\text{var}m_o^2/m_o^2$ is more important than $\text{var}m_o^2$ by itself and Eddington (1933, p. 280) expressed the same opinion. Czuber also proves that, for the normal distribution, that relative error is minimal for the derived estimator.

7) The precision of the estimator of the sample variance. Gauss (§ 40) directly calculates the boundaries of the $\text{var}m_o^2$ by the fourth moment of the errors and indicates that for the normal distribution

$$\text{var}m_o^2 = 2m^4/(n - k).$$

He erred in calculating the abovementioned boundaries. In addition, his formulas should include the unknown magnitude $E\varepsilon_i^2$ (ε_i are the observational errors) rather than m^2 . But this formula shows that m_o^2 is a consistent estimator of the sample variance; and this property persists in the general case.

Many years later Bertrand (1888a) criticized the Gauss formula for m^2 . Tacitly assuming the normal distribution, he gives an example in which his own estimate of σ^2 is less than the Gaussian. He forgot, however, that the Gauss formula ensures an unbiased estimate whereas his own estimate was biased. Then, he calculates σ^2 but forgets the Gauss formula for the normal distribution. It is this episode that led Czuber to the discussion above.

8) Other topics. Gauss also determines the variance of a linear function of the estimators of the unknowns (which are not independent) and mentions expedient procedures for further calculations after additional data become known or after the weights of some observations are changed.

9) Another manner of adjusting observations. In the supplement (1828a) to his memoir Gauss describes the adjustment of observations by the MLsq according to the pattern of conditional observations^{5,16}. In geodetic practice, it is often expedient to issue from the directly measured magnitudes and *conditional equations* rather than from the observational equations. Sometimes both kinds of equations are used at the same time, but I leave this case aside and consider now a (later) typical chain of, say, 10 triangles of triangulation. Each angle is measured as are the lengths of two extreme sides (baselines) whose directions (azimuths) are determined by astronomical observations. The observational errors are such that both the baselines and the azimuths may be considered exact; only the angles are adjusted. Each measured angle q_i provides an equation

$$x_i - q_i = v_i$$

where the first term is the true value of the angle and the right side is the sought correction. Now, the condition of closing the first triangle (I disregard its excess) is

$$x_1 + x_2 + x_3 - 180^\circ = 0.$$

Extremely simple is also the condition that demands that the azimuth of the first baseline plus the algebraic sum of the appropriate angles is equal to the azimuth of the second baseline. The sine theorem is needed for the transition from the first baseline to the second one, but a first approximation is achieved by introducing the measured angles so that the required trigonometric equation is linearized. All the conditions can be written as

$$[av] + w_1 = 0, [bv] + w_2 = 0, \text{ etc.}$$

They should be exactly fulfilled and the number of the terms in the square brackets is either three, or more. That depends on the number

of the triangles in the chain. The adjustment proper consists in determining the conditional minimum of $[vv]$ with the usual application of the Lagrange multipliers and the corrections v_i are determined through these multipliers. Strangely enough, only Helmert (1872, p. 197) was the first to explain this.

10) **A new exposition of the memoir** (Sheynin 2012a; 2014). Gauss was able to derive formulas of the sample variance in the very beginning of the memoir, just after he introduced the variance: the required conditions (linearity of the initial equations, (physical) independence of their free terms and unbiasedness of the estimators of the unknowns) are not connected with the further exposition. And the MLSq directly follows from those formulas! Hundreds of textbooks described the MLSq as justified by Gauss in 1809: such an approach is incomparably easier. Now it is quite possible to follow the memoir of 1823. Its very existence was barely known, see Eisenhart (1964, p. 24) above. Even previously Fisher (1925, p. 260) involuntarily testified as much: the MLSq

Is a special application of the method of maximum likelihood.

The most eminent scientists (Boltzmann 1896/1909, p. 570; Chebyshev) were barely acquainted with the work of Gauss.

Many authors beginning with Gauss himself derived the formula of the sample variance which is not difficult. The main point, however, is that the proof does not depend on the condition of least squares. On the contrary, this condition can now be introduced at once since it means minimum variance. The Gaussian formulas for constructing and solving the normal equations and calculation of the weights of the estimators and of their linear functions are still useful.

Gauss twice justifies least squares (of which I only left the second one), but why does not he even hint at this fact? I can only quote Kronecker (1901, p. 42) and Stewart (Gauss 1823b – 1828/1995, p. 235):

The method of exposition in the “Disquisitiones [Arithmeticae”, 1801] as in his works in general is Euclidean. He formulates and proves theorems and diligently gets rid of all the traces of his train of thoughts which led him to his results. This dogmatic form was certainly the reason for his works remaining for so long incomprehensible.

Gauss can be as enigmatic to us as he was to his contemporaries.

Gauss himself said so. His eminent biographer, Sartorius von Waltershausen (1856/1965, p. 82) testifies: He *used to say* that, after constructing a good building, the *scaffolding* should not be seen. And he often remarked that his method of description *strongly hindered* readers *less experienced* in mathematics.

Finally, I note Gauss' words (letter to W. Olbers 30.7.1806): *Meine Wahlspruch [motto] ist aut Caesar, aut nihil.*

The second substantiation of the MLSq can be accomplished by the notions of multidimensional geometry (Kolmogorov 1946; Hald 1998,

pp. 473 – 474). Nevertheless, the new exposition of the memoir of 1823 is essential, and it appeared more than 200 years after its publication!

Kolmogorov (p. 64) also believed that the formula for m^2 is its definition. Much earlier Tsinger (1862, § 33) stated that it already “concealed” the MLSq. This however, was only a hint at the real possibility of understanding Gauss. Harter (1977, p. 28) states almost the same.

5.4. Practical considerations

1) The number of observations. Given the required precision, how many observations were necessary? Formulas of the theory of errors ignore systematic errors. Therefore, only after concluding all work (for example, the work necessary for a chain of triangulation) will the practitioner understand something about them. But, if its angles are measured a definite number of times (determined by trial and error) with a certain type of instruments, under favourable conditions and in an *usual* region, and certain regulations are met, a certain degree of precision can be expected. And so, by the end of the 19th century or later some countries introduced rigid programmes of observation for primary triangulation (Bomford 1971, p. 24).

Gauss had to manage otherwise, and manage he did. He continued to measure each angle until becoming sure that no more work was needed. Incidentally, Schreiber (1879, p. 141) confirmed this statement by referring to the *Protokollen* of Gauss. Some of these *Protokollen* are reprinted in the *geodetic* Band of the *Gauss Werke*, in Bd. 6 if I remember correctly.

Gauss’ former student, Gerling (1839, pp. 166 – 167) kept to the same attitude. He indicated that after some work the observer convinces himself that

Jedes weitere Fortsetzen ... nur verlorenen Arbeit sein würde.

Then, Bessel, in a letter to Airy of 1833, see his *Abhandlungen*, Bd. 3, 1876, pp. 462 – 465, noticed:

Continual oscillations within the limits of unavoidable imperfections are ... agreeing with the very nature of results derived from observations.

Clarke (1880, pp. 18 and 52), Dorsey & Eisenhart (1969, p. 53) agree that the number of measurements should not exceed a certain boundary. They, as well as Cournot (1843, §§ 132 and 138), justify this fact by the presence of unavoidable systematic errors (and some dependence between observations).

2) Precision of observations. The best formula estimates the sample variance but in any case $[vv]$ is a random variable that somewhat differs from its expectation which, strictly speaking, is meant in that best formula. For this reason (and especially if the number of observations is small), at least in one instance Gauss (letter of 17.4. 1844 to Gerling; Schäfer 1927, p. 687) added together his observations at several stations and calculated a single value for the sample variance. He explained that, without any reasonable cause, the variances for separate stations essentially differed. In other letters to Gerling and Bessel Gauss noted that for a small number of

observations his formula was untrustworthy. Modern authors (Ku 1967, p. 309) concur.

3) Rejection of outliers. In modern times, many scientists attempted to study the ensuing problem. Laplace (1818, p. 534) reasonably stated that

Pour appliquer avec succès les formules de probabilité aux observations géodésiques, il faut rapporter fidèlement toutes celles que l'on admettrait si elles étaient isolées, et n'en rejeter aucune par la seule considération qu'elle s'éloigne un peu des autres.

Gauss (letter to Olbers of 3.5.1827 /W-8, pp. 152 – 153) indicated:

Zu einer erfolgreichen Anwendung der Wahrscheinlichkeitsrechnung auf Beobachtungen ist allemal umfassende Sachkenntnis von höchster Wichtigkeit. Wo diese fehlt, ist das Ausschließen wegen größerer Differenz immer misslich, wenn nicht die Anzahl der ... Beobachtungen sehr groß ist. ... Halte man es wie man will, mache aber zum Gesetz, nichts zu verschweigen, damit andere nach Gefallen auch anders rechnen können.

Gauss also added that rejection can lead to overestimation of the precision of observations.

Here, now, is a new viewpoint of Barnett & Lewis (1978, p. 360), the authors of a book on this subject:

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

4) Calculations. Owing to his exceptionally convenient notation and his apt method of successive elimination of the unknowns, Gauss was much better able to solve his systems. Without even an arithmometer, Gauss carried out difficult calculations; once he solved a system of 55 normal equations (letter to Olbers of 1826; W-9, p. 320). For other examples see Sheynin (1979, p. 53). His preparatory work (station adjustment; compilation of the initial equations and of the normals themselves) was very considerable as well.

Sometimes Gauss applied iterative calculations, a non-cyclic one-step process (letter to Gerling of 1823; W-9, pp. 278 – 281), also see Forsythe (1951) and Sheynin (1963). The first to put on record this fact, in 1843, was Gerling himself. Then, Gauss (1809b, § 185) left an interesting qualitative remark. He stated that “it is often sufficient” to calculate approximately the coefficients of the normal equations. The American astronomer Bond (1857) and Newcomb (1897a, p. 31) applied Gauss’ advice.

As a calculator of the highest calibre (Maennchen 1918/1930, p. 3),

Gauss was often led to his discoveries by means of mentally agonizing precise calculations [...]; we find [in his works] substantial tables whose compilation will in itself occupy the whole working life of some calculators of the usual stamp.

Gauss made some mistakes in his computations possibly because, first, he did not invariably check them, see for example Gerardy (1977) or his own methodological note (1823c) where the signs of dx and dy were wrong. Second, Gauss calculated “unusually fast” (Maennchen 1918/1930, p. 65ff). This caused mistakes and additional difficulties in proving that he applied the MLSq before 1805.

Maennchen did not study Gauss' geodetic calculations possibly because in his time the solution of systems of linear equations did not yet attract the attention of mathematicians.

When compiling a certain table of mortality, Gauss (W-8, pp. 155 – 156) somehow calculates the values of exponential functions b^n and c^n for $n = 3$ and $7(5)97$ with $\lg b = 0.039097$ and $\lg c = -0.0042225$.

Notes

Notes 1, 2, 5, 10, 12, 14, 15 not needed

5.3. Later Gauss (1845, p. 143) remarked that, for independent observations, the application of the arithmetic mean was *in allgemeinen vollkommen richtig* and had resulted in *glänzende Resultate* in natural sciences.

5.4. This is Gauss's indirect definition of (usual) random errors.

5.6. When determining the constant in the normal law, Gauss remarked that the appropriate integral was first calculated by Laplace and repeated this statement in § 182. Legendre, in his letter of 1809 to Gauss, noted that that calculation was rather due to Euler. For Gauss, it was a bit late for inserting a correction in the text of *Theoria motus*.

5.7. In 1809, Gauss naturally considered normally distributed errors. In 1823, he generalised his formulas.

5.8. Gauss noted the possibility of estimating the precision of all the unknowns by repeated solutions of normal equations. *Einige Rechner* applied this procedure (Gauss 1823b, § 31) but (1823a, p. 197) that, for a large number of unknowns,

Durch dieses kunstlose Verfahren ... in Rücksicht auf Kürze der Rechnung nichts gewonnen wird.

5.9. Gauss applied absolute moments so that his K_n is wrong. And mK_n is the mean rather than the most probable value of S_n .

5.11. Gauss voiced the same opinion in some of his letters, especially to Bessel on 28.2.1839 (Plackett 1972, p. 287):

I must consider it less important in every way to determine that value of an unknown parameter for which the probability is largest, although infinitely small, rather than that value by relying on which one is playing the least disadvantageous game.

5.13. Unbiasedness is a feature of the entire Gaussian theory of errors.

5.16. The approach to adjustment by the method of conditional observations is extremely important for practitioners although it does not contain any essentially new ideas and Gauss (1828) described it.

6. From Gauss to Helmert and beyond

I also dwell on Helmert in § 7, but it became necessary to go here beyond him.

6.1. Some new work

1) Bessel. I repeat that he picked up the term *Theorie der Fehler* and formally introduced the probable error. He, or at least one of his students (Rosenberger 1828) described his method of adjusting indirect observations connected by conditions. This awkward procedure presents no essential innovations, but Bessel had thus initiated adjustment for observations separated into groups.

Bessel (1838 a) attempted to prove the CLT. In § 10 he listed 13 independent sources of error inherent in measuring the zenith distance of a star and thus illustrated the possibility of applying this theorem to observations themselves^{6.1}.

In § 2 Bessel noted the existence of an observational error with an antimodal density. He did not thus refute the (possible) normality of

the total error of observation and neither did practitioners notice this error.

In § 7 Bessel proved that the normal law is stable, i. e., that the sum of two (and therefore of any finite number of) normally distributed random variables is again normal. Laplace and Gauss knew it but had not provided the appropriate formula, let alone justify it^{6.2}.

In § 11 Bessel argued that astronomical observations are normally distributed. He presents three series of Bradley's observations, 300, 300 and 470 in number, and states that their errors almost precisely obey normal distributions. However, modulo both large and small errors occurred there more often and intermediate errors less frequently than required by the appropriate normal laws.

He is wrong and it is difficult to believe that he is mistaken (especially see below). Moreover, he thus misses the opportunity to discover an example of long series of not quite normally distributed errors of precise observations. Later, scientists, in the first place Newcomb (1886), gradually discovered such series.

Bessel's contribution included a proof of a version of the CLT (only Liapunov and Markov rigorously proved it). Bessel states that, given more observations, the deviation from normality disappears. Did not he notice that he thus undermines the essence of that theorem? Did not he formulate his wrong conclusion to save that proof (and thus to achieve the impossible)? Many commentators up to Idelson (1947, § 33.1) repeated his celebrated inference without noticing the forgery.

Cauchy (1853a, 1853b) suggested a method for solving linear equations, see Linnik (1958, § 14.5). He calculated the efficiency of the appropriate estimators for the cases of one and two unknowns^{6.3}. Idelson (1947, § 21) noted that the method of Cauchy is applied (at least?) in two national time services.

2) Bienaymé. He (1852) indicated that the estimators of the unknowns should be jointly efficient. Then, he (1853, p. 313) declared that the choice of the mean square error as a measure of precision was not arbitrary *comme le croyait Gauss*, but with respect to Gauss that remark is most likely wrong. Heyde & Seneta (1977) described Bienaymé's contribution to statistics.

6.2. Physics, chemistry, meteorology

As seen above, the theory of errors originated and developed in response to the needs of astronomy and geodesy. Lambert's *Photometria* (1760) was an outstanding exception. Physics and chemistry followed. Paucker (1819) devoted an elementary booklet to the MLSq and Strecker (1846) applied that method for the determination of the atomic weights of two elements^{6.4}. Clausius described the MLSq in a manuscript of 1857 (Schneider 1975, p. 248, note 28) and Maxwell (1869) applied error-theoretic considerations while studying the distribution of molecular velocities. Fechner (1860) founded psychophysics and thus introduced the statistical method into physics although not in the crucial direction. He also somewhat furthered the error theory. Not later than in 1870 (Kohlrausch) elements of the error theory were introduced into *practical physics*.

Below, I describe the situation in the three *new* branches of natural science.

1) The number of observations (see also § 5.4). Recall (§ 2.5) that Boyle contrasted the quality and the number of experiments and believed that the second factor was hardly important. In spite of Simpson's efforts (Ibidem) this opinion or tradition of choosing the *best* observation persisted in some branches of natural science^{6.5}. Joule (1849) determined five values for the mechanical equivalent of heat from equations of the type $ax = l$ but used only one of them. The chosen value belonged to the experiment with the maximal number of observations and, at the same time, the maximal (therefore, the best) value of a ^{6.6}.

Mendoza (1991, p. 283) ignorantly downgraded astronomy and geodesy as compared with physics. Note however that a triangulation is measured only once whereas constants of natural science were determined in at least several places.

2) Weighing of observations. Observations (say, s in number) made under the same conditions will probably be corrupted by systematic errors in much the same way and it is unwise to assign weight to their arithmetic mean equal (or proportional) to s . Metrology is an exception (Eisenhart 1963/1969, p. 31): the unavoidable tiny systematic error is included in the standard. Below, I provide a sufficient account of his statement.

And so, measure under variable conditions, see Gerling's attitude in § 5.4. A proper procedure is needed when processing pendulum observations made at almost the same latitude: they can be combined and given the summary weight but only if their longitudes differ.

Mendeleev (1872, p. 144) determined the empirical relation between the density of a gas and its pressure (thus attempting to refine the Boyle – Mariotte law) and preferred to

Make a few but precise and repeated observations at several significantly different pressures. ... Amassing observations made at various closely spaced pressures not only presents many obstacles, but in addition increases the errors of [determining the unknowns].

Mendeleev apparently thought that some systematic errors depended on the value of the pressure.

3) Dependence between observations. Meteorology presents obvious difficulties. Correlation theory did not owe its birth to this science^{6.7}, but it is opportune to note Lamont (1867, p. 245):

Ich schon vor dreißig Jahren angefangen habe, die Unterschiede gleichzeitiger Beobachtungen [made in different localities] an die Stelle der gewöhnlichen meteorologischen Constanten zu substituieren, und ich habe jetzt noch die Überzeugung, dass dieser Weg der geeignetste ist, um die Meteorologie als mathematische Disciplin auszubilden^{6.8}.

In the 19th century it became known that densities in meteorology were asymmetric and Meyer (1891, p. 32) decided that the theory of errors is not applicable to meteorology. However, mathematical statistics does not leave aside the treatment of asymmetric series of observations, and K. Pearson (1898) used Meyer's material to illustrate his theory of skew curves. Meyer could have also argued that the scatter of a meteorological element was not random.

4) Ignorance. Gauss's second justification of the MLSq remained hardly known as Eisenhart (§ 5.3-10) noted in 1964. Worse: many natural scientists and even mathematicians were hardly familiar with the theory of errors. Thus, in 1826 – 1830 Ivory published a number of papers^{6,9} on the adjustment of pendulum observations and only gradually mastered the MLSq. But, most importantly, the Gauss classical formula or the sample variance was often described somewhat wrongly (Chebyshev 1880/1936, pp. 249 – 250).

To ignorance I also ascribe a few badly studied methods of the adjustment of direct observations, such as based on the use of their range or of some of its modifications, i. e., on the neglect of almost all the available data^{6,10}.

6.3. The normal law

1) The success. For many decades the normal law had been regarded as the law of error. The first Gaussian justification of least squares (1809) was incomparably easier than the second (1823) and led to the normal law. That same law became entrenched in natural science including anthropometry, the exponential function was handy and more or less reasonably described the scatter of observational errors and the CLT, in spite of its poor proof, lent it an air of respectability. Finally, it was stable.

Then, Maxwell (1860)^{6,12} stated that the distribution of molecular velocities, appropriate to a gas in equilibrium, is normal. In astronomy, from the mid-19th century to 1896, the stellar motions were thought to be normally distributed (Sheynin 1984a, § 8.4). And on Bessel's fraudulent inference (§ 6.1) astronomical observations were held to be normally distributed.

In natural science, the normal density appeared also as the law governing the errors *faites par la nature* (Quetelet 1853, pp. 64 – 65). He mentioned chest measurements of soldiers^{6,13}. And in 1873 he maintained that the normal law was *une des plus générales de la nature animée* (Sheynin 1986, p. 313).

Many observations in natural sciences cannot be described by the normal curve, but they are not random in the usual sense (cf. Meyer's possible reasoning above). Here is Quetelet (1846, p. 168) on atmospheric pressure:

L'abaissement du mercure au-dessous de la moyenne est en général plus grand que son élévation au-dessus de ce terme. Les exemples où la courbe de possibilité perd de sa symétrie, sont assez fréquents; ils méritent d'autant plus d'être étudiés, que ce défaut de symétrie tient toujours à des causes plus ou moins curieuses, dont on peut apprécier l'influence.

This means that the happy-go-lucky Quetelet then considered the normal law as the general rule. Not surprisingly, several authors commented on the popular belief in the normal law of error. Thus, Poincaré (1912, p. 171) repeated an oral jocular statement of G. Lippmann, an author of a treatise on thermodynamics:

Les expérimentateurs s'imaginent que c'est un théorème de mathématiques, et les mathématiciens que c'est un fait expérimental.

Lippmann described the prevailing attitude^{6,14}.

2) The opposition. Bessel (§ 6.1) missed the opportunity to oppose the normal law. Bienaymé (1853, p. 313) declared that the exponential *N'est qu'un moyen d'approximation très-commode, mais qui pourrait être remplacé par d'autres formules.*

But it was Newcomb (1886, p. 343) who initiated the real opposition. He argued that the cases in which the errors follow the normal law were *quite exceptional* and that (p. 345) *in certain classes of important observations* the proportion of large errors was so great that

No separation into normal and abnormal observations was possible.

He mentioned his earlier contribution (1882, p. 382):

That any general collection of observations of transits of Mercury must be a mixture of observations with different probable errors was made evident to the writer by his observations of ... 1878.

While offering his celebrated chi-square test, Pearson (1900, p. 353) harshly commented on the contemporaneous treatment of astronomical and geodetic observations (and target shooting). He mentioned *current textbooks of the theory of errors* and noted that the normal law was usually derived analytically (apparently, by applying the CLT) and that the authors

Give as a rule some meagre data of how it fits actual observations. ... Perhaps the greatest defaulter in this respect is the late Sir George Biddell Airy.

6.4. The normal law modified

I treat the attempts to improve on the normal law. At least in some cases a new universal law of errors was vainly aimed at. Only one author used a frequency curve of the Gram – Charlier Type A and even he later abandoned it.

Cournot (1843) was the first to discuss a series of observations of unequal precision. If n_i observations have densities $f_i(x)$, then (§ 81) the density $f(x)$ of the entire series of observations will be the generalised arithmetic value of those densities. Cournot had not specified either their type or the differences between them.

Density $f(x)$ is not a sum but a mixture of densities, and even if the individual densities are normal, it is not stable.

De Morgan (1864) generalised the normal law but his attempt was unbelievably wrong.^{6.15, 6.16}

From 1873 to 1887 several authors (Peirce; Stone in *Monthly Notices Roy. Astron. Soc.*, 1873 – 1874; Glaisher, *Ibidem*, 1874; Edgeworth, *Phil. Mag.*, 1883 and 1887; Newcomb) stated that a series of observations can obey normal laws with differing measures of precision^{6.17}. They did not mention Cournot. Harter (1977) described their ideas and efforts and I only discuss Newcomb (1886, p. 351). He adopted *a very probable hypothesis*, that the law of error was a mixture of normal distributions with various measures of precision which occurred with respective probabilities. The parameter h of the ensuing normal law became a discrete random variable but his proposal required subjective decisions.

Newcomb also introduced some simplifications and Hulme & Symms (1939, p. 644) noted that they led to the choice of the location parameter by the principle of maximum likelihood.

Two authors modified Newcomb's proposal (Lehmann-Filhés and Ogorodnikov) but their efforts only amounted to mathematical exercises. True, at the same time Ogorodnikov made some interesting remarks. Thus, he (1928, p. 16) noted that no observational series with a negative excess were known. On the contrary, Kemnitz (1957) found many such series of geodetic observations and explained that official instructions required rejection of some outliers which led to the appearance of truncated normal laws.

Pearson (1894) did not mention Newcomb but investigated a related problem, the dissection of abnormal densities into normal curves. He (p. 74) proved that

A curve which breaks up into two normal components can break up in one way, and one way only.

But the dissection required the solution of an equation of the ninth order.

Eddington (1933, p. 277) quite simply proved that the excess of the Newcomb's distribution was positive (and, therefore, in particular, not normal. Idelson (1947, p. 309) called Eddington's theorem (and wrongly, Ogorodnikov's proposition)

One of the most important results of the contemporary theory of errors.

6.5. The theory of errors and statistics

1) The true value of a measured constant. This is a usual expression. Fourier (1826/1890, p. 534) defined it as the limit of the arithmetic mean of the appropriate observations (actually, the mean of a large number of them). Many authors independently and without recalling Fourier introduced the same definition and Eisenhart (1963/1969, p. 31) formulated the unavoidable corollary: the mean residual systematic error had to be included in that true value.

Mathematical statistics (Fisher 1922, pp. 309 – 310) introduced instead the notions of consistency, efficiency and sufficiency of statistical estimators, but he himself, and Gauss and Hald, to mention only them, continued to apply the ineradicable true value as well.

Just as the theory of errors, theoretical (not mathematical) statistics studies systematic errors or, more generally, structures in the data. Thus, Halley, in 1701, drew lines of equal magnetic declinations for Northern Atlantic; Humboldt, in 1817, introduced isotherms and originated climatology^{6.18}; and Galton, in 1863, discovered the existence of anticyclones. Statistics therefore studies mean conditions or states, and even vague notions as the number of yearly births in a nation^{6.19}. A new term, theory of means (Sheynin 1986, pp. 311 – 312) had appeared^{6.20}. One of the last authors to apply it was Hilbert (1901, § 6).

2) Frequency curves. At the turn of the 19th century Pearson and his associates began to fit frequency curves to observations and therefore to derive the parameter of those curves rather than to ascertain mean states (much less, to determine true values). This new approach was in keeping with the general development of mathematics: it introduced a

new mathematical object and new terminology. Thus, the arithmetic mean estimates the parameter of location of the appropriate density; that mean became *sample mean*, mean square error was replaced by *standard deviation*, but the Gaussian apt notation $[ab]$, $[bc1]$ etc. was regrettably forgotten. That, however, mostly happened since Laplace conceitedly refused to apply it.

3) Correlation theory. Its development further estranged error theory from statistics. Here, however, is a walk on thin ice. According to Gauss (§ 5.3-4), dependence between observations is caused by the presence of their common errors and Pearson (1920, p. 199) referred to Galton and declared that

Correlation must be the consequence of the variations of the two organs being partly due to common causes,

And Eisenhart (1978, p. 382) reasonably stated that much of the *Mathematical machinery that Gauss had devised ... was immediately applicable to correlation analysis.*

But then, statisticians are more likely to consider relations of cause and effect.

Kapteyn (1912) wished to quantify the connection between two functions with partly coinciding observations, mostly for astronomy, but his paper was ignored or at least forgotten. Nevertheless, his, or actually Gaussian opinion was (and is?) independently and even intuitively pronounced by a number of geodesists.

4) Bridging the gap between the theory of errors and statistics? Hardly possible since the distribution of observational errors is generally unknown.

Notes

6.1. A bit earlier Hagen (1837) introduced *the theory of elementary errors*. He (p. 34) thought that there were infinitely many of them, all modulo equal with equal possibility of both their signs. I doubt whether his theory was useful.

6.2. Bessel (§ 3) remarked that the random variables should be independent but did not repeat this remark when proving his theorem. Several authors confirmed his forgotten (unnoticed?) remark and Seidel (1863, p. 326) formulated it without proof and wrote out a wrong formula for the measure of precision of the sum of two normal laws.

Czuber (1903, p. 23) referred to Pizzetti and Lindelöf and proved that *Wenn die unabhängigen Beobachtungsfehler ... einzeln das [normal] Gesetz befolgen, so unterliegt eine homogene lineare Funktion ... derselben einem Gesetz der gleichen Form.*

And this proposition is *ein Hauptsatz* of the theory of errors. Sampson (1913, p. 170) repeated the proof of that proposition, cited two predecessors and attached some importance to the *reproduction of form* but was unable to define *form*.

6.3. Denote the estimators of the same magnitude obtained by some method and by the MLSq by α and β . Then the efficiency of the former estimator is the ratio of their variances, less or equal to unity, cf. § 5.3-5.

6.4. Somewhat earlier Courmot (1843, § 136, note) suggested to apply *Cette théorie à la détermination des poids atomiques ou des équivalents chimiques.*

He obliquely specified that theory as the theory of errors.

6.5. Cf. also Mill (1843/1886, p. 353):

A very slight improvement in the data by better observations, or by taking into fuller consideration the special circumstances of the case, is of more use than the most elaborate application of the calculus of probabilities founded on the data in their previous state of inferiority.

But why not say that both approaches were needed?

6.6. Joule possibly selected the *best* experiment beforehand and chose the number of observations accordingly. The rejection of four cases was then of no importance.

Cf. Mendeleev's (1895, p. 159) later statement:

When, however, one of the numbers certainly ensures a higher guarantee of precision than the other ones, it alone should be taken into account, leaving the numbers, certainly reflecting either worse conditions of experiment and observation or any cause for doubt, without any notice. ... To consider worse numbers taking them with some weight is tantamount to deliberately corrupting the best number.

Cf. end of § 1.3.

6.7. Without introducing any measure of correlation, Seidel (1865; 1866) made the first ever quantitative stochastic study of the dependence between several factors (number of cases of typhoid fever, the level of subsoil water and, in 1866, in addition, on precipitation). Weiling (1975) first noticed his work. See Sheynin (1982, §§ 7.4.2 – 7.4.3).

Perhaps even earlier statisticians used to apply a common sense procedure for the same purpose. Dependence was ascertained if numbers varied steadily in response to an increase or decrease of an argument. Mortality from amputations increased with the number of beds (with the worsening of conditions) in hospitals (J. Y. Simpson 1869 – 1870, p. 399) and the probability of the defendant's conviction depended on his social status and education (Quetelet 1836, t. 2, p. 313).

6.8. Quetelet (1849, chapter 4, p. 53) remarked that observations at three stations in Brussel and vicinity

Donnent des differences ... qui procèdent dans un ordre conforme à celui que la théorie des probabilités assigne aux erreurs accidentelles.

6.9. In a letter of 15.3.1827 to Olbers Gauss (Schilling 1909, pp. 475 – 476) called Ivory a *scharfsinnigen* (acute) mathematician.

6.10. Statisticians have formulas for calculating non-parametric confidence intervals for the population median which take into account all the observations.

6.11. At least in Russia geodesists were familiar with the second Gaussian justification of the MLSq since Markov resolutely supported it. However, at the same time Markov (1899/1951, p. 246) contended that that method had no optimal properties, so its justification was to no avail!

Neyman (1934, p. 595) mistakenly attributed to Markov the second Gaussian justification of least squares and David & Neyman (1938) even proved an *extension of the Markov theorem* again due to Gauss. Finally, however, Neyman 1938/1952, p. 228) noticed his mistake. Nevertheless, the mysterious *Gauss – Markov theorem* is still alive and kicking. Scheffé (1959, p. 14) invented that term although Plackett (1949, p. 460) earlier noticed Neyman's mistake.

6.12. His proof was somewhat defective, see § 4.5.

6.13. Cf. Edgeworth (1885a, p. 140):

Observations are different copies of one original; statistics are different originals affording one generic portrait.

6.14. Laplace was prepared to recognise the normal law as the law of error only because he superficially traced the proof of the CLT.

6.15. It was Pearson (1894, p. 93) who introduced the term excess. In the same contribution he also defined *standard deviation* and *normal curve* (pp. 75 and 72).

6.16. De Morgan introduced a frequency which took negative values and considered an event with probability 2.5 but these facts did not bother him at all. Much worse (Sophia De Morgan 1882, p. 147) in a letter (date unknown) he maintained that the sine and cosine of infinity vanish and tangent of infinity is a dual complex number.

6.17. Edgeworth (1883): he applies a special term, *probability curve*, for the normal density but, in general, densities are *facility-curves*. Then, any estimator is *evil* (p. 361), a term adopted by Newcomb (1886) in a restricted sense. Lastly, Edgeworth somehow believes that one of the integrals of two densities over $[0, x]$ is larger than the other for every value of x .

6.18. No less than eight other natural scientific disciplines connected with statistics have emerged in ca. 1815 – 1915: geography of plants; public hygiene (the predecessor of ecology); stellar statistics; epidemiology; zoogeography; psychophysics; kinetic theory of gases; and biometry.

6.19. Just two questions. Does *nation* include foreigners living in the country or

its own nationals residing abroad? How to allow for the difference between registration figures and actual births?

6.20. Cf. the title of Hauber's study (1830 – 1831) of the error theory!

7. Helmert

For many decades Helmert's treatise (1872) remained one of a very few best sources for a study of the previous state of the theory of errors and of the MLSq. It was also important for mathematical statistics (Sheynin 1995). Not without reason the third, posthumous edition of his treatise appeared in 1924. By 1907, when its second edition was published, Helmert's interests shifted to the study of the figure of the Earth.

7.1. Rejection of outliers

Jordan (1877) approximated the normal law of error by an even trinomial on $[0, M]$. He calculated all the three parameters in terms of M . Other simple calculations led to an approximation of M by the variance m of the trinomial, $M = 2.65m$, and Jordan suggested that observations deviating more than $3m$ from their mean be rejected. That was the celebrated *three-sigma rule*.

Helmert (1877) objected. Mainly, he (p. 143) argued that M depended on the number of observations n , but that he accepted that rule for $n = 10 - 100^{7.1}$.

7.2. Revealing systematic errors

Helmert (1872, p. 257; 1875c, pp. 147 and 151; 1905) recommended to single out the mean value of the systematic error. To this end, he (1905) proposed several tests by tacitly adhering to the normal law.

7.3. The Abbe criterion for same end

Abbe (1863, pp. 80 – 81) introduced a test for revealing systematic influences. He considered two functions of the errors of observation and their ratio μ which was sensitive to *gesetzmässig wirkenden Ursachen*. Helmert (1905, changed Abbe's statistic μ in two different ways and introduced residuals instead of errors. Adjustment of observations by the MLSq means that its condition $[vv] = \min$ is a *Zwang, wie ein systematischer Fehlereinfluß*. Its investigation proved difficult and Helmert restricted it to direct observations and the normal law.

7.4. Sums of natural powers of errors

Helmert (1875a; 1876b) studied powers m of such sums of n terms for observational errors distributed normally or uniformly or arbitrarily in the asymptotic case. In his main article of 1876^{7.2} he considered

An arbitrary distribution and a finite n ;

The uniform distribution for $n = 1$ and 2 and $m = 1, 2, 3$. The case of $n = 2$ was again difficult.

The normal distribution, $m = 2$ and finite n

He then derived the chi-squared distribution by induction, see Hald (1960, pp. 258 – 261). Pearson (1931) took pains to note that Helmert had preceded him in the derivation of the chi-squared distribution, but did not mention Helmert's most appropriate contribution (1876b).

A limit theorem for an arbitrary distribution.

Helmert followed Poisson (1837, p. 267) and Glaisher (1872), both of whom had considered the case of $m = 1$. He obtained the Gaussian limit theorem (§ 5.2-2) but not conclusively.

Limit theorems for the uniform and the normal distributions.

The measure of precision for the normal distribution.

7.5. The chi-squared distribution

Helmert (1875a) first published his discovery without justifying it. However, in 1852 Bienaymé (Heyde & Seneta 1977, § 4.3) derived that distribution but not for observational errors. Abbe obtained it for the theory of errors (§ 7.3) but Helmert somehow did not mention him. Kendall (1971) is a modern description of Abbe's discovery.

7.6. The Peters formula

Peters (1856) derived a formula for the mean absolute error of unit weight for the case of n direct and normally distributed observations in terms of the differences v_i between the observations and their arithmetical mean. Helmert (1875b) proved that formula anew since Peters (tacitly) assumed that the v_i were independent. He considered four cases in the two first of which the Peters formula was however confirmed.

Direct observations, $n = 2$.

Denote the errors of observations by ε_i . Helmert calculated $E|\varepsilon|$.

Same, but n arbitrary.

Helmert had to calculate a multiple integral by applying the Dirichlet discontinuity factor. It is much easier to note a formula (David 1957, p. 27) for $E|v_i|$ in case of the normal distribution.

Indirect observations with m unknowns (so that $m > 1$). Helmert only indicated that the Peters formula underestimated the mean absolute error.

The precision of the Peters formula for n direct observations (Helmert 1876a).

This difficult investigation was theoretically necessary, and independently repeated by Fisher (1920, p. 761), but unimportant since covered by Gauss.

7.7. The Gauss formula

Gauss (§ 5.3-6) proved the formula for the square of the mean square error of observations

$$m^2 = [vv]/(n - k)$$

where n and k were the numbers of the observations and unknowns. He also derived bounds for the variance $\text{var}m^2$

$$\begin{aligned} & 2(v_4 - 2s^4)/(n - k); \\ & [1/(n - k)](v_4 - s^4) + (k/n)(3s^4 - v_4) \end{aligned}$$

where v_4 was the fourth moment of the errors and $s^2 = Em^2$. Without application of any new methods Helmert (1904) discovered that the lower boundary was wrong and Kolmogorov et al (1947) independently repeated his finding. Here is the final result; Maltzev (1947) proved that the lower bound is attainable. For non-negative and

then non-positive $(v_4 - 3s^4)$, the product $(n - k)\text{var}m_o^2$ is contained within, respectively, boundaries

$$[(v_4 - s^4) - (k/n)(v_4 - 3s^4); (v_4 - s^4)] \\ [(v_4 - s^4); (v_4 - s^4) + (k/n)(3s^4 - v_4)].$$

7.8. Anticipation of the Student – Fisher theorem

Helmert (1876b) indicates that for the normal distribution, denoting the observational errors and their mean by ε_i and ε ,

$$P = n(h/\sqrt{\pi})^n \exp[-h^2([vv] + n\varepsilon^2)] dv_1 dv_2 \dots dv_{n-1} d\varepsilon.$$

This formula shows that, for the normal distribution, $[vv]$, and, therefore, the variance as well, and the arithmetic mean are independent. Helmert thus proved the important Student – Fisher theorem but did not pay any attention to it. Kruskal (1946) mentioned several modern derivations of that formula and offered his own inductive proof^{7.3}.

7.9. The precision of the mean square error

In addition to the just mentioned formula, Helmert (1876a) studied the mean square error, m , although only for the normal distribution.

He noted that for small values of n $\text{var}m_o^2$ does not estimate the precision of formula for m^2 in § 7.7 well enough and derived the formula for

$$E(m - [vv]/\sqrt{n-1})^2$$

in terms of n and the Γ function. He issues from the probability of the values of v_i , $i = 1, 2, \dots, (n-1)$,

$$P = \sqrt{n}(h/\sqrt{\pi})^{n-1} \exp(-h^2[vv]) dv_1 dv_2 \dots dv_{n-1}$$

which follows from the formula of § 7,8, notes that the probability $P(\varepsilon \leq [vv] \leq \varepsilon + d\varepsilon)$ is equal to the appropriate integral, and introduces new variables

$$t_1 = \sqrt{2}(v_1 + 1/2v_2 + 1/2v_3 + 1/2v_4 + \dots + 1/2v_{n-1}), \\ t_2 = \sqrt{(3/2)}(v_2 + 1/3v_3 + 1/3v_4 + \dots + 1/3v_{n-1}), \\ t_3 = \sqrt{(4/3)}(v_3 + 1/4v_4 + \dots + 1/4v_{n-1}), \dots, \\ t_{n-1} = \sqrt{[n(n-1)]} v_{n-1}.$$

Note that $[vv] = [tt]$ where, however, the first sum consists of n terms and the second one, of $(n-1)$ terms, and the Jacobian of the transformation is $\sqrt{n}$. The derivation of the formula in terms of the Γ function now follows immediately since Helmert knows the χ^2 distribution. Taken together, the transformations from $\{\varepsilon\}$ to $\{v\}$ and from $\{v\}$ to $\{t\}$ are called after him.

7.10. Is unbiasedness necessary?

Sprott (1978, p. 199) remarked that unbiasedness in estimating parameters is hardly necessary and sometimes impossible.

Furthermore, although the sample variance is unbiased, the practically applied mean square error is biased. In this respect, Helmert's study (§ 7.9) was therefore important but possibly forgotten.

Bertrand (1888c) first challenged the celebrated Gauss formula by providing an example of a *better* estimator of precision but he failed to notice that his estimator was biased (and he missed the opportunity to apply the apt Gaussian calculations).

Czuber (1891b, p. 460) took up Bertrand's example and discussed it with Helmert:

Zur Wahrnehmung dieses Fehlschlusses gelangte ich gelegentlich einer Besprechung mit ... Helmert, welcher mir seine Bedenken gegen die obige Aufstellung äußerste und bald auch das Grund des eigentümlichen Resultates erkannte, das, wenn richtig, einen gewichtigen Einwand gegen die Gauss'sche Theorie bilden müsste. Er lagte darin, dass Bertrand die Unsicherheit der [Gaussian] Formel nach dem absoluten Betrage ihres mittleren Fehlerquadrates beurteilt, statt, wie es sein muss, den relativen Betrag ... zugrunde zu legen.

Thus, bias did not trouble either Czuber or Helmert. Gauss however proved that, again for normally distributed errors, $\text{var}m^2/m^2$ was minimal for his estimator. In this case the Gauss celebrated formula provides not only the least variance $\text{var}m^2$ but also the least relative variance. For the practitioner this is important. Eddington (1933, p. 280), for one, also preferred the relative over the absolute variance.

At present, bias of statistical estimators is tolerated, at least to a certain extent, see above. It would be apparently prudent to adopt the same attitude in treating observations, the more so since the mean square error is practically applied.

Notes

7.1. Modern authors (Dixon 1962; Kuskal 1960, p. 348) confess that they have no general answer to rejection. Some statistical tests can help, but no rule is better than its premises whose validity is difficult to check. A special inseparable problem is the decision about the possible rejection of an observation which certainly belongs to an alien population. Barnett & Lewis (1964), see § 5.4, apparently closed the problem, at least for some long time. The popular tests for rejection were the three sigma rule (§ 7.2) and the proposal of Chauvenet (1863, vol. 2, pp. 558 – 566) who presumed normally distributed errors. At present, especially since normality is often questioned, his proposal is deservedly forgotten.

7.2. I mention the first paper in the beginning of § 7.5.

7.3. Another author who came close to the Student – Fisher theorem was Bertrand (1888a; 1888b), and Heyde & Seneta (1977, p. 67 note) noticed the latter source.

Bertrand's treatise (1888d) is impregnated with its non-constructive negative (and often unjustified) attitude towards the theory of probability and treatment of observations. And at least once he (pp. 325 – 326) wrongly alleged that Cournot supposed that judges decide their cases independently one from another.

However, Bertrand exerted a strong (too strong) influence upon Poincaré, and, its spirit and inattention to Laplace and Bienaymé notwithstanding, on the revival of the interest of French scientists in probability (Bru & Jongmans 2001).

8. Stable laws

Mathematicians began to study stable laws of distribution in the 1920's. Lévy was cofounder and the sole author who (mistakenly) argued that they were necessary for establishing the theory of errors

anew. His account is mostly contained in the two unmethodically compiled contributions (1924; 1925) which I attempt to systematise.

8.1. Random errors

Lévy (1924, p. 51; 1925, p. 278) certainly thought that their mean value was zero, that (1924, p. 50) they were *indépendantes et très petites* or at least (1925, p. 278) that they appeared *comme la somme* of such errors.

He (1925, pp. 70 – 71 and 278) twice stated that random errors were normally distributed but then he (p. 75) also argued that

Elle n'obéira qu'à peu près à la loi de Gauss

and (p. 279) concluded:

En définitive, l'erreur accidentelle obéit à la loi de Gauss d'autant plus exactement que les conditions [of the CLT] sont plus exactement vérifiées.

Late in life Lévy (1970, p. 71) maintained that in 1919 he had only

Un vague souvenir du fait que les erreurs accidentelles obéissent à la loi de Gauss.

Mostly, however, Lévy was concerned with non-normal laws (and even with peculiar stable laws with index less than unity (see § 8.5) and it seems that, in spite of his definitive remark above, he was mainly discussing observations corrupted by systematic influences.

8.2. Precision of observations

Precision can be comprehensively described only by means of the appropriate law of error (1924, pp. 78 – 79; 1925, pp. 75 – 78).

Correct, but hardly possible to apply. Those who *prétendent fonder la théorie des erreurs* on the concept of precision of observations were wrong since precision is not a *notion première* (1925, p. 74).

Accordingly, Lévy (1924, p. 77; 1925, pp. 80 and 284 – 285) disapprovingly mentioned Bienaymé (1853), who had denied the practical importance of the Cauchy distribution: [sound] observations cannot obey it since precision is measured by the variance. In essence, however, Lévy attacked Laplace and Gauss: he (1924, p. 77) mentioned them as well not forgetting Bienaymé either. The *fausseté* of Laplace and Gauss, as he argued,

Aurait dû apparaître lorsqu'en 1853 Cauchy attire l'attention to stable laws and to his distribution in particular. Lévy thus conditioned the possibility of a plausible estimation of the precision of observations by the existence of a stable law of error (without any restrictions!).

8.3. The mean square error

Lévy considered true values ξ_i rather than deviations from the arithmetic mean and tacitly and certainly thought that the mean square error was the square root of $[\xi\xi]/n$. This statistic, as he (1925, p. 75) confessed, see also his earlier contribution (1924, p. 52), corresponded to the *idée le plus simple* and its application was *assez naturel*. Again (p. 74),

Il semble qu'en effet on ne puisse pas choisir un meilleur paramètre.

Then (1925, p. 77), the mean square error provided

Faute de mieux une certaine idée de l'ordre de grandeur de l'erreur.

However (p 61), other estimators, for example the mean value of the power p/n with p ayant une valeur positive quelconque of the sum of $|\xi_i|$ are also possible. And (p. 78) under the Gaussian law the mean square error was not better than any *paramètre défini d'une autre manière*. For near-normal densities it was still important to use exactly that estimator (same p. 78 and p. 282). Lévy had not justified that last-made pronouncement and (1925, p. 77) argued that the variance should be somehow explained:

Une théorie déduite d'axiomes introduits arbitrairement ne saurait avoir aucune valeur.

And so, the sample variance is a convenient estimator of precision, but other measures of precision can also be applied and in any case (§ 8.5) a comprehensive estimation is impossible without the knowledge of the appropriate law of error, But the greatest trouble with Lévy is that both the index of stable laws and the stability itself remain unknown. I return to this circumstance below.

8.4. New concept of precision

Lévy (1924, p. 73) proposed to estimate the precision of a random error ξ (of an observation corrupted by that error) by a parameter that indicated

L'ordre de grandeur de l'erreur à laquelle on doit s'attendre en valeur absolue.

On p. 75 he noted that the parameter was defined to within an arbitrary multiplier, see p. 78:

Considérer la notion de paramètre de précision comme intuitive, c'est admettre qu'on peut définir par un seul nombre les avantages d'une méthode de mesure.

For the normal law the parameter of the sample mean is $\sqrt{n}$ times less than that of x_i (Lévy 1925, p. 280) whereas its *module de précision* ($h = 1/\alpha^2$) est alors n fois plus grand^{8.1}.

8.5. Stable laws

Lévy offered a definition of stable laws in terms of their characteristic functions, but I am only interested in its corollary (Lévy 1924, p. 69; 1925, p. 258): given, independent and identically distributed errors $\xi_1, \xi_2, \dots, \xi_n$, and positive numbers $a_1, a_2, \dots, a_n$. If there exists such a positive number A that

$$A^\alpha = a_1^\alpha + a_2^\alpha + \dots + a_n^\alpha$$

with $[a\xi]/A$ having the same distribution as ξ_i , then this distribution is stable. The initial definition of stability (Lévy 1924, p. 70; 1925, p. 255) additionally imposes two conditions. First, $0 < \alpha < 1$ and, second, requires that the variance of a stable law is finite if $\alpha = 2$ and infinite otherwise.

For $a_i = 1/n$

$$A = n^{-(\alpha-1)/\alpha}, [a\xi]/A = \bar{\xi}/A$$

has the same distribution as any ξ_i . If, for example, $\alpha = 2$, then the mean $\bar{\xi}$ is distributed as ξ_i/n whereas $\alpha = 1$ leads to the mean $\bar{\xi}$ having

the same distribution as ξ_i . These two cases correspond to the normal law and the Cauchy distribution respectively.

The importance of stability consists in that (Lévy 1925, pp. 78 and 282)

Les moyennes ... calculées avec différents systèmes de coefficients ne donneront lieu à des erreurs du même type, et, par suite, ne seront facilement comparables au point de vue de la précision que si ce type est stable.

Suppose that the law of error is indeed stable. What then? If $\alpha = 2$, the MLSq holds:

La loi de Gauss est bien la seule pour laquelle cette méthode s'applique

(1925, p. 79)^{8.2}. If $1 < \alpha < 2$ the weight of observation i should be proportional to a_i to the power of $-\alpha/(\alpha - 1)$. Here a_i is the appropriate parameter of precision (1924, pp. 75 – 76; 1925, p. 283). In this case (1924, p. 77; 1925, p. 285) adjustment of observations, as compared with the MLSq, should be done *avec quelques modifications* (1924). Suppose that $a_i = \text{const}$ and that consequently the mean ξ has the same distribution as $\xi_i/n^{1/3}$ for $\alpha = 3/2$. Then introduce posterior weights decreasing towards the tails (1924, p. 77).

Indeed, since the variance is infinite, large errors are more dangerous than in the previous case ($\alpha = 2$) and their influence should be diminished. But Lévy did not use these calculations.

The other cases, $\alpha = 1$ and $0 < \alpha < 1$, are still left. If $\alpha = 1$, choose arbitrary weights and calculate a generalised arithmetic mean whose precision will not however be higher than that of a single observation. But why bother? Why not choose any single observation and ignore all the others?

If $0 < \alpha < 1$ the mean is worse than a single observation, so (Lévy 1924, p. 76, also see 1925, pp. 79 and 284)

On peut aussi écarter, dans une proportion déterminée, les plus grands et les plus petits nombres trouvés, et prendre la moyenne des nombres conservés.

The simplest procedure, Lévy continues, is to retain a half of the observations or even a third of them.

And so, the law of error should be stable, otherwise adjustment is dangerous. But is the law stable^{8.3}? Moreover, how to distinguish between stable laws having index less and greater than unity? Nevertheless, Lévy's advice to trim suspect series of observation was followed, see Elashoff & Elashoff 1978, p. 253):

There is more to gain than to lose by discarding some extreme observations when long tails are possible^{8.4}.

Lévy's attempt to introduce stable laws into the theory of errors was damnedly useless not to say stupid and more so since his knowledge of the theory of errors was quite inadequate. I discovered a regrettably unjustified remark allegedly made by Neyman: he felt not quite himself with Lévy's methods.

Notes

8.1. The expression $1/\alpha^2$ is difficult to understand.

8.2. He obviously meant the arithmetic mean rather than this method, cf. below. Later Lévy (1929, p. 30) came to regard least squares (again, the mean) more favourably and stated that it can also be applied, if, beginning with some n , it

Conduit à prendre une valeur d'autant plus exacte que n est plus grand.

Apparently he thus allowed to apply the mean in case of stable laws with index greater than unity, but how to check the fulfilment of these conditions?

8.3. Sufficiently simple expressions for the densities of stable laws are known only for the cases of $\alpha = 2$, 1, and $1/2$ (Zolotarev 1984, pp. 30 – 31).

8.4. Later Lévy (1929, p. 31) voiced a similar opinion without mentioning stability.

9. The determinate theory of errors

It was swallowed by experimental design and preliminary data analysis and the text below describes the history of these new disciplines, cf. § 0.1.

Suppose that an unknown constant w is a function f of several other constants whose observations are corrupted by given errors. Then the error of estimating w is the differential of f which therefore should be reasonably chosen. Intersection in geodesy provides a simple example: the form of the triangle of intersection should ensure the optimal (or reasonably precise) coordinates of the unknown point.

In treating observations it is necessary to decrease the influence of systematic errors, and, once more, preliminary data analysis is necessary.

9.1. The eighteenth century

Ancient astronomers knew well enough that observations are corrupted by errors and should be made under optimal circumstances and had some understanding about the difference between random and systematic errors. All the more this applied to scholars of the 16th and 17th centuries, but the determinate error theory originated in the 18th century. Cotes (1722) solved 28 problems which connected the errors of the various elements of plane or spherical triangles with each other. He thus showed the effect of errors on indirectly determined sides of the triangles. His work became widely known and Condamine (1751, p. 91), for example, had applied *la théorie de Cotes*.

According to Lambert (1765a, § 321), who did not refer to Cotes, the study of the errors of functions of observed magnitudes constituted the subject of the *Theorie der Folgen* (of consequences of errors, of the determinate theory of errors). In §§ 340 – 426 he derived the optimal types of standard geodetic figures.

Daniel Bernoulli (§ 2.7) isolated the two kinds of error, understandably in a rather narrow way, but nevertheless made an important and necessary step by studying how they influence the results of observation.

Mayer and Boscovich derived differential formulas which connected errors of instruments with the ensuing errors of the observed magnitudes (for example, of the moment of the observed passage of a star across the meridian). Proverbio (1988) described their work and noted that such formulas ensure confidence in the prior rectification of instruments, and, I would add, enable to compile optimal programmes of observation.

After 1752 Mayer (who died in 1762) invented the repeating theodolite which made it possible to diminish greatly the error of reading-off the result of sighting of a target.

9.2. Laplace

Many of his works contain pronouncements on optimal programmes of observation, on the influence of errors on final results and on the design of experiments in general. Thus, he (1821) described from this angle a method for determining the orbits of comets and devoted his memoir (1784) to general physical considerations, again in the same connection.

A related example is his estimation of the mass of Jupiter (1816, p. 518): he declared that even after a century of new observations *discutées de la même manière* his result will not change more than by *un centième de sa valeur*. This remark possibly saves his opinion but diminishes its importance. And Poisson (1837, p. 316) noted that observations *d'une autre nature* proved that Laplace's theory which underlay his calculations was wrong.

Laplace's Supplements 2 (1818) and 3 (ca. 1819) to his *Théorie* are more specific. There, he studied the precision of the length of a meridian arc which was determined by a chain of triangles. In the second instance he additionally considered chains which consisted of congruent isosceles triangles and estimated the precision of trigonometric levelling (based on measures of zenith distances).

Laplace's notation is awkward and its explanation (1818) insufficient. He had not introduced weights or variances of the adjusted elements, he worked with integrals starting from the normal law of error (in Suppl. 3, also from even laws). The Gaussian non-parametric approach of 1823 is ignored. But Laplace was the first who studied chains of triangulations.

Kepler (§ 1.9) apparently applied elements of the method of minimax. Now, Laplace (1792, p. 506) studied observations which determine the figure of the Earth:

L'ellipse déterminée ... dont sert à reconnaître si la supposition d'une figure elliptique est dans les limites des erreurs des observations. Mais elle n'est pas celle que les degrés mesurés indiquent avec le plus vraisemblance.

This is quite in keeping with what I believed was Kepler's attitude. Elsewhere, however, Laplace (1812, p. 351) stated that he had discovered *par beaucoup d'exemples* that the methods of minimax and least squares lead to slightly differing results. Was his statement ever checked?

Eventually the method of minimax was applied to approximate non-linear (and non-algebraic) functions, in particular by Poncelet and by Chebyshev who studied the transformation of rotation into rectilinear movement (Goussac 1961).

9.3. Gauss and Bessel

Such scholars as Hipparchus, Tycho and Bradley are justly considered observers of the highest calibre, but it were Gauss and Bessel who originated the modern stage in experimental science. Not without reason Newcomb (Schmeidler 1984, pp. 32 – 33) mentioned the *German school of practical astronomy* although he connected it

only with Bessel (in this connection Gauss was barely known since his appropriate work mostly consisted of official reports):

The fundamental ideas of this school was that the instrument is indicted ... for every possible fault, and not exonerated till it has proved itself correct in every point. The methods of determining the possible errors of an instrument were developed by Bessel with ingenuity and precision of geometric method.

Gauss and Bessel also gave much thought to observation itself.

Gauss detected the main systematic errors: those caused by lateral reflection, by the errors of graduation of the limbs of theodolites and inherent in the method of repetition, also see § 6.2-2. Much of that is contained in his correspondence, see his *Werke*, Bd. 9. Recall (§ 5.4-1) that Gauss formulated his ideas on the optimal number of observations and of course he invented the heliotrope (the solar mirror), see same volume of his *Werke*.

The Gauss method of determining the difference between two approximately equal weights (A and B) deserves special attention since he essentially improved on Borda (Helmert 1872, pp. 47 – 49). Borda weighed A and B alternatively and had to introduce an additional unknown whereas Gauss weighed them at the same time, did not need any additional unknown and thus increased Borda's precision twofold. Pukelsheim (1993, p. 427) noticed Helmert's description and connected the Gauss method with modern ideas on designing weighing experiments. In 1836 and 1839 Gauss explained his innovation in letters to Schumacher (Peters, Bd. 3, 1861, pp. 99 – 101, 268, 272 – 275 and 330 – 333).

Bessel, together with J. J. Baeyer, carried out the triangulation of East Prussia and, alone, described it (1838b). In particular, he discussed the investigation and use of metallic bars for measuring baselines; the examination of theodolites and the observation of horizontal angles and zenith distances; and the appropriate astronomical observations.

Bessel (1839) concerns the first topic and ought to be somewhat described. A measuring bar, several feet long, is supported at two points situated at equal distances from its middle. The weight of the bar bends it and changes its length, and the amount of change naturally depends on the position of the supporting points. So where should you place these points? Bessel formulated this problem and solved it by means of appropriate differential equations. But was his investigation interesting for civil engineers (if they existed) of his time? Certainly not for modern specialists.

The personal equation (1823). The recorded moment of the passage of a star through the cross-heirs of an astronomical instrument strongly depends on the psychological habits of the observer, and two astronomers will hardly ever record the same moment, even approximately. And Bessel reasonably began his paper by mentioning Maskelyne: about 1796, he had sacked his assistant whose recorded time of such passages essentially differed from his own record.

And so, Bessel calculated such differences for several pairs of astronomers and made the previously unknown inference stated

above. He determined the mean value of systematic differences between such paired measurements.

Bessel's treatment of the differences between himself and one of the other astronomers, Walbeck, is shocking, and, what is hardly known, *the other Bessel* was a happy-go-lucky fellow (cf. § 6.1!).

The differences between them amounted to

1.145, 0.985, 1.010 and 1.025 sec

yet Bessel (p. 301) blatantly declared that their mean value

Kaum einige Hundertteile einer Sekunde zweifelhaft sein kann.

9.5. Helmert

Actually following Laplace (§ 9.3-1), Helmert (1868) studied configurations of various geodetic figures to determine their optimal unknown elements. For example, he examined quadrilateral base nets whose shorter diagonal was the measured baseline and the longer, calculated, diagonal was included in a triangulation as a side of its first or last triangle^{9.1}.

Helmert (pp. 1 and 60) broke fresh ground. In accordance with the not yet existing linear programming he formulated his aim as

Einen notwendigen Genauigkeit of geodetic systems] mit möglichst wenig Zeit und Geld zu erreichen

or as achieving more precise results *bei gleicher Mühe*.

Helmert noted that it was expedient to leave some angles of a particular geodetic system unmeasured, but practice required measurement of each angle, at least as a check.

Schreiber (1882) and Bruns (1882; 1886) followed his work and a related modern study is Grafarend & Harland (1973), which however seems to me purely theoretic, and Friedrich (1937) is an intermediate link.

Helmert (1886, pp. 1 and 68) was also the first to facilitate the adjustment of large networks of triangulations by replacing their parts by geodesics. He worked with an extremely complicated triangulation completed during several decades so that first and foremost he had to devise a reasonable *Näherungsverfahren*.

Much later Krasovsky put Helmert's innovation to better use by taking advantage of the harmonious system of Soviet triangulation (Krasovsky's merit as well). See Zakatov or Sakatov (1957, pp. 438 – 440).

Note

9.1. The short baselines were connected with *their* triangulations by base nets. Schwerd (1822) was the first to introduce them. Because of the great difficulties of directly measuring long distances at least to the mid-20th century more or less complicated base nets often had to be measured.

Bibliography

AHES = *Arch. Hist. Ex. Sci.*

AN = *Astron. Nachr.*

C. r. = *C. r. Acad. Sci. Paris*

MNRAS = *Monthly Notices Roy. Astron. Soc.*

OC = *Oeuvr. Compl.*

Stigler = Stigler (1980)

Studies 1, 2 = *Studies ...* (1970 – 1977)

ZfV = *Z. f. Vermessungswesen*

Aaboe A., De Solla Price D. S. (1964), Qualitative measurements in antiquity. In *L'aventure de la science. Mélanges A. Koyré*, t. 1, pp. 1 – 20. Paris.

Abbe C. (1871); Historical note on the method of least squares. *Amer. J. Sci. Arts*, vol. 1, pp. 411 – 415. Reprint: Stigler.

Abbe E. (1863), Über die Gesetzmässigkeit in der Verteilung der Fehler bei Beobachtungsreihen. *Ges. Abh.*, Bd. 2, pp. 55 – 81. Hildesheim, 1989

Adrain R. (1808), Research concerning the probabilities of the errors which happen in making observations. Reprint: Stigler.

--- (1818a), Investigation of the figure of the earth and of the gravity in different latitudes. Reprint: Stigler.

--- (1818b), Research concerning the mean diameter of the earth. Reprint: Stigler.

Al-Biruni (1967), *Determination of the coordinates of position* etc. Beirut.

--- (1983), On the ratios between metals and precious stones. In Al-Khasini (1983).

Al-Khasini (1983), *Kniga vesov mudrosti* (Book on the scales of wisdom). In *Nauchnoe nasledstvo*, vol. 6, pp. 15 – 140.

Archimedes (1925), *Die Sandzahl*. In author's *Über schwimmende Körper und Sandzahl*, pp. 67 – 82. Leipzig.

Aristarchus (1959; Greek and Engl.), On the sizes and distances of the sun and the moon. In Heath T. *Aristarchus of Samos*, pp. 353 – 414. Oxford.

Barnett V., Lewis T. (1978), *Outliers in statistical data*. Chichester, 1984.

Berggren J. L. (1991), Ptolemy's maps of earth and the heavens. A new interpretation. *AHES*, vol. 43, pp. 133 – 144.

Bernoulli Daniel (1770 – 1771), *Mensura sortis*. *Werke*, Bd. 2, pp. 326 – 360. Basel, 1982.

--- (1778, Latin), The most probable choice between several discrepant observations etc. *Biometrika*, 1961; Studies 1, pp. 157 – 167.

--- (1780), Specimen philosophicum de compensationibus horologicis. *Werke*, Bd. 2, pp. 376 – 390.

Bernoulli Johann III (1789), Milieu. English translation in *Festschrift for Lucien Le Cam*. New York, 1997, pp. 358 – 367.

Bertrand J. (1888a), Sur l'évaluation a posteriori de l'confiance méritée par la moyenne. *C. r.*, t. 106, pp. 887 – 891.

--- (1888b), Sur l'erreur à craindre dans l'évaluation des trois angles d'un triangle. *C. r.*, t. 106, pp. 967 – 970.

--- (1888c), Sur les conséquences de l'égalité acceptée etc. *C. r.*, t. 6, pp. 1259 – 1263.

--- (1888d), *Calcul des probabilités*. Second edition 1907. Reprints of first edition, New York, 1970, 1972.

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

--- (1818), *Fundamenta astronomiae*. Königsberg.

--- (1820), Beschreibung des auf das Königsberger Sternwarte. *Astron. Jahrb. für 1823*, pp. 161 – 168. Berlin.

--- (1823), Persönliche Gleichung bei Durchgangsbeobachtungen. *Abh.*, Bd. 3, pp. 300 – 303.

--- (1826), Methode die Thermometer zu berichtigen. *Ibidem*, pp. 226 – 233.

--- (read 1832), Über den gegenwärtigen Standpunkt der Astronomie. In author's *Populäre Vorlesungen*, pp. 1 – 33. Hamburg, 1848.

--- (1833), Letter to G. B. Airy. *Abh.*, Bd. 3, pp. 462 – 465.

--- (1838a) Untersuchungen über die Wahrscheinlichkeit der Beobachtungsfehler. *Abh.*, Bd. 2, pp. 372 – 391.

--- (1838b), *Gradmessung in Ostpreussen*. *Abh.*, Bd. 3, pp. 62 – 138. Parts of book.

--- (1839), Einfluss der Schwere auf die Figur eines, auf zwei Punkten von gleicher Höhe auflegenden Stabes. *Abh.*, Bd. 2, pp. 275 – 282.

--- (1876), *Abhandlungen*, Bde 1 – 3. Leipzig.

Bienaymé I. J. (1852), Sur la probabilité des erreurs d'après la méthode des moindres carrés. *J. math. pures appl.*, sér. 1, t. 17, pp. 33 – 78.

--- (1853), Considérations à l'appui de la découverte de Laplace etc. *C. r.*, t. 37, pp. 309 – 324. Reprint: *J. math. pures appl.*, sér. 2, t. 12, 1867, pp. 158 – 176.

Biermann K.-R. (1956), Spezielle Untersuchungen zur Kombinatorik durch G. W. Leibniz. *Forschungen u. Fortschritte*, Bd. 30, pp. 169 – 172.

--- (1966), Über die Beziehungen zwischen Gauss und Bessel. *Mitt. Gauss-Ges. Göttingen*, Bd. 3, pp. 7 – 20.

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

Bomford (1971), *Geodesy*. Oxford. Previous editions 1952, 1962.

Bortkevich V. I. (Bortkiewicz L. von), Kritische Betrachtungen zur theoretischen Statistik, Tl. 2. *Jhrb. Nationalökon. u. Statistik*, 3. Folge, Bd. 10, pp. 321 – 360.

Boscovich R. J. (1757), *De literaria expeditione per pontificum itionem*. See Cubranic (1961).

--- (1758), *Philosophiae naturalis theoria*. Latin and Engl. Chicago – London, 1922.

--- (unpubl., no date), *De calculo probabilitatum*. Boscovich archive, Dept rare books and Sp. collections. Univ. Calif. Library, MS 62.

Boyle R. (posth. publ. 1772), Physico-chymical essay. *Works*, vol. 1, pp. 359 – 376. Reprint of volume: Hildesheim, 1965.

Bradley J. (1750), Letter ... concerning the apparent motion observed in some of the fixed stars. In Rigaud S. P. *Misc. works and corr. of J. Bradley*, p. 17 – 41. Oxford, 1832.

Bru B. Jongmans F. (2001), Bertrand. In Heyde, Seneta (2001, pp. 185 -189).

Bruns H. (1886), Über eine Aufgabe der Ausgleichung. *Abh. Sächs. Ges. Wiss., math.-phys. Kl.*, Bd. 13, No. 7, pp. 517 – 563.

Buniakovsky V. Ya. (1846), *Osnovania matematicheskoi teorii veroiatnostei* (Principles of math. theory prob.) Petersburg.

Caspar M., von Dyck W. (1930), *Kepler in seinen Briefen*, Bd. 2. München – Berlin.

Cauchy A. L. (1853a), Sur l'évaluation d'inconnues déterminées par un grand nombre d'équations approximatives. OC, sér. 1, t. 12, pp. 36 – 46. Paris, 1900.

--- (1853b), Sur la nouvelle méthode d'interpolation comparée à la méthode des moindres carrés. Ibidem, pp. 68 – 79.

Chauvenet W. (1863), *Manual of spherical and practical astronomy*, vols. 1 – 2. New York, 1960, reprint of edition of 1891.

Chebotarev A. S. (1955), *Geodesia*, vol. 1. Moscow.

Chebyshev P. L. (read 1880), *Teoria veroiatnostei* (Theory of probability). Moscow, 1936. Published from notes taken by A. M. Liapunov. Enormous number of math. misprints.

Clarke A. R. (1880), *Geodesy*. Oxford.

Condamine C. M. de la (1751), *Mesure des trois premiers degrés du méridien*. Paris.

Coolidge J. L. (1926), Adrain and the beginnings of American mathematics. *Amer. math. monthly*, vol. 33, No. 2, pp. 61 – 76.

Cotes R. (1722), Aestimatio errorum in mixta mathesi etc. English translation in Gowing R. (1983), *Roger Cotes, natural philosopher*. Cambridge.

Cournot A. A. (1843), *Exposition de la théorie des chances et des probabilités*. Paris, 1984. Editor, B. Bru. **S, G**, 54.

Cramér H. (1946), *Mathematical methods of statistics*. Princeton. A large number of printings.

Cubranic R. (1961), *Geodetski rad Boscovica* (Geod. work of Boscovich). Zagreb. Contains reprint and Serbo-Croatian translation of Boscovich (1757).

Czuber E. (1890), Zur Theorie der Beobachtungsfehler. *Monatsh. Math. Phys.*, Bd. 1, pp. 457 – 464.

--- (1891a), *Theorie der Beobachtungsfehler*. Leipzig.

--- (1891b), Zur Kritik einer Gausschen Formel. *Monatsh. Math. Phys.*, Bd. 2, pp. 459 – 464.

--- (1903), Über einen Satz der Fehlertheorie und seine Anwendung. *Jahresber. deutsch Math. Vereinigung*, Bd. 12, pp. 23 – 30.

David H. A. (1957), Some notes on the statistical papers of F. R. Helmert. *Bull. Stat. soc. New South Wales*, No. 19, pp. 25 – 28. Reprinted from the same source in 1954.

David F. N., Neyman J. (1938), Extension of the Markoff theorem on least squares. *Stat. Res. Mem.*, vol. 2, pp. 105 – 117.

- Davidov A. Yu. (1857), The theory of mean magnitudes. In *Rech i otchet v torzhestvennom sobranii Moskovskogo universiteta* (Orations and report grand meeting Mosc. Univ.), separate paging. Moscow.
- Delambre J. B. J. (1827), *Histoire de l'astronomie du dix-huitième siècle*. Paris.
- De Moivre A. (1712), De mensura sortis. English translation, De mensura sortis or measurement of risk. *Intern stat. rev.*, vol. 52, 1984, pp. 236 – 262. Commentary: A. Hald, pp. 229 – 236.
- (1730), *Miscellanea analytica*. London. Paris, 2009.
- (1733, Latin; author's translation), Method of approximating the sum of the terms of the binomial $(a + b)^n$ etc. Incorporated in the *Doctrine* (1738 and, expanded, in 1756, pp. 243 – 254).
- (1743), *Annuities on lives*. London. First edition, 1725. Also in author's book (1756, pp. 261 – 328).
- (1718, 1738, 1756), *Doctrine of chances*. Reprint of last edition: New York, 1967.
- De Montessus R. (1903), Une paradoxe du calcul des probabilités. *Nouv. Annales Math.*, sér. 4, t. 3, pp. 21 – 31.
- De Morgan A. (1864), On the theory of errors of observation. *Trans. Cambr. Phil. Soc.*, vol. 10, pp. 409 – 427.
- De Morgan Sophia Elizabeth (1882), *Memoir of A. De Morgan*. London.
- Descartes R. (1637), *Discours de la méthode. Oeuvr.*, t. 6, 1965, pp. 1 – 78. Paris.
- De Vries H. (ca. 1905), Evidence of evolution. *Annual Rept Smithsonian Instn* for 1904, pp. 389 – 396.
- De Vries W. F. M. (2001), Meaningful measures etc. *Intern. Stat. Rev.*, vol. 69, pp. 313 – 331.
- Dixon W. J. (1962), Rejections of observations. In *Contributions to order statistics*, pp. 299 – 342. Editors, A. E. Sarhan, B. G. Greenberg. New York – London.
- Dorsey N. E., Eisenhart C. (1969), On absolute measurements. In Ku (1969, pp. 49 – 55).
- Dutka J. (1990), Adrain and the method of least squares. *AHES*, vol. 41, pp. 171 – 184.
- Eddington A. S. (1933), Notes on the method of least squares. *Proc. Phys. Soc.*, vol. 45, pp. 271 – 287.
- Edgeworth F. Y. (1883), The method of least squares. *Phil. Mag.*, ser. 5, vol. 16, pp. 360 – 375. Reprint in author's *Writings in probability, statistics and economics*, vols. 1 – 3. Cheltenham, 1996. Editor, C. R. McCann Jr.
- Edwards A. W. F. (1987), *Pascal's arithmetical triangle*. Baltimore, 2002.
- Eggenberger J. (1894), Darstellung des Bernoullischen Theorems etc. *Mitt. naturforsch. Ges. Bern*, No. 1305 – 1334 für 1893, pp. 110 – 182. Also separate publication: Berlin, 1906.
- Ehrenfest P., Ehrenfest T. (1907), Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem. In Ehrenfest P. (1959), *Coll. Scient. Papers*. Amsterdam, pp. 146 – 149.
- Eisenhart C. (1961), Boscovich and the combination of observations. Reprint: *Studies* 2, pp. 88 – 100.
- (1963), Realistic evaluation of the precision and accuracy of instrument calibration systems. In Ku (1969, pp. 21 – 47).
- (1964), The meaning of *least* in least squares. *J. Wash. Acad. Sci.*, vol. 54, pp. 24 – 33.
- (1976), Discussion of invited papers on history of statistics. *Bull. Intern. Stat. Inst.*, vol. 46, No. 2, pp. 355 – 357.
- (1978), Gauss. In Kruskal, Tanur (1978, vol. 1, pp. 378 – 386).
- (1983), Laws of error. *Enc. stat. sciences*, vol. 4, pp. 530 – 566. Editors, S. Kotz, N. L. Johnson. New York.
- Elashoff, Janet D., Elashoff R. M. (1978), Effects of errors in statistical assumptions. In Kruskal, Tanur (1978, vol. 1, pp. 229 – 250).
- Euler L. (1749), Recherches sur la question des inégalités du mouvement de Saturne et de Jupiter. *Opera omnia*, ser. 2, t. 25, pp. 45 – 157. Turici, 1960.
- (1778, Latin), Observations on the foregoing dissertation of [Daniel] Bernoulli. *Biometrika*, 1961. *Studies* 1, pp. 167 – 172.
- Farebrother R.W. (1985), Statistical estimation of the standard linear model,

- 1756 – 1853. *Proc. First Tampere Sem. Linear Models* 1983, pp. 77 – 99. Tampere.
- Fischer P. (1845), *Lehrbuch der höheren Geodäsie*. Darmstadt.
- Fisher R. A. (1920), Mathematical examination of the methods of determining the accuracy of an observation. *MNRAS*, vol. 80, pp. 758 – 770.
- (1925), *Statistical methods for research workers*. Reprint of the fourteenth edition (1973) in author's *Stat. methods, experimental design and scient. inference* of 1990 with separate paging.
- (1939), "Student". *Annals eug.*, vol. 9, pp. 1 – 9.
- (1951), Statistics. In *Scient. thought in the 20th century*, pp. 31 – 55. Editor A. E. Heath. London.
- Forsythe G. E. (1951), Gauss to Gerling on relaxation. *Math. tables and other aids to comp.*, vol. 5, No. 36, pp. 255 – 258.
- Fourier J. B. J. (1826), Sur les résultats moyen. *Oeuvr.*, t. 2, pp. 525 – 545. Paris.
- Freudenthal H., Steiner H.-G. (1966), Aus der Geschichte der Wahrscheinlichkeitstheorie und der mathematischen Statistik. In *Grundzüge der Mathematik*, Bd. 4, pp. 149 – 195. Editors H. Behnke et al. Göttingen.
- Friedrich K. (1937), Allgemeine ... Lösung für die Aufgabe der kleinsten Absolutsummen. *ZfV*, Bd. 66, pp. 305 – 320, 337 – 358.
- Galilei G. (1632, Italian), *Dialogue concerning the two chief world systems*. Berkeley – Los Angeles, 1967.
- (1718), Sopra le scoperte dei dadi. Translation: Thoughts about dice games. In F. N. David, *Games, gods, and gambling*, pp. 192 – 195. London.
- Gauss C. F. (1809a), *Theoria motus* Selbstanzeige. In Gauss (1887, pp. 204 – 205).
- (1809b), *Theoria motus* ... German translation of appropriate sections: Ibidem, pp. 92 – 117).
- (1811), *Disquisitio* ... German translation: Ibidem, pp. 118 – 128.
- (1816), Bestimmung der Genauigkeit der Beobachtungen. Reprint: Ibidem, pp. 129 – 138.
- (1821), *Theoria combinationis*, pt. 1. Selbstanzeige. Reprint: Ibidem, pp. 190 – 195.
- (1822), Anwendung der Wahrscheinlichkeitsrechnung auf eine Aufgabe der praktischen Geometrie. Reprint: Ibidem, pp. 139 – 144.
- (1823a), *Theoria combinationis*, pt. 2. Selbstanzeige. Reprint: Ibidem, pp. 195 – 199.
- (1823b), *Theoria combinationis*, pts 1 – 2. German translation: Ibidem, pp. 1 – 53.
- (1826), Selbstanzeige to Gauss (1828). Reprint: Ibidem, pp. 200 – 204.
- (1828a), Supplementum theoriae combinationis. German translation: Ibidem, pp. 54 – 91.
- (1828b), Bestimmung des Breitenunterschiedes zwischen den Sternwarten etc. *W-9*, pp. 5 – 64. **S, G**, 72.
- (1845), Anwendung der Wahrscheinlichkeitsrechnung auf die Bestimmung der Bilanz für Witwenkassen [pt. 2]. *Nachlass. W-4*, pp. 125 – 157.
- (1855), *Méthode des moindres carrés*. Traduit par J. Bertrand. Paris.
- (1863 – 1930), *Werke*, Bde 1 – 12. Göttingen. Reprint: Hildesheim, 1973 – 1981.
- (1887), *Abhandlungen zur Methode der kleinsten Quadrate*. Hrsg A. Börsch, P. Simon. Berlin. Latest edition: Vaduz, 1998.
- (1863 – 1930), *Werke*, Bde 1 – 12. Göttingen. Reprint: Hildesheim, 1973 – 1981.
- (1975 – 1976), *Werke, Ergänzungsreihe*, Bde 1 – 5. Reprint of his correspondence. Bd. 1, with Bessel. Bd. 3, with Gerling. Bd. 4, with Olbers. Bd. 5, with Schumacher.
- Gerardy T. (1977), Die Anfänge von Gauss' geodätische Tätigkeit. *ZfV*, Bd. 102, pp. 1 – 20.
- Gerling Ch. L. (1839), *Beiträge zur Geographie Kurhessens*. Kassel.
- Gingerich O. (1983), Ptolemy, Copernicus and Kepler. In *Great Ideas Today*, pp. 137 – 180. Editors M. J. Adler, J. van Doren. Chicago.
- Glaisher G. W. L. (1872), On the law of facility of errors of observation. *Mem. Roy. Astron. Soc.*, vol. 39, pp. 75 – 124.

Goussac A. A. (1961, Russian), La préhistoire et les débuts de la théorie de la représentation approximative des fonctions. *Istoriko-matematich. issledovania*, vol. 14, pp. 289 – 348.

Gower B. (1993), Boscovich on probabilistic reasoning and the combination of observations. In *Boscovich. Vita e attività scient.*, pp. 263 – 279. Editor P. Bursill-Hall. Roma.

Grafarend E., Harland P. (1973), *Optimale Design geodätischer Netze*. München.

Hagen G. (1837, 1867, 1882), *Gründzüge der Wahrscheinlichkeitsrechnung*. Berlin.

Hald A. (1960), *Statistical theory with engineering applications*. Fourth printing. New York – London.

--- (1990), *History of probability and statistics and their applications before 1750*. New York.

--- (1998), *History of math. statistics from 1750 to 1930*. New York.

Harter H. L. (1977, date of preface), *Chronological annotated bibliography on order statistics*, vol. 1. No place. Published by US Air Force and its subunits.

Hauber C. Fr. (1830 – 1832), Theorie der mittleren Werte. *Z. Phys. Math.*, Bd. 8, pp. 25 – 56, 147 – 179, 295 – 316; Bd. 98, pp. 302 – 322; Bd. 10, pp. 425 – 457.

Hegel G. W. F. (1812), *Wissenschaft der Logik*, Tl. 1. Hamburg, 1978.

Helmert F. R. (1868), Studien über rationelle Vermessungen im Gebiete der höhern Geodäsie. *Z. Math. Phys.*, Bd. 13, pp. 73 – 120, 163 – 186. References to separate publication: Leipzig, 1868.

--- (1872, 1907, 1924), *Ausgleichsrechnung nach der Methode der kleinsten Quadrate*. Leipzig.

--- (1875a), Über die Berechnung der wahrscheinlichen Fehlers. *Z. Math. Phys.*, Bd. 20, pp. 300 – 303.

--- (1875b), Über die Formeln für den Durchschnittsfehler. *AN*, Bd. 85, pp. 353 – 366.

--- (1875c), Discussion der Beobachtungsfehler in Koppe's Vermessung für die Gotthardtunnelachse. *ZfV*, Bd. 5, pp. 146 – 155.

--- (1876a), Genauigkeit des Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers. *AN*, Bd. 88, pp. 113 – 132.

--- (1876b), Über die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler. *Z. Math. Phys.*, Bd. 21, pp. 192 – 218.

--- (1877), Über den Maximalfehler einer Beobachtung. *ZfV*, Bd. 6, pp. 131 – 147.

--- (1886), *Lotabweichungen*, Tl. 1. Berlin.

--- (1904), Zur Ableitung der Formel von Gauss für den mittleren Beobachtungsfehler und ihrer Genauigkeit. *Sitz. Ber. Kgl. Preuss. Akad. Wiss. Berlin*, Hlbbd 1, pp. 950 – 964. Shorter version (1904), *ZfV*, Bd. 33, pp. 577 – 587.

--- (1905), Über die Genauigkeit der Kriterien des Zufalls bei Beobachtungsreihen, *Sitz. Ber. Kgl. Preuss. Akad. Wiss. Berlin*, Hlbbd. 1, pp. 594 – 612.

Herschel W. (1805), On the direction and the motion of the Sun. *Scient. papers*, vol. 2, pp. 317 – 331. London, 1912, 2003.

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

Hilbert D. (1901), Mathematische Probleme. *Ges. Abh.*, Bd. 3, pp. 290 – 329. Berlin, 1935.

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

Hulme H. R. (1940), Statistical theory of errors. *MNRAS*, vol. 100, pp. 303 – 314.

Hulme H. R., Symms L. S. T. (1939), The law of error and the combination of observations. *Ibidem*, vol. 99, pp. 642 – 649.

Huygens C. (1699), Correspondence with Lodewijk Huygens. OC, t. 6. La Haye, 1895.

Idelson N. I. (1947), *Sposob naimenshikh kvadratov* etc. (MLSq). Moscow. S, G, 58 (only chapter 1)

Ivory J. (1830), On the figure of the earth. *Phil Mag.*, new ser., vol. 7, pp. 412 – 416.

Jordan W. (1877), Über den Maximalfehler einer Beobachtung. *ZfV*, Bd. 6, pp. 35 – 40.

- Joule J. P. (1849), On the mechanical equivalent of heat. *Phil. Trans. Roy. Soc.* for 1850, pp. 61 – 82.
- Kac M. (1939), On a characterisation of the normal distribution. In author's *Probability, number theory and stat. physics*, pp. 77 – 79. Cambridge, Mass.
- Kapteyn J. C. (1912), Definition of the correlation coefficient. *MNRAS*, vol. 72, pp. 518 – 525.
- Kemnitz Yu. V. (1957, Russian), The density of the errors of measurement. *Geodezia i kartografia*, No. 10, pp. 21 – 29.
- Kendall M. G. (1971), The work of E. Abbe. *Biometrika*. Studies 2, pp. 331 – 335.
- Kepler J. (1606), *Über den neuen Stern* etc. Würzburg, 2006.
- (1609, Latin), *New Astronomy*. Cambridge, 1992, 2015.
- Keynes J. M. (1921), *Treatise on probability*. Reprint: *Coll. writings*, vol. 8. London.
- Kohlrausch F. (1870), *Leitfaden praktischen Physik*. Leipzig.
- (1609, Latin), *New Astronomy*. Cambridge, 1992, 2015.
- Kolmogorov A. N. (1931, Russian), La méthode de la médiane dans la théorie des erreurs. *Matematich. sbornik*, vol. 38, No. 3 – 4, pp. 47 – 49.
- (1946, Russian), Justification of the method of least squares. *Sel. works*, vol. 2, pp. 285 – 302.
- (1985 – 1986, Russian), *Selected works*, vols. 1 – 2. Dordrecht, 1991 – 1992.
- Kolmogorov et al (1947, Russian), A formula of Gauss in the method of least squares. *Sel. works*, vol. 2, 1992, pp. 303 – 308.
- Kruskal W. H. (1946), Helmert's distribution. *Amer. math. monthly*, vol. 53, pp. 435 – 438.
- (1960), Some remarks on wild observations. In Ku (1969, pp. 346 – 348).
- (1978), Formulas, numbers, words: statistics in prose. In *New directions for methodology of soc. and behavioural sci.*, No. 9, pp. 93 – 102. Editor D. Fiske. San Francisco.
- Kruskal W. H., Tanur J. M., Editors (1978), *Intern. Enc. of Statistics*, vols. 1 – 2. New York – London.
- Ku H. H. (1967), Statistical concepts in metrology. In author's (1969, pp. 296 – 330).
- Editor (1969); *Precision measurement and calibration*, *Nat. Bureau Standards. Papers on stat. concepts and procedures*. NBS Sp. Publ. 300, vol. 1. Washington.
- Lagrange J. L. (1776), Sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations. OC, t. 2, pp. 173 – 234. Paris.
- Lambert J. H. (1760), *Photometria*. Augsburg.
- (1765a), Anmerkungen und Zusätze zur praktischen Geometrie. In author's *Beiträge zum Gebrauche der Mathematik und deren Anwendung*, Tl. 1, pp. 1 – 313. Berlin.
- (1765b), Theorie der Zuverlässigkeit der Beobachtungen und Versuche. *Ibidem*, pp. 424 – 488.
- Lamont J. (1867), Über die Bedeutung arithmetischer Mittelwerte in der Meteorologie. *Z. Öster. Ges. Met.*, Bd. 2, No. 11, pp. 241 – 247.
- Lancaster H. O. (1966), Forerunners of the Pearson chi-square. *Austr. J. Stat.*, vol. 8, pp. 117 – 126.
- Laplace P. S. (1784), Sur la chaleur. OC, t. 10, 1894, pp. 149 – 200. Paris.
- (1792), *Sur quelques points du système du monde*. OC, t. 11, pp. 477 – 558. Paris, 1895.
- (1798; 1798; ca. 1803; 1805; 1825), *Traité de Mécanique céleste*, tt. 1 – 5. OC, tt. 1 – 5, 1878 – 1882. Paris.
- (1812), *Théorie analytique des probabilités*. OC, t. 7, 1886. Paris.
- Supplements 1 – 3, 1816, 1818, ca. 1819; *Ibidem*, pp. 497 – 530, 531 – 580, 581 – 616.
- (1814, French), *Philosophical essay on probabilities*, 1995. New York. Translator A. I. Dale.
- (1821), Sur la détermination des orbites des comètes. OC, t. 13, pp. 285 – 267, 1904. Paris.
- (1827), Sur le flux et reflux lunaire atmosphérique. OC, t. 13. Paris, 1904, pp. 342 – 358. An almost identical text is in a manuscript: OC, t. 5, pp. 489 – 505, 1882. Paris.

- (1832, 1966), *Celestial mechanics*, vol. 2. New York. Translation of same volume by N. Bowditch.
- Legendre A. M. (1805), *Nouvelles méthodes pour la détermination des orbites des comètes*. Paris.
- (1820), *Nouvelles méthodes ...*, Supplement 2. Paris.
- Lehmann E. L. (1959), *Testing statistical hypotheses*. New York – London.
- Lehmann-Filhès R. (1887), Über abnorme Fehlverteilung. *AN*, Bd. 117, pp. 121 – 132.
- Lévy P. (1924), La loi de Gauss et les lois exceptionnelles. *Bull. Soc. Math. France*, No. 2, pp. 49 – 85.
- (1925), *Calcul des probabilités*. Paris
- (1929), Sur quelques travaux relatifs a la théorie des erreurs. *Bull. Sci. Math.*, sér. 2, t. 53, pp. 11 – 32.
- (1970), *Quelques aspects de la pensée d'un mathématicien*. Paris.
- Linnik Yu. V. (1952, Russian), Comments on the classical derivation of the Maxwellian law. *Doklady Akad. Nauk SSSR*, vol. 85, pp. 1251 – 1254.
- (1958, 1962, Russian), *Method of least squares*. Oxford, 1961.
- Linnik Yu. V. et al (1951, Russian), Essay on the work of Markov on number theory and theory of probability. In Markov (1951, pp. 614 – 640).
- Lipschitz R. (1890), Sur la combinaison des observations. *C. r.*, t. 111, pp. 163 – 166.
- Lourie S. Ya. (1934, Russian), Approximate calculations in ancient Greece. *Arkhir istorii nauki i tekhniki*, No. 4, pp. 21 – 46
- Maire C., Boscovich R. (1770), *Voyage astronomique et géographique dans l'Etat de l'Eglise*. Paris.
- Maistrov L. E. (1967, Russian), *Probability theory*. New York – London.
- Maltsev A. I. (1947, Russian), Remark on the work of A. N. Kolmogorov et al. *Izvestia Akad. Nauk SSSR*, ser. math., vol. 11, pp. 567 – 578.
- Markov A. A. (1899, Russian), The law of large numbers and the method of least squares. In Markov (1951, pp. 231 – 251).
- (1900, 1908, 1913, 1924), *Ischislenie veroiyatnostei* (Calculus of probabilities). Petersburg, Moscow. German translation (1912), Leipzig – Berlin.
- (1951), *Izbrannye trudy* (Sel. works). No place.
- May K. O. (1972), Gauss. *Dict. Scient. Biogr.*, vol. 5, pp. 298 – 315.
- Mayer T. (1750), Abhandlung über die Umwälzung des Mondes um seine Achse. *Kosmogogr. Nachr. u. Samml. für 1748*, pp. 52 – 183.
- Meadowcroft L. V. (1920), On Laplace's theorem on simultaneous errors. *Messenger math.*, vol. 50, pp. 40 – 48.
- Mendeleev D. I. (1872, Russian), On the compressibility of gases. *Sochinenia*, vol. 6, pp. 128 – 171. Leningrad – Moscow.
- (1895), On the weight of a definite volume of water. *Soch.*, vol. 22, pp. 105 – 171.
- Mendoza E. (1991), Physics, chemistry and the theory of errors. *Arch. intern. hist. sci.*, vol. 41, pp. 282 – 306.
- Merian P. (1830), D. Huber. In *Verh. allg. schweiz. Ges. ges. Naturwiss. in ihrer 16. Jahresversamml. zu St. Gallen. 1830*, pp. 145 – 152. St. Gallen.
- Merriman M. (1877), List of writings relating to the method of least squares. Reprint: Stigler 1.
- Meyer H. (1891), *Anleitung zur Bearbeitung meteorologischer Beobachtungen*. Berlin.
- Mill J. B. (1843), *System of logic*. London, 1886.
- Neugebauer O. (1948), Mathematical methods in ancient astronomy. In author's *Astronomy and history. Sel. essays*, pp. 99 – 127. New York, 1983.
- (1950), The alleged Babylonian discovery of the precession of the equinoxes. *Ibidem*, pp. 247 – 254.
- Neumann J. von et al (1941), The mean square successive difference. *Annals math. stat.*, vol. 12, pp. 153 – 162.
- Newcomb S. (1882), Discussions and results of observations. *Astron. Papers Amer. ephemeris*, vol. 1, pp. 363 – 487.
- (1886), A generalized theory of combination of observations. *Amer. J. Math.*, vol. 8, pp. 343 – 366. Reprint Stigler 2,
- Newton R. R. (1977), *The crime of Clausius Ptolemy*. Baltimore – London.

- Neyman J. (1934), On the two different aspects of the representative method. *J. Roy. Stat. Soc.*, vol. 97, pp. 558 – 625.
- Ogorodnikov K. F. (1928), A method for combining observations. *Astron. Zh.* (Moscow), vol. 5, No. 1, pp. 1 – 21.
- (1929a), On the occurrence of discordant observations. *MNRAS*, vol. 88, pp. 523 – 532.
- (1929b), On a general method of treating observations. *Astron. Zh.* (Moscow), vol. 6, pp. 226 – 244.
- Paucker M. G. (1819), *Über die Anwendung der Methode der kleinsten Quadratsumme auf physikalische Beobachtungen*. [A programme for a gymnasium.] Mitau.
- Pearson K. (1894), On the dissection of asymmetrical frequency curves. *Phil. Trans. Roy. Soc.*, vol. A185, pt. 1, pp. 71 – 110.
- (1900), On a criterion etc. *Phil. Mag.* Reprint in author's *Early stat. papers*, pp. 359 – 357. Cambridge, 1956.
- (1920), Notes on the history of correlation. *Biometrika*. Studies 1, pp. 185 – 205.
- (1931), Historical note on the distribution of the standard deviations. *Biometrika*, vol. 23, pp. 416 – 418.
- (1978), *History of statistics in the 17th and 18th centuries*. Lectures 1921 – 1933. Editor E. S. Pearson. London.
- Peters C. A. F. (1856), Über die Bestimmung des wahrscheinlichen Fehlers. *AN*, Bd. 44, pp. 29 – 32.
- Peters C. A. F., Hrsg (1860 – 1865), *Briefwechsel zwischen Gauss und Schumacher*, Bde. 1 – 6. Altona.
- Petrov V. V. (1954, Russian), On the method of least squares and its extreme properties. *Uspekhi matematich. nauk*, vol. 9, No. 1, pp. 41 – 62.
- Plackett R. L. (1949), Historical note on the method of least squares. *Biometrika*, vol. 36, pp. 458 – 460.
- (1950), Some theorems in least squares. *Biometrika*, vol. 37, pp. 149 – 157.
- (1958), The principle of the arithmetic mean. *Biometrika*, Studies 2, pp. 121 – 126.
- (1972), Discovery of the method of least squares. *Biometrika*, Studies 2, pp. 279 – 291.
- Poincaré H. (1896, 1912, 1923), *Calcul des probabilités*. Paris.
- (1921), Résumé analytique [of his own works]. In *Math. heritage of H. Poincaré*. *Proc. Symp. Indiana Univ.* 1980, pp. 257 – 357. Editor, F. E. Browder. Providence, RI.
- Poisson S. D. (1837), *Recherches sur la probabilité des jugements*. Paris. **S, G**, 53.
- Price D. J. (1955), Medieval land surveying and topographical maps. *Geogr. J.*, vol. 121, pt. 1, pp. 1 – 10.
- Proverbio E. (1988), Boscovich's determination of instrumental errors in observation. *AHES*, vol. 38, pp. 135 – 152.
- Ptolemy (1956, Greek and English), *Tetrabiblos*. London.
- (1984, English), *Almagest*. London.
- Pukelsheim F. (1993), *Optimal design of experiments*. New York.
- Quetelet A. (1836), *Sur l'homme*, tt. 1 – 2. Bruxelles.
- (1846), *Lettres sur la théorie des probabilités*. Bruxelles.
- (1849), *Sur le climat de la Belgique*, t. 1. Bruxelles.
- (1853), *Théorie des probabilités*. Bruxelles.
- Rabinovich N. L. (1974), Early antecedents of error theory. *ASHES*, vol. 13, pp. 348 – 358.
- Rosenberger O. A. (1828), Über die ... vorgenommene Gradmessung. *AN*, Bd. 6, pp. 1 – 32.
- Sabine E. (1821), Account of experiments to determine the acceleration of the pendulum. *Trans. Roy. Soc.*, pp. 163 – 1909.
- Sampson R. A. (1913), On the law of distribution of errors. *Proc. Fifth Intern. Congr. mathematicians*, vol. 2, pp. 163 – 173.
- Schäfer G., Hrsg (1927), *Briefwechsel zwischen Gauss und Gerling*. Berlin.
- Schilling C., Hrsg (1900 – 1909), *W. Olbers. Sein Leben und sein Werk*, Bd. 2, Abt. 1 – 2. *Briefwechsel zwischen Gauss und Olbers*. Berlin.

- Schmeidler F. (1984), *Leben und Werk des Königsberger Astronomen F. W. Bessel*. Kelkheim/T.
- Schneider I., Hrsg (1988), *Die Entwicklung der Wahrscheinlichkeitstheorie von den Anfängen bis 1933*. Darmstadt.
- Schreiber O. (1879), Richtungsbeobachtungen und Winkelbeobachtungen. *ZfV*, Bd. 8, pp. 97 – 149.
- (1882), Die Anordnung der Winkelbeobachtungen im Göttinger Basisnetz. *ZfV*, Bd. 11, pp. 129 – 161.
- Schwerd F. M. (1822), *Die kleine Speyerer Basis*. Speyer.
- Seal H. L. (1949), Historical development of the use of generating functions in probability theory. *Mitt. Vereinigung Schweiz. Versicherungsmathematiker*, Bd. 49, pp. 209 – 228. *Studies* 2, pp. 67 – 86.
- (1967), Historical development of the Gauss linear model. *Biometrika*, *Studies* 1, pp. 207 – 230.
- Seidel L. (1863), Über eine Anwendung der Wahrscheinlichkeitsrechnung. *Sitz. Ber. Bayer. Akad. Wiss.*, Bd. 2 für 1863, pp. 320 – 350.
- (1865), Über den ... Zusammenhang ... zwischen der Häufigkeit der Typhuserkrankungen und dem Stande ... des Grundwassers. *Z. Biol.*, Bd. 1, pp. 221 – 236.
- (1866), Vergleichung der Schwankungen der Regenmengen mit der Schwankungen der Häufigkeit des Typhus. *Ibidem*, Bd. 2, pp. 45 – 177.
- Shevchenko M. (1988), The star catalogue of Ptolemy. *Istoriko-astronomich. issledovania*, vol. 20, pp. 167 – 186.
- Sheynin O. B.** Papers 1965, 1967, 1972c, 1975b, 1978, 1992 are in Russian. Many later publications are listed in the Bibliography in my *History of probability*, 2017.
- Sheynin O. B.
- 1963, Adjustment of a trilateration figure by frame structure analogue. *Empire surv. rev.*, vol. 17, No. 127, pp. 55 – 56.
- 1965, On the work of R. Adrain in the theory of errors. *Istoriko-matematich. issledovania*, vol. 16, 325 – 336.
- 1966, Origin of the theory of errors. *Nature*, vol. 211, pp. 1003 – 1004.
- 1967, On the history of adjustment of indirect observations. *Izvestia vuzov. Geod. i aerofotos'emka*, No. 3, pp. 25 – 32.
- 1971a, Newton and the classical theory of probability. *AHES*, vol. 7, pp. 217 – 243.
- 1971b, Lambert's work in probability. *Ibidem*, pp. 244 – 256.
- 1972a, On the mathematical treatment of observations by Euler. *AHES*, vol. 9, pp. 45 – 56.
- 1972b, Daniel Bernoulli's work on probability. *Studies* 2, pp. 105 – 132.
- 1972c, Theory of probability. Chapter in *Istoria matematiki s drevneishikh vremen do nachala 19go veka* (Hist. math. from most ancient times to beginning of 19th c.), vol. 3, pp. 126 – 152. Editor, A. P. Youshkevich. Co-author L. E. Maistrov.
- 1973a, Finite random sums. *AHES*, vol. 9, pp. 275 – 305.
- 1973b, Boscovich's work on probability. *Ibidem*, pp. 306 – 324.
- 1973c, Math. treatment of astronomical observations. *AHES*, vol. 11, pp. 97 – 126.
- 1975a, Kepler as a statistician. *Bull. Intern. Stat. Inst.*, vol. 46, pp. 341 – 354.
- 1975b, publicator, Liapunov's manuscript *On the Gauss formula for estimating the precision of observations*. *Istoriko-matematich. issledovania*, vol. 20, pp. 319 – 328.
- 1977, Laplace's theory of errors. *AHES*, vol. 17, pp. 1 – 61.
- 1978, Theory of probability. A chapter in *Math. of the 19th c.*, pp. 211 – 288. Editors, A. N. Kolmogorov, A. P. Youshkevich. Basel, 1992, 2001. Coauthor B. V. Gnedenko.
- 1979, Gauss and the theory of errors. *AHES*, vol. 20, pp. 21 – 72.
- 1982, On the history of medical statistics. *AHES*, vol. 26, pp. 241 – 286.
- 1984a, On the history of the statistical method in astronomy. *AHES*, vol. 29, pp. 151 – 199.
- 1984b, On the history of the statistical method in meteorology. *AHES*, vol. 31, pp. 53 – 95.
- 1986, Quetelet as a statistician. *AHES*, vol. 36, pp. 281 – 325.
- 1989, Markov's work on probability. *AHES*, vol. 39, pp. 337 – 377; vol. 40,

p. 387.

1991, The notion of randomness from Aristotle to Poincaré. *Mathématiques, informatique et sci. humaines*, année 29, No. 114, pp. 41 – 55.

1992, Al-Biruni and the mathematical treatment of observations. *Arabic sciences and philos.*, vol. 2, pp. 299 – 306.

1993a, On the history of the principle of least squares. *AHES*, vol. 46, pp. 39 – 54.

1993b, Treatment of observations in early astronomy. *Ibidem*, pp. 153 – 192.

1994a, Theory of errors. In *Companion Enc. Hist. Phil. Math. Sciences*, vol. 2, pp. 1315 – 1324. Editor, I. Grattan-Guinness. London – New York.

1994b, Gauss and geodetic observations. *AHES*, vol. 46, pp. 253 – 283.

1994c, Ivory's treatment of pendulum observations. *Hist. math.*, vol. 21, pp. 174 – 184.

1994d, Chebyshev's lectures on the theory of probability. *AHES*, vol. 46, pp. 321 – 340.

1994e, Bertrand's work on probability. *AHES*, vol. 48, pp. 155 – 199.

1995a, Helmert's work in the theory of errors. *AHES*, vol. 49, pp. 73 – 104.

1995b, Density curves in the theory of errors. *Ibidem*, pp. 163 – 196.

Shoosmith E. (1985), Th. Simpson and the arithmetic mean. *Hist. math.*, vol. 12, pp. 352 – 355.

Short J. (1763), Second paper concerning the parallax of the Sun. *Phil. Trans. Roy. Soc. Abridged*, 1809, vol. 12, pp. 22 – 37.

Simpson J. Y. (1869 – 1870), Hospitalism. *Works*, vol. 2, pp. 289 – 405. Edinburgh, 1871.

Simpson T. (1740), *Nature and laws of chance*. London.

--- (1756), On the advantage of taking the mean etc. *Phil. Trans. Roy. Soc.*, vol. 49, pt. 1 for 1755, pp. 82 – 93.

--- (1757), Later version of same. In author's *Misc. tracts on some curious subjects*, pp. 64 – 75. London, 1757.

--- (1775), *Doctrine of annuities and reversions*. London.

Spieß W. (1939), Kann man für D. Huber Ansprüche als Erfinder der Methode der kleinsten Quadrate geltend machen? *Schweiz. Z. Vermessungswesen u. Kulturtechnik*, Bd. 37, pp. 11 – 17 and 21 – 23.

Sprott D. A. (1978), Gauss's contributions to statistics. *Hist. math.*, vol. 5, pp. 183 – 203.

Stigler S. M. (1980), *American contributions to mathematical statistics in the 19th c.*, vols. 1 – 2. New York. No general paging.

--- (1984), Boscovich, Simpson and a 1760 note. *Biometrika*, vol. 1, pp. 615 – 620.

--- (1986), *History of statistics*. Cambridge, Mass.

--- (1999), *Statistics on the table*. Cambridge, Mass. Coll. papers.

Strasser G. (1957), *Ellipsoidische Parameter der Erdfigur (1800 – 1950)*. München.

Strecker G. (1846), Über die Atomgewichte des Silbers und Kohlenstoffs. *Annalen Chem. Pharm.*, Bd. 59, pp. 265 – 284.

Studies in the history of statistics and probability (1970 – 1977), vols. 1 – 2.

Editors E. S. Pearson, M. G. Kendall (vol. 1), Sir Maurice Kendall, R. L. Plackett (vol. 2). London.

Subbotin M. F. (1956, Russian), Gauss's astronomical and geodetic work. In *K. F. Gauss*, pp. 243 – 310. Moscow.

Tilling Laura (1975), Early experimental graphs. *Brit. J. Hist. Sci.*, vol. 8, pp. 193 – 213.

Todhunter I. (1865), *History of the math. theory of probability*. London. New York, 1949, 1965.

--- (1869), On the method of least squares. *Trans. Cambr. Phil. Soc.*, vol. 11, pp. 219 – 238.

Toomer G. J. (1974) Hipparchus and the distances of the sun and the moon. *AHES*, vol. 14, pp. 126 – 142.

Vasiliev A. V. (1885), *Teoria veroyatnostei*. Kazan. Lithographic edition.

Walker Helen M. (1928), Relation of Plana and Bravais to theory of correlation. *Isis*, vol. 10, pp. 466 – 484.

Wesley W. G. (1978), Accuracy of Tycho Brahe's instruments. *J. Hist. Astron.*, vol. 9, pp. 466 – 484.

Whittaker E. T., Robinson G. (1924), *Calculus of observations*. London. Later edition, 1960.

Wilson G. (1984), Sources of Ptolemy's parameters. *J. Hist. Astron.*, vol. 15, pp. 37 – 47.

Wolf R. (1858), Daniel Huber von Basel. In *Biographien zur Kulturgeschichte der Schweiz*, 1. Cyclus, pp. 441 – 462.

Yaroshenko S. (1893) Sur la méthode des moindres carrés. *Bull. Sci. Math.*, sér. 2, t. 17, pp. 113 – 125.

Zakatov (Sakatov), P. B. (1953, 1964, Russian), *Lehrbuch der höheren Geodäsie*. Berlin, 1957.

Zolotarev V. M. (1984), *Ustoichivye zakony i ikh primeneniya* (Stable laws and their applications). Moscow.

Index of Names

Acknowledgement not covered. My own name is absent. The numbers refer to subsections rather than to pages. N- i stands for Notes to chapter (paragraph) i

Aaboe, A. 1.3

Abbe, C. 4.5

Abbe, E. 7.3, 7.5

Adrain, R. 4.5, N-4

Airy G. B. 2.2, 5.4, 6.3

Al-Biruni (973 – 1048) 1.3, 1.5, 1.6, 1.9, N-1

Al-Battani 1.6

Baeyer J. J. 9.3

Barnett N. 5.4, N-7

Beer, A. 1. 1.5

Berggren, J. L. 1.5

Bernoulli, Daniel 1.9, 2.7, 2.8, N-2, 5.1, 9.1

Bernoulli, Jacob N-1, 3.1

Bernoulli, Johann III 2.7

Bertrand, J. 5.1, 5.3, 7.10, N-7

Bessel, F. W. 0.4, 0.5, 2.3. N-2, 4.6, 5.2 – 5.4, 6.1 – 6.3, N-6, 7.5, 8.2, 9.3

Bienaymé, I. J. 5.3, N-5

Biermann, K.-R. 4.6

Binet, J. 4.1

Bohlmann G. 5.3

Boltzmann, L. 5.3

Bolyai, J. 4.6

Bolyai, W. 4.6

Bomford, G. 5.2 – 5.4

Bond, G. P. 5.4

Borda, J. C. 9.3

Boscovich, R. J. 0.4, 2.3, 2.4, 4.5, 5.1, 9.1

Bowditch, N. 2.4

Boyle, R. 2.5, 6.2

Bradley, J. 1. 2.5, 6.1, 9.3

Brahe, T. 1.8, 1.9, 2.5. 9.3

Bravais, A. N-4

Brendel, M. 4.6

Bru, B. N-7

Bruns, H. 9.4

Buniakovsky, V. Ya. N-1

Caspar M. 1.9

Cauchy, A. L. 6.1, 8.2, 8.5

Charlier, C. V. 6.4

Chauvenet, W. N-7

Chebyshev, P. L. 5.3, 6.2, 9.2

Cicero M. T. (106 – 43 BC) 1.9, N-1

Clarke, A. E. 5.4

Clausius, R. 6.2

Condamine, C. M. de la 9.1

Condorcet, M. J. A. N. 2.1, N-2
 Coolidge, J. L. N-4
 Cotes, R. 2.1, 9.1
 Cournot, A. A. 5.4, 6.4, N-6, N-7
 Cramér, H. 5.2, 5.3
 Cubranic, 2.4
 Czuber, E. 5.3, N-6, 7.10
D'Alembert, J. Le Rond, N-2
 David, F. N-6
 David, H. A. 7.6
 Delambre, J. B. J. 2.1
 De Moivre, A. 2.5, 2.7, N-2, 3.1
 De Morgan, A. 6.4, N-6
 De Morgan, Sophia N-6
 Descartes, R. N-2
 De Solla Price, D. J. 1.3
 Dirichlet, G. Lejeune 7.6
 Dixon, N. J. N-7
 Dutka, J. N-3, 4.3, 4.5,
 Dyck, W. von 1.9, 4.3, 4.6
Eddington, A. S. 5.3, 6.4, 7.10
 Edgeworth, F. Y. 6.1, N-6
 Eisenhart, C. 1.9, 5.3, 5.4, 6.2, 6.5
 Elashoff, J. D. 8.5
 Elashoff, R. M. 8.5
 Encke, J. F. 5.1, 5.3
 Erathosthenes (ca. 276 – 194BC) 1.1
 Euclid 5.3
 Euler, L. 0.4, 0.5, N-1, 2.1, 2.8
Farebrother, R. W. 0.5
 Fechner, G. T. 6.2
 Filliben 1.9
 Fischer, P. 2.3
 Fisher, R. A. 5.3, 6.5, 7.6, 7.8, N-7
 Forsythe, G. E. 5.4
 Fourier, J. B. J. 5.1, 6.5
 Friedrich, K. 9.4
Galilei, G. 0.4, .0.5, 1.6, 1.7, N-1, 2.4
 Galle, A. 4.6
 Galton, F. 5.3, 6.5
 Gauss, C. F. 0.4 – 0.6, 2.2, 2.3, 4.4 – 4.6, N-4, 5, 6.1, 6.5, N-6, 7.6, 7.7, 7.10, 8.1,
 8.2, 8.5, 9.3
 Gerardy, T. 4.6, 5.4
 Gerling, C. L., 5.4, 6.2
 Gingerich, O. 1.4
 Glaisher, J. W. L. 6.4, 7.4
 Gowing, R. 2.1
 Grafarend, E. 9.4
 Gram, J. P. 6.4
 Graunt, J. 0.4, 1.8
 Gusak (Gousac) A. A. 9.2
Hagen, G. N-6
 Hald, A. 0.5, 1.7, 4.4, 5.3, 6.3, 7.4
 Halley, E. 6.5
 Harland, P. 9.4
 Hauber, C. Fr. N-6
 Helmert, F. R. 0.4, 0.5, N-2, 5.1 – 5.3, 6, 7, 9.3, 9.4
 Heyde, C. C. 6.1, 7.5, N-7
 Hilbert, D. 6.5
 Hipparchus (180 or 190 – 125 BC) 1.1, 1.2, 1.4, 1.5, 9.3
 Hippocrates (460 – 377 or 356 BC) N-1
 Hogan, E. R. 4.5, N-4

Huber, D. 4.3
 Hulme H. R. 6.4
 Humboldt, A. 6.5
 Huygens, C. N-2
 Huygens, L. N-2
Idelson, N. I. 0.5, 5.3, 6.1, 6.4
 Ivory, J. 6.2, N-6
Jacobi, C. G. J. 4.1
 Jordan, W. 7.1
 Joule, J. P. 6.2, N-6
Kac, M. 4.5
 Kapteyn, J. C. 6.5
 Kemnitz, Yu. V. 6.4
 Kendall, M. G. (Sir Maurice) 7.5
 Kepler, J. 0.3, 0.5, 1.4, 1.6, 1.8, 1.9, 2.1, N-2, 9.2
 Keynes, J. M. 3.1
 Kohlrausch, F. 6.2
 Kolmogorov, A. N. 5.3, 7.7
 Krasovsky, F. N. 4.5, 9.4
 Kronecker, L. 5.3
 Kruskal, W. 7.8, N-7
 Ku, H. H. 5.4
Lagrange, J. L. 0.4, 2.6, N-2, 5.3
 Lambert, J. H. 0.4, 0.5, N-1, 2.3, 2.5, N-2, 6.2, 9.1
 Lamont, J. 6.2
 Laplace, P. S. 0.4 – 0.6, 1.4, 1.9, 2, 2.3, 2.4, 2.7, N-2, 3, 3.1, N-5, 6.5, N-6, N-7, 8.2, 9.2, 9.4
 Legendre, A. M. 4.3, 4.4, 4.6, N-4, N-5
 Lehmann-Filhès, R. 6.4
 Leibniz, G. W. 2.4
 Lévy, P. 0.1, 0.2, 8
 Lewis, T. 5.4, N-7
 Liagre, J. B. J. 2.3
 Liapunov, A. M. 6.1
 Lindelöf, E. N-6
 Linnik, Yu. V. 4.5, 6.1
Maire, C. 2.4
 Maistrov L. E. 1.7, N-1
 Maltsev, A. I. 7.7
 Mariotte, E. 1. 6.2
 Markov, A. A. 6.1, N-6
 Marsden, B. G. 4.6
 Maskelyne, N. 9.3
 Maxwell, J. C. 4.5, 6.2, 6.3
 May, K. O. 4.6
 Mayer, T. 2.2, 2.3, 6.3
 Mendeleev, D. I. 5.2, 6.2, N-6
 Mendoza, E. 6.2
 Merian, P. 4.3
 Merriman, M. 5.1
 Meyer, H. 6.2
 Mill, J. S. N-6
 Mises, R. von 3.1
Neugebauer, O. 1.1, 1.3, 1.4
 Newcomb, S. 1.4, 5.2, 5.4, 6.1, 6.3, 6.4, N-6, 9.3
 Newton, I. 0.3, N-2, 3.1
 Neyman, J. N-6, 8.5
Ogorodnikov, K. F. 6.4
 Olbers, H. W. 4.5, 4.6, 5.1, 5.3, 5.4, N-6
Paucker, M. G. 6.2
 Pearson, E. S. N-2
 Pearson, K. 0, 2.5, 2.6, N-2, N-4, 5.3, 6.2, 6.4, 6.5, N-6, 7.4

Peirce, G. B. 6.4
 Peters, C. A. F. 2.2, 7.6, 9.3
 Picard, J. 2.1
 Pizzetti, P. N-6
 Plackett, R. L. 1.8, 4.6, N-5, N-6
 Plana, G. A. A. N-4
 Poincaré, H. 0.1, 1.2, 6.3, N-7
 Poisson, S.-D. 2.5, 4.6, 7.4, N-7, 9.2
 Poncelet, J. V. 9.2
 Price, D. J.
 Proverbio, E. 9.1
 Ptolemy, C. (Ilc.) 0.4, 0.5, 1.1 – 1.6, 1.9, N-2
 Pukelsheim, F. 9.3
Quetelet, A. 6.3, N-6
Robinson, G. 6.1
 Rosenberger, O. A. 6.1
Sabine, E. N-2
 Sampson, E. A. N-6
 Sartorius von Waltershausen, W. de 5.3
 Schäfer, G. 5.4
 Scheffé, H. N-6
 Schilling, C. 4.5, 5.1, N-6
 Schmeidler, F. 9.3
 Schneider, I. 6.2
 Schreiber, O. 5.4, 9.4
 Schumacher, H. 2.2, 5.3, 9.3
 Schwerd, F. M. N-9
 Seal, H. L. N-2, 5.3
 Seidel, L. N-6
 Seneta, E. 6.1, 7.5, N-7
 Shevchenko, M. N-9
 Short, J. 2.7
 Simpson, J. Y. N-6
 Simpson, T. 0.4, 0.5, 2.4 – 2.6, N-2, 6.2
 Spieß, W. 4.3
 Sprott D. A. 5.3, 7.10
 Stewart G. W. 5.3
 Stigler, S. M. 0.5, 4.5
 Stone, E. J. 6.4
 #Strasser, G. 4.5
 Strecker, A. 6.2
 Struve, L. O. 5.2
 Student (Gosset, W. S.) 5.3, 7.8, N-7
 Subbotin, M. F. 0.4
 Symms S. L. T. 6.4
Tilling, Laura N-2
 Todhunter, I. N-2
 Toomer, G. J. 1.2
 Tsinger, V. Ya. 5.3
Vasiliev, A. V. 2.5
Walbeck 9.3
 Walker, Helen M. N-4
 Weiling, F. N-6
 Wesley, W. G. 1.8
 Whiteside, D. T. N-2
 Wilson, C. 1.4
 Wolf, Abr. 4.3
Yarochenko, S. P. 5.3
Zach, F. X. von 4.6
 Zakatov (Sakatov), P. S. 9.4
 Zoch, R. T. 5.1
 Zolotarev, V. M. N-8

II

Alphonse De Candolle

On a dominant language for science

Chapter 5 of *Histoire des sciences et des savants depuis deux siècles*.
Genève, 1873. Translation: *Annals and Mag. of Nat. Hist.*, ser. 4, No. 11,
date not indicated. Reprint of translation:
Annual Rept Smithsonian Instn for 1874, 1875, pp. 239 – 248

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 not one of the modern languages presented at that time a sufficiently rich literature to become its rival. But at a later date the Reformation disturbed the unity of the Romish influence. Italian, Spanish, French and English gained successfully regular idioms, and became rich in literary productions of every kind; and at last, 80 or a 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 to that, 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. This 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, special and technical part, where the number of words is limited and the construction 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 of 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.

If we attempt to translate into the Latin of Linnaeus sentences in modern floras, written in English or German, we quickly perceive a want of clearness. In English, the word *smooth* applies equally to *glaber* and *laevis*¹. In German, the construction of sentences indicating generic or other characters is sometimes so obscure that I have found it impossible, in certain cases, 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 language² which should be to all nations what writing is to 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. 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 countries³. 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 17th and 18th 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 [in four] advanced in his study of the language. If a discussion were sustained in French, if books were written or translations 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 the 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. Second, it must be spoken by a 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 50 or a hundred years offer all these conditions united. The language is 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, when it is easy to give in a few words and a few figures.

At the present time the population stands thus (*Almanach* 1871). English-speaking peoples in England, 31 *mln*; in the United States. 40 *mln*; in Canada etc., 4 *mln*; in Australia and New Zealand, 2 *mln*. Total, 77 *mln*.

German-speaking peoples in Germany and a portion of Austria, 60 *mln*; in Switzerland (German cantons), 12 *mln*. Total, 62 *mln*.

French-speaking peoples on France, 36.5 *mln*; in Belgium (French portion), 2.5 *mln*; in Switzerland (French cantons), 0.5 *mln*; in Algeria and the colonies, 1 *mln*. Total, 40.5 *mln*.

Now, judging by the increase that has taken place in the present century, we may estimate the probable growth of population as follows (*Almanach* 1870, p. 1039)⁴.

In England it doubles in 50 years and therefore, in a century, in 1970, it will be 124 *mln*. In the United States, in Canada, in Australia, it doubles in 25 and therefore it will be 736 *mln*. Probable total of the English-speaking race in 1970, 860 *mln*.

In Germany the northern population doubles in 56 – 60 years and that of the south in 167 years⁵. Let us suppose 100 years for the average. It will probably be in 1970, for the countries of German speech, about 124 *mln*.

In the French-speaking countries the population doubles in about 140 years. In 1970, therefore, it will probably amount to 69.5 *mln*.

Thus the three principal languages spoken at the present time will be spoken in a century hence with the following progression, increase from – to, respectively

77 – 860; 62 – 124; 40.5 – 69.5 *mln*.

The individuals speaking German will form 1/7, and those speaking French, 1/12 or 1/13 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 those two countries]. 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 letters and sciences is so much advanced as in Europe, and it is probable that, for a 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 Irish⁶. 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 a 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, grammatical simplicity is 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 a more easy and convenient method of acting or speaking has been arrived at, it is naturally preferred. Besides, civilization encourages individual activity, and this necessitates 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 an ever 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 a building the entrance-door of which is on the third storey. It is no longer a possible language, even in poetry.

Modern languages have not all, to the same degree, the advantage now demanded, of clearness, simplicity and briefness. The French language has shorter words and less complicated verbs than the Italian, and this in all probability has contributed to its success. The German has not undergone the modern revolution by which each 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 sometimes 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*. It makes use only of indispensable and natural tenses: the present, the past, the future and the conditional. There is no arbitrary distinction 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 read⁷.

The pronunciation is not well articulated, not well defined. I shall not go as far as Madame Sand [Aurore Dupin] 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 wrier, is a well-expressed language, just 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. Do 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 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. He replied:

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

Before the events of 1870⁸, a great Alsatian manufacturer sent his son to study at Zürich. I was curious to know the reason why. And he said:

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 a 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 one. I have also observed families where the two languages 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 a 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 than 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 needs only be said that the simpler a language is, the easier it is to be learned, and the quicker can it be made available for practical employment. The English language has another advantage in family use, its literature is the one 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 a 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, Australian and Americans thus appears to be 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 greater need to suppress than to add to the vocabulary. No writer for three centuries past has employed nearly so 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. Why re-establish them? A people so economical in its use of words does not require more than one term for each thing⁹,

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 800 or a thousand millions of civilized men? To other languages it would be as a vast mirror, in which each would become reflected, thanks to newspapers and translations, and all friends of intellectual culture would have a convenient medium for interchange of ideas. It would be rendering an immense service to future races, and at the same time the authors and men of science of English-speaking race would give a strong impulsion to their own ideas. The American, 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?¹⁰

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 reappeared 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 trace¹¹. 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 has itself felt the influence of the universal decline. Americans and Australians possess many flourishing schools, and 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 classical 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 should 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 Britain¹². 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 peoples. 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 organise a meeting of delegates from the principal universities of the Three Kingdoms, 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 probably 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 a book-buying people, and especially given to the study of quasi-scientific [popular scientific?] books, as is proved by the fact of the extensive sale which they command.

In support of this assertion, I may quote the Baron Férussac's review of Wood's *Index Testaceologiens* in the *Bull. Sci. Nat.* Paris, 1829, p. 375. He remarks:

We observe with interest the number of subscribers that exist in England for an octave 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 1000 copies were sold! Since 1829 the sale of scientific books has much increased as is shown, for example, by the many editions of the works of Lyell and other naturalists, each edition being of 1000 copies.

Most scientific 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 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 in English to them, as Thorell on European spiders [...]. The Danes and Dutch often publish their scientific papers in French as [...] who themselves read and write English, but it appears they regard French as the polite language of courts and forget that courtiers generally have a contempt for science, and that they should look among the people for their readers.

It is to be observed that DeCandolle 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 in the *Fortnightly Review* for March 1873, p. 346 which contains some interesting observations on DeCandolle's work

Notes

1. In botany, the word *glaber* means bald, not hairy and is applied in other parts as well as the head. And *levis*, smooth, not rough, but I know they have been both carelessly translated *smooth*, as DeCandolle implies. J. E. G.

2. Artificial languages had been constructed and the most popular is Esperanto, but they did not replace any natural language. A few lines below, concerning the Chinese language, I note that it has seven very different dialects. O. S.

3. A strange statement. O. S.

4. No notice is here taken of the English-speaking people in India and [or] the East. J.E.G.

DeCandolle made a grave mistake by depending on that *Almanach*. Anyway, it is patently impossible to extrapolate the number of inhabitants of a region for a hundred years. O. S.

5. This great difference between the two parts of Germany required an explanation. O. S.

6. Those Irish were mostly uneducated peasants emigrating to the U.S.A. en masse in the mid-19th century because of starvation at home. O. S.

7. Surprised, on one occasion, by the slowness with which intelligent English children learn reading [aloud], 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

8. Until 1870 means until the beginning of the French – Prussian war. O. S.

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

10. This conclusion was in conflict with common sense and history disproves it. And there certainly was no *responsibility* of preserving the language of England itself. Actually, there are three dialects of a single language and seven dialects exist in England itself. And in various places simplified English is used, especially the Pidgin English.

The late Professor Truesdell, a most eminent physicist and mechanic and historian of these sciences, was also a great student of American English. In a letter to me he remarked that he feels himself sitting in the last trench of the defenders of their language against its corruption by newcomers. O. S.

11. The *savages* have indeed all but disappeared. First, the *pale-faced* were carriers of weak forms of deadly diseases unknown to the natives. Second, those natives had been living under invariable conditions and the sudden and violent changes morally destroyed them and led them to excessive drinking (Darwin). Finally, the pale-faced who needed more land actively promoted their disappearance. O. S.

12. This opinion proved wrong.

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

Lyell Ch. (1797 – 1875), geologist

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

Three kingdoms, dated. England, Scotland, Northern Ireland

Bibliography

Almanach (1870), *Almanach de Gotha*. Gotha. This yearly reference book existed from the 18th century to 1944.

Galton Fr. (1873), On the causes which operate to create scientific men. *Fortnightly Rev.*, vol. 13, pp. 345 – 351.

Lalande J. de (1802/1803), *Bibliographie astronomique*, Osnabrück, 1985.

Sheynin O. (1980), On the history of the statistical method in biology. *Arch. hist. ex. sci.*, vol.22, pp. 323 – 371.

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

Alphonse DeCandolle (1806 – 1893) was a botanist, a co-founder of the geography of plants. He therefore possessed some knowledge of

statistics. Here are two of his (French) statements about that science (Sheynin 1980, p. 332; 1986, p. 286):

Numbers should not be accumulated but subordinated to the laws of logic and common sense.

We should understand how to combine numbers and calculate by a method which leads to definite results.

In 1776 both Daniel Bernoulli and Lambert published astronomical papers in German, and Lalande (1802/1803, 1985, p. 539) commented: (French) astronomers should now study German.

I commented on DeCandolle's work in my Notes, and now I only add that in 1885 there appeared a reprint, and in 1911, a German translation of his *Histoire*.

III

E. Czuber, *Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung*, Bde. 1 – 2. Leipzig, 1908 – 1910. Second edition. Review: J. M. Keynes, *J. Roy. Stat. Soc.*, vol. 74, No. 6, 1911, pp. 643 – 647

Prof. Czuber is to be congratulated upon the completion of an exhaustive treatise on the mathematical theory of probability and of statistics, embodying the substance of the greater part of his investigation on these subjects which have been published during the last 27 years. In spite of his 900 pages, his treatment is extremely compressed, and the great length of his book is due to the very wide range which he brings into direct relation with the fundamental theorems of mathematical probability. This new and greatly enlarged edition must long remain the standard treatise on the topics with which it deals. There is no work in English which covers at all the same ground, and it greatly excels in grasp and thoroughness the French treatises which most challenge comparison with it.

By expanding into two volumes the single volume of the first edition (1903) Czuber has been able to find room for considerable addition, and it may be convenient to readers of the earlier edition to indicate briefly the main alterations which have been made. In the first part which deals with the pure theory of some 40 pages have been added which are mainly directed towards strengthening the philosophical side. The second part on the theory of error[s] and the combination of observations, is virtually unchanged.

The third part, *Kollektivmasslehre*, which deals with statistical frequencies, is entirely new. The fourth part, *Mathematische Statistik*, of which only the first section out of three is concerned with general theory. The second section deals with mortality statistics, and the third with invalidity. All through this part the additions are considerable. The philosophical treatment in the first section is enlarged, there is a fuller discussion than before of mortality tables, and the account of the statistics of invalidity and of their relation to those of mortality is mainly new. The fifth part, on the mathematical foundations of life insurance, is also enlarged, and there are, in particular, new subsections on invalidity insurance and on State insurance.

The foregoing summary shows that the work of most general interest is in the three parts which comprise the first volume, and in the first section of the fourth part which covers the first 78 pages of the second volume. The rest of the second volume is mainly concerned with methods of technical detail with which students of insurance and mortality statistics need alone occupy themselves. For such students there is an advantage, no doubt, in bringing this detail into close connection with more fundamental theorems. But the selection of these particular applications to the exclusion of others for very full treatment in what is a general treatise is, from the point of view of the general student, rather arbitrary¹.

With regard to fundamental questions in the first part, Czuber adopts what is probably the best course in a treatise which is mainly

mathematical. He attempts no very searching analysis into philosophical difficulties, but assumes after brief discussions, which are often illuminating, the conclusions which reflective common sense can reasonably expect that philosophers will ultimately justify. He does not solve any of the more perplexing problems in the philosophy of probability, but he almost invariably adopts the provisional conclusions which, at the present stage of the discussion, it seems on general grounds most reasonable to hold. Since the first edition was published he has moved from what we may term German influences, somewhat further from what we may term the *disjunctive theory*, and somewhat further from what we may term the *disjunctive theory* and somewhat nearer to the *relativity theory*.

According to the *disjunctive theory*, which was originally propounded, I think, by F. A. Lange, and which has found a good many supporters in Germany though not elsewhere, probability is based in a very fundamental sense upon disjunctive judgements or hypotheses. According to the *relativity theory*, on the other hand, emphasis is laid rather upon the evidence, on which the probability is based and to which it must be referred. The frequency theory, according to which probabilities have a very intimate relation in every case to statistical frequencies, and which, originally propounded by Leslie Ellis, has found many supporters in England though not elsewhere, Czuber repudiates.

The theory of geometrical probabilities, which furnish the main examples of probabilities for which the number of alternatives is not finite, is dealt with fully. The method of mean values and a very representative collection of examples reproduced from Czuber's early memoir² on this subject is dealt with fully. The paradoxes and contradictions which not infrequently arise in these cases, Bertrand's example, for instance, in which several discrepant calculations can be found for the probability that a chord of a circle taken at random shall be greater than the side of the inscribed equilateral triangle, he attributed to ambiguity in the data, in the interpretation, in the above example, of the expression *a chord of a circle taken at random*. But he is not able to show precisely where the ambiguity lies, or why examples of this type lead to contradictory conclusions in some cases and not in others.

Laplace's Rule of Succession is dealt with carefully and is stated in a form which does not justify the more surprising of the conclusions which have been sometimes derived from it³. But the treatment, at the beginning of the second volume, of inductive probabilities and of their derivation from data of statistical frequency, is not very satisfactory. It is not possible to deal with this far-reaching question in a review, but it seems that Czuber's method, which does not much differ from the treatment of the classical writers on probability, disguises the fact that these statistical inductions do not differ fundamentally from any other kind of induction and permits him to attribute excessively high probabilities on evidence which would be admittedly insufficient in the case of other types of scientific induction.

Consider the following example. In 1866 – 1877 there were born in Austria 4,311,076 males and 4,052,193 females; in 1877 – 1894 there

were 6,533,961 male births; what, on this evidence, are the probable limits of the number of female births? It is contrary to common sense⁴ to conclude on this evidence alone, as Czuber does, that there is a probability of 45,249/45,250 that the number of female births in the second period will be between 6,118,361 and 6,164,813.

In more detail, there are a number of the ingenious examples in algebraic probability, for which the *Educational Times* used to be famous, and which are still sometimes set in examination papers. The solution of example xiii, which gives the probability that, if votes are drawn out one by one from a ballot box containing a votes in favour of A, and b votes in favour of B, $a > b$, A will be ahead at every stage of the scrutiny, is especially satisfying in its simplicity. Of more importance is Czuber's reproduction of Chebyshev's very remarkable theorem, from which Bernoulli's theorem and Poisson's theorem can be derived as special cases. This result is reached rigorously and without approximation by means of the simplest algebra, without the use of the differential calculus. Apart from the beauty and simplicity of the proof, the theorem is so valuable and so little known that it may be worth while to quote the result [the Chebyshev theorem is generally known to specialists and I omit the quotation.]

Czuber makes no reference, however, to any other of Chebyshev's interesting contributions to the theory of probability. Much of his work, which was mostly published previous to 1870, appeared originally in Russian; and although his most important theorems were reproduced from time to time in *J. reine angew. Math.* and *J. math. pures appl.*, it (?) was not easily accessible until the publication at St. Petersburg of the collected edition of his works in French, which was completed in 1907. His theorems are consequently not nearly so well known as they deserve to be.

All through the book there have been added numerous references to the latest German literature on the subject, a feature of very great value to the English reader. With the development of the subject in France and England, Czuber is less exhaustively acquainted. There are some brief references to Pearson's methods for fitting frequency curves to statistical series, and to Edgeworth's recent treatment of the law of error⁵. But of the modern theory of correlation and [or] of the central position which this now holds in English statistical theory there is no hint. This is a very notable omission, for no one is better equipped than Czuber for giving some account of the opinion of Continental thinkers on these modern developments⁶.

But the task of mastering the numerous papers and memoirs, scattered through a great variety of journals, in which the theory of correlation must at present be sought, may well prove too baffling for anyone who has not been brought up amongst them, and perhaps English statisticians ought to wait until they have presented their work in the compact and lucid form in which Czuber presents his, before they can expect German thinkers to pass judgement on it. At any rate a comparison between the subject matter of Czuber's or any other of the recent German treatises on *Kollektivmasslehre* and that of Yule's recent *Introduction to the Theory of Statistics* shows very remarkably on what different lines the best recent statisticians in the two countries

have been advancing. Czuber's methods are in direct line of descent from those of the classical writers on Probability and Error, and they possess the style and lucidity which such a history naturally gives them. But the reader must feel that these methods have reached their limit of accomplishment, and that nothing very novel can result from attempts to perfect them further.

Recent English contributions, on the other hand, fragmentary and often obscure or inaccurate though they now are, seem to have within them the seeds of further development, and to carry the methods of mathematical statistics into new fields. At present the advantage is with Czuber. With sanity of judgement in matters of philosophy and polished mathematical technique, he summarizes for us and completes those modes of statistical enquiry which were evolved during the past century out of the ideas which Laplace and Gauss had originated.

Notes

1. Statistics of population in its entirety is regrettably missing.
2. Cf. Czuber (1884a; 1884b),
3. In translation (Laplace 1814/1995, p. 11):
When an event has happened any number of times running, the probability that it will happen again next time is equal to this number increased by 1, divided by the same number increased by 2.
4. An explanation is lacking. The main objection concerns the highly probable change in the conditions of life.
5. Czuber (Bd. 2, p. 21 and Bd. 1, p. 25, in the reprint of 1968) described those papers, Pearson (1896) and Edgeworth (1901).
6. See however Slutsky (1912). And later Keynes himself (1921) highly praised Chuprov.

The style of the review is bad. The structure of the sentences is not thought out and in many cases it is only with some difficulty that the reader understands to what source is Keynes referring to. And now, three essential points.

The *ballot problem* as was described in the *Educational Times*. Bertrand (1888) studied it. It has many applications (Feller 1950/1968, § 1 in Chapter 3) and Takasz (1982) traced its history to De Moivre.

In 1916, Markov criticized the *correlation theory*. Indeed, without the knowledge of the appropriate parent distribution the sample correlation coefficient is not trustworthy (Sheynin 2017, p. 230).

The classical probability theory has not reached its *limit of accomplishment*. Indeed, that theory is the study of the laws of randomness by considering random variables. The entire development of the theory of probability might therefore be described as an ever fuller use of the power of the concepts of random variable and its expectation. And Markov began to study dependent variables and Markov chains even before 1911, before the date of the Keynes' review.

Bibliography

- Bertrand J. (1888), *Calcul des probabilités*. Second edition, practically coinciding with the first, 1907. Reprints: New York, 1970, 1972.
- Czuber E. (1884a), Zur Theorie der geometrischen Wahrscheinlichkeiten. *Sitz. Ber. Kais. Akad. Wiss. Wien, Math.-Naturwiss. Kl.*, Bd. 90, pp. 719 – 724.
- (1884b), *Geometrische Wahrscheinlichkeiten und Mittelwerte*. Leipzig.

Edgeworth F. Y. (1905), The law of error. In his *Writings in prob., statistics and economics*, vol. 1. Cheltenham, 1996, pp. 325 – 383.

Feller W. (1950), *Introduction to probability theory and its applications*, vol. 1. Third edition, 1968. New York – London.

Kotz S., Johnson N. L., Editors (1982 – 1989), *Enc. of statistical sciences*, vols. 1 – 16 with ingle paging. Hobokan, New Jersey, 2006. Second edition.

Keynes J. M. (1921), *Theory of probability*. In his *Coll. writings*, vol. 8. London, 1973.

Lange F. A . (1877), *Logische Studien*. Iserlohn.

Laplace P. S. (1814, French), *Philosophical essay on probabilities*. New York, 1995. Translator A. I. Dale.

Pearson K. (1896), Skew variation etc. *Phil. Trans. Roy. Soc.*, vol. A186, pp. 343 – 414.

Sheynin O. (2017), *Theory of probability. Historical essay*. Berlin. **S, G**, 10.

Slutsky E. E. (1912), *Teoriya korreliatsii*. Kiev. **S, G**, 38.

Takasz L. (1982), On the classical ruin problems. In Kotz & Johnson (2006, vol. 1, pp. 183 – 188).

Yule G. U. (1911), *Introduction to the theory of statistics*. London. Many later editions.

IV

John Cournos

London under the Bolsheviks

Russian Liberation Committee, [Publication] No. 4, 1919

Foreword

Thus is not a fantasy, in spite of the title. It is a true and precise picture of Petrograd during the first few months of the Bolshevik Revolution. The author has placed the scenes in London merely to emphasize the realities of the Bolshevik nightmare, to bring it home to those who do not quite realise the nature of the Russian upheaval. If the result appears fantastic, let the reader beware of blaming it on the imagination of the author. The author would feel flattered, were it not that, unfortunately, he is too well aware of what his eyes have seen.

After all, the first few months were but a mild prelude to the orgies of horror which followed. Other witnesses, who have arrived more recently, are in a position to make the scenes set forth here appear heaven-like in comparison.

But there will always be doubting Thomases. It is to be hoped that the opportunity will never come for their eyes to be opened, for the sake of those millions who would be the inevitable sufferers.

On returning from Russia lately, having endured the discomforts of life so long, both in Russia and on my journey, my first thought was to enjoy a few luxuries of civilized life. From a good hot bath I therefore proceeded to a good hot meal and from the meal to the tune of four or five glasses of port (I kept no account) to what appeared to me a luxurious bed. Perhaps I had eaten too much, perhaps I had had too much port (I don't know which it was) but I had a dream. A dream? Perhaps I had better call it a nightmare. I dreamt that London was under the Bolsheviks.

I dreamed that one sunny winter morning I was walking down the Strand. It was undoubtedly the Strand, for who had seen the Strand once, can fail to recognize it again? And yet there was something about the street that made it seem strangely and utterly different. It wasn't altogether that the buildings appeared faded and shabby in the sunlight, looking as though they had not been scraped or painted for a long time, or that the street was sloppy and dirty, being filled with little mud banks of uncleared snow and little pools of dirty water which collected in the caved-in pavements. No, it wasn't that. Though, to be sure, the dirt was appalling.

It was perhaps the atmosphere of the place that was different. I am not referring to the stench, which was indeed horrible, but rather to the general character of the street, hard to describe. Instead of the usual steady outpouring, in both directions of streams of people with bright cheery faces, the sidewalks were filled with loafers standing about inert, hands in their pockets, and with shabbily dressed, hungry-looking men and women moving in a slothful pace and an aimless

kind of way. At intervals I ran into small groups of people, mostly workmen and soldiers, discussing something, interrupting one another, and gesticulating in an alarming fashion, at times almost coming to blows.

All the faces were sullen, and not a few threatening. Some looked at me with curiosity, others with open hostility. For a long time I wondered why, until observing a pair of ferocious eyes glancing at my new bright tan boots with particular fierceness and avidity. I suddenly realized that this general scrutiny of me by all passers-by was due to my being dressed better than the others, also to the fact that I wore a clean white shirt and a collar. Whereupon, desiring to attract less attention, I deliberately waded through one of the puddles, while a passing omnibus coming along squeaking for lack of oil, covered the rest of me with its flying smirch. In trying to escape this foul shower I slipped and fell, and, turning my ankle, I gave a cry of pain and made a wry face. Just then an omnibus stopped to unload passengers, and everyone in it looked at me and laughed, not in the usual good-natured way but maliciously. And no one came to give me a hand, no one seemed to care how long I lay in my helpless state. I had time to observe the omnibus. It had evidently once been red, but was now a dirty faded brown, streaked with rust. The passengers filled not only all the seats, both inside and out, but the gangways and platform as well, they also sat on the steps leading to the top, and every passenger wishing to get in or out had to fight his way. When it started away again, squeaking as before, four or five passengers were hanging on the platform step. A soldier running to catch it knocked over a woman and did not stop to apologise. She filled the air with loud imprecations. The crowd laughed. That woman and I, lying in the dirt, might have been no more than two poor mongrels.

At last I picked myself up and walked on, first having helped the woman up, much to her amazement. What struck me as strange was the number of kerb merchants, men, women and children, ranged along the whole way, selling newspapers, chocolate, cigarettes, matches, etc. Among these merchants were a number of well-dressed, refined-looking women, and Army officers wearing the Victoria Cross and other honours, but without their shoulder straps. I approached one of these on the pretext of buying a pack of cigarettes, but actually to discover the reason of their degradation.

I picked up a ten-cigarette packet, and, seeing the price, one shilling, marked on the box, I took a half-crown out of my purse, and handing it to the vendor, waited for my change. The man examined the coin with astonishment and said:

It's a treat to see real money again; all the same, I must trouble you for another five bob.

It was my turn to be astonished:

What, seven-and six for ten cigarettes! And the packet is marked only one shilling.

That's the old price, but in these illegal days ... Besides, I stood in the queue for about four hours to get them. Ah, we are an unfortunate people, a benighted people.

Much puzzled by his words, I fumbled in my pocket, and at last drew out of my wallet a five-pound Bank of England note and said:

I am sorry, my dear man, but this is the smallest I have.

To my great astonishment he remarked:

That's just about enough to get one a couple of meals without wine, mind you. Well, you know what money is, rubbish.

While he was talking, he put his hand in his pocket and drew therefrom a little pile of green and yellow notes, each no larger than three by two inches. He picked out two of them and gave them to me, then plunged his hand into a bag at his side and drew out a handful of dirty, dilapidated notes and postage stamps of all denominations. Seeing me look mystified at the green notes in my hands, each of which, I had time to learn, was worth forty shillings, the man muttered apologetically:

I am sorry to have to give you these Macdonalds, but it's the only kind I've got. The old notes are getting scarcer and scarcer.

Then he proceeded giving the rest of my change in sixpenny notes and in postage stamps, which I noted had no gum on the back, but bore instead an inscription to the effect that they were available as currency in place of the customary copper coins.

More and more astonished, I begged the soldier-merchant for an explanation.

What do you mean by calling these green notes the Macdonalds?

Hardly less astonished was my informer at my question:

You must be a stranger in these parts. Just arrived, I suppose, from some happier country than this. Surely you've heard of our Revolution, how our royalty and the old Government were overthrown and a Provisional Government appointed, with Ramsay Macdonald at its head. These are called Macdonalds because they were issued during his administration.

You say were, from which I may assume that he and his Government are no longer in power?

No – worse luck! They were bad enough, well-intentioned but weak. They dilly-dallied with everything and made a mess of things generally. But they were nothing compared with the present Government. You see, when the second Revolution came, Macdonald and his advisers were overthrown and MacLenin and Trotsman took their places. Macdonald fled to Scotland, and came back with an army under General Haig, but was defeated when he reached St. Albans. He has fled, and is hiding somewhere. MacLenin's crowd have so far arrested about 500 men who resemble him.

Thunderstruck by these unexpected tidings, I said:

Tell me. I see that you have the Victoria Cross, the D.S.O., the Military Cross and the Legion of Honour. How is it you are not wearing your shoulder straps, and how have you come down to such a mean trade?

Here came the sad response:

Well may you ask the question, but I am not the only one. There are hundreds like me in the same boat. The comrades, as they call

themselves, have done that for me, and I've saved about 50 of the bounders myself when the Germans were upon us near Amiens, so help me God! You see, when MacLenin and Trotsman disorganised our fine army with their pernicious propaganda, they preached to them that all men were alike, that one man was not better than another, and that sort of thing. And so we officers, every man of us, had our stars taken from us by those comrade fellows. I was a captain, and I've come to this. But we've come off easily, compared with some of the famous civilians. There's Viscount Grey, and Asquith, and Lloyd George and ... every blessed one of them is kept in the Tower, they are called counter-revolutionaries and liable to be shot any day. And not only they, but all the cultured chaps are having a hard time of it. Well, I dare say you've heard of H. G. Wells – to think of a brainy man like that in jail, counter-revolutionary, that's what he is called. It would be easier to give you a list of those who are not in jail. Look at that nice girl over yonder selling papers. She spent all her time doing good among the poor. And see that chap over there in spats selling shoestrings, he invented ---

Suddenly the rapid staccato of machine-gun fire sounded through the air, drawing nearer. There was commotion among the crowd, until then listless.

My acquaintance pulled me along by the arm down an Underground entrance:

Take cover. It's those comrade bounders coming.

In spite of expostulations, I remained standing on the stairs and glanced furtively down the street, which had suddenly become deserted. An extraordinary sight met my eyes: a *tank* came rolling up, as it were, at full speed, firing its guns indiscriminately in all directions. Its occupants were invisible. I ducked quickly below. And asked turning to my acquaintance:

What is the meaning of all this?

Meaning of it? Nothing! It's just a couple of comrades out to give the people a scare and to show them who's boss of the town.

He pronounced the word comrades with contempt. We ventured up the stairs again. He pointed to the tank, now in the distance. It appeared to be going slowly now, in zig-zag fashion, still firing its broadsides,

Look! They are probably drunk. There has been a sacking of wine-cellars this past fortnight or so. They do what they like and there's no one to stop them, for they've got all the weapons. And there's no police. Five thousand robberies at night, my friend. I'd advise you, old chap, not to go out at night in your best clothes. A good suit of clothes fetches as much as 50 pounds, and these comrade chaps don't stop at undressing you in the street, and leaving you in your undershirt, be the night ever so cold. They shoot you afterwards if they don't like your face. They can get more bullets where the first came from. They have no respect for women either. I've known them to drag a woman out of a taxi for to get her glad rags, as the Yanks would say. What can you expect when you have an ex-criminal for Chief of Police and Conscientious Objector as General of the Red Army? That's a strange

thing about some of these pacifists: they objected to killing Germans, but they don't mind killing Englishmen. Red Army is a good name for them. Only this morning I ran across a pool of blood in Trafalgar Square. The usual thing. People are foolish to be walking in such a lonely spot at night.

A great fear began to gnaw its way into my heart. Much depressed, I left my companion and made my way towards the Embankment. For the first time I noticed that many buildings were decorated with flags, which, judging from their appearance, must have been at one time red, but which were now faded by the rains to a dirty black, and were so battered as to make their original shape almost unrecognisable. And, gazing upon these once-proud emblems of revolution, I reflected:

They must have been once as bright as the hopes of the people, and the people's hopes must be now as faded and as tattered as these flags.

Such, as I began more and more to discover, was indeed the case. Upon reaching the Embankment, I barely had time to glance at the wonderful river, which had been made historic by English free men, and had hardly more than noted its now processions of barges that I once used to know, when rude sounds of lashing as with a whip assailed my ears, sounds punctuated by loud, foul curses. I turned to look, and to my amazement witnessed a sight I had never witnessed in London before, to wit: a thin, quite emaciated horse, which had evidently slipped under its heavy burden on the ill-kept pavement, was lying with its legs bent under it, and, straining under the regular and well-laid-on full-swing lashes of the driver's whip, it was making desperate efforts to rise, without success. And this failure on the beast's part aroused those foul-mouthed imprecations already alluded to, impossible of repetition here. Incensed at the brutal treatment of the helpless beast, I wished to cry out, I wished to run and put my fingers on the driver's throat, but, as happens in dreams, I remained rooted to the spot, unable even to cry out. Presently I heard a voice, that of a bystander;

The rotter! He wouldn't have dared to do that during the old regime. That's what comes of not having any police. Any old bounder can do as he likes, and there's no one to stop him.

An old gentleman with a kindly face that belonged to gone-by days intervened:

The poor old beast has worn out his shoes, that's how he came to slip. And new shoes cost a fortune. It's pretty rough on the horse, whichever way you look at it.

A little old woman in an anxious voice asked:

What's become of the Society for the Prevention of Cruelty to Animals?

A man's gruff voice replied:

Disbanded, of course. Counter-revolutionaries.

The crowd grew, and everybody talked, but the man kept on beating the horse. Sickened by the sight I moved on. Had I known that I was making for more gruesome sights I would have remained where I was.

The whole Embankment had an unfamiliar appearance. Some of the railings, which in the old days I often had leant upon to look at my

beloved river, were broken down, the Embankment itself was badly in need of repairs. Apart from this, at short intervals, the sidewalk was piled up with discarded tins and all sorts of refuse. The river too looked dirty, and large collections were floating with the tide. With horror I noticed a dead human body floating peacefully along, seawards.

My reflections were again suddenly disturbed, this time by rifle-firing. I ran with the crowd towards Westminster Bridge. I paused for a moment with many other spectators near the railing, and what I saw chilled me to the marrow.

A man had just been thrown from the bridge. Striking the water, he disappeared, then came up again, and was swimming for the Surrey side bank. Upon seeing the man swim, the men on the bridge began to throw sticks and stones at him. Eluding these, he swam on. Then one or two shots were fired which missed him. Presently he began to clamber up the muddy bank close to the new County Council Hall. Whereupon the men on the bridge, evidently not wishing to be robbed of their prey, ran down to the bank, and opened a fusillade on their struggling victim. Wounded, he nevertheless managed somehow to get on his knees and began to pray to his tormentors to spare his life. Unheeding, they continued firing, until they finished him.

Astonished at the whole proceedings, I turned to a bystander and asked him why the poor man was so ill-treated.

The usual thing, robbed somebody, I suppose. What else are you to expect when there are no police and no courts of justice? We are under mob law, my friend, When it isn't the mob, it's the Red Guard. There's little to choose between them. Why, sometimes they are not even sure they've got the right man. Merely to suspect a man is enough. That's what they call summary justice. It's worse than in the Middle Ages. Just look how they had riddled the man with shot. Well, you know, in the Middle Ages, if a man was condemned to death and he had managed to save himself in some way or other, they looked upon it as an intervention of Providence, and let the man off, and even today there are places where if a man is condemned to be hanged and the rope breaks, the man's let off. But you saw for yourself: they meant to drown the man, they chucked him into the river, he saved himself. But no, that wasn't enough, they had to shoot him. That's what we have come to in this 20th century. Of course, the Bolshevik authorities pretend they don't look with favour upon these dealings, all the same they do their best to encourage them.

An interruption came in the man's speech, for rifle-shots rang out again. Most of the crowd scattered. I and the man who had been talking to me took refuge in a doorway just across the road, close to the Westminster Underground Station. We ventured to peep out from behind one of the pillars which supported the door arch. But for the whispered explanations of my companion what I saw would have been a mystery to me.

A long procession of men and women, evidently of all conditions of life, quiet and grave in appearance, bearing large red banners, was moving with dignified slowness from the direction of Victoria Street towards the Houses of Parliament. The banners bore such inscriptions

as *Hail to the Popular Assembly, The Popular Assembly will save England, All Parties Unite Round the Popular Assembly*, and the like.

To my utter amazement, I saw several men, hardly more than a dozen in number, each wearing a red band on his arm (which, I was to learn, was the mark of the Red Guard), advancing, in a crouching position, gun in hand, towards the approaching procession and shooting into it. Apparently, the crowd that formed the procession was unarmed. At all events, no one offered any resistance. Indeed, upon seeing some of its members fall, it scattered precipitately.

I inquired of my companion:

What is the meaning of all this? Why are these men shooting into the crowd, which seems harmless enough?

The man replied sadly:

Surely you know that the Popular Assembly elected by nearly all parties to save England, was to meet today in the old House of Commons, and these men and women, representing the best elements in England, were making a peaceful demonstration in honour of the event, but the Bolshevik Red Guards, representing a mere minority consisting mostly of the worst elements, have got orders from MacLenin and Trotsman to keep the Assembly from meeting at all costs. They have forbidden all demonstrations in favour of the Assembly.

I don't understand. You say the majority in England has decided in favour of the Assembly. How is it possible then for a small minority to controvert the will of the people?

He replied apologetically, as if his own honour were in question:

What are we to do? They've managed to get hold of the war machinery. They have got hold of all the arms. We can do nothing against force. Besides, they offer the attractions to the criminals and rowdies, who terrorise the population. They are wolves in sheep's clothing, for on the pretence of taking care of the interests of the downtrodden poor they are filling their own pockets with boot. Their one object in life is to live without working. That's all very well, while there's something left of accumulated wealth, but what will they do when that's gone? Look at my case. I was comparatively speaking a poor man before the Revolution. I had a printer's shop, and because I employed a printer's devil I was immediately dubbed employer of labour, and, therefore, a boorjooy, that is, member of the bourgeoisie, hence entitled only to a quarter of a food ration.

I have worked all my life hard like a nigger and have managed to get for my family a small house of five rooms, and what's all my efforts come to? The new Bolshevik House Act means that I shall have to give up to these brigands and good-for-nothings three of my rooms. Yet there was a time when an Englishman's house was known to be his castle.

I remarked sympathetically:

That's hard lines my friend. Come and have a drink with me.

My acquaintance replied:

I should like to oblige, but a drink is not to be had for love or money. To sell drink is illegal, though to be sure it's done. It's under

the head of the Defence of the Revolution Acts, commonly known as Dora Number Two. The first Dora whom she has displaced, is an amateur compared with this new creature. There are only two ways of getting a drink, either by paying a small fortune by way of a back door or by joining the rowdies who are sacking the wine-cellars. There, do you hear? Those shots must come from Piccadilly. They are sacking the wine cellars there today. Come, let's see the fun. But you must keep a sharp eye and dodge a stray bullet. When you hear shots it means that they are drunk and letting off their guns just for a lark.

It had grown dark by now. The streets, dark enough in war time, seemed even darker than usual. My companion informed me that this was due partly to the disorganisation of the railways which had failed to bring in the requisite fuel, partly to the fact that the working-men were holding another of their meetings that day to discuss the extermination of the *boorjoos*.

We made our way along the narrow streets, dodging here and there small but reckless parties of men with guns. We strayed into a restaurant. The place was dark and cold for want of fuel, but dimly lit up with candles, which were struck in empty bottles standing on the tables. It was filled mostly with *comrades*, though here and there a pale *boorjooy* might be distinguished among them. We found a table.

Waiter, I called. Everyone turned and looked fiercely in my direction. My companion whispered:

Good God! You mustn't use that word. The Bolsheviks have decreed that you mustn't call a man a waiter, but an officiator. It's ridiculous, but you must conform if you value your life.

So I called Officiator! This seemed to calm my neighbours.

The officiator officiated. That is to say, he brought us everything that my purse would permit: a cupful of a dark, dirty-looking liquid which he called coffee; no milk was available, and a piece of bread two inches square with a small pat of margarine. For this I was charged for my friend and myself 16s, so far had money depreciated in value. The stench and the smoke of the place were horrible. Presently a newsboy came in shouting the names of his papers: the *Red Gazette*, the *Red Standard*, the *Red Voice*, the *Red Dawn*. Every newspaper seemed to be Red in this new England.

Upon inquiry I learned from my vis-à-vis that the old Conservative papers like *The Times* and the *Morning Post* had been suppressed long ago, that the liberal papers like the *Manchester Guardian* and the *Daily News* were not long ago in following suit, that even such Radical Socialist newspapers as the *Herald* and the *Labour Leader* were now considered by the new political leaders hopelessly *bourgeois* and *counter-revolutionary*. I was further to learn that the editors of these newspapers were in the Tower of London awaiting trial by the Revolutionary Tribunal. Freedom of speech had ceased to exist.

The things I had learnt in the course of the day depressed me, and, holding what would have been considered in the old days as Liberal, even Radical opinions, I suddenly realised that I would now be deemed by the new regime to be an excessively dangerous character for holding opinions so palpably Conservative and old-fashioned.

With this fear upon me I decided to visit an old friend of mine, and seek some comfort in his counsel.

With fear and trembling I made my way cautiously through the dark streets, for I could still hear an occasional rifle-shot, and at last reached the Queen Alexandra Chambers, now renamed the Karl Marx Flats. The large double doors, open at this hour in the old days, for it was not much later than half-past eight, were now shut, indeed locked. In response to my ring a tiny door, no larger than that of a cage, opened just above the letter-box, and through the opening peeped out the gleaming barrel of a revolver.

Don't shoot! I cried out in alarm. The revolver retreated into the darkness behind the door. *Who is it?* said a voice which I recognised as that of the housekeeper, an old man.

It is I, Mr. S..., don't you remember me? Mr. Thornton's old friend. Does Mr. Thornton still live here?

An eye looked out at me through the little aperture. The old man said more softly: *So it is really you? And there is no one with you?*

The key turned in the lock, and the door opened, admitting me.

I thought it might be a band of those comrades. You can't be too careful nowadays. This sentry business is no joke, let me tell you. These are dark days, governor.

The lift not running, I stumbled up the dark staircase. My friend lived on the fifth storey. I rang the bell. For some time there was no answer. My hearing being particularly acute, I heard, at last, someone approaching, as if in slippers on tip-toe. I knocked. I heard a timorous feminine voice on the other side of the door: *Who is it?* I replied in reassuring tones *A friend.*

The door opened slightly, on a chain. A woman's frightened eyes looked out in the candle-light. I recognised my friend's wife. *Oh, it's you!* She cried, recognising me at the same time. *I was so frightened. I thought it was ... The comrades!* I finished the sentence for her. She let me in, and locked and bolted the door.

My friends found me a place at their supper table, and put a plate of thin vegetable soup and some bread crumbs before me. My host said most apologetically:

I am sorry I have nothing better to offer you. I live like a bird of the fields.

You mean to say that you, a great literary celebrity, who have given England of your best, have come to this?

The great man smiled sadly.

Yes, my friend. I am regarded as a parasite, a boorjooy, a counter-revolutionary. I am a man past 50, all my work is behind me. My contribution to literature counts for naught, simply because I cannot subscribe to the political opinions of these proletariat tyrants. Let me whisper a secret to you. These men hate the artist and the man of intellect far more than they do the capitalist. For they can take away the wealth of the rich man, as, in fact, they already have done, but they cannot take away for their own from the inexhaustible capital that is in a man's brain. They can only shoot me, which they probably will do, but they cannot confiscate my intellect.

He laughed ironically, and there was a suspicion of triumph in his voice: *My brain shall die with me*. He invited me to remain for the night, suggesting that it was dangerous to be out after dark. I had not lain in bed long, thinking over the day's adventures, when I was startled by a loud knocking on the outer door. Presently a number of men burst into my room, and, ordering me to dress myself, hustled me into a taxi, and drove off. It seemed ages, that journey in the taxi. We arrived at a building, which I recognised as the Tower. I was taken up the little spiral staircase of stone, leading to a famous prison-cell, which I remembered once having paid sixpence to see, and left there.

I don't remember how long I remained there. The time seemed immeasurable. Then I was led down again, and into the courtyard. I was placed against the wall, and found myself looking into the barrel of a revolver. It must have been but a moment, but again it seemed like ages. I cried out ... I woke up. ... The perspiration ran down my body in streams ... I took it for blood at first,

Now I was fully awake. I put on my dressing-gown and my slippers, and walked over to the window. I looked out on the Strand. There was nothing unusual in the street. The bright red omnibuses sped their way in their accustomed fashion, and streams of people moved restlessly on and on. A Union Jack fluttered in the breeze, and caught the glint of the sun. I rang for the waiter, and ordered bacon and eggs to be brought into my room. It was good to be back in England.

An announcement by the Russian Liberation Committee (RLC) was placed on the next page. Here it is:

If you are interested in the overthrow of Bolshevism; if you want to help to save civilization from the menace of Anarchy, write to the Literary Secretary, RLC, [address, telephone].

Some explanation

Asquit (1862 – 1928), a state figure

Gray (1862 – 1928)

Douglas Haig (1861 – 1928), field marshal

Ramsay MacDonald (1866 – 1937), a political and state figure, prime minister (three times)

Amien, France, place of great battle of 1918

Orders: The author lists three English orders (DSO = Distinguished Service Order) and a French order

Red Gazette, a Leningrad city newspaper, the only Soviet periodical to publish Chuprov's obituary

Printer's devil, a wretched fellow employed in a printer's shop

Tank in inverted commas: the word was still unusual

The story is ridiculous. The year was 1917 and the world war was still going on, but London was its usual self (end of paper).

I do not mention minor misunderstandings, but here are the real faults

Officers of the pre-revolutionary army wearing orders. (Readers ought to replace mentally foreign by Russian orders). Quite

impossible: soldiers demoralised by *MacLenin* would have killed them at once.

Irrespective of the exact meaning of *Popular Assembly*, killing participants of a demonstration to support it was hardly imaginable, although cruelties were commonplace. Just the same, shooting an innocent foreigner was only possible as a most serious mistake. A literary celebrity apparently avoided even a hint of criticism and would be left to die of the harshest conditions of life.

Depreciation of sound foreign money seems highly improbable. It is well known that Chuprov had been living in ruined post-war Germany for about five years by highly valued royalties for publications abroad (beyond Germany).

The whole story stinks to the skies.

Harald Cramér

Richard von Mises' work in probability and statistics

Annals math. stat., vol. 24, 1953, pp. 657 – 662

Professor Richard von Mises of Harvard University died in June, 1953, shortly after his 70th birthday. A native of Austria, he took his doctor's degree at the Technical University of Vienna in 1907, and then acted as a lecturer and professor and director of the Institute for applied mathematics of the University of Berlin. The Hitler regime, depriving the German universities of so many of their best men, brought von Mises to Istanbul, and finally, in 1939, to Harvard. There, he first was professor of mathematics, and in 1944 became Gordon McKay professor of aerodynamics, and applied mathematics.

Von Mises was of those men who have both the ability and the energy requisite for taking an active and creative interest in many widely different fields. He has made outstanding contributions to subjects as heterogeneous as literary criticism, positivistic philosophy, aerodynamics and probability. In this short notice we shall be concerned exclusively with those of his works that belong to the field of probability and mathematical statistics.

It is well known that Mises is one of the significant names in the history of the tremendous development that has taken place in this field during the last thirty years. As can be seen from the appended *Selected Bibliography*, his works on probability and mathematical statistics range from books and papers on the foundations of probability which, of course, always represented one of his main interests, to investigations dealing with special problems in various statistical applications. Only a few of these works can be mentioned here, but it will be attempted to characterize and to follow up some of the main lines of thought, along which his contributions seem to group themselves.

In 1919 the two basic papers (37) and (40) which were practically his first publications on probability appeared almost simultaneously. The first of these was concerned with the general theorem in mathematical probability for which a year later Polya was to propose the now well known name, *central limit theorem* (CLT). The second paper, on the other hand, gave the first exposition of Mises' views with respect to the foundations of probability theory. Each of these two papers was to become the first in a long series of works, and the two groups of works thus initiated may perhaps be looked upon as containing the most important of Mises' contributions to the subject.

To judge the two basic papers correctly, it is necessary to realise the situation in mathematical probability theory about the year 1919. Since the appearance of the classical treatise of Laplace, a few mathematicians, Chebyshev, Markov, [Liapunov], Borel, and some others had done important work in the field, but the conceptual foundations on which the whole subject rested were still obscure.

There was no commonly accepted definition of mathematical probability, and in so far as there were any definitions at all, they were clearly inadequate for the numerous applications that were made in fields such as population statistics, molecular physics, and many others. Moreover, with few exceptions, mainly belonging to the French and Russian schools, writers on probability did not seem to feel under any obligation to conform to the standards of rigor that were regarded as obvious in other parts of mathematics. The admirable work of Liapunov on the CLT seemed to be entirely unknown among mathematicians.

In the introductions to his two abovementioned papers Mises reviews the situation and arrives at the conclusion which seems entirely justified:

Today, probability theory is not a mathematical science.

He then develops his own programme for building up probability theory as a mathematical science. He starts from the thesis that

Probability theory should be regarded as the mathematical theory of a group of observed phenomena, in the same way as, for example, geometry and theoretical mechanics.

Just as, for example, geometry gives an idealized mathematical picture of the large bulk of our observations with respect to the configuration and position of bodies in space, so probability theory should be constructed to provide a mathematical model of the statistical regularities observed in cases where a given experiment or observation may be repeated a large number of times under similar conditions.

Starting from this thesis, Mises (40) develops his system of foundations, which soon became familiar to all probabilists. We find here the concept of a collective¹, the definition of mathematical probability as the limit of a frequency ratio, and the two fundamental postulates, requiring the existence of the limiting values of the relevant frequencies, and their invariance under any place selection. It is shown how the main rules for operating with probabilities can be deduced from these basic principles, and a system of classification of the operations used in probability theory is worked out.

The publication of (40) aroused a great deal of interest among mathematicians, statisticians and philosophers. Quite naturally, opinions were divided, and even if the basic view of probability theory as a mathematical theory of random phenomena was, on the whole, completely endorsed by most mathematicians and civilians, the collective concept and the two postulates were severely criticized by many authors.

An extensive literature grew up about these questions, and Mises himself took an active part in the discussion. Besides, in a number of papers dealing with special problems particularly related to the second postulate, the foundations of the subject are discussed in his two treatises (75; 127) and, above all, in his well known book (64) which has been translated from the original German edition into English, Russian and Spanish. There, he gives a detailed exposition of his system, intended for non-mathematical readers, and also his replies to

various criticisms, and his comments on alternative systems proposed by others.

It is particularly interesting to read in his book, as well as in the discussion (117) with Doob, his comments on the measure-theoretic approach to the subject favoured by a certain number of contemporary mathematicians and statisticians. Even though this approach, as pointed out, for example, in the well known book by Kolmogorov, starts from a conception of the object and character of probability theory, which is very close to the one advocated by Mises himself, he takes a strongly critical position against this method of introducing the concept of mathematical probability and formulating the axioms expressing its basic properties.

In a chapter [of (64)] with an expressive title *A part of the theory of sets? No!* he declares that probability theory

Remains in all circumstances a theory of certain observable phenomena, which are idealized in the concept of a collective.

So far, many of his opponents would be prepared to agree, at least partly. But then he goes on to say that, from this point of view, he finds it impossible to

Concede the existence of a separate concept of probability based on the theory of sets, which is sometimes said to contradict the concept of probability based on the notion of relative frequency ... there can be also no question of reconciling these two concepts.

This seems to be a final expression of his standpoint in the question. The discussion will undoubtedly be continued during many years to come, but however the question of the best choice of axiomatic foundations of probability theory may be decided (if it will ever be decided at all), it will always stand out as the great achievement of Mises to have been the first to draw general attention to the problem, to have indicated the way along which its possible solutions should be sought, and to have given one such solution.

Let us now pass to the second main group of Mises' writings in our field, the one that begins with (37). As already mentioned, this paper is concerned with the CLT and with certain other problems belonging to the same general range of ideas. A proof of the asymptotic normality of the distribution of a sum of independent random variables is given under certain rather restrictive conditions. There is a detailed discussion of the asymptotic behaviour of the distribution of the sum in the case when the distribution of the terms belongs to one of those simple classes that are usually encountered in the applications. Similar results are given in respect of the asymptotic behaviour of the posterior distributions obtained by applying the Bayes theorem to a sample of observations.

Mises returned to this subject in a great number of his works, continually improving his results and extending the field of problems considered. The most important papers belonging to this group are (87; 89; 102; 105; and 126). With respect to the CLT itself, his main results were superseded by others, but he soon generalized the problem in a very interesting way where he was able to find important new results, and which still seems to open possibilities for further research. We shall briefly characterize the problem considered in this

group of his writings, which may be said to culminate in (126). Let $U(x_1, \dots, x_n)$ be a symmetric function of the independent random variables x_i , which are assumed to have a common distribution function $F(x)$. (Mises considers the general case of unequal distributions.) Then U may be regarded as a function $V(S_n)$ of the repartition function $S_n = S_n(x)$ where $nS_n(x)$ denotes for every x the number of $x_i \leq x$. It is assumed that the function V can be defined on a convex domain D in the space of all distribution functions including the given distribution function F as well as all possible repartitions S_n with $n = 1, 2, \dots$. It is required to study the asymptotic behaviour of the distribution of random variable $V(S_n)$ as n tends to infinity.

Following Volterra, Mises defines the derivatives of the function V at any given point of D and shows that, subject to certain general regularity conditions, the distribution of V is asymptotically normal if the first derivative of V at point F is different from zero. This covers, among others, the classical case of the sum of the x_i and also most ordinary statistics based on moments. When the first derivative vanishes, certain non-normal limiting distributions are obtained. Mises discusses in detail the case (including among others) the chi-squared statistic) when there is a non-vanishing second derivative. The limiting distribution in this case is shown to be intimately related to the Fredholm determinant of a certain symmetric kernel. As already mentioned, interesting problems in this direction still seem to be open for research.

Finally, we shall only briefly mention some other main groups among the works of Mises. In (96; 100; 111 and 112) the relations between various moments and the characteristics of a probability distribution and between these characteristics and the values of the corresponding distribution function are studied and some important inequalities are given. The papers (109; 119 and 121) are concerned with the Bayes theorem and its various applications, a subject which also receives great attention in (75) and (127).

Mises did not sympathize with the tendency in contemporary mathematical statistics to avoid the use of the Bayes theorem and [or] the concept of prior probability². In the works belonging to this group, he discusses the application of the theorem to various problems including the problem of testing statistical hypotheses, where, according to him, it leads to more reliable results than the methods now currently employed by mathematical statisticians.

As a glance at the list of publications will show, there are many works of Mises in this field that have not been mentioned at all. The majority of these papers are [is] concerned with various applications of probability theory, in fields as diverse as physics, genetics, demography and actuarial science.

This brief and insufficient review of a small part of the writings of Richard von Mises will certainly be enough to give the reader a strong impression of an active and powerful scientific personality. Those who knew him saw, in addition, many other sides of his personality, giving him a human charm that his friends will never forget.

Selected Bibliography

I am greatly indebted to Dr. Hilda Geiringer Mises for kindly compiling this list of publications. The numbers indicate the position of the works within a complete list of the writings of Mises,

[This list was apparently based on Mises' own collection of manuscripts and reprints. A slightly more detailed list of his publications (*Sel. Papers*, vol. 2. Providence, RI, 1964) was based on that same collection. It was too difficult to check whether or not do the numbers in both lists coincide, so I have to reprint the first list (appended to Cramér's paper). Those *Selected Papers* contain 20 items.]

Abbreviation. AMS = *Annals math. stat.*

14. Über die Grundbegriffe der Kollektivmasslehre. *Jahresber. deutschen math. Vereinig.*, Bd. 21, 1912, pp. 9 – 20.

33. Einfache und exacte Ableitung des Maxwellschen Geschwindigkeitsverteilungssatzes, *Phys. Z.*, Bd. 19, 1918, pp. 81 – 86.

35. Über die Ganzzahligkeit der Atomgewichte und verwandte Fragen. *Phys. Z.*, Bd. 19, 1918, pp. 490 – 500.

37. Fundamentalsätze der Wahrscheinlichkeitsrechnung. *Math. Z.*, Bd. 4, 1919, pp. 1 – 97.

39. Marbes Gleichförmigkeit in der Welt und die Wahrscheinlichkeitsrechnung. *Naturwissenschaften*, Bd. 7, 1919, pp. 168 – 175, 186 – 192, 205 – 209.

40. Grundlagen der Wahrscheinlichkeitsrechnung. *Math. Z.*, Bd. 5, 1919, pp. 52 – 99.

44. Ausschaltung der Ergodenhypothese in der physikalischen Statistik. Bd. 21, 1920, pp. 225 – 232, 256 – 262.

48. Über die Wahrscheinlichkeit seltener Ereignisse. *Z. angew. Math., Mech.*, Bd. 1, 1921, pp. 121 – 124.

49. Das Problem der Iterationen. Ibidem, pp. 298 – 307.

53. Über die Variationsbreite einer Beobachtungsreihe. *Sitz. Ber. Berliner math. Ges.*, Bd. 22, 1923, pp. 3 – 8.

61. Über das Gesetz der großen Zahlen und die Häufigkeitstheorie der Wahrscheinlichkeit. *Naturwissenschaften*, Bd. 15, 1927, pp. 497 – 502.

64. *Wahrscheinlichkeit, Statistik und Wahrheit*. Wien, 1928, 1936, 1951. English, Russian and Spanish translations: London, 1939; Moscow, 1930; Buenos Aires and Mexico, 1946.

67. Über kausale und statistische Gesetzmässigkeit in der Physik. *Naturwissenschaften*, Bd. 18, 1930, pp. 145 – 153.

73. Über die Weinbergische Geschwistermethode. *Assekuranz Jahrbuch*. Wien – Leipzig, 1931, pp. 40 – 52.

74. Über einige Abschätzungen von Erwartungswerten. *J. reine angew. Math.*, Bd. 165, 1931, pp. 184 – 193.

75. *Vorlesungen aus dem Gebiete der angew. Mathematik*, Bd. 1. Leipzig – Wien 1931.

77. Altersschichtung und Bevölkerungszahl in Deutschland. *Naturwissenschaften*, Bd. 20, 1932, pp. 59 – 62.

78. Théorie des probabilités. Fondements et applications. *Ann. Inst. H. Poincaré*, t. 3, 1932, pp. 137 – 190.

79. Fragen der Wahrscheinlichkeitsrechnung. *Verh. Intern math. Kongr.*, Bd. 2, 1932. Zürich, pp. 221 – 228.

80. Über die Vorausberechnung von Umfang und Altersschichtung der Bevölkerung Deutschlands. *Blätter Versicherungsmath. (Beibl. Z. ges. Versicherungswiss.)*, Bd. 2, 1933, pp. 359 – 371.)

81. Über Zahlenfolgen, die en Kollektiv-ähnliches Verhalten zeigen. *Math. Ann.*, Bd. 108, 1933, pp. 757 – 772.

84. Generalizzazione di un teorema sulla probabilita della summa di un numero illimitato di variabili casuali. *Giorn. Ist. Ital. Attuari*, t. 5, 1934, 15 pp. French translation: *Actes Congr. Interbalcanique mathématiciens Athens 1934*. Athens, 1935.

86. Problème de deux races. *Rec. math.*, t. 41, 1934, pp. 359 – 389.

87. Deux nouveaux théorèmes de limite dans le calcul des probabilités. *Rev. Fac. Sci. Univ. Istanbul*, t. 1, 1935, pp. 61 – 80.

89. Die Gesetze der großen Zahl für statistische Funktionen. *Monatshefte f. Math. u. Phys.*, Bd. 43, 1936, pp. 105 – 128.

92. Sull concetto de probabilità fondato sul limite di frequenze relative. *Giorn. Ist. Ital. Attuari*, t. 7, 1936, pp. 235 – 255.
93. Les lois de probabilité pour les fonctions statistiques. *Ann. Inst. H. Poincaré*, t. 6, 1936, pp. 185 – 212.
94. La distribution de la plus grande de n valeurs. *Rev. Math. Union Interbalcan*, t. 1, 1936, pp. 141 – 160.
96. Bestimmung einer Verteilung durch ihre ersten Momente. *Skand. Aktuarietids.*, t. 20, 1937, pp. 220 – 243.
97. Über Aufteilungs- und Besetzungswahrscheinlichkeiten. *Rev. Fac. Sci. Univ. Istanbul*, t. 4, 1938/1939, pp. 145 – 163.
98. Note on deduced probability distributions, *Bull. Amer. Math. Soc.*, vol. 44, 1938, pp. 81 – 83.
100. Sur une inégalité pour les moments d'une distribution quasi-convexe. *Bull. Sci. Math.*, t. 42 (62?), 1938, pp. 68 – 71.
102. Généralisation des théorèmes de limite classiques. Colloque consacre à la théorie de probabilités. *Genève. Actual. Scient.*, t. 736, 1938, pp. 61 – 68.
104. Quelques remarques sur les fondements du calcul des probabilités. *Ibidem*, t. 735, 1938, pp. 57 – 66.
105. Sur les fonctions statistiques. Soc. math. de France. *Conf. de la Reunion Intern. de Math.* Paris, 1937, pp. 1 – 8.
109. Modification of the Bayes theorem. AMS, vol. 9, 1938, pp. 256 – 259.
111. The limits of a distribution function if two expected values are given. AMS, vol. 10, 1939, pp. 99 – 104.
112. An inequality for the moments of a discontinuous distribution. *Skand. Aktuarietids.*, t. 22, 1939, pp. 32 – 36.
116. On the probability of a set of games and the foundations of probability theory. *Revista de Ciencias*, t. 47, No. 453, 1946, pp. 435 – 456.
117. On the foundations of probability and statistics. AMS, vol. 12, 1941, pp. 191 – 205.
119. On the correct use of the Bayes formula. AMS, vol. 13, 1942, pp. 156 – 165.
121. On the problem of testing hypotheses. AMS, vol. 14, 1943, pp. 238 – 252.
124. On the classification of observation data into distinct groups. AMS, vol. 16, 1945, pp. 68 – 73.
126. On the asymptotic distribution of differentiable statistical functions. AMS, vol. 18, 1947, pp. 309 – 348.
127. *Mathematical theory of probability and statistics*. Harvard Univ., Graduate school of engineering. Sp. publ. No. 1, 1946, mimeogr., 320 pp.
138. Sur les fondements du calcul des probabilités. In *Collection de Logique Mathématique*, sér. B, I, pp. 16 – 29. Paris, 1952
139. Théorie et application des fonctions statistiques. *Rendiconti di matematica e delle sue applicazioni*, ser. 5, t. 11, 1952, 37 pp.
141. Über die J. von Neumannsche Theorie der Spiele. *Math. Nachr.*, Bd. 9, 1953, pp. 363 – 378.

This list shows at once that Mises' writings are to a large extent forgotten. Cramér's discussion of his frequentist theory is insufficient, see Khinchin's study of 1961 in my translation (*Science in context*, vol. 17, 2004, pp. 391 – 422). Caution: the Editor of my translation is responsible for some disappointing points; my *Machian* became *Machist*!

VI
R. Frisch, F. C. Mills, Ch. F. Roos

Constitution of the Econometric Society

Econometrica, vol. 1, 1933, pp. 106 – 108

1. Scope of the Society

The Econometric Society is an international society for the advancement of economic science in its relation to statistics and mathematics. The Society shall operate as a completely disinterested, scientific organization without political, social, financial, or nationalistic bias. Its main object shall be to promote studies that aim at a unification of the theoretical-quantitative and the empirical – quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking similar to that which has come to dominate in the natural sciences. Any activity which promises ultimately to further such unification of theoretical and factual studies in economics shall be within the sphere of interest of the Society,

2. Members

The Society shall consist of two classes of members: regular members and Fellows. Members of either class shall be subject to election. To become a regular member, a person must be proposed to the Council by two members of the Society. Once a year the Council shall nominate new members and these nominations shall be voted upon by mail by all members. No person can be elected to membership unless he is nominated by the Council.

During the first year following the organization meeting of the Society, the Council shall have the authority to invite eligible persons to become charter members.

3. Fellows

All Fellows of the Society shall be nominated by the Council and elected by mail-vote of the Fellows. Such nominations may be made at any time. To be eligible for such nomination a person must have published original contributions in economic theory or to such statistical, mathematical or accounting analyses as have a definite bearing on problems in economic theory, and must have been a member of the Society for at least one year. Each year the Council shall offer the members an opportunity to suggest nominees for fellowships.

The first group of Fellows shall be elected by the Council.

A meeting of the Fellows shall be held each year at a place and time designated at the previous annual meeting of the Fellows. In default of action by the Fellows, the Council shall have the power to fix the place and date of such meeting. The annual meeting of the Fellows shall represent the highest authority of the Society. In important cases, the annual meeting of the Fellows may decide by a simple majority vote to subject any decision of the meeting to ratification by mail-vote of all Fellows.

4. Officers

The President of the Society shall be elected for one year by ballot by the Council. The President shall act as chairman of the Council. The first President shall be elected by those present at the organization meeting of the Society.

The Council may elect one or more of its members to act as Vice-Presidents of the Society for a period of one year.

The Council shall elect a Secretary and a Treasurer to serve for a period of one year. These two offices may be combined. The Treasurer shall present financial statements to the Council, which, in turn, shall submit the statements to the annual meeting of the Fellows.

5. The Council

The Council shall consist of no less than fifteen members. The President, the Secretary, and the Treasurer shall be ex-officio members of the Council. The other members of the Council shall be elected with a view to represent the various geographic areas in which the Society has members. These latter members of the Council shall be elected for three years by the Fellows upon nomination by the Council, the offices for approximately 1/3 of them terminating each year. The first Council shall be elected by those present at the organization meeting of the Society.

6. Activities

Any activities which fall within the sphere of interest of the Society may be authorized by the Council, such as local scientific meetings, international meetings under the auspices of the Council, and the issuance of publications reporting the activities of the Society or containing other matters of economic interest.

7. Financial organization

The dues for regular members and the dues for Fellows shall be fixed by the Council. Such dues shall be moderate in amount. Bequests and gifts may be received.

8. Amendment of the Constitution

Amendments to this Constitution must be approved by a simple majority vote of those Fellows present at any annual meeting of the Society or by a two-thirds majority of those voting in a mail-vote taken among all Fellows.

This Constitution was adopted by general vote of those present at the organization meeting of the Society in Cleveland, Ohio, U.S.A. on Dec. 30, 1930, and there decided upon the final phrasing.

On p. 445 a list of Fellows is appended. It is 30-strong and signed by Irving Fisher, Chairman of Council

VI

R. Frisch, F. C. Mills, Ch. F. Roos

Constitution of the Econometric Society

Econometrica, vol. 1, 1933, pp. 106 – 108

1. Scope of the Society

The Econometric Society is an international society for the advancement of economic science in its relation to statistics and mathematics. The Society shall operate as a completely disinterested, scientific organization without political, social, financial, or nationalistic bias. Its main object shall be to promote studies that aim at a unification of the theoretical-quantitative and the empirical – quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking similar to that which has come to dominate in the natural sciences. Any activity which promises ultimately to further such unification of theoretical and factual studies in economics shall be within the sphere of interest of the Society,

2. Members

The Society shall consist of two classes of members: regular members and Fellows. Members of either class shall be subject to election. To become a regular member, a person must be proposed to the Council by two members of the Society. Once a year the Council shall nominate new members and these nominations shall be voted upon by mail by all members. No person can be elected to membership unless he is nominated by the Council.

During the first year following the organization meeting of the Society, the Council shall have the authority to invite eligible persons to become charter members.

3. Fellows

All Fellows of the Society shall be nominated by the Council and elected by mail-vote of the Fellows. Such nominations may be made at any time. To be eligible for such nomination a person must have published original contributions in economic theory or to such statistical, mathematical or accounting analyses as have a definite bearing on problems in economic theory, and must have been a member of the Society for at least one year. Each year the Council shall offer the members an opportunity to suggest nominees for fellowships.

The first group of Fellows shall be elected by the Council.

A meeting of the Fellows shall be held each year at a place and time designated at the previous annual meeting of the Fellows. In default of action by the Fellows, the Council shall have the power to fix the place and date of such meeting. The annual meeting of the Fellows shall represent the highest authority of the Society. In important cases, the annual meeting of the Fellows may decide by a simple majority

vote to subject any decision of the meeting to ratification by mail-vote of all Fellows.

4. Officers

The President of the Society shall be elected for one year by ballot by the Council. The President shall act as chairman of the Council. The first President shall be elected by those present at the organization meeting of the Society.

The Council may elect one or more of its members to act as Vice-Presidents of the Society for a period of one year.

The Council shall elect a Secretary and a Treasurer to serve for a period of one year. These two offices may be combined. The Treasurer shall present financial statements to the Council, which, in turn, shall submit the statements to the annual meeting of the Fellows.

5. The Council

The Council shall consist of no less than fifteen members. The President, the Secretary, and the Treasurer shall be ex-officio members of the Council. The other members of the Council shall be elected with a view to represent the various geographic areas in which the Society has members. These latter members of the Council shall be elected for three years by the Fellows upon nomination by the Council, the offices for approximately 1/3 of them terminating each year. The first Council shall be elected by those present at the organization meeting of the Society.

6. Activities

Any activities which fall within the sphere of interest of the Society may be authorized by the Council, such as local scientific meetings, international meetings under the auspices of the Council, and the issuance of publications reporting the activities of the Society or containing other matters of economic interest.

7. Financial organization

The dues for regular members and the dues for Fellows shall be fixed by the Council. Such dues shall be moderate in amount. Bequests and gifts may be received.

8. Amendment of the Constitution

Amendments to this Constitution must be approved by a simple majority vote of those Fellows present at any annual meeting of the Society, or by a two-thirds majority of those voting in a mail-vote taken among all Fellows.

This Constitution was adopted by general vote of those present at the organization meeting of the Society in Cleveland, Ohio, U.S.A. on Dec. 30, 1930, and there decided upon the final phrasing.

On p. 445 a list of Fellows is appended. It is 30-strong and signed by Irving Fisher, Chairman of Council

VII
R. Frisch
Editorial (excerpt)

Econometrica, vol. 1, 1933, pp. 1 – 2

Two years have now elapsed since the founding of the Econometric Society. Although the Society has purposely given little publicity to its affairs during these years of organization, inquiries and suggestions have come from many quarters manifesting a readiness for, and a keen expectation of something along the lines now followed by the Econometric Society. A source of potential energy much larger than originally anticipated by the founders of the Society seems to exist, only waiting to be released and directed into econometric work.

This is the reason why the Society has decided to establish its own journal, *Econometrica* will be its name. It will appear quarterly and this is its first issue. A word of explanation regarding the term *econometrics* may be in order. Its definition is implied in the statement of the scope of the Society, in Section I of the Constitution, which reads [see item VI].

This emphasis on the quantitative aspect of economic problems has a profound significance. Economic life is a complex network of relations operating in all directions. Therefore, so long as we confine ourselves to statements in general terms about one economic factor having an *effect* on some other factor, almost any sort of relations may be selected, postulated as a law, and *explained* by a plausible argument.

Thus, there exists a real danger of advancing statements and conclusions which, although true as tendencies in a very restricted sense, is nevertheless thoroughly inadequate or even misleading if offered as an explanation of the situation. To use an extreme illustration, they may be just as deceptive as to say that a man who tries to row a boat will be driven backward because of the pressure exerted by his feet. The rowboat situation is not, of course, explained by finding out *that* there exists a pressure to one direction or another, but only by comparing the relative *magnitudes* of a number of pressures and counter-pressures. It is this comparison of magnitudes that gives a real significance to the analysis. Many, if not most of the situations we have to face in economics are just of this source. The full usefulness of a large and important group of our economic analyses will come therefore only as we succeed in formulating the discussions in quantitative terms.

But there are several aspects of the quantitative approach to economics, and no single one of these aspects taken by itself should be confounded with econometrics. Thus, econometrics is by no means the same as economic statistics. Nor is it identical with what we call general economic theory, although a considerable portion of this theory has a definitely quantitative character. Nor should econometrics be taken as synonymous with the application of mathematics to economics. Experience has shown that each of these three viewpoints, those of statistics, economic theory and mathematics is a necessary but not by itself a sufficient condition for a real understanding of the quantitative relations in modern economic life. It

is the *unification* of all three that is powerful. And it is this unification that constitutes econometrics.

This unification is more necessary today than at any previous stage in economics. Statistical information is currently accumulating at an unprecedented rate. But no amount of statistical information, however complete and exact, can by itself explain economic phenomena. If we are not to get lost in the overwhelming, bewildering mass of statistical data that are now becoming available, we need the guidance and help of a powerful theoretical framework. Without this no significant interpretation and coordination of our observations will be possible.

The theoretical structure that shall help us out in this situation must however be more precise, more realistic, and in many respects more complex than any heretofore available. Theory, in formulating its abstract quantitative notions, must be inspired to a larger extent by the technique of observation. And fresh statistical and other factual studies must be the healthy element of disturbance that constantly threatens and disquiets the theorist and prevents him from coming to rest on some on some inherited, obsolete set of assumptions.

This mutual penetration of quantitative economic theory and statistical observation is the essence of econometrics. And therein lies the need for mathematics, both in the formulation of the principles of economic theory and in the technique of handling the statistical data. Mathematics is certainly not a magic procedure which in itself can solve the riddles of modern economic life, as is believed by some enthusiasts.

VIII J. Schumpeter

The common sense of econometrics

Econometrica, vol. 1, 1933, pp. 5 – 12

The aims of this Journal and of the Society, of which it is to be the organ, have been stated above by the Editor with that brevity and precision which are characteristic of every statement of a sound case. What I have to add by way of comment and simplification, will, I hope, confirm the impression that there is nothing startling or paradoxical about our venture, but that it grows naturally out of the present situation of our science. We do not wish to revive controversies about general questions of *method*, but simply to present and discuss the results of our work. We do not impose any credo, scientific or otherwise, and we *have* no common credo beyond holding: first, that economics is a science, and second, that this science has one very important quantitative aspect. We are no sect. Nor are we a *school*. For all possible differences of opinion on individual problems, which can at all exist among economists, do, and I hope always will, exist among us.

Like everything else, economic life may be looked at from a great, from a great, strictly speaking infinite number of standpoints. Only some of them belong to the realm of science, still fewer admit of, or require the use of quantitative methods. Many non-quantitative aspects are and always have been more interesting to most minds. Much of what we want to know about economic phenomena can be discovered and stated without any technical, let alone mathematical refinements upon ordinary modes of thought, and without elaborate treatment of statistical figures. Nothing is farther from our minds than any acrimonious belief in the exclusive excellence of mathematical methods, or any wish to belittle the work of historians, ethnologists, sociologists, etc. We do not want to fight anyone, or, beyond dilettantism, anything. We want to serve as best as we can.

1. Economics, the quantitative science

There is, however, one sense in which economics is the most quantitative, not only of *social* or *moral* sciences, but of *all* sciences, physics not excluded. For mass, velocity, current and the like *can* undoubtedly be measured, but to do so we must always invent a distinct process of measurement. This must be done before we can deal with these phenomena *numerically*. Some of the most fundamental economic facts, on the contrary, already present themselves to our observation as quantities made numerical by life itself. They carry meaning only by virtue of their numerical character. There would be movement even if we were unable to turn it into measurable quantity, but there cannot be prices independent of the numerical expression of every one of them, and of definite numerical relations among all of them.

Econometrics is nothing but the explicit recognition of this rather obvious fact, and the attempt to face the consequences of it. We might

even go so far as to say that by virtue of it every economist is an econometrician whether he wants to be or not, provided he deals with this sector of our science and not, for example, with the history of organization of enterprise, the cultural aspects of economic life, economic motive, the philosophy of private property, and so on. It is easy to understand why explicit recognition of this fact should have been so difficult, and why it has taken so long to come about. Philosophers who have at all times delighted in classifying sciences, have always felt uneasy about the precise place to be allocated to economics as a whole. As it was, they practically followed the empirical dividing line between *natural* and *moral* sciences, and classed economics with the latter. And there, of course, the quantitative aspect or sector of our science found but uncongenial ground.

Another reason was that economic problems have most of the time been approached in a practical spirit, either indifferent or hostile to the claims of scientific habits of thought. No science thrives, however, in the atmosphere of direct practical aim, and even practical results are but the by-products of disinterested work at the problem for the problem's sake. We should still be without most of the conveniences of modern life if physicists had been as eager for immediate *application* as most economists are and always have been. This accounts for the neglect of econometrics, as well as for the unsatisfactory state of our science in general. Nobody who craves for quick and short answers to burning questions of the day will care to entangle himself in difficulties which only patient labour can clear in the course of many years.

Nevertheless, the quantitative character of the subject was bound to assert itself. It is one of the most striking facts about the history of our science, that most (and if we exclude historians, *all*) of those men who we are justified in calling great economists, invariably display a remarkably mathematical turn of mind, even when they are entirely ignorant of anything beyond the quantitative technique at the command of a schoolboy: Quesnay, Ricardo, Böhm-Bawerk, are instances in point.

Nor this is all. If econometricians have any wish to imitate other people and to glory in heroic ancestry, they may with justice claim the great name of Sir William Petty as their own. The second part of the 17th century is full of vigorous ventures into the field of econometrics, one has but to point to Gregory King's statistical demand curve. It is a question of some interest how it was possible that such hopeful beginnings could have failed to inspire further work, and how their results could have left to linger in the dusk, although they were by no means forgotten, the reference to *King's rule* being part of the stock in trade of almost every stale textbook written ever since.

In the sphere of monetary phenomena and its neighbourhood, quantitative and even numerical analysis became established practice as far back as the 16th century, mainly in Italy, and this tradition was never lost again. Passages in the Italian writers of the 18th century, such as Beccaria, Carli, Verri and others, sounding very familiar to the modern ear. What we have before us there is nothing less than a

conscious attempt to weld into one indivisible argument theorems and statistical facts.

And, provided we leave out the word *conscious*, we find substantially the same tendency in whatever piece of the work of our predecessors we choose to look at. To give but one example. We are accustomed to scoff at the literature of the time-honoured controversy on value. But what else is at the bottom of it, overlaid, it is true, by heavy masses of speculative verbiage, if not the truly scientific search for an economic unit of measurement, or of several such units adapted to various classes of phenomena? There were no more un-empiric speculations about it than there is about every science in its infancy. Nor is there less connection with such statistical materials as each epoch could command than we are entitled to expect, as everyone will admit who has taken the trouble of perusing Ricardo's reply to Bosanquet.

2. Later developments

Essentially quantitative analysis, but crippled by the lack both of appropriate technique and of adequate statistical material – this is the diagnosis we arrive at when we study the work of economists up to that time, when Mill's principles were fairly representative of what our science had to give. This, too, is the element of truth which emerges from the hostile phraseology which we are in the habit of using about *classical doctrine*. Obviously, therefore, what our society stands for is anything but an innovation. It is no more than a conscious endeavour to remove obstructions to the flow of a stream which has been running ever since man began to think and write about economic life.

To see in their full significance the conditions which have made it desirable, and indeed necessary to form, under the banner of Econometrics a *coalition* of the different types of economists who are to join hands in our Society, we must, however, now glance at later developments. The phase that could, until about ten years ago, called the *modern* phase of economics, admits of description in terms of three facts and their consequences: first, the rapid growth of our wealth of statistical and other material; second, the progress of statistical technique at our command (which so far as it grew up largely outside our own field and without reference to our needs, was a stroke of good fortune, very much as a lift in another man's car is to the wanderer on a dusty road); and third, the emergence of a theoretical engine very much superior to the old one. True, on none of these heads are we, however be, satisfied; on every one of them, it seems to me, the real thing is still to come, and present performance calls for apology rather than congratulation. Yet it would not only be ungenerous but positively false to deny the importance of what has been achieved, or the possibilities which begin to loom in the future.

In all this the econometric line stands out clearly. It was definitively established that economic theory involves quantity, and therefore requires the only language or method available to deal with quantitative argument as soon as it outstrips its most primitive stage.

To W. St. Jevons belongs the honour of having spoken one of those simple messages, which at times seem to focus both past and future history and to become milestones visible ever after. It was he who said in the introduction to his *Theory* (1871):

It is clear that Economics, if it is to be a science at all, must be a mathematical one.

But still higher tribute [than to Jevons] is due to Cournot who, without encouragement or lead, in what was then a most uncongenial environment, in 1838 fully anticipated the econometric programme by his *Recherches*, one of the most striking achievements of true genius, to which we pay respect to this day by nearly always starting out from them.

Of course, it would be superfluous to emphasize the paramount importance of that great teacher of ours whose exposition of exact theory sprang from his head as did Minerva from the head of Jove. What I want to stress is that he constructed his analytic apparatus with a clear perception of the ultimate econometric goal, every part of it being thought out to it to grip statistical fact when the time should come. In this he went much farther than Jevons. This reads like a paradox because Jevons actually worked on figures, as in the matter of index numbers. But within the precincts of pure theory itself, he seems much less concerned with that goal than Cournot, and it is much more difficult for the numerical horse to jump Jevons' fences than it is to trot on Cournot's road.

In our pantheon, J. H. von Thünen's place is side by side with Cournot's. It is not only, indeed not even primarily, the idea of marginal productivity which it is important to mention her, but Thünen's peculiar relation to a set of facts, which is as vital to econometrics as statistics in the narrow sense of the word. Thünen pointed out that cost accounting, bookkeeping and neighbouring headings cover a mass of material which economists have entirely neglected. This neglect has indeed been so complete that the specialists of *Business Administration* now actually have begun to build their own theoretical houses which will wall them, in spite of the fact that both groups of workers to a great extent (a signal instance is the matter of cost curves) till the same ground. It is clear that economists cannot indefinitely do without that vast reservoir of fact, nor cost accountants, bookkeepers, and so on, do without the economists' contribution. And, looking back, we see now that as early as 1826 Thünen's book could have taught us *how theory grows out of observation of business practice*.

I for one shall always look up to Léon Walras as the greatest of all economists. In his theory of equilibrium he gave a powerful basis to all our work. It is true that while he made the decisive step in the *quantitative*, he failed to move in the *numerical* line, the *junction* of which two is characteristic of econometrics. But we have been taught of late to look more hopefully even on the *numerical* possibilities of that most general and most abstract art of our science which is equilibrium theory in the Walras sense. And this fact similarly indicates the econometric claims of the work of Auspitz and Lieben,

of Knut Wicksell, of Francis Y. Edgeworth, and of Walras' great successor in Lausanne, Vilfredo Pareto.

In a somewhat different sense, we may finally claim as our own that greatest of all teachers of economics, Alfred Marshall. With some of us, it has become a custom to speak of him as the exponent of *neoclassical* doctrine. This is not the place to show how it came to pass (not without some fault of Marshall himself) that so utterly unjust, and in fact meaningless a label was affixed to his name. But I wish to emphasize first [of all], that nobody can peruse his address on *The old generation of economists and the new* without discovering, though not perhaps without some surprise, how clearly our programme stood before his mind. Nor is it possible for anyone who knows how to read his *Principles* in the light of his *Industry and trade* to define what he really strove to accomplish in any but econometric terms. Most important of all, he always worked with an eye to statistical application, and he was at his best as a *theorist* when constructing those handy tools, like elasticity, quasi-rent, external and internal economics and so on, which are so many bridges between the island of pure theory and the terra firma of business practice and business statistics.

I do not wish to speak of any living economists. But the readers would probably not forgive me if I failed to make two exceptions and to mention the pioneer work of Irving Fisher and Henry L. Moore.

3. The present state

All these achievements were, so to say the least, enough as a good start and to build from. And indeed, work full of promise has been done in our line during the last two decades, work which makes us feel, when we now look at the Walrasian system, very much as we feel when beholding the model of a motor car constructed 40 years ago. But still, most of us undoubtedly do agree in finding the present state of our science disappointing, not only in comparison with the achievements of other sciences, but also in comparison with what our science could fairly be expected to perform. There are many reasons for this, but some of them only, having special bearing on the mission of this Society, call for attention here.

Reasoning on economic facts means, and always meant, within a very important sector, quantitative reasoning. And there is no *logical* breach between quantitative reasoning of an elementary character, and quantitative reasoning of the kind involving the use of *higher mathematics*. But nothing makes a greater *practical* breach in the evolution of a science than the introduction of a habit of thought which has so far been foreign to the recognized equipment of the specialist, and which at the same time is inaccessible except by strenuous effort. When the necessity of proceeding to the use of more refined mathematical methods, both in economic theory and in statistics, became apparent to some, the majority even of those economists who did work the quantitative sector refused to follow. At first they laughed. They do so no longer. Integrals cease by and by to be as hieroglyphs to them. Many of them try to understand and have

made their peace with us, while reserving their right to criticise our results and to object to mathematical excesses. But this is not the full cooperation we need. Even in this improved situation, economics lacks that broad expanse of professional common ground which, in the case of physics, transmits acquired results to the general public. Beginners are bewildered by this unsettled situation. Energy is being wasted and the real business of the science hampered. Recent progress, and still more than *actual* progress wide *possibilities* of it, has drawn to our field a most promising host of newcomers. But the old situation being fundamentally changed, we had no uniform training to offer them. Hence the lack of coordination of the work. The new men came to face our problems from very different angles and with very different acquirements, full of impatience to clear the ground and to build entirely anew. The man whom nature had moulded to delight in unadulterated fact, whether he worked in a statistical bureau or did field work, often knew little and cared less for that engine of analysis, which we call *economic theory*, or for refined statistical technique. On the other hand, the master of this technique, feeling its power and seeing the material to grip with it, tried to rush at his own kind of regularities or generalizations. And the theorist, conscious of his own task, refused more often than was wise to accept the work of the other two types as anything but (possible) verifications of his theorems. But, although uncoordinated, the growth has been tropical. It might be expected to settle down and bear fruit in time, but there is chaos for the present, in which only a very experienced eye can see an underlying tendency working its way slowly though powerfully towards a good common to all.

4. The programme

The common sense of the programme of our Society centres in the question: Can we not do better than this? Surely it would not be a reasonable policy to sit down and wait till, in the end, things find their level by themselves, and meanwhile to allow econometricians of all countries to fight single-handed their uphill battle. What we want to create is, first, a forum for econometric endeavour of all kinds wide enough to be hampered by the weight of an audience which keeps discussion in the ante-rooms of the real points at issue, and forces every speaker or writer to go every time over the same preliminaries.

On this forum, which we think of as international, we want, second, to create a spirit and a habit of cooperation among men of different types of mind by means of discussions of concrete problems of a quantitative and, as far as may be, numerical character. The individual problems themselves are, as it were, to teach us how they want to be handled. We want to learn how to help each other, and to understand *why*, and precisely *where*, we ourselves, theorists, statisticians, collectors of facts, or our neighbours, do somehow not quite get to where we want to be. No general discussion on principles of scientific method can teach us that. We have had enough of it. We know it leads nowhere, and only leaves the parties to the contest where they were before, still more exasperated perhaps by that gentle rudeness it is

customary to administer to each other on such occasions. No general arguments of this kind ever carry conviction to the man who means real work. But, confronted with clear-cut questions, most of us will, we hope, be found to be ready to accept the only competent judgement on, and the only relevant criterion of scientific method, that is the judgement or criterion of the result. There is high remedial virtue in quantitative argument and exact proof. That part of our differences, no matter whether great or small, which is due to mutual misunderstanding, will vanish automatically as soon as we show each other, in detail and in practice, how our tools work and where they need to be improved. And metaphysical acerbity and sweeping verdicts will vanish with it. Theoretic and *factual* research will of themselves find their right proportions, and we may not unreasonably expect to agree in the end on the right kind of theory and the right kind of fact and the methods of treating them, not postulating anything about them by programme, but evolving them, let us hope, by positive achievement.

We should not indulge in high hopes of producing rapidly results of immediate use to economic policy of business practice. Our aims are first and last scientific. We do not stress the numerical aspect just because we think that it leads right up to the core of the burning questions of the day, but rather because we expect, from constant endeavour to cope with the difficulties of numerical work a wholesome discipline, the suggestion of new points of view, and helps in building up the economic theory of the future. But we believe of course that indirectly the quantitative approach will be of great practical consequence. The only way to a position in which our science might give positive advice on a large scale to politicians and businessmen, leads through quantitative work. For as long as we are unable to put our arguments into figures, the voice of our science, although occasionally it may help to dispel gross errors, will never be heard by practical men. They are, by instinct, econometricians all of them, in their distrust of anything not amenable to exact proof.

Notes

1. But what is the difference between numerical and quantitative?
2. Marginal productivity is a generalization of the Ricardo theory of rent a particular case of the general theory of determining value.
3. Quasi-rent is the difference between the sale price and the production cost. External economy is connected with attempts to reconcile competition with decrease of production cost. Internal economy is connected with the decrease of production cost and increase of the volume of production.

The figures mentioned

Auspitz R. (1837 – 1906), published two books on the theory of pricing, 1887 and 1889, co-author R. Lieben
 Beccaria C. (1738 – 1794), lawyer, philosopher
 Bosanquet B. (1848 – 1923), philosopher
 Carli G. R. (1720 – 1795), astronomer and economist
 King G. (1650 – 1712), statistician
 Moore H. L. (1869 – 1958), economist
 Quesnay Fr. (1694 – 1774), founder of the physiocracy school
 Verri P. (1728 – 1797), philosopher, economist, historian

Wicksell K. (1851 – 1926), economist

Bibliography

Cournot A. A. (1838), *Recherches sur les principes mathématiques de la théorie des richesses*. Paris, 1890.

Jevons W. (1871), *Theory of political economy*. London – New York.

King C. (1801), *Natural and political observations upon the state and condition of England in 1696*. London.

Marshall A. (1890), *Principles of economics*. London.

--- (1897), *The old generation of economists and the new*. Boston.

--- (1919), *Industry and trade*. London.

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

Thünen J. H. (1826), *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*. Hamburg.

Schumpeter did not even hint at the possibility of some anticipation of econometrics in the writings of Russian or Japanese authors.

In 1910 – 1911 Bortkiewicz published a treatise on Marxist econometrics. His presentation, as it occurred in all of his contributions, was dry, and owing to his fact his attempt was ignored.

In the Soviet Union, econometrics began to break through only in 1960, and only in a concealed way. In 1959, leading Soviet statisticians refused to recognize it (Sheynin 1998, p. 542). Naturally, they had not even thought about the experience gained abroad.

VIII

Thor Andersson

Statistics or chaos

Nordic Stat. J., vol. 1, 1929, pp. 3 – 32

The moral of the tale is the power of reason, its decisive influence on the life of humanity. The great conquerors from Alexander to Caesar, and from Caesar to Napoleon, influenced profoundly the lives of subsequent generations. But the total effect of this influence shrinks to insignificance if compared to the entire transformation of human mentality produced by the long line of men of thought from Thales to the present day, men individually powerless, but ultimately the rulers of the world. A. N. Whitehead

Laws of nature have turned out to be in some cases human conventions, in others mere statistical averages. Bertrand Russell

[1] The Assyrians, Babylonians and Egyptians had certainly made great progress in the use of mechanical devices for moving great loads, in the construction of scales, and of pumps. Their measuring instruments were well developed, and acute observations were made, but of systematic, scientific investigation there is no evidence. The Greeks received many results and suggestions from Asia Minor, Mesopotamia, and Egypt but their achievements are essentially their own.

Ionia in Asia Minor was the cradle of free speculation. The history of European science and European philosophy begins in Ionia. Here (in the 6th and 5th centuries BC) the early philosophers by using their reason sought to penetrate into the origin and structure of the world. They could not of course free their minds entirely from received notions, but they began the work of destroying orthodox views and religious faiths. Xenophanes may specially be named among those pioneers of thought (though he was not the most important or the ablest), because the toleration of his teaching illustrates the freedom of the atmosphere in which those men lived. He went about from city to city, calling in question on moral grounds the popular beliefs about the gods and goddesses, and ridiculing the anthropomorphic conceptions which the Greeks had formed of their divinities:

If oxen had hands and the capacities of men, they would make gods in the shape of oxen.

This attack on received theology was an attack on the veracity of the old poets, especially Homer, who was considered the highest authority on mythology. Xenophanes criticized him severely for ascribing to the gods acts which, committed by men, would be considered highly disgraceful. We do not hear that any attempt was made to restrain him from thus assailing traditional beliefs and branding Homer as immoral. We must remember that the Homeric poems were never supposed to be the word of God. It has been said

that Homer was the Bible of the Greeks. This remark exactly misses the truth. The Greeks fortunately had no Bible, and this fact was both an expression and an important condition of their freedom. Homer's poems were secular, not religious, and it may be noted that they are freer from immorality and savagery than sacred books that one could mention. Their authority was immense, but it was not binding like the authority of a sacred book. And Homeric criticism was never hampered like Biblical criticism.

In this connection, as Bury says, notice may be taken of another expression and condition of freedom, the absence of sacerdotism [of priests endowed with special power]. The priests of the temples never became powerful castes, tyrannizing over the community in their own interests and able to silence voices raised against religious beliefs. The civil authorities kept the general control of public worship in their own hands, and, if some priestly families might have considerable influence, yet as a rule the priests were virtually state servants whose voice carried no weight except concerning the technical details of ritual.

These circumstances have been of fundamental importance to the science of the Greeks. In science they had to build from the ground. Other peoples had extensive knowledge and highly developed arts. But only among the Greeks there existed the true scientific method characterized by free investigation, rational interpretation, verification or rectification by systematic and repeated observation, and controlled deduction from accepted principles.

[2] As long as the circumstances, under which the scientific activity of the Greeks made its luminous start, continued, as long continued this activity too. Certainly it is possible that scientific progress could continue from one era to another through the genius of special teachers and students without regard to political and social circumstances. Such is not however the witness of history. For progress in science the men of genius are indispensable, but in no country and at no time they alone have been capable to make science flourish under circumstances so unfavourable as those prevailing during the first centuries after the acceptance of Christianity as the state religion of the Roman world empire.

Certainly, during the days of its first greatness, the Roman empire had a general rule to tolerate, throughout the empire, all religions and all opinions. Blasphemy was not punished. This principle was expressed in the maxim of the emperor Tiberius:

If the gods are insulted, let them see to it themselves.

An exception to the rule of tolerance was made in the case of the Christian sect, and the treatment of the oriental religion may be said to have inaugurated religion persecution in Europe. The exclusiveness and intolerance of the Jews could not be endured by the tolerant pagans. The fanaticism of the Jews and of the Christians became the main cause of the persecution. The Christians having become themselves the foremost in the state, they have, in their turn, made use of every means in the power of religion and state to persecute their antagonists with the most brutal recklessness and the most exquisite instruments of torture. Thus ceased the conditions under which the

Greeks had been able, even during the Roman empire, to continue, with certain success, their promising scientific activity. The acceptance of Christianity by Constantinus the Great inaugurated a millennium in which, as Bury says, reason was enchained, thought was enslaved, and knowledge made no progress.

The famous Danish describer of mathematics and natural sciences in the classic antiquity, Heiberg, declares that

The Greek professional science, to which all European professional science ultimately goes back, originates from the Ionians, who have personified their yearnings for knowledge in the national hero Odysseus, and of whose realistic mind and sharp observation the Homeric poems bear witness. Their contact with the old civilisation in Egypt and Mesopotamia can, at most, have given them fresh suggestions and new material. To make science of it was not in the power of the orient, captured, as it was, in religion. This was only in the power of the spiritually free Greece.

[3] If also those parts of mathematics, which are the foremost toils of the statistics, have been noticed by the Greeks, is not yet known. Perhaps the still undiscovered sources of Diophantos' works and these works themselves, if complete, would possibly give some allusion to the fact that the principal problems of the combinatorics did not lie wholly outside mathematics and other sciences of the Greeks. Do not such spiritually great men as Heraclitus and Democritus meet in the beginning of the Greek science? Is it not probable that the sentence that the world and everything in it are changing every instant or the atomic theory of the universe allude to a possibility that also the means for the demonstration of the extensiveness and validity of these theories have been, already in early phases of the Greek science, subjected to its investigations and examinations?

Already the Greek science can thus have begun the building up of the combinatorial analysis, which is the very foundation of investigations in the theory of probability and accordingly also in the theory of statistics. The work having been performed later from statistical observations, within the sphere of combinatorics has not yet been suitable to make clear the importance of this science to statistics, Combinatorics have in fact not gained the attention, neither from statisticians, nor from mathematicians. They will possibly get in the continued building up of the statistical science.

Not even that part of the statistics which is of such an importance that the declaration of independence of the statistical science by Lexis is founded on it¹, or the theory of probability has yet developed so highly that it can, in all respects, fill its place as the excellent auxiliary science of the statistics. The relation between mathematics and statistics is that between an auxiliary science and an independent science. The in no way rare idea that statistics would be nothing else than applied mathematics, is wrong. When statistical phenomena are to be estimated or when conclusions are to be drawn from statistical observations, it is not sufficient to apply purely mathematical theorems, for it is undisputable that such theorems had a disadvantageous effect in many cases, as in the English biometrics. They can even lead to mathematical errors². Statisticians must have a

clear notion also of the individual state of the material used in an investigation, a notion which is overlooked almost too often by the mathematicians. Therefore, we may even witness now and then that mathematicians of good reputation attack a statistical problem in a wholly incorrect way. The simplest solution of mathematical problems is not always reached in the way proposed by professional mathematicians.

The fact that even many eminent mathematicians have not been successful in their investigations trips within the statistical boundaries is chiefly due to their tendency to confine themselves exclusively to mathematics. Czuber³ has pointed out that mathematics cannot be applied in a reasonable way in a certain field until one has gained a complete knowledge of this field, but mathematicians appear to be more or less unfamiliar with this statement. The essential thing in the statistical work has been unknown to them, and the mathematical method called to their assistance has been of very little use, however fine it may have been, not to say that it had areally prejudicial effect, a consequence which has also been observed.

The statistical exaggerations and extravagances incurred by mathematicians lacking the necessary statistical insight have contributed, in turn, to the state of opposition which may be noticed in many cases between mathematics and statistics, even if it does not exist between mathematicians and statisticians for the reason that they are scientists in the modern sense of the word

[4] The statistical scientists are well aware of the fact that mathematics is indispensable, and that it has already done a great deal and will possibly do still more for the work of statistical science and its future development. But there are a countless number of so-called statisticians all over the world, who do not have the right to bear the title, if it is taken in the modern scientific sense. At present, in many states and especially in international scientific institutions this crowd of so-called statisticians drowns the voices of real statisticians. This is for instance the case in Wargentin's and Lexis' native countries as well as in the International Statistical Institute and in the parodic statistical workshop at Geneva. The crowd of the so-called statisticians is constituted at present in an almost too high proportion by persons having no scientific training or knowledge in the subjects which were indispensable even for statisticians of yesterday, that is, in the first place, in scientific statistics and physics and in essential parts of, above all, mathematics and geography. They are strangers to the sciences which is the essential thing in the human development, the science without which this development will have no future. Even those who have gained juridical learning can scarcely pretend to have an inductive scientific training, a fact pointed out already in 1860 by such an eminent expert on the relations between statistics and the law as Brougham.

One of the most eminent persons in the history of statistics is the noble Robert Meyer, who established the Austrian financial statistics in a paragon manner, and afterwards was appointed minister of finance and finally president of the Austrian central statistical committee. In the latter position the lack of sufficient mathematical

learning embarrassed him in such a degree that he, without any hesitation, took up the study of those parts of mathematics with which he was not conversant. This decision was fulfilled in such a way that this master of general statistics, as he calls statistics with no special application of mathematics, could make a striking statement in an article *A word on mathematical and general statistics* in *Festschrift* (1914) published in the author's year of death, a statement which must be regarded as the last word in the fight which is still carried on by those who are not able or willing to penetrate into the present science in general or into the present statistical science in particular. Meyer sums up his opinion in the following word which hit the point exactly:

The re-established connection between, on the one hand, mathematical statistics, applied for a long time in insurance work and in biological measurements only, and on the other hand statistical problems of general character must be saluted as an extremely valuable enrichment of statistical methods. It maintains this value irrespective of the possible limitations and definitions of the tasks of statistics. In fact, it seems to me that its whole bearing does not become apparent until we give the widest sense to statistics and this for the reason that even the descriptive representation of concrete circumstances can be made deeper and more complete under certain conditions by taking assistance from mathematical methods. This may be applied to the representation of concrete salary conditions the estimation of the degree of labour employment.

There are many different opinions, even among friends of mathematics as to the practical advantages a widened use of mathematics can get for statistics. There are all degrees from enthusiastic adherents to cold doubters. In every case it must be kept in sight that the inner connections I have alluded to with apparently rather remote questions about the technics of collection and adaptation offer scarcely surveyable possibilities of development.

From what has been said it is my opinion evident that we must give up the present sharp distinction of mathematical statistics from the general, and that we must accept the mathematical doctrines as an integral part of the statistical theory. In other countries this development is already accomplished by leading geniuses. We exclude ourselves from the general progress if we leave these facts out of sight.

With regard to contents and task the so-called mathematical statistics (?) does not stay in any opposite relation to statistics in its entirety or to general statistics. From a practical point of view the attempts to refer the so-called mathematical statistics to a special discipline cannot therefore be defended. Such attempts are injurious to the further development of statistics. At present they are also chiefly made by so-called statisticians who lack necessary scientific pre-education at a time when the statistical science built on the ground of the theory of probability is foremost. These so-called statisticians also manifest their likewise lacking knowledge of the toil of mathematicians when they propose a difference between general statistics and something they want to call *statistical mathematics*^{5a}. The last name is a sufficient proof of the confusion of ideas by the proposers and by their

supporters and of their lacking understanding of the scientific character of statistical work from the ground to the top, during whose building, and only there, according to their opinion high mathematics can be effectively used.

From the very design statistical work is scientific. It is in no way so that

The scientific work begins first when the material has left the workshops of official statistics,

as the Swedish Geddes committee has said. If, from the very beginning, the material has not been treated from scientific points of view and by scientific methods, there cannot of course exist the same bearing foundation for further scientific work as when science has been from the start the leader and examiner of the collection and treatment of the material. Above all, science must be decisive from the very start. At such representative investigations the scientific claims on planners and performers are very great. The representative methods here coming into use offer from a scientific point of view the greatest difficulties to practical statistical work.

Only a statistician who has gained a thorough scientific training and knows also the finest auxiliaries of mathematics is able to conquer these difficulties. Where these, as is still rather often the case, are lacking, his knowledge of mathematics enables him to procure the auxiliary means and methods that a statistician cannot dispense with for the performance of his work and art. The terms mathematical statistics or statistical mathematics are rubbish, their maintenance may imply the assent that there is a type of statistics which could pretend to be science although it is not built on the ground of the theory of probability. It is the everlasting merit of the great Lexis that he, when publishing well 50 years ago, the independence act of statistical science, laid such a firm foundation that the scientific development, at least that of the nearest future, will not be likely to derange it essentially. Therefore, when statistics are (?) introduced as a special discipline in the higher educational institutions, above all in the highest, the epithet mathematical cannot and should not be inserted.

[5] Where there is no central authority over official statistics, departmental independence inevitably leads to duplication, overlapping and incongruity,

is the description of official statistics 50 years ago in the country which, all since the days of John Graunt, has given the science of statistics so many valuable contributions to its present property. This description holds for official statistics in almost every state of the world, at least in certain periods. The slow and gradual development of the statistical organisation has contributed in a very high degree to the fact that none of the older states in the Old World has reached the centralisation of the entire governmental statistical activity into a scientific institution which may serve as the brain of the whole administration. The work for the centralisation of official statistics is carried on in many states and has been successful in many places where eminent politicians and statisticians have been the foremost promoters of centralisation.

The slow progress of centralisation is also depending on the tardy development of statistical science and on its often infinitesimal interest in practical statistics. Not all the representatives of scientific statistics yet realise that practical statistics consists in scientific work for practical and theoretical purposes. Nor do all leaders of practical statistical work embrace the opinion expressed by the former director of the Norwegian statistical central bureau, Rygg:

The work of a statistical leader is necessarily a miscellany of administrative and scientific occupations. The richer his equipment in the last-mentioned respect the better the work he is able to offer. His projects may not only concern the momentary need, but he ought to partake in the work for the future, and thus he might collect material which is instantly of no great interest.

It would be better if problems demanding great attention and carefulness were treated in a central office, declared the eminent Australian statistician Hayter 50 years ago in the statistical committee of Great Britain to which reference has been made previously. It is necessary to have a big central institution where the best care could be given to the steadily increasing number of statistical branches of which the modern state has an incessantly growing need. This care must extend over the whole field of statistics, and the aims must be to procure the theoretically and practically best results. Excellent statisticians are needed. The type of man needed for higher statistical work is the scientific type, Cohlán says. Hayter declared to the committee that if the statistician has his heart in his work, he will procure the best information he can.

[6] The great importance which comparison has for good statistics, whose value so greatly depends on the fact that statistics of one district can be fully compared with the corresponding statistics of another district, not only in the country in question, but with that of all countries, and not only for the present, but for all times, has already for more than 3/4 of a century caused efforts to effect an international comparative statistics. The efforts then resulted in the organisation of the International Statistical Congress which had its first meeting in 1853 and its ninth and last in 1876 [in Petersburg].

The most important merits of the Congress were the effecting of a connection between official and scientific statistics, from which may be derived the germs of the development which has lately taken place in the history of statistics in general, and the effecting of improvements in statistical offices in the majority of countries under the influence of western culture. So long as the Congress refrained from a direct influence on the development of the administrative statistics of the different countries, it met with success. But when its permanent committee, appointed in 1873, five years later, tried to enlarge its sphere of activity and seemed to seek a direct influence on the development of the administrative statistics of various countries, several of the principal countries of Europe (Germany, England, Russia, Sweden and many more) declined further participation on the work of the committee. That ended the tale of the Congress.

The successor of the Congress was the International Statistical Institute (ISI) founded in 1885. The Institute is a private society with

200 members at the most besides a maximum 20 honorary members. According to the statutes such persons should be elected who have distinguished themselves in administrative or scientific statistics. In accordance with the statutes, members of the Institute consist of a considerable majority of administrative statisticians.

Lowsky (1925) says that

Prudence in the measurement of possibilities is the first virtue in a statesman.

These possibilities (economic and political) a statesman can hardly control without statistics and the measurement of these possibilities is the nucleus of the science of statistics. Though it may doubtless still belong to the exceptions, that the political leaders of the various states have this first virtue they still lack, far too often, the power of placing the respective countries' most eminent statisticians at the head of national statistics. The world's administrative statistics is now also such that no country, possibly with the exception of Iceland, has a vital statistics corresponding to the possibilities of the country, and to the present requirements of statistical science. Regarding the other great sphere, when it concerns the population question, or agricultural statistics, Professor East, one of the world's most celebrated experts in this sphere, said at the World's Population Conference, that the reports circulated by the International Institute of Agriculture are not satisfactory for the critical-minded statistician. Only a few countries have reasonably sound systems of agricultural bookkeeping, and these systems having been developed independently are not comparable in type or accuracy. The data from the rest of the world consist of indifferent guesses. Such is the world's present statistics, although the ISI has existed for nearly half a century during which the science of statistics has made a gigantic progress.

[7] The ISI's attitude towards science may be illustrated by the blackballing of the statistical master, to whom this volume is dedicated, the first time he was up for election. At the session of the ISI in Rome, 1926, sampling was subject to exhaustive treatment. Notwithstanding the many and long reports, the great Chuprov's contributions were the most important. The Editor of this journal wrote to him asking for his statements. He answered:

I have given them to the Institute and have no copy, so you must wait until the Bulletin comes.

The *Bulletin* came with twelve lines instead of the full report he had handed it. Chuprov himself said that he was satisfied with this report. So the lost papers may, on the strength of this statement, be said to be nearly irreplaceable, above all at a time like the present, when the employment of sampling has nearly enough become a thing of fashion, especially among those who lack insight into the great difficulties this method's employment offers in general. This method can hardly be handled by others than masters of statistics. Not even in the practical activity of the ISI's own office does the science of statistics get its rights. The office is almost a parody of the plans of its original proposer.

In all times and in all countries the population is the chief subject of statistics. Census is the most necessary and, at the same time, the most

difficult operation. Population is the measure of the strength, the source of the wealth and the political thermometer of the power of states, said Francois Neufchâteau already in the sixth year of the French Republic. The investigations of agricultural statistics comprise the most important parts of national fortune. In spite of all this the international statistical activity in the form of Congress or Institute during more than 3/4 of a century has not yet reached such a development that it has given, or even indicated with certainty the foundation on which future population statistics has to build or how statistics concerning cultivation and produce of the soil are to be arranged.

The ISI has not taken yet any real and far reaching attempts to make the population register. Extensive reports on this institution are given in many essays in this volume, the basis of not only the future population statistics but also of all other statistics and thus even the basis of a state administration, whose chief contents shall be no longer property but man, the administrator of property. The ISI displayed so little interest for, or estimation of statistics on the most important sources of living of the population that the international endeavours in this field had to procure a special organisation.

No more has the ISI reached any results of fundamental importance for commercial statistics. An extensive international list of goods in which each commercial specimen is included under a given number, such a statistical record does not yet exist. As long as this is the case, commercial statistics has no uniform scientific basis. On the whole, ISI has been an indifferent spectator of the proceeding and development of commercial statistics of the world, a development which has no contact with science and is influenced by the interests of the states or special groups. The present activity of the international bureau of commercial statistics at Brussel is one of the many distressing consequences of the Versailles peace. Its reports, in which there are no traces of influence from the principle (perhaps more important for commercial than for any other branch of statistics) that the value of a report also depends, in a very considerable degree, on the swiftness with which it is published, are nothing but parodies on good international statistics.

Moreover, they have an injurious effect by impeding the establishment of commercial statistics of which the world is in such an urgent need in the present political and economic chaos. From their continued existence the conclusion may be drawn that the chief commercial leaders and pioneers have an insufficient understanding of what is the sole way out of the present chaos and thus cannot find it. The great international chamber of commerce often blows the trumpet for an appropriate classification of the production of commercial specimens according to the natural qualifications of the different states.

But how is such a planned development realised as long as the indispensable statistics does not exist? Even when this melody was played at Stockholm last year no sounds were distinguished out of the orchestra of the chamber from the instrument which is made to and should give the leading tune, that is the all-embracing, scientifically established commercial statistics with pervasively exact data about the

origin and consumption country of the goods. Still in the mentioned year the international chamber of commerce had no definite plan for statistics on world commerce. This statistics is the logical basis for the activity of the chamber.

[8] The ISI and its activity have reached such a point that one of its vice-residents appointed quite recently emphasized, when laying down his thanks for the charge entrusted to him, that he will do his utmost in collaboration with the president and the two [other?] vice-presidents to enable the Institute not only to hold its position within the boundaries of existing international institutions but to develop [it] according to present demands. The portion of this statement devoted to criticism is left out of the report from the congress concerned.

For the continued existence the ISI is in urgent need of a complete reorganization and a new scheme for its work. Members must be statisticians in the modern sense, if the purpose is that of promoting the theory and practice of scientific statistics. For the maintenance of the ISI the majority of the clerks in the workshops producing statistical tables must go. It is the steadily increasing statistical activity from Geneva that now drives the ISI to debates concerning its future existence. To lay a secure basis for this future existence it has been proposed to divide the international statistical work by limiting the work in the Institute and that in Geneva so that its scientific and critical part should be left to the ISI while the editorial should be the duty of Geneva. But this proposal can scarcely be earnestly deliberated by those who know the science of statistics and the activity of the ISI and Geneva up to now. By the dilettante activity of its bureau the ISI has revealed its incapability for the scientific part of the statistical work at Geneva, which in turn must be penetrated by scientific investigations if carried on with good results. It is a fact that many eminent persons among recent members of the ISI are no longer opposed to the idea that the Institute must close its doors.

[9] When the League of Nations took up its great statistical investigations no scientific basis was laid for this activity from the very beginning. Its leaders, however, have realized the importance of statistical science for practice of statistics. One of their first tasks, therefore, was to call in for the planning of practical work even (?) persons whom they believed to be experts on the theory of statistics. The secretary general and many other members of the ISI have been among these experts. Their proposals, however, have been almost too often of such character that they have been laid ad acta [set aside]. But the most prominent of statisticians of the present time have not, with a few exceptions, been called in. Thus, the first and foremost among all of them in the theory as well as in the practice of statistics, Chuprov, has never been consulted. Could they get a more apt person for indicating the norm for the statistical work of the League, directing it simultaneously on a sound scientific basis towards its high ethical aims for the world peace?

In his relation (?) of the origin, structure and working of the League of Nations, Howard Ellis also deals with its statistical activity, even if not in the same meritorious way as he treats many other phases of its work. It is said that the economic and financial section is probably the

only centre in the world where information of public finances, central banks, of armaments etc. is collected from nearly all the principal financial administrations. It is then issued as reports that are, and have to be not only complete but rigorously accurate since they have to stand the scrutiny of some 50 odd Treasury departments. It is not probable that the great financial statistician Robert Meyer would have been able to join in that judgement, and a modern financial statistician can add that these reports do not offer all that they will be able to give if the great possibilities of the League were exhausted and if the leadership was given to modern statistical science.

The health section issues a number of publications, is also responsible for two series of handbooks on, respectively, the methods for comparing vital statistics in use in a number of countries; and the general organisation of a number of national health services. The latter series is summed up in the *Annual Health Yearbook*, which shows the progress achieved in public health matters in the countries under review. Here again, one says, the secretariat through its competent section, has become an unique source of reliable and authoritative information from all over the world. It is issued in a clear, practical and readable form for the use of both private individuals and official and semi-official organizations.

Such a judgement is given on a yearbook whose greatest merit seems to be that it is extraordinarily well suited to teach us how statistics must not be. Uniformity, scrutiny and complete comparability which should first and foremost characterise every such relation comprising many different countries are wholly lacking here. The section shows its lacking competence to organize and lead a work which is otherwise particularly needed, and the staff does not display any remarkable skill in treating the comparatively simple statistical problems occurring here. For a rational treatment of statistics of health it is necessary, first and foremost, to have a good knowledge of man under healthy conditions which is often lacking by the physician. He who is competent to deal with general population statistics, also owns the best requisites for a good treatment of statistics of health.

The enormous statistical work, nearly the biggest in the world, performed by the League is, to a too high degree lacking a scientific basis. Against the fundamental rule of good practical statistics and centralisation the work is going on in three different sections so that one is not always conversant with what another has published. Instead of being a world paragon for good universal statistics the statistics of the League has contributed to the lowering of the claims on good statistics in a world whose escape from chaos seems to depend more than on something else on good statistics and on empiric social ethics statistically founded. At the same time the world-embracing statistics of the League renders difficult and sometimes wholly restrains the continuation and further development of the good statistics occurring here and there in the world.

[10] The mentioned report on statistics concerning laws and criminality given in 1860 by Brougham to the International Statistical Congress in London begins with the following introductory remarks by that renowned statesman;

It is hardly necessary that I should begin ... by stating the infinite importance of this subject. Judicial statistics give the means, and the only means of ascertaining the effects of any laws that exist, of any description, civil or criminal, in any country, and the effects produced by a change in any of these laws. No science of legislation can be said to exist, nothing that deserves the name of an inductive science, if we have not the means of knowing what the effects of any law existing, or changed, or altered, or revoked have been. And what the effects have been, what the influence has been of any of these changes which from time to time are made for the improvement of our laws. Or what the existing defects of those laws are, how ancient soever [however] to which we are accustomed, without a thorough examination of the administration of those laws in the courts of justice in which they are administered. Without that full examination we are entirely in the dark upon their tendency, their effects, or their excellence, – all is perfectly unknown to us, – we have not the facts upon which our conclusions must be founded.

So, without judicial statistics, without a statement of the operation in minute detail of all the laws that exist and all the changes that are made in those laws, we cannot be said to possess anything like a inductive science, such as Lord Bacon, I will not say established, for it had been established before him, though partially, but as he reduced it to a system, the science of induction, from facts ascertained.

However brilliant were Brougham's arguments for the importance of statistics to the judicial system, the judicial statistics has not yet reached such a degree of development that it can approximately fulfil the duties he assigned to it. In most states the judicial statistics is scarcely more than a mechanical registration of the resolutions of the judicial institutions. Only a faint attention is paid to the aim and purpose of the resolutions, and the data of the judicial statistics are almost never suited to form a basis fit for international comparisons or even for comparisons between different periods of time in the same country.

By way of example the Swedish statistics on intemperance accounts the number of delicts condemned during the year. From a social scientific point of view it is impossible to use the number of delicts of intemperance in a country. Not only the social notion of intemperance has altered much during the latest generation, but great changes have also taken place in the judicial system with regard to the cases of alcoholism. Since the Supreme Court of the kingdom has ascribed full evidence to the statement of the arresting policeman against the arrested person this one has in many cases been placed wholly without rights. Even if any premium is no longer paid for the arrest, as was still the case some years ago, the lawlessness continues as long as the arrested person has no right to have his condition of temperance examined by those who might be considered of greater capability to judge ... Even other circumstances not seldom exercise a great influence on the number of delicts. Thus, it has been told that, when, at a case of intemperance at Stockholm about 20 years ago, the respective inspector of police together with the police sergeant in

service at the occasion asked the superintendent to suspend the ordained indictment, they got the answer:

If there had ben no political regards to take, there would never had been any indictment.

Still at the beginning of this century the superintendent of police in the second town of the kingdom [Gothenburg] urged his inspectors to persecute his adversaries of opinion.

The related here circumstances are in no way limited to the country where Pufendorf was once active. In all countries that have in some respect or other been influenced by Pufendorf's works similar circumstances are still not rare. Judicial statistics is not yet capable to perform the tasks assigned by Brougham and so they will scarcely ever become in the police state which is still the prevailing form of government even in most of the democracies of present time.

The judicial science which was in Brougham's mind has not become a reality yet, and it cannot be foretold if or when it happens. There exists no supreme court deciding over the transactions of all nations, no international judicature, although centuries' work has been devoted to establish it. A mere court without the help of other activity cannot protect the world from war.

[11] The League of Nations has no absolute conception of justice. It proceeds by patient and continuous discussion, shifting perhaps from one forum to another until solution is reached satisfactory to all. Together with the effort to maintain peace the administrative work of the League, such as relating mandates, minorities, etc., and the effect upon public opinion throughout the world must be regarded as equally significant for the future. A clear distinction cannot be made between the political and non-political activity of the League. But whatever it tends to, its leaders must never forget the words cited above, the conclusive words of England's great scientist Whitehead in his series of lectures at Harvard University on science and the modern world.

The League of Nations must appeal to the forum of science if its activity is to have any future, that is, to a forum where the first place is given to the science which has the greatest importance for investigation and critical estimation of peoples and states and their condition in all the respects concerned in the activity of the League. This science is nowadays statistics, both on account of its general condition and its content. If science, after all, means careful and continued observation of phenomena which the ordinary man regards carelessly and superficially, then statistics, the science that teaches how this observation must and should be made, will occupy the highest rank among sciences.

The chief content of statistics is man. It is the great task of statistics to investigate human activity in physiological, economic and ethical respect and represent the results in such a form that the social activity of man gets the scientific basis which is still lacking. Social activity must be directed and settled by the empirically established social ethics resting on the foundation built by statistics and towering over the edifice for social sciences planned by the great Lexis.

A place in the foreground is also given to statistics in the second sentence below the headline of this article, cited from Whitehead's

famous collaborator, Bertrand Russell, in their great work (1910 – 1913). There, they say that [the author repeats the quotation]. On whatever reasonable tasks man sets his mind, he cannot nowadays do, in almost any circumstances, without the science of statistics. This science is of the very greatest need when man is struggling with the greatest of all tasks, the task of conquering himself and forming a world association where it is no longer necessary to think or speak of war.

[12] The science of statistics is the Monitor of peace, thanks to which man will maintain and develop his humanity in the fight against all the Merrimacs [all the men of war] that the imperialism, that mixture of trade and nationalist rivalries, will surely try to produce for a long time still.

A statistical forum is the condition for a world peace. How, when and where this forum should be erected are problems requiring a speedy solution. Here, I only want to touch the first question and to answer it by saying that this forum might be such that its watchers, led by and responsible to science and for the rest in perfect personal liberty, will be enabled to procure and examine the material of knowledge necessary for the universal peace. The international statistical testing institute, or whatever its name may be, will have to examine the material which will be made the basis for the solution of the great problems of humanity as well as it will have to procure that material. The more swiftly and safely the institute can perform these tasks the sooner statistics will be the saviour from chaos.

For a soon establishment of the statistical forum it is necessary to enlighten the people about the great aim and importance of statistics. Public opinion is not yet sufficiently alive to the importance of obtaining accurate information. There is not sufficient pressure on the governments to make them devote adequate attention to the organisation of statistics. If this be the case, many of the governments consisting of hand-to-mouth politicians will have to resign and the practical politics will be penetrated and supported by statistics and social ethics on the basis of experience.

Education is the acquisition of the art of utilisation of knowledge. There is only one subject-matter for education and that is life in all its manifestations. Statistics which gives the best knowledge has its given place in a future educational system as the basis for communication of knowledge about life, even in the form of society the future of which is depending on its members' knowledge of the subjects which form the chief contents of statistics. The first schoolmaster who has been a minister of education in Wargentin's country had maintained this service for some years when the proposal to introduce statistics even in the people's school was made to him. He did not reject the proposal and declared that this was the most important scheme proposed to him since he had become a minister. People who are unable to use properly statistics are not yet capable of self-government, as Oldham says. To be thoroughly capable to use something one must know it. The self-governing state consists of self-mastering individuals who begin to learn statistics already in the elementary school, where place

can be made also through a reformation of the mathematical instruction.

Notes

1. Independence (of the appropriate events) was forcibly stated by Knies (1850) and indirectly even by Butto (1808, p. XI):

Statistics is a science concerned with the ability to understand and evaluate statistical data, to collect and treat them.

2. Apparently, leading to statistical errors. Somewhat below the author formulates a similar statement, but neither is hardly true in the literal sense.

3. Here is a statement by Czuber from another source; he did not provide an exact reference:

Actuaries will err if they completely restrict their deliberations by mathematics. For reasonably applying mathematics it is necessary to study thoroughly the terrain.

About 1841, Gauss (*Werke*, B d. 12, pp. 201 – 204) preceded Czuber: *If only based on numbers, applications of the theory of probability can be greatly mistaken.*

4. Meyer is insufficiently known, and Andersson's information about him is valuable. He published three books and many papers in the *Hdbuch der Staatswissenschaften* and *Öster. Staatswörterbuch*.

5. Fisher is the main creator of modern mathematical statistics and Andersson's opinion about this discipline is utterly wrong.

6. Andersson greatly exaggerated Lexis' merits.

7. Quetelet is known to have been attempting intensively to standardize statistics. Here is his forceful statement (1846, p. 364):

When two different states are meant, it seems that they are glad to prevent any kind of comparison [of their statistics].

Schlözer (1804, §§ 14-3 and 15-12) and even Leibniz recommended comparisons between states and studies of time changes in a given state.

8. When discussing the history of the ISC, the author did not explain its structure, and many details remain incomprehensible. Zahn (1934, pp. 3 – 4) and Nixon (1960, p. 9, Note 1) describe the obliged self-dissolution of the ISC, and Zahn noted that it was more or less an official body and therefore yielded to political influence of some states. Its requirement to standardize national statistics had been interpreted as interference. Outstanding statisticians became convinced enemies of the ISC.

With reference to the Vienna newspaper *Allgemeine Zeitung* for 14 June 1885, Nixon remarked in 1876, since the Franco – Prussian relations worsened, Bismarck forbid Prussian statisticians to participate in the work of the standing committee of the ISC. I note that Saenger (1934/1935, p. 452) remarked but did not justify that Bismarck had regarded statistics superfluous.

9. The first volume of that periodical was devoted to Bortkiewicz in connection with his sixtieth birthday, but the described decision can apparently be explained by his individualistic attitude.

10. Reports of Scandinavian countries were included.

11. Cf. Schlözer (1804, § 15.11): *Statistics and despotism do not get along together.*

12. Andersson did not know about the French judicial statistics or Poisson (1837), see Sheynin (2017).

13. If understood literally, this is impossible.

Constantine the Great (ca. 285 – 337), emperor in 306 – 337.
Before death became a Christian.

Democritus (460 – 380 BC), philosopher

Heraclitus (530 – 470 BC), philosopher

Tiberius Claudius Nero (42BC – 37), emperor from 14.

Xenophanes (sixth – fifth centuries BC), philosopher, founder of the Eliats

Brougham H. P. (1778 – 1868), state figure, lawyer

Bury J. B. (1861 – 1927), philosopher

Heiberg J.L. (1791 – 1860)

Rankiaer C. (1860 – 1938), botanist

Whitehead A. N. (1860 – 1947), mathematician, philosopher

Ionia, region in western Anatolia

Bibliography

Butte W. (1808), *Die Statistik als Wissenschaft*. Landshut.

Chuprov A. A. (1922), Review of the report of C. Gini (in French) *On the problem of raw materials and foodstuffs* published by the League of Nations. *Metron*, t. 2, No. 1 – 2

Festschrift (1914), *Festschrift für Franz Klein*. Wien.

Knies C. G. A. (1850), *Die Statistik als selbstständige Wissenschaft*. Kassel.

Lawsky H. J. (1925), *Grammar of Politics*. London. Routledge, 2014.

Nixon J. W. (1960), *History of the International Statistical Institute, 1885 – 1960*. The Hague.

Pufendorf S. (1673), *Über die Pflicht des Menschen und des Bürgers nach dem Gesetz der Natur*. Frankfurt/Main, 1994.

Quetelet A. (1846), *Lettres sur la théorie des probabilités*. Bruxelles.

Saenger K. (1934/1935), Das Preussische statistische Landesamt 1805 – 1934. *Allg. stat. Archiv*, Bd. 24, pp. 445 – 460.

Whitehead A. N., Russel B. (1910 – 1913), *Principia mathematica*. Cambridge. Reprint of 2nd edition: Cambridge, 1978.

Zahn Fr. (1934), *50 années de l'Institut International de Statistique*. München.