Studies

in the History of Statistics and Probability

Vol. 18

Compiled by Oscar Sheynin

Berlin
2020

Contents

I am the author of all the contributions listed below

oscar.sheynin@gmail.com

Notation

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

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

I

On the Prehistory of the Theory of Probability

Arch. Hist. Ex. Sci., vol. 12, N. 2, 1974, pp. 97 – 141

1. Introduction

Evidently, none of the traditional sciences busies itself about the accidental, says ARISTOTLE[1], continuing that *this* (the accidental) *none of the recognized sciences considers, but only sophistic,* and repeats himself m other places[2]. However, this opinion is wide of the mark since neither does the modern theory of probability busy itself with chance, but rather with the laws of chance, with the probable[2a]. And ARISTOTLE describes rhetoric as an art of persuasion based on probabilities (§ 3.2). Moreover, reasoning on the probable abound in various sciences in antiquity. The study of this aspect of various sciences before the origin of the theory of probability (i. e., before the second half of the 17th century) is attempted in §§ 3 – 8 while § 2 is devoted mainly to ARISTOTLE and his notable commentator, THOMAS AQUINAS. § 9 is a short account of appropriate philosophical reasoning in the new time with a special reference to JAKOB BERNOULLI, and general conclusions are formulated in § 10.

It proved difficult to incorporate the prehistory of the theory of errors into this article, and it has been dealt with separately[3].

2. Randomness and Probability in Antique Philosophy

2.1. Prior to ARlSTOTLE. The period prior to ARISTOTLE is known partly, possibly even mostly, from his own account[4]. Thus (196a 10)

Certainly the early physicists found no place for chance among the causes which they recognized. … This is strange, whether they supposed that there is no such thing as chance or whether they … omitted to mention it and that too when they sometimes used it, as Empedocles does. … He tells us that most of the parts of animals came to be by chance.

There are some too who ascribe this heavenly sphere and all the worlds to spontaneity. … This statement might well cause surprise. For they are asserting that chance is not responsible for the existence or generation of animals and plants, nature or mind yet they assert that the heavenly sphere and the divinest of visible things arose spontaneously.

Others, continues ARISTOTLE (196b 5), *believe that chance is a cause, ... but that it is inscrutable to human intelligence. …* Still others (195b 35) *say that nothing happens by chance*.

In each case the unnamed person(s) is DEMOCRITUS, says the editor of ARISTOTLE. The ambiguity of DEMOCRITUS in his attitude towards chance, which ARISTOTL'E does not mention, seems genuine. This is proved by comparing the opinions of various ancient commentators, including PLATO. Indeed, some of them thought that, according to DEMOCRITUS, people made a fetish of chance so as to conceal their own want of sense while others ascribed to him the explanation of everything by chance[5].

2.2. ARISTOTLE. He was the first to attempt an explanation of chance. He included this explanation in the general context of his teaching of causes[6], repeatedly mentioned chance, and also accidents and coincidents. An accident, says ARISTOTLE[7], actually introducing this term into classical philosophy, is *something which may possibly either belong or not belong to any one and the selfsame thing*.

As to chance (and change), these[8] are *characteristic of the perishable things of the earth*. Some effects can be caused incidentally, i.e., by spontaneity and chance (198a 5); chance is opposed to mind and reason[9] and its cause *cannot be determined*[10].

The products of chance and fortune are opposed to what is, or comes to be, always or usually.

Similar assertions are made about accidents[12], chance conjunctions[13] and coincidences[14].

ARISTOTLE also distinguishes between chance and spontaneity[15]:

Chance and what results from chance are appropriate to agents that are capable of good fortune and of moral action generally. Therefore necessarily chance is in the sphere of moral actions.

The spontaneous on the other hand is found both in the lower animals and in many inanimate objects.

Possibly the best account of ARISTOTLE'S concept of chance is given in a rather rare source by JULIENNE JUNKERSFELD[16]: by chance ARISTOTLE means something which takes place

Occasionally; has the character of an end; is such that it might have been the object of a natural or of a rational appetite; was not in fact the object of any appetite but came into being by accident.

As shown above, the circularity of this definition is indeed present in ARISTOTLE'S writings. JUNKERSFELD (p. 78) also mentions the essential difference between chance as understood by ARISTOTLE and in the modern sense: in the latter case chance is not connected with intention (or, rather, with its non-fulfilment), but it seems that by referring only to COURNOT she overestimates his role in the formation of the concept of chance.

I mention now three examples of chance events given by ARISTOTLE[17] (JUNKERSFELD mentions a dozen):

Digging a hole for a plant, someone finds a treasure (not a rusty nail, which can hardly be *the object of a natural or of a rational appetite*). ARISTOTLE calls this an accident in its first meaning (which actually is the same as chance). I notice that the difficulty of dividing unexpected events into remarkable and usual is here considerably less than in a similar division of outcomes of a random event in natural science needed for evaluating the probability of *remarkable* outcomes.

A meeting of two persons which takes place *by chance*.

Mistakes *in the operations of nature* which give rise to *monstrosities*. The first departure of nature from the *type is that the offspring should become female instead of male; … as it is possible for the male sometimes not to prevail over the female … either through youth or age or some other such cause.*

Thus ARISTOTLE attributes the birth of a female to chance, adding though that this is a natural necessity. Then, in general, for ARISTOTLE the outcome of a chance event depends on rather small changes in the chain (chains) of previous events. Of course, this is true only if his connection of chance with non-fulfilment of intention is disregarded. This general illustration is the same as was used by POINCARÉ[18]: a chance event takes place when, in conditions of an unstable equilibrium, very slight causes determine considerable effects.

ARISTOTLE connects the chance occurrence of sex with natural necessity, i. e. with a definite (optimal) ratio of males and females for any given species. Though he does not elaborate, this seems to be the first statement connecting chance and necessity.

ARISTOTLE'S writings also contain reasoning on the probable. A probability[19]

Is a generally approved proposition: what men know to happen or not to happen, to be or not to be, for the most part thus and thus ..., e. g. the envious hate.

Formulating what actually is a rather weak corollary of the strong law of large numbers, ARISTOTLE[20] notices that *what is improbable does happen … therefore it is probable that improbable things will happen*. On the other hand, ARISTOTLE understands that a very rare event either in games of chance or in nature is impossible (see § 5). He also describes rhetoric as an art of persuasion based on probabilities (§ 3.2) and, speaking about persuasion in poetry and anticipating ARCESILAS and CARNEADES, he even introduces a rudimentary scale of subjective probabilities[21]: *a likely impossibility is always preferable to an unconvincing possibility*.

Several times ARISTOTLE mentions luck[22] and fortune[23]. His understanding of these concepts is that they qualitatively express deviations from reasonable expectation, which is another qualitative notion. Even if expectation is understood as a numerical measure in the sense of the classical theory of probability, ARISTOTLE'S luck and fortune will still differ from the concepts used by Jakob BERNOULLI: although in one place ARISTOTLE (1361 b – 1362 a) holds that luck equally means luck accruing beyond expectation or escape from expected evil, elsewhere he (1207a 30) supposes that

Good fortune would seem to consist to a greater extent and more properly in the obtaining of good … while the escaping of evil is a piece of good fortune indirectly.

ARISTOTLE in fact qualifies his own opinion on the impossibility of a science of the accidental (§ 1) when he[24] attributes strategy and navigation to *matters involving art, but into which chance largely enters*. However, he refutes his own qualification by saying (1247a 20) that in navigation *not the cleverest are the most fortunate, but it is as in throwing dice*.

Does this really mean that there can be no science of navigation? A saner opinion is (PLATO[25]) that, though

Chance is almost everything in the arts of the … pilot, and the physician (§ 6), and the general[26] *… yet in a storm there must surely be a great advantage in having the aid of the pilot's art*.

S. SAMBURSKY[27] quotes SIMPLICIUS as saying that PLATO (and ARISTOTLE) called natural science the science of the probable (*ecotologia*). His failure to cite the specific passage is much to be regretted.

Other interesting features in ARISTOTLE are, first, his actual admission of the probable in biology and medicine (§ 6.2.2) and, second, his use of the mean as a moral category. Thus[28], *temperance and courage are destroyed by excess and defect, and preserved by the mean*.

In medicine (§ 6.2) the mean was considered as the ideal state (of health), while in the theory of errors[29] and games of chance (§ 5) the (arithmetic) mean came to be considered as possessing definite stochastic properties.

2.3. EPICURUS and LUCRETIUS. In EPICURUS' opinion[30] atoms *recoil and recoil, whenever they chance to be checked by the interlacing with others.*

This is an attempt at a qualitative chance explanation of a physical phenomenon. In another source[31] EPICURUS is quoted as saying that

It …. may be that according to the diversity of the regions traversed (by the stars) *in some places there are uniform tracts of air, forcing them forward … in others these tracts present such irregularities as cause the motions observed,* i. e.*, the wanderings of certain stars and the regular movement of certain other stars.*

This passage possibly means that EPICURUS concerned himself with laws of chance. See also § 8.1.1.

Swerves of atoms neither the moment nor the direction of which can be known beforehand are essential in LUCRETIUS' physics[32] and even serve to explain naively the occurrence of free will. In the absence of any quantitative description they also serve as a chance mechanism which brings about determinate effects. As in ARISTOTLE (§ 2.2) and, possibly, in EPICURUS (see above) this again is a connection of chance and necessity.

Some modern authors hold that ancient atomists did not consider random events. Thus, bearing in mind philosophers before ARISTOTLE, B. RUSSELL[33] noticed:

There is considerable reason to think that weight was not an original property of the atoms of Leucippus and Democritus. It seems more probable that, on their view, atoms were originally moving at random. … As a result of collisions, collections of atoms came to form vortices.

It was common in antiquity to reproach the atomists with attributing everything to chance. They were, on the contrary, strict determinists. Democritus explicitly denied that anything can happen by chance ... Leucippus … is known to have said <u>Naught happens for nothing, but everything from a ground and of ncessity</u>*. It is true that he gave no reason why the world should originally have been as it was; this, perhaps, might have been attributed to chance. …*

Causation must start from something, and wherever it starts no cause can be assigned for the initial datum. … The Creator Himself is unaccounted for.

Even if this was meant to apply also to post-ARISTOTLE philosophers, there still remains room to see both chance and causation in their works, and it seems that almost every philosopher, at least until and including I. KANT[34], chose to find in them either the first or the second.

2.4. THOMAS AQUINAS. He was one of the main commentators of ARISTOTLE. Striving towards his general goal, which was to unite faith and reason and to adapt pagan ARISTOTLE to Christianity, he could not fail to study chance.

THOMAS' commentary on ARISTOTLE'S *Physica* seems to suggest that he was the advocate of the Philosopher's understanding of chance[35]. THOMAS' main work[36] also contains explanations of the concept of chance:

The effects willed by God happen contingently … because God has prepared contingent causes for them.

Casual and chance events are such as proceed from their causes in the minority of cases and are quite unknown.

He connects chance with hindrances[37]:

Some causes are so ordered to their effects as to produce them not of necessity but in the majority of cases, and in the minority to fail in producing them, which is due to some hindering cause.

An example of such a hindering cause at work is *the production of woman* which THOMAS[38] explains by referring to ARISTOTLE (§ 2. 2) and repeating his connection between chance and necessity:

On the other hand, in the relation to the universal nature, woman is not misbegotten, but is included in nature's intention as ordered to the work of generation.

It is really difficult to see how ARISTOTLE, much less THOMAS, combined chance (which takes place *occasionally*) with the production of females. Even in the absence of any vital statistics it was possibly known that the numbers of men and women in any "normal" locality do not differ significantly, at least not in a proportion which would justify the attribution of the production of women to occasional chance effects.

Another reasoning on the connection between chance and necessity occurs in the same source[39]:

.. *Contingence arises from matter, for contingency is a potency to be or not to be, and potency pertains to matter. But necessity results from form, because whatever is consequent on form is of necessity in the subject. But matter is the principle of individuation, whereas the universal comes from the abstraction of the form from the particular matter. … The contingent … is known directly by sense and indirectly by the intellect, while the universal and necessary principles of contingent things are known by the intellect. Hence if we consider the objects of science in their universal principles, then all science is of necessary things. But if we consider the things themselves, then some sciences are of necessary things, some of contingent things*.

Thus, widening the concept of chance to correspond to individualization and uttering a dialectical comparison of individual with universal, THOMAS departs from ARISTOTLE (§ 1). However, his bifurcation of science seems not to be elaborated: on the one hand, no science is interested in the individual; on the other hand, each science has to do with different levels of abstraction so that something general (e. g., a triangle) is only individual as regards the next stage of abstraction (e. g., a polygon).

THOMAS[40] also tells us of his bifurcation of the accidental:

Accidents which are altogether accidental are neglected by every art, by reason of their uncertainty and infinity. But accidents of this kind are not what we call circumstances, because circumstances … are in a kind of contact with (the act) … *Proper accidents … come under the considerations of art*.

This is not quite clear, but it is at least possible to say that likewise the theory of probability can do nothing with, to paraphrase THOMAS, randomness which is altogether random (i. e., possesses no law of distribution).

BYRNE[41] considered all the works of THOMAS from the point of view of probability. *For the medieval*, he notices on p. xxiii (and also for ARISTOTLE, see § 2 2),

It is an opinion which is or is not probable, or is more or less probable; and the notion of opinion refers not only to an <u>*objective*</u> *proposition but to a subjective commitment to that proposition*.

Probable opinions and conjectures, says BYRNE, are THOMAS' grounds for proceedings of law courts (see § 3.2). His is also an

Appeal to a kind of moral law of large numbers[42]*, so that it is more probable that a given group will do that to which it is inclined by a heavenly body than that one single man would so act. … Therefore, astrological predictions are verified ut in pluribus*.

As it seems, this appeal is rather to a simple corollary of the law of large numbers. See also §§ 7 and 9.1.

BYRNE (Ibidem, pp. 202 – 208) also holds that THOMAS used a rudimentary frequency theory of probability but his arguments are not sufficiently convincing. Lastly, in a more general sense he holds (p. 296) that there is a similarity (A) between THOMAS' theory of probability and the modern logical theory of probability and (B) between his theory of contingency and the modern (MISES-type?) frequency theory of probability.

THOMAS' writings certainly and essentially influenced scholars of subsequent centuries but it seems that his influence was mainly indirect, brought about through religion, jurisprudence etc., while the formation of the theory of probability proper, as of any other scientific discipline, was mainly caused by direct practical requirements. BYRNE himself emphasizes not THOMAS' direct influence but rather the (possibly unconscious) continuity of ideas and points out that his work is a contribution to the global, largely unsolved problem of studying the correlation between medieval and modern science.

3. Jurisprudence

This title should properly read *Randomness and probability in jurisprudence*, but I have here (and similarly in §§ 4 – 8) chosen the shorter version.

3.1. Ordeals. AL-BIRUNI'S account[43] of law procedures in the 11th century India includes evidence of what apparently contradicts modern standards of presumption of innocence:

If the suitor is not able to prove his claim, the defendant must swear There are many kinds of the oath, in accordance with the value of the object of the claim.

Various kinds of the oath should have corresponded to different expectations of gain from an unjust verdict: the higher the value of the claim, the less should be the probability of impunity of a perjury. In a sense this possibly was the case but there did not seem to exist any definite scale of (subjective or objective) probabilities, the less so as everybody obviously believed in divine defence of the right:

If the object of claim was of some importance, the accused was invited to drink some kind of a liquid which in case he spoke the truth would do him no harm.

AL-BIRUNI mentioned still higher sorts of ordeals including the carry of iron

So hot that it is near melting point on hand there being nothing between the hand and the iron save a broad leaf of some plant and under it some few … corns of rice in the husks.

Degrees of ordeals were connected neither with the degree of proof achieved by the parties concerned, nor with the corresponding physical suffering: one of the highest ordeals as described by AL-BIRUNI was to change one's weight in practically no time, an effect hardly accomplishable but at least painless to attempt.

Ordeals had been known everywhere and, in particular, they are described in legal documents of medieval Georgia and Russia. The following is a commentary on ordeals the world over[44]:

We have not to speak of trial; we have to speak of proof. The old modes of proof might be reduced to two, ordeals and oaths, both were appeals to the supernatural. The history of ordeals is a long chapter in the history of mankind. … Men of many, if not all, races have carried the red-hot iron or performed some similar feat in proof of their innocence.

Among our own forefathers the two most fashionable methods of obtaining an indicium Dei were that which adjured a pool of water to receive the innocent and that which regarded a burnt hand as a proof of guilt. Such evidence as we have seems to show that the ordeal of hot iron was so arranged as to give the accused a considerable chance of escape.

Having referred to 13th century documents of a Hungarian monastery, the authors conclude: *it was about an even chance whether the ordeal of hot iron succeeded or failed.*

Thus AL-BIRUNI'S testimony seems to be correct, at least essentially. However, a more "simple" system of ordeals in India is described elsewhere[45]:

§114. *Or the* (judge) *may cause the* (party) to carry fire or to dive under water, or severally to touch the heads of his wives and children.

§ 115. *He whom the blazing fire burns not, whom the water forces not* (?) *to come* (quickly) *up, who meets with no speedy misfortune, must be held innocent on* (the strength of) *his oath.*

§ 116. *For formerly when Vatsa was accused by his brother, the fire burned not even a hair* (of his) *by reason of his veracity.*

In contrast to AL-BIRUNI the two latter sources do not suggest an unconditional divine defence of the right: though God did not allow his hand to be burnt, He left him to swim or sink on his own. How then was it possible to distinguish between the innocent supernaturally remaining under water and the guilty sinking under water quite naturally?

In other words, the intervention of the supernatural introduces additional difficulties in the study of ordeals and it seems impossible to say whether the legal profession (and society in general) really imagined any probabilities attached to the possible outcomes of ordeals.

On the other hand, at least in countries influenced by the catholic church, THOMAS AQUINAS' teaching on miracles may have been widely known. According to him[46], *those things, which God does outside those causes which we know, are called miracles.*

THOMAS attributes various objective ranks of greatness to miracles (art. 8, p. 545) and subdivides each rank into degrees *according to the different ways in which the power of nature is surpassed*. He does not elaborate. But in any case the lowest rank of miracles is when

A thing surpasses nature's power in the measure and order in which it is done, as when a man is cured of a fever suddenly.

It seems possible to place a successful outcome of some ordeals on a par with accomplishment of miracles of the lower rank so that after all the legal profession might have developed a tradition of comparing one or another ordeal with evidence (and probability) of guilt. That successful outcomes were really possible is testified by POLLOCK & MAITLAND (see above) but, and this is the worst point, were not these outcomes so many cooked-up frauds?

.. The innocent believed that God will help him to get safely through the ordeal. Not so it was with KEPLER'S mother[47], a suspect witch, almost sentenced to death (1621):

Sie jedoch ohngeachtet aller ernstlicher erinnerung und Betrawungen der beschuldigten Hexerey und allzuemit uff die Knie nidergefallen, ein Vater unser gebetten, und darauff vermeldendt, Gott solle alda ein Zeichen thuen, wann Sie ein Hexin oder Unholden seye.

Wasn't KEPLER himself, for all his piety, instrumental in reversing the usual procedure (God should have given a sign *wann Sie kein Hexin … seye*!) virtually giving his mother a safe escape?

3.2. Probabilities in Law-Courts. According to SAMBURSKY[48], SOCRATES held that in law-courts *men care nothing about truth, but only about conviction, and this is based on probability*.

ARISTOTLE[49] paid special attention to application of rhetoric in law and, even without emphasizing the matter, spoke about probabilities:

If you have no witnesses … you will argue that the judges must decide from what is probable. … If you have witnesses, and the other man has not, you will argue that probabilities cannot be put on their trial and that we could do without the evidence of witnesses altogether if we need do no more than balance the pleas advanced on either side.

THOMAS AQUINAS[50] also discussed probability in law:

In the business affairs of men, there is no such thing as demonstrative and infallible proof, and we must be content with a certain conjectural probability. … Consequently, although it is quite possible for two or three witnesses to agree to a falsehood, yet it is neither easy nor probable that they succeed in so doing; therefore their testimony is taken as being true. (A reference to ST. AUGUSTINE follows.)

Another line of development begins in India[51]:

On a conflict of witnesses the king shall accept (as true) *the* (evidence of the) *majority; if* (the conflicting parties) *are equal in number,* (that of) *those distinguished by good qualities.*

Then comes another point (§ 108, p. 273):

The witness (in law-suits pertaining to loans), *to whom, within seven days after he has given evidence, happens* (a misfortune through) *sickness, a fire, or the death of a relative, shall be made to pay the debt and a fine.*

Thus it seems that ordeals and oaths comprised the second stage of law-suits, being ordered on the basis of existing evidence (on the balance of probabilities). As to the last point, it is nothing less than possibly one of the first criteria for distinguishing between randomness and divine intervention (determination).

It is also possible that precisely in jurisprudence the first concept of errors of the first and second kind, not yet formalized and, in particular, unconnected with probabilities, came to be used. Thus, ARISTOTLE[52] held that

Any one of us would prefer to pass a sentence acquitting a wrong-doer rather than condemning as guilty one who is innocent[53].

Similar assertions occur also in THOMAS AQUINAS[54]:

It is better to risk being deceived about others more often by having a good opinion of them than to risk misjudging someone even rarely by being suspicious of others.

The peril (to others) *that exists so long as they* (the criminals) *are alive is greater and more certain than the good which might be expected from their rehabilitation. Moreover, even at the very moment of death they have the opportunity to repent and be converted to God. If therefore, … even at the moment of death their hearts do not turn from malice, it can be estimated with sufficient probability that they would never turn away from malice.*

I shall now skip over to LEIBNIZ[55], a lawyer, who described differences of probabilities such as became established in law-courts of his day:

Es gibt … mehr als halb vollständige Beweise, bei denen, man dem, der sich auf die stützt, die Ergänzung durch den Eid gestattet (*man nennt das iuramentum suppletorium*); *außerdem gibt es weniger als halbvollständige Beweise, bei denen man im Gegensatz dazu denjenigen zum Eid zulässt, der den Tatbestand abstreitet, damit er sich reinwasche* (*man nennt das iuramentum purgationis*). *Außerdem gibt es viele Grade von Vermutungen und Indizien.*

Distinguishing between four of them, LEIBNIZ concludes:

Alle Verfahrensformen in der Rechtsprechung sind in der Tat nicht anderes als Arten der Logik, die auf Rechtsfragen angewandt werden.

Auch die Mediziner haben für ihre Symptome und Indikationen viele Grade und Unterschiede, wie man bei ihnen sehen kann.

Though he does not mention probabilities, they are meant to be present, as is testified by LEIBNIZ' *Arten der Logik* (see a related passage from LEIBNIZ in § 5).

A relevant passage occurs in a contribution by E. NAGEL[55a]. In the Middle Ages two witnesses were demanded for a *full proof,* while a doubtful witness counted for *less than half*.

Until now I discussed trials as such. However, trials are based upon laws which, in a context of a given social system, were also worked out so as to correspond with the usual, probable behaviour of men. In the words of THOMAS[56]

(1) *The lawgiver cannot have in view every single case, he shapes the law according to what happens most frequently*.

(2) *In appointing the punishment for theft the Law considered what would be likely to happen most frequently*.

However, the history of legal punishments is almost the history of society in general and I leave the problem at that.

It is well known that both LAPLACE and POISSON busied themselves with applications of probability in jurisprudence but their work is outside my scope.

3.3. Application of Means. It seems that in lawsuits and courts of arbitration arithmetical means of estimates made by different persons were widely used. Such, at least, is the testimony of G. CARDANO[57], who, while describing conditions of a game of chance, notices that *this mean is composed of extremes* (is the semirange), *not as in lawsuits, and valuations, and the like*.

Also, a definite testimony is due to LEIBNIZ[58]:

Die Grundlage, auf welche man baute (in the theory of probability) *kommt auf die Prostapherese zurück, das heißt darauf, dass man ein arithmetisches Mittel zwischen mehreren gleich annehmbaren Voraussetzungen nimmt. Und unsere Bauern bedienen sich auf Grund ihrer natürlichen Mathematik dieses Verfahren seit langem. Wenn eine Hinterlassenschaft oder ein Grundstück verkauft werden soll, so bilden sie drei Gruppen von Schätzern … jede Gruppe stellt einen Taxwert des fraglichen Gutes auf … nimmt man die Summe dieser drei Schätzungen* (an example follows) *und teilt sie durch drei. … Dies ist das Axiom aequalibus aequalia, gleiche Voraussetzungen müssen gleichermaßen in Betracht gezogen werden*.

4. Fine Arts

Ending § 3 with what amounted to the use of the stochastic aspect of the theory of means in jurisprudence, I shall now describe the same subject as manifested in sculpture, one of the branches of fine arts.

It is generally known that the mathematical teaching of proportions was universally applied in antique and Renaissance architecture and, also, that so early a scholar as VITRUVIUS systematically measured proportions of the human body.

L. B. ALBERTI (1404 – 1472), the scholar, architect, sculptor, musician and writer, resumed such measurements. Not only did he develop the procedure of measurement using a specially devised ruler, the *exempeda*, but, what is much more important, he came to use mean dimensions of various models[59]:

I want to establish not the particulars of this man or that one, but as far as possible, that exact beauty granted by Nature and given, as if in select portions, to many bodies. ... I have therefore chosen many bodies which are reputed to be the most beautiful by those who are knowledgeable, and I have taken the measures and proportions of all of these, comparing and eliminating the excesses of the extremes (sic!). *I have selected from many bodies and models those mean proportions which seem to me most praiseworthy.*

Thus, ALBERTI really understood the idea of taking statistical means, and the introduction of this statistical method into fine arts is due precisely to him. To be sure, his method cannot be directly compared with that of QUETELET who, in the 19th century, introduced the concept of *l'homme moyen* (anticipated by BUFFON[60]). QUETELET'S is the concept of man in general, though awkward and even impossible anthropometrically[61], a standard, at least according to QUETELET and his followers, of *moyen* social and moral qualities of man. On the other hand, ALBERTI'S is the concept of a statistically most beautiful man, of *l'homme moyen* taken not simply out of beautiful men in general, but, it seems, out of beautiful men of almost the same constitution, of such a *l'homme moyen* as is completely useless beyond fine arts.

Being a man of education ALBERTI was also an outstanding geodesist[62]:

Regiomontan hat in seinem Briefe vom … 1464 zwei Männer als besonders zuverlässige Beobachter genannt, Toscanelli und Alberti.

It would be extremely interesting to study ALBERTI'S methods of treating astronomical and/or geodetic observations, a problem, which, since I found no original sources, I could not solve.

As to ALBERTI's general place in the history of fine arts[63],

He was the first theorist to advance the system of proportions (in art) *beyond medieval standards and beyond the classical system as well. His two "rules*" (for measuring proportions of bodies) *and his exempeda system of mensuration* (measurement of the dimensions of the parts of the body in terms of the length of the whole) *are all original. Among artists-theoreticians, Leonardo and Dürer both incorporated and developed Alberti's exempeda system, and F. Giazgi described it in his speculative <u>Harmonia mundi totius</u>.*

LEONARDO DA VINCI[64] employed a statistical method resembling that of ALBERTI:

(1) *Look around you and take the best parts of many beautiful faces. … So select beauties … and fix them in your mind.*

(2) *Look at many men of 3 braccia, and out of the larger number who are alike in their limbs. Choose one of those who are most graceful and take your measurements.*

A method of obtaining a "mean" photograph of few kindred persons, or people of a certain nationality or occupation (or criminals) is due to GALTON[65]. According to his idea, extremely popular at the time, composite photographs, as he called them, serve as a means for general cognition of the relevant statistical object. I do not venture to pronounce any opinion about the practical value of this method.

5. Games of Chance

Games of chance are known to have existed at the outset of civilization[66] as had also the drawings of lots[67]. In themselves, these games did not essentially facilitate the development of either combinatorial techniques or of the idea of randomness and probability. The contributory reasons were the imperfection of ordinary dice and the belief in supernatural intervention[68] as well as the comparative complexity of' the rules of many games. Thus, in a throw of four astragali, not only the total number of points, but also the manner of composition of this total had to be considered.

However, by the middle of the 17th century mathematicians of the highest calibre (PASCAL, FERMAT) became interested in the stochastic aspect of games of chance and precisely such games provided them with an opportunity to introduce first numerical (mathematical) notions pertaining to probability.

Later, games of chance were studied by such scholars as HUYGENS, JAKOB BERNOULLI and DE MOIVRE, partly in response to social demands of the day, but also in accordance with the intrinsic logic of development of probability. Possibly bearing in mind that simple games of chance provided mathematicians with precise natural examples of problems, whose solution simultaneously led to the development of general principles and theory, HUYGENS[69] said:

Je veux croire qu'en considérant ces choses plus attentivement, le lecteur apercevra bientôt qu'il ne s'agit pas ici d'un simple feu d'esprit, mais qu'on y jette les fondements d'une spéculation fort intéressante et profonde. Les Problèmes appartenant à cette Matière ne seront pas, me semble-t-il, jugés plus faciles que ceux de Diophante, mais on les trouvera peut-être plus amusants attendu qu'ils renferment quelque chose de plus que de simples propriétés des nombres.

Understandably, HUYGENS underestimated the importance of the *simples propriétés des nombres*. However, it is interesting to notice that in earlier times games of chance were used to provide examples of design rather than of chance. Even so they were used by ARISTOTLE and KEPLER to prove that certain events in nature were designed rather than produced by chance.

Ten thousand Coan throws in succession (whatever this means) *with the dice are impossible*, says ARISTOTLE[70], so that *it is difficult* (impossible) *to conceive that the pace of each star should be exactly proportioned* (by chance) *to the size of its circle* (Ibidem, 289b 22).

Similarly, discussing the appearance of a new star, KEPLER[71] supposes this to be no chance event:

Hingegen will ich aber auch mit denjenigen nicht gemeinschafft haben wölliche diese zusammenstimmung aller dings in Wind schlagen und darfür halten das es des blinden glücks schuld das dieser newe sterne eben gerad diß Jahr Monat tag und ort der grossen conjunction getroffen habe. Dan ob wol war (*zum exempel*) *ein jeder gerader wolgemachter würffel sechs felder hat und eins so wol fallen* *khan als das andere jedoch wan ein anzahl spieler jeder mit vier oder fünff Würffeln nur einen einzigen Wurff thuen sollen und einem under jnen füele das Sechsen auff allen würffeln so wiirde man ein sollichen nit unbillich wegen einer verborgenen kunst verdacht haben und es schwärlich dem glück zuschreiben: angesehen das wol hundert tausendt würffe geschehen möchten ehe wieder einer auff diese weise geriethe. Derowegen … diese wunderbarliche eintreffung der zeit und ort nit gern dem blinden glück zuschreiben wollte: zumahl weil die erscheinung selbsten eines newens Sternens fiir sich allein (auch ohne betrachtung der zeit und ort) nit ein gemein ding ist wie ein spiel wurff sondern ein grosses wunder desgleichen vor unsern zeiten nie erhört oder gelesen worden*.

Such reasoning essentially depends on the correctness of dividing events into remarkable and usual ones. Highly relevant criticisms of stochastic considerations made by a no lesser person than LAPLACE *lui-même* are due to C0URNOT[72], and I say this once more in § 7 in connection with KEPLER's inference about the influence of celestial aspects on meteorological phenomena.

At least one methodological difficulty inherent in the related problem of enumerating usual events is that they are not readily noticed[73]:

Die sublunarische Natur aber, an die fortwährenden unharmonischen Konfigurationen gewöhnt, achtet sie für nichts, weil sie nichts Neues für sie sind. Auch einen harmonischen Winkel ist sie jedoch so gespannt, wie wenn er allein da wäre. So wird es auch übersehen, wenn ein Prognostikum tausendmal irrt; wenn er aber eine mal einschlagt, so hält man das für besonders beachtenswert und aller Mund spricht rühmend davon.

Another, indirect reference to games of chance occurs in KEPLER[74] in connection with his astronomical calculations:

Man muss (*nach der Regula falsi*) *… eine doppelte, sozusagen quadratische Annahme, d. h. in der Tat eine unmathematische Glückspielmethode anwenden*.

That iterative methods did not in KEPLER'S view belong to mathematics proper is evident and is possibly understood by his weariness occasioned by them, but, what is here more important, mathematicians, and, for that matter, after the advent of the theory of probability, astronomers hardly continued to use expressions such as *unmathematische Glückspielmethode*.

However, the main point in the early history of games of chance seems to be that they promoted the general, possibly intuitive idea of definite stochastic properties possessed by mean outcomes.

The mean possible number of points in a throw of dice or astragali served as an estimate of reasonable luck, as a measure of expectation. Thus, CARDANO[75] noticed that for a usual die the mean number of points is 3.5 and that, if more than one ace shows up in a throw of four astragali, that throw is called *the dog, because whatever the other dice may be, the throw cannot exceed the average number*.

An astragalus is a small bone in the ankle of animals and to its four faces numbers 1, 3, 4 and 6 were usually attributed so that the mean outcome of a throw of four astragali is 14. Curiously enough, CARDANO does not mention that, owing to the asymmetry of the astragalus, the probabilities of the occurrences of its different faces are unequal. On the other hand, this very omission seems to strengthen my thesis formulated above.

CARDANO quite often refers to the mean outcome, in particular concerning throws of three usual die, so that ORE[76] seems to be right in saying that the reasoning on the mean outcome was one of CARDANO'S main arguments. That, as also pointed out by ORE, this argument sometimes led him to erroneous conclusions, is in this context not so important.

The dissemination of possibly unconscious ideas on the advantage of mean outcomes is testified by one of GALILEI'S notes[77] where he tells us that gamblers of his day considered that in a throw of three die 10 or 11 points are more advantageous than 9 or 12 points. Supposing that the gamblers intuitively compared the conditional probabilities of events A = (10 or 11 points) with B = (9 or 12 points), one comes to

$$P_A = P\{A/A \text{ or } B\} = 27/52, \quad P_B = P\{B/A \text{ or } B\} = 25/52.$$

And $\Delta P = 2/52 = 0.0385$ was not an insurmountable obstacle for the statistical discovery of the fact that $P_A > P_B$. This fact did really appear plausible even in the absence of statistical observations since both (10 and 11), and (9 and 12) are equidistant from 10.5, the mean (impossible) outcome with the first distance being less than the second one.

An extremely interesting and generally known opinion of LEIBNIZ is concerned with the development of (statistical) decision theory, which came into being only recently (occasioned, moreover, by practical requirements other than games of chance)[78]:

Ich habe schon mehr als einmal gesagt, dass man eine neue Art Logik braucht, die die Grade der Wahrscheinlichkeit behandelt. … Es wäre gut, wenn derjenige, der diesen Gegenstand behandeln will, die Untersuchung der Glücksspiele weiter verfolgte. Und im allgemeinen würde ich wünschen, dass ein gelehrter Mathematiker ein umfängliches Werk über alle Arten von Spielen mit genauer Beschreibung und guter Begründung schreiben wollte, was von großem Nutzen wäre, um die Erfindungskunst zu vervollkommen, da der menschliche Geist besser bei den Spielen als bei den ernsteren Gegenständen in Erscheinung tritt.

6. Biology and Medicine

6.1. Biology. The role of randomness in biology came to be systematically studied only after DARWIN[79] although he himself did not admit it[80]:

I have hitherto sometimes spoken as if the variations had been due to chance. This, of course, is a wholly incorrect expression, but it serves to acknowledge plainly our ignorance of the cause of each particular variation.

Before DARWIN'S time biologists and scholars in general did admit that many biological facts, including rather important ones, were occasioned by chance and biology really was one of the sciences in which chance had been explicitly spoken of. But chance was not yet considered to be the starting moment of (the yet barely discovered) evolution.

ARISTOTLE (§ 2.2) himself attributed the occurrence of one or another sex of the offspring in animals to chance and noticed the connection of randomness and necessity in this phenomenon. The study of relative frequencies of births of both sexes in man played a decisive role in the development of the theory of probability at least during the whole of the 18th century. An additional point in this connection is a regrettably unsubstantiated assertion[81], that

The notion of probability seems to have been mentioned first in China by Sun-Tze about 200 B. C. in connection with the probability that a birth would be that of a boy or a girl.

Explicit utterances on the variations inside a given species are due to W. HARVEY and at least once he attributes them to chance[82]:

To me the form of the egg has never appeared to have aught to do with the engenderment of the chick, but to be a mere accident; and to this conclusion I come the rather when I see such diversities in the shapes of the eggs of different hens.

I do not venture to agree with the first part of his assertion but at least the second one is significant. Still more significant is the statement of the same author[83] that accident is the motive power of the generation of some *creatures*:

Creatures that arise spontaneously are called automatic … because they have their origin from accident, the spontaneous act of nature.

That the theory of spontaneous generation has been abandoned ages ago makes no difference here.

Another author who mentioned variations was KEPLER. Although not a botanist, he nevertheless seems to pronounce a generally accepted opinion[84]:

Finden sich wol einzehle Früchte und Blumen, die 7, 9 oder 11 Fächer oder Blätter haben, wann die species in individuis gemeiniglich variert, aber kein species findet sich nicht, die diese Zahl beständig halte.

HUYGENS[85] maintained that diversities in animals and plants separated in space are due to the divine will:

Il lui a plu … d'établir une certaine diversité de formes entre nos animaux et plantes et les organismes d'outre-mer.

On the other hand, LAPLACE[86], with his celebrated avoidance of the divine, was prepared to accept a tendency of (chance ?) changes in time:

Mais tant d'espèces d'animaux éteintes dont M. Cuvier a su reconnaitre … l'organisation dans les nombreux ossements fossiles qu'il a décrits, n'indiquent-elles pas dans la nature une tendance à changer les choses mêmes les plus fixes en apparence?

6.2. Medicine

6.2.1. Hippocrates. HIPPOCRATIC writings contain a large number of case histories and, as a rule, each one of them ends with a commentary such as[87]

It is probable that, by means of … this patient was cured (or, alternatively, *It is probable that the death … is to be attributed to* …).

The use of qualitative stochastic considerations is his rule, e. g[88].

To speak in general terms, all cases of fractured bones are less dangerous than those in which …

This kind of reasoning seems to form the basis of HIPPOCRATIC art of medicine. Thus, after an apparently matter-of-fact statement[89]

It appears … that the coming on of summer should have done good in these cases. … And yet the summer … was not of itself well constituted, for it became suddenly hot.

HIPPOCRATES gives a general counsel (Ibidem, § 16):

I look upon it as being a great part of the art to be able to judge properly of that which has been written. For he that knows and makes a proper use of these things, would appear to me not likely to commit any great mistake in the art.

He then explains that, knowing the climate and other external conditions, the physician would be able to foretell *the order of the critical days* and to know *when and how* to administer medicine. Obviously this means that the physician ought to know the probable course and outcome of a disease and to act accordingly.

That the constitution and general condition of the patient should be also taken into account is also stated[90]:

(1) *The separation of denuded bones is quicker or slower, according to the mode of treatment; something, too, depends upon whether the compression be stronger or weaker, and whether the nerves, flesh … are quicker or slower in becoming blackened and in dying … it is impossible to define accurately the time at which each of these cases will terminate.*

(2) *Men's constitutions differ much from one another as to the facility or difficulty with which dislocations are reduced.*

Thus, HIPPOCRATES understands that men differ from one another in their response to medical treatment and in time necessary for their convalescence. No such terms as chance or randomness are used, but implicitly they are certainly present.

Even more interesting, though, is the occurrence of what could be called qualitative correlation[91]:

(1) *Persons who are naturally very fat are apt to die earlier than those who are slender.*

(2) *Those who are accustomed to endure habitual labours, although they be weak or old, bear them better than strong and young persons who have not been so accustomed.*

Consider, e. g., the first example. Let the excessive weight of a given person be measured along the x axis and his longevity along the y axis. Also, let segment $[a, b]$ on the x axis represent the domain of slenderness and point c be the beginning of fatness. By plotting results of a large group of observations it would be possible to obtain a quantitative (correlative) measure of longevity as against weight of body.

It is needless to say that HIPPOCRATES should be credited with nothing more than an understanding of the existence of some statistical relationship between weight and longevity. In the absence of any vital statistics he could have hardly thought of anything else.

6.2.2. Aristotle. Interesting statements occur also in his works[92]:

(1) *Why is it that, though the diseases due to bile occur in the summer, … acute diseases due to bile occur rather in the winter*?

(2) *Generally speaking, the change which occurs when a warm, dry summer follows … on a wet spring, being violent has a deleterious effect upon the body.*

(3) *Why is it that deaths are particularly likely to occur during the hundred days following each solstice*?

He discusses meteorological phenomena, which, according to his opinion, accompany solstices and are responsible for the increased mortality.

(4) *Why is it that fair men and white horses usually have grey eyes*?

Lacking, however, are statements about chance differences between patients in their response to medical treatment.

6.2.3. GALEN. Developing HIPPOCRATES' doctrine, GALEN devotes a special chapter of his *Hygiene*[93] to explain that different men require different ways of life and medical treatment. He also uses stochastic reasoning such as[94]

For the person with perfect constitution of body, who both chooses a free life and in it never goes to any excess … is not very likely to fall into any very pathological conditions.

Chance repeatedly enters into his considerations, sometimes explicitly and sometimes almost so: the body, says he[95],

Has two sources of deterioration, one intrinsic and spontaneous, the other extrinsic and accidental. Of those things which affect it from without (some) *contacts are occasional, irregular, and not inevitable.*

Also, GALEN recognizes chance in medicine as understood by POINCARÉ (see § 2.2)[96]:

In those who are healthy the body does not alter even from extreme causes; but in old men even the smallest causes produce the greatest change.

At the same time GALEN firmly believed in the divine design of man. His utterances to this effect are numerous[97] and in one instance[98] he attributes chance explanations to sophists (see § 1 for a similar point of view of ARISTOTLE).

GALEN did vaguely suppose that nature is prone to changes[99]:

Nature is a constructive artist and … the substance of things is always tending towards unity and also towards alteration because its own parts act upon and are acted upon by one another.

However, he hardly thought these changes to be random[100]:

If the universe is not originated, it is in no danger of decay, nor is it open to chance happenings and disorders … Whoso believes … that the world is originated … comes to blasphemy.

Thus, chance occurs in biology, but not in the world in general so that GALEN'S philosophical outlook seems to be hardly consistent.

The idea of means, as understood in moral philosophy by ARISTOTLE, was described in § 2.2. GALEN also uses the same idea, although in a medical context[101]:

(1) *A good constitution* (is) *a mean between extremes.*

(2) *If the exact mean of all the extremes were in all parts of the body, this would be the best to observe as being the symmetry most suitable for all activities*.

Possibly his statements are not altogether correct but at least they show how the idea of means (in this case, the mean seems to be the semirange) came to be advocated in science. See also § 5 where another aspect of the same idea is described.

To my understanding, GALEN says the same thing once more, this time adding that chance deviations from the mean condition should be small[102]:

Health is a sort of harmony … all harmony is accomplished and manifested in a two-fold fashion, first in coming to perfection and second in deviating slightly from this absolute perfection.

GALEN'S writing *On medical experience* (note 100) is devoted to the discussion of empiricism versus dogmatism in medicine and his general conclusion (§ 31, p. 153) is that

Empiricism suffices to discover everything used in healing.

Possibly this discussion is not as interesting as a discussion of induction versus deduction would have been, but GALEN (§§ 16 – 18) also makes remarkable comments on the paradox of the heap, the solution of which (§ 18, p. 121) *it is idle to demand*. The paradox, as seen in a medical context, or, rather, in that of medical statistics, occurs because of the fact that response to medical treatment is random (§ 15, p. 112):

Experience has shown that what has produced a like result in three cases can produce the reverse in three others … a thing seen may be seen exactly as before, and yet belong to those things which are of both kinds (*amfidoxos*) *or to those things which happen often, or to those … which take place but rarely.*

Therefore (p. 113),

What is to prevent the medicine which is being tested from having a given effect on two (three ?) *hundred people and the reverse effect on twenty others, and that of the first six people who were seen at first and on whom the remedy took effect, three belong to the three hundred and three to the twenty without your being able to know which three belong to the three hundred, and which to the twenty, even if you were a soothsayer you must needs wait until you see the seventh and the eighth, or, to put it shortly, very many people in succession.*

Understanding that the response to medical treatment is in a sense random, GALEN strives for discretion and, it seems, is not prepared to entrust himself to the relative frequencies of the both possible responses. However, this meant that he did not give any positive recommendation about the quantitative assessment of data. But at least the formulation of the classical problem of medical statistics (that of testing a given medicine), is due to him as is also the remark that statistics has to do with the quantitative assessment of the "heaping" of data. And precisely these philosophical aspects of statistics justify assertions such as[103]

Statistics has long had a neighbourly relation with philosophy of science in the epistemological city, although statistics has usually been more modest in scope and more pragmatic in outlook. In a strict sense, statistics is part of philosophy of science, but in fact the two areas are usually studied separately.

I have described writings of HIPPOCRATES and GALEN. IBN SINA, the third great man in the history of medicine, has no new ideas to offer but at least he confirms HIPPOCRATES' opinion concerning both stochastic considerations and significance of symptoms[104]. See also § 3.2 for LEIBNIZ' discussion of the *Symptome und Indikationen* in medicine.

7. Astrology

From a modern point of view astrology is nothing but a pseudoscience. There were, however, astrologers, scholars of the highest calibre included, who strove to discover connections between heaven and earth, sincerely believing them to exist.

That their delusion was not so evident as it could now seem to be is testified by the fact that heaven does, after all, influence earth (e. g., ocean tides are occasioned by the sun and the moon).

I describe KEPLER'S opinion on astrology. He was the man whose attempt, if unsuccessful, to base astrology on new methodological principles should be specifically mentioned. Indeed, his astrological activities had been noticed before[105]:

Er wollte einen Unterschied machen zwischen dem chaldäischen, sternguckerischen Aberglauben und der <u>*Physik*</u>*, d. h. der auf Erfahrung begründeten reinen Wissenschaft, die nach seiner Überzeugung einen gewissen Zusammenhang zwischen den Himmelerscheinungen und dem irdischen Geschehen bestätige*.

And, in more detail[106],

Dass er sich schon damals mit (astrology) *beschäftigte, hat seinen Grund nicht nur darin, dass er als Landschaftsmathematiker ... den jährlichen Kalender anfertigen musste … oder gar darin, dass er sich mit dem astrologischen Handwerk nur einen erwünschten Nebenverdienst habe verschaffen wollen, ohne selbst an das zu glauben, was er sagte. Er hat oft genug durch die Tat bewiesen, dass er um der Wahrheit willen ideale und materielle Opfer bringen konnte, und er hat seine Überzeugung nie um äußere Vorteile verkauft. Und wenn auch das närrische Töchterlein Astrologie bisweilen der Mutter Astronomie Mittel verschaffen musste, so ließ er doch das Töchterlein alles nur so vorbringen, dass er es seiner Überzeugung nach vertreten konnte. Der Glaube an eine Beeinflussung des irdischen Geschehens durch die Erscheinungen am … Himmel war vielmehr in seiner Zeit so allgemein verbreitet und wurde von so vielen Männer, die er hochachtete, vertreten, dass er auch ihn anstecken musste. Dass er aber auch hier schon von Anfang an mit kritischem Verstand vorging beweisen die Ausführungen, die sich in den … Briefen finden. …*

Dass eine Einwirkung des Himmels auf die Wettererscheinungen, aber auch auf die menschliche Seele erfolgt, das ist ihm eine Tatsache der Erfahrung.

KEPLER'S original intention was to become a theologian[107]:

Ich wollte Theologe werden; lange war ich in Unruhe. Nun aber sehet, wie Gott durch mein Bemühen auch in der Astronomie gefeiert wird.

It is to be doubted however, that he would have been successful as a theologian. In his own words[108]

Auf die Meinungen der Heiligen über diese natürlichen Dinge antworte ich mit einzigen Wort: In der Theologie gilt das Gewicht der Autoritäten, in der Philosophie aber das der Vernunftgründe.

It clearly seems that his sympathies lie with the *Vernunftgründe* and, in any case, his main occupation happened to be the *hochvernünftige* astronomy and her daughter, *the närrisches astrology*[109]:

Es ist wol diese Astrologia ein närrisches Töchterlin aber lieber Gott, wo wolt ihr Mutter die hochvernünftige Astronomia bleiben, wann sie diese ihre närrische Tochter nit hette, ist doch die Welt noch viel närrischer, und so närrisch, dass demselben zu ihren selbst frommen diese alte verständige Mutter die Astronomia durch der Tochter Narrentaydung, weil sie zumal auch einen Spiegel hat, nur eyngeschwatzt und eyngelogen werden muss.

Uns seind sonsten der Mathematicorum salaria so seltzam und so gering, dass die Mutter gewisslich Hunger leyden müsste, wann die Tochter nichts erwürbe.

KEPLER never renounced his moral principles. Therefore, somewhat developing CASPAR'S opinion as stated above, I hold, that, according to KEPLER, it is possible, and maybe even reasonable, for an applied science (astrology) to keep its theoretical counterpart (astronomy). And KEPLER did suppose that astrology is a science. Moreover, he considered himself to be the actual founder of astrology as a science[110]:

Ich habe nach gegen (H. RÖSLIN and P. FESELIUS) *mit zwei … Schriften gestellt. … Daraufhin erhielt ich Briefe von Gelehrten, worin sie bezeugten, dass die Astrologen jetzt erst durch mich in eine reinere Lehre eingeführt würden.*

According to KEPLER'S astrology, heaven does influence earth. However, and this is his main point, this influence is a tendency rather than a fatal drive. That this point of view was not due to KEPLER alone is proved by TYCHO BRAHE'S preface to a book published in 1591 by one of his students and recently reprinted with a commentary by J. CHRISTIANSON[111]. TYCHO is there quoted as supposing that *human beings are influenced by heaven in a lesser degree than animals.*

In another source[112] TYCHO is said to have

Criticized astrologers who drew improper conclusions based on superstition and error rather than astrology itself, which he considered a science for which both accurate knowledge of the course of the stars and experience gained from signs seen in the elementary world were needed.

It thus seems reasonable that TYCHO hardly thought that heaven influenced earth directly, fatally. Also, it is quite possible that many scholars held a similar point of view (see § 2.4 for a relevant passage from THOMAS AQUINAS) but at least KEPLER did have to defend this point of view, a fact witnessed by his *Tertius interveniens*[113] devoted to this end. A second relevant and, for that matter, more important argument in favour of KEPLER is that he applied, or attempted to apply, the principle of correlative influence to explain not only men’s dispositions (and, sometimes, fates) but also phenomena pertaining to natural science; see below.

The tendency of the heavenly influence is explained by him thus[114]:

Meine Gestirne waren … nicht der morgendliche Merkur … sondern Kopernikus und Tycho Brahe, ohne dessen Beobachtungsjournale alles, was ich bis heute in helles Licht gerückt habe, in Finsternis begraben läge. …

Die einzige Wirkung der Geburtskonstellation bestand darin, dass sie jene Flämmchen der angeborenen Anlage und der Urteilskraft geschnenzt und den Wissendurst vermehrt hat; kurz, sie hat den Geist und die genannten Seelenvermögen nicht inspiriert, sondern nur geweckt.

No wonder that[115]

(1) *Die Astrologi können futura contingentia nicht vorsagen.*

(2) *Wahr ist es von dem großen Theil, aber nicht von allem, was die Astrologi fürgeben: wahr ist es von den individuis, aber nicht von der generalitet, die in alle individua eyngetheilt ist.*

(3) *So cörperlich und so greifflich gehet es nicht zu, dass Himmel und Erde einander anrühreten, wie die Räder in einer Uhr.*

(4) *Die Handlungen seyend … nicht mehr influentia coelestis, sondern actio naturae, in quam coelum influxit.*

Ordinary men expected astrologers to foretell important events in their lives. Having neither the possibility of accomplishing this request nor the desire to disappoint his readers, much less to deceive them, KEPLER had to discontinue compilation of astrological almanacs[116]:

Die Astrologi keine besondere Spraach haben, sondern die Wort bey dem gemeinen Mann entlehnen müssen, so wil der gemeine Mann sie nicht anderst verstehen, dann wie er gewohnet, weiss nichts von den abstractionibus generalium, siehet nur auff die concreta, Lebt offt einer Calender in einem Zutreffenden Fall, auf welchen der author nie gedacht … ich endtlich hab auffhören Calender zu schreiben.

His general conclusion is[117]:

Ein Astrologus, der nur den Himmel sihet und von … zwischenursachen nicht weiss, nur allein probabiliter, nit Messungsweiss, das ist, ein klein wenig mehr dann nichts, von dem letzten Erfolg vorsagen könne.

It is my understanding that *probabiliter* is a lame substitute for something like *almost at a guess*. These thoughts naturally led KEPLER to comparisons between astrology and medicine, another science based on the probable (§ 6)[118]:

(1) *Astrologia in einem bessern Verstande ist mit ihrer experientia so gewiss als Medicina Botanica, und muss man sich bey der orten der Aberglauben erwehren.*

(2) *Nit allein die Astrologi, sondern auch Medici bissweilen krumme Wege gehen müssen, zu einem guten Intent zu gelangen.*

The *krumme Wege* of physicians are mentioned: dissection of (snatched) bodies (criminal on both counts) and recommendation of the use of contraceptives to avoid venereal diseases. The *krumme Wege* of astrologers are not mentioned but possibly KEPLER thought that astrology infringed on theological customs and by-laws (see also below).

(3) *Seynd die Medicinalische illationes nicht alle so gewiss als ein vorsagung der Finsternuss, und bleibt demnach, dass in der Astrologia auch wol etliche illationes aus der Erfahrenheit geschehen können, welche gleich so gewiss, als wann ein Medicus einer Person, die etwan gehling ihre Gedächtnuss und Sinne verlohren, und doch baldt wider gesundt worden, angedeutet, sie wisse nun welches Todts sie sterben werde.*

(4) *Wann ein Medicus seiner Patienten Geneses und Crises so fleissig auffgezeichnet hette, so fleissig ich diese 16 jahr das Wetter auffgezeichnet habe. … Er müste aber mit seiner Experientz fürsichtig handeln und sich keinen Patienten mit falschem Bericht betriegen lassen*.

It was the established practice of astrologers to single out several *aspects*, i. e. remarkable mutual positions of the sun, the moon and the planets and at least a few commentaries on the connections between aspects[119] and meteorological phenomena on earth are to be found in KEFLER[120]:

(1) *Ich habe … bemerkt, dass mit großer Regelmäßigkeit der Zustand der Luft gestört wird, so oft Planeten entweder in Konjunktion treten oder nach der herkömmlichen Lehre der Astrologen Aspekte bilden. Anderseits habe ich bemerkt, dass meistens Ruhe herrscht, wenn keine oder nur wenige Aspekte einfallen oder wenn sie sich rasch vollziehen und vorübergehen*.

(2) *Ist es auch möglich, dass es in der heißen Zone zur Zeit der Aspekte mehr regnet als an Tagen, die frei von Aspekten sind.*

Bibliographie Kepleriana contains no traces of KEPLER'S meteorological activities. KEPLER did describe them in his correspondence[121] but in his time meteorology did not yet concern itself with numerical data and for this reason, if for no other, the correspondence contains no quantitative correlation between the aspects and meteorological phenomena on earth.

I have already quoted KEPLER'S opinion concerning the *krumme Wege* of astrologers. That he possibly meant infringement on theology is indirectly illustrated by a passage which ends on a cautious note resembling that of J. BERNOULLI[122]:

Wenn ein vom Dach herabfallender Ziegel einen Vorübergehenden trifft, … wenn die Geburtskonstellation für solche Ereignisse, die dem Aufgabenbereich des Schützengels angehören, Anzeichen enthält, so müssen sich aus dieser Konstellation für diese Betreuung hindernde oder umgekehrt fördernde Wirkungen ergeben. Ob diese Meinung nicht der Gottesverehrung widerspricht, das mögen die Theologen entscheiden.

Another extremely interesting feature of KEPLER'S astrology is his presumption of the divine care for mankind, one of the corollaries of which is[123] that illegitimate children, whose birth just as well follows the divine will, are endowed with reason no less than their legitimate counterparts and have the same right to live (so as, presumably, to foster the multiplication of mankind). If one bears in mind that astrology had also to do with the general destiny of nations as decided by the prevailing aspects, and, also, by geographical conditions, etc. (see KEPLER'S prudent qualification about the *Zwischenursachen*) and that KEPLER, for one, contributed quite a few astrological almanacs, it follows that both in his presumptions and aims KEPLER the astrologer resembles founders of political arithmetic (J. GRAUNT, W. PETTY). Howevér, to call KEPLER their precursor would be far-fetched: completely lacking in his works are statistical data on population, which were to form the basis of political arithmetic. But hardly anyone could suppose that such data had existed in feudal and subdivided Germany.

8. Astronomy

8.1. KEPLER. Possibly the first stochastic reasoning concerning astronomy is due to ARISTOTLE (§ 5) and it is interesting to notice that, beginning at least with KEPLER (see § 5 and below) and NEWTON[124], similar arguments about the impossibility of chance's governing the system of the world (or heaven in general) were time and again pronounced by various European scholars[125].

KEPLER, who devoted all his life to the discovery of general laws of nature, bitterly disclaimed randomness[126]:

Mais qu'est-ce que le hasard? Pas autre chose qu'une idole, et la plus détestable des idoles; pas autre chose que le mépris du Dieu souverain et tout puissant, ainsi que du monde très parfait sorti de ses mains.

In a letter to HERWART dated Jan. 13, 1606 KEPLER pointed out the origin of his philosophical outlook[127]:

Da ich es hierbei mit den Philosophen zu tun haben werde betreffs der Begriffe Schicksal, Bestimmung, Zufall, wenn ich das wunderbare Zusammentreffen einer großen Konjunktion mit dem neuen Stern behandle (see also § 5), *habe ich mich an die Lektüre von Augustinus' Werk über den Gottesstaat gemacht*.

AUGUSTINUS[128] has little to say about chance, so it seems that KEPLER'S reasoning above was rather occasioned by ancient thinkers in general (and, of course, by his own astronomical discoveries).

8.1.1. Eccentricies. KEPLER did have to find room for the work of random causes, first and foremost in connection with the eccentricities of the planetary orbits. By this term, even after having discovered his first law of planetary motion, KEPLER meant the eccentric position of the sun as measured from the centre of a circular orbit[129].

Eccentricities caused much trouble to KEPLER even in his juvenile work[130]. Attempting to explain the general construction of the system of the world, KEPLER inserted the five regular solids between the spheres of the six then known planets[131], but the existence of eccentricities, and for that matter unequal to one another, much worried him[132]:

Die Ursache der Excentrizitäten wie auch ihrer Unterschiede noch nicht erforscht ist.

On p. 108, Kap. 17, he formulates the corresponding problem for those who *Lust dazu haben*:

Die Ursachen der … Excentrizitäten aus den entsprechenden Körpern (the regular solids) *ableiten. Da nämlich auch diese Abweichungen nicht aufs Geratewohl und ohne Grund gerade in* *dieser Größe von Gott den einzelnen Planeten zugemessen worden sind.*

In the second edition of his *Mysterium* Kepler added *Anmerkungen* to almost each chapter and among them are[133]

(1) *Man konnte noch nicht die Ursache der Excentrizitäten, man wusste nicht, warum die Excentrizität bei den einzelnen Planeten gerade so groß ist …*

(2) *Ich habe die Größe der Excentrizitäten … erforscht, ich habe in der Harmonik die Ursache der Excentrizitäten aufgedeckt.*

(3) *Ich habe gesucht, und, siehe da, ich fand die vorzüglichsten Ursachen* (a reference to book 5 of the *Harmonik* is given).

What then, occurred before 1621, the date of the publication of the second edition of the *Mysterium*? In his main work[134] KEPLER first ascribes the eccentricities to random ifluences from without:

Die Beispiele aus der Natur und die … Verwandtschaft zwischen den himmlischen und irdischen Erscheinungen bezeugen laut, dass die Wirkungen eines einfachen Körpers um so einfacher sind je allgemeiner sie sind, und dass Verschiedenheiten … (*wie … die Exzentrizität*) *von Ursachen herrühren, die von außen her hinzutreten.*

Thus, because of external hindrances, says KEPLER, rivers cannot flow directly to the centre of the Earth.

Das Abwärtsgleiten, die Stagnation, … das alles rührt von den angegebenen Ursachen her, die zufällig von außen her hinzukommen.

On p. 244 (Kap. 39) his similar reasoning is that the planets, being situated at enormous distances from the sun, just could not accurately trace their prescribed circular paths. But then (p. 268, Kap. 45) he refutes himself:

Denn nachdem ich im Kap. 39 *in größter Verlegenheit war, weil ich keine hinreichend wahrscheinliche Ursache dafür anführen konnte, wie die Planetenbahn zu einem vollkommenen Kreis wird* (*immer musste ich der Kraft im Planetenkörper eine absurde Eigenschaft beilegen*)*, nun aber auf Grund der Beobachtungen entdeckt hatte, dass die Planetenbahn gar nicht vollkommen kreisförmig ist.*

However, and this is the important thing, KEPLER repeatedly returns to the principle of outside influences. Thus, in 1616 he writes[135]:

Beweise ich, dass die Ungleichförmigkeit der Bewegung der Natur der Planetenkugeln entspricht, also physikalisch ist. Außerdem beweise ich, dass und in der sublunarischen Natur und in den mechanischen Bewegungen Beispiele für eine solche regelmäßige Ungleichförmigkeit der Himmelbewegungen zu Verfügung stehen, also wiederum dass diese regelmäßige Ungleichförmigkeit physikalisch ist.

A few years pass and KEPLER explicates his thoughts in more detail[136]:

If the celestial movements were the work of mind, as the ancients believed, then the conclusion that the routes of the planets are perfectly circular would be plausible.

But the celestial movements are … the work of … nature and this is not proved by anything more validly than by the observation of the astronomers, who … find, that the elliptical figure of revolution is left *in the real and very true movement of the planet, and the ellipse bears witness to the natural bodily power and to the emanation and magnitude of its form.*

.. *Because in addition to mind there was then need of natural and animal faculties also for the sake of movement; those faculties followed their own bent …* (and) *did many things from material necessity. So it is not surprising if those faculties, which are mingled together, could not attain perfection completely. The ancients themselves admit that the routes of the planets are eccentric, which seems to be a much greater deformity than the ellipse.*

KEPLER'S general reference to the ancients seems inconclusive and I feel it opportune to refer additionally to EPICURUS (§ 2.3) with his similar thoughts about outside influences. Similar reasoning, in a somewhat clumsy style which smacks of mysticism, may be found also in book 5 of his *Harmonices Mundi*[137]. The title of Kap. 9 of this book is

Dass die Exzentrizitäten bei den einzelnen Planeten ihren Ursprung in der Vorsorge für die Harmonien zwischen ihren Bewegungen haben and in this chapter (p. 317) one finds:

Der ... himmlische Werkmeister höchstselber die harmonischen Proportionen … mit den fünf räumlichen regulären Figuren verbunden hat, um aus den beiden Figurenklassen ein einziges vollkommenstes Urbild des Himmels zu formen. … Die Maße der Exzentrizitäten der einzelnen Bahnen zum Zweck einer entsprechenden Regelung der Körperbewegungen enthalten waren.

Actually here (and explicitly on p. 319, V. Satz) KEPLER is referring to his second law of planetary motions: the measures of eccentricities (and here eccentricities could be meant only as pertaining to elliptical orbits) are predetermined so as to proportion the movements of the planets. It is precisely this that KEPLER means by saying (p. 316) that *Es können also … die Gesamtharmonien aller sechs Planeten nicht von ungefähr auftreten.*

It remains uncertain why KEPLER failed to refer to his second law in the *Epitome*. But in any case, it is my opinion that KEPLER never rejected the idea that the circular motions of the planets are predetermined while the elliptical motions are occasioned by a relatively small corruption due to *the natural and animal faculties* etc.

Moreover, the values of eccentricities of these motions are also predetermined to proportion the movements of different planets, but this, so to speak, is a predetermination of a second order. Dismissing these faculties, one discovers that KEPLER, for all his negation of randomness, ascribed the elliptical deviation from the circle to random influences.

I do not know just how much KANT and LAPLACE borrowed from KEPLER but at least it was their point of view also that the irregularities in the system of the world were occasioned by variety (randomness)[138]:

(1) *Da es denn ein gar zu glückliches Ungefähr sein würde, wenn gerade alle Planeten ganz genau in der Mitte zwischen diesen zwei Seiten in der Fläche der Beziehung selber sich zu bilden anfangen*

sollten … überhaupt die Vielheit der Umstände, die an jeglicher Naturbeschaffenheit Antheil nehmen, eine abgemessene Regelmäßigkeit nicht verstattet.

(2) *Woher sind ihre Umläufe nicht vollkommen zirkelrund. … Ist es nicht klar einzusehen, dass diejenige Ursache, welche die Laufbahnen der Himmelskörper gestellt hat, es nicht völlig hat ausrichten können. Ist nicht das gewöhnliche Verfahren der Natur hieran zu erkennen, welches durch die Dazwischenkunst der verschiedenen Mitwirkungen allemal von der ganz abgemessenen Bestimmung abweichend gemacht wird*?

(3) *Si le système solaire s'était formé avec une parfaite régularité, les orbite des corps … seraient des cercles, dont les plans … coincideraient avec le plan de l'équateur solaire. Mais on conçoit que les variétés sans nombre qui ont dû exister dans la température et la densité des diverses parties de ces grandes masses ont produit les excentricités de leurs orbites, et les déviations de leurs mouvements du plan de cet équateur*.

Thus, KEPLER's *natural und animal faculties* are soberly called *variétés dans la température et la densité*!

8.1.2. End of the World. Another important subject in KEPLER's writings is his speculation *über den astronomischen Anfang und das astronomische Ende der Welt*[139].

In the first edition of the *Mysterium*, assuming an incorrect form of the as yet unknown third law of planetary motions, KEPLER rejects any possibility of a simultaneous return of all the planets to their *Anfangslage*, their position at the moment of creation (Kap. 23, p. 144):

Die Exzentrizität stehe in rationalem Verhältnis zum Bahnhalbmesser, dann werden diese untereinander irrational, da sie sich verhalten wie die Radien der In- und Umkugeln der Körper (the regular solids), *die in irrationalem Verhältnis zueinander stehen … Nun aber stehen die Bewegungen zu den Radien in rationalem Verhältnis, also sind die Bewegungen unter sich irrational, und kehren daher nie wieder zur Anfangslage zurück, auch wenn sie unendlich viele Jahrhunderte dauern würden.*

According to KEPLER, this means that the end of the world is impossible. He refers to this argument in a letter to D. FABRICIUS[140]:

In der Tat entsteht auch jeder einzelne Aspekt in den Ephemeriden aus seiner eigenen Ursache; sie treten aber in keinerlei Ordnung auf; auch wenn die Welt 100,000 Jahre dauerte, so würde nie dieselbe Reihenfolge wiederkehren, wie in meinem Mysterium (*Kap. 23*) *steht.*

Then, in the second edition of *Mysterium*, bearing in mind (the correct form of) the third law of planetary motions, KEPLER adds:

Nun aber haben wir bereits dieses Fundament umgestoßen, insofern die Verhältnisse der Himmelsbahnen nicht allein von dem fünf Körpern herrühren. Es fragt sich also, was nun von dem vorliegenden Satz zu halten ist. Gibt es eine vollständige Wiederkehr aller Bewegungen? Ich sage nein, obwohl jener Beweisgrund umgestoßen ist.

His explanation, which follows on p. 146 with no reference being given, is the same as the one in N. ORESME[142]:

It is probable (*verisimile*) *that two proposed unknown ratios are incommensurable because if many unknown ratios are proposed it is most probable that any* (one) *would be incommensurable to any* (other).

Knowing nothing about the (KEPLERIAN) laws of planetary motions, ORESME[143] expressed an opinion much like KEPLER'S:

It is probable that in any instant the celestial bodies are related in such a way that they were never so related in the past, nor will so be related at any time in the future.

The arguments of ORESME and KEPLER are interesting as being the first, if naive, ones in which stochastic considerations are applied to an abstract mathematical notion.

Of course, neither of them knew that a dynamical system will eventually return arbitrarily close to its original state.

8.2. GALILEI. Denying any possibility that a heavenly body can move irregularly, GALILEI, as it seems, simultaneously denies randomness[144]:

Those lines are called regular which, having a fixed and definite description, have been susceptible of definition and of having their qualities and properties demonstrated. … But irregular lines are those which have no determinacy whatever and are indefinite and casual, and hence indefinable; … to say, "Such events take place by reason of an irregular line" is the same as saying "I do not know why they occur." The introduction of such lines is in no way superior to the sympathy, antipathy, occult properties, influences, and other terms employed by some philosophers as a cloak for the correct reply, which would be "I do not know".

If this was meant to be directed against some of KEPLER'S utterances it will possibly explain, even if to a small extent, the puzzle which GALILEI'S failure to recognize the KEPLERIAN laws of planetary motions still poses.

However, as also was the case with KEPLER, and, for that matter, with any natural scientist worth his salt, GALILEI'S denial of randomness did not prevent him from separating regularity and randomness in an observed natural phenomenon.

Exactly this was his main problem when he studied the behaviour of solar spots[145]. Their rotation with the sun itself was the regular component and their movement relative to the sun's disk was the random component of the observed phenomenon of their (general) movement. GALILEI arrived at a rather accurate estimate of the period of rotation of the sun (one lunar month; the modern estimate is 24.5 – 26.5 days).

9. Modern Philosophy: Chance and Its Laws

9.1. Chance. Many eminent modern philosophers just did not recognize chance, or, to put it more cautiously, reduced chance to intersections of determinate chains of events. Thus, it was T. HOBBES' opinion[146] that

(1) *Generally all contingents have their necessary causes … but are called contingent in respect of other events upon which they do not depend; as the rain … shall be … from necessary causes, but we think*

… it happens by chance, because we do not yet perceive the causes thereof.

(2) *By contingent, men … mean that which hath not for cause anything that we perceive; … when a traveller meets with a shower, the journey had a cause, and the rain had a cause sufficient to produce it; but because the journey caused not the rain, nor the rain the journey, we say they were contingent one to another.*

Much the same was the opinion of B. SPINOZA[147], G. W. LEIBNIZ[148], VOLTAIRE[149], C. A. HELVETIUS[150], P. H. T. D'H0LBACH[151] and D. HUME[152]. As to LEIBNIZ, he[153] adds that the *zufällig* has

Ihre Ursprung von einer überwiegenden Ursache her, die zwar neiget, aber nicht zwinget.

See §§ 2.4 and 7 for a similar kind of influence of celestial bodies on man as understood by THOMAS AQUINAS and KEPLER.

The point of view of POINCARÉ that randomness has to do with phenomena in which slight causes have considerable effect (§ 2.2) had also been pronounced by various earlier scholars, at least as applied to history. This was the opinion of HELVETIUS, B. PASCAL, possibly laughing in his sleeve, HOLBACH, and, to a certain extent, VOLTA1RE[154] of whom I shall quote only PASCAL:

Le nez de Cléopatre s'il eût été plus court toute la face de la terre aurait changé.

Modern philosophers used stochastic arguments either to explain the origin of the world and/or of animals and man or to refute such explanations. All these explanations (and refutations) were, of course, qualitative, and the theory of probability did not come into being because of them. However, a short account of them is justified at least by the interesting concomitant considerations.

Those against a chance origin of the world usually cited as examples of similar impossible feats a chance composition of a lengthy book[155]. Those against the origin of life by chance either took their opinion for granted[156] or thought it to be even less possible than that of the world[157]. On the other hand, perhaps HELVETIUS[158] did imagine that the world originated by chance:

La nature, par ses combinaisons, enfante des soleils

while R. DESCARTES[159], though with a qualifying remark

Il est bien plus vraisemblable que, des le commencement, Dieu l'a rendu tel qu'il devoit être,

attempted to explain the origin of the world by chance[160]:

(1) *Quelque inégalité et confusion que nous puissons supposer que Dieu ait mise au commencement entre les parties de la Matière, il faut, suivant les loix qu'il a imposées a la Nature, que par après elles se soient reduites presque toutes à une grosseur et à un mouvement mediocre.*

(2) *Je me résolus de … parler … de ce* [monde] *qui arriveroit dans un nouveau, si Dieu créoit maintenant quelque part, … assez de matière pour le composer, et qu'il agitât diversement et sans ordre les diverses parties de cette matière, en sorte qu'il en composât un chaos aussi confus que le poètes en puissent feindre, et que par après il ne fit*

autre chose que prêter son concours ordinaire à la nature, et à laisser agir suivant les lois qu'il a établies.
… Je montrai comment la plus grande part de la matière de ce chaos devoit, en suite de ces lois, se disposer et s'arranger d'une certaine façon … comment cependant quelques-unes de ses parties devoient composer une terre, et quelques-unes des planètes.

Against this general background, it is all the more interesting to notice passages, which, it seems, show that at least some scholars were prepared to accept randomness in a more general sense than that confined to blind chance (= to the uniform distribution)[161]:

(1) *Seroit-on bien étonné, s'il avoit dans un cornet cent mille dés, d'en voir sortir cent mille six de suite? si ces dés étoient tous pipés on cesseroit d'en être surpris. … Les molecules de la inatière peuvent être comparées à des dés pipés … ces molécules étant essentiellenent variees par elles-mêmes et par leurs combinaisons, elles sont pipées, pour ainsi dire, d'une infinité de façons différentes.*

(2) *En conséquence de cette sensibilité sourde et de la différence des configurations, il n'y aurait eu pour une molécule organique quelconque qu'une situation la plus commode de toutes, qu'elle aurait sans cesse cherchée par une inquiétude automate.*

9.2. Randomness and Necessity; Dialectics. An argument about randomness and necessity, possibly the first one in recent times, is due to KANT[162]:

Der Zufall im Einzelnen nichts desto weniger einer Regel im Ganzen unterworfen ist.

G. W. F. HEGEL[163] developed this idea:

Diese Einheit der Möglichkeit und Wirklichkeit ist die Zufälligkeit. [Das Zufällige] *ist … unmittelbare Wirklichkeit; es hat keinen Grund … Das Zufällige ist aber … das Wirkliche als ein nur Mögliches, … es hat einen Grund.*

… Die Einheit der Notwendigkeit und Zufälligkeit … ist die absolute Wirklichkeit zu nennen.

HEGEL's style is ponderous and the passages quoted above remained virtually unnoticed at least until the development of stochastic ideas in biology took place. F. ENGELS[164] stressed the importance of a dialectical understanding of randomness and necessity for natural science as well as the actual if unconscious use of it by DARWIN.

9.3. An Attempt of Formalizing the Concept of Disorder. I have already described J. H. LAMBERT'S work on probability including his lone attempt to formalize the concept of randomness and the connection of randomness with disorder, i. e. to solve a problem which has been attacked once more only in our time[165].

LAMBERT posed much the same problem in a memoir[166] which I did not notice before. Referring to C. WOLFF, who, as is well known, having started from LEIBNIZ' general ideas, attempted to develop a comprehensive system of knowledge, and, also, to his student, A. G. BAUMGARTEN, LAMBERT applies mathematics to *metaphysical* objects. Considering (§ 7 of pt. 1) the development of $\sqrt{12}$, LAMBERT notices *dans ces nombres un ordre de liaison* (*ordre* *légal*, as he also calls it) which means that each digit *occupe nécessairement sa place*. However,

Il est également vrai aussi, qu'il n'y a absolument point d'ordre de ressemblance (*ordre local*), *et qu'ils se succedent comme jettés au hazard. … Aussi le calcul des probabilités y est parfaitmnent applicable*.

Then LAMBERT estimates disorder in number series. Considering permutations of elements (§ 11) he introduces a measure of disorder equal to the sum of products of the value of each element by its distance from its proper place. Thus, the disorder of a series 4, 3, 1, 2 is equal to

$$4\ 3+3\ 1+1\ 2+2\ 2 = 21.$$

Passing to the multidimensional case, LAMBERT (§ 13) considers the order in the arrangement of n books of a given library:
Les livres s'y classifient d'abord suivant les sciences; ensuite on a égard a leur ancienneté, au format, ...

Assuming that these aspects of classification possess weights a, b, c, LAMBERT (§ 15) estimates the *order* of the library: if the arrangement of m books complies with each condition (aspect), while that of p, q, r books ($m + p + q + r = n$) complies with conditions (1) and (2), (1) and (3), and only (1) correspondingly, the *order* will be

$$\frac{m(a+b+c)+p(a+b)+q(a+c)+ra}{n(a+b+c)}.$$

A curious utterance is contained in § 16:

Mais si à cet égard on repasse la plupart des Institutions de Chymie, où y trouvera un ordre d'un degré bien inférieur; et quand il s'agit des écrits où l'ordre est = 0, *c'est aux alchymistes qu'il faut s'adresser*.

The first part of the memoir ends with the calculation of the optimal place of an element in a given series. If, according to three contradictory rules, this element is displaced by x, $m – x$ and $n – x$ respectively, and if these rules possess weights A, B, and C, the *degré de défaut d'ordre* will be

$$y = Ax + B(m – x) + C(n – x).$$

The minimal y is sought in three different cases, i.e. when $A + B$ is more or less than, or equal to C.

The second part of the memoir is an attempt to formulate rules for arranging elements of a given series. Beginning with preliminary considerations, LAMBERT (§§ 8 – 18) studies one-dimensional series; § 23 is devoted to two-dimensional series, while in § 24 he formulates numerous rules which, on his view, should be followed. Each rule, says LAMBERT, has to do only with *l'ordre de ressemblance*, to which, and only to which, stochastic considerations are applicable (see above).

One-dimensional series, it is said in rule No. 1, should be symmetric relative to the central symbol, which, moreover, should not occur anywhere else. (LAMBERT uses symbols rather than numbers or digits.)Then (rules No. 4 and 5),

Les objets qui different en espèce doivent encore différer en nombre.

… *Chaque variété doit être rachetée par quelque ressemblance, et réciproquement chaque ressemblance doit être contrebalancée par quelque variété.*

Thus, LAMBERT the mathematician finally gave way to LAMBERT the artist: his rules obviously have nothing to do with any mathematical understanding of randomness. Moreover, stochastic considerations are mentioned only in passing. Nevertheless, this memoir should be mentioned in the general context of LAMBERT'S heroic attempt to quantify disorder and randomness.

9.4. Theory of Probability and Mathematical Statistics. The direct source of inspiration for J. BERNOULLI'S *Ars Conjectandi*[167] is the celebrated *Port-Royal* where many of BERNOULLI'S initial assumptions and even some of his examples are present. Nevertheless, the idea of quantifying inductive inference and its translation into mathematical formulae are due to BERNOULLI.

What is called the classical probability of an event,

$$p = \frac{m}{n} \qquad (1)$$

where m is the number of favourable cases and n is the total number of all (equally possible) cases, was known even to CARDANO and to KEPLER (§ 8.2). Simultaneously with (1) a second, if tacit, statistical definition

$$p = \frac{\mu}{\nu} \qquad (2)$$

where μ is the number of occurrences of a certain event in ν independent trials with a constant probability (!) of success came to be used.

In distinction from R. CARNAP[163], who stresses the difference between (1) and (2), I notice that precisely the combination of these definitions of probability beginning with the *A. C.* formed the basis for the development of the theory of probability, possibly until the middle of the 19th century.

Consider, indeed, BERNOULLI'S law of large numbers: if the probability p of the occurrence of an event in each trial is constant and in ν trials this event occurs μ times, it follows that

$$\lim P(|\frac{\mu}{\nu} - p| < \varepsilon) = 1, \ \nu \to \infty. \qquad (3)$$

This law expresses a relation between the two probabilities (1) and (2). BERNOULLI himself says so almost explicitly: in his view, the experimental method of determining the number of cases, i. e., of determining probability (2), is not new or unusual. However[169],

Man muss … ob ein bestimmter Grad der Gewissheit, das wahre Verhältnis der Fälle gefunden zu haben, vorhanden ist, welcher auch bei beliebiger Vermehrung der Beobachtungen niemals überschritten werden kann.

Because of (3) apprehensions are groundless and the probabilities (1) and (2) may be applied on a par. As to the doubts of *einiger Gelehrten* (p. 92), caused by the fact that the terms of fraction (2), in distinction from those of fraction (1), are *unbestimmt und unsicher*, *unendlich* and, moreover, *nicht beständig*, BERNOULLI refutes them (p. 93): both fractions

Sind hinsichtlich unserer Erkenntnis gleich ungewiss und unbestimmt. ... Auch zwischen zwei unendlich großen Zahlen ein bestimmtes Verhältnis bestehen kann. … Bisweilen neue Beobachtungen angestellt werden müssen.

Interesting in this refutation is the assumption of an infinite number of trials. VON MISES subsequently postulated this assumption, while in mathematical statistics, at least from A. A. MARKOV'S work onward, a finite group of observations is interpreted as a sample from an infinite population.

The *Gelehrter* to whom BERNOULLI refers is, presumably, just one: LEIBNIZ, in correspondence with whom BERNOULLI[170] had outlined the content of the fourth part of his as yet unpublished work and, in particular, had confided his concepts about the use of statistical probabilities. LEIBNIZ, at least initially, disagreed, which is all the more strange in that it was he who proposed that a stochastic logic be developed (§ 5). My own understanding of this fact is that LEIBNIZ may have been prepared to weigh delicate subjective opinions and probabilities rather than to enumerate successful and unsuccessful trials in (2).

The law of large numbers and the limit theorems of the theory of probability are generally used in natural science and precisely this use is a definite proof that a science of the accidental, or, rather, of the laws of randomness (§ 1), of a quantitative estimation of the gradual transition of isolated statistical observations into a representative *heap* (§ 6.2.3), does exist.

It seems that from a philosophical point of view both the law of large numbers and the (stochastic) limit theorems establish relations between deductive and inductive methods, i. e. of the theory of probability and mathematical statistics.

One of the possible consequences is that the history of mathematical statistics, if not of probability, should begin with J. BERNOULLI. That the borderline between probability and statistics should be drawn in this wise is also the opinion of CARNAP[171], while mathematicians evidently stress the similarity of both these disciplines[172]. For a natural scientist this last opinion may be less attractive.

10. General Conclusions

ARISTOTLE connected the concept of randomness with non-fulfilment of intention. Disregarding this connection, it is possible to say that his is an understanding of a chance event as dependent on small changes in chains of previous events; which takes place when slight causes determine considerable effects (POINCARÉ).

In a biological context ARISTOTLE offers a dialectical statement connecting randomness and necessity. However, the laws of chance being of course still unheard of, ARISTOTLE denied any possibility of studying randomness.

Jurisprudence had always been based on probabilities and in LEIBNIZ' times (possibly even earlier) the probability $p = 1/2$ became officially recognized in law.

The first concept, though not formalized and unconnected with probabilities, of errors of the first and second kind, now generally used in statistics, perhaps came to be used precisely in jurisprudence. Also in jurisprudence a naive criterion for distinguishing between randomness and divine intervention (determination) had been formulated, possibly for the first time in the history of mankind.
In lawsuits and courts of arbitration arithmetical means of estimates made by different persons were widely used. This stochastic aspect of the theory of means had also been manifested in sculpture (ALBERTI, 15th century) and in games of chance.

There is evidence (CARDANO) that gamblers regarded the mean possible outcome of a throw of astragali as an estimate of reasonable luck, and (GALILEI, indirectly) that of a throw of dice as being most advantageous. Games of chance did not essentially facilitate the development of either combinatorial techniques or of the idea of randomness and probability, but they were used to facilitate reasoning on design versus chance (ARISTOTLE, KEPLER). Also, they provided an opportunity to introduce first numerical notions pertaining to probability (PASCAL, FERMAT).

The role of randomness in biology became apparent only after DARWIN but even before him biologists (HARVEY) and scholars in general did admit that many important biological phenomena were occasioned by chance.

Medicine, as developed by HIPPOCRATES and GALEN, was a science of the probable and HIPPOCRATES even used what amounts to qualitative correlation, while GALEN emphasized the connection of health with mean constitution, mean of all extremes, etc., which is another aspect of the theory of means, and of the general idea that the mean of something is most advantageous. Also, GALEN formulated the problem of testing medicine. Though denying any possibility of a numerical (statistical) solution of this problem, he at least noticed its connection with the general problem (paradox) of the heap, i. e., the connection of statistics with philosophy of experimental science. Lastly, GALEN, also in a medical context, admitted randomness in the sense later offered by POINCARÉ.

Astrology, notably at the hands of KEPLER, had been understood as a study of tendencies and of correlation and it is in this connection that he compared astrology with medicine. As regards foretelling tendencies in the future of nations, the aims of KEPLER resembled those of later political arithmeticians (GRAUNT, PETTY), but of course he did not yet possess statistical data, therefore could not use it and could not be called their precursor.

In astronomy, KEPLER, though denouncing randomness, had to recognize it to explain eccentricities of planetary orbits. His arguments on this subject could be seen as a development of a purely qualitative reasoning of EPICURUS and perhaps it was just in his writings that KANT and LAPLACE later picked up a similar point of view.
In distinction from ARISTOTLE, who had to do with isolated acts of randomness (and whose understanding of it I have linked with that of POINCARÉ), EPICURUS and KEPLER, and, also, KANT and LAPLACE distinguished random influences such as occasioned by causes whose general effect is relatively small, like a noise superimposed on a determinate phenomenon. The first understanding of randomness leads to the uniform distribution (POINCARÈ), while the second one leads to the normal distribution (the central limit theorem, one of whose sufficient conditions is, that, in a sense, each random component is small).

Another point of interest in KEPLER'S (and ORESME'S) astronomy is his understanding that two numbers, taken "at random", are most possibly irrationally related to each other, a fact which he used to prove that the end of the world is practically impossible.

GALILEI, also denouncing randomness, nevertheless successfully separated the regular rotation of solar spots with the sun itself from the random component (i. e., from their proper movement relative to the sun's disk) of the observed phenomenon (of the general movement of solar spots).

Modern philosophers usually reduced randomness to intersections of determinate chains of events; also, an understanding of randomness similar to that of POINCARÉ did exist, but it was related to the general history of mankind rather than to phenomena in nature.

With the outstanding exception of DESCARTES, who described a qualitative picture of a chance origin of the world, almost no one believed in such an origin, the less so as the central limit theorem, according to which a certain order is to be expected even out of blind chance (out of randomness in the sense of uniform distribution), remained of course unknown. And exactly blind chance was the patent notion of randomness. With the exception of HOLBACH and DIDEROT, whose opinion happened to be either unnoticed or at least forgotten, no philosopher thought of any other kind of randomness, e. g., of random quantities possessing binomial distributions, which had been used by A. DE MOIVRE and N. BERNOULLI. Even HOLBACH and DIDEROT did not formulate their reasoning in a mathematical way.

LAMBERT'S lone attempt to formalize the notions of randomness and disorder, an attempt made ahead of time and therefore doomed to failure, is described in part as are also the outstanding achievements of JAKOB BERNOULLI in quantifying inductive inference (quantifying and in the statistical sense solving the paradox of the heap) and thus starting the history of mathematical statistics. A broad picture of the history of randomness and probability has been attempted, and the most general conclusion is that randomness did enter into philosophical and astronomical systems built by most eminent scholars and that, even in antiquity, there did exist sciences of the probable, if not of the accidental.

The main weakness of this article, as I see it, is that I have been unable to connect randomness, as understood by scholars of the past, with what seems to be a modern point of view as developed by A. N. KOLMOGOROV, R. J. SOLOMONOFF and P. MARTIN-LÖF[173] (the "nearer" is the density function of a certain random quantity to the uniform law, the "more random" is this quantity). However, this point of view seems to be pronounced (and, for that matter, only implicitly),
on a heutistic level only; even so, it does not seem to be generally accepted.

Addendum (added in proof)

[1] YU. A. DANILOV & YA. A. SMORODINSKY, J. Kepler: from *Misterium* to *De Harmonice*. *Uspekhi physic. Nauk,* vol. 109, No. 1, 1973, pp 175 – 209 (in Russian).
[2] O. GINGERICH, Kepler. *Dict. Scient. Biogr.* (see note 35), vol. 7, 1973, pp. 289 – 312.
[3] N. L. RABINOVITCH, *Probability and statistical inference in ancient and medieval Jewish literature*. Toronto, 1973.
[4] A. C. CROMBIE, Avicenna's influence on the medieval scientific tradition. In *Avicenna: scientist and philosopher*. Ed. G. M. WICKENS. London, 1952, pp. 84 – 107.

Reference [4], brought to my attention by S. M . STIGLER describes AVICENNA'S (= IBN SINA's) methodology of experimentation.
I have noticed that LAPLACE, in his Sur les probabilités (see note 146, p. 480), introduced what is now called DIRAC'S delta function.

Acknowledgement. Acknowledgement is due to Dr. R. JAQUEL, Professor W. KRUSKAL, Mr. A. H. LANE, Dr. N. L. RABINOVITCH and Dr. I. SCHNEIDER for copies of books and xerox copies and/or reprints of articles. Doing my best, I still have no doubt that Professor C. TRUESDELL will have to correct my English again. Lastly, it is high time to acknowledge the great moral support of Professor A. P. YOUSHKEVITCH, which I have enjoyed now for quite a few years.

Notes

1. *Metaphysica*, 1064b 15. Throughout this article I refer to the English edition of ARISTOTLE (vols. 1 – 12, editor D. Ross). References are given to pages and lines only. Also, it seemed unnecessary to quality the reference to ARISTOTLE by dividing the ARISTOTELIAN *Corpus* into ARISTOTLE proper and pseudo-ARISTOTLE, a distinction which for that matter seems not to have been established with certainty.
2. *Metaphys.*, 1026b,1027a; *Ethica Eudemia*, 1247b.
2a. In the words of F. SCHILLER,
Der Weise sucht das vertraute Gesetz in des Zufalls grausenden Wandern, Sucht den ruhenden Pol in der Erscheinungen Flucht.
(*Der Spaziergang*, 1795. *Werke*, Bd. 2. Leipzig, 1955, p. 710.)
Noticed by A. VASILIEV (*Vestnik Evropy*, 1892, vol. 27, No. 10. In Russian.)
3. O. B. SHEYNIN, Mathematical treatment of astronomical observations. In this collection.
4. *Physica*, 195b – 196b.
5. C. B. BAILEY, *The Greek atomists and Epicurus*. Oxford, 1928, pp. 121 and 139; *Democritus in his fragments and testimonies by ancients*. No place, 1935. Ed. G. K. BAMMEL (in Russian). See pp. 57, 61, 129.
6. *Analytica posteriora*, 94a; *Phys*. 194b; *Metaphys*., 983a, 1013a, 1044a.
7. E. g., *Topica*, 102b 6. See also *Anal. post.*, 73b and *Topica*, 120b
8. *De partibus animalium*, 641b 15.
9. *Magna moralia*, 1207a 5.
10. *Rhetorica*, 1369a 31, see also *Ethica Eud.*, 1247b.
11. *De Caelo*, 283b 1, see also *Phys*. 196b and *Rhetorica*, 1369a.
12. *Metaphys*, 1025a, 1065a.
13. *Anal. post*., 87b.
14. *Parva naturalia*, 463 b.
15. *Phys*., 197b 0 and 197b 14. See also 197a 5.
16. *The Aristotelian-Thomistic concept of chance*. Notre Dame, Indiana, 1945. See p. 22.
17. *Metaphys*., 1025a; 196b 30; and (example 3), *Phys*., 199b 1 and *De generatione animalium*, 767b 5.
18. *Science et méthode*. Paris, 1906. English translation of the relevant chapter: pp. 1380 – 1394 of *World of mathematics*, vol. 2. New York, 1956, ed. J. R. NEWMAN.
19. *Analytica priora*, 70a 0.
20. *Rhetorica*, 1402a 5, see also *De Poetica*, 1461 b.
21. *De Poet*., 1460a 25.
22. *Metaphys*. 1065a; *Rhetorica*, 1361 b.
23. *Magna moralia*, 1206b and 1270a.
24. *Ethica Eud*., 1247a.
25. V. CIOFFARI, *Fortune and fate from Democritus to St. Thomas Aquinas*. New York, 1935. See p. 30.
26. Interesting reasoning on the military art is contained in an ancient Chinese book *Sun-Tsi* written by SUN-BIN in the 4th century B. C. (see chap. 1, § 9 of this book in *Drevnekitaiskaia Filosofia* (*Ancient Chinese philosophy*), vol. 1. Moscow, 1972, in Russian, p. 203):
Who even before battle gains victory by military estimations has many chances … Who has many chances (actually) *gains victory, who has few chances does not gain victory, all the less he who has no chances at all*.
This means that a rare event is practically impossible, a fact which, it seems, still causes headache to philosophers trying to explain it.
Another statistical argument appears in the same source (p. 214) which was possibly written in the 4th or 5th century B. C.:
Jan-Chzhu said: A hundred years is the upper limit of the length of human life. Out of a thousand men even one is unable to reach the age of hundred.
I return to this source in notes 28 and 68. The transcriptions of Chinese names and titles of books which I use in all three places are likely wrong.
27. On the possible and probable in ancient Greece. *Osiris*, vol. 12, 1956,

pp. 35 – 48. See p. 37.
28. *Ethica Nicomachea*, 1104a 24. For similar assertions see 1107b – 1108b, 1133b and, also, *Magna Moralia* and *Ethica Eud.*

A similar understanding of the mean courses of action and behaviour as the best ones is in ancient Chinese literature, See chap. 52 – 53 (*Doctrine of the mean*) and 59 (*Behaviour of scholars*) of the book *Li Tszi* (*Book of rites*, or, alternatively, *A tract on the norms of behaviour*). Compiled in the 4th – 1st centuries B. C., it is one of the books of the CONFUCIUS canon. The *Doctrine* … is attributed to TSZY SY, a pupil of CONFUCIUS. See *Ancient Chinese philosophy*, vol. 2. Moscow, 1973, pp. 119 – 140 (in Russian).
29. See note 3.
30. C. BAILEY, *Epicurus the extant remains*. Oxford, 1926. See p. 25.
31. *Philosophers speak for themselves*, vol. 1. Ed., T. V. SMITH. Chicago, 1956. See p. 140.
32. *De rerum natura*, book 2, lines 216 – 224, 251 – 262, 292 – 293. Engl. translation by H. A. J. MUNRO in *Great books of western world* (hereafter: *Great books*). Chicago, 1952, vol. 12, pp. 1 – 97.
33. *History of western philosophy*. London, 1962. See p. 83.
34. E. g., *Nicht der ungefähre Zusammenlauf der Atomen des Lucrez hat die Welt gebildet* (*Allgemeine Naturgeschichte und Theorie des Himmels* etc., 1755. *Ges. Schriften*, Bd. 1. Berlin, 1910, pp. 215 – 368. See p. 334).
35. E. F. BYRNE, *Probability and opinion*. The Hague, 1968. General information about THOMAS and his works is found in W. A. WALLACE, Aquinas. *Dict. scient. biogr.*, vol. 1, 1970, pp. 196 – 200.
36. *Summa Theologica*. Engl. transl.: *Great books*, vols. 19 and 20. This writing is subdivided into numerous tracts and references to THOMAS' works are given below precisely to these tracts. Passages immediately following are from *Treatise of God*, Q. 19, art. 8; *Great books*, vol. 19, p. 116, and *Treatise on the angels*, Q. 57, art. 3;
Ibidem, p. 297.
37. *Treatise on the divine government*, Q. 115, art. 6; Ibidem, p. 592.
38. *Treatise on man*, Q. 92, art. 1; Ibidem, p. 489.
39. Q. 86, art. 3; Ibidem, p. 463. It is instructive to notice the opinion of D. DIDEROT on the same subject:

Nous les [diverses sciences] *divisons en trois classes, relativement à leur objet: en sciences nécessaires, telles que la métaphysique, les matématiques* … 2° *en sciences contingentes; l'on comprendra sous ce titre la science des esprits créés et des corps*; 3° *en arbitraires, et sous cette dernière classe l'on peut ranger la grammaire*, [une] *partie de la logique, qui dépend des mots, signes de nos pensées*.

See Induction (one of the articles from the *Encyclopédie*) on pp. 206 – 216 in *Oeuvr. Compl.*, t. 15. Paris, 1876. Quotation from p. 208. On p. 212 DIDEROT attributes architecture, painting, music, etc. to the third class of sciences.
40. *Treatise on human acts*, Q. 7, art: 2. *Great books*, vol. 19, p. 653.
41. See note 35.
42. BYRNE, p. 210. See also my § 7.
43. *Alberuni's India*. Delhi, vols. 1 – 2, 1964. Ed. E. C. SACHAN. See vol. 2, chap. 70 (pp. 158 – 160). The three quotations below are all from this small chapter.
44. F. POLLOCK & F. W. MAITLAND, *History of English law before the time of Edward I*, vols. 1 – 2. Cambridge, 1898 (2nd ed.). See vol. 2, p. 598.
45 *Laws of Manu*. Ed. G. BÜHLER. Oxford, 1886 this being vol. 25 of *Sacred books of the East*. See p. 274 and further. Written some time between 2nd century B. C. and 2nd century A. D. (*Enc. Brit.*, vol. 14, 1965, p. 812).
46. *Treatise on the divine government*. Q. 105, art. 7. *Great books*, vol. 19, p. 544.
47. *Judicium matris Kepleri*, see pp. 361 – 562 in J. KEPLER, *Opera omnia*, t. 8, pt. 1. Francofurti a. M., 1870. Quotation from pp. 549 – 550.
48. See p. 36 of article mentioned in note 27.
49. *Rhetorica*, 1376a 19.
50. *Treatise on law*, Q. 105, art. 2. *Great books*, vol. 20, p. 314.
51. *Laws of Manu* (see note 45), § 73, p. 267.
52. *Problemata*, 951 b 0.

53. It is such humane utterances that make distasteful, to say the least, comparison of ARISTOTLE with HITLER:

Aristotle's works, though tough going, did not require, or did not seem to require, anything but common sense to understand them. Like Hitler, Aristotle never told anyone anything they did not already believe.

See (an otherwise excellent book) J. BERNAL, *Science in history*. London, 1957. Quotation from p. 148.

54. BYRNE (see note 35), pp. 223 and 226. See also THOMAS AQUINAS, *Treatise on faith* etc., Q. 25, art. 6. *Great books*, vol. 20, p. 505.

55. *Neue Abhandlungen über den menschlichen Verstand, Nouveaux essais sur l'entendement humain* (bilingual edition), 1765, Bd. 2. Frankfurt/Main, 1961. See book 4, chap. 16, pp. 511 and 513. As to LEIBNIZ' *Grade von Vermutungen und Indizien* I refer to two curious articles (N. CLIFF, *Adverbs multiply adjectives* and J. B. KRUSKAL, *The meaning of words*, both from *Statistics: a guide to the unknown*. Chief ed. JUDITH M. TANUR, San Francisco, 1972, pp. 176 – 184 and 185 – 194, respectively whose authors attempted at, and, as it seems, succeeded in, quantifying (in a statistical sense) usual qualitative characteristics of men's social behaviour. It remains of course an open question whether LEIBNIZ actually thought of any such procedures but it seems that at least the spirit of LEIBNIZ' general endeavours is not here violated.

55a. *Principles of the theory of probability*. *Intern. enc. of unified science*, vol. 1, No. 6 (the whole issue). Chicago, 1939.

56. *Treatise on law*, Q. 96, art. 6; *Great books*, vol. 20, p. 235 and Q. 105, art. 2; Ibidem, p. 314.

57. *Book on games of chance*. Transl. by S. H. GOULD. In: O. ORE, *Cardano, the gambling scholar*. Princeton, 1953, pp. 181 – 241. See § 20, p. 215.

58. *Neue Abhandl*. (see note 55), pp. 513 and 515.

59. JOAN GADOL, *L. B. Alberti, universal man of the early Renaissance*. Chicago – London, 1969. See p. 82. The Italian text is on pp. 116 – 117 of ALBERTI'S *Della pittura e della statua*. Milano, 1804.

60. *Histoire naturelle*, *Suppl*., t. 4. Paris, 1777. See § 8.

61. J. BERTRAND, Preface to *Calcul des probabilités*. Paris, 1888.

62. M. CANTOR, *Vorlesungen über Geschichte der Mathematik*. Leipzig, 1913. See chap. 56, p. 292.

63. J. GADOL, see note 59, p. 80 (footnote).

64. *Scritti letterari*, *Literary works*, vol. 1. Bilingual edition. Ed. J. P. RICHTER. London, 1939. See §§ 587 and 309, respectively. Noticed by GADOL.

65. K. PEARSON, *Life, letters and labours of F. Galton*, vol. 2. Cambridge, 1924. See chap. 12.

66. F. N. DAVID, Dicing and gaming (a note on the history of probability). *Biometrika*, 1955, vol. 42, pp. 1 – 15; M. G. KENDALL, The beginnings of a probability calculus. Ibidem, 1956, vol. 43, pp. 1 – 14. Both articles reprinted in *Studies in the history of statistics and probability*. Ed., E. S. PEARSON & M. G. KENDALL. London, 1970, pp. 1 – 17 and 19 – 34 respectively.

67. This procedure was widely used in social life, as is testified by no lesser authorities than PLATO and ARISTOTLE. Drawing of lots in religious life is described by A. M. HASOVER, Random mechanisms in Talmudic literature. *Biometrika*, 1967, vol. 54, pp. 316 – 321. Reprint: *Studies* (see note 66), pp. 39 – 43. Also, stochastic reasoning in religion is described in several recent articles by N. L. RABINOVITCH (e. g., *Biometrika*, 1969, vol. 56, No. 2, pp. 437 – 441). His latest contribution on this subject is to appear in *Isis*.

68. The outcome of a given throw of dice could well have been considered to carry information about the divine will so that hardly anyone would venture to study possibilities of various outcomes. On the other hand, the probe of the divine will was often attempted by other methods and at least in ancient China such methods were possibly combined with a naive statistical approach:

Fortune tellers (should be) *appointed, so that they should tell fortune by* (studying) *tortoises' shells and stems of milfoil. Fortune* (should be) *told by three men, and the words of two of them* (whose answers coincide) *should be followed*.

Quotation from *Shu Tein*, or *Shan Shu*, a book usually attributed to CONFUCIUS. See chap. *The great law*, p. 108 of *Ancient Chinese philosophy* (see note 26). The

meaning of the title of the book, as given by its Russian translator, is *Book of history*, or, alternatively, *Book of documents*.
69. Letter to F. VAN SCHOOTEN dated Apr 27, 1657. *Oeuvr. Compl*., t. 14. La Haye, 1920. See p. 56. In turn, A. DE MOIVRE cautioned would-be gamblers: *This Doctrine is so far from encouraging Play, that it is rather a Guard against it* (*Doctrine of chances*. London, 1718, 1738, 1756. See Dedication to Lord CARPENTER prefixed to the 2nd edition. The two last editions were reprinted in 1967, in London and New York, respectively.

Also, CARDANO (see note 57), though he does not say anything similar, at least advises his readers as to where, when, with whom etc. it is admissible to play. A strongly-worded accusation of gambling was written by A. DE MORGAN (Theory of probabilities. *Enc. metropolitana*, vol. 2. London, 1845, pp. 393 – 490, see § 31, p. 406): *a society of* (gamblers) *must finally come to ruin, for their transactions* … (do not) *increase the value of any thing they handle … And herein consists the villainy of their occupation, that any permanent increase of their fund must come from unequal play … when a man plays for more than he can prudently spend … with a man or a firm which lives by gambling, there is a fool on one side and a rogue on the other.*
70. *De Caelo*, 292a 30.
71. *Bericht vom neuen Stern* (1604). *Ges. Werke*, Bd. 1. Hrsg. M. CASPAR. München, 1938, pp 391 – 399. See p. 397.
72. *Exposition de la théorie des chances et des probabilités*. Paris, 1843, 1984. Editor, B. Bru. **S, G,** 54. See chap. 12. See also § 2.2.1 of my contribution on NEWTON (note 79) and § 42 of E. BOREL *Le hazard*, Paris, 1914.
73. J. KEPLER, *Harmonices Mundi* (1619). *Welt-Harmonik*. München-Berlin, 1939. Hrsg. M. CASPAR. See book 4, chap. 4, p. 229. See also similar opinion of LAPLACE in his *Essai philosophique* (1814), English translation by A. I. DALE, *Philosophical essay on probabilities*. New York, 1995, p. 99.
74. Brief an HERWART, July 12, 1600. See M. CASPAR &W. VON DYCK, *J. Kepler in seinen Briefen*, Bde. 1 – 2. München – Berlin, 1930. Quotation from p. 139 of Bd. 1. This source contains letters from KEPLER'S correspondence, either translated into, or originally written in German. The whole correspondence (mainly in Latin) is published in KEPLER'S *Ges. Werke*, Bd. 13 – 18. München, 1945 – 1959.
75. § 32, pp. 240 – 241 of his book (see note 57).
76. See p. 145 of his book (see note 57).
77. Sopra la scoperte dei dadi. Engl. transl. by E. H. THORNE appended to F. N. DAVID, *Games, gods and gambling*. London, 1962 (pp 192 – 195).
78. *Neue Abhandl.* (see note 55), p. 515; also, his correspondence with J. BERNOULLI (see note 170).
79. O. B. SHEYNIN, Newton and the classical theory of probability. This *Archive*, 1971, vol. 7, No. 3, pp. 217 – 243.
80. *Origin of species* (1859). Cambridge, Mass, 1964. Reprint of original edition. See chap 5, p. 131.
81. H. H. WOLFENDEN, *Fundamental principles of mathematical statistics*. Toronto, 1942. See p. 181.
82. *Anatomical exercises on the generation of animals*. Originally published in 1651 in Latin. *Great books*, vol. 28, pp. 329 – 496. Transl. R. WILLIS. See ex. 59, p. 462.
83. Ibidem, ex. 1, p. 338.
84. *Tertius interveniens. Das ist Warnung an etliche Theologos, Medicos und Philosophos* (1610). *Ges. Werke*, Bd. 4. Hrsg. M. CASPAR, F. HAMMER. München, 1941, pp. 149 – 258. See § 59, p. 204.
85. *Cosmotheoros* (1698). *Oeuvr. Compl.*, t. 21. La Haye, 1944, p. 653 – 842. Latin-French edition. See p. 702.
86. *Exposition du système du monde*. *Oeuvr. Compl*., t. 6. Paris, 1884. (Reprint of the 1835 edition.) See livre 5, chap. 6, p. 480. A cautious remark is also due to D. DIDEROT (*De l'interprétation de la nature. Oeuvr. Compl.*, t. 2. Paris, 1875, pp. 1 – 62. See § 58, p. 57):

Si la foi ne nous apprenait que les animaux sont sortis des mains du Créateur tels que nous les voyons; et s'il était permis d'avoir la moindre incertitude sur leur commencement et sur leur fin, le philosophe abandonné à ses conjectures ne

pourrait-il pas soupçonner que l'animalité avait de toute éternité ses éléments particuliers … qu'il est arrivé à ces éléments de se réunir … que l'embryon formé de ces éléments a passé par une infinité d'organisations et de développements.
87. *Of the epidemics. Great books*, vol. 10, pp. 44 – 63. See book 3, sect. 1, pp. 54 – 55. This and other *writings which now go by the name of the Hippocratic Collection* were translated by F. ADAMS.
88. *On fractures*. Ibidem, pp. 74 – 91. See § 45, p. 90.
89. *Of the epidemics*, See book 3, sect. 3, § 15 (p. 59).
90. *On the articulations. Great books*, vol. 10, pp. 91 – 121. See § 69, p. 116 and § 71, p. 117.
91. *Aphorisms*. Ibidem, pp. 131 – 144. See sect. 2, No. 44 and No. 49.
92. *Problemata*, 859b 5; 860a 5; 862b 5; 892a 0.
93. *Hygiene* (*De sanitata tuenda*). Springfield, Ill., 1951. Transl. by R. M. GREEN. See chap. 11 of book 5.
94. Ibidem, book 3, chap. 10, p. 132.
95. Ibidem, book 1, chap. 4, p. 11.
96. Ibidem, book 5, chap. 4, p. 202.
97. *On natural faculties. Great books*, vol. 10, pp. 167 – 215. Transl. by A. J. BROCK. See book 3, § 3, p. 200 and § 8, p. 206. Also, *On parts of body*. (I have seen the Russian translation by S. P. KONDRATIEV. Moscow, 1971.)
98. *On parts of body*, book 14, chap. 4, p. 465 of Russian edition.
99. *On natural faculties*, book 1, § 14, p. 177.
100. *On medical experience*. London, 1946. Transl. by R. WALZER. See § 19, pp. 122 – 123.
101. *Hygiene* (see note 93); title of chap. 6 (book 1), p. 20 and pp. 20 – 21 of this chapter.
102. Ibidem, book 1, chap. 5, p. 13.
103. W. KRUSKAL, Statistics – the field. In: *Intern. Enc. Social Sciences*, ed. D. L. SILLS, vol. 15. New York, 1968, pp. 206 – 224 (see p. 216). An instructive paper on the same subject is W. S. B. WOOLHOUSE, On the philosophy of statistics. *J. Inst. Actuaries and Assurance Mag*., vol. 17, 1873, pp. 37 – 56. He describes the early history of the (Royal) Statistical Society established in 1834 and quotes from its official document (p. 37):

The Statistical Society will consider it to be the first and most essential rule of its conduct to exclude carefully all opinion from its … publications, to confine its attention rigorously to facts.

WOOLHOUSE (p. 39) also notices that the emblem chosen for the Statistical Society, a wheatsheaf, reminds statisticians that they should be content, so to speak, with

Binding up … sheaves of wheat for others to thrash out! *These absurd restrictions have been necessarily disregarded in numerous papers.*

The wheatsheaf still remains the emblem of that Society in whose periodicals, however, new definitions of statistics are to be seen nowadays:

Statistics covers all those branches of the mathematical sciences, which are applied to the analysis and understanding of numerical observations, particularly those affected by chance
(H. O. LANCASTER, Problems in the bibliography of statistics. *J. Roy. Stat. Soc.* vol. A133, pt. 3, 1970, pp. 409 – 441, see p. 411.)
104. *Canon of medicine*, vol. 4. Tashkent, 1960 (in Russian). See pt. 1, art. 2, § 52 and pt. 2, art. 1, § 24.
105. M. CASPAR, *J. Kepler.* Stuttgart, 1958. See p. 209.
106. M. CASPAR, see p. 22* of his commentary to KEPLER'S *Welt-Harmonik* (note 73).
107. Brief an MÄSTLIN dated Oct. 3, 1595. M. CASPAR et al (see note 74), Bd. 1, pp. 17 – 24. Quotation from p. 24.
108. *Neue Astronomie* (1609, in Latin). München – Berlin, 1929. Hrsg. M. CASPAR. See Einleitung, p. 33.
109. *Tertius interveniens* (note 84), § 7, p. 161. KEPLER'S understanding of the interrelation of astronomy and astrology is also in his preface to the *Rudolphine tables* (1627). An English translation by O. GINGERICH & W. WALDERMAN is

on pp. 360 – 373 of *Q. J. Roy. Astron. Soc.*, vol. 13, No. 3, 1972, appended to GINGERICH'S article devoted to KEPLER (pp. 346 – 359).
110. *Welt-Harmonik* (note 73), book 4, chap. 6, p. 248.
111. *Tycho Brahe's cosmology from the Astrologia of 1591. Isis*, 1968, vol. 59, No. 3 (198), pp. 312 – 318. The book is P. FLEMLOSE'S *Astrologia* (1591) reprinted in 1644. Tvcho's preface is reprinted by CHRISTIANSON on pp. 316 – 318.
KEPLER himself named TYCHO as his forerunner (see preface to his *Rudolphine tables*, p. 368 of the translation mentioned in note 109):
Tycho knew how to distinguish with most accurate judgement the general effects of the stars from the events themselves in individual human matters.
112. C. D. HELLMAN, Brahe. *Dict. Scient. Biogr.*, vol. 2, pp. 401 – 416, 1970. Quotation from p. 410. FLEMLOSE, also says HELLMAN on the same page,
Gave 399 short rules for weather prediction from the appearance of the sky, sun, moon, and stars, or animal behaviour … daily weather record was kept at Hven in 1582 – 1597.
KEPLER'S attitude towards connections between heaven and meteorological phenomena on earth is discussed below.
113. See note 84.
114. *Welt-Harmonik*, book 4, chap. 7, pp. 269 – 270. See also KEPLER'S Brief an HERWART dated Apr. 9 – 10, 1599. M. CASPAR et al (note 4), Bd. 1, p. 105. In the same place of the *Welt-Harmonik* one finds:
Von dem in der Himmelsmitte erhöhten Jupiter rührt es her, dass ich mehr Gefallen finde … an der Physik als an der Geometrie.
115. *Tertius interveniens*, title of § 55 as given in the *Register*; § 36, p. 179; § 57, p. 200; § 81, p. 220.
116. Ibidem, § 133, p. 253.
117. Ibidem, § 74, p. 217.
118. Ibidem, titles of §§ 12 and 114; § 15, p. 165; § 70, p. 214.
119. The division of certain celestial phenomena into remarkable and usual seems to present the same difficulties as the similar division of events in cases of calculating probabilities of remarkable events (see also § 5). In any case, KEPLER (*Welt-Harmonik*, book 4, chap. 5, Satz 14, pp. 246 – 247) "added" one aspect to those already recognized. Also, KEPLER (*De fundamentis astrologiae certioribus*, § 40) did not fail to notice that the aspects are only *auf der Erde*:
Dieses Vermögen aber, welches den Aspekten ihre Kraft verleiht, ruht nicht in den Sternen selber. Denn die Aspekte … sind auf der Erde und … nicht aus der Bewegung der Sterne wesentlich hervorgehen, sondern die sich ergibt aus der zufälligen Lage je zweier Sterne zur Erde.
Quoted from p. 93 of *Die Astrologie. Ein Auswahl*. Hrsg. H. A. STRAUSS &. S. STRAUSS-KLOEBE. München-Berlin, 1926.
120. *Welt-Harmonik*, book 4, chap. 7, pp. 256 and 263.
121. J. B. H. BROCARD, *Essai sur la météorologie de Kepler*. Grenoble 1880. 2 vols. Extrait du *Bull. de la Soc. de stat. des sci. naturelles du dépt de l'Isère*. This rare source, inaccurately mentioned in the *Bibliographia Kepleriana*, contains translations of relevant passages from KEPLER'S correspondence and comments on the meteorological works of J. WERNER and TYCHO BRAHE. I have seen only one Extrait (pp. 281 – 313).
122. *Welt-Harmonik*, book 4, chap. 7, pp. 272 – 273. BERNOULLI'S *Ars conjectandi* (1713; reprint: Bruxelles, 1968) contains a following passage:
Darüber aber, wie sich diese Gewissheit des zukünftigen Seins mit der Zufälligkeit und der Unabhängigkeit der wirkenden Ursachen verträgt, mögen andere streiten, pt. 4, cap. 1, p. 72 of the German translation by R. HAUSSNER (Leipzig, 1899, *Ostwalds Klassiker* No. 107 and 108). There exists also an English translation, with passages from the LEIBNIZ – BERNOULLI correspondence, by BING SUNG (*Techn. rept* No. 2, 1966, Dept of statistics, Harvard University).
Another point in which KEPLER resembles BERNOULLI, or, rather, vice versa, is his attitude towards a question put to him by someone (*Tertius interveniens*, § 115, p. 238):
Einer … mich bete, ich sollte ihm sagen, ob sein Freundt in fernen Landen lebendt oder todt were … Und ich … sagte ihm ja oder nein, so were ich ein Ariolus und er ein Verbrecher an Gottes Gebott.

BERNOULLI (*Ars conjectandi*, pt. 4, chap. 2, p. 77 of the German translation) is not really interested whether the man is or is not alive. His is the only possible problem: *ob man ihn für todt erklären könne* and he seems to be quite prepared to weigh corresponding probabilities against each other. This was the beginning of "moral" applications of probability completely lacking in KEPLER'S works.
123. *Tertius interveniens*, § 4.
124. See note 79.
125. See, e. g., I. KANT, *Der einzig mögliche Beweisgrund zu einer Demonstration des Daseins Gottes* (1763). *Ges. Schriften*, Bd. 2. Berlin, 1912, pp. 63 – 163. On p. 123 one finds:

Es kann nichts dem Gedanken von einem göttlichen Urheber des Universum nachtheiliger und zugleich unvernünftiger sein, als wenn man bereit ist eine große und fruchtbare Regel der Anständigkeit … und Übereinstimmung dem ungefähren Zufall beizumessen.

See however § 9.2 and note 162.
126. *De stelle nova in pede Serpentarii* (1606). *Ges. Werke*, Bd. 1, p. 284. French transl. by P. SERVIEN, *Science et hasard*. Paris, 1952, p. 132.
127. M. CASPAR et al (see note 74), Bd. 1, p. 261.
128. *The City of God. Great books*, vol. 18, pp. 129 – 618. Transl. by M. Dods.
129. *Welt Harmonik*, p. 317 and p. 17* of M. CASPAR'S commentary; also, CASPAR'S *Alphabetisches Verzeichnis der ... Fachausdrücke* on p. 64* of the *Neue Astronomie* (note 108).
130. *Mysterium Cosmographicum. Das Weltgeheimnis* (1596 und 1621, in Latin). Hrsg. M. CASPAR. Augsburg, 1923.
131. See also my paper Mathematical treatment …(note 3).
132. *Mysterium*, p. 111 of Kap. 18.
133. Anm. 3 und 7 zum Kap. 18, Anm. 3 zum Kap. 17 (pp. 117, 118 und 109 respectively).
134. *Neue Astronomie*, see note 108. Quotation from Kap. 38, p. 238.
135. M. CASPAR et al (see note 74), Bd. 2, p. 66, this being KEPLER'S *Randbemerkung* to a letter from MÄSTLIN dated Oct. 1, 1616.
136. *Epitome of Copernican astronomy* (1618 – 1621). Engl. transl. of books 4 and 5 (1620 and 1621, respectively) by C. G. WALLIS. *Great books*, vol. 16, pp. 845 – 1004. See book 4, pt. 3, § 1, p. 932.
137. See note 73. Other places of interest, besides those quoted in my text, are on pp. 343 and 349.
138. *Allgem. Naturgesch.* (note 34), pp. 269 and 337; *Exp. syst. monde* (note 86), Note 7, p. 504.
139. This is the title of Kap. 23 of his *Mysterium.*
140. M. CASPAR et al. (note 74), Bd. 1, p. 261 (letter dated Oct. 11, 1605).
141. Anm. 5 zum Kap. 23, p. 145.
142. *De proportionibus proportionum.* Transl. by E. GRANT. Madison, 1966. See chap. 3, prop. 10, p. 247.
143. *Ad pauca respieientes*, pt. 2, prop. 17. Ibidem, p. 422.
144. *The Assayer* (1623). Transl. by S. DRAKE. In *Controversy on the comets of 1618* (G. GALILEI, H. GRASSI, M. GUIDUCCI, J. KEPLER). Philadelphia, 1960, pp. 151 – 336. See § 11, p. 197.
145. *History and demonstrations concerning sunspots and their phenomena* etc. (1613). Transl. with intro. and notes by S. DRAKE. Garden City, New York, 1957, pp. 88 – 144.
146. *Elements of philosophy*, pt. 2, chap. 10, § 5. *English works*, vol. 1. London, 1839, p. 130. *Of liberty and necessity* (1646). *English works*, vol. 4. London, 1840, pp. 229 – 278 (reprint of the 1652 edition). See p. 259.

Curiously enough, among those who related randomness to ignorance was J. BERNOULLI (see p. 235 of my article, mentioned in note 79), whose outstanding achievements are discussed in § 9.4. Much the same was the opinion of LAPLACE (Recherches sur l'intégration des équations différentielles etc, 1776. *Oeuvr. Compl.*, t. 8. Paris, 1891, pp. 69 – 197. See § 25, p. 145):

Le hasard n'a … aucune réalité en lui-méme; ce n'est qu'un terme propre à désigner notre ignorance sur la manière dont les différentes parties d'un phénomène se coordonnent entre elles et avec le reste de la Nature.

See also his *Mémoire sur les probabilités* (1781). *Oeuvr. compl.*, t. 9. Paris, 1893, pp. 383 – 485 (§ 2, p. 385). However, he seems to have avoided this subject in his *Essai philosophique*.
147. *Principien der cartesischen Philosophie*. (1663). 1. Tl., Cap. 3 of the *Anhang*. *Sämmtl. Werke*, Bd. 1, Hrsg. B. AUERBACH. Stuttgart, 1871 (pp. 97 – 99); *Appendix* to pt. 1 of *Ethics* (posth. publ. in 1677). *Chief works*, vols. 1 – 2 bound as one. New York, 1951 (reprint of the 1883 edition), p. 78 of vol. 2; *Ethics*, prop. 29 (p. 68 of vol. 2 of the C*hief works*) and note 1 to prop. 33 (p. 71).
148. *Theodicee* (1710). Hannover-Leipzig, 1744. See p 511 of § 302.
149. Lettre troisième of his *Lettres de Memmius à Cicéron* (1771). *Oeuvr. Compl.*, t. 28. Paris, Garnier Frères, 1879, pp. 440 – 463. See p. 441; Chaine ou génération des événements. *Oeuvr. compl*., t. 18. Paris, 1878, pp. 125 – 127.
150. *Le vrai sens du système de la nature*. Londres, 1774. See chap. 5, p. 11; chap. 21, pp 63 – 64; *De l'homme* (1773). *Oeuvr. Compl.*, t. 2. Paris, 1818, p. 33.
151. *Le bon sens* (1772; published under the name of Le curé MESLIER). Paris, 1881. See § 38, pp. 40 – 41; *Système de la nature* etc. Paris, 1770. (Published under the name of MIRABAUD.) See pp. 137 and 141 of pt. 2.
152. D. HUME, *Treatise on human nature*, vol. 1 (1740). London, 1874. See pp. 424 and 428; *Enquiry concerning the human understanding* etc. Oxford, 1902 (reprint of the 1777 edition), chapt. 6, pp. 56 – 59.
An argument resembling reduction of chance to intersections of chains of determinate events is in Indian philosophy (S. K. BELVALKAR & R. D. RANADE, *History of Indian philosophy*, vol. 2. Poona, 1927, see p. 458):
The Yadrichchhä or Chance theory, which is the ultima ratio of Scepticism, when confronted with the argument from … design, is … illustrated by the familiar … maxim:
The crow had no idea that its perch would cause the palm-branch to break, and the palm-branch had no idea that it would be broken by the crow's perch: but it all happened by pure Chance.
153. Anmerkungen über das Buch von dem Ursprunge des Bösen, § 14. Bound to the *Theodicee* (note 148), pp. 710 – 767. See p. 732.
154. *De l'homme* (see note 150), p. 32; *Pensées. Oeuvr. Compl.* Paris, 1963, pp. 493 – 649. See p 549, fragment 413–162; *Système de la nature* (see note 151), chap. 12 of pt. 1, p. 214; *Chaine ou génération*. (see note 149).
155. R. Boyle, *Some considerations touching the usefulness of experimental natural philosophy*, pt. 1, essay 4. *Works*, vol. 2. London, 1772, pp. 36 – 49. See p. 43; D. DIDROT, *Pensées philosophiques* (1746). *Oeuvr. Compl*., t. 1. Paris, 1875, pp. 124 – 155. See § 21, p. 135; HELVETIUS *Le vrai sens* (see note 150) chap. 21, p. 64. For a similar argument put forth in the *Port-Royal* see my article mentioned in note 79. In a private communication Dr. N. L. RABINOVITCH informs me of a similar reasoning by the 10th century Jewish thinker BAHYA IBN PAKUDA. An obsolete, ARISTOTELIAN argument was offered by SPINOZA in letter No. 56 (60) to H. BOXEL. *Sämmtl. Werke*, Bd. 5. Stuttgart, 1841, pp. 350 – 358. See p. 351.
156. I. BARROW, *Exposition on the creed*. *Theological works*, vol. 6. Oxford, 1830, pp. 71 – 442. See p. 99. Also, B. NIEUWENTIT, *The religious philosopher* etc., vols. 1 – 2. London, 1718 – 1719. Transl. J. CHAMBERLAYNE. For a similar opinion of GALEN see § 6.2.3.
157. I. KANT, *Allg. Naturgeschichte* (see note 34), p. 230; VOLTAIRE *Homélies prononcées à Londres en 1765* (publ. in 1767). Première homélie. *Oeuvr. Compl*., t. 26. Paris, 1879, pp 315 – 354. See p. 316.
158. *Le vrai sens* (see note 150), chap. 3, p. 8.
159. *Discours de la méthode* (1637). *Oeuvr. choisies*, nouv. éd. Paris, Garnier Frères, No date, pp. 1 – 59. See p. 34.
160. *Le monde* (1664). *Oeuvres*, t. 11. Paris, 1909, pp. 1 – 118. See chap. 8, p. 48; *Discours*, p. 32 and further.
161. HOLBACH, *Syst. nat.* (see note 151), pt. 2, pp. 138 – 139; DIDEROT, *De l'interpr. nat.* (see note 86), § 51, p. 49.
162. *Kritik der reinen Vernunft*. (1781). *Werke*, Bd. 3. Berlin, 1911. See p. 508. Also KANT's anticipation of the LAPLACEAN determinism observed in *große Zahlen* of individual acts of free will, see my article R. J. Boscovich's work on probability. This *Archive*, 1973, vol. 9, No. 4 – 5, p. 320 (footnote).

163. *Wissenschaft der Logik*. 2 Tl. Leipzig, 1934. (*Sämtl. Werke*, Bd. 4). See pp. 173, 174 and 180.
164. *Dialektik der Natur*. Berlin, 1958. See *Zufälligkeit und Notwendigkeit* (pp. 231 – 235) from the *Notizen und Fragmente*. This unfinished book, written mainly during 1873 – 1882, was first published in 1925.
165. J. H. Lambert's work on probability. This *Archive*, vol. 7, 1971, pp. 244 – 256.
166. Essai de taxéométrie ou sur la mesure de l'ordre. *Nouv. Mém. Acad. Roy. Sci. et Belles–Lettres Berlin*, 1770 (1772), pp. 327 – 342 and 1773 (1775), pp. 347 – 368.
167. See note 122. Quotations below are from the German translation of pt. 4.
168. *Logical foundations of probability*. Chicago, 1951.
169. *Ars Conjectandi*, pp. 90 – 91.
170. The correspondence is published in LEIBNIZ' *Ges. Werke*, Bd. 3/1. Halle, 1855. German translations of important parts are given by C. GIN1, *Schweiz. Z. f. Volkswirtschaft u. Statistik*, 1946, 82. Jg., No. 5; pp. 401 – 413. English translations are mentioned in note 122.
171. *Logical foundations* (note 168), § 49.
172. J. NEYMAN, *First course in probability and statistics*. New York, 1950. See § 1.3.1; S. S. WILKS, *Mathematical statistics*. New York, 1962. See Preface.
173. *Selecta mathematica*, Bd. 2. Hrsg., K. JACOBS. (Coll. of articles). Berlin, 1970.

Afterword

I have somewhat corrected my English and see now that the Editor was unable to revise it thoroughly. But I will now offer more important remarks and, first of all, much additional information is in my later papers, and in particular in my forgotten paper of 1983 (p. 174).

First, the bibliographic information is dated. I mention just two new important sources: Kepler (1609/1992, 2015) and Lambert (1965 – 2007). Readers can find many other new sources in my book Sheynin (2017).

Second, the moral climate of the scientific community, at least in my field of work, has drastically worsened and researchers are more actively than ever before engaged in the scientific rat race. *Publish or perish*! Detestable books are put out by venerable publishers (Sheynin 2006) partly since reviewing is barely recognized as scientific work. Why? Because such is the opinion of the tribe of scientometricians who reigns supreme, reigns by its sledgehammer law (Sheynin 2018). It, that tribe, defies criticism while the pillars of science, the national academies, keep arrogantly silent.

Third, here are some remarks about my text.

On § 3: see Franklin (2001). **On § 8.1.1.** The explanation of the form of planetary orbits by Kant and Laplace are wrong. Laplace was ignorant of Newton's finding, see Sheynin (2017, p. 119). O**n § 8.1.2.** Problems about rational/irrational numbers have no connection with distances between heavenly bodies. **On § 9.3.** Lambert conditioned the application of the theory of probability by the existence of *local disorder* (= irregularity) of the studied elements. Mises specified this condition (which Lambert, in the second part of his memoir, declared necessary and sufficient) by assuming that randomness can only be discussed in a *collective*. **On § 9.4.** Statistical probability (2) depends on the lame definition of the classical probability (1) and the remark on § 9.3 is valid here also.

And, as a general remark on randomness, I stress that chaotic movement ought to be nowadays discussed. Finally, as a general remark on chance, I quote Ecclesiastes 9:11: *all is decided by chance, by being in the right place at the right time*.

Franklin J. (2001), *The Science of Conjecture.* Baltimore.
Lambert J. H. (1965 – 2020), *Philosophische Schriften*, Bde. 1 – 10 + Supplement volume. Hildesheim.
Kepler J. (1609/1992, 2015), *New Astronomy*. Cambridge.
Sheynin O. (1983), Corrections and short notes on my papers. This *Archive*, vol. 1983, pp. 171 – 195.
--- (2006), Reviews. *Historia Scientiarum*, vol. 16, pp. 206 – 209 and 212 – 214.
--- (2017), *Theory of Probability*. *Historical Essay*. Berlin. **S, G,** 10.
--- (2018), History of mathematics. Some thoughts about the present situation. *Silesian Stat. Rev*., No. 16 (22), pp. 127 – 156.

II

Poisson and statistics

Math. Scientist, vol. 37, 2012, pp. 149 – 150

Leaving aside Poisson's important contributions to probability theory, I point out two lesser known facts (Sheynin 1978, §§ 2.5 and 7.2). Poisson introduced both the notions of a random variable (calling it by a provisional term) and a distribution function. More to the point, he emphasized the need to check the significance of empirical discrepancies, and studied criminal statistics following and enlarging on the work of Laplace. Thus, Poisson (1837, p. 4) introduced the mean prior probability of a defendant's guilt, statistically justified but not applicable to any particular individual. Finally, Poisson furthered, although indirectly, stochastic applications to medicine.

In spite of numerous criticisms, mostly directed against the tacitly assumed independence of the verdicts reached by jurors or judges, (which Laplace mentioned in passing), his work was important. Heyde and Seneta (1977, p. 31) noted that in the 19th century *there was a surge of activity stimulated by Poisson* (which seems doubtful) and that (p. 34) in the 1970s *there appeared papers which had reexamined Poisson's work and placed it in a modern setting*.

Poisson's former student, Gavarret, had taken to medicine and published the first book (Gavarret 1840) on what was later termed medical statistics. He (p. XIII) acknowledged his debt to *The lectures and writings of the illustrious geometer* and described the normal approximation of the binomial distribution and the estimation of admissible discrepancies of frequencies in Poisson trials. Gavarret (1840, for example on p. 194) was also the first to stress the importance of checking the null hypothesis, as it was later called, in medicine, but actually in natural sciences generally.

This innovation can be seen as a logical consequence of Poisson's insistence on testing empirical discrepancies. His approach has been an extremely important feature of statistics which followed the work of Wilhelm Lexis in the 1870s.

Gavarret's book became well known, although, on the other hand, it was not mentioned in the literature pertaining to the breakthrough in surgery that took place in the mid-century, i.e. to the introduction of anaesthesia and antiseptic measures (Sheynin 1982, § 6.1).
This is not surprising, as both Poisson and Gavarret only considered large numbers of trials and Poisson (1837, p. vi, in a footnote to the T*able des matières*) declared that

Medicine will not become an art or a science without basing itself on numerous observations.

From the context of his entire book, he meant numerical observations and his opinion was at least corroborated by epidemiologists.

The German physician Liebermeister (ca. 1877, pp. 935 – 940) resolutely opposed that condition. He argued that in therapeutics it was impossible to collect so many observations (and even mentioned hundreds of thousands of them, in an obvious exaggeration), but he did not say that they should have been separated into groups according to the patients' characteristics. Indeed, statisticians cannot avoid small samples, and Liebermeister was a pioneer of medical statistics. Freudenthal and Steiner (1966, pp. 181 – 182) mistakenly attributed to Gavarret rather than Liebermeister the transition from almost absolute certainty with large samples to reasonable probability with small samples.

References

FREUDENTHAL, H., STEINER, H.-G. (1966). Die Anfänge der Wahrscheinlichkeitsrechnung. In *Grundzüge der Mathematik*, Bd. 4, Eds H. Behnke et al. Göttingen, pp. 149 – 195.

GAVARRET, J. (1840). *Principes généraux de statistique médicale*. Paris.

HEYDE, C. C., SENETA, E. (1977). *I. J. Bienaymé. Statistical Theory Anticipated*. New York.

LIEBERMEISTER, C. (ca. 1877). *Über Wahrscheinlichkeitsrechnung in Anwendung auf therapeutische Statistik*. In *Sammlung Klinischer Vorträge* (Innere Medizin No. 31 – 61), Bd. 110. Leipzig, pp. 935 – 962.

POISSON, S.-D. (1837, 2003). *Recherches sur la probabilité des jugements* etc. Paris. **S, G,** 53.

SHEYNIN, O. (1978). Poisson's work in probability. *Arch. Hist. Exact Sci*. vol. 18, pp. 245 – 300.

SHEYNIN, O. (1982). On the history of medical statistics. Ibidem, vol. 26, pp. 241 – 286.

III

Simon Newcomb as a Statistician

Historia Scientiarum, vol. 12, 2002, pp. 142 – 167

1. Introduction

Newcomb (1835 – 1909) was "the most honoured American scientist of his time" (Marsden 1974, p. 33). Known and respected worldwide, he published many contributions in Europe, corresponded with leading European scholars and was a member of prestigious scientific societies at home and abroad. Some of his letters are kept in Berlin and in London (correspondence with Karl Pearson).[1]
From my point of view, Newcomb's greatest merit consisted in discussing and combining the observations of the sun, moon and planets obtained at the main observatories of the world.[2]

For Newcomb's biography see Benjamin (1910), Brown (1910), Campbell (1924) and Marsden (1974). The first two of these authors (pp. 376 and 347 respectively) as well as later commentators stated that Newcomb had to process more than 62 thousand observations of the sun and the planets and that his work included a complete revision of the constants of astronomy. He thus performed a titanic amount of calculations without any aids except for logarithmic tables.[3]
Newcomb's work involved him in studies of the observations made from Ptolemy onward,[4] e.g. [13; 22; 28], and he might well be called a historian of astronomy.

Marsden (1974, p. 33) maintained that Newcomb "had great respect but no particular love, for observational work". Nevertheless, he measured the velocity of light (§ 3.l) and in many cases he [10; 28; 33; 34; 44] discussed the errors of observation which corrupted either the ancient records or contemporary findings.

In addition to planetary astronomy, Newcomb studied the structure of the starry heaven and thus contributed to stellar statistics. He regarded that discipline as a "new branch of astronomical science" [47], see Sheynin (1984, pp. 186 and 187), and held a high opinion of Seeliger, see e. g. Newcomb [55, p. 227]. Drawing on an unpublished letter of that astronomer dated 1904, Paul (l993, p. 78) reports however that he "was not mutually impressed with Newcomb".

Newcomb also published contributions on mathematics,[5] treatment of observations, population statistics, meteorology, and even political economy. In this last mentioned field, he published 42 important writings,[6] many of them controversial (Fisher 1909).

In several addresses and papers Newcomb showed himself as a humane scholar concerned about both abstract science and the future of his native country. The Anthropological Society of Washington awarded him one of its two prizes for his essay [31] on the duties of an American citizen. Back in 1876 Newcomb [18] published a still interesting essay on the history of the abstract science in the United States, from which I am now quoting (pp. 117 and 122):

It might seem entirely feasible to make our country the leader of the world in science at no very remote day.

No want from which our nation suffers is more urgent than that of a wider diffusion of the ideas and modes of thought of the exact sciences.

I am describing Newcomb's contributions, whether astronomical or not, from the angle of the theory of errors and statistics; I understand the latter term as both mathematical statistics and probability theory. My predecessors, whom I mention in the main text below, have only studied a few of the relevant sources (mostly [26]). I myself sketched a mere outline of my present subject (Sheynin 1984) and discussed the same contribution [26] (Sheynin 1995, §§ 4.3 – 4.6).

2. Popular and Expository Writings

2.1. Popular Writings. Newcomb was an eager populariser of science. Up to 1918, his *Popular Astronomy* [21] went through 13 editions. It was translated into three languages and its German version, "bearbeitet" and eventually "völlig umgearbeitet" by Rud. Engelmann, appeared eight times up to 1948. The third German edition which I saw included a discussion of the statistical method in astronomy and contained short biographies of ancient and modem astronomers.

In 1893 – 1895, the *Johnson's Universal Cyclopedia* carried 72 of his items on astronomy, mathematics and physics and its later edition of 1901, the *Universal Cyclopedia and Atlas*, included 81 such items. Even apart from that source, Newcomb published about 75 popular contributions and reviews of the works of his contemporaries, notably R.A. Proctor. He [11] maintained that that astronomer was an able mathematician possessing a good style but liable to error[7]. Later he [17] stated that some of Proctor's essays were "fugitive".

Newcomb presided at the International Congress of Arts and Sciences (St. Louis, 1904) and delivered, as a highly appropriate opening address there, a general and concise review of the development of science over the centuries [50].[8] Newcomb thought [18, p. 106] that the pure and the applied science should not be divorced, but he [39, p. 87] also stated that

The astronomer is moved by the love of knowledge for its own sake, and not for the sake of its application. Yet he is proud to know that his science has been worth more to mankind than it has cost.

I do not discuss Newcomb's three notes on life insurance which were intended for a broader circle of readers. He devoted another note to a problem in geometric probability.

2.2. General Statements. Newcomb's popular writings contain general statements which I describe with minimal comment.

1. Newcomb [45 p. 413] thought that the universe "has existed forever", and that (p. 416) "matter itself is eternal", but he somehow believed that the universe will some time cease to exist (p. 414).

2. Newcomb [16, p. 838] believed that "we probably see the limit of its [of the universe's] densest portion". Elsewhere he [37, p. 8] declared that the evidence suggested that "we actually see the boundary of our universe". Still later he [45, p. 403] corrected himself: only the densest part of the universe had a boundary; otherwise "the heavens would blaze with the light of the noonday sun”. He [5, p. 377; 45, p. 412] rejected Struve’s belief that the interstellar space absorbed light; nowadays this phenomenon is recognized. In addition, Newcomb [45, p. 410] declared that, in “an absolutely infinite system of stars”, “the attractive force … would be infinitely great in some direction or another”.

3. On the contrary, Newcomb [l6, p. 840] doubted that the mutual attraction of the stars was “sufficient to prevent them from all flying away from each other in virtue of the proper motions.” He [19, p. 102] also noted that it was not established “whether the motions of the planets accord perfectly with the theory of gravitation” and mentioned Le Verrier. Newcomb made similar remarks elsewhere. Le Verrier’s discovery led to the conclusion that the law of gravitation “by itself fails to account for all the astronomical phenomena” (Marsden 1995, p. 211).

4. While discussing solar eclipses, Newcomb [51, p. 35] concluded that

The act of rejecting the results of gravitational theory in order to secure the best possible representation of these supposed [total] *eclipses is subject to judgements similar to those pronounced by the Congregation of Cardinals against the doctrine of Galileo.*

This statement is indeed topical! Following others, I (1993, pp. 178 – 179) ridiculed the absurd ideas put forward by A. T. Fomenko et al. and partly based on similar arguments. In essence, they attempted to rewrite the entire history of mankind, and they still continue in the same vein. Novikov (2000) and Zaliznyak (2000) destructively criticize them and regret that Kolmogorov had died: he would have prevented the growth of such rubbish! Highly relevant is Fomenko’s arbitrary opinion (Zaliznyak, p. 164) that the solar eclipse of 431 BC had been total. Cf. Newcomb’s prudent discussion of ancient lunar eclipses (§ 3.1, Note 16). Moreover, he [22, pp. 32 – 33] examined the very same eclipse of 431 BC as recorded by Thucydides and concluded that

The probability in favour of the totality of this eclipse is as great as in the case of any other of those under consideration, though not sufficient to justify the introduction of an equation founded on it.

Later Newcomb [55, p. 232] still remained undecided.

5. Newcomb [47, p. 122] made an unfortunate remark about chemical elements. These, he asserted, “might merge into each other by insensible gradations”. Radioactivity had then been just discovered, but he mentioned three ordinary elements.

2.3. Theory of Probability. In this field, Newcomb published several expositive writings. He [25] called the theory “the art of judging in cases where only probable evidence can be obtained”, and considered it “by far the most slippery [subject] with which the mathematician or logician has to deal”. I would rather say that probability theory studies the laws of chance; Newcomb himself (Ibidem) mentioned the “law of averages” [of large numbers] as “one of the most curious and important results of this calculus”[9].

Elsewhere Newcomb [1] discussed the theory of probability in much more detail and already then he [l, vol. 1, p. 136] offered its similar definition. He (Ibidem, p. 235) introduced ‘the value of expectation" for the discrete finite case and proved that it was expressed in the usual way. He thus followed Huygens and Jakob Bernoulli, whereas, beginning with De Moivre, the expectation (and not its "value") is being defined rather than derived.**10**

True, late in life Newcomb [53; p. 540] borrowed Poincaré's definition (1912, p. 62) according to which "the probable value of a doubtful [of a discrete random] magnitude" taking values $x_1, x_2, \ldots, x_n$ with probabilities $p_1, p_2, \ldots, p_n$ respectively was equal to $[px]$**11**. This, of course, was the expectation of that magnitude, which, curiously enough, Poincaré (p. 57) defined in exactly the same way!

Expectations had not been adequately utilized in probability or its applications. Clausius (1889 – 1891, p. 71) quite unnecessarily proved that $E(\xi/E\xi) = 1$ (modern notation); his ξ was the velocity of a molecule. Then, Poincaré (1912, p. 63) did not see fit to provide the formula for the "probable" [the mean] value of a product of independent random variables.**12** So it is hardly strange that Newcomb [l5, pp. 270 – 271] went on to prove, unnecessarily and restrictively, that

$$E[(s + r_1)(s + r_2) =$$

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(s+r_1)(s+r_2)\exp[-(r_1^2+r_2^2)drdr_2 = s^2$$

where s was a systematic error and r_1 and r_2 were "independent accidental errors".

Two more points. First, as noticed by Stigler (1978, p. 252), Newcomb [1, vol. 3, p. 123], when discussing a problem concerning drawings with replacement, expressed the idea of sufficiency, an important modern notion of mathematical statistics. Second, on the next page, he wrote down the sum of a finite number of terms of the type $1/n^{10}$ as the difference between the corresponding infinite sum and the remainder. He apparently calculated the former by the formula involving the Bernoulli number B_{10} and the Euler – MacLaurin summation formula in the latter case, but he did not explain anything. In another note Newcomb [3, p. 436] distinguished between subjective and objective probabilities. In actual fact he thus followed Cournot and Poisson (Sheynin 1978, p. 249), but he hardly thought about that issue when solving astronomical problems. His note as well as another one [4] were largely devoted to the calculation of the probability of the existence of close stars, see § 4.1.4.

Newcomb [3, p. 440] also mentioned his plans for compiling an apparently never published "paper on the application of the theory of probabilities to natural phenomena".

3. The Theory of Errors

In adjusting observations Newcomb had to consider the venerable problems of minimizing the influence of systematic errors, of assigning weights to observations and to tackle issues now directly pertaining to mathematical statistics. The treatment of observations made under unfavourable conditions prompted him to originate a "generalized" theory (§ 3.4). He obviously excluded such circumstances when advocating "greatest possible diversity" [32, p. 134].

3.1. Systematic and Irregular Errors. The study and the elimination of non-random errors might be related to the aims of preliminary data analysis, an important chapter of theoretical statistics. Newcomb studied systematic errors of observation and especially [33, p. 187; 34; 44] the variation of the personal equation with the magnitude of the observed stars.

In one instance, after examining his measurements of the velocity of light, Newcomb [27] rejected two out of three of his observational series because of suspected systematic errors and much later he [41, p. 3] maintained that "bad consequences" might follow if "all existing material of any kind" be utilized. Dorsey (1944, p. 54), however, concluded that his doubts were unfounded: all three series should have been retained.[13]

At the same time Newcomb [40, p. 133] thought, with good reason, that "the practice of correction for systematic errors has, in recent times, been carried too far" – for such errors, as he himself explained, "which cannot be plausibly traced to a known cause". And in one case he [12, p. 22] decided that the systematic errors were "so small that they may be treated as accidental." Newcomb invariably attempted to estimate the influence of systematic errors in all of his materials. Thus, when reducing the "principal original catalogues to a homogeneous system", Newcomb [12, p. 5] thought, first and foremost, of assigning weights to each catalogue by considering "the probable freedom of the instruments from sources of systematic error" (p. 8).

At the final stage of his calculations he assigned new weights chosen with regard to the possible random errors.[14]

I adduce two examples of systematic (or, in the second case, periodic or irregular) influences considered by Newcomb.

1. Newcomb [35] studied the centennial proper motion in declination of stars separated into four groups in accord with their magnitudes. He assumed that their motions were "as likely to be in one direction as in another" and concluded that the mean proper motion in each group was due to the solar motion. In essence, this statement was tantamount to revealing a systematic error of observation.

2. Newcomb [15, p. 268] investigated "whether the sun's diameter is subject to any changes" by considering the observations made at Greenwich and Washington during 1862 – 1870 (and concluded, on p. 277, that no "sensible variability" was likely). Here is his reasoning. Let $u_1, u_2, \ldots, u_n$ and $v_1, v_2, \ldots, v_n$ with mean values $\bar{u}$ and $\bar{v}$ be such observations. If, during day i, the sun was larger (smaller) than the average, the probability of positive (negative) residuals $(u_i - \bar{u})$ and $(v_i - \bar{v})$ would be $p = 1/2 \pm \alpha$, $\alpha > 0$, and, of the coincidence of the signs of these residuals, $(1/2 + 2\alpha^2)$. Newcomb had not however provided a pertinent quantitative rule.[15] Then, he (p. 270) stated that, when the residuals were "purely accidental" [and independent], the mean value of $(u_i - \bar{u})(v_i - \bar{v})$ will tend to zero. This is wrong: the mean value tends to $\bar{u}\ \bar{v}$ [in probability]!

In 1865 – 1866, Seidel, an astronomer and mathematician, studied the dependence of typhoid fever on some meteorological factor(s), see Sheynin (1982, § 7.4). Following him, Newcomb could have counted the number of the agreements and disagreements of the signs of his residuals and compared that empirical evidence with the results conforming to the symmetric binomial distribution. Seidel (1863) also applied a few reasonable tests for revealing systematic errors, and, at the same time, Ernst Abbe offered his celebrated criterion intended for the same purpose. Newcomb's aim was, however, different, and I do not understand why he [15, p. 270] went on to calculate the mean value of $(s + r_1)(s + r_2)$ for a constant s and independent random variables r_1 and r_2, see § 2.3.

3.2. Rejection of Outliers. Even modern mathematical statistics is unable to solve definitively this issue. Newcomb [24, p. 368] rejected observations "of unknown observers at stations whose longitudes would be difficult to determine"; a similar statement is in [28, p. 376]. He [24, p. 372] also "thr[ew] aside or corrected" an observation "when it appear[ed] probable that it did not or could not correspond to the general mean". Later, however, he [32, p. 186] maintained that deviating or doubtful observations should be rejected only in the presence of a "well established cause of systematic error". And when the distinction between a systematic error (to be somehow corrected), or, possibly, a blunder[16], and a legitimate large random error was too difficult, he [26 p. 344 – 345; 40, p. 134] was naturally at a loss. In the first case and in [66, pp. 212ff] Newcomb remarked that the rcjection of outliers led to abrupt changes depending on the decision made and was therefore not good enough. Once he [36, p. 16] stated, indicating the abrupt change, that rejection was a "matter which has to be left entirely to the judgement of the investigator". Nevertheless, he [28, pp. 263 – 264] remarked that there had been "a decided bias in each generation of astronomers towards depending upon a few recent observations to the exclusion of past ones." He himself, Newcomb added, was also guilty.

Elsewhere Newcomb [26, p. 344] reasonably and even insufficiently criticized the Peirce test for rejecting observations calling it the "best known" criterion. I suspect, however, that at the time the much simpler Chauvenet test published in 1863 already superseded it.

3.3. Various Estimators. In addition to the arithmetic mean Newcomb applied other estimators of the "location parameter".

l. "Where there is a large deviation from the normal law of error[17] the mean values should be determined mainly from the quantities near the centre or [arithmetic] mean of the series" [35, p. 43] but he added that

A more rigorous mode of proceeding would have been to assign a graduated series of weights, converging to zero at the two limits
of the central group of observations. He thus recommended a generalized arithmetic mean. He formulated similar ideas elsewhere [28, p. 374, 32, pp. 84 and 168; 36, p. 17].[18] In the last-mentioned contribution Newcomb again restricted his recommendation to non-normal laws. His attitude and his qualification still seem reasonable

and it is worthwhile to note that some natural scientists thought that the asymmetry of observational series precluded the "Fehler-rechnung" (Meyer 1891, p. 32). See related material in § 3.4.

2. Newcomb [38, p. 162] noted that the "statistical mean" was "independent of the distribution of the quantities in magnitudes". He reasonably applied it when being confronted with a "marked inequality in the distribution". That mean was actually the median, the "valeur médiane", as Cournot (1843,§ 34) named it. In another contribution of the same year Newcomb [36, p. 19] called the same estimator the "mid-point". In neither case did he mention the other term!

Boscovich and then Laplace applied the median much earlier than Newcomb did (Sheynin 1977, § 8.3). The median is a robust estimator, a statistic, robust against possible deviations from the assumed law of distribution. Its application as also the introduction of appropriate posterior weights makes the rejection of outliers much less harmless; nevertheless, the median had not at all superseded the arithmetic mean.

3. Newcomb [55, pp. 212 – 214] described a special procedure for diminishing the weights of the outliers. Let observations be x_i, i = l, 2, …, n, and $\bar{x}$, their arithmetic mean. Then, as he proposed, the corresponding weights will be

$$p_i = \frac{c}{\max(|x_i - \bar{x}|,\ c)}, \qquad (3.3.1)$$

where $c > 0$. Observations of the central group will thus have unit weights and the weights of the deviating observations will diminish. Newcomb noted that the parameter c had to be somehow selected; in addition, iterations will be necessary for consecutively correcting $\bar{x}$.

Stiglier (l973, p. 413) connected this formula with a lesser known modern estimator.

From Gauss onward, the proper measure of precision of the observations had been the variance or its root, the standard deviation (two modern terms)[19]. Astronomers, however, were keeping to a statistically inferior but intuitive appealing probable error. The ratio of the two last-mentioned measures varies with the appropriate law of distribution but in some cases Newcomb [41, p. 3; 53, p. 542; 54, p 167] calculated either of them by issuing from the other one without any comment, as though the pertinent law were normal.

True, Newcomb [24, p, 381] once stated that the normal law should have indeed been assumed. Elsewhere he[l2, p, 17] maintained that the calculation of the probable error by issuing from the mean square error was "more elegant" but that it seemed to him that "too much weight is [thus] assigned to those large residuals the presence or absence of which is a matter of chance", see also [35, p. 43]. I return to this issue in § 3.4. Once Newcomb [24] rather based his computation on the mean absolute error to which the probable error was proportional "and only a little smaller". The latter is indeed equal to 0.845 of the former, but, again, only for the normal distribution (Gauss 1816, § 5).

3.4. A Mixture of Normal Distributions. Cournot (1843, § 132) was the first to recommend the application of a mixture of densities, but he did not elaborate; it is not even clear whether he thought of normal laws. For some details see Sheynin (1995, p. 177). Now, Newcomb [26] assumed that the law of error in astronomical observations was a mixture of n normal distributions with measures of precision h_i occurring with corresponding probabilities p_i, so that the parameter h had thus become a discrete random variable. The unknown values of p_i, h_i and n had to be assigned subjectively[20]. Newcomb then suggested that the observed quantity should be estimated by issuing from that mixture and the principle of maximum likelihood. I (1995, pp. 179 – 182) described his proposal in somewhat more detail and dwelt on its generalizations by Lehmann-Filhés (in 1887) and Ogorodnikov (in 1928 and 1929) and I also mentioned Eddington who had proved, in 1933, that the proposed mixture was not normal, which was possibly unknown to Newcomb!

The proposals of the three authors were hardly interesting for the practitioner, but it is worthwhile to study Newcomb's justification of his innovation. At least from 1872 onward he began to realize that observations did not obey the normal law; in any case, he then expressed his apprehension for the undue influence of large residuals on the estimate of the observed quantity. Later he [24, p. 382] stated that,

Owing to unfavourable circumstances under which observations are frequently made, … the arithmetical mean does not necessarily give the most probable result … any collection of observations of transits of Mercury must be a mixture of observations with different probable errors.

It was this contribution [24] that led Newcomb to his law. Nevertheless, he [26 p. 351] was much too optimistic in hoping to modify "the usually assumed law in order that it may be applicable to all cases whatever".

Newcomb (p. 346) remarked that his innovation led to a generalized arithmetic mean with posterior weights decreasing toward the tails of the adopted distribution (cf. § 3.3.1). Although he had not indicated this feature when providing the necessary formulas (p. 357), it was seen in his numerical example on p. 361.

3.5. The Difference between Empirical Values. Two independent series of observations whose errors obey the (usually normal) distribution with given variances σ_1^2 and σ_2^2 are made. It is required to decide whether the corresponding arithmetic means essentially differ from each other. Newcomb and other astronomers had to solve this problem that now belongs to mathematical statistics.

Suppose that the distribution is indeed normal, that the two series consist of n_1 and n_2 observations respectively, that $\sigma_1 = \sigma_2$ and that $n_1 = n_2 = 100$. Then

the admissible $|\bar{u} - \bar{v}| \le (\sigma_1^2/n_1 + \sigma_2^2/n_2) = 0.22zS$

where S is the sum of the equally probable errors and z is determined in accord with the chosen value of

$$I = \frac{1}{2\pi}\int_0^z \exp(-x^2/2)dx.$$

For $2I = 0.95$ we have $z \approx 1.96$ and, approximately, $|\bar{u} - \bar{v}| \le 0.45S$. If, however, $|\bar{u} - \bar{v}| \le S$ (see below), $z \approx 5$ and $I \approx 1$; this requirement seems too rigid.

Newcomb had not considered the similar problem about the difference between two variances. True, in several cases he noted that it was significant and attempted to act accordingly (cf. § 3.4). Now, however, I describe an instance in which he had not commented that $|\sigma_1 - \sigma_2|$ should have been neglected.

Newcomb [14] investigated whether the then already suspected unevenness in the rotation of the earth had caused the inequalities of long period in the motion of the moon. He asked the Pulkovo astronomer Sergei Glasenapp, who had been studying the eclipses of the first satellite of Jupiter, to check his hypothesis, and he incorporated the latter's answer in his paper.

Glasenapp compared [14, p. 164] the probable errors 9.89 and 9.09 with 9.77 and 8.97 respectively[21] characterising the two possible answers and concluded (unreasonably, in my opinion) that the second case provided a "somewhat better" correspondence between theory and observation.

Then Glasenapp (p. 166) stated that the difference between the values of certain quantities, m_1 and m_2, and k_1 and k_2, was "larger than their probable errors allow it". Now, the actual figures were

$\Delta m = 0.044$; sum of probable errors of m_1 and m_2, 0.036;
$\Delta k = 0.17$; the similar sum, 0.25.

Newcomb only concluded that, although the unevenness in the rotation of the earth was thus likely corroborated, mostly by other, independent qualitative comparisons also made by Glasenapp, his colleague had not proved that it really caused the remaining errors in the moon's place. He performed additional calculations based on an empirical approach and decided that the hypothesis tested remained "worthy of reception" (p. 170).

Newcomb himself [38, p. 165], somewhat more definitely than Glasenapp did, maintained that two empirical values differed essentially if their difference was "much greater than the sum of their probable errors", and it seems that at the time this was the standard criterion.

.

In 1903 Markov (Sheynin 1989, p. 351) agreed with F. A. Bredikhin's orally expressed rule according to which "in order to admit the reality of a computed quantity, it should at least twice numerically exceed its probable error". I chose to differ, and, moreover, this case does not apply to the problem under discussion, but at least the (hardly Bredikhin's) test confirms that astronomers had been using statistical rules of thumb[22]. Newcomb, however, did not comply. In one instance he [28, p. 401] concluded that a certain constant was $k = 0.23 \pm 0.15$. Elsewhere he [43, p. 9] listed four computed values with their mean (not probable) errors two of which were

$$0.297 \pm 0.43 \text{ and } 0.05 \pm 0.92$$

and similar cases are elsewhere [55, p. 219]. Regrettably, Newcomb had not commented.

4. Statistics

4.1. Laws of Distribution. Noting that the first pages of logarithmic tables wear out "much faster" than the last ones, Newcomb [23] set out to derive the probability that the first significant digits of empirically obtained numbers will be n_1, n_2, … Such numbers were "ratios of quantities"; a measurement, as I understand him, shows the ratio of the relevant magnitude to the appropriate unit.

Going over to logarithms with base a and considering only their mantissas, Newcomb had to study the expression

$$a^{(s_1-s_2)(s_3-s_4)+\dots}$$

and take into account only the positive fractional portion of the exponent. Each difference in the exponent corresponded to an empirical number, and s_1, s_2, … were selected "at random".

"Whatever be the original law of arrangement" of the s_i, 2^n in number, Newcomb stated, "the fractions will approach to an equal distribution" around a circumference as n increased. Without any ado, he formulated his main conclusion:

The law of probability of the occurrence of numbers is such that all mantissas of their logarithms are equally probable.

Several authors, some of them without mentioning Newcomb (Benford 1938, Feller 1971, pp. 63 – 64), proved this proposition which Raimi (1976, p. 536) considered Newcomb's "inspired guess". Raimi also noted that it was not universal. For my part, 1 connect Newcomb's statement with his remark about the distribution of the asteroids (Item 2 below) and note that it heuristically resembles the celebrated Weyl theorem according to which the terms of the sequence $\{nx\}$ where x is an irrational number, $n = 1, 2, \dots$ and the braces mean "drop the integral part", are uniformly distributed on an unit segment. Again, it is remarkable that Newcomb's statement means that, in the sense of the information theory, each empirical number tends to provide one and the same information.

I continue with several astronomical examples the second of which is likely connected with the material just discussed.

l. Newcomb [6] qualitatively studied the distribution of the, parameters of the orbits of the 57 first asteroids (in longitude, for the perihelia and nodes; by magnitude, for the eccentricities and inclinations). For the first two parameters he also compared the actual and the "probable" distributions without explaining how he derived the latter.

2. Newcomb began his next relevant article [7] by maintaining that the inequalities in longitude of the nodes and perihelia of the small planets had "frequently been a subject for remarks and speculation", for "foolish speculation", as he wrote in the covering letter to the Editor, C. A. F. Peters, on July 31, 1862.

Newcomb's account is extremely difficult to understand, mostly because of careless notation, but it seems that he (p. 212) intuitively came to the following statement: The independent angles $(B_1 + b_1t)$, $(B_2 + b_2t)$, … will be "pretty evenly distributed around the circle [the circumference] if their number is great". Here, t denotes time and the other magnitudes are constant. Newcomb admitted that at some moments the uniformness might disappear, although not in the case of an infinitely many asteroids. His qualification remark seems, however, wrong.

Newcomb then makes use of the uniformness for deriving the probability that the longitudes of the perihelion and node of an asteroid of a given mean distance is in any particular quadrant and provides stochastic considerations simple in principle but again more difficult to follow in detail.

Here is an example (pp. 214 – 215) concerning a certain parameter σ. It

May be regarded as a constant for all values of x up to a certain limit More exactly, when $\sigma = 0$, *we have also* $d\sigma/dx = 0$. *Again, as we approach* [some value of x], σ *will gradually vanish*.

The simplest hypothesis, Newcomb continues, is that $d\sigma/dx$ is "continually negative".

3. While studying proper stellar motions, Newcomb [46, p. 166], see Sheynin (1984, pp. 182 – 183), supposed, as a "simplest" initial hypothesis, that their projections on an arbitrary axis were normally distributed. As a corollary, he correctly provided, although without any explanation, the density law of their projections on an arbitrary plane and their own distribution. Both laws were connected with the chi-squared distribution and their derivation is not elementary[23].

On p. 167 Newcomb provided a table

Of the adopted distribution of the linear speeds of stars, relative to the Sun, and projected on a plane.

For such projections, at least for angular motions, he derived (above) the density

$$\varphi(x) = Cx\exp(-x^2/b^2), \text{ or } C(x-a)\exp[-(x-a)^2/b^2]$$

but he had not explained how he estimated b. And he also stated that the "distribution should follow the exponential law"[24].

4. In 1767, Michell, assuming that the stars were scattered "by mere chance" over the sky, attempted to derive the probability that two of them were close to each other. I (1984, § 5) discussed his calculations and the work of later commentators including Newcomb [1, vol. 1, pp. 137 – 138; 3, pp. 437 – 439] who made use of the Poisson distribution.

He [4] also solved a related problem where he determined the probability of the mutual inclination of two great circles randomly situated on a sphere, or, rather, that the distance a between their poles was $n < a < m$. Newcomb stated that in this case the probability sought was equal to the ratio of two areas, of the spherical zone just containing the poles of the two circles and the hemisphere, that is, to

$$P = \cos n - \cos m.$$

Elsewhere he [48, p. 13] noted that

When a point is taken at random on a sphere, all real [?] *values of the cosines of its distance from a fixed point are equally probable*.

In this latter instance Newcomb effectively maintained that the density function of a was

$$\varphi(x) = \sin x.$$

Indeed, this function is proportional to the length of the locus of random points having distance x, and it is now easy to see that $\cos a$ has a uniform density. Newcomb [48] also correctly stated that the probable [the mean] value of $\cos a$ [a blunder: he meant $\cos^2 a$] was 1/3. He derived this equality from a simple but apparently insufficient geometric condition; nevertheless, it is again easy to see now, that, if a random variable is uniformly distributed on [0; l], its square has mean value 1/3.

Laplace, as Newcomb [4] remarked in his previous paper, had wrongly solved the earlier problem. He (1812, p. 261) assumed that, rather than (3),

$$P = (m - n)/90.$$

Cournot (1843,§ 148) offered yet another solution:

$$P(\alpha \le a \le \alpha + d\alpha) = \sin\alpha d\alpha.$$

He apparently thought that the second pole could be situated anywhere on a circumference of a circle with radius proportional to $\sin\alpha$.

5. Many times Newcomb compared empirical and theoretical distributions and noted whether the discrepancies were reasonably small. He accomplished this work in the examples 1 – 3 above as well as in at least two other cases.

a) Newcomb [42] studied the mean daily motion (μ) of asteroids. He selected 354 of them having motion $600'' \le \mu \le 1000''$, arranged them in 40 groups with μ =600 – 610, 610 – 620, …, 990 – 1000″, and, without explanation, adduced a table of probabilities for a group to contain 0, l, 2, …, 19 small planets. These probabilities conformed to the Poisson distribution

$$\varphi(x) = e^{-\lambda}\lambda^x/x!$$

for λ = 354/40 = 8.85 to within 0.002.

b) Newcomb [35] compared the law of distribution of the stars by their motion with two normal laws differing in "probable errors" [variances]. Both the choice of these variances and the comparison were qualitative.

4.2. Sex Ratio at Birth. In his pertinent study, Newcomb [49] drew on some American genealogical data[25] and on the statistics of twin births in France and Berlin. He noted that the American Civil War had "not the slightest influence" on that ratio (p. 27) and that the American records were imperfect since a large family might have consisted, over the years, of the same father and two [or more] mothers so that the ages of the parents were not sufficiently clear (p. 26).

Newcomb (p. 21) suggested that after conception there might have been "a series of causes" tending

To make one sex or the other more probable, until, gradually, the sex is definitely determined.

He thus thought about an event whose probability changed with time which was a step toward introducing random functions. Newcomb (p. 29ff) then assumed that there existed three classes of families, h_l, $h_2/2$ and $h_2/2$ in relative numbers, $h_l + h_2 = 1$, producing male children with probabilities p, $p + \alpha$ and $p - \alpha$ ($p \approx 1/2$) respectively. By means of the binomial distribution he obtained the probability that a family with n children will have s boys and r girls ($s + r = n$)

$$P = C_n^s[\frac{1}{2^n} + \frac{(r-s)^2}{2^{n-1}} h_1\alpha^2].$$

Necessarily considering $h_1\alpha^2$ as a single unknown, assuming $h_1 = 1$ and making use of statistical data, he arrived at $\alpha = 0.27$. He (p. 18) then somehow applied the same value of α for studying twin births, although, complying with his explanation of the production of sex, he held that unisex twins were more probable than their counterparts, and, by implication, the parameter α should have been larger for them. The magnitude h_1 remained unknown.

On Nov. 14, 1904, Newcomb arranged for sending a copy of his study to Pearson calling it "a paper very much in your line" and asking him to "examine it carefully". He repeated his request on Nov. 21 indicating that in his case "the ordinary system of comparing causes and effects" was "impracticable". He did not elaborate. On Jan. 2, 1905, Pearson answered that he had "duly received & read" Newcomb's memoir. Pleading lack of time, he found himself "unable to write at length", but he thought that the work "might be strengthened had it been accompanied by the probable errors of the quantities involved".

4.3. Time Series. These are sequences of empirical numbers corresponding to certain instants of time. Until the 1920s, elementary methods were used (by Newcomb also) for studying such series. After that, stochastic models began to be applied.

1. When investigating the variation of terrestrial latitudes Newcomb [29] issued from Chandler's empirically derived period $T = 427$ days

= l.17 years. Assuming a uniform movement of the poles along a circumference, he compared eight properly reduced observations of maximal latitude during 1865 – 1890 with the calculated epochs.

The observations followed each other after $\Delta t \approx$ l.17n years with n = 6, 3, …, but I might as well suppose that n = 1. Newcomb stated that, "were the observed epochs not periodically recurring", the mean discrepancy "excluding the first and the last of the series" would be 0.29 years. Noting that only one result was greater in absolute value than 0.29 whereas all the rest ones fell "within an octant of the observed epoch", he concluded that "the chances are about 400 to 1 against this being the result of chance", and that the existence of period T = 430 (not 427!) was proved "beyond reasonable doubt".

The total length of segments each representing 1/8 of the distance between adjacent equidistant points 1, 2, …, 7, 8 will be 14:8 = 7:4, or 1/4 of the length of [1; 8]. The probability that exactly seven out of the eight observational points will be situated on these segments is therefore

$$P = 8 - (1/4)^7 \cdot 3/4 = 1/2730$$

which does not at all corroborate Newcomb's estimate (but all the more confirms his final conclusion). If the "left" neighbourhood of point 1, and the "right" neighbourhood of point 8 be added, then P = 1/1420, which does not change the situation.

I shall now check the other estimate, 0.29. Arrange eight other points $a, b, \dots, g, h$ uniformly on segment [1; 8] in such a way that [1; a] = [a; b] = … = [g; h] = [h; 8]. Then the mean distance between a new point and the nearest initial point will be 0.275; or, multiplied by 1.17, 0.32, but not 0.29.

2. The essence of Newcomb's contribution [43] was the derivation of the periodicity (T) of sunspots. He collected their observations covering 26 periods (1610 – 1889) and considered the maxima and the minima of the phenomenon as well as the "half-tide phases", half the sums of the numbers of the sunspots "corresponding to the year of minimum and the following maximum, or vice versa" (p. 4).

Assuming some zero phase and a provisional periodicity T_0 Newcomb formed observational equations

$$x + n\Delta T_0 = \Delta t$$

with x being the correction to the assumed zero phase, Δt, the difference between the observed and the computed epochs, and n, the appropriate number of periods. His preliminary solution (p. 6) of these equations by the method of least squares (which he omitted) yielded

$$T = 11.13 \text{ years.}$$

Then Newcomb found several "definite" solutions but the preliminary value of T persisted (p. 10). He had not expressly explained the aims of the preliminary adjustment; obviously, it was needed for studying any peculiar features of the observational results (which he indeed indicated) and assigning them proper weights. What I fail to understand is the separate adjustment of each of the four phases (only two of which were independent) and the calculation of the mean of the four periodicities thus obtained. It is thought nowadays that $T \approx 11$ years, but that a strict periodicity does not exist.

3. Newcomb [52] attempted to devise a test for correlation (loosely understood) among terrestrial temperatures by means of their simultaneous measurements

.

$$v_{1j}, v_{2j}, \ldots, v_{nj}, j = 1, 2, \ldots, r$$

in r regions of the earth. He did not go beyond the limits of the classical error theory and his main assumption was that the products of two independent observations might be neglected.

In a few preliminary sections Newcomb considered time series possibly having a main periodic component. Let

$$a_{0j}, a_{1j}, \ldots, a_{nj}, j = 1, 2, \ldots, r$$

be sequences of equally distant terms selected from the series. Then, as he assumed,

$$a_{1j} = a_{0j}x_1 + e_{1j}, a_{1j} = a_{0j}x_2 + e_{2j}, \ldots, a_{nj} = a_{0j}x_n + e_{nj},$$

with the e's being random terms. The unknown x's might then be calculated (each one separately) from n normal equations. If $a_{0j} > 0$, then, as Newcomb remarked, the unknowns will increase until "the completion of the period", assuming that it existed and did not exceed the time interval covered by the selected sequence; if no period existed, then (p. 321)

$(1/n)(x_1 + x_2 + \ldots + x_n)$ will converge [in probability] to zero.

I doubt whether his recommendations were ever applied.

4. Newcomb [54, pp 167 – 170] studied the moon's mean motion in longitude from 1621 to 1908. He graphically represented the observed points and the curve of "the pure theory", showed the mean error of the "true [observed] curve" at each point by additional curves; he obviously interpolated both the observations and the errors so as to obtain continuous lines. In addition, he constructed the curve of the "great fluctuation" of long period although without insisting on its reality and he (p. 168)[26] maintained that the

Minor deviations during the past 100 years may be empirically represented by a trigonometrical series,

but he did not justify his conclusions or adduce his calculations.

Newcomb regarded the observed fluctuations as "the most enigmatic phenomenon presented by the celestial motions", "perhaps due to fluctuations in the Earth's speed of rotation" (p. 168), cf. end of § 3.5.

5. The Method of Least Squares, MLSq

The initial data treated by least squares are represented by a redundant system of equations such as

$$a_i x_1 + b_i x_2 + \ldots + s_i = 0,\ i = 1, 2, \ldots, n$$

in k unknowns ($k < n$) where the coefficients are given by the appropriate theory and the observed free terms s_i are "physically" independent. Hence, the system is inconsistent and is adjusted under the condition

$$\sum v_i^2 = \min$$

where the v's are the residual free terms. The errors of s_i are assumed to be only random; their normal distribution presents the most favourable, but not necessary case.

The condition above leads to a system of k normal equations

$$[av] = [aa]x_1 + [ab]x_2 + \ldots + [as] = 0$$
$$[bv] = [ab]x_1 + [bb]x_2 + \ldots + [bs] = 0, \text{ etc.}$$

In accord with the two methods of forming the initial equations, they are called observational or conditional. Unlike geodesists, astronomers applied only the former, but always called the equations conditional.

In 1823, when providing his definitive justification of the MLSq Gauss proved that for unimodal distributions (for one and the same distribution in each case) it yielded unbiased estimators with least variance. True, systematic errors and blunders, dependence between observations and their incorrect weighting were not accounted for by the classical error theory.

Strange as it is, Newcomb, although referring to that justification [26, p. 348n], shared the wrong opinion of many other scientists (Eisenhart 1964, p. 24) that the MLSq was inseparable from the normal distribution [32, p. 82; 38, p. 161].

5.1. Non-Standard Procedures. Newcomb undoubtedly understood that formal rules should not be strictly adhered to Thus [28, p. 372], observations are made "under widely different conditions", they have' "very different moduli of precision", and the MLSq "as usually applied, will not give the best result". Here is an example.

Bond (1857) adopted certain approximations when calculating [*av*], [*bv*], …, , but I saw only one contribution, [36, p. 31], whose author, Newcomb, applied his recommendations. Interesting enough, Bond followed and referred to Gauss (1809, § 185) who had dropped a few pertinent lines. Gauss remarked that it was often possible to apply factors more convenient for calculation and little differing from those indicated by the theory for transition to the normal equations.

When dealing with nine thousand equations in 20 unknowns, Newcomb [32, p. 48] recommended to drop "all superfluous decimals” in their coefficients and claimed that “no serious harm would result”. This statement should have been better formulated; in essence, however, the influence of rounding off on the end results of such calculations was then not yet studied, cf. § 5.2.

Newcomb (Ibidem, p. 52) also stated that small coefficients in the systems of normal equations might be neglected, but he did not come up with any definite criterion.

Newcomb [13, p. 167] formed 89 observational equations, solved the corresponding normal system in five unknowns and remarked that he, likely wanting to eliminate some systematic influences, made “a farther approximation … by solving the equations given by the residuals”. Without justifying this approach, he only provided the least-squares, and the final solutions. During the second stage of this work, Newcomb apparently treated the observational, rather than the normal equations: once the latter are solved, a further attempt with the v’s replacing the free terms s_i in the sums $[as]$, $[bs]$, … leads to a zero solution.

Newcomb had not explained either the method of performing the second stage of his calculations, or even why was the first stage insufficient. In any case, however, as he stated, the final residuals [nevertheless] testified to some unrevealed systematic errors of observation.

Until now, all the weights, p_i, of the normal equations were supposed equal; otherwise, however, these equations should be multiplied by $\sqrt{p_i}$ respectively so that the coefficients of these equations become $[paa]$, $[pab]$. … In one instance Newcomb [55, p. 208] stated without explanation that an equivalent procedure was only necessary when the probable errors and the coefficients of the equations were large.

Another unusual case concerned an unknown, λ, subject to fluctuations. Newcomb [55, pp. 124 – 125] subdivided his equations into time-groups and assumed that in each of these the real λ did not essentially differ from its value increasing uniformly with time.

5.2. Ill-Conditioned Systems. Marian Kowalski, a Professor at Kazan University, published his study of the movements of Neptun (1856)[27], and Newcomb [8] made use of his work. Kowalski (p. 179) formed four normal equations whose terms he calculated with six significant digits. The first two provided (in my notation)

$$x_1 = \alpha_1 x_3 + \beta_1 x_4,\ x_2 = \alpha_2 x_3 + \beta_2 x_4,$$

and he used these expressions to form

$$[ac]x_1 + [bc]x_2 \text{ and } [ad]x_1 + [bd]x_2.$$

Kowalski noted that, with minor discrepancies, he thus derived the third and the fourth normal equations respectively and he reasonably concluded that his normal system included only two independent equations.

Newcomb [8, p. 4] stated that the outcome of Kowalski's calculations was "the necessary result of the mode of treating equations by the MLSq". On his next two pages he presented his further comments. Suppose, he began, that a non-redundant system in k unknowns is

$$a_1x + b_1y + \ldots + n_1 = 0,\ a_2x + b_2y + \ldots + n_2 = 0,\ \ldots$$

then

$$x = \frac{A_1}{R}n_1 + \frac{A_2}{R}n_2 + \ldots,\ y = \frac{B_1}{R}n_1 + \frac{B_2}{R}n_2 + \ldots,$$

where R is the determinant of the system and A_i and B_i are the appropriate "partial determinants". For a redundant system, he continued, R should be replaced by the sum of the squares of all the determinants corresponding to the solution of all the subsystems of k equations each, and A_1, A_2, … become "certain powers and products of the partial determinants".

Only then Newcomb mentioned the MLSq and stated that the partial determinants, when applying it, will be

Very small if the equations are nearly equivalent to a number less than that of the unknown quantities; that is, if they can be put into the form

$$X = n_1,\ Y = n_2,\ \ldots,$$
$$\alpha_1X + ß_1Y + \ldots + \rho_1 = n_4,\ \alpha_2X + ß_2Y + \ldots + \rho_2 = n_5,\ \ldots$$

where the set $\{X, Y, \ldots\}$ consists of less than k elements.

Newcomb next calls a system in k unknowns identical if it contains less than k independent equations and formulates the following theorem:

If a system of equations differs from identity by a very small quantity, the normal equations derived from them will be identical to small quantities of the second order.

Their solution, he continues, should then be carried out with "nearly twice as many decimals as are necessary in the original coefficients". In such cases Newcomb recommended to transform the observational equations by introducing new unknowns instead of the initial ones. As usual, he expressed himself loosely ("equations … equivalent to a number")**[28]** and described his thoughts much too concisely.

Now, Jacobi (1841) had proved what Newcomb stated, viz., that the least-squares solution of the initial equations was the same as the solution derived from the partial subsystems when taking the appropriate weighted means; such coefficients as Ai/R, Bi/R, … should allow for the weights of the partial solutions. Again, Newcomb correctly maintained that the MLSq might aggravate the situation: in case of ill-conditioned observational equations it is better to proceed without forming the normal equations, applying for example successive approximations (Gentle 1998, p. 111). It is also known that the determinant of a system of normal equations is necessarily positive, but that for ill-conditioned observational equations it might become small, and, owing to the unavoidable "play" of the last significant digits, even negative (Idelson 1947, p. 39). According to my calculation, the determinant of Kowalski's normal equations was $R = -60$!

Again, Faddejew & Faddejewa (1970, p. 151) state, without however, offering any quantitative rule, that if R is much less than its boundary

$$R^2 \le \left[\prod_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 \right],$$

where a_{ij} are the coefficients of the normal system (the Hadamard inequality), the system is ill-conditioned.

5.3. A Mechanical Representation. In 1873, Newcomb had published a note on the mechanical representation of the MLSq verbatim reproduced and discussed by Farebrother (1999, pp. 168 – 169). The latter also dwelt on Newcomb's companion note of the same year and linked both notes with some previous work of other authors. A mechanical interpretation of the treatment of observations appeared in 1722 (Roger Cotes) and his methodological approach led to the so-called method of geodetic relaxation one of whose forms is due to Gauss. See Sheynin (1963).

6. Some Conclusions

In addition to what I had to say in § 1, I note that Newcomb worked during the period when the theory of probability had been largely neglected and this explains his careless terminology and clumsy calculation of expectations (§ 2.3). Worse, he (§ 3.1-2) did not indicate that in probability theory the notion of limit should be specified[29]. On the other hand, he (§§ 4.1 – 4.2) made some interesting stochastic findings.

Newcomb certainly mastered the theory of errors and applied it (including its MLSq) non-formally and successfully. He himself noted that the assigned weights of the observations were "the result of judgement more than of computation" [32, p. 21], cf. the subjective nature of his generalized law of error (§ 3.4), and that (p. 85) his "policy"was to obtain as many as possible independent observations of the quantities sought; by implication, these statements meant that the collection of sound observations followed by their reasonable treatment was the essence of his pertinent work.

I do not understand however why he was obviously satisfied with some of his results providing the quantities sought with great probable errors (§ 3.5) and I also remark that he had not mentioned Helmert's classical treatise (1872). Much later Newcomb [55, p. 41] expressed his gratitude to the German scientist for acquainting him with the latest values of the parameters of the geoid.

7. Appendix

7.1. Newcomb and Pearson. I mentioned Pearson in §§ 2.1, 3.1 (Note 14) and 4.2. Here, I discuss other relevant issues, again drawing on Newcomb's correspondence with him.

l. Newcomb – Pearson, 27.6.1903, 21.11.1904 and l.11.1907, respectively.

a) *I believe you are the one living author whose productions I nearly always read, when I have time and can get at them, and with whom I hold imaginary interviews while I am reading.*

b) *I know of no one more competent than yourself to appreciate this attempt* [see my § 4.2]; *and any opinion or discussion of the paper which you might express would have great weight. If favourable, it might be decisive in enabling me to develop and extend general methods in dealing with statistical data in all the sciences where they are available.*

c) *I am getting more and more interested in almost every branch of the work you are pursuing, and postpone taking an active part in it only because I have to complete some astronomical investigations.*

Newcomb would have been pleased to know whether Pearson had any new information about the "grouping in the results of the roulette at Monte Carlo", cf. Pearson (1894).

2. Newcomb – Pearson, 21.11.1904. Newcomb enlarges on his idea

Of an institute for reducing and working out the results of scientific observations … In meteorology, for example, we should begin with the observation most appropriate to test meteorological theories, or the general phenomena of meteorology, their periodicity and their dependence on any such terrestrial cause as the sun's radiation.

No other details are presented. Did Newcomb envisage a national, or international institution? In 1915, or early 1916, the Russian statistician Chuprov thought about the creation of an institute for the statistical study of Russia (Sheynin 1990/2011, p. 130).

3. Newcomb – Pearson, 1.11.1907

When … you published [communicated] *Miss Gibson's paper … I was minded to write you expressing my pleasure that you were extending your statistical methods into astronomy, but pointing out that the method adopted … was not likely to lead to any conclusive result. This,… from the meagreness and uncertainty of the data and the omission to consider relations known a priori among the quantities classified. … having noticed the recent discussion in* <u>Nature</u> *I venture to submit a few remarks …*

When we seek to find a correlation between two systems of observed quantities, it is requisite to a certain result that the quantities of each series be not in the nature of purely accidental ones and that there be something we can consider definite. Examples are when either system is the result of random sampling, or when the number of quantities is sufficiently large to establish some law among the magnitudes, even when pure accidental.

In the case of stellar parallaxes regarded simply as observed quantities without reference to known conditions affecting their value, neither of these requirements is satisfied. …

In order to reach definite results in this field, the known relations between magnitudes, distances, and parallaxes must be taken as the *basis of the investigation. Moreover, the adopted method must be that of trial from hypotheses, by deduction and comparison with observation rather than by pure induction. … It seems to me the only method by which we can obtain results is that of making hypotheses as to the several distributions, and comparing the results with our observations, so as to derive the system of hypotheses which will best accord with what we learn from observation. No general result applicable to the totality of the stars, or to any portion of the universe outside our very limited means of measurement, can be reached by pure induction from the extremely imperfect results of observations which are so far available. …*

I am sending you ... some short and rather desultory papers of mine … especially [46]. *May I invite your attention to the results of this paper, especially as set forth in the last two problems? Can we not obtain a coefficient of correlation from the relations between parallax and proper motion given by the two equations* [46, pp. 168 – 169]

for any parallax π, *mean proper mot.* = 6.78π,

for any proper mot. μ, *mean parallax* = 0.064μ,

that will be more definite than any to be derived inductively from the observations?

I (1984, § 9.2.2) discussed this topic and provided the necessary references but I did not then see the Newcomb paper [46]. Now, I note, first of all, that no answer to the last-quoted letter is available, perhaps missing. Second, Newcomb's criticism carries weight. True, however, his remark about the knowledge of the relations between magnitudes, distances and parallaxes was difficult to understand: the concept of mean distance (or parallax) of the stars of a given magnitude hardly made sense. Then, before calculating the coefficients of some correlations one should decide whether he is dealing with random variables and specify his hypotheses about their laws of distribution.

Newcomb's desire to apply the newest statistical findings is commendable, but the correlation theory was then not sufficiently developed.

7.2. Natural Scientists and the Treatment of Observations. When treating observations, Newcomb often applied non-standard tricks (§§ 3.3, 3.4, 5.1) and I hold that he thus followed a certain tradition. Indeed,

l) Ptolemy is notorious for his free use or rejection of observations. However, if they are heavily corrupted by error, it might happen that their mean is no better than any one of them selected at random, recall the Cauchy distribution.

Ptolemy's cartographic work apparently testifies that he was attempting to present plausible results rather than secure mathematical consistency (Berggren 1991, pp. 135 – 136). A related fact pertains even to the Middle Ages (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'.

2) Kepler (Sheynin 1993, p. 167) adjusted the Tychonian observations by corrupting them by small arbitrary quantities and thus possibly made use of some elements of statistical simulation.

3) Fechner, who originated psychophysics, outlined a theory of treating observational series in natural sciences. His tools and his mathematical approach were primitive, but he influenced von Mises (1972, p. 26).

4) Gauss was a natural scientist as well as a mathematician. He certainly applied various roundabout methods; in § 3 I mentioned his remark about the possibility of a non-strict formation of the normal equations. Then, Gauss did not adhere to any definite rules when measuring angles in the field: he rather continued his work until feeling sure that further attempts were futile (Sheynin 1994, p. 263). Finally, Gauss did not unquestionably apply his own formula for estimating precision of observations. At least once, noting that the conditions of observation were the same for several stations, he added together the data for all of them and calculated a common mean square error (Ibidem, p. 266).

Modern statisticians and astronomers (Marsden 1995, p. 185) doubt that Gauss had actually applied (as he claimed he did) the MLSq before Legendre. I (1999, § 4.1) have somewhat disproved these doubts and I also indicated that several factors could have made any reconstruction of his calculations hardly possible. Now I emphasize the significance of one of those factors (use of short cuts) and stress that for Gauss the *application of least squares* was not at all a cut and dried procedure.

Acknowledgement. I am grateful to Professor Curtis Wilson who kindly sent me photostat copies of some of Newcomb writings and to the scientific bodies mentioned in Note 1 for permission to quote their archival sources.

Notes

1. Staatsbibliothek zu Berlin – Preussischer Kulturbesitz, Handschriftabteilung, Darmstaedter J 1871 (l1), Newcomb; University College London, No. 773/7. Below, I quote the letters from both these sources. Ludwig Darmstaedter (1846 – 1927) was a chemist and a collector of autographs some of which (usually, short autobiographies) were sent to him in response to his requests. Thus it apparently came about that the Newcomb papers in Berlin include the following form/letter (without date): "American Journal of Mathematics. Johns Hopkins University. Baltimore, MD. I like to make everyone happy when I can. Simon Newcomb". Nevertheless, his readers could have hardly felt themselves happy with his style and I do not agree with Brown (1910, p. 344) according to whom "clearness and freedom from unessential detail" characterized "everything" Newcomb had written.

2. Cf. Quetelet (1845, p. 225): "Le but principal de la statistique" is to render different materials comparable.

3. 1 quote from one of his last papers [53, p. 544]: "a serious problem is that of summing perhaps 100 periodic terms with coefficients not differing greatly from 0".01. I have devised a machine for this purpose the description of which must form the subject of another publication".

4. I should have considered his opinion [22, p. 20] when discussing the debates around that great astronomer (Sheynin 1993, § 3.8): " all of Ptolemy's <u>Almagest</u> seems to me to breathe an air of perfect sincerity".

5. His presidential address before the American Mathematical Society in 1893 testified that he had "devoted some thought to modern ideas on hyperspace, group theory, projective geometry, and the like" (Brown 1910, p. 351).

6. Here and below, the bibliographic information is chiefly due to Archibald (1924).

7. In a letter to F. A. T. Winnecke of Aug. 7, 1871, Newcomb wrote: "If you … want to laugh, read Proctor['s] *Sun* [that appeared in the same year], chapter 1". The same volume of *Nature* (vol. 4, 1871) contained Proctor's answers to Newcomb and in one of these he (p. 183) commented on Newcomb's determination of the solar parallax: "Not only I, but Sir John Herschel, as well as the Council of the Astronomical Society would seem to have done Prof. Newcomb less than justice."

8. The Congress was successful; from the scholars who reported there I mention Boltzmann and Kapteyn. Newcomb vainly invited Pearson to speak on "Methodology of science" (his letters to Pearson of June 27 and July 3, 1903), undoubtedly on the strength of Pearson's *Grammar of Science* published in 1892 and further work. The letters were written in spite of Pearson's earlier (June 26) refusal to come caused by his financial difficulties and fear of leaving his Department under "less complete supervision". At the time, Pearson held the chair of applied mathematics and mechanics; he became head of the Department of applied mathematics, University College London, in 1907.

9. In a few cases Newcomb [49, pp. 29 and 30] used other careless expressions as *method*, or *theories of probabilities.*

10. Instead of "expectation" Newcomb [l, vol. 3, p. 343] once used the unfortunate expression "value of probability".

11. Here and below, I use the Gauss notation

$$p_1x_1 + p_2x_2 + \ldots + p_nx_n = [px].$$

12. Whitworth (1901, p. 205) provided this formula; I have not seen the previous editions of his book.

13. Newcomb's final result, as estimated by Dorsey, was $v = 299.71 – 299.86 \cdot 10^3$ *km/sec* in vacuo. In 1924 – 1926, Michelson, who had also participated in Newcomb's work, arrived at $v = 299.774 \cdot 10^3$ *km/sec.*

14. The treatment of observations in [12] conformed to the classical error theory. Nevertheless, on June 27, 1903, in a letter to Pearson, Newcomb wrote that the latter "sixteen months ago" (apparently, in Pearson (1902)) "had developed so fully some ideas which I had enumerated in general form … but only incidentally, in a paper on the Right Ascensions of the Fundamental Stars" [12]. Pearson (1902, pp. 291ff) concluded that a very considerable correlation of judgements between two observers can arise even when they were working independently. He briefly discussed the law of error and the applicability of the "current theory of errors" without however changing it. The classical error theory is still with us.

15. I shall therefore call the relevant checks or tests, founded on reasonable considerations, qualitative. In most cases, quantitative criteria were not yet discovered.

16. Blunders and insufficiently definite registration corrupted many observations, especially in ancient times. Newcomb [20, pp. 402ff], for example, rejected seven out of the 19 lunar eclipses reported by Ptolemy, and concluded (p. 404) that even the records of Tycho Brahe of the same phenomenon "are so confused that it is impossible to obtain any definite result from them". Mendeleev (Sheynin 1999b, pp. 61 – 62) acted in a similar way.

17. Newcomb [35] was one of the first to use that term, normal distribution, cf. Sheynin (1984, p. 183, note 47).

18. After deriving the probable error characterizing his series by utilizing only its middlemost part, Newcomb [35, 36] compared his observations with appropriate normal distributions.

19. The latter term was called mean square error, which astronomers "generally designated … as the mean error" (Newcomb [53, p. 540]). He [30, p. 49] even made use of a loose and unfortunate expression "probable mean error" bearing in mind some reasonable common estimate of several mean square errors. He [48, p, 14; 52, p. 324] fared no better when introducing a "probable mean deviation" or even a "probable equation" (obtained by least squares) [52, p. 322].

20. When providing an example Newcomb (p. 359) stated that he chose four values of *h* "by several trials" and that, although the precision parameter was really a continuous magnitude, it was possible to keep to the discrete case.

21. In accord with a venerable but illogical tradition (Sheynin 1984, p. 183, note 47), Glasenapp calculated the errors to four significant digits!

22. Mendeleev applied the same criterion at least from 1860 onward (Sheynin 1996b, p. 64). From 1877 onward geodesists have been using the celebrated three-sigma rule.

23. On p. 166 Newcomb defined "as the unit of distance that of a star having a parallax of 1".

24. Newcomb (p. 166) also made use of the expression "usual exponential law of error".

25. He mentioned an Andrew Newcomb who had died "about 1650" but did not expressly connect himself with that person.

26. Later Newcomb [55, plate facing p. 210] reproduced these graphs.

27. It appeared in a collection of papers that also included Lobachevsky's *Pangéomètrie*.

28. Elsewhere Newcomb [9] had even written "probability of 589", "of 411", etc., but perhaps the German compositor paid no attention to the decimal dots without the zeros, likely inserted by Newcomb.

28. Appropriately referring to Laplace's memoir of 1786, Molina (1930, p. 386) remarked that the Master "had in mind the fundamental difference between the idea of a limit as used in pure mathematics" and in probability.

References

Simon Newcomb

1. Notes on the Theory of Probability. *Math. Monthly*, vol. 1, 1859, pp. 136 – 139, 233 – 235, 331 – 335; vol. 2, 1860, pp. 134 – 140, 272 – 275; vol. 3, 1861, pp. 119 – 125, 343 – 349.

2. On the Secular Variations and Mutual Relations of the Orbits of the Asteroids. Abstract. *Proc. Amer. Acad. Arts and Sciences*, vol. 4, 1860, for 1857 – 1860, pp. 417 – 418.

3. [Discussion of the Principles of Probability Theory]. Ibidem, pp. 433 – 440.

4. Solution of Problem. *Math. Monthly*, vol. 3, 1861, pp. 68 – 69.

5. Modern Theoretical Astronomy. *North Amer. Rev*., vol. 93, 1861, pp. 367 – 390.

6. On the Secular Variations and Mutual Relations of the Orbits of the Asteroids. *Mem. Amer. Acad. Arts and Sciences*, New ser., vol. 8, pt. 1, 1861, pp. 123 – 152.

7. Determination of the Law of Distribution of the Nodes and Perihelia of the Small Planets. *Astron. Nachr*., Bd. 58, 1862, pp. 210 – 220.

8. Investigation of the Orbit of Neptune. *Smithsonian Contr. to Knowledge*, vol. 15, 1867. Separate paging.

9. Comparison of the Actual and Probable Distribution in Longitude of the Nodes and Perihelia of 105 Small Planets. *Astr. Nachr*., Bd. 73, 1869, pp. 278 – 288.

10. On the Mode of Observing the Coming Transits of Venus. *Amer. J. Sci*., vol. 50, 1870, pp. 74 – 83.

11. [Reviews of Contemporary Popular Books]. *Nature*, vol. 4, 1871, pp. 41 – 43.

12. *On the Right Ascensions of the Equatorial Fundamental Stars*. Washington, 1872.

13. Investigation of the Orbit of Uranus. *Smithsonian Contr. to Knowledge*, vol. 19, 1874. Separate paging.

14. On the Probable Variability of the Earth's Axial Rotation As Investigated by Mr. Glasenapp. *Amer. J. Sci*., Ser. 3, vol. 8 (108), 1874, pp. 161 – 170.

15. On the Possible Periodic Changes of the Sun's Apparent Diameter. Ibidem, pp. 268 – 27. Co-author, Edw. S. Holden.

16. Some Talks of an Astronomer. *Harper's Mag*., vol. 49, 1874, pp. 693 – 707, 825 – 841.

17. Review of R. A. Proctor, Transits of Venus [1874, 1875, 1883]. *Nation*, vol. 20, 1875, p. 230.

18. Abstract Science in America, 1776 – 1876. *North Amer. Rev*. vol. 122, 1876, pp. 88 – 123.

19. Recent Astronomical Progress. Ibidem, vol. 123, 1876, pp. 86 – 122.
20. On the Mean Motion of the Moon. *Amer. J. Sci*., ser. 3, vol. 14 (114), 1877, pp. 401 – 410.
21. *Popular Astronomy*. New York, 1878.
22. Researches on the Motion of the Moon, pt. 1. *Wash. Observations* for 1875, 1878, Appendix 2.
23. Note on the Frequency of Use of the Different Digits in Natural Numbers. *Amer. J. Math*., vol. 4, 1881, pp. 39 – 40.
24. Discussion and Results of Observations on Transits of Mercury from 1677 to 1881. *Astron. Papers*, vol. 1, 1882, pp. 363 – 487.
25. Probability. In *Johnson's New Universal Cyclopedia*, vol. 3, pt. 2, 1884, p. 1426.
26. A Generalized Theory of the Combination of Observations. *Amer. J Math.*, vol. 8, 1886, pp. 343 – 366. Reprint; Stigler (1980, vol. 2).
27. Measures of the Velocity of Light. *Astron. Papers*, vol. 2, 1891, pp. 107 – 230.
28. Discussion of Observations of the Transits of Venus in 1761 and 1769. Ibidem, pp. 259 – 405.
29. On the Law and the Period of the Variation of Terrestrial Latitudes. *Astron. Nachr*., Bd. 130, 1892, pp. 1 – 6.
30. Remarks on Mr. Chandler's Law of Variation of Terrestrial Latitudes. *Astron. J*., vol. 12, 1892, pp. 49 – 50.
31. The Elements which Make Up the Most Useful Citizen of the United States. *Anthropologist*, vol. 7, 1894, pp. 345 – 351.
32. *The Elements of the Four Inner Planets and the Fundamental Constants of Astronomy*. Washington, 1895.
33. On the Value of the Precessional Constant. *Astron. J*., vol. 15, 1895, pp. 185 – 188.
34. On the Variation of the Personal Equation with the Magnitude of the Star Observed. Ibidem, vol. 16, 1896, pp. 65 – 67.
35. On the Solar Motion As a Gauge of Stellar Distances. Ibidem, vol. 17, 1896, pp. 41 – 44.
36. A New Determination of the Precessional Constant. *Astron. Papers*, vol. 8, 1897, pp. 1 – 76.
37. *The Problems of Astronomy*. N. p., Univ. of Pennsylvania, 1897.
38. A New Determination of the Precessional Motion. *Astron. J*., vol. 17, 1897, pp. 161 – 167.
39. Aspects of American Astronomy. *Annual Report Smithsonian Instn* for 1897, 1898, pp. 85 – 99.
40. Catalogue of the Fundamental Stars for the Epochs 1875 and 1900. *Astron. Papers*, vol. 8, 1898, pp. 77 – 403.
41. Remarks on Prof. Boss's Third Paper on the Precessional Motion. *Astron. J*., vol. 19, 1898, pp. 2 – 3.
42. On the Distribution of the Mean Motions of the Minor Planets. Ibidem, vol. 20, 1900, pp. 165 – 166.
43. On the Period of the Solar Spots. *Astrophys. J*., vol. 13, 1901, pp. 1 – 14.
44. On the Cordoba Durchmusterung and Some Conclusions Derived from It. *Astron. J*., vol. 22, 1901, pp. 21 – 26.
45. The Problem of the Universe. *Intern. Monthly*, vol. 5, 1901, pp. 395 – 417.
46. On the Statistical Relations among the Parallaxes and the Proper Motions of the Stars. *Astron. J*., vol. 22, 1902, pp. 165 – 169.
47. The Universe As an Organism. *Science*, New Ser., vol. 17, 1903, pp. 121 – 129.
48. *On the Position of the Galactic and Other Principal Planes toward Which the Stars Tend To Crowd. Carnegie Instn of Washington*, Publ. 10, 1904.
49. *Statistical Inquiry into the Probability of Causes of Sex in Human Offspring*. Ibidem, Publ. 11, 1904.
50. Evolution of the Scientific Investigator. *Annual Report Smithsonian Instn* for 1904, 1905, pp. 221 – 233.
51. Note on the Astronomical Value of Ancient Statements of Solar Eclipses. *Monthly Notices Roy. Astron. Soc*., vol. 66, 1906, pp. 34 – 35.

52. Search for Fluctuations in the Sun’s Thermal Radiation through Their Influence on Terrestrial Temperature. *Trans. Amer. Phil. Soc.*, New Ser., vol. 21, 1908, pp. 309 – 387.
53. Considerations on the Form and Arrangement of New Tables of the Moon. *Monthly Notices Roy. Astron. Soc.*, vol. 68, 1908, pp. 538 – 544.
55. Fluctuations in the Moon’s Mean Motion. Ibidem, vol. 69, 1909, pp. 164 – 169.
56. Researches on the Motion of the Moon, pt. 2. *Astron. Papers*, vol. 9, 1912, pp. 5 – 29.

Other Authors

Archibald, R. C. (1924), S. Newcomb, Bibliography. *Biogr. Mem. Nat. Acad. Sci.*, vol. 17, pp. 19 – 69.
Benford, F. (1938), The Law of Anomalous Numbers. *Proc. Amer. Phil. Soc.*, vol. 78, pp. 551 – 572.
Benjamin, M. (1910), S. Newcomb. In *Leading American Men of Science*. Editor, D. S. Jordan. New York, Holt, pp. 363 – 389.
Berggren, J. L. (1991), Ptolemy’s Map of Earth and the Heavens: a New Interpretation. *Arch. Hist. Ex. Sci.*, vol. 43, pp. 133 – 144.
Bond, G. P. (1857), On the Use of Equivalent Factors in the Method of Least Squares. *Mem. Amer. Acad. Arts and Sciences*, New Ser., vol. 6, pt. 1, pp. 179 – 212.
Brown, E. W. (1910), Newcomb. *Bull. Amer. Math. Soc.*, vol. 16, pp. 341 – 355.
Campbell, W. W. (1924), S. Newcomb. *Biogr. Mem. Nat. Acad. Sci.*, vol. 17, pp. 1 – 18.
Clausius, R. (1889 – 1891), *Die kinetische Theorie der Gase*. Braunschweig, Vieweg & Sohn.
Cournot, A. A. (1843), *Exposition de la théorie des chances et des probabilités*. Reprint: Paris, Vrin, 1984. Editor, B. Bru. **S, G,** 54.
Dorsey, N. E. (1944), The Velocity of Light. *Trans. Amer. Phil. Soc.*, New Ser, vol. 34, pt. 1 (the whole part).
Eisenhart, C. (1964), The Meaning of “Least” in Least Squares. *J. Wash. Acad. Sci.*, vol. 54, pp. 24 – 33.
Faddejew, D. K., Faddejewa, V. N. (1970), *Numerische Methoden der linearen Algebra*. München – Wien, Oldenburg. Orig. publ. in Russian, 1960 and 1963. First German edition, 1964.
Farebrother, R. W. (1999), *Fitting Linear Relationships*. New York, Springer.
Feller, W. (1971), *Introduction to Probability Theory and Its Applications*, vol. 2. New York, Wiley. Second edition.
Fisher, I. (1909), Obituary. S. Newcomb. *Econ. J.*, vol. 19, pp. 641 – 644.
Gauss, C. F. (1809), *Theoria motus*. German transl. in author’s *Abhandlungen zur Methode der kleinsten Quadrate*. Hrsg. A. Börsch, P. Simon. Berlin, 1887, pp. 92 – 117. Latest reprint: Vaduz, Sändig, 1998. English translation: 1865, 2009.
Gauss, C. F. (1816), Bestimmung der Genauigkeit der Beobachtungen. Ibidem, pp. 129 – 138.
Gentle, J. E. (1998), *Numerical Linear Algebra for Applications in Statistics*. New York, Springer.
Helmert, F. R. (1872), *Ausgleichungsrechnung nach der Methode der kleinsten Quadrate*. Leipzig, Teubner. Two more editions (1907 and 1924).
Idelson, N. I. (1947), *Sposob Naimenshikh Kvadratov* [Method of Least Squares]. Moscow, Geodezizdat. In Russian. **S, G,** 58.
Jacobi, C. G. J. (1841), *Über die Bildung und die Eigenschaften der Determinanten*. Leipzig, Engelmann, 1896. Orig. published in Latin.
Kowalski, M. (1856), Recherches sur les mouvements de Neptune. *Sbomik Statei Imp. Kazan Univ.*, vol. 1, pp. 97 – 278.
Laplace, P. S. (1812), *Théorie analytique des probabilités. Oeuvr. compl.*, t. 7. Paris, Gauthier-Villars, 1886.
Marsden, B. G. (1974), Newcomb. *Dict. Scient. Biogr.*, vol. 10. New York, Scribner, pp. 33 – 36.
Marsden, B. G. (1995), 18th and 19th Century Developments in the Theory and Practice of Orbit Determination. In *General History of Astronomy*, vol. 2B. Editors, R. Taton, C. Wilson, Cambridge, Univ. Press, pp. 181 – 190.

Meyer, Hugo (1891), *Anleitung zur Bearbeitung meteorologischer Beobachtungen.* Berlin, Springer.
Mises, R. von (1972), *Wahrscheinlichkeit, Statistik und Wahrheit.* Wien – New York, Springer. First edition, 1928. English translation: New York, 1981.
Molina, E. C. (1930), The Theory of Probability: Some Comments on Laplace's *Théorie analytique. Bull. Amer. Math. Soc.*, vol. 36, pp. 369 – 392.
Morando, B. (1995), The Golden Age of Celestial Mechanics. In *General History of Astronomy*, vol. 2B. Editors, R. Taton, C. Wilson. Cambridge, University Press, pp. 211 – 239.
Novikov, S. P. (2000), Pseudo-History and Pseudo-Mathematics: Fantasy in Our Life. *Uspekhi Matematich. Nauk*, vol. 55, pp. 159 – 161. In Russian. From 1945, this periodical is being translated as *Russian Math. Surveys*.
Paul, E. R. (1993), *The Milky Way Galaxy and Statistical Cosmology 1890 – 1924*. Cambridge, University Press.
Pearson, K. (1894), Science and Monte Carlo. *Fortnightly Rev.*, New Ser., vol. 55, pp. 183 – 193.
Pearson, K. (1902), On the Mathematical Theory of Errors of Judgement with Special Reference to the Personal Equation. *Phil. Trans. Roy. Soc.*, vol. A198, pp. 235 – 299.
Poincaré, H. (1912), Calcul des probabilités (1896). Reprints of second edition of 1912: 1923 and Paris, Gabay, 1987.
Price, D. J. (1955), Medieval Land Surveying and Topographical Maps. *Geogr. J.*, vol. 121, pt. 1, pp. 1 – 10.
Quetelet, A. (1845), Sur l'appréciation des documents statistiques. *Bull. Comm. Centr. de Statistique Belg.*, t. 2, pp. 205 – 286.
Raimi, R. (1976), The First Digit problem. *Amer. Math. Monthly*, vol. 83, pp. 521 – 538.
Seidel, L. (1863), Resultate photometrischer Messungen. *Abh. math.-phys. Kl. Kgl. Bayer. Akad. Wiss. München*, Bd. 9, No. 3, pp. 419 – 609.
Sheynin, O. B. (1963), Adjustment of a Trilateration Figure by Frame Structure Analogue. *Survey Rev.*, vol. 17, pp. 55 – 56. **S, G,** 110.
Sheynin, O. B. (1977), Laplace's Theory of Errors. *Arch. Hist. Ex. Sci.*, vol. 17, pp. 1 – 61.
Sheynin, O. B. (1978), Poisson's Work in Probability. Ibidem, vol. 18, pp. 245 – 300.
Sheynin, O. B. (1982), On the History of Medical Statistics. Ibidem, vol. 26, pp. 241 – 286.
Sheynin, O. B. (1984), On the History of the Statistical Method in Astronomy. Ibidem, vol. 29, pp. 151 – 199.
Sheynin, O. B. (1989), Markov's Work in Probability. Ibidem, vol. 39, pp. 337 – 377; vol. 40, p. 387.
Sheynin, O. B. (1993), Treatment of Observations in Early Astronomy. Ibidem, vol. 46, pp. 153 – 192.
Sheynin, O. B. (1994), Gauss and Geodetic Observations. In this collection.
Sheynin, O. B. (1995), Density Curves in the Theory of Errors. *Arch. Hist. Ex. Sci.*, vol. 49, pp. 163 – 196.
Sheynin, O. B. (1996a), *Chuprov. Life, Work. Correspondence*. Göttingen, 2011. Vandenhoeck & Ruprecht. Orig. publ. in Russian in 1990.
Sheynin, O. B. (1996b), Mendeleev and the Mathematical Treatment of Observations in Natural Science. *Hist. Math.*, vol. 23, pp. 54 – 67.
Sheynin, O. B. (1999), Discovery of the Principle of least Squares. *Hist. Scientiarum*, vol. 8, pp. 249 – 264. **S, G,** 112.
Stigler, S. M. (1973), Simon Newcomb, Percy Daniel] and the History of Robust Estimation 1885 – 1920. *J. Amer. Stat. Assoc.*, vol. 68, pp. 872 – 879. Reprinted (1977) in *Studies in the History of Statistics and Probability*, vol. 2. Eds, Sir Maurice Kendall & R. L. Plackett. London, Griffin, pp. 410 – 417.
Stigler, S. M. (1978), Mathematical Statistics in the Early States. *Annals Stat.*, vol. 6, pp. 239 – 265.
Stigler, S. M., editor (1980), *American Contributions to Mathematical Statistics in the 19th Century*, vols. 1 – 2. New York.

Whitworth, W. A. (1867), *Choice und Chance*. New York, Hafner, 1959 (reprint of the edition of 1901).
Zaliznyak, A. A. (2000), Linguistics according to A. T. Fomenko. *Uspekhi Matematich. Nauk*, vol. 55, pp. 162 – 188.

Afterword

Comment on § 7.1-2. See Methods for promoting research in the exact sciences. Carnegie Instn of Washington, *Yearbook* No. 3 for 1904, 1905, pp. 179 – 193.

On § 7.1-3. Newcomb's reasoning applies to science in general rather than only to astronomy. It is a destructive argument against the empirical approach to science characteristic of Pearson and should be known to historians of statistics and probability.

IV

Mathematical Treatment of Astronomical Observations (Historical Essay)

Archive Hist. Ex. Sci., vol. 11, 1973, pp. 97 – 126

1. Introduction

Mathematical treatment of astronomical observations has been considered by numerous authors, but usually within the framework of astronomy proper. I consider this subject from the point of view of the classical theory of errors and mathematical statistics.

The classical theory of errors is separated into two parts, the determinate and stochastic. The former aims at securing most reliable experimental results by means of specially designed experiments (by the choice of best time intervals and other conditions for astronomical observations). Formally, this is the theory of prior calculation of errors of functions of measured quantities, which properly originated with the differential calculus. It is now included in the prehistory of the design of experiments and it is precisely in this, the determinate part of the theory of errors, that various errors of measurements were first mentioned, implicit statements about the inevitability of errors in general were made and methods for their elimination described.

Some authors[1] believe that the bifurcation of errors into random and systematic is crude and cannot be justified from the metrological point of view. However, it is extremely important at least for other branches of science as well as for a historian because, intuitively understood, it led to the first notions of the ways in which various errors influence measurements. I shall use this bifurcation of errors understanding the systematic errors as those that, depending upon the circumstances of observation, vary according to one or another determinate but unknown law (or, in particular, remain constant). This division is in accord with the point of view of at least some physicists[2] and was adapted by GAUSS[3] who denoted the two kinds as *irregulares seu fortuiti* and *constantes seu regulares.*

A detailed description of observational errors would have been a research in itself, hardly possible here. One such research pertaining to antique astronomy is due to P. COLLINDER[4], to whom I refer below. Also, some important primary sources pertaining to the 17th and 18th centuries are mentioned in § 4.

§ 2 is devoted to the design of experiments (in the meaning explained above), § 4, to the stochastic treatment of observations while § 3 is an account of the “intermediate” problem of selecting observations. Even from this explanation it is evident that such a separation is not devoid of defects and, in particular, the general chronological description becomes divided between these main sections. A synopsis (§ 5) may serve as a means for overcoming this division.

The subject-matter of the main sections is carried through to the middle of the 18th century, and a sketch of subsequent events is given. It is my belief that the general outline of this paper with prominence given to such scholars as PTOLEMY, AL-BIRUNI and KEPLER may be considered sufficiently new and that the argumentation thereof as sufficiently sound.

2. Design of Experiments

2.1. Ptolemy. The simplest notions pertaining to the determinate part of the theory of errors are found in PTOLEMY'S *Almagest*[5].

(a) PTOLEMY emphasized the need of applying observations separated in time (Book 3, § 1, pp. 78 and 81):

The sun's year is found by as many observations as possible taken over a rather long interval.

The period of return will be gotten the more accurately the longer the time between the observations compared (because the error of the length of the year is divided by the corresponding time interval).

(b) PTOLEMY repeatedly noticed the preference of one method of observation and/or one set of circumstances of observation over another and, also, the shortcomings of certain types of observation. To quote him (Book 4 § 1 p. 108; Book 5, § 2, p. 144; Book 8, § 6, pp. 269 and 271 respectively):

1. *It is only by means of* (observations of lunar eclipses) *that the positions of the moon can be found in an accurate way, since the other kinds of observations … can, because of the moon's parallaxes, be very deceptive.*

2. *At conjunctions and full moons there was either little or no appreciable discrepancy and only such as the moon's parallaxes could account for; and ... about the first and third quarters ... most* (of the discrepancy occurs) *when the moon in its mean courses effects the greatest difference of first anomaly … we see … that it is necessary to suppose the moon's epicycle is borne on an eccentric circle ...*

3. *The observers and the atmosphere for the places observed can make the time of the first glimpse unlike and unsure* (*as I know from trial and from the differences in the observations*).

4. *The same angular distances appear to the eye greater near the horizon and smaller near the culminations.*

(c) PTOLEMY also demanded that various errors (actually, both systematic and random) be eliminated as completely as possible (Book 4, § 9, p. 135 – 136):

In the correction of the (moon's) *mean course we first looked for lunar eclipses from among those accurately recorded, as far apart in time as possible, in which the magnitudes of the shadows were equal, near the same node with both shadows either on the southern or northern side, and in which, moreover, the moon was at the same distance from the earth.*

However, continues PTOLEMY (pp. 136 – 137), eclipses were also used where the nodes were no longer the same but opposite. Elsewhere (Book 9, § 2, p. 273) he states:

We have used for the demonstrations of each planet only those observations which cannot be disputed, that is those taken at contact or great proximity with the stars …, and above all those taken with the astrolabe where the eye is lined up with the diametrically opposite sights in the circles, sees on every side equal angular distances by means of similar arcs.

In particular, PTOLEMY noticed (Book 6, § 9, p. 215) that

The periodic return (in HIPPARCHUS' observations) *would have erred by the* 2° *of both mistakes together if both had happened to involve a difference towards the greater or towards the less. But … by happy chance the one makes the return fail of completion and the other makes it exceed …*

PTOLEMY himself would have therefore designed his observations so as not to rely on *happy chance*.

(d) Lastly, PTOLEMY repeatedly considers the effect of various (systematic) errors (Book 5, § 10, p. 162, Book 6, § 9, p. 214 and Book 9, § 2, p. 271, respectively):

1. *It is reasonable to suspect that at times an appreciable difference occurs at the conjunctions, full moons, and eclipses accompanying them, because of the moon's eccentric circle … we shall try and show that* (though the epicycle's centre does not always fall exactly on the eccentric's apogee, this difference) *can produce no error worth mentioning*.

2. *The time from the beginning of the eclipse to the middle is not always equal to the time from the middle to the end. … But supposing these times not unequal would work no discrepancy perceptible to sense in the appearances.*

3. *The stations cannot indicate the exact time, since the planet's local motion remains imperceptible; and the apparitions not only make the places immediately disappear along with the stars as they are seen for the first or last time, but also can be utterly misleading because of the differences in the atmosphere and in the eye of the observer*.

One more passage on systematic errors from the *Almagest* is quoted below.

PTOLEMY did not distinguish explicitly between random and systematic errors, a bifurcation due to D. BERNOULLI[6]; neither did he formulate the stochastic properties of random errors. Lastly, he did not devote a special section of his book to this subject. However, PTOLEMY did possess a clear-cut notion about various errors of observation and about the different effects which they produce and did leave proposals for selecting methods and circumstances of observations and for combining different observations.

A. AABOE and D. J. DE SOLLA PRICE[7] considered antique astronomical observations:

In the pre-telescopic era there is a curious paradox that even a well-graduated device (their estimate of the error of graduation is 5') *for measuring celestial angles is hardly a match for the naked and unaided eye judiciously applied* when, for example, an observation consists of registrating that a certain planet *falls so-and-so many moonwidths from some other star or from … the midpoint of a line joining two stars.*

The function of smaller antique instruments, they suppose, was to serve as a means for avoiding calculations while

The characteristic type of measurement depended not on instrumental perfection but on the correct choice of crucial phenomena.

This is an extremely interesting article. However, I object to their regard of antique observations as *qualitative*. Counting the number of moonwidths (see above) is, after all, a purely *quantitative* procedure. Also, neither the authors nor COLLINDER[8] who refers to their article, notices that, as regards the effort toward accuracy, observational methods in antique astronomy do not present any peculiarity. In particular, if PTOLEMY did not follow the principle of regular observations (the authors' conclusion), he at least strove for accurate results by other methods at his disposal. However, PTOLEMY does mention regular observations, including those of HIPPARCHUS (Book 3, § 1, p. 78 of the *Almagest*):

But since a suspected inequality in the periods of … this return (of the sun), *suspected through continuous and successive observations, more or less worried Hipparchus, we shall try … to show … by the continuous … observations we have made … that these periods are not unequal. For we find them differing by no appreciable amount from the additional quarter day, but at times by about as much as could be attributed to the error due to the construction or position of the instruments.*

A passage on systematic errors continues thus (p. 79):

Even if the position or discrimination of the instruments is inaccurate by only 1/3,600 *of the circle … the sun makes up for this advance in latitude by shifting* 1/4° *in longitude along the ecliptic*.

For some twenty centuries PT0LEMY’S observations have been an object of discussion. This is the first reason why I discuss KEPLER’S opinion of PTOLEMY. The second is, of course, KEPLER’S own position in experimental science.

First, KEPLER[9] appreciates PT0LEMY’S observations:

Wenn ich aber … geglaubt habe, Ptolemäus sei bei seiner Annahme einer blinden Vermutung gefolgt, so verhält sich dies anders. Denn er hätte sie mit einem vortrefflichen Beweis auf Grund einer geeigneten Beobachtung erhärten können, … nur das möchte man bei dem Meister vermissen, dass er nicht jene Beobachtungen mit einem Beweis der Nachwelt überliefert hat.

Even if evidence is lacking, KEPLER[10] is ready to support PTOLEMY (and HIPPARCHUS):

Ich kann mich freilich nicht davon überzeugen, dass Hipparch und Ptolemäus ihr Augenmerk auf den Moment des Eintritts (der Sonne in den Krebs) *selber gerichtet haben, ohne auf die Zwischenpunkte zu achten. Ich glaube eher, dass sie den ganzen Sommer über fleißig die Deklinationen der Sonne miteinander verglichen und den zeitlichen Mittelpunkt zwischen den Momenten gleicher Deklination für den wahren Eintritt … gewählt haben.*

At least this would have been KEPLER'S own mode of action. See also similar methods of treating observations by AL-BIRUNI (§ 2.2). To return to KEPLER’S account (p. 387):

Wenn er (PTOLEMÄUS) *also auch nur eine einzige Beobachtung überliefert, um die Methode aufzuzeigen* (!), *so dürfen wir doch glauben, dass er mehrere Beobachtungen in Betracht gezogen hat*.

However, KEPLER simultaneously accuses PTOLEMY of following presupposed ideas and even, in one case, of falsification

(Kap. 14, pp. 132 – 133, Kap. 66, pp. 373 – 374 and Kap. 70, p. 396, respectively):

1. *Um Fehler durch Fehler zu stützen, verrenkte er seinen Epizykel aus der parallelen Lage und wählte sich, nicht im Vertrauen auf Beobachtungen, die ihm in nicht sehr großer Zahl vorlagen, noch nach den Größenangaben, wo Beobachtungen vorlagen* (*da er ihrer Zuverlässigkeit misstraute*), *mittlere Werte aus, indem er extreme Werte im Zweifel zog*.

2. *Ptolemäus sowie seine Nachfolger höchst verwickelte Neigungs-, Beugungs- und Drehungsbewegungen ausgedacht haben. Das heißt aber nicht, die Wahrheit durch Beobachtung ermitteln, sondern vielmehr nach einer falschen vorgefassten Vorstellung Beobachtungen erfinden. Man muss ihm das freilich hingehen lassen, weil ihm nur wenige Beobachtungen zur Verfügung standen.*

3. *Man … bedenke, dass Ptolemäus das Verhältnis seiner Bahnen so gefälscht hat, dass diese Beobachtung gedeckt ist*.

KEPLER'S account with its declarations in favour of PTOLEMY and the accusatory ones, with their *freilich hingehen lassen*, seems to be accurate. One may well ask how to reconcile PTOLEMY'S regular observations with *ihm nur wenige Beobachtungen zur Verfügung standen.* A possible explanation is that most observations (the worst ones, which PTOLEMY *mißtrauent*) were usually made for corroboration only (see the opinion of AABOE and DE SOLLA PRICE above and, also, AL-BIRUNI'S mode of action, § 3).

2.2. Al–Biruni. Statements that regular observations had been made or were needed occur also in AL-BIRUNI[11]. In several places this author tells us about his own regular observations; on p. 65 he testifies that AL-BATTANI declared that he had repeated his observations over many years. Possibly more interesting though is that AL-BIRUNI maintains that regular observations are needed so as to predict dangerous landslides etc., an aim far from possible even nowadays (p. 32):

Latitudes may be changed sensibly by that movement (of masses over the earth's surface) … *or a dangerous displacement may be produced which can cause havoc and destruction. Therefore latitudes should be continually observed and examined*.

It is my opinion that, with the naked and unaided eye (§ 2.1) or otherwise, regular observations had been a common feature even of antique astronomy. However, the observer usually enjoyed an unrestricted liberty about the use or tacit rejection of any of his observations (§ 3). As to the avoidance of calculations in antiquity (§ 2.1), AABOE and DE SOLLA PRICE seem to be completely right. As even AL-BIRUNI says, calculations are difficult and introduce additional errors, besides (a new feature), they are forbidden by the Islamic law. So (p 51), measurements are only approximate, and the

Approximation is also due to the extraction of square roots in the calculation … and the lack of refined methods for the calculation of some quantities … hence very minute defects, in operations including sines, lead to defective composite approximations. I do not use (methods requiring computations) *except for exploring the truth* *behind the veils, and for comparing results arrived at by different methods, to feel more confident about a derived result.*

Similar statements are made at least in seven other places (pp. 58, 96, 105, 115, 152, 191 and 237), and I quote from the first and the fifth of them, respectively:

1. *The first method is more reliable, because it depends on direct observation and does not involve any computation.*

2. *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.*

It seems that AL-BIRUNI was the first to reason on the propagation of computational (random!) errors (see also § 4) and on the combined effect of observational and computational errors, though of course he was unable to evaluate numerically such combined effects.

Lastly comes the passage about the Islamic law (p. 259):

When one determines accurately the longitude and the latitude of one's town, one can compute the (moment of the) *rise of the dawn which ushers in the beginning of the fast. ... Further, one can determine the times for the visibility of the new moons, though the Islamic law commands their determination with the naked eye and not by computation, because the Prophet said: "We are people who neither write nor compute." Hence, the month is so, and so, and so, showing out his ten fingers thrice.*

Hardly any astronomer wholly complied with such restrictions, but at least here is an additional general argument against computations. Now I notice additional passages from Al-BIRUNI who also **(a)** demands elimination or decrease of systematic errors, **(b)** compares the accuracy of different methods of observation, and **(c)** reveals an error of astronomical calculations due to the inaccuracy of the mathematical model used.

(a) On p. 39 AL–BIRUNI comments on time intervals between observations:

The wider the intervals … the more reliable the (end) *result.* A similar passage from the *Almagest* is in § 2.1.

On pp. 76 – 77 Al-BIRUNI discusses the consequences of an instrumental error and concludes that observers *should keep alert*, pursue research *without impatience or boredom* etc.

Bearing in mind determinations of longitudinal differences between cities he notices (p. 155) that both

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.*

AL-BIRUNI is here concerned with obtaining comparable results *from every pair of opposite times*, a problem methodologically similar to that of isolating systematic influences from repeated measurements (§ 2.4). Also, *the impression* of the lunar eclipse in both cities, AL-BIRUNI concludes on p. 131, is almost or completely identical, which means that systematic errors of observation are almost eliminated. He also comments on the elimination of systematic errors from distances between cities (p. 199):

By common consent, the calculators increase the distance by its sixth. It should not necessarily be so, because the amount of this increase depends on deviations whose number is not definitely known, and whose extents are also undetermined.

He himself increases the distances concerned by various amounts rather than by a predetermined fraction, which more consistent with common sense (p. 191):

He, who has experienced the technicalities of observations, knows that the correction (of terrestrial longitudes) *by the method of distances, when they are assessed carefully by distinguishing between level and mountainous ones, and by studying the nature of the slopes, the number of curves und the extent of their curvatures, – if it is not superior to the correction obtained by observation of lunar eclipses, it is not inferior to it.*

(a) and **(b)** Discussing observations of eclipses, AL–BIRUNI (p. 129) concludes that, because of various errors (mostly systematic) peculiar to solar eclipses, one ought to prefer observations of lunar eclipses.

(c) Discussing certain solar observations, AL-BIRUNI (pp. 115 – 116) notices that

The discrepancy between (one of them) *and the established amount of the constant sought is intolerable. It is partly due to the assumed equality between* (arcs), *on account of the equality of the two time intervals* (between observations). *That equality cannot hold unless the middle observation is made when the sun is exactly in apogee, or in perigee.*

2.3. India (12th century). This is a short aside from astronomy.

An Indian writing[12] contains a rule for calculating the volume of an irregular earth excavation. The volume is considered equal to the product of the mean measures of the length, width and depth of the excavation with the measures taken at different places. A 16th century comment by GÁNÉZA is that

The greater the number of the places (of measurement), *the nearer will the mean measure be to the truth and the more exact will be the consequent computation*.

Another comment by the same author is that the rule applies to excavations, whose sides are trapezia. Whatever is meant by the sides, I am inclined to suppose that GÁNÉZA thought of irregular trapezia. The rule itself seems to contain no stochastic reasoning, but it is very possible that even ancient scholars did understand the general idea of taking the mean, and adjusting observations of a series of quantities (not of one and the same quantity) to decrease the influence of systematic errors and of the inexactitude of the mathematical model.

Applications of mean measures in calculations of areas can be traced to ancient Babylonia[13] where the area of a quadrangle was held to equal the product of the half-sums of its opposite sides. This method of calculation was used rather often (Ibidem) in cases either of inexact rectangles (inexactitude of the mathematical model) or of rough terrain with the measures of the opposite sides unequally influenced by systematic errors.

GÄNÉZA’S commentary is considered in § 4.

2.4. European History. I pass to the 16th and 17th centuries and describe considerations of different scholars, but see those of NEWTON in Sheynin (1971), this *Archive*, vol. 7.

W. GILBERT[14] noticed the influence of temperature on one of his experiments:

This experiment is best [better] *made in winter and in a cold atmosphere ... than in summer and in warm climates.*

Elsewhere in the same writing (Book 4, chap. 12, p. 86) he mentions some errors of measurement:

The whole trick consists in proper use of the instruments by which the sun's position is ascertained … for either the hand trembles, or the eyesight is defective, or the instrument does not work aright.

GALILEO[15] several times reasons on the methodology of experiments; in particular, he takes up experiments on the free fall of bodies[16]. Also, comparing voyages over the Mediterranean, he[17] says:

Keeping a special record and account of the days of departure and arrivals of ships ... I discovered … that the returns here (at Venice) *were made in proportionately less time than those in the opposite direction, in a ratio of 25%. Thus we see that on the whole the east winds are stronger than those from the west.*

His inference does not seem penetrating, but it is obviously made in accord with the rules of the then non-existent classical theory of errors (mathematical treatment of the so-called *Wiederholungs*-measurements with the isolation of the mean value of a systematic influence, of the unequal influence of winds on this occasion) and his end result, on which it is of course difficult to check, is stated with a reasonable degree of accuracy (25 rather than, e. g., 26%).

But by far more interesting example is GALILEO'S treatment of sun spots, see § 8.2 of *Prehistory*… in this collection.

Some of KEPLER'S arguments from his *Neue Astronomie* are described in § 2.1. Two other passages from the same source (Kap. 28, p. 209 and Kap. 51, p. 311, respectively) testify also that he was concerned with the design of experiments:

1. *Denn auch bei der herkömmlichen Art, wie man auf der Erde die Entfernungen von Dingen misst, erhält man die Entfernung eines Punktes um so sicherer, je weiter die Standpunkte voneinander abstehen.*

It would have been more accurate to discuss the shape of the triangle formed by the three points, or the ratio of its sides.

2. *Alle drei Beobachtungen sind angestellt worden, als Mars im Osten stand, keine, wenn er im Westen stand. Es fehlen nämlich weitere Beobachtungen.*

Other places of interest are two of KEPLER'S letters to HERWART[18] and, also, Part 3 of his *Neue Stereometrie*[19] where he offers recommendations about the measurement of the content of barrels. Lastly, his *Somnium*[20] contains recommendations to observers of solar eclipses:

Observers should … be warned that the paper which receives the image of the eclipsed sun must be protected from all disturbances and must always be placed at the same distance from the hole and at right *angles to the ray coming through it. For if the paper bends, the circumferences of the bright image are distorted, and degenerate from circles into ellipses. Accordingly, let the disputant verify whether he took adequate precaution against this defect.*

C. HUYGENS studied possibilities for decreasing errors of clockworks and for more accurately observing the free fall of bodies (Ibidem, Part 4, Proposition 26).

Rules for eliminating systematic errors are also given in writings pertaining to nautical astronomy which became extremely important by the beginning of the era of great geographical discoveries (end of the 15th century). Thus[22],

For knowledge of the true height of the Sun (the Astrolobe not hanging upright), do thus: if the Astrolobe be truely marked, marke the diversitie, that being knowne, rebate from the greatest heigth halfe the diversitie, or else adde unto the lesser heigth halfe the diversitie, and that shall be the true heigth of the Sunne, although that the Astrolobe doth not hang upright.

The foregoing description deals with the prehistory of my subject, and it is only with the advent of the differential calculus that the real history begins. The first special research is due to R. COTES[23], who calculated the errors of the different elements of plane and spherical triangles calculated from their observed (and thus error-burdened) angles and sides. For the era of arc measurements which commenced in the first half of the 18th century, his writing appeared extremely timely and, for example, C. M. DE LA CONDAMINE[24] referred to COTES:

Après d'assez longues reeherches, aux quelles j'ai appliqué la théorie de M. Cotes ... je me suis convaincu ...
(a passage on the propagation of errors follows).

The further history of this subject is directly connected with the general geodetic activity in 19th century Europe and should be properly considered in the framework of the history of geodesy.

3. Selection of Data

3.1. Ptolemy. According to modern standards, each observation, including those rejected, should be recorded and dealt with to ensure the highest possible degree of objectivity. The history of selecting observations begins at least from PTOLEMY who not only tolerated but even recommended to select the best observations.

Thus he[25] will

Demonstrate this lunar anomaly … first using three of the oldest eclipses and then again three from the present very accurately observed by ourselves.

Also, in other places (Book 4, § 6, p. 123 and Book 10, § 4, p. 316):

1. *Of the three eclipses we have chosen from those most carefully observed by us …*

2. ***F****or the periodic movements of the star ... we took two sure observations from among our own and from among the old ones.*

Did PTOLEMY select observations just for confirmation of observations made by HIPPARCHUS? Some scholars thought so[26], although KEPLER (see § 2.1) was of a different opinion. In any case, if this was sometimes PTOLEMY'S goal, it would mean he had a higher opinion of HIPPARCHUS' observations than of his own. Moreover, as implicitly follows from § 2.1, his general goal seems to have been to select observations least influenced by random and systematic errors or, to put it otherwise, to reject inferior observations capable of corrupting the end results.

PT0LEMY'S correct point of view gave birth to an established tradition of freedom, excessive from a modern point of view, of selecting observations while leaving the rejected (= unused) observations unknown to anyone except the astronomer himself.

3.2. China (8th Century). It seems that a similar selection of observations was carried out in the 8th century by Chinese astronomers[27] during

A large-scale attempt to define terrestrial units (*li*) *of measure in terms of an unvarying astronomical or geodetic constant* (an early attempt at a metric system!) … *the ground distances between the more distant stations were not measured but assessed by extrapolation on the basis of the result for the short … line of stations.*

NEEDHAM similarly assessed the astronomical part of the work and concluded (p. 51):

In all probability I-Hsing (one of the astronomers) *thought it very undesirable to admit a mass of raw data showing considerable scatter, and not being able to assess it statistically, he used it only to satisfy himself that his calculated values came about were they should, indeed he probably believed that they were much more reliable than most of the observations*
(of the worst observations?).

3. Falsification of Observations. Another source of falsification (in the modern sense) of observations is mentioned by AL-BIRUNl[28] whose account (pp. 169 – 170) seems to indicate that this source had been far from unusual:

I have found in some books that the scientists measured the longitudes of towns by observations of eclipses. … I am not sure whether this is a genuine report of what was obtained by observation, or whether it is just an example made up for illustration, after the longitudinal difference had been obtained (by other methods).

I emphasize, though, that astronomers hardly aimed at deliberate deception; rather, they followed established traditions, and it is in this context that the possible veneration of HIPPARCHUS' observations by PTOLEMY should be seen (see also § 3.1). Of course I do not discuss extreme cases such as the one reported by AL-BIRUNI in another source[29]:

Abu Muhammad al-Nasafi (?) *pretended* (?) *that he had made observations while, in fact, he is a plagiarising liar and an impostor to the craft* (of astronomy).

The question marks were inserted by the translators.

It is instructive in this connection to compare astronomy with experimental science in general[30]:

The scientific literature of the 17th century, and not only of the 17th century, is full of … fictitious experiments … that have not been made, and are even impossible to make.

This comment is made in connection with PASCAL who has not given a complete and exact account of the experiments that he made or imagined. The author does not deny that PASCAL'S (hydrostatic) experiments were actually carried out, but he supposes that as compared with the account they were made on a rather modest scale.

3.4. European History (17th and 18th Centuries). Speaking of these centuries, one naturally takes notices of FLAMSTEED[31]:

He does not appear to have taken the mean of several observations for a more correct result where more than one observation of a star has been reduced, he has generally assumed that result which seemed to him most satisfactory at the time, without any regard to the rest. Neither, in fact, did he reduce the whole (nor anything like the whole) of his observations: many day's work having been wholly omitted in his computation-book. And, moreover, many of the results, which have been actually computed … have not been inserted in any of his MS catalogues.

This is trustworthy evidence, but in my opinion it does not contain the whole truth. First, FLAMSTEED experienced hardships[32], a fact which should well be taken into account. Second, foiling NEWTON's expectations, FLAMSTEED never hurried to publish his observations[33], and at least in one instance[34] he intended to use his observations for his own *private use*, then *resolved to communicate them to the ingenious* (rather than to the scientific community at large!). Third, FLAMSTEED[35] repeatedly emphasized the need to have trustworthy observations:

I … give you the sun's diameters, of which I esteem the first, third and fourth too large, by reason of my impractisedness … the rest I esteem very accurate, yet will not build upon them till I have made some further trials with an exacter micrometer.

The general impression seems to be that, in cases noticed by BAILY, FLAMSTEED just did not consider his work finished. A particular case of the use of the arithmetic mean by FLAMSTEED is described by R. L. PLACKETT[36] who, partly following J. L. E. DREYER[37], describes the practice adopted by TYCHO BRAHE. It occcurs that, like FLAMSTEED, TYCHO disregarded some of his observations.

On the other hand, BRADLEY never failed to consider all of his observations; in one case[38] he even calculated the mean of 120 of them, although this was not his general rule[39]:

When several observations have been taken of the same star within a few days of each other, I have either set down the mean result, or that observation which best agrees with it.

However, one way or the other, BRADLEY seems to be the first staunch originator of the idea of using whole sets of observations, and he said just this when introducing his new discovery, the phenomenon of nutation (Ibidem, p. 17):

This (discovery) *points out to us the great advantage of cultivating* (astronomy) *as well as every other branch of natural knowledge, by a regular series of observations and experiments.*

BRADLEY's point of view did not immediately become universally accepted, a fact proved by the appearance of T.

SIMPSON's memoir[40] whose aim was *to refute … some persons, of considerable note,* (who) *have been of opinion, and even publickly maintained, that one single observation, taken with due care, was as much to be relied on as the mean of a great number* (of them).

SIMPSON did prove that, at least for the uniform and triangular distributions, the arithmetic mean is preferable to a single observation[41]. However, the first reasoning on the mean is due to Cotes, according to whom[42] the centre of gravity of the observations (= of points in space) is the most probable and most reliable value of the constant sought. Several authors have described his qualitative reasoning, modernized it and noticed its relation to the method of least squares (MLSq)[43], while LAPLACE[44] remarked that

La règle de Cotes fut suivie par tous les calculateurs.

3.5. European History (19th Century, Sketch). After the construction of the classical theory of errors, which took place from the middle of the 18th century until the beginning of the 19th century, the ancient qualitative selection of observations had ceased to be considered acceptable and became gradually superseded by rejection of outlying observations according to certain quantitative criteria (the first criteria were introduced by B. PEIRCE, 1852, and W. CHAUVENET, in 1863) or by posterior weighting of observations, a procedure known even in the 18th century.[45]

Both the rejection of outlying observations and their posterior weighting, which brought into account stochastic considerations, had been time and again severely criticized, and the discussion of observations still remains an extremely delicate problem. In particular, modern counterparts of posterior weighted means are the so-called best linear estimators of the location parameter.

I begin the account with GAUSS[46]:

Zu einer erfolgreichen Anwendung der Wahrscheinlichkeits-rechnung 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 vorhandenen Beobachtungen sehr groß ist.

Generally speaking, continues GAUSS, there is a case for rejection, but

Halte man es wie man will, mache aber zum Gesetz, nichts zu verschweigen, damit andere nach Gefallen auch anders rechnen können.

Also, *wenn man mit dem Ausschließen zu schnell bei der Hand ist*, there exists a risk of overestimating the accuracy of observations.

Actually GAUSS went much further, postulating the principle of the arithmetic mean[47]:

It has been customary certainly to regard as an axiom the hypothesis that if any quantity has been determined by several direct observations, made under the same circumstances and with equal care (this, then, is GAUSS' definition of *gleichgenaue* observations, a notion of post-GAUSSIAN classical literature. He continued):

The arithmetical mean of the observed values affords the most probable value, if not rigorously, yet very nearly at least, so that it is always most safe to adhere to it.

Lesser known is another GAUSS' opinion on the mean[48]:

Pflegen die Erfahrungsdata selten in der reinen Gestalt sondern fast immer mehr oder weniger behaftet mit Störungen oder Schwankungen, die in ihrem Wechsel keiner Regel gehorchen, und man sucht dann den daraus entstehenden Nachteil so viel thunlich zu vermindern, dass man aus vielen einzelnen Resultaten das Mittel nimmt. Man rechnet darauf, dass bei einer solchen Benutzung einer großen Zahl von Fällen die zufälligen Schwankungen einander größentheils compensiren, und legt dann dem Mittelwerthe eine desto größere Zuverlässigkeit bei, je mehr partielle Resultate zugezogen sind. Dieses ist auch im allgemeinen vollkommen richtig, und durch consequente weitere Entwicklung und unsichtige Ausbeutung dieses Princips sind besonders in der Naturwissenschaften nicht selten die belohnendsten Früchte, selbst glänzende Resultate, gewonnen. Allein die Sicherheit des Grundprincips beruhet auf einer wesentlichen Bedingung, die häufig genug, auch von Gelehrten von Fach außer Acht gelassen wird, und die darin besteht, dass die an den einzelnen Beobachtungen haftenden regellosen Störungen oder Schwankungen von einander ganz unabhängig sein müssen. Das Urtheil, ob eine solche Unabhängigkeit vorhanden sei oder nicht, kann zuweilen sehr schwierig und ohne tiefes Eindringen in das Sachverhältnis unmöglich sein, und wenn darüber Zweifel zurückbleiben, so wird auch das den Endresultaten beizulegende Gewicht ein precäres sein.

Here GAUSS demands that the observations be mutually independent. He formulated this restriction before, at least twice[49], but it seems that emphasis is added only in this instance.

Thus, demanding equal rights for each observation, GAUSS was moreover inclined against rejection. Other prominent authors of the middle of the 19th century compared rejection of observations with their downright falsification. Such was the opinion of G. HAGEN[50], C. L. GERLING[51] and, possibly, V. YA. STRUVE[52] whom I quote in this order:

1. *Die Täuschung, die man durch Verschweigen von Messungen begeht, lässt sich eben so wenig entschuldigen, als wenn man Messungen fälschen oder fingiren wollte.*

2. *Jede Beobachtung, die nicht einen entschiedenen protocollarischen Verdachtsgrund gegen sich hat, habe ich als einen Zeugen für die Wahrheit zu betrachten, und eben so wenig wie ich den Zeugen torquiren darf, bis er sagt, was ich gesagt haben will, ebenso wenig darf ich auch ohne weiteres sein Zeugnis verwerfen, weil dasselbe von den übrigen bedeutend abweicht.*

3. *Discrepancies* (in sums of the angles of triangles) *furnish a veritable measure of probable errors only in case observations have been made without slightest prejudice and when no measurements have been rejected and substituted for new ones to guarantee a more satisfactory concordance in the sums of the angles.*

It would have been more accurate to take into account all conditions, including those provided by astronomical observations and base measurements. Struve continues:

The history of geodesy really presents a few examples of such work in which the concordance of the sums testifies to accuracy almost impossible to attain.

Strictly speaking, those were opinions against rejection as accomplished on subjective grounds. The use of stochastic criteria of rejection were also criticized time and again, but the climate of opinion gradually became milder[53]:

The question of rejection reduces to a question of common sense. Certainly the judgement of an experienced observer should be allowed considerable influence. … The judgement can undoubtedly be aided by one or more tests based on the theory of probability but any test which requires an inordinate amount of calculation seems hardly to be worth while, and the testimony of any criterion which is based upon a complicated hypothesis should be accepted with extreme caution.

Now I quote MENDELEYEV's[54] negative opinion about posterior weighting of observations:

A mean of various determinations could, and sometimes even should, be taken, but only when the relative merits of the (separate) *determinations are either completely unknown or could in no wise be distinctly deduced*[55]; *however, when one of the numbers* (of the determinations) *certainly secures more guarantees of accuracy than the other ones, it alone ought to be taken with complete disregard of the* (other) *numbers which undoubtedly represent either worse conditions of experiment … or leave room for doubt. … To consider worse numbers, taking them with some* (even small) *"weight" is tantamount to a deliberate corruption of the best of the numbers.*

But how to judge the accuracy of the various determinations? How to know that this or that determination is worse? The cursed problem is unsolved, but at least the general attitude of MENDELEYEV, the director of Russia's Bureau of weights and measures from 1893 until his death (1907), possibly reflects the essential difference between conditions of observations in astronomy and geodesy on the one hand and in metrology on the other hand.

3.6. Conclusion. I give word to a modern author[56]:

In former centuries the astronomer selected from among his observations those that seemed the best; this made him liable to bias or inclined to select such data as showed a possibly unreal agreement. (An example concerning TYCHO BRAHE follows.) *In the 17th century scientists like Huygens and Picard realized that the average of a number of equivalent measurements would be better than one of a couple of selected from them, and in the 18th century this averaging came more and more into use, all the more so since the concept of chance or probability of errors as a quantitative character had gradually become clearer. … A new attitude was brought into being, typical of the nineteenth-century scientist toward his material: it was no longer a mass of data from which he selected what he wanted, but it was a protocol of an examination of nature.*

I deal with the stochastic nature of averaging in § 4 and notice relevant qualitative opinions of a number of scholars. I also note that

PANNEKOEK did not substantiate the reference to HUYGENS. Here is the only instance known to me in which he[57] took a mean of numerical values:

Il paraît après tout le plus raisonnable d'admettre que, comme la Terre est placée entre Mars et Vénus par rapport aux distances, elle occupe également une place intermédiaire par rapport a la grandeur. … le diamètre de Mars est 1/166 du diamètre du Soleil, et celui de Vénus 1/84. Prenant done pour diamètre de la Terre la moyenne (arithmétique) *de ces deux diamètres, nous trouvons qu'il est 1/111 de celui du Soleil.*

En vérité, le diamètre de la terre est a celui du Soleil comme 1 est a 109 environ.

That was the (absolutely wrong) editorial comment on the remarkable accuracy of a hardly justified prediction. The Earth is even a bit larger than Venus.

Here are my own general remarks. Astronomers always attempted to deduce the "best" result from their observations; they more or less intuitively thought that a deliberate omission of some of their observations, stronger influenced by systematic and, possibly, random errors, was sometimes desirable; also, in case a large systematic error influenced each observation, which is a rather natural supposition, astronomers would possibly infer, reasonably enough, that numerous observations are just useless.

Speaking now about the 17th century, it is my opinion that the general improvement of observational techniques (invention of the telescope, vernier, level and crosshairs) led to the increase in the accuracy of observations by a whole order; thus, the errors of angle measurements decreased from minutes to seconds. This could have well influenced astronomers to regard the mean of a couple of observations, or even one observation, as sufficient for any practical purpose. Exactly this fact possibly explains why BRADLEY'S ideas (see above) did not immediately become universally accepted.

However, new and fundamental problems of the natural science, in particular, the deduction of the figure of the earth systematically studied from the middle of the 18th century, compelled astronomers to reconsider their approach and make full use of their observations. A second example of a problem, pertaining though to astronomy proper, which equally demanded utmost accuracy in the treatment of observations, was the problem of discovering nutation. It was in this connection that BRADLEY, the discoverer of both aberration and nutation, remarked[58]:

Science … had acquired such extraordinary advancement, that future ages seemed to have little room left for making any great imrovements. But, in fact, we find the case to be very different; for, as we advance in the means of making more nice inquiries, new points generally offer themselves that demand our attention.

As stated above, the quantitative, stochastic discussion of observations remains an extremely delicate procedure. The knowledge of corresponding laws of distributions (or at least of the nature of various random errors) and of the systematic errors is essential, and the appropriate studies nowadays belong to the realm of mathematical statistics.

4. Stochastic Treatment of Observations

4.1. Ptolemy. I have shown (§ 2.1) that already PTOLEMY clearly understood that obviously *gleichgenaue* observations (see GAUSS' definition of this notion in § 3.5) can be rather unequally influenced by systematic errors and that their elimination may be achieved at least partially by an appropriate combination of observations. It is more difficult to find out when combination of observations came to be used for the simultaneous elimination of random errors.

An example from PTOLEMY'S *Almagest* suggests that he thought that the "true" value of his observed constant was between the values of his two observations. Thus, supposing that the value of an astronomical constant lies between 47°40' and 47°45' (Book 1, § 12, p. 26), he took this constant to be $_{11/83}3600 = 47°42'39''$, a value previously accepted by ERATOSTHENES and HIPPARCHUS. It is an open guess at PTOLEMY'S choice in case of a lack of predecessors. But, even as it is, his decision does not violate common sense, differs but insignificantly from the choice of the arithmetic mean of the two bounds, 47°42′30", and is consistent with the properties of usual random errors (equal probability of positive and negative errors and greater probability of lesser errors).

4.2. Al-Biruni. Much more interesting are the remarks of AL-BIRUNI[59], notably (a) almost explicit statements about random errors and (b) a reasoning on the adjustment of direct observations.

(a) On p. 51 is a statement on the inevitability of (random) errors:

The same thing comes out in different amounts (= observations scatter), *because celestial observation is a very delicate matter; it requires precise measurements. … The approximation is also due to the extraction of square roots etc*., see § 2.2.

On pp. 155 – 156 AL-BIRUNI repeats his assertion, on this occasion in respect to time keeping:

Some measure (time) *with precision by continuous motions which are empirically equal in equal times and, as a rule, this has been done by the use of water. But it is subject to variation in many respects. For instance, the purity and density depend on its sources. ... Also, it is subject to accidental variations* (sic!), *by variation in the quality of the air. … Man has preferred the motions of sand to it.*

Lastly, it is AL-BIRUNI'S opinion (p. 83) that (random) errors can be both in excess or in defect and it is really possible that he supposes both these cases are equally likely:

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 estimators is due to the instrument.*

(b) The last passage also points to his general method of adjustment. As I see it, it consists of collecting several observations, discussing them qualitatively and choosing a more or less comfortable and common-sense single value for the constant sought. Thus, on p. 237 he says:

I shall rely on this 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 and (thus) *corroborates it*.

One or another corroboration of results is typical (p. 156):

If altitudes of fixed stars are observed (for measuring time), *though they are numerous, the corroboration of evidence obtained from some of them with that obtained from others leads to better accuracy*.

As it seems, corroboration of evidence means here a consideration of at least several observations. However, there is also a second use for corroboration, viz., for rough checking only. Thus, on pp. 46 – 51 he records four observations of solar altitudes and azimuths at Jurjaniya, the capital of Khwarism, from whence different values of its latitude (φ) are computed. It is on this occasion that AL-BIRUNI reasons on the inevitability of errors, see above item (a). Then he (p. 51) puts on record a fifth observation, that of the altitude of summer solstice. This crucial observation, simplest to be made and requiring almost no calculations, leads to $\varphi = 42°17'$, which, says AL-BIRUNI disregarding his previous observations, is his *reliable estimate*.

Returning to actual adjustment, I notice that in at least two instances AL-BIRUNI is more definite. Thus, on p. 168:

As to the halvng of the interval between the two times, it is a rule of procedure which has been adopted by calculators for the purpose of minimizing the errors of observation, so that the time calculated will be between the upper and the lower bounds.

He actually repeats himself (p. 203), this time, however, choosing the half-range:

As to the latitude of Baghdad, different observations have found that it is neither less than 33°20′ *nor greater than* 33°30′, *and the approved one is* 33°25′, *because it also the mean between those two*.

I do not think, however, that this was his general rule. Then follow the actual adjustments of longitudes and latitudes of different cities. AL-BIRUNI’S general presentation is rather obscure (tables are of course lacking and explanations insufficient), but it seems that hardly any essential further information can be gleaned. But we may say that AL-BIRUNI’S were the first explicit (qualitative) statements about propagation of computational errors, almost explicit statements about the inevitability of random errors of observation whose typical properties he possibly knew. Also, like PTOLEMY (§ 2.2), he clearly understood the essence of systematic errors of observation and strove to exclude their influence. Lastly, he repeatedly adjusted direct observations. He did not seem to adhere to any definite rule (e. g., to the choice of the arithmetic mean), but at least his general mode of action was sufficiently sound (use of inferior observations for corroboration only, choice of reasonable estimators, etc.).

4.3. India (16th Century). In § 2.3, I quoted a 16th century commentary on Lilávati, a 12th century Indian writing. It is now necessary to notice that the commentary lends itself to formalization. Let (for a two-dimensional case) ξ and η be (independent) random quantities, the (variable) length and width of an excavation, with expectations $E\xi$ and $E\eta$. Then, the more the number of measurements, the closer, according to the law of large numbers, will be the mean measures to those expectations, and the more accurate the calculated expectation of the area, $E\xi\eta$, the natural measure of the area of excavation[60].

4.4. European History (16th and 17th Centuries)

4.4.1. Galileo. He was the first to formulate the stochastic properties of (random) errors and some ensuing propositions[61].

His reasoning, purely metaphysical (= ARISTOTELIAN) in form of presentation, but no doubt substantiated by his own experimental experience, is that errors are inevitable, positive and negative errors are equally probable, lesser errors are more probable, the greater portion of observations is concentrated in the vicinity of the true value of the observed constant and that outlying observations should be rejected. He also proposed a method of treating indirect observations resembling the method of averages[62].

4.4.2. Kepler. His achievements seem to be known even less than those of GALILEO. It is generally known that KEPLER compared the conformity of the two main systems of the world, those of PTOLEMY and COPERNICUS (and the intermediate one, due to TYCHO BRAHE) with observations. After deciding in favour of COPERNICUS, he solved the second problem, viz., that of comparing possible closed curves as orbits of planets. I do not mention here his other two laws. However, KEPLER'S calculations, although discussed by historians of astronomy, should also be explained from a formal, mathematical point of view complete with statistical research. As it is, however, I present only a glimpse of his work.

KEPLER'S scientific outlook developed under the prevalent influence of ancient science, which is also evident from his attitude towards adjustment of direct observations. It is true that, in a businesslike manner, he chose the arithmetic mean (of two observations) as the estimator of constants sought in at least two instances[63]. However, he did not seem to be a staunch adherent of any definite estimator. In itself, this is no crime: from a modern point of view there is a case for choosing different estimators in accord with the corresponding laws of distribution. But KEPLER'S attitude, as described below, seems to be determined by ancient traditions. Thus, he says (Ibidem, Kap. 32, p. 219 with confirmation in Kap. 46, p. 273),

Das arithmetische Mittel aber zwischen Größen, deren Unterschied klein ist, ist nur um einen unmerklichen Betrag größer als das geometrische Mittel.

What KEPLER does not mention here is that the geometric mean is not additive and difficult to compute. Elsewhere (Ibidem, Kap. 10, p. 113) he collects four observations of the right ascension of Mars and, with no explanation given, assumes the *Mittlerer Betrag recht und schlecht* (*medium ex aequo et bono*) to be 134°24′33″. This passage is all the more interesting because on the very same page occurs one of the mentioned above cases in which KEPLER used the usual arithmetic mean.

Thus far the adjustment of direct observations. What I mean by ancient traditions is the complete lack of any mention of this subject in KEPLER'S deliberate account[64] of a book published by BODINI in 1586 and devoted to a curious description of the use of the three chief means (arithmetic, geometric and harmonic) in public life. Thus, KEPLER (p. 178) says

Vergleicht Bodini die Demokratie mit der arithmetischen Proportion, die Aristokratie mit der geometrischen und die Monarchie mit der harmonischen.

Consider with KEPLER three series

$$3, 9, 5, 10, 17, 38 \qquad (1)$$
$$6, 12, 8, 13, 20, 41 \qquad (2)$$
$$9, 27, 15, 30, 51, 114 \qquad (3)$$

The differences between corresponding terms of (2) and (1), i. e. 6 – 3 , 12 – 9 etc., are all equal, which is a property of consecutive terms of an arithmetic progression, a series directly connected with the arithmetic mean. On the other hand, a similar connection holds between the quotients of the corresponding terms of (3) and (1) and the geometric mean. KEPLER'S commentary is that

1. If die *Zuwüchse aller Zahlen … gleich sind, so will das Volk in der Republik, dass Lasten, Vorteile, Ehre … für alle gleich seien.*

2. But in an aristocratic society *die Zuwüchse der Zahlen den Zahlen selber angleicht.*

I am mentioning all this, to point out a heuristic connection between BODINI'S discussion of the geometric mean and D. BERNOULLI'S principle of moral expectation[65] with its flavour of aristocracy, and, second, to notice an example of KEPLER'S political views.

More interesting is KEPLER'S reasoning on the adjustment of indirect observations[66]. His data were the observations of TYCHO BRAHE who was able to reveal errors of 8′:

Für uns, denen die göttlche Güte in Tycho Brahe einen so sorgsamen Beobachter geschenkt hat, aus dessen Beobachtungen der Fehler der Ptolemäischen Rechnung im Betrag von 8′ sich verrät, geziemt es sich, dass wir dankbaren Sinnes diese Wohltat Gottes anerkennen und ausnützen. … Da jener Fehler aber jetzt nicht vernachlässigt werden durfte, so wiesen allein diese 8′ den Weg zur Erneuerung der ganzen Astronomie.

Even after establishing the heliocentric system of the world, a feat finally achieved only by the discovery of annual parallaxes of stars, KEPLER'S work still remained magnificent; for one thing, he had no reputed procedure of adjustment, e. g., the MLSq, at his disposal. He had to manage otherwise, and manage he did. He adjusted observations, correcting them by small arbitrary quantities compatible with the accuracy of the TYCHONIAN observations (see above).

Though KEPLER did not formulate any rules for this mode of action, he clearly recognized the necessity of utmost discretion (Ibidem, Kap. 26, p. 197):

Man könnte nun diese Freiheit, mit der ich an den gegebenen Größen kleine Änderungen anbringe, beargwöhnen und glauben, mit *dieser Freiheit zu ändern, was uns an den Beobachtungen nicht gefällt, könne man schließlich auch auf die ganze Exzentrizitet Tychos gelangen. Nun, man probiere das, und wenn man dann die Änderungen mit den unsrigen vergleicht, urteile man, welche von beiden innerhalb der Grenze der Beobachtungsfehler liegen. Ja, man hüte sich wohl, dass man sieh nicht, durch den Erfolg eines einzigen derartigen Schrittes gehoben, nachher bei den weiteren Schritten um so mehr blamiert sieht, wenn man das Apogäum der Sonne in entlegensten Örtern findet.*

Passages on the influence of errors occur in his *Neue Astronomie* at least twice (in Chapters 47 and 53) and in many instances (Chapters 49, 50, 53, 69, etc.) discrepancies are shown with their signs. There is no direct evidence of KEPLER'S thoughts about compensation of errors, but he[67] pronounced his opinion on a related topic:

Da bringe an ein gemein Orth zusammen 1900 alter ... Guldenthaler … was nun etwan ein Thaler zu schwer, das ist der ander zu leicht, dass es also auf … 100 Pfund … nichts aufftrügt.

This opinion is of course founded on what could be called the most rudimentary form of the law of large numbers[68] but KEPLER should have mentioned the mean weight.

Then, KEPLER noticed the inevitability of errors (*Neue Astronomie*, Kap. 10, p. 114):

Die Unstimmigkeit … habe ich deswegen angeführt, um zu zeigen, dass auch der Beobachtung selber eine Unsicherheit von etlichen Minuten anhaftet, wenn sie nicht durchaus mit größter Sorgfalt unter den günstigsten Umständen angestellt wird.

What KEPLER does not say is that even then some *Unsicherheit* still remains. Especially interesting is Chapter 51 of the *Neue Astronomie* where KEPLER adjusted indirect observations, a procedure possibly unheard of before even in usual land surveying. Discussing his observations, he demands that every one of them be taken into consideration (p. 313):

Da beim ersten und dritten Ort nahezu Übereinstimmung besteht, könnte jemand, der nicht weiter denkt, glauben, man müsse sich an diese Örter halten und die (two) *anderen auf irgendeine Weise mit ihnen in Übereinstimmung bringen.*

This is a denunciation of tampering with observations and, therefore an additional evidence of KEPLER'S discretion in treating observations.

When calculating corrections to angles of a quadrangle, he notices (Ibidem) that all the angles are of the same order so that there is no need to allow for very acute angles, whose small correction would have largely influenced the linear dimensions of the figure. Similar reasoning occurs also in Chapter 53, where KEPLER calculates changes of linear dimensions corresponding to unit changes in angles, and it really seems that he was prepared to discuss the advantages of different ways of relative weighting of measured angles in geodetic figures or at least the construction of the so-called base conditions in triangulation chains.

Problems pertinent to the theory of errors occur also in KEPLER'S *Mysterium Cosmographicum*[69]. His idea, as is well known, was to explain the general construction of the system of the world by inserting the five regular solids between the planet spheres of the six then known planets. He thus explained (at least to himself) the existence of just six planets and their mutual disposition. The greatest difficulties KEPLER encountered were of course those occasioned by the non-circularity of the orbits, which compelled him to introduce the *Dicke der Bahnen* (Kap. 18, p. 111). In any case, KEPLER'S calculations resulted in constructing the five solids and recording the discrepancies of his scheme as compared with observations, see the data from Kap. 21, pp. 134 – 135:

Discrepancies between calculations and observations

1. Saturn-Jupiter (cube) 2
2. Jupiter-Mars (tetrahedron) – 16
3. Mars-Earth (dodecahedron) 36
4. Earth-Venus (icosahedron) 43
5. Venus-Mercury (octahedron) 4

KEPLER noticed that four (out of five) discrepancies were positive and one negative, and he divided all of them into three groups according to their absolute values. Thus, two of them are small, another two are large and one is intermediate. All this is sound, but now KEPLER introduces duality: the cube and the octahedron are dual, and therefore the corresponding discrepancies are positive (and small); dual, also, are the dodecahedron and the icosahedron, and also positive (and large) are the corresponding discrepancies; lastly, the tetrahedron is dual to itself, and the corresponding discrepancy is the only negative (and intermediate).

This is hardly convincing, but clearly KEPLER attempts to interpret qualitatively the results obtained, and at least he remains loyal to his crazy theory of regular solids!

It is of course possible to notice the absence of any quantitative measure of the degree of fitness. However, measures of accuracy in the theory of errors first appear in the works of J. H. LAMBERT[70], whereas mathematical statistics had to wait for such measures until the end of the 19th century.

4.5. European History (18th Century and Gauss). Further achievements were due to LAMBERT and SIMPSON, neither of whom referred to GALILEI or KEPLER. Separate passages related to the stochastic theory of errors also exist in the writings of scholars who participated in arc measurements, where, again, there is no reference to GALILEI or KEPLER. In these writings, as also in those of BOUGUER and MAUPERTUIS, which I do not cite, there is no lack of numerous passages partly related to my subject but more appropriate in studies of the history of astronomy and geodesy.

Now, I quote the more relevant passages.

J. PICARD[71] explained discrepancies between observations as *un effet du hazard*. However, he added, *nous ne sommes pas fort éloignez de la vraye mesure du degré* (of the meridian). P. DE LA HIRE[72] called the mean of three observations *la vraye*. PICARD in other passages[73] repeatedly took a weighted mean of several observations calling it *la véritable*. G. D. CASSINI[74] noticed that *petites* (random) *erreurs sont presque inévitables*. He could well have omitted the *presque*.

C. M. DE LA CONDAMINE[75] inserted a special passage on the mean:

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 altere: puisque l'excès, ou le defaut de celles-ci se partageant entr'elles et toutes les autres, changeroit peu le résultat.

Returning to the postulate of the arithmetic mean (§ 3.5), I ought to add now that it was in actual use and even pronounced by at least one scholar (CONDAMINE) prior to GAUSS. That the arithmetic mean, to quote GAUSS (§ 3.5), affords *the most probable value*, is not exactly so. That value is the mode, the point of absolute maximum of the corresponding density curve. According to the law of large numbers, the arithmetic mean of the observations converges in probability to the expectation of the location parameter which, generally speaking, does not coincide with the mode. However, GAUSS actually restricted himself to unimodal symmetric densities[76], for which the two quantities coincide.

GAUSS applied this postulate for deducing the normal distribution and the principle of least squares (Ibidem, §§ 177 – 179). It is generally known that he later abandoned this approach and gave a different substantiation of the principle of least squares, independent of the postulate of the arithmetic mean (and of the normal distribution)[77]. Nevertheless, because of its elegance, of the extreme difficulty of understanding his main memoir of 1823, and, of course, because, generally speaking, the normal distribution did hold (approximately) in astronomy and geodesy, his first approach came to be widely known and even over-popularized so that possibly down even to our time some astronomers and physicists believe that the MLSq is inalienably connected with the normal distribution. The real connection of this method with the distribution of the corresponding observational errors is due to the statistical properties of the least squares' estimators.

It is impossible to say whether GAUSS read one of D. BERNOULLI'S memoirs on probability, published in St. Petersburg, with a commentary by EULER[78]. This commentary includes what amounts to a heuristic introduction of the MLSq, and, what is more, both of these writings taken together contain ideas sufficient for GAUSS' first deduction of the principle of least squares. However, EULER refutes BERNOULLI, so that if GAUSS did read their writings it would have been necessary for him to separate the ideas and to reassemble the whole reasoning. For a long time GAUSS himself felt that the MLSq was due to someone else, but he was never able to remember the reference[79].

GAUSS' celebrated rival, A. M. LEGENDRE[80], qualitatively founded the MLSq:

De tous les principes qu'on peut proposer pour cet objet (adjustrment of observations), *je pense qu'il n'en pas de plus général, de plus exact, ni d'une application plus facile que celui ... Par ce moyen il s'établit entre les erreurs une sorte d'équilibre qui empéchant les extrêmes de prévaloir, est très-propre a faire connoître l'état du système le plus proche de la vérite.*

Noticing the arbitrariness of the MLSq, LEGENDRE claimed further that when using it the *erreurs* (residuals) extrêmes, *sans avoir égard à leurs signes, soient renfermées dans les limites les plus étroites qu'il est possible.*

The LEGENDRE – GAUSS dispute over priority has been described time and again, but it seems that this unfounded claim regarding the *erreurs extrêmes* has been overlooked. Actually it is the generalized principle of least squares

$$\lim(v_1^{2k} + v_2^{2k} + ... + v_n^{2k}) = \min,\ k \to \infty$$

where v_i are the observational errors or, rather, the residuals of the corresponding system of n linear equations, which lead to this *minimax principle*[82] due to EULER[83].

5. Synopsis

5.1. Design of Experiments. PTOLEMY possessed a clear-cut notion about various errors of observations and about the different effect they produced, and left proposals concerning the selection of methods and circumstances of observations and the combination of different observations with each other to exclude systematic errors.

AL-BIRUNI advocated the use of regular series of observations both for astronomy proper and for practical applications. Noticing the existence of (random) computational errors and their propagation (see also § 5.3), he preferred direct methods of observations with as little subsequent computation as possible.

In ancient Babylonia and in the 12th century India approximate formulas for calculations of areas of land plots and volumes of excavations were in use, and it seems possible that these formulas were a means for allowing for the inaccuracy of the assumed mathematical models and for partially excluding systematic errors of measurement. In 16th – 17th century Europe a number of' scholars including GALILEO and KEPLER noticed the effect of systematic errors on observations and the necessity of "designing" experiments. Rules for excluding systematic errors are also found in 15th century writings pertaining to nautical astronomy.

5.2. Selection of Data. PTOLEMY selected the best observations disregarding those inferior which led to the excessive freedom of treating observations.

In 8th century China a part of observations made in connection with an arc measurement were disregarded and replaced by estimates computed from the other observations. This mode of action was not restricted to astronomy: physical experiments in 17th century as though carried out were often purely fictitious.

The first scientist of the new age to advocate the use of regular series of observations was possibly BRADLEY. His point of view, occasioned by his own fundamental discoveries in astronomy, did not win general recognition, possibly because of a sharp increase in the accuracy of observations in the 17th century. The advent of the era of arc measurements with new fundamental problems and theoretical research in probability proper, changed the climate of opinion so that the principle of the arithmetic mean became even postulated (COTES, CONDAMINE, GAUSS). Selection of observations in the framework of the classical theory of errors had been tried out in accord with stochastic, quantitative criteria (beginning from the second half of the 19th century) and with posterior weighting of observations (a milder procedure used even in the 18th century). Based on stochastic criteria or otherwise, the rejection of observations and also posterior weighting of them have been repeatedly criticized, and the selection of observations (and their treatment in general) remains an extremely delicate problem.

5.3. Stochastic Treatment of Observations. The inevitability of random errors had been noticed even by AL-BIRUNI. He also adjusted direct observations by a qualitative approach which is not at variance with stochastic properties of usual random errors of observations.

AL-BIRUNI used inferior observations for corroboration only, and noticed the propagation of various (also computational) errors. Adjusting direct observations, KEPLER used different estimators. He adjusted the TYCHONIAN observations by corrupting them by small arbitrary quantities. It seems that KEPLER understood the stochastic properties of usual random errors and applied his *corrections* with utmost discretion. A reasonable qualitative attempt at an explanation of residuals in his *Mysterium Cosmographicum* is marred by a non-sensical discussion in the framework of his crazy general theory of regular solids.

GALILEI was the first to formulate most important propositions now regarded as the cornerstone of the classical theory of errors. Some such propositions were also independently pronounced by a number of French savants of the 18th century.

Addendum (added in proof)

A Russian edition of AL-BIRUNI's *Qanun al-Masudi* is about to appear in Tashkent and I am indebted to Professor B. A. ROSENFELD for the possibility of commenting on the manuscript of this translation.

AL-BIRUNI discusses treatment of observations (pt. 4, chap. 1, pp. 364 – 367 of the manuscript), notices the existence of random errors of observations (pt. 6, chap. 4, pp. 631 – 632 and chap. 6, p. 636; pt. 9, chap. 2/1, p. 991), compares the accuracy of different methods of observations (pt. 5, chap. 1, p. 509; pt. 6, chap. 2, pp. 614 – 615 and chap. 7, pp. 657 – 658) and of course deals with the design of experiments (pt. 4, chap. 15, pp. 446 and 449; pt. 6, chap. 6, p. 637 and chap. 8, p. 669; pt. 7, chap. 5, p. 778).

Acknowledgement. My subject has been considered by a number of authors, some of whom are repeatedly mentioned. However, my

paper is written from a more mathematical point of view and is of wider scope. I did not use directly the two unpublished reports written in the last few years by C. EISENHART[84], but his work gave me moral encouragement and I am indebted to him for sending me copies of both his reports as well as of a provisional draft of a part of a subsequent article, to whose publication I look forward to with great interest.

Special features of his second report are references to English translations of the works of PTOLEMY and KEPLER about which I did not know previously. As a matter of fact, it was after reading EISENHART'S draft that I finally understood that the proper place for a preliminary version of my own, a. rather short article on the same subject, was in the dustbin. Lastly, I have used extensively my unpublished candidate thesis (1967) on the history of the theory of errors.

Dr. EISENHART has kindly sent me comments on this paper, but unfortunately they arrived too late to be taken fully into account, although they have helped me correct or improve quite a few passages. He objects that far too often I quote the texts only in part. Indeed, in search of brevity I may have somewhat sacrificed clarity.

Notes

1. NALIMOV, V. V., *Primenenie matematicheskoi statistiki pri analise veshchestv* (Application of mathematical statistics in analysis of chemical substances). Moscow, 1960. See p. 20.

2. Anonymous, Errors, of observations. *Fizich. Enz. slovar* (*Phys. enc. dict.*), vol. 4, pp. 77 – 78. Moscow, 1965. An instructive though non-mathematical treatment of experimental errors is STUDENT's Errors of routine analysis. *Biometrika*, vol. 19, pt. 1 – 2, 1927, pp. 151 – 164.

3. Theoria combinationis observationum etc. (1823). *Werke*, Bd. 4, pp. 3 – 53. Göttingen, 1873. See § 1.

4. On the accuracy of astronomical observations in antiquity. *Univ. of Gothenburg Astron. Notes* No. 10, 1968, pp. 3 – 36.

5. English edition (transl. by R. C. TALIAFERRO): *Great books of western world*, vol. 16, pp. 5 – 465. Chicago, 1952. Hereafter: *Great books*.

6. SHEYNIN, O. B., D. Bernoulli's work on probability. In *Studies in history of statistics and probability*, vol. 2. London, 1977. Editors, Sir Maurice Kendall, R. L. Plackett, pp. 105 – 132. First published in 1972 in an extremely rare source.

7. Qualitative measurements in antiquity. In: *L'aventure de la science* (*Mélanges A. Koyré*, t. 1), pp. 1 – 20. Paris, 1964. Quotations from pp. 2 and 3.

8. See note 4.

9. *Neue Astronomie*. Übers. und eingeleit M. CASPAR. München – Berlin, 1929. Originally published in Latin (1609).

KEPLER is referred to throughout this article and I feel it possible to add that, though denouncing randomness, he actually had to recognize it to explain eccentricities of planetary orbits. Also, following, though not referring to N. ORESME, KEPLER supposed that two numbers, taken "at random", are most possibly irrationally related to each other, a fact which he used to prove that the end of the world is practically impossible. Lastly, KEPLER'S astrology is a study of tendencies, of qualitative correlation between heaven and earth. I am incorporating these features of KEPLER'S writings in a separate article [Prehistory …, in this collection.].

10. Ibidem, Kap. 69, p. 386.

11. *Determination of the coordinates of positions for the correction of distances between cities.* Transl. JAMIL ALI. Beirut, 1967. There exists also a Russian translation by P. G. BULGAKOV (AL-BIRUNI'S *Sel. works*, vol. 3. Tashkent, 1966).

12. BHÁSCARA, *Lilávati*. Transl. with comments in H. T. COLEBROOKE, *Algebra und mensuration from the Sanscrit of Brahmegupta and Bháscara*. London, 1817, pp. 1 – 128. See Chap. 7, §§ 217 – 218, p. 97. COLEBROOKE gives the date of Lilávati as the middle of the 12^{th} century and that of the comment (by GÁNÉZA) as the 16^{th} century.

13. VEIMAN, A. A. *Shumero-Babylonian mathematics of the 3^{d} – 1^{st} millennia* B. C. Moscow, 1961 (in Russian). See p. 204.

14. *On the loadstone* etc. In: *Great books*, vol. 28, pp. 1 – 121. Transl. P. F. MOTTELAY. Quotation from Book 3, Chap. 12, p. 72. Originally published in Latin (1600).

15. *Dialogue concerning the two chief world systems*, etc. Berkeley and Los Angeles, 1962. See pp. 362, 383 and 388. Originally published in Italian (1632). Transl. S. DRAKE.

16. *Dialogues concerning the two new sciences*. In: *Great books*, vol. 18, pp. 128 – 260. Transl. H. CREW & A. DE SALVIO. Originally published in Latin (1638).

17. *Dialogue*, etc. (note 15), p. 440.

17a. *History and demonstrations concerning sunspots* etc. (1613). Transl. with introduction and notes by S. DRAKE. In: *Discoveries and opinions of Galileo*. Garden City, New York, 1957, pp. 88 – 144.

18. CASPAR, M. & W. VON DYCK, *J. Kepler in seinen Briefen*. München-Berlin, Bd. 1 – 2, 1930. See letters dated 16 Dec 1598, pp. 87 – 92 and 24 Nov. 1607, pp. 298 – 302 in Bd. 1.

19. *Neue Stereometrie der Fässer*. Leipzig, 1908. Originally published in Latin (1615).

20. *Kepler's Somnium*. Transl. E. ROSEN. Madison – London, 1967. Originally published in Latin. See KEPLER'S note 223 on p. 142. The notes, written in 1620 – 1630, occupy pp. 30 – 147.

21. *Horologium oscillatorium sive de motu pendularum* (1673). *Oeuvr. Compl.*, t. 18. La Haye, 1934, pp. 27 – 438. See part 1 of the memoir.

22. BOURNE, W., *A regiment for the sea and other writings on navigation*. E. G. R. TAYLOR (editor). Cambridge, 1963. The quotation is from p. 208 of the *Regiment* (1574) which occupies pp. 135 – 314.

Also connected with nautical astronomy was M. V. LOMONOSOV'S reasoning on the errors of compasses and clockworks and on their elimination, see *Meditationes de via navis in mari certius determinanda*; reported in 1759; *Poln. sobr. soch*. (Complete works), vol. 4. Moscow & Leningrad, 1955, pp. 187 – 319, in original Latin with transl. into Russian by YA. M. BOROVSKY. Another version of the same report, in Russian, is on. pp. 123 – 186.

23. Aestimatio errorum in mixta mathesi per variationes partium trianguli plani et sphaerici. (1722). *Opera misc.* London, 1768, pp. 10 – 68.

24. *Mesure des trois premiers degrés du méridien*. Paris, 1751. See p. 91.

25. *Almagest*, see note 5. Quotation from Book 4, § 6, p. 123. A similar passage is on p. 129.

26. For a short discussion see COLLINDER (note 4).

27. BEER, A., et al. An 8^{th} century meridian line etc. *Vistas in astronomy*, vol. 4, 1961, pp. 3 – 28. Also J. NEEDHAM, *Science and civilisation in China*, vol. 4, pt. 1. Cambridge, 1962. The relevant place occupies pp. 42 – 55 of the latter source, and the quotation below is from its p. 47.

28. See note 11.

29. *On transits*. Transl. M. SAFFOURI & A. IFRAM with comm. by E. S. KENNEDY. American Univ. of Beirut, 1959. See p. 27.

30. KOYRÉ, A., Pascal-Savant. In his *Metaphysics and measurement. Collection of essays*. Cambridge (Mass), 1968, pp. 131 – 156. Originally published in French (1956). Quotations from pp. 150 and 152.

31. BAILY, F., *An account of the Revd J. Flamsteed* etc. London, 1835. Quotation from p. 376.

32. FLAMSTEED, J., *Observations of Jupiter's transits* etc. (1673). *Phil. Trans. Roy. Soc. London 1665 – 1800 abridged*. London, 1809, vol. 2, p. 65.

33. BERRY, A., *Short history of astronomy.* London, 1898. Reprint: New York, 1961. See § 198.

34. *Correspondence of scientific men of the 17th century*, vol. 2. Oxford, 1841. See FLAMSTEED's letter to Lord BROUNKER dated 24 Nov 1669 (pp. 76 – 90), partly published m 1670. Reference is to p. 78.

35. Ibidem, See FLAMSTEED's letter to COLLlNS dated 10 Feb. 1672 (pp. 129 – 131). Quotation from pp. 129 – 130. Other letters to COLLINS could be also referred to, e. g., that dated 1 Apr. 1672 (pp. 131 – 133). The same source contains an accusation of RICCIOLI (letter to COLLINS dated July 16, 1670, p. 97):

In his catalogue of the fixed stars … he was so conscious of his own defects, he scarce has dared any where to recede (from the TYCHONIAN catalogue).

A similar accusation of RICCIOLI had been made by PICARD (p. 189 of his writing mentioned in note 71): *De deux observations il a choisi la première comme celle qui s'accomadoit mieux à son calcul*.

36. The principle of the arithmetic mean. *Biometrika*, vol. 45, 1958, pp. 130 – 135. Reprinted in *Studies in the history of statistics and probability*. E. S. PEARSON & M. G. KENDALL, editors. London, 1970, pp. 121 – 126.

37. *Tycho Brahe* etc. Edinburgh, 1890.

38. BRADLEY's loose paper, in S. P RIGAUD, *Misc. works and correspondence of J. Bradley.* Oxford, 1832, p. 78. Originally published by MASKELYNE in vol. 77 of the *Phil. Trans. Roy. Soc. London.*

39. BRADLEY, J., A letter concerning an apparent motion observed in some of the fixed stars. In S. P. RIGAUD (see note 38), pp. 17 – 41. Quotation from p. 29. Originally published in vol. 45 of the *Phil. Trans. Roy. Soc. London*.

40. On the advantage of taking the mean etc. *Phil. Trans. Roy. Soc. London*, vol. 49, pt. 1, 1756, for 1755, pp. 82 – 93. Quotation from p. 82.

Except for R. BOYLE I had been unable to trace those persons, and even BOYLE, besides not being an astronomer, died more than half a century before the publication of SIMPSON'S memoir. This is BOYLE'S opinion (A physico-chymical essay. *Works*, vol. 1. London, 1772, pp. 359 – 376, see p. 376):

Experiments ought to be estimated by their value, not their number; a single experiment … may as well deserve an entire treatise … As one of those large and orient pearls … may outvalue a very great number of those little pearls, that are to be bought by the ounce.

The BOYLE – SIMPSON controversy of course reflects the difference between conditions of observations in different branches of experimental science (see also § 3.5). Then, BOYLE seems to refute himself (*A proemial essay*, 1661. Ibidem, pp. 299 – 318, see p. 307):

What pleased me for a while, as fairly comporting with the observations was soon after disgraced by some further or new experiment.

41. SHEYNIN, O. B., Finite random sums (Historical essay). This *Archive*, vol. 9, 1973, pp. 275 –305.

42. See pp. 57 – 58 of his writing mentioned in note 23.

43. IVORY, J., On the method of least squares. *Phil. Mag. and J.*, vol. 65, No. 321, 1825. pp. 3 – 10.

44. *Essai philosophique sur les probabilités.* (1814). *Oeuvr. Compl.*, t. 7, pt. 1. Paris, 1886. See p. CL.

45. SHEYNIN, O. B., On the mathematical treatment of observations by L. Euler. This *Archive*, vol. 9, 1972, pp. 45 – 56.

46. Brief nach OLBERS dated 3 May 1827. In *Werke*, Bd. 8. Leipzig, 1900, pp. 152 – 153.

47. *Theoria motus* etc. (1809), § 177. Quotation from English translation by C. H. DAVIS (Boston, 1857).

48. Nachlass (Anwendung der Wahrscheinlichkeitsrechnung auf die Bestimmung der Bilanz für Witwenkassen, 1845). *Werke*, Bd. 4, pp. 119 – 183. Göttingen, 1873. Quotation from p. 143.

49. *Theoria motus* (cf. note 47), § 175; Theoria combinationis (see note 3), § 18. It is extremely instructive to notice the essential difference between the classical theory of errors and mathematical statistics as regards such fundamental motions as independence. See K. PEARSON, Notes on the history of correlation. *Biometrika*, vol. 13, 1920, pp. 25 – 45. Reprinted in *Studies in the history of statistics* etc. (see note 36), pp. 185 – 206.

50. *Grundzüge der Wahrscheinlichkeitsrechnung.* (1837). Berlin, 1867. See § 19, p. 50.

51. *Die Ausgleichungs-Rechnung der practischen Geometrie* etc. Hamburg – Gotha, 1843. See § 28, p. 68.

52. *The arc of the meridian* (1856 – 1861). Moscow, 1957. Reprint of selected chapters, in Russian. See § 32, pp. 127 – 128 (§ 37 of the 1861 edition). French edition, tt. 1 – 2, 1857 – 1860. Author´s name in German and French: F. G. W. Struve.

53. RIDER, P. R., Criteria for rejection of observations. *Washington Univ. studies*, new ser., Science and technology, No. 8, 1933. Quotation from pp. 21 – 22.

54. On the weight of a definite volume of water (1895). *Poln. sobr. soch.* (Complete Works), vol. 22. Leningrad – Moscow, 1950, pp. 105 – 171 (in Russian). Quotation from pp. 159 – 160.

55. In A. N. KOLMOGOROV'S opinion it is reasonable to choose the median in both these cases, see his La méthode de la médiane dans la théorie des erreurs. *Matematich. Zbornik*, t. 38, No. 3 – 4, 1931, pp. 47 – 49. In Russian; title of contribution also in French, French summary on p. 50.

56. PANNEKOEK, A., *History of astronomy*. London – New York, 1961. Quotation from pp. 339 – 340.

57. *Systema Saturnium.* (1659). *Oeuvr. Compl.*, t. 15. La Haye, 1925, pp. 209 – 388. Quotation from p. 346.

58. See S. P. RIGAUD (note 38), p. 17. (From BRADLEY'S *Letter … concerning an apparent motion* etc., see note 39).

59. See footnote 11.

60. Not needed.

61. *Dialogue* (see footnote 15).

62. On the method of averages see my contribution mentioned in note 45.

63. *Neue Astronomie* (see note 9), Kap. 10, p. 113 and Kap. 69, p. 386.

64. *Weltharmonik*. Hrsg. M. CASPAR. München – Berlin, 1939. Originally published in Latin (1619). See Buch 3, Kap. 16, pp. 175 – 195.

65. See note 6.

66. KEPLER, J., *Neue Astronomie*, Kap. 19, p. 166.

67. An den Senat von Ulm, 30 July 1627. In *Kepler in seinen Briefen* (see note 18), pp. 244 – 249 of Bd. 2. Quotation from p. 248.

68. See my contribution mentioned in note 60.

69. *Mysterium cosmographicum. Das Weltgeheimnis*. Hrsg. M. CASPAR. Augsburg, 1923. Originally published in Latin, 1596 and 1621.

70. SHEYNIN, O. B., J. H. Lambert's work on probability. This *Archive*, vol. 7, 1971, pp. 244 – 256.

71. *Mesure de la terre* (1671). *Mém. Acad. Roy. Sciences 1666 – 1699*, t. 7, pt. 1. Paris, 1729, pp. 133 – 190. Quotation from p. 175. This volume consists almost exclusively of reprints.

72. *Observations faites en Provence et à Lyon en 1682* (1693). Reprinted in same volume (note 71), pp. 413 – 428. Quotation from p. 414.

73. Observations astronomiques faites … en 1672, 1673, 1674 (1693). Reprinted in same volume, pp. 329 – 347. See pp. 330, 334, 335, 343.

74. *Observations astronomiques faites … en 1672.* First published in the volume of 1729 (see note 71), pp. 349 – 375. Quotation from p. 350.

75. See p. 223 of the book mentioned in note 24.

76. *Theoria motus*, § 175 (see also note 47).

77. *Theoria combinationis*, pt. 1 (see note 3).

78. See my contributions mentioned in notes 6 and 45.

79. GALLE, A., Über die geodätische Arbeiten von Gauss. In GAUSS' *Werke*, Bd. 11, Abt. 2. Berlin, 1924 – 1929 (separate paging).

80. *Nouvelles méthodes pour la détermination des orbites des comètes*. Paris, 1805 and 1806. Two Supplements, 1806 and 1820. Quotations in text from p. 72.

81. The most recent contribution is R. L. PLACKETT, The discovery of the method of least squares. *Biometrika*, vol. 59, No. 2, 1972, pp. 239 – 251. PLACKETT inserted numerous passages from the correspondence of several scholars, LEGENDRE and GAUSS included. Here is one more, from LEGENDRE'S anonymously published Second supplement to his *Nouvelles méthodes*. ... The place and date of publication (Paris, 1820) are given on p. 78, after the end of the main text. Then (p. 79 – 80) comes a *Note par M.** from which I quote:

L'auteur parlant de sa Méthode des moindres carrés (1805). *Cependant, comme un géomètre fort célèbre n'a pas hésite à s'approprier cette méthode, dans un ouvrage imprimé en 1809, nous croyons devoir nous arréter … sur cette prétention que tout lecteur impartial pourra qualifier ensuite du nom qu 'il jugera convenable.*

LEGENDRE then quotes GAUSS' claim to his use of the MLSq from 1795 and continues:

Si cela n'est pas très décisif, c'est du moins fort clair, et surtout fort commode. Dans ce système, l'histoire des sciences s'écrira beaucoup plus aisément, une découverte pourra bien ne plus appartenir à celui qui l'aura faite, mais qu'importe! Elle appartiendra toujours à quelqu'un, à celui qui aura trouvé bon de la revendiquer sans titre, à partir de l'époque la plus reculée. De bonne foi, un pareil système est-il admissible? On ne l'avait pense jusqu'ici dans aucune circonstance, comme l'Histoire des Mathématiques en fait foi à chaque page; on regardait la propriété d'une découverte comme invariablement assurée à celui qui, le premier, la mettait au jour et toute prétention contraire à ses droits, pouvant l'exposer lui-même à des soupçons d'une nature injurieuse, exigeait l'appui de titres précis et authentiques.

LEGENDRE goes on, pronouncing similar claims in the second subject, the theory of numbers.

82. See C. F. GAUSS, *Theoria motus*, § 186 (see also note 47) and P. S. LAPLACE, *Théorie analytique* … (1812), Book 2, Chap. 4, § 24. *Oeuvr. Compl.*, t. 7, pt. 2. Paris 1886.

83. See my contribution mentioned in note 45.

84. *The background and evolution of the method of least squares*. Report distributed at the 34th Session of the International Statistical Institute. Ottawa, 1963; *The development of the concept of the best mean of a set of measurements from antiquity to the present day*. Presidential address, 131st Annual meeting of the American Statistical Association, Fort Collins, Colorado, 1971.

Afterword

I do not repeat some statements from the Afterword to *Prehistory* … (in this collection) and begin by commenting on my text. First, concerning Bradley's opinion about new discoveries (end of § 3.6) I quote Descartes (1637/1982, p. 63):

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

Kepler (see § 4.4.2) indirectly proved that in his time or perhaps somewhat earlier the arithmetic mean became *the letter of the law*, see Sheynin (2017, § 1.2.4, p. 32). Numerous new sources are included in the Bibliography appended to that book of mine. Here, I only mention new translations of Kepler and Gowing (1983). The problem of outlying observations (§ 3.5) seems insoluble (Barnett & Lewis, 1978, p. 360):

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?

Lastly, I recommend readers to look up Jakob Bernoulli and Bayes (!) in the appropriate chapters of that same book.

Barnett V., Lewis T. (1978), *Outliers in Statistical Data*. Chichester, 1984.
Descartes R. (1637), Discours de la méthode. *Oeuvr.*, t. 6. Paris, 1982.
Gowing R. (1983), *Roger Cotes – Natural Philosopher*. Cambridge.

Sheynin O. (2017), *Theory of Probability. Historical Essay*. Berlin. **S, G,** 10.

V

Gauss and geodetic observations

Arch. Hist. Ex. Sci., vol. 46, 1994, pp. 253 – 283

1. Introduction

1.1. The aim and the scope of this paper. I (1979; 1988) have studied GAUSS'S work on the theory of errors, and my present aim is to study methods of estimating the precision of observations and related considerations (§§ 3 and 4) and the motion of dependence of observations (§ 5). I devote a special section (§ 6) to attempts, made by several authors, to present alternative views on the treatment of observations. Consequently, I had to discuss the goals of the theory of errors and to explain the present situation of its stochastic branch relative to mathematical statistics (§ 2). In treating the second topic of § 2 I also found it advisable to consider the metrological point of view.

Naturally, I dwell on the results and opinions of many scholars of the 18^{th} – 20^{th} centuries. However, since I (1971; 1973b, § 1.2.2, 1977) have already treated their work elsewhere, only a few words on LAMBERT or LAPLACE are in order here. In summary, one may say that:

In the theory of errors, LAMBERT was one of GAUSS'S main predecessors, but (understandably), in contrast to the latter, he did not engage in, or study triangulation (VOGLER 1902, p. 14). This fact obviously limited the scope of his studies.

LAMBERT was one of the first to estimate the precision of observations, but he did not relate his measures of precision to the number of the involved observations.

LAPLACE and GAUSS were the founders of the stochastic branch of the theory of errors. LAPLACE centred his research on the limiting case (see however § 6.3) whereas GAUSS studied the treatment of a finite number of observations.

1.2. The precision of calculations was not yet studied. GAUSS had to adjust large networks of triangulation (SHEYNIN 1979, p. 53), but he did not publish anything on the propagation of errors in geodetic calculations, nor, evidently, did he touch on this subject in his correspondence. This testifies once again to the fact that numerical analysis was not then really developed. For example, it was hardly possible to estimate the loss of accuracy in calculations.

Note that DELAMBRE (1814b, p. 309) solved systems of several thousand equations repeatedly combining them to obtain his unknowns. Elsewhere, he (1814a, p. iv) tells us that LEGENDRE'S remark to the effect that the choice of the arithmetic mean is a consequence of the principle of least squares

Autoriser les astronomes à prendre la somme de plusieurs centaines d'observations pour en former une équation finale …; à réunir ainsi plusieurs groups d'équations pour en former autant *d'équations finales qu'on le jugera convenable, et auxquelles enfin on appliquera la méthode des moindres carrés* … [1]

Furthermore, HARTER (1977, p. 33) noticed that HAUBER (1832) had recommended a hybrid method in which the number of equations is reduced to a manageable number (still greater than the number of quantities to be determined) by combining subsets of them as in the method of averages, and then solving the resulting equations by the method of least squares (MLSq).

Finally, BREEN (1849) described a similar procedure. I shall only say that calculations presented a formidable task not only for GAUSS and that the application of the hybrid method obviously hindered the estimation of precision.

In geodetic calculations, GAUSS retained more digits than necessary. Thus he (1828, § 23) calculated angles to within 0.″001 whereas their mean square error was 2. ″7 (he gave it as 2.″7440).[2] I am sure that GAUSS knew that the precision implied by such figures was fictitious and I believe that he was simply striving for *Schönheit und Rundung* (§ 5.3).

My remarks are also valid with regard to BESSEL (1838, § 87).

2. The theory of errors

2.1. Its two branches. In the classical theory of errors, two problems were studied:

One. To find most advantageous conditions and best methods for performing experiments and/or observations in natural science.

Two. To find optimal stochastic methods of treating observations.

The first problem was solved by determinate means by using differential calculus, and the relevant branch of the theory of errors was incorporated into the later experimental science in general. Thus, astronomers had to find out how best to eliminate systematic errors, and to study how the form of triangles in a chain of triangulation influenced the error of the final result, the length of a meridian arc.

The second problem, which demanded a stochastic approach, involved the determination of the best mean value(s) of observations and the estimation of their precision. LAPLACE and GAUSS were the real founders of the theory of errors, although many scholars (SIMPSON, LAMBERT, DANIEL BERNOULLI) preceded them. Moreover, GAUSS and BESSEL initiated the determinate branch of the theory of errors by putting forward the concept of the fullest possible analysis of instruments, and by making allowances for the influence of instrumental errors (SHEYNIN 1979, § 6.5). Now I note that the new approach was not restricted to astronomy. Thus, GAUSS investigated the errors of physical and geophysical instruments (see vol. 5 of his *Werke)*, and BESSEL had to pay attention to the measurement of meteorological elements. While investigating thermometers he (1826, p. 226) formulated a general statement:

Im Allgemeinen verdienen die Berichtigungen aller Instrumente durch Rechnung einen entschiedenen Vorzug vor den auf mechanischen Mitteln beruhenden; ich glaube sogar, dass die Verfertiger der Instrumente keineswegs verantwortlich sind für Alles, was der Besitzer selbst prüfen und dessen Verbesserung er selbst in

Zahlen bestimmen kann; … [*Seitens der Verfertiger*] *muss aller Fleiß angewandt werden, die Instrumente so einzurichten, dass ihre Prüfung in allen Theilen möglich wird* …

I additionally quote three authors: BESSEL [letter to GAUSS 15.6.1818, GAUSS (1975, p. 272)]; OLBERS [letter to GAUSS 17.7.1821; SCHILLING (1909, p. 123)]; and NEWCOMB (SCHMEIDLER 1984, pp. 32 – 33):

1) *Wir verdanken Ihnen den größten Theil der heutigen Verfeinerung der Astronomie, nicht nur wegen Ihrer kleinsten Quadrate, sondern auch wegen der Erweckung des Sinns für Feinheit, der seit Bradley's Zeit von der Erde verschwunden zu sein schien und erst seit l8 Jahren wieder erschien. Wir sind erst jetzt auf den Punkt gekommen, kleinen Fehlern oder Abweichungen außer den Grenzen der Wahrscheinlichkeit mit derselben Aufmerksamkeit nachzuspüren als früher großen* …

Did BESSEL single out the year 1800 just for the sake of simplicity? Forbes (1978, p. 177) holds that after 1819 GAUSS

Appreciated more than ever the need to subject every major astronomical instrument to very careful investigations …

He also mentions the British astronomer P0ND (cf. my Item 3 below).

2) *Auch praktische Astronomie, worin Sie und Bessel gewissermaßen Epoche machen.*

3) NEWCOMB credited BESSEL with founding the German school of practical astronomy:

The fundamental idea 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.
NEWCOMB did not refer to GAUSS. He mentioned POND, but did not elaborate. Of course, GAUSS and BESSEL had forerunners, such as TYCHO and BRADLEY, and, in ancient times, HIPPARCHUS. '

2.2. Its present situation. The stochastic branch of the theory of errors is now thought to be incorporated into mathematical statistics. Accordingly, its results are interpreted in statistical terms and the theory itself is hardly mentioned. Thus, the *Enc. of Statistical Sciences* (Kotz & JOHNSON 1982 – 1988) refers to the error theory only twice, each time in passing. Witness also the situation in technical academic institutions. TAYLOR (1982, Preface) states that error analysis (he does not use the classical term) is *often the most abused and neglected part* of introductory college courses in experimental physics *of the sort usually taken … in the sciences of engineering*. Nevertheless, the theory of errors is still needed in geodesy (if not astronomy) and is indispensable to metrology. In essence, the peculiarity of the error theory largely results from the fact that there is no clear-cut boundary between systematic and random errors (GAUSS 1823, § 1) and, consequently, between its two branches.

COLCLOUGH (1987) recently stated that the theory of errors cannot dispense with the motion of systematic error. Indicating that there is another approach, whose partisans treat all errors in the same way, he denied this *randomatic* theory.[3] He referred to GIACOMO (1981) who had discussed the randomatic approach to the treatment of observations at the *Bureau Intern. Poids et Mesures*.

Much can be picked up from EISENHART (1963) who paid attention to the study of systematic errors (cf. the quotation from DORSEY & EISENHART in § 4.1) and stated, on p. 53, that the obtained measures of precision and accuracy should be put on record separately rather than combined in one or another way. GAUSS (1823, §§ 8 and 15) held that, if possible, the constant component of the total error should be isolated. In metrology, it is now indeed customary to distinguish between precision and accuracy: the former term relates to the internal consistency of observations (i.e., to the influence of random errors) whereas the latter is a measure of bias (EISENHART 1968, p. 69).

Another notion of the classical theory of errors is that of the *true value* of the unknown constant. In the 19th century, astronomers used this term without explanation. Although mentioning it on occasion (§ 4.3), GAUSS, as it seems, did not favour it. COTES, in 1722 (posthumous publication), was the first to introduce it. Evidently, it is difficult to say what, for example, is the true value of an angle in a chain of triangles. Similar difficulties exist in metrology (EISENHART 1963, pp. 30 – 31). Nevertheless, generally speaking, when the number of observations is increased indefinitely, their arithmetic mean approaches a certain value. By definition, when an *exemplary* method of measurement is used, the true value sought is considered equal to this limiting mean (Ibidem, p. 30). Thus, true value has been specified rather than discarded.

WHITTAKER & ROBINSON (1958, p. 215n) and GLEISBERG (1964) offered the same definition without mentioning the *exemplar*. However, it was FOURIER (1826, p. 534) who first stated that

C'est cette quantité fixe … [the limit of the arithmetic mean] *que nous avons en vue comme le véritable objet de la recherche*.
[On that notion see Sheynin 2007.]

3. The precision of observations

3.1. Redundant observations increase precision (Mayer). Suppose that l_i, i = l, 2, …, n are measured directly, that x, y, z, … are the constants sought, and that

$$a_ix+b_iy+c_iz + \ldots + l_i = 0. \qquad (1)$$

Then, obviously, along with solving this system it is necessary to estimate the precision of its unknowns.[4] MEYER, in 1750 (STIGLER 1986, p. 23), was the first to estimate the precision of one of his unknowns (the one which most interested him). Having 27 equations (1) in three unknowns, he first solved some subsystem of three equations calculating (x_1; y_1; z_1). Then he aggregated all his equations in three groups of nine equations each. He assumed that in each group the sum of the residuals vanished and was thus able to solve these three groups simultaneously,[5] calculating (x_2; y_2, z_2). The two values of x (say) differed by Δx and MAYER concluded in essence that the error of x was 27:3 = 9 times less than $|\Delta x|$.

MAYER'S inference was wrong: as STIGLER noted, precision increased only √9 = 3 times. Nevertheless, to repeat, he had no predecessors and, in addition, he attempted to solve his problem without knowing how to estimate the precision of the l_i's.

3.2. The error of the sum of observations (Daniel Bernoulli). DANIEL BERNOULLI (1780) noted, although did not prove, that the error of the sum of modulo equal errors, equally likely to be either positive or negative, increased as the root of their number. He did not say that the observations should be independent. Much later FOURIER (1829, p. 568) made a similar statement assuming that the errors were bounded rather than modulo equal.

The proof was not difficult even for BERNOULLI. Given

$$H=h_1+h_2+\ldots+h_n$$

one has

$$\Delta H^2=\Delta h_1^2+\Delta h_2^2+\ldots+\Delta h_n^2+2\Delta h_1\Delta h_2+\ldots+2\Delta h_{n-1}\Delta h_n$$

and, in the mean, since the mean values of Δh_i vanished,

$$|\Delta H| = |\Delta h|\sqrt{n}.$$

FOURIER (1829, p. 574) also derived the formula for the mean error of a function of several observations. In § 3.6 I note that his formula is due to GAUSS. Of course, similar relations were known even to COTES, but the concept of the mean square error, properly speaking, originated with GAUSS.

3.3. The weight of the arithmetic mean (Gauss). Supposing that the measure of precision (*gradus praecisionis, Grad der Genauigkeit*) of observations x_i (notation changed) were proportional to e_i GAUSS (1809, § 173) wrote out the arithmetic mean of n such observations,

$$\bar{x}=\frac{e_1^2x_1+e_2^2x_2+\ldots+e_n^2x_n}{e_1^2+e_2^2+\ldots+e_n^2} \qquad (2)$$

and stated that its measure of precision was proportional to the square root of the denominator of (2), call this root (3) so that

Vier oder neun gleich genaue Beobachtungen erforderlich sind, wenn sich das Mittel der doppelten oder dreifachen Genauigkeit erfreuen soll, und so weiter.

That the appropriate parameter of the normal distribution could have been considered *als das Maass für die Genauigkeit* (Gauss 1809, § 178) was, practically speaking, not so important.

Of course, e_i^2 were the weights of the observations (*pondus, Gewicht*), a term which GAUSS introduced later (letter to BESSEL 27.1.1819; GAUSS 1975, p. 294), GAUSS to OLBERS 14.4.1819; SCHILLING (1900, p. 722); and GAUSS 1823, §7).

GAUSS did not prove his statement, but referred (at least with regard to using the mean value of $\bar{x}$) to a later section of his work. He introduced the notion of random error only in § 175 (although he did mention this term earlier in § 173).

The relevant later section was obviously § 181, where, in accord with his general context, GAUSS considered the case of normally distributed errors. If, again in changed notation, observations provide

$$e_i x = m_i,\ i=1, 2, ..., n, \tag{4}$$

then

$$x = \frac{m_1 e_1 + m_2 e_2 + ... + m_n e_n}{e_1^2 + e_2^2 + ... + e_n^2}. \tag{5}$$

Formula (5) might seem strange. If so, note that equations (4) lead to a single normal equation

$$[ee]x - [em] = 0.$$

If the original observations are assumed to have unit precision, that of x is equal to (3). In both formulas (2) and (5) the weight of the unknown, given weights of x_i, (or m_i), is obviously the square of expression (3).

3.4. The weight of indirect observations (Gauss). In the same contribution GAUSS (1809, §§ 183 – 184) determined the relative measure of precision of the unknowns of system (1) thus making an important step in the right direction. His reasoning is not easy to follow,[6] and I shall therefore corroborate it using his own example. GAUSS considered four initial equations in three unknowns and formed the appropriate normal equations:

$$27x + 6y - 88 = 0$$
$$6x + 15y + z - 70 = 0$$
$$y + 54z - 107 = 0$$

He calculated the unknowns and estimated their relative precision as 4.96; 3.69; and 7.34. According to a standard method of calculation (GAUSS 1823, § 21; HELMERT 1872, pp. 81 – 89) the relative weights of the unknowns are equal to

$$p_x = 1/Q_{11},\ p_y = 1/Q_{22},\ p_z = 1/Q_{33}. \tag{6}$$

With quantities Q_{ii} and Q_{ij} being determined from three systems:

$$[aa]Q_{i1} + [ab]Q_{i2} + [ac]Q_{i3} = q$$
$$[aa]Q_{i1} + [bb]Q_{i2} + [bc]Q_{i3} = r$$
$$[aa]Q_{i1} + [bc]Q_{i2} + [cc]Q_{i3} = s$$

(notation [aa], [ab], etc. is standard). In the first system, $i = q = 1$, $r = s = 0$; in the second one, $i = 2$, $q = s = 0$, $r = 1$; and in the third system, $i = 3$, $q = r = 0$, $s = 1$. Quantities Q with $i \neq j$ are also important, but not in this case. According to my calculations,

$$p_x = 24.60 = 4.96^2,\ p_y = 13.65 = 3.69^2 \text{ and } p_z = 53.93 = 7.34^2.$$

In other words, GAUSS determined the square roots of weights.

ENCKE (1834 – 1836, pp. 157 – 160) took up GAUSS'S example, formed and solved the normal equations and calculated the weights of the unknowns.

3.5. The probable error. After LAMBERT, the first to estimate the precision of directly observed quantities, the l_i's in system (1), was DANIEL BERNOULLI. In 1780, assuming a normal distribution for the errors of pendulum observations, he described the precision of the measurements by their probable error although without defining or naming it.

BERNOULLI modelled his investigation on his earlier work on the ratio of male and female births (1770 – 1771). It was there that he first introduced the normal distribution and used probable deviation.[7]

DELAMBRE (1814a, p. iii) mentioned the term (or was it a loose expression?) *l'erreur probable. Tout ce qu'on peut prétendre alors,* he stated, was that the errors remaining after adjusting observations [the residuals] will not surpass *l'erreur probable des observations*. DELAMBRE next wrote about least squares in such a way that readers could have thought that this new method ensured the fulfilment of his wish.[8]

BESSEL (1816, pp. 141 – 142) formally introduced the probable error and GAUSS (1816) studied various methods of calculating the probable error of normally distributed observations. I have described this material (1979, §4; 1983, p. 177) and quoted the latter's opinion to the effect that the probable error should be altogether abandoned. At the time, this was a pipedream.

GAUSS himself sometimes made use of the probable error; see his letters to OLBERS of *etwa* 19.5.1819 (SCHILLING 1900, p. 726) and to SCHUMACHER of 14.8.1825 and between 14.7 and 8.9.1826 (PETERS 1860 – 1865, pp. 30 and 65 of Bd. 2). Furthermore, in one case, see letter to OLBERS 26 – 30.7.1825; SCHILLING (1909, pp. 424 – 425) GAUSS remarked that the probable error of his observations was *fast genau* ± 0."140, so that *1 gegen 1 gewettet werden kann* etc.[9]

I conclude that at least once GAUSS was unable to resist what I would call the temptation of the probable error, but I cannot understand why he applied this measure in his published paper of 1828 (**S, G,** 72).

3.6. The mean square error (Gauss). GAUSS (1823, § 38) derived the formula for the mean square error of indirect observations

$$m = \sqrt{\frac{[vv]}{n-\rho}} \qquad (7)$$

where the v_i's were the residuals of equations (1) and ρ was the number of the unknowns. Of course, $\rho = 1$ corresponded to the case of direct observations.

Note that in a letter to BESSEL 18.5.1814 GAUSS (1975, p. 191) still incorrectly took the denominator of formula (7) as n.

For the related work of LAPLACE see SHEYNIN (1977, § 7.1). Note also that LAPLACE always estimated the precision of his calculations. Thus (*Théorie analytique*, chapter 6; SHEYNIN 1977, § 2.5), in studying the function

$$y = x^p(1 - x)^q$$

of an unknown probability x with observed p and q, he chose $a = p/(p + q)$ as the [asymptotically unbiased] estimator of x and calculated $P(|x - a| < \alpha)$. The function y represented, for example, the probability of p male, and q female births with very large p and q.

GAUSS did not directly introduce the mean square error of the arithmetic mean. FOURIER (1826, p. 541) stated that this measure was equal to $m\sqrt{2/n}$ (his main estimator was thus $m\sqrt{2}$), but he calculated m as GAUSS did previously, taking $p = 0$ instead of 1.

Not being satisfied with his new finding, GAUSS (§ 40) estimated the bounds for varm^2. I have put on record (SHEYNIN 1979, p. 45) that HELMERT, in 1904, had improved GAUSS'S estimate. In 1947, three authors, including KOLMOGOROV, independently arrived at HELMERT'S result.[10]

LIAPUNOV (1975, posthumous publication) derived formula (7) stating that it was not proven rigorously. Evidently he saw it in a later treatise rather than in GAUSS'S *Theoria combinationis*. LIAPUNOV did not publish his proof and his manuscript appeared posthumously. His efforts were not lost: he showed that, in modern terminology, m^2 was a consistent estimator of the variance, a proposition that can now be proved simply by referring to an appropriate theorem due to KHINCHIN.

The date of LIAPUNOV'S discovery remains unknown, but possibly it was made at, and marked the beginning of his fundamental work on probability.

In a letter to OLBERS 17.6.1824 (SCHILLING 1909, p. 317) GAUSS remarked that his *Theoria combinationis* provided *leichte Mittel … die mathematische Güte eines Dreiecksystems methodisch zu würdigen*. Obviously he had in mind his § 21 where he derived formulas for the precision of the unknowns of system (1).

It occurred that for observations of equal weight

$$m_x = m\sqrt{Q_{11}},\; m_y = m\sqrt{Q_{22}},\; m_z = m\sqrt{Q_{33}}\,,$$

see my formulas (6) and (7). Indeed, GAUSS thus made it possible to estimate the errors (or weights) of the unknowns. Moreover, it became possible to calculate the Q_{ii}'s (and, therefore, to estimate the relative precision of the unknowns) even before making the observations, just by using the rough measurements provided during reconnaissance. Note that GAUSS could have said that the *leichte Mittel* became available in 1809 rather than in 1823.

Finally, the *Theoria combinationis* contained a proof of the now known formula for the weight (P) of the arithmetic mean of observations $x_1, x_2, \ldots, x_n$ having weights $p_1, p_2, \ldots, p_n$:

$$P = p_1 + p_2 + \ldots + p_n. \qquad (8)$$

GAUSS (§ 22) derived this formula by noting that direct observations (one unknown) constituted a particular case of indirect observations (with several unknowns). He also provided an independent demonstration (§ 18). In the same section (§ 18) GAUSS derived the mean error of a linear function of several independent observations. It is quite possible that even before the time of GAUSS astronomers sensed formula (8), but it was GAUSS who substantiated it. Cf. § 6.3.

4. Practical considerations

4.1. The Number of Observations. When should an astronomer stop his work on a given station? May he guide himself, in estimating the precision of his observations, by their internal consistency? I have stated (SHEYNIN 1979, p. 51) that with respect to observations GAUSS did not adhere to any definite rules, and that after his time, striving to eliminate systematic errors, national geodetic services in several countries have established rigid programmes of observation.

I shall now elaborate on my first remark, and begin by quoting SCHREIBER (1879, p. 141) whose statement was noticed by JORDAN (1882, p. 11):

Aus seiner [GAUSS'S] *mir vorliegenden Protokollen geht vielmehr hervor, dass er auf jeder Station so lange gemessen hat, bis er meinte, dass jeder Winkel sein Recht bekommen habe. Er hat dann … die hervorgehenden Richtungswerthe als gleichgewichtig und von einander unabhängig in die Systemausgleichung eingeführt.*

SCHREIBER went on to agree with GAUSS'S attitude but did not explain it. And he could have referred to the latter's one-time student, GERLING (1839, pp. 166 – 167) who declared that

Kommt man … früher oder später an die Gränze, um welche [the result] *in gewissen geringen Oscillationen herumschwanken, und überzeugt sich, dass jedes weitere Fortsetzen der Repetition nur verlorene Arbeit seyn würde.*

And, further, in the same context:

So aber habe ich es, nach dem Beispiel von Gauss, regelmässig immer gemacht.

COURNOT (1843, pp. 151 and 162) and even BAYES (STIGLER 1986, pp. 94 – 95; DALE 1991, pp. 313 – 315) were of the same opinion. Both scholars indicated that the chances for the same error of observation in excess or defect were not exactly equal [that systematic errors were unavoidable]. Another reason was that, generally, the observations were not strictly independent.

It is opportune to compare GAUSS'S attitude with a modern opinion (DORSEY & EISENHART 1969, p. 53): the experimenter

Will proceed to change, one by one, every condition,[11] *... that seems by any chance likely to affect his result,*

and thus eliminate known systematic errors as much as possible; and he will *take long series of observations* to eliminate systematic errors from unsuspected causes.[12] They conclude: he will presently

Feel justified in saying that he feels, or believes, … that his … work indicates that the quaesitum does not depart from his … definitive value by more than so-and-so, meaning thereby, … that he has found no reason for believing that the departure exceeds that amount.

Previous authors (CLARKE 1880, pp. 18 and 52; CAMPBELL 1928, p. 164) came to a similar conclusion.

Did GAUSS change the conditions affecting his observations? Did he, for example, measure angles both in the morning and in the evening? I have not been able to find an exhaustive answer to this question. However, it is known, see for example his letter to BESSEL 29.10.1843; GAUSS (1903, pp. 494 ff) that, to eliminate a certain instrumental error, GAUSS measured both the angle sought, A, and its complement, 360° – A. Then, GERLING (1839, pp. 14 – 15) testified that he had observed at least

Die Punkte, welche zu verschiedenen Tageszeiten oder bei heiterem und bedecktem Himmel Lich -Phasen befürchten ließen, so viel möglich auch unter diesen verschiedenen Umständen.

I can also mention STRUVE (1824, p. 433), though his attitude was probably independent. He measured angles 32 times each by a repetition theodolite.

Ich aber nicht einen Winkel 32 Mal gleich nacheinander nahm, sondern nach jedem 4 fachen Winkel ablass und mit der Beobachtung der verschiedenen Winkel abwechselte und jeden unter verschiedenen atmosphärischen Umständen beobachtet zu haben.

GAUSS did not measure all his angles the same number of times; indeed, there are instances even of striking differences between these numbers (GAUSS 1903, pp. 278 – 281). But then, was it proper for him to enter these angles in a general adjustment, as SCHREIBER (above) noted, with equal weights? This is a delicate question and, obviously, the observer himself should answer it. Also cf. GERLING (1839, p. 167):

Insofern also ist allen Richtungen einerlei Gewicht beizugeben, wenn gleich nach Verschiedenheit der Umstände mitunter eine sehr verschiedene Anzahl von Beobachtungen dazu gehörte, um in ihrer Festlegung die oben bezeichnete Grenze zu erreichen.

See above his statement from the same source and page.

4.2. Estimating the precision. From the practical point of view, formula (7) is not always sufficiently sensitive. BERTRAND (1888, p. 274), for example, indicated that inaccurate observations can be internally consistent. GAUSS himself was fully aware of this fact. In several letters (to OLBERS 29.1.1822; SCHILLING (1909, p. 164) and 14.5.1826; GAUSS (1903, pp. 320 – 322), to BESSEL 15.11.1822; GAUSS (1975, p. 407), and to BOHNENBERGER 16.11.1823; GAUSS (1903, pp. 364 – 367) he noted that a small deviation of the sum of the angles in a triangle from its theoretical value did not yet testify to the worthiness of the observations. It was too easy to obtain small discrepancies by manipulating the observations, GAUSS stated. In his letter to BESSEL he remarked that this deviation

Zuweilen dazu gedient haben mag, wenn auch nicht die Beobachtungen zu verfälschen, doch etwas zu wählen (*man bemerkt eine Tendenz dazu selbst bei Delambre*).[13]

GAUSS added that falsification was easier to detect, and the influence of systematic errors was more fully revealed, in the case of a braced quadrilateral or of a system of triangles with a common vertex. See also GAUSS (1826).

STRUVE (1831, p. 86; 1860, p. 145), without referring to GAUSS, corroborated his opinion concerning the selection of observations. He added, however, that eccentric measurements, to which astronomers often had to resort, led to a *völlige Unbefangenheit*.

Another relevant point is the case of a small number of observations. In a letter to BESSEL 19.4.1821 GAUSS (1975, p. 382) stated that *es immer misslich ist, auf wenige … Beobachtungen ein Resultat zu gründen*. He expressed his views more emphatically in letters to GERLING 17.4.1844 and 29.1.1847 (SCHÄFER 1927, pp. 687 and 744). In the second case he wrote

Wo man nicht Rechnung auf eine <u>große</u> Anzahl von Erfahrungen stützen kann, soll man sich lieber bloß an eine nur auf Kenntnis des <u>Sachverhältnisses</u> stützende Schätzung halten.

Perhaps even more interesting was his earlier letter to GERLING. Here he tabulated the magnitudes ($n - \rho$), [vv], and m^2, see formula (7), for 18 triangulation stations. The 18 values of m^2 ranged from 0.″19 to 5.″1, for which there was *gar kein Grund … die Sache ist lediglich die, dass aus einem so kleinen n* [*so kleinen* ($n - \rho$), as it happened on several stations] *sich nichts Sicheres folgern lässt*.

GAUSS added together the data for stations 1 – 5; 6 – 9; and 10 – 18 obtaining (n — ρ) = 47; 48; and 47, and m^2 = 3.″1; 3.″1 and 2.″8 respectively. In neither letter did he recall his formula for estimating the variance of m^2 (§ 3.6)! Cf. KU (1967, p. 309):

An estimate of the standard deviation based only on a small number of measurements cannot be considered as convincing evidence.

4.3. The limits of security. ENCKE (1834 – 1836/1888, pp. 43 – 47) was (one of?) the first to use this short-lived term, *Grenze der Sicherheit*. What he meant was the *wahrscheinlichen Grenzen der wahren Werthe* (GAUSS 1816, §§ 4 – 7) so that the *Sicherheit* was directly connected with the probable error.

Both GERLING, in a letter to GAUSS 19.2.1838 (SCHÄFER 1927, p. 522) and BESSEL, writing to OLBERS on 28.6.1839 (ERMAN 1852, Bd. 2, p. 441), mentioned the same term in a businesslike manner, without any comment. What GAUSS used was a certain confidence interval with confidence level 0.50.

4.4. Rejection of outlying observations. When should the astronomer reject an outlying observation in the absence of any prior evidence against it? Modern authors (DIXON 1962; KRUSKAL 1960, p. 348) confess that they do not know any general answer. Obviously, statistical tests may help the astronomer to decide what to do, but no test is better than its premises whose validity it is difficult to check. A special inseparable problem is to decide whether or not to reject an observation which does not (cannot) belong to the same population as all the other ones. KRUSKAL concludes:

My own practice is to carry out an analysis both with and without the suspect observations. If the broad conclusions of the two analyses are quite different, I should view any conclusions from the experiment with very great caution.

BARNETT & LEWIS (1984) contributed a whole book to *Outliers in Statistical Data*. They conclude (p. 360) that the problem of treating outliers obviously *arous*[es] *more interest today than it has ever done*, and that

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

I shall now briefly discuss the relevant developments in the early 19th century. GAUSS (1900, pp. 152 – 153), in a letter to OLBERS dated 3 May 1827, stated:

Zu einer erfolgreichen Anwendung der Wahrscheinlichkeits-rechnung 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 vorhandenen Beobachtungen sehr groß ist.

Generally speaking, GAUSS continued, there is a case for rejection, but

Halte man es wie man will, mache aber zum Gesetz, nichts zu verschweigen, damit andere nach Gefallen auch anders rechnen können. … wenn man mit dem Ausschließen zu schnell bei der Hand ist,

the precision of the observations can be overestimated.

It seems that KRUSKAL'S experience (above) is quite in line with GAUSS'S attitude.

GERLING (1843, p. 68) stated that with regard to rejecting observations *bleibt der Beobachter … ganz auf sein praktisches Gefühl verwiesen.*

The difficulty of the problem of rejection is easy to understand: the theory of probability applies only to what will happen after a great number of trials whereas the observer has to decide what to do with a single set of observations (corrupted by systematic errors as well).

The first quantitative recommendations concerning the rejection of observations were made by FOURIER (1824) and JORDAN (in 1877), see CZUBER (1891, pp. 207 – 211) and HARTER (1977). JORDAN introduced the celebrated rule of *three sigma* whereas FOURIER (1826, p. 543) believed that *le triple de g* [= $m\sqrt{2/n}$, see § 3.6] *est la limite des plus grandes erreurs*. He did not mention rejection. Furthermore, his statement had to do with the arithmetic mean rather than with individual observations.

5. Dependence of observations

5.1. Independence of observations. Mathematicians of the 19th century are known for tacitly assuming that they studied independent events. However, stronger statements (GNEDENKO & SHEYNIN 1992, p. 225) claiming that up to the end of that century no one mentioned this assumption, were wrong. Indeed, DE MOIVRE (1756, p. 5), for one, while formulating the multiplication theorem noted that he considered independent events, but I restrict my study to the treatment of observations.

LAPLACE mentioned independence of observations in the *Duxiéme Supplèment* to his *Théorie analytique* and in a memoir published in 1827 (SHEYNIN 1977, p. 11). This is not to say that he was infallible: STIGLER (1986, p. 151) noticed that, in a memoir of 1823, LAPLACE had failed to take into account the dependence of certain observations.

In several cases GAUSS (1809,§ 175; 1823, §§ 15 and 19; 1828, § 22) assumed that observations were independent. In other instances (GAUSS 1823, § 39) he wrote that (in modern notation), for errors of observations x and y,

$E(x^m y^n) = Ex^m Ey^n$. Yet elsewhere GAUSS (1845, p. 143) declared that independence was an essential condition *die häufig genug . .. von Gelehrten von Fach außer Acht gelassen wird.*

In 1805 LEGENDRE (STIGLER 1986, p. 59) had to solve a system of four equations

$$a_i x + b_i y + l_i = v_i - v_{i+1},\ i = 1, 2, 3, 4$$

where the v's were the errors of (five) observations. He did not proceed directly (taking $v_i - v_{i+1} = w_i$,). Instead, he introduced a fifth equation (identity), $v_3 = v_3$ and thus got rid of the *physical dependence* of his equations, see Sheynin (1993, § 3.6).

5.2. Dependence of observations (Gauss). GAUSS understood dependence in the same way as LEGENDRE tacitly did. Indeed he (1823, § 18) stated that two functions of observations were not independent if [at least] one of their arguments was common to both of them.[14] I have described his reasoning in connection with a remarkable proposal made by KAPTEYN (§ 5.4). Now, in addition, I can also refer to GAUSS (1845, p. 143):

Die Sicherkeit des Grundprincips [of taking the arithmetic mean] *beruhet auf einer wesentlichen Bedingung … die darin besteht, dass die an den einzelnen Beobachtungen … haftenden regellosen Störungen oder Schwankungen von einander ganz unabhängig sein müssen. Das Urteil, ob eine solche Unabhängigkeit vorhanden sei oder nicht, kann zuweilen sehr schwierig … sein, und wenn darüber Zweifel zurückbleiben, so wird auch das den Endresultaten beizulegende Gewicht ein precäres sein.*

I have quoted this passage (SHEYNIN 1973, p. 112) without, however, fully realizing its importance. Note that here GAUSS considered observations themselves rather than their functions as in 1823 (above). Discussing dependence, GAUSS (1826; 1828, § 3) singled out a special circumstance. I quote from the first source:

Sind … auch die Fälle nicht selten, … wo … die gegenseitige Abhängigkeit der beobachteten Grössen … durch gewisse Bedingungsgleichungen gegeben ist.

BESSEL (1838, p. 132), GERLING (1843, p. 26) and SCHREIBER (1882, p. 134) concurred with GAUSS without mentioning him or explaining their (common) attitude.

Consider an example: all three angles of a plane triangle are measured independently. However, since their sum should be equal to 180° (this is the condition), the angles are not independent. The only explanation, or, rather, specification that seems possible is that the observed angles remain independent whereas the adjusted angles, being corrupted by a common residual error, are interdependent.

On the heuristic level GAUSS'S understanding of independence of observations coincides with the axiomatic definition of independent events. In the axiomatic theory of probability events A and B are independent if

$$P(\mathrm{AB}) = P(\mathrm{A})P(\mathrm{B}),$$

whereas, according to GAUSS, this formula means that observations (not events) A and B have no common errors.

5.3. Dependence of observations (later authors and geodetic practice). GAUSS'S implicit definition of dependence took root in astronomy, geodesy and metrology (below). German authors likely knew his ideas. Others, as I presume, came to the same conclusion all by themselves, although PUISSANT (1832) might have noted LEGENDRE'S treatment of dependent equations (§ 5.1).

One of the essential problems confronting observers in geodesy was station adjustment. Suppose that an observer, standing at a certain point (station), measures the directions to points A, B, C, and D, and that, owing to unfavourable meteorological conditions, he has to take only two directions at a time.

How then should he adjust his observations to present them as a series of independent measurements? Indeed, only such measurements should be entered in the general adjustment of his network. This is a special problem and I discuss it insofar as it bears on my subject.

The first whom I mention is HANSEN (1831, p. 191). He stated that, in MAUPERTUIS'S triangulation, two angles were not independent since each of them contained one and the same angle as its separately measured component.

BESSEL (1838, Chapter 3) noted that, in considering the general adjustment of triangulation, GAUSS (1828) had restricted his study to the case of independently observed directions. He then went on to discuss the general case.

Much more interesting is SCHREIBER'S explanation (l882, p. 134):

Die beobachteten Werthe [directions A, B, … observed at a certain station] *sind unabhängig von einander, wenn jeder aus besonderen Beobachtungen abgeleitet ist. Sind also eine oder mehrere Beobachtungen, die zur Herleitung von A gedient haben, auch zu der von B benutzt werden, so sind A und B abhängig von einander.*

SCHREIBER'S terminology is not sufficiently precise: *beobachteten Werthe aus Beobachtungen*. His idea is, however, clear. He continues (p. 135):

Die durch Stationsausgleichung abgeleiteten Winkel- oder Richtungswerthe sind demnach im Allgemeinen von einander abhängig. Es sollen dann aber auch nicht diese, sondern die unmittelbaren Beobachtungen oder die aus solchen gebildeten, von einander unabhängigen Mittel mit den Bestimmungen A, B, identifizirt werden.

Thus, in adjusting networks or chains of triangulation, corrections should be made to magnitudes which may be considered independent. The same principle still governs the rigorous treatment of traverses: their measured elements are angles and sides (legs), and exactly these elements should be corrected in the adjustment.

Another example was provided by the layout of the chains of triangulation both in India (CLARKE 1880, p. 257) and in the former USSR (though not in the USA): baselines were measured at the vertices of the quadrilaterals formed by these chains. This practice ensured a higher degree of independence (in GAUSS'S sense) of the chains and lent S*chönheit und Rundung* to the whole system (Gauss's expression, which he used in a letter to OLBERS 8.7.1824 (GAUSS 1903, p. 371), in discussing his efforts to avoid acute angles in his triangulation).

I am unable to prove that the layout was planned in either case in accordance with GAUSS'S views. However, having graduated (in 1951) from the Moscow Geodetic Institute, I remember well enough that we, students, were told that common errors led to dependence of observations and that, consequently, baselines should be measured at the intersections of the chains of triangulation. The American practice was justly looked down upon. True, GAUSS was not mentioned on these occasions, but he was often remembered with utmost respect. It is hardly amiss to add that two volumes of GAUSS'S *Selected Geodetic Works* have since been published in Russian (in 1957 – 1958).

I adduce pronouncements of several more authors and begin by noting that AIRY (1879, pp. 51 and 60 – 70) invented a special term, *entangled measures*, to describe observations corrupted, in part, by common errors.

1) PUISSANT (1832). On p. 125 he writes out the formula for the error of a function of several observations whose errors are *indépendantes les unes des autres*. Then, on p. 128, he applies this formula (which is his *seconde règle*) to a certain case *en supposant qu'aucune loi ne lie ces erreurs* [of the observations] *entre elles, et en se conformant par conséquent à la seconde règle ci-dessus*.

2) BERTRAND (1888, p. 264):

Ces valeurs ne sont pas indépendantes. Les erreurs commises … sont liées l'une et l'autre. … La théorie des moyennes n'est pas applicable.[15]

3) CZUBER (1891, p. 2). While introducing *zwei Gattungen von Fehlerursachen*, he stated:

Der einen Gattung von Ursachen gegenüber erscheinen die einzelnen Beobachtungen als völlig unabhängige Ereignisse, d. h. die Wirkung dieser Ursachen ist durch Umstände bedingt, welche von einer Beobachtung zur nächsten sich ändern und mit dieser selbst in keinem nachweisbaren Zusammenhange stehen.

Obviously, in adopting GAUSS'S definition CZUBER stretched it mercilessly.[16]

4) KU (1967, p. 311), a metrologist:

A sequence of measurements showing a trend or pattern are not independent measurements.

5.4. Dependence and correlation. Statisticians did share GAUSS's view in that common causes lead to dependence (LANCASTER 1972, p. 300):

Galton, Weldon and Pearson all believed that correlation or lack of independence in the distribution of attributes was usually brought about by the possession of random elements in common.

Indeed, PEARSON (1920, p. 199) quoted GALTON'S general pronouncement read in 1888 to this effect:

Co-relation must be the consequence of the variations of the two organs being partly due to common causes.

The mathematical theory of correlation, as developed later, necessarily shed all direct links with biology. Furthermore, when studying correlation, modern statisticians are more likely to discuss relations of cause and effect (spurious correlation brought about by common causes is an exception). Thus, the problem of dependence became one of the topics where statistics departed from the theory of errors. PEARSON'S categorical statement (1920, p. 187) was a step in the wrong direction. He declared that

There is no trace in Gauss' work of observed physical variables being, apart from equations of condition, associated organically which is the fundamental conception of correlation.

GAUSS did not study variables; furthermore, PEARSON'S phrase is difficult to understand and I would restate it as follows: Equations of condition unite the observed magnitudes; as considered by GAUSS, however, these magnitudes were not connected with one another, whereas statisticians study variables associated organically with each other.

PEARSON continued (p. 192):

We must … hold that they [GAUSS and BRAVAIS] *contributed nothing of real importance to the problem of correlation*.

Regrettably, statisticians of the 19^{th} century did not notice GAUSS'S idea of common causes leading to dependence and thus did not arrive at correlation theory earlier than they actually did. In offering this remark I am not thwarted by GALTON'S statement made in 1908 that the error theory aims at getting rid of, or at allowing for errors, whereas he, GALTON, desires to study them.[17]

EISENHART (1978, p. 382), who quoted GALTON, adds that

When Karl Pearson and G. Udny Yale began to develop the mathematical theory of correlation … they found that much of the mathematical machinery that Gauss had devised for finding <u>*best*</u>

values for the parameters of empirical formulas [?] *by the method of least squares, was immediately applicable in correlation analysis*.

In regretting PEARSON'S stiff attitude and the ensuing gap between statistics and the error theory I have followed SEAL (1967, p. 208).

In 1912, being dissatisfied with the (statistical) theory of correlation, KAPTEYN made an attempt to introduce (after Pearson) correlation theory into astronomy. Or, rather, he aimed at quantifying the connection between two functions depending on partly coinciding observations and, consequently, defined a new coefficient of correlation (SHEYN1N 1984, § 9.2.1). KAPTEYN'S work remained unnoticed, perhaps because he did not mention GAUSS. A second possible reason is that World War I was about to begin. I do not think that his attempt can lead to essential development, but, at the very least, it can be used to estimate the degree of dependence of observations in geodesy. Of course, it showed once again the viability of GAUSS'S idea.

6. Alternative views

Here I describe the attempts made by several authors to challenge, or substantiate the initial propositions and the approaches to the theory of errors. One such approach is stochastic; it is based on limit theorems and is associated with LAPLACE. The other one, which I venture to call astronomical, was initiated by GAUSS. It connects the theory with optimal properties of statistical estimators and does not depend on the number of observations. Furthermore, his *Theoria combinationis* makes no use of the theory of probability at all.

At least up to the 1920's, most distinguished scholars spoke out against the astronomical approach. Their pronouncements are virtually forgotten, and, for the time being, I restrict my attention to one example.

LÉVY (1925, p. 74) maintained that *certains auteurs* had committed a mistake by considering

La notion de précision d'une mesure comme une notion première sur laquelle ils prétendent fonder la théorie des erreurs.

LÉVY thought that the initial notion was the law of probability of observational errors. He did not refer to GAUSS at all, thus perpetuating the nasty tradition of former French mathematicians, cf. § 6.3.

6.1. The Arithmetic Mean. GAUSS (1809, § 177), in his first substantiation of the MLSq, assumed that the arithmetic mean of a series of observations was the most probable value of the constant sought. This *postulatum* (BERTRAND 1888, p. 176) met with disapproval not only because it led to a single law of error (the normal law), but on other grounds as well. DE MORGAN (1845) and GLAISHER (1872, p. 102) held that extreme observations were obviously worse than those in the middle of a series. NEWCOMB (1886) voiced the same opinion adding that the mean was sensitive to a possible rejection of an outlier. Such criticisms were not new at all: even DANIEL BERNOULLI (1778, §§ 2 and 5) asserted that the mean was advisable only in the case of equal probability of all errors and assumed that small errors were more probable than large ones.

PEARSON (1978, p. 268), while commenting on BERNOULLI, aptly remarked that, being more probable, small errors would be more frequent and *have their due weight in the arithmetical mean.* Note, however, that NEWCOMB was right in that the median is, of course, safer than the sample mean. For the comparison of these two estimators see SHEYNIN (1977, § 8.3) and STIGLER (1973). I do not discuss this comparison here. The more general problem of introducing order statistics deserves to be considered separately.

PEARSON'S remark effectively refutes the idea of posterior weights (weights, assigned in accordance with the location of the corresponding observations in the series). In addition, if the law of distribution of the errors is symmetric, the introduction of such weights will not essentially change the usual arithmetic mean.

GLAISHER (1872, p. 87) put on record a previously neglected (and still forgotten) point:

In effect, Gauss's view [was] *that the arithmetic mean is practically the best mode of combining … observations. … he was very* [?] *far from asserting … that the arithmetical mean is the most probable value of the quantity observed.*

Indeed, here are GAUSS'S own words (1809, § 177):

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 Werth gebe.

ENCKE (1832) attempted to base the choice of the arithmetic mean on deterministic axioms.[18] KNOBLOCH (1985, pp. 580 – 586) reviewed many similar attempts; also see ZOCH (1935 – 1937). The constructive aspect of this approach laid the foundation for the modern theory of invariant tests and estimators (LEHMANN 1959, chapter 6).

The following passage from GAUSS'S letter to ENCKE of 23.8.1831 (GAUSS 1900, pp. 145 – 146) describes his opinion about this line of development:

Nicht ohne Interesse habe ich aus Ihrem Briefe den Gang gesehen, den Sie zur Rechtfertigung des Verfahrens, das arithmetische Mittel zu nehmen, eingeschlagen haben. Ich finde diesen Gang sehr beifallswerth, insofern auf die Frage, was zu thun sei, eine von allen Betrachtungen der Wahrscheinlichkeitsrechnung ganz unabhängige Antwort gegeben werden soll. Nur kann ich nicht wohl einräumen, das, was man auf diese Art erhält, den wahrscheinlichsten Werth zu nennen. In der That ist die Aufgabe, den wahrscheinlichsten Werth zu finden, eine mathematisch ganz bestimmte, die aber ihrer Natur nach die Kenntnis des Fehlergesetzes voraussetzt und nur in dem einzigsten Falle, wo dieses durch die Form e^{-kxx} *ausgedrückt wird, auf die arithmetischen Mittel führt.*

6.2. The Estimate of Precision. Denote the observations of constant a by $x_1, x_2, \ldots, x_n$ ($x_1 \le x_2 \le \ldots \le x_n$). Then (BERVI 1899) it is sufficient to estimate the precision by the equality

$$P(x_1 < a < x_n) = 1 - (1/2)^{n-1}. \qquad (9)$$

BERVI'S opinion was hardly justified. First, the length of the interval [x_1; x_n] tends to increase with n so that formula (9) becomes useless. Second, BERVI did not discuss indirect observations. Obviously, he did not account for the single approach to estimating precision in the two cases (direct and indirect observations) which is characteristic of the GAUSSIAN theory of errors.

KORNFELD (1955), whose note was communicated by an eminent physicist (LEONTOVICH), repeated BERVI'S proposal without referring to him.

Of course, statisticians began to calculate confidence intervals for quantiles (e. g, for medians of series of observations) long before 1955, but they did not suggest rejecting all other methods of estimating precision. I do not mention such measures of precision as, for example,

$$\frac{\sum v_i}{n-\rho}, \ \frac{\sum |v_i|}{n-\rho},$$

notation as in formula (7). These were mostly recommended in order to make computations easier. Of course, the use of one or another estimator should be determined by the law of distribution of the appropriate errors, but *usually* both the arithmetic mean and the mean square error reign supreme.

Nevertheless DELAMBRE (1912, pp. 39 and 235), in estimating the precision of direct observations, introduced measures similar to (10), only the denominators were n instead of (n – ρ). In spite of this obvious mistake (also made by GAUSS in 1814, see § 3.6), his measures were fit for comparing the precision of two or more series of measurements of differing lengths. Since DELAMBRE (p. 258) referred to a book published in 1818, he must have written his contribution between 1818 and 1822, the year of his death.

6.3. The astronomical approach.

Gauss and Laplace are representatives of two absolutely different opinions on [approaches to] *the meaning of the method of least squares* [MLSq]. *In Laplace's work we find a rigorous* [!] *and impartial study of this problem. His analysis shows that the results of the MLSq enjoy a more or less substantial probability only when the number of observations is large whereas Gauss attempted to attach absolute meaning to this method* [a damned lie!], *using extraneous considerations. If we turn our attention to the fact that all the essence of the Theory of chances is contained in the law of large numbers,*[19] *and that all the properties of random phenomena take real importance only when the number of trials is large, it would not be difficult to perceive the correctness of the Laplacean inference. However, when the number of observations is limited, we cannot at all reckon upon the mutual cancellation of errors and any combination of observations can … lead as much to the increase of errors as to their decrease*.

A detailed comment on this strange pronouncement made by TZINGER (1862, p. 1), the future President of the Moscow

Mathematical Society (1886 – 1891), is hardly needed. Nevertheless, I emphasize two points. First, GAUSS'S approach (both in 1809 and 1823) was *astronomical* (§ 6) and nowadays we know that the least-squares estimators possess certain optimal statistical properties. Second, LAPLACE, in following the stochastic approach, did not restrict his attention to the limiting case. Here are two statements from his *Essai* (LAPLACE 1814/1995, pp. 45 and 47):

The total weight of the result of the various systems is the sum of their individual weights.

Each angle [of a triangle] *should be decreased by the third of their sum in order that the weight* etc.

This second pronouncement was contained in the edition of 1816 of the *Essai*, and it might have been said that LAPLACE even preceded GAUSS in stating the principle of maximum weight. However, his was only an elementary example and it is important to consider what meaning did he attach to the notion of weight. Thus, the first statement was accompanied by a remark to the effect that the appropriate density was normal. Cf. my discussion of this point (1977, pp. 50 – 52).

It is easy to contrast LAPLACE and GAUSS the way TZINGER did, but it is much more correct to say that the Masters founded the theory of errors by tacitly studying each other's contributions and filling in the missing pieces. TZINGER was not the only, or the first one to reject the astronomical approach. In 1852 B1ENAYMÉ (HEYDE & SENETA 1977, § 4.3) voiced a similar opinion. In general, French mathematicians sided with LAPLACE; even POISSON (1833), in his obituary of LEGENDRE, properly referred to LAPLACE but failed to mention GAUSS.

6.4. The first substantiation of the MLSq. It is generally known that GAUSS himself found fault in his earlier justification of the method.The first, or one of the first to concur with him was GALLOWAY (1839). However, as early as in 1825, IVORY (1825, p. 7) spoke out in favour of substantiating least squares by the principle of least variance. He (p. 9) remarked that it *has been usual to introduce the doctrine of probabilities in order to explain this theory*. Such explanations, as IVORY held, were impossible without knowing *the general function for expressing the probability of an error*. Furthermore (p. 81), *there is even no proof* that this function does not change with time.

Consequently, he attempted to demonstrate the MLSq *without having recourse to the doctrine of probabilities*. This remark is also important. As I indicated in § 6, GAUSS was another who, in 1823, did not base his work on the theory of probability.

LAPLACE'S stochastic approach (§ 6) obviously evades IVORY'S criticism. However, IVORY (1825, p. 10), not failing to mention it, somehow brushes it aside, and on pp. 164 – 165 attempts to prove that LAPLACE'S derivation of least squares (in spite of its professed invariance with respect to the law of distribution of the errors of observation) leads to the normal law. This is a misunderstanding. It is known that LAPLACE based his work on the principle of minimal absolute expectation of errors in the unknowns sought, whereas IVORY attributed to him the principle of maximum likelihood.[20] No wonder that he successfully arrived at his conclusion! Later authors dismissed IVORY'S contribution. CZUBER (1891, pp. 301 – 304), partly supporting his work on ELLIS'S remark (below), correctly pointed out his logical mistakes. I would add that IVORY did not make any use of GAUSS'S "Theoria combinationis", perhaps not having read it yet.

GAUSS himself, in a letter to OLBERS of 15.3.1827, see SCHILLING (1909, p. 475), called IVORY'S demonstration a *petitio principii*. He considered IVORY'S paper (1825) *unter aller Kritik* and continued: *Welche Verworrenheit, Unklarheit und völliger Mangel logischer Bündigkeit*. GAUSS adduced only one definite argument (the one later voiced by ELLIS and repeated by CZUBER) to the effect that IVORY'S derivation of the principle of least squares on pp. 163 – 164 could have just as well led to an indefinite number of other principles.[21]

I turn now to a certain point in GAUSS'S *Theoria motus*. GAUSS (§ 179) stated that, just as his *postulatum*, the principle of least squares *muss überall … als Axiom gelten*. Quoting this pronouncement, I (1979, p. 31) called it a *certain deviation from his* [Gauss's] *main train of thought.* WATERHOUSE (1990, pp. 45 – 46) insists, however, that GAUSS made a

Quite accurate summary of his argument and indeed proved that the assumption about arithmetic means implies the more general statement.

He also admonishes me for omitting GAUSS'S words *mit demselben Recht* from the quotation above. I cannot agree, and for my part I should believe that Gauss *was never really satisfied with his first substantiation*. Cf. GLAISHER'S remark ((§ 6.1) from which it follows that the normal law emerges as the law of error only in some ideal case.

From the later criticisms of GAUSS'S substantiation I single out MERRIMAN'S comment (1877, p. 165). He noticed that GAUSS'S [normal] distribution was not strictly *a law of facility of error but only a law of distribution of residuals*. Indeed, and, in addition, GAUSS (1809, § 175) formulated the properties of usual random errors, but made use of them only indirectly, through the principle of the arithmetic mean.

At least one author (HENKE 1894, p. 44) stated that the principle of least squares *vielleicht nicht mit Unrecht* might be left unjustified. On p. 67 he added:

Würde ich für zweckmässig halten wenn man zur Begründung der Methode der kleinsten Quadrate die Wahrscheinlichkeitstheorie überhaupt nicht mehr in Anspruch nehmen würde.[22]

Finally, CAMPBELL (1928, pp, 156 – 167) published a fierce attack against the Gaussian theory of errors [against the MLSq as developed by GAUSS]. lt seems, however, that he never heard of the second substantiation of the method by GAUSS. Accordingly, CAMPBELL was hopelessly late, but his mistake shows once again how popular was GAUSS'S *Theoria motus*.

6.5. The second substantiation of the MLSq. This contribution, *Theoria motus*, had introduced the principle of least squares in such an elegant way that, despite GAUSS'S mature thoughts, astronomers continued to adhere to it since, by and large, their observations did obey the normal law. [See also Sheynin (2017, § 9A.4-10, p. 148)].

BERTRAND (1888, p. 267) noted that GAUSS'S second substantiation still demanded a fixed density function $\varphi(x)$ of errors. Indeed,

$$\varphi(x) = \varphi(0) + x\varphi'(0) + \tfrac{1}{2}x^2 \varphi''(0) + \ldots$$

and, for small errors x and even densities,

$$\varphi(x) = \varphi(0) + \tfrac{1}{2}x^2 \varphi''(0) = a + bx^2.$$

The only answer to this remark is that, obviously, nothing better than GAUSS'S second substantiation had as yet been offered. Later objections to this substantiation, of which I mention two, are interesting at least by themselves. POINCARÉ (1896, p. 188), a natural scientist as well as a mathematician, called GAUSS'S change of mind *assez étrange*. And recently HAR'IER (1977, p. 28) stated that

Gauss's second exposition seems … to be no more satisfactory than his first. In each case he starts from a postulate, plausible but not universally valid, which leads inexorably to the foregone conclusion.
The author is perhaps justified in his criticism and, as is clear from the title of his book, he was hardly obliged to add that the least-squares estimators nevertheless possess certain optimal statistical properties; cf. § 6.3.

Acknowledgements. This paper presents part of a research programme on the history of the theory of errors, performed at the Mathematical Institute of the University of Cologne, with the support of the Axel-Springer-Stiftung. In reading the MS of this paper, Professors J. PFANZAGL and R. L. PLACKETT offered valuable comments.

Notes

1. The corollary which DELAMBRE mentioned indeed had intuitive appeal, but he evidently made too much of it.

2. KARL PEARSON was also in the habit of computing an excessive number of digits. In 1936, E. B. ROESSLER (*Science*, vol. 84, pp. 289 – 290) pointed out that among statisticians no uniformity of practice exists in the retention of significant figures and that, accordingly, a very misleading impression of the accuracy of the results can be created. He offered several examples, one of them pertaining to R. A. FISHER. A discussion in the same volume of *Science* followed, see pp. 437, 483 – 484, and 574 – 575.

3. In a special note prefixing COLCLOUGH'S paper, the Editor (p. 167) called his subject controversial and promised to publish an article by R. COLLE presenting the alternative view. Obviously, however, this contribution did not appear either in the same periodical or elsewhere.

4. Or, rather, the precision of the estimators of its unknowns. I shall not repeat this remark anymore.

5. Cf. the hybrid method of solving systems (1) in § 1.2.

6. Readers can follow up KOLMOGOROV'S non-parametric substantiation (1946) of least squares which made use of the notions of the n-dimensional vector geometry. KOLMOGOROV (pp. 59 – 60) attributed to GAUSS the statement that, for normal errors of observation with known variance, the normed errors of the least-squares estimators have standard normal distribution. He did not provide a definite reference and I believe that it should have been GAUSS (1809, § 182). It seems, however, that this particular proposition is not really important since the variance is seldom known. It was STUDENT who determined the appropriate distribution in the usual case.

7. The case of equal probabilities of an event either happening or failing to happen in a given number of trials interested DE MOIVRE (1756, e. g., Problem 5) whereas LODEWIIK and CHRISTIAAN HUYGENS, in their correspondence with each other, introduced the probable duration of life (1669).

8. STIGLER (1986, p. 140n) pointed out a downright mistake made by DELAMBRE(1810, p. 182). LEGENDRE'S method, the latter maintained, *consiste à égaler à zéro la somme des carrés de toutes les erreurs …*

9. Note the double sign. GAUSS also adduced the probable error of another group of his observations and calculated the mean of the two errors, but in both these cases he prefixed the error with the sign "plus".

10. I also note that CRAMER (1946, § 27.4) derived the formula

$$\operatorname{var} m^2 = \frac{\mu_4 - \mu_2^2}{n} - \frac{2(\mu_4 - \mu_2^2)}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}$$

where $m^2 = [vv]/n$ and μ_2 is the second central moment;

11. This attitude was abandoned in FISHER'S design of experiments.

12. Here also the authors obviously thought of changing the conditions of observation.

13. DELAMBRE'S historical research was mentioned in GAUSS's correspondence no less than nine times (on six occasions by GAUSS himself), each time negatively. See my paper (1993, § 3.7, Note 33).

14. In § 19 he added that those functions were linear, otherwise his statement contradicts the Student – Fisher theorem on the independence of the sample variance and the arithmetic mean.

15. For a long time natural scientists including astronomers used this term rather than *theory of errors* (SHEYNIN 1986, pp. 310 – 312). LAMBERT'S *Theorie der Fehler* was forgotten, as I thought, until the mid-19th century, when it was put into circulation anew. Now, I correct myself: our present term was also used by BESSEL (1820, p. 166; 1838, p. 36), again independently of LAMBERT. In the second instance BESSEL wrote *Theorie der zufälligen Beobachtungsfehler*.

16. CZUBER'S contribution is useful even now and I feel it necessary to add that he was one of the first (after Poisson) to introduce, on p. 7, the integral distribution function $\Phi(x)$ defining the density as the derivative of $\Phi(x)$. True, he did not make any use of that function.

17. While estimating the precision and accuracy of observations, the theory of errors thereby studies the relevant errors.

18. I also quote CHEBYSHEV (1936, p. 233):

If two observations are available, it is possible to assume as obvious that their best combination will be the arithmetical mean, because in this case nothing allows us to prefer one of the observed quantity to the other one. … However, it is not possible to say the same about three or more observations.

19. This statement is much too strong.

20. IVORY (p. 164) states that LAPLACE, in his *Théorie analytique* (p. 319), arrived at the MLSq by considering the matter a little differently. Obviously, he referred to the edition of 1820 of LAPLACE'S classic. At present, the standard edition is that of 1886 (*Oeuvr. Compl.*, t. 7) and the corresponding reference should be to p. 323. However, on p. 324 LAPLACE goes on to say *On peut parvenir au même résultat de cette manière* and expounds his main train of thought. More: IVORY'S claim that LAPLACE'S [preliminary] derivation of least squares was similar to his own was also misleading.

21. In the same letter to OLBERS, GAUSS criticized, with some reservation, IVORY'S contribution *über die Pendellängen*. In 1826 – 1830 IVORY published a few papers (all of them in the same periodical, the *London, Edinburgh und Dublin*

Philosophical Magazine und Journal) on the derivation of the earth's eccentricity by pendulum observations, but GAUSS'S criticism was once more justified.

22. That GERLING (1843) introduced the principle of least squares axiomatically is not really interesting since he compiled his book as a practical treatise.

References

C. F. GAUSS: Works

(1809) *Theoria motus* … German transl.: Aus der Theorie der Bewegung der Himmelkörper etc. In GAUSS (1887, pp. 92 – 117).
(1816) Bestimmung der Genauigkeit der Beobachtungen. Ibidem, pp. 129 – 138.
(1823) Theoria combinationis. German transl.: Theorie der den kleinsten Fehlern unterworfenen Combination der Beobachtungen. Ibidem, pp. 1 – 53.
(1826) Selbstanzeige of GAUSS (1828). Ibidem, pp. 200 – 204.
(1828) Supplementum theoriae combinationis. German transl.: Theorie der Combination etc., Ergänzüng. Ibidem, pp. 54 – 91.
(1845) Anwendung der Wahrscheinlichkeitsrechnung auf die Bestimmung der Bilanz für Witwenkassen [2.] Nachlass. *Werke*, Bd. 4. Göttingen, 1880, pp. 125 – 157.
(1887) *Abhandlungen zur Methode der kleinsten Quadrate*. Hrsg., A. BÖRSCH & P. SIMON. Berlin. Vaduz, 1998.
(1900, 1903) *Werke*, Bde. 8, 9. Göttingen – Leipzig.
(1975) *Werke. Ergänzungsreihe*, Bd. 1. Hildesheim.

C. F. GAUSS: CORRESPONDENCE

[GAUSS, BESSEL] (1880), *Briefwechsel zwischen Gauss und Bessel*. Leipzig. Reprinted: GAUSS (1975). Quotations in my text are from this reprint.
PETERS, C. A. F. (Hrsg.) (1860 – 1865), *Briefwechsel zwischen Gauss und Schumacher*, Bde 1 – 6. Altona.
SCHÄFER, G. (Hrsg.) (1927), *Briefwechsel zwischen Gauss und Gerling*. Berlin.
SCHILLING, C. (1900 – 1909), *W. Olbers. Sein Leben und sein Werk*. Bd. 2, Abt. 1 – 2, this being *Briefwechsel zwischen Gauss und Olbers*. Berlin.

OTHER AUTHORS

AIRY, G. B. ([1861] 1879), *On the Algebraic and Numerical Theory of Errors of Observations* etc. London.
BARNETT, V. & LEWIS, T. ([1978] 1984), *Outliers in Statistical Data*. Chichester a. o.
BERNOULLI, D. (1770 – 1771), Mensura sortis ad fortuitam successionem rerum naturaliter contingentium applicata (BERNOULLI 1982, pp. 326 – 360).
BERNOULLI, D. (1778), Dijudicatio … English transl.: The most probable choice etc. See KENDALL (1961).
BERNOULLI, D. (1780), Specimen philosophicum de compensationibus horologicis etc. (BERNOULLI 1982, pp. 376 – 390)**.**
BERNOULLI, D. (1982), *Werke*, Bd. 2. Hrsg., B. L. VAN DER WAERDEN. Basel.
BERTRAND, J. (1888), *Calcul des probabilités*. Paris. Reprints: 1889, 1907, 1970, 1972.
BERVI, N. V. (1899), Determining the most probable value of the observed object apart from Gauss's postulate. *Bull. Imp. Moskovskoe Obshchestvo Liubitelei Estestvosnania, antropologii i etnografii*, Otdelenie fizich. nauk, vol. 10, No. 1, pp. 41 – 45 (in Russian).
BESSEL, F. W. (1816), Untersuchungen über die Bahn des Olbersschen Kometen. *Abh. Preuss. Akad. Wiss.* [Berlin], math. Kl., 1812 – 1813, pp. 119 – 160.
BESSEL, F. W. (1820), Beschreibung des auf des Königsberger Sternwarte. *Astron. Jahrb.* für 1823, pp. 161– 168. Berlin.
BESSEL, F. W. (1826), Methode die Thermometer zu berichtigen. *Abhandlungen*, Bd. 3. Leipzig, 1876, pp. 226 – 233.
BESSEL, F. W. (1838), *Gradmessung in Ostpreußen*. Berlin.
BREEN, H. (1849), Correction of Lindenau's elements of the orbit of Venus. *Monthly Notices Roy. Astron. Soc*., vol. 9, pp. 49 – 51.
CAMPBELL, N. R. (1928), *An Account of the Principles of Measurement and Calculation*. London a. o.

CHEBYSHEV, P. L. (1936), *Theory of Probability*. Lectures read in 1879 – 1880 as written down by A. M. LIAPUNOV. Ed., A. N. KRYLOV. Moscow – Leningrad. (In Russian.)
CLARKE, A. R. (1880), *Geodesy*. Oxford.
COLCLOUGH, A. R. (1987), Two theories of experimental error. *J. Res. Nat. Bureau Stand.*, vol. 92, No. 3, pp. 167 – 185.
COURNOT, A. A. ([1843] 1984), *Exposition de la théorie des chances et des probabilités*. Ed., B. BRU. Paris.
CRAMER, H. (1946), *Mathematical Methods of Statistics*. Princeton.
CZUBER, E. (1891), *Theorie der Beobachtungsfehler*. Leipzig.
DALE, A. I. (1991), T. Bayes's work on infinite series. *Hist. Math.*, vol. 18, pp. 312 – 327.
DELAMBRE, J. B. J. (1810), *Rapport historique sur les progrès des sciences mathématiques depuis 1789 et sur leur état actuel*. Paris. Reprint: Amsterdam, 1966.
DELAMBRE, J. B. J. (1814a), Analyse des travaux de la Classe des sci. math. et phys. de 1'Institut, pendant l'année 1811. *Mém. cl. math. et phys. Inst. Imp. de France* année 1811, pt. 2, first paging, i – lxxviii. Paris.
DELAMBRE, J. B. J. (1814b), *Astronomie théorique et pratique*, t. 2. Paris**.**
DELAMBRE, J. B. J. (1912), *Grandeur et figure de la terre*. Paris.
DE MOIVRE, A. (1718/1756), *Doctrine of Chances*. London**.**
DE MORGAN, A. (1845), Theory of probabilities. In *Enc. Metropolitana*. Pure sciences, vol. 2, pp. 393 – 490. London.
DIXON, W. J. (1962), Rejection of observations. In: *Contributions to Order Statistics*. Eds A. E. SARHAN & B. G. GREENBERG. New York – London, pp. 299 – 342.
DORSEY, N. E. & C. EISENHART (1969), On absolute measurements. (KU 1969, pp. 49 – 55.)
DREYER, J. L. E. (1890), *Tycho Brahe*. Edinburgh.
EISENHART, C. (1963), Realistic evaluation of the precision and accuracy of instrument calibration. (KU 1969, pp. 21 – 47.)
EISENHART, C. (1968), Expression of the uncertainties of final results. (KU 1969, pp. 69 – 72.)
EISENHART, C. (1978), Gauss. In: *Intern. Enc. of Statistics*, vol. 1, pp. 378 – 386. Eds. W. H. KRUSKAL & JUDITH M. TANUR. New York – London.
ENCKE, J. F. (1832), Über die Bergründung der Methode der kleinsten Quadrate. *Abh. Kgl. Akad. Wiss. zu Berlin*, Math. Kl., 1831, pp. 73 – 78.
ENCKE, J. F. (183 – 1836), Über die Methode der kleinsten Quadrate. *Ges. math. und astron. Abh.*, Bd. 2, pp. 1 – 200. Berlin, 1888.
ERMAN, Ad. (Hrsg.) (1852), *Briefwechsel zwischen Olbers und Bessel*, Bde. 1, 2. Leipzig.
FORBES, E. G. (1978), The astronomical work of C. F. Gauss. *Hist. Math.*, vol. 5, No. 2, p. 167 – 181.
FOURIER, J. B. J. (1824), Régle usuelle pour la recherche des résultats moyens etc. *Bull. sci. math., astron., phys. et chim.* this being *Bull. universelle des sci.*, premier sect., t. 2, pp. 88 – 89. A supplement written by DEFLERS (pp. 89 – 90) contains a numerical example. DEFLERS also states that FOURIER read his note at the *Société philomatique* le 3 juil, dernier.
FOURIER, J. B. J. (1826), Mémoire sur les résultats moyens etc. (FOURIER 1890, pp. 525 – 545). **S, G,** 88.
FOURIER, J. B. J. (1829), Second mémoire sur les résultats moyens etc. (FOURIER 1890, pp. 551 – 590). **S, G,** 88.
FOURIER, J. B. J. (1890), *Oeuvres*, t. 2. Pari**s.**
GALLOWAY, T. (1839), *A Treatise on Probability*. Edinburgh.
GERLING, Ch. L. (1839), *Beiträge zur Geographie Kurhessens* etc. Cassel.
GERLING, Ch. L. (1843), *Die Ausgleichungsrechnung* etc. Hamburg – Gotha.
GIACOMO, P. (1981), News from the BIPM. *Metrologia*, vol. 17, pp. 69 – 74.
GLAISHER, J. W. L. (1872), On the law of facility of errors of observations etc. *M.em Roy. Astron. Soc.*, vol. 39, pp. 75 – 124.
GLEISSBERG, W. (1964), Zur Begründung des Auftretens zufälliger Beobachtungsfehler. *Sterne*, Bd. 40, No. 5 – 6, 105 – 108.

GNEDENKO, B. V. & O. B. SHEYNIN ([1978] 1992), The theory of probability. In: *Mathematics of the Nineteenth Century*, [vol. 1], pp. 211 – 282. Eds, A. N. KOLMOGOROV & A. P. YUSHKEVICH. Basel. Transl. from Russian.
HANSEN [**P. A.**], (1831), Über die Anwendung der Wahrscheinlichkeitsrechnung auf geodätische Vermessungen etc. *Astron. Nachr.*, Bd. 9, No. 202, pp. 189 – 204.
HARTER, H. L. (1977, date of preface), *A Chronological Annotated Bibliography on Order Statistics*, vol. 1. No place. Publ. by the US Air Force and several of its sub-units.
HAUBER, C. Fr. (1832), Theorie der mittleren Werthe. *Z. f. Phys. Math.*, Bd. 10, pp. 425 – 457.
HELMERT, F. R. (1872), *Die Ausgleichungsrechnung* etc. Leipzig.
HENKE, R. (1868/1894), *Über die Methode der kleinsten Quadrate*. Leipzig.
HEYDE, C. C. & E. SENETA (1977), *I. J. Bienaymé*. New York.
HUYGENS, C. ([1669] 1895), Correspondance. *Oeuvr. compl*., t. 6, pp. 531 – 532. La Haye.
IVORY, J. (1825 – 1826), On the method of least squares. *Lond., Edinb. Dublin Philos. Mag. & J*., vol. 65, pp. 1 – 10, 81 – 88, 161 – 168; vol. 68, pp. 161 – 165.
JORDAN, W. (Hrsg.) (1882), *Höhere Geodäsie und Topographie des Deutschen Reiches.* Stuttgart.
KAPTEYN, J. C. (1912), Definition of the correlation-coeflicient. *Monthly Notices Roy Astron. Soc*., vol. 72, No. 6, pp. 518 – 525.
KENDALL, M. G. (1961), Daniel Bernoulli on maximum likelihood. Incorporates English transl. of BERNOULLI (1778). *Biometrika*, vol. 48, pp. 1 – 18. (**E. S. PEARSON & M. G. KENDALL** 1970, pp. 155 – 172).
KENDALL, SIR MAURICE & R. L. PLACKETT (Eds) (1977), *Studies in the History of Statistics and Probability*, vol. 2. London.
KNOBLOCH, E. (1985), Zu Grundlagenproblematik der Fehlertheorie. In: *Festschrift für Helmuth Gericke*. Hrsg. M. FOLKERTS et al. Stuttgart, pp. 561 – 590.
KOLMOGOROV, A. N. (1946), On the substantiation of the method of least squares. *Uspekhi math. nauk*, vol. 1, No. 1, pp. 57 – 70 (in Russian)**.** From 1945 that periodical is being translated as *Russian Math. Surveys.*
KORNFELD, M. (1955), On the theory of errors. *Doklady Akademii Nauk SSSR,* vol. 103, No. 2, pp. 213 – 214 (in Russian).
KOTZ, S. & N. L. JOHNSON (Eds) (1982 – 1988), *Encyclopedia of Statistical Sciences*, vols 1 – 9. New York.
KRUSKAL, W. H. (1960), Some remarks on wild observations. (KU 1969, pp. 346 – 348).
KU, H. H. (1967), Statistical concepts in metrology. (KU 1969, pp. 296 – 330).
KU, H. H. (Ed.) (1969), *Precision measurement und calibration.* Selected Nat. Bureau of Standards; papers on statistical concepts and procedures. *NES Sp. Publl. 300,* vol. 1. Washington.
LANCASTER, H. O. (1972), Development of the motion of statistical dependence. (KENDALL & PLACKETT 1977, pp. 293 – 308.)
LAPLACE, P. S. ([1814] 1820), *Essai philosophique sur les probabilités.* The edition of 1820 was reprinted in LAPLACE'S *Oeuvr. compl.*, t 7, No. 1. Paris, 1886, with separate paging. In my text, I refer to the *Oeuvr. compl.* English transl. by A. I. Dale: New York, 1995. *Philosophical Essay on Probabilities.*
LEHMANN, E. L. (1959), Testing statistical hypotheses. New York – London.
LÉVY, P. (1925), *Calcul des probabilités*. Paris. '
LIAPUNOV, A. M. (1975), On Gauss's formula for estimating the measure of precision of observations. *Istoriko-mathematicheskie issledovania*, vol. 20, pp. 319 – 328, in Russian. Posth. publ. with my comments.
MERRIMAN, M. (1877), List of writings relating to the method of least squares etc. Reprint: Stigler (1980, vol. 1).
NEWCOMB, S. (1886), A generalized theory of the combination of observations etc. *Amer. J. Math.*, vol. 8, pp. 343 – 366. Reprint: Stigler (1980, vol. 2).
PEARSON, E. S. & M. G. KENDALL (Eds) (1970), *Studies in the History of Statistics and Probability*, vol. 1. London.
PEARSON, K. (1920), Note on the history of correlation. *Biometrika*, vol. 13, pp. 25 – 45. (PEARSON & KENDALL 1970, pp. 185 – 205.)

PEARSON, K. (1978), *The History of Statistics in the 17th and 18th Centuries*. Lectures 1921 – 1933. Posthumous publ. by E. S. PEARSON. London.
POINCARE, H. (1896) *Calcul des probabilités*. Paris, 1912, 1923, 1987.
POISSON, S. D. (1833), Discours prononcé aux funérailles de M. Legendre. *J. reine und angew. Math*., Bd. 10, pp. 360 – 363. **S, G,** 58.
PUISSANT, L. (1832), Deuxiéme mémoire sur l'application du calcul des probabilités aux mesures géodésiques. *Mém. Acad. Roy. Sci. de l'Inst. de France*, t. 11, pp. 123 – 156.
SCHMEIDLER, F. (1984), *Leben und Werk des Königsberger Astronomen F. W. Bessel*. Kelkheim/T.
SCHREIBER, [**O.**] (1879), Richtungsbeobachtungen und Winkelbeobachtungen. *Z. für Vermessungswesen*, Bd. 8, pp. 97 – 149.
SCHREIBER, [**O.**] (1882), Die Anordnung der Winkelbeobachtungen im Göttinger Basisnetz. Ibidem, Bd. 11, pp. 129 – 161.
SEAL, H. L. (1967), The historical development of the Gauss linear model. *Biometrika*, vol. 54, pp. 1 – 24. (PEARSON & KENDALL 1970, pp. 207—230.)
SHEYNIN, O. B. (1971), J. H. Lambert's work on probability. *Arch. Hist. Ex. Sci*., vol. 7, No. 3, pp. 244 – 256.
SHEYNIN, O. B. (1973), Mathematical treatment of astronomical observations etc. In this collection.
SHEYNIN, O. B. (1977), Laplace's theory of errors. Ibidem, vol. 17, No. 1, pp. 1 – 61.
SHEYNIN, O. B. (1979), C. F. Gauss and the theory of errors. Ibidem, vol. 20, No. 1, pp. 21 – 72.'
SHEYNIN, O. B. (1983), Corrections and short notes on my papers. Ibidem, vol. 28, No. 2, pp. 171 – 195.
SHEYNIN, O. B. (1984), On the history of the statistical method in astronomy. Ibidem, vol. 29, No. 2, pp. 151 – 199.
SHEYNIN, O. B. (1986), Quetelet as a statistician. Ibidem, vol. 36, No. 4, pp. 281 – 325.
SHEYNIN, O. B. (1988), C. F. Gauss and the chi-square distribution. *NTM Schriftenreihe Gesch. Naturwiss., Technik, Med*., Bd. 25, pp. 21 – 22.
SHEYNIN, O. B. (1993), Treatment of observations in early astronomy. *Arch. Hist. Ex. Sci*., vol. 46, pp. 153 – 192.
SHEYNIN, O. B. (2007), True value of a measured constant and the theory of errors. *Historia Scientiarum*, vol. 17, pp. 38 – 48.
SHEYNIN, O. B. (2017), *Theory of Probability. Historical Essay*. Berlin. **S, G,** 10.
STIGLER, S. M. (1973), Laplace, Fisher, and the discovery of the concept of sufficiency. *Biometrika*, vol. 60, pp. 439 – 445. (KENDALL & PLACKETT 1977, pp. 271 – 277.)
STIGLER, S. M., Ed. (1980), *American Contributions to Mathematical Statistics in the 19th Century*, vols. 1 – 2. New York. Lacks single paging.
STIGLER, S. M. (1986), *The History of Statistics*. Cambridge, Mass**.**
STRUVE, F. G. W. (1824), Über das Universalinstrument etc. *Astron. Nachr*., Bd. 2, pp. 431 – 440.
STRUVE, F. G. W. (1831), *Breitengradmessung in den Ostseeprovinzen Russlands*, Tl. 1. Dorpat.
STRUVE, F. G. W. (1860), *Arc du méridien*, t. l, Pétérsbour**g.**
TAYLOR, JOHN R. (1982), *An Introduction to Error Analysis*. Mill Valley, Calif.
TZINGER, V. YA. (1862), *Method naimenshikh kvadratov*. [Method of least squares.] Thesis. Moscow, in Russian.
VOGLER, C. A. (1902), *Lambert und die practische Geometrie*. Berlin.
WATERHOUSE, W. C. (1990), Gauss's first argument for least squares. *Arch. Hist. Ex. Sci*.,vol. 41, No. 1, pp. 41 – 52.
WHITTAKER, E. T. & G. ROBINSON (1924/1958), *Calculus of Observations.* London – Glasgow.
ZOCH, R. T. (1935—1937), On the postulate of the arithmetical mean. *Annals Math. Stat*., vol. 6, pp. 171 – 182; vol. 8, pp. 177 – 178.

VI

Gauss, Bessel and the adjustment of triangulation

Historia Scientiarum, vol. 11, No. 2, 2001, pp. 168 – 175

I am dwelling on Bessel's indirect conflict with Gauss over the adjustment of triangulation, a subject overlooked by Biermann [1966] and other commentators, e.g., May [1972]. I begin by touching on Biermann's paper who discussed the relations between the two scholars and provide my own relevant material in §§ 3 and 4. I also have to explain the two main patterns of adjusting geodetic measurements (§ 2), to formulate some conclusions and offer related considerations (§ 5). Finally, I devote § 6 to several new unexpected dramatis personae.

1. Biermann [1966]

Bessel, as well as other scientists, felt that Gauss should have publicly acknowledged the work done by his predecessors. In my context, the evident example concerns Legendre's priority in discovering the principle of least squares (PrLSq)[1], and Bessel (Felber 1994, p. 174, this being Bessel's letter to Humboldt of 19.4.1844) rhetorically asked, "Warum citiren nun viele Gauss und schweigen von Legendre?" True, he did not fail to note either the [much] greater significance of Gauss's publication [of 1809], or that Gauss had communicated the PrLSq to him (to Bessel) before Legendre's memoir appeared in print.[2]

Bessel was disappointed at Gauss's delays in publication and at his devoting time to geodetic work in the field; moreover, in 1837 and 1839 Bessel spoke his mind and thus embittered Gauss.

Only in passing Biermann (p. 15) mentions the meeting between Gauss and Bessel in 1825. Without documenting his remark, he states that Gauss's "herbeigewünschten Aussprache" did not take place "da noch mehrere Astronomen anwesend waren". Bruhns (1869, p. 108n) reported something else:

Das Zusammentreffen zwischen Gauss und Bessel wird weder in dem Briefwechsel zwischen Gauss und Schumacher, noch zwischen Olbers und Bessel erwähnt; ich habe die Nachricht von einem Ohrenzeugen, der auch erwähnte, dass Gauss Bessel wegen wissenschaftlicher Meinungsverschiedenheit sehr hart angelassen hätte.

The subject of disagreement remains however unknown.

2. Adjustment of geodetic observations

Two main versions of adjusting redundant geodetic observations in accord with the PrLSq were elaborated in the early 19th century. Since the approximate values of the magnitudes sought were almost always known, both issued from linear systems of "physically" (hence, linearly) independent equations.

1. The pattern of indirect observations [Legendre 1805]. Given, a system of equations

$$L_i = a_i x_1 + b_i x_2 + \ldots + s_i = 0,\ i = 1, 2, \ldots, n \qquad (1)$$

with coefficients indicated by the appropriate theory and measured free terms. An acceptable solution will render

$$a_i x_1 + b_i x_2 + \ldots + s_i = v_i \qquad (2)$$

with small residuals v_i The least-squares solution is obtained from the condition

$$W = \sum v_i^2 = \min \qquad (3)$$

which leads to the "normal" equations

$$\frac{\partial W}{\partial x_1} = \frac{\partial W}{\partial x_2} = \ldots = 0.$$

Two examples are: a) Determination of the two parameters of the earth's ellipsoid of rotation by redundant meridian arc measurements;

b) Determination of the unknown coordinates of a station in the field by redundant measurements of angles at this station between other stations with given coordinates (the so-called Potenot problem) [Gauss 1823].

2. The pattern of conditioned observations. Equations (1) and additional linear conditions

$$M_j = \alpha_j x_1 + ß_j x_2 + \ldots + q_j = 0,\ j = 1, 2, \ldots, + r \qquad (4)$$

are given, and the requirement (3) now becomes

$$W + \sum \lambda j M j = \min.$$

The problem is thus solved by the classical method of Lagrange multipliers. When describing his forthcorning pertinent contribution (of 1828) Gauss [1826, p. 200] had indeed noted that

Dieser Fall von dem anderen [see Item 1] *nicht wesentlich, sondern bloß in der Form verschieden ist.*

The free terms in (4) are the calculated discrepancies between observation and theory; for example, between the sum of the three measured angles of a triangle and the theoretical sum, or a similar difference between the measured length of the second baseline (considered error-free) and the same length calculated from the first baseline through the connecting chain of triangles.

When adjusting a triangulation, most or even all the equations (1) are of the form

$$x_j - q_j = 0 \qquad (5)$$

where $q_j > 0$ is the measured (but not yet adjusted) value of angle j and pattern 2 becomes simple enough which explains why it became popular in geodesy.

Strange as it seems, the first to describe the adjustment of conditional observations as shown above was apparently Helmert [1872, p. 197]. Neither Gauss, nor Bessel (my §§ 3 and 4) provided a readily understandable account.

I ought to emphasize that Adrain [1808], described an adjustment of a traverse with measured azimuths and lengths of sides. In actual fact, he issued from equations of the type (4) with equations (5) being implied. His notation and exposition were clumsy; his contribution became known only by the end of the 19th century; and most commentators reasonably paid attention to Admin's main result, to his (not at all rigorous) derivation of the normal law of error rather than to his adjustment procedure. Hogan [1977] apparently proved that Adrain's paper had actually appeared in 1809. Also see Hammer [1900], Sheynin [1965] and Dutka [1990].

3. Galloway [1844]

Thomas Galloway (1796 – 1851) was a mathematician, astronomer and actuary, Fellow of the Royal Society. Here is his letter to Bessel. London, March 12, 1844.

My dear Sir, I have had the pleasure of receiving your kind letter dated the 22nd of January last,[3] *and I beg to thank you, very sincerely, for the remarks with which you have honoured my Memoir on a portion of our English Survey. I*[n] *drawing up this paper I have no hope of being able to throw any new light on a subject which had passed through your hands; but it occurred to me that at this particular time I might possibly do some small service to Geodesy by calling the attention of my countrymen, through the Astronomical Society, to your methods, and giving an easy example of their application.*

At present the results of the observations made in the course of our Ordnance Survey for determining the British Arc of Meridian, through its whole length from Dunnose to Balta (*in Shetland*) *are preparing* [are being prepared] *for publication; and we have a Survey going forward at the Cape of Good Hope from which we may expect to have, ere long, a considerable extension of Lacaille's Arc.*[4] *As there is every reason to expect that in these two great operations the observations, both astronomical and geodetic, will have a very high degree of precision, it is particularly desirable that the computations should be made in such a manner as to preserve to* [for] *the results all the weight which the excellence of the observations is capable of giving them; and this, I am of opinion, can only be done by adopting the methods of computation which you have so fully explained in the* <u>Gradmessung</u> [Bessel 1838a]. *Fortunately, these two Surveys are under the direction of men who understand their subject too well to require any assistance, but it is of some advantage to computers to have the method explained in their own language, and an example worked out at length from Observations with which they are familiar.*

With respect to the two passages which you have been so good as to point out as not quite accurate, I must beg to apologize for them as follows: In alluding to Struve's[5] *method of observing my object was simply to call attention to the important advantage which arises from observing every signal visible from the station where the instrument is placed, instead of taking only particular angles, as was the usual practice in the early part of our Survey. Having little practical acquaintance with these matters, and this mode of observation not having been used with respect to the observations I have undertaken to calculate, I did not perhaps sufficiently advert to the difficulty which will immediately arise when the signals are not all visible contemporaneously, and which is only to be overcome by having recourse to the method you have explained in* §§ *22 and 26 of the* <u>*Gradmessung*</u> [see my § 5].

As to my citation of your measure of the Prussian degree as an Example of <u>*Gauss's formula*</u>, *it is very clear that I have carelessly and inadvertently used the words* <u>*Gauss's formula*</u> *for "general method". I feel greatly obliged to you for noticing this error, for as the volume of the Memoirs in which the paper will appear is not yet published, the mistake may be corrected in the* <u>*Errata*</u>. *Such mistakes are unfortunate, inasmuch as the*[y] *tend to introduce confusion into the history of discovery.*

At the last meeting of the Royal Astronomical Society the translation of a Letter from you to Sir John Herschel, on the effect of gravity in obtaining the shape of a meridian circle, was read and [one word undecipherable] *with great interest by the members present.*

I quote now from Galloway's memoir [1846, pp. 28, 29 and 60n] respectively.

a) … *previously to the publication of Gauss's* <u>Supplementum</u> [1828], *an equivalent and similar method was applied by Rosenberger* [1827] … *and he* … *states that the method which he followed had been long before communicated to him by Bessel*.[6]

b) *The best example of the application of Gauss's formulae is given by Bessel, in his invaluable work on the measure of the meridional degree in Prussia* [Bessel 1838a], *a work in which every thing connected with the subject of geodetic measurements, and the comparison of celestial and terrestrial arcs, is treated in the fullest detail, and according to the best methods at present known. Another example has also been recently given by Bessel in the calculation of the triangles for extending the French arc of meridian through Spain* [Bessel 1841].[7]

c) *The form in which theory is here stated is the same as given by Rosenberger* [1827] …; *a fuller development is however given by Gauss* [1828].

This note seems to have been superfluous. Finally, here is the *Errata* to the volume of the *Memoirs*, concerning p. 29:

for Gauss's formula, <u>*read*</u> *the general method. The formulae in the work referred to are due to Bessel*.

In the last instance, Galloway again apparently thought of Rosenberger.

4. Gerling [1861]

Gauss's student, Christian Ludwig Gerling (1788 – 1864), was an astronomer and geodesist. In 1861, about 15 years after Bessel's death, he commented on two priority conflicts concerning the MLSq. First he touched on the Legendre – Gauss strife and mistakenly accused von Zach of failing to defend Gauss [Sheynin 1999, pp. 257 – 258 and p. 264, note 7].

Gerling then went on to describe Bessel's attempts to establish his priority over Gauss in the adjustment of triangulation by least squares. It occurs that in 1843, soon after Gerling's treatise [1843] had appeared, he became involved in an "ausführliche polemische Correspondenz" with Bessel. The latter did not even mention Gauss and claimed that "er [Bessel] doch zuerst über die Sache habe drucken lassen". He was offended by Gerling [1843, p. 32] who had adduced a list of literature mentioning Legendre and Gauss and noting that Encke had published an essay[8] studying their work "mit Rücksicht auf Arbeiten von Bessel und Hansen".

Gerling reasonably explained to Bessel that his list only contained the sources which he had used in compiling his treatise, but Bessel remained unsatisfied. Gerling also correctly indicated that although [Gauss 1828] certainly appeared later than [Rosenberger 1827], the description of the pertinent method was contained already in [Gauss 1826], and in Gauss's letter to Gerling of 30.10.1823 [Gauss 1927, pp. 300 – 302], as I would add.

Lastly, Bessel was vexed at Gerling's failure to note his merits "um Entwickelung der Wahrscheinlichkeits-Theorie". Gerling explained that he had not required the use of the probable error (applied by Bessel in 1815 and formally introduced by him in 1816) but Bessel apparently remained_unhappy.[9] Bessel hardly mentioned his attempt at proving the central limit theorem (my§ l, note 1) to Gerling because the latter's treatise did not dwell on such topics at all, and it could be doubted whether Bessel's final complaint had anything to do with probability theory excepting the probable error.

5. Conclusions and Related Considerations

Thus, both Galloway (§ 3) and Gerling (§ 4) referred to Bessel himself through Rosenberger [1827], and to Gauss [1828] with Galloway properly mentioning "an equivalent and similar method" and Gerling additionally citing [Gauss 1826]. Bessel's claim seems therefore unfounded or nearly so. In addition, he applied the PrLSq much more formally than Gauss did; and he could have well emphasized this difference instead of stressing the identity of their methods.

Such an attitude would have not however be advantageous for him. Indeed, when measuring angles, Gauss hardly ever kept to any definite number of observations, but he assigned equal weights to all the angles.[10] On the other hand, Bessel [1838a, §§ 15 and 34] measured directions and included every visible (but therefore not every adjacent) signal in one single set [Ibidem, §§ 22 and 26; mentioned by Galloway, see my § 3]. He then formally calculated the (differing) weights of the observed directions, and, unlike Gauss, jointly adjusted all the measurements of a given network. His calculations thus had to be more difficult whereas the unavoidable systematic errors likely corrupted the final results more than they did when the triangulation was being adjusted in two stages (first, the stations, and then the network as a whole).[11]

Because of obvious organizational considerations, other practitioners never followed the Gaussian method of observations. The number of necessary measurements had to be established beforehand. An additional argument here was that otherwise systematic errors would have been excluded to a lesser extent. However, networks continued to be adjusted in two stages, and in this respect Bessel had, figuratively speaking, lost to Gauss; his approach was all but forgotten.

6. New figures

Gerling [1861], see my § 4, concluded his report by stating that *Schleiermacher in Darmstadt … auch selbstständig die Ausgleichungsformeln aufgestellt hat*.

He referred to Eckhardt without providing an exact bibliographic description. Here, however, it is [Eckhardt 1835, pp. 131 – 132]:

Die Formeln zur Einführung der kleinsten Quadrate bei geodätischen Rechnungen waren von Schleiermacher, nach eigenen Ansichten, mit einer Genauigkeit und Eleganz, entwickelt worden, wie man sie von diesem gewandten Analysten nicht anders gewohnt ist, und der Wahrheit zur Steuer muss ich bekennen, ohne jedoch ein Prioritätsrecht für meinen Freund hierdurch in Anspruch nehmen zu wollen, dass alle diese Formeln bereits längst in meinen Händen waren, ehe diese Anwendung öffentlich zur Sprache kam. Nur die überaus weitläufigen Rechnungen, welche die Methode der kleinsten Quadrate bei Dreiecksnetzen von einiger Ausdehnung verursacht, hatten mich abgehalten, früher davon Gebrauch zu machen, bis ich endlich so glücklich war, in Herrn Dr. Hügel einen Gehilfen zu erhalten, der mit den nothwendigen theoretischen Kenntnissen ausgerüstet, die erforderliche Gewandtheit im Zahlenrechnen verband und Beharrlichkeit genug besaß, sich einer so langwierigen Arbeit mit Ausdauer hinzugeben.

After consulting the relevant volumes of two celebrated sources, Poggendorf's *Biographisch-Literarisches Handwörterbuch* and the Royal Society *Catalogue of Scientific Papers,* I think that

1. Hügel (whoever he was) hardly published anything on my subject.

2. The same holds for Ludwig Schleiermacher (1785 – 1844), an author of a contribution on the influence of refraction on "Kreismicrometer Beobachtungen" which appeared in 1808.

Anyway, Eckhardt did not say that Schleiermacher (or he himself) had introduced the pattern of conditional observations, whereas the pattern of indirect observations was first publicly described by Legendre [1805]. Nevertheless, recalling that Gauss published a methodological note [1823] on the adjustment of a simple geodetic construction, it is worthwhile to put on record that Schleiermacher was apparently one of the first to work successfully on such problems.

Notes

l. Hardly remembered is Legendre 's later article [1814] where he reprinted his original text. Bessel himself could have well been annoyed by Gauss's response to his study [1838b] of the behaviour of observational errors. Gauss (letter to Hessel 28.2.1839, W-8, pp. 146 – 147) indicated that he had read it "mit grossem Interesse" largely caused by the "Darstellung" rather than by the "Sache selbst" which was familiar to him since many years ago. Bessel's work was mainly devoted to proving a version of the central limit theorem (a later term); according to modern standards, it was proved by Chebyshev (not in full rigor), then by Markov and Liapunov. W-i is, generally, my abbreviation for Gauss's *Werke*, Bd. i.

2. With regard to the latter statement I repeat [Sheynin 1993, p. 51] that Bessel first made it in 1832, publicly. Bessel 's attitude towards Gauss had indeed changed over the years. On 29.9.1812, in a letter to Olbers (*Briefwechsel zwischen W. Olbers und F. W. Bessel*, Bd. 1, Leipzig, 1852, p. 345. Hrsg., Ad. Erman), he stated that Gauss "war doch der Erfinder der moindres quarrés."

3. Sent to John Herschel, also see below. In my context, the published extract from this letter is not interesting.

4. N. L. Lacaille (1713 – 1762), a French astronomer. In 1739 – 1740 he verified the measurement of the meridian arc from Dunkirk to Perpignan.

5. Vasily Jakovlevich, or (Friedrich Georg) Wilhelm, von Struve (1793 – 1864), a Russian astronomer and geodesist of German extraction.

6. 1 quote this author from his p. 1:

Bereits vor längerer Zeit teilte mir mein hochverehrter Lehrer, Herr Professor Bessel, eine ihm eigenthümliche Methode mit, aus geodätischen Vermessungen die wahrscheinlichsten Resultate herzuleiten.

7. With regard to adjustment procedures this contribution contained nothing new.

8. Encke [1834 – 1836, p. 186] mentioned Rosenberger [1827] and Gauss [1828] on a par but followed the latter.

9. Now, in 1861, Gerling was able to mention Gauss's negative opinion about the probable error voiced in a letter to Schumacher of 2.2.1825 (*W*-8, p. 143). On its history see Sheynin [1979, § 4.4]. I did not find a single reference to this error in the *Gradmessung* [Bessel 1838a]; on the contrary, it contained computations of the mean square error not less than in eight sections.

10. See extracts from his field observations in Bd. 9 of his *Werke*. I discussed this point elsewhere [Sheynin 1979, § 6.2; 1994, § 4.1; 1996, pp. 97 – 99]. And at least in one instance Gauss [1927, p. 687, this being his letter to Gerling of 17.4.1844] estimated the overall precision of his observations at several stations rather than at each of them separately. Not trusting his own formula when the number of observations was small (and, in any case, systematic errors were also present), he preferred, apparently with due justification, to consider all of them equally reliable.

11. For large networks it is therefore better to accomplish this second stage step by step: adjust each chain and replace it by the appropriate geodetic line; adjust all these lines taken together; go back to the chains and adjust them definitively, cf. Sakatov [1957, p. 438]. The introduction of geodetic lines instead of chains is due to Helmert [Sheynin 1995, § 3.3].

Bibliography

Adrain, R., 1808: "Research concerning the probabilities of errors which happen in making observations." Reprinted in Stigler [1980, vol. 1]. **S, G,** 111.

Bessel, F. W., 1838a: *Gradmessung in Ostpreussen*. Berlin.

Bessel, F. W., 1838b: "Untersuchungen über die Wahrscheinlichkeit der Beobachtungsfehler". *Abh*., Bd. 2, pp. 372 – 391.

Bessel, F. W., 1841: "Über einen Fehler in der Berechnung der französischen Gradmessung. *Abh*., Bd. 3, pp. 55 – 62.

Bessel, F. W.. 1876: *Abhandlungen*, Bde 1 – 3. Leipzig.

Biermann, K.-R., 1966: "Über die Beziehungen zwischen Gauss und Bessel". *Mitt. Gauss-Ges. Göttingen*, No. 3, pp. 7 – 20.

Bruhns, C., 1869: *J. F. Encke*. Leipzig.

Dutka, J., 1990: "Adrain and the method of least squares". *Arch. Hist. Ex. Sci*. vol. 41, pp. 171 – 184.

Eckhardt, C. L. P., 1835: “Vorläufige Nachricht von den geodätischen Operationen zur Verbindung der Observatorien”. *Astron. Nachr.* Bd. 12, pp. 129 – 134. Also in 1834: *Arch. ges. Naturlehre,* Bd. 26 (= *Arch. Chem. Meteorol.,* Bd. 8), pp. 297 – 308 which contained nothing else on my subject.
Encke, J. F., 1834 – 1836: Über die Methode der kleinste Quadrate”. Reprinted in author’s *Ges. math. und astron. Abh.,* Bd. 2. Berlin, 1888, pp. 1 – 200.
Felber, H.-J., Editor, 1994: *Briefwechsel zwischen A. von Humboldt und F. W. Bessel*. Berlin.
Galloway, T., 1844: “Letter to F. W. Bessel of 12.3.1844”. Archiv, Berlin-Brandenburgische Akademie der Wissenschaft. Bessel papers, No. 229.
Galloway, T., 1846: “On the application of the method of least squares to the determination of the most probable error of observation in a portion of the Ordnance Survey of England". *Mem. Roy. Astron. Soc.*, vol. 15, pp. 23 – 69.
Gauss, C. F., 1823: “Anwendung der Wahrscheinlichkeitsrechnung auf eine Aufgabe der practischen Geometrie". *Abh*., pp. 139 – 144
G**auss, C. F,** 1826: “Supplementum Theoria Combinationis, Selbstanzeige”. *Abh*., pp. 200 – 204.
Gauss, C. F., 1828: “Supplementum Theoria Combinationis". German translation: *Abh*., pp. 54 – 91.
Gauss, C. F., 1870 – 1930: *Werke,* 12 Bde. Göttingen a.o.
Gauss, C. F., 1887: *Abhandlungen zur Methode der kleinsten Quadrate*. Hrsg, A. Boersch, P. Simon. Latest reprint: Vaduz, 1998.
Gauss, C. F., 1927: *Briefwechsel zwischen C. E Gauss und Ch. L. Gerling.* Hrsg., C. Schaefer. Reprinted in Gauss’s *Werke*, Ergänzungsreihe, Bd. 3. Hildesheim, 1975.
Gerling, Ch. L. 1843. *Die Ausgleichungsrechnung* etc. Hamburg – Gotha.
Gerling, Ch. L. 1861: “Notiz m Betreff der Prioritäts-Verhältnisse in Beziehung auf die Methode der kleinsten Quadrate". *Nachr. Georg-August Univ. und Kgl. Ges. Wiss. Göttingen,* pp. 273 – 275.
Hammer, E., 1900: “Zur Geschichte der Ausgleichungsrechnung”. *Z. Vermessungswesen*, Bd. 29, pp. 613 – 628.
Helmert, F. R., 1872: *Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate*. Leipzig. Later editions: 1907 and 1924.
Legendre, A. M., 1805: *Nouvelles méthodes pour la détermination des orbites des comètes*, Appendice. Paris.
Legendre, A. M., 1814. “Méthode des moindres quarrés”. *Mém. Cl. sci. math. et phys. Acad. Sci. Paris,* t. 11, pt. 2, année 1810, pp. 149 – 154. Lu 24 Sept. 1811.
May, K. O., 1972: “Gauss”. *Dict. Scient. Biogr.*, vol. 5, p. 298 – 315.
Rosenberger, O. A., 1827: “Über die, auf Veranstaltung der französischen Academie, während der Jahre 1736 und 1737 in Schweden vorgenommene Gradmessung”. *Astron. Nachr*., Bd. 6, No. 12, pp.1 – 32.
Sakatov, P. S., 1957: *Lehrbuch der höheren Geodäsie*. Berlin. Orig. published in Russian, in 1953.
Sheynin, O., 1965: “On the work of Adrain in the theory of errors”. In Russian. **S, G,** 1.
Sheynin, O., 1979: “Gauss and the theory of errors”. *Arch. Hist. Ex. Sci.* vol. 20, pp. 21 – 72.
Sheynin, O., 1993: “On the history of the principle of least squares”. Ibidem, vol. 46, pp. 39 – 54.
Sheynin, O., 1994: “Gauss and geodetic observations”. In this collection.
Sheynin, O., 1995: “Helmert’s work in the theory of errors”. *Arch. Hist. Ex. Sci*, vol. 49, pp. 73 – 104.
Sheynin, O., 1996: *The History of the Theory of Errors*. Egelsbach.
Sheynin, O., 1999: “The Discovery of the principle of least squares”. *Hist. Scientiarum*, vol. 8, pp. 249 – 264. **S, G,** 112.
Stigler, S. M., Editor, 1980: *American Contributions to Mathematical Statistics in the 19th Century*, vols. 1 – 2. New York. No general paging. Reprints of original papers of many authors

VII

The Theory of Probability: Definition and Relation to Statistics

Arch. Hist. Ex. Sci., vol. 52, 1998, pp. 99 – 108

l. Introduction

How did the founders of the theory of probability define it and its aims? What were, and what are its relations with statistics? I discuss these questions in §§ 2 and 3; and, in § 4, I dwell on the stochastic theory of errors, which was once a chapter of probability and a source of ideas for statistics, but later, as I shall argue, became a separate entity under the dominion of statistics.

2. The Theory of Probability

2.1. The First Definition. Pascal (1963, pp. 101 – 103) was the first to suggest a name for the new discipline whose elements emerged as a scientific topic in his correspondence with Permat. In a letter of 1654 to the *Académie Parisienne des Sciences* (the predecessor of the official Academy) he wrote about his desire to compile a treatise devoted to the geometry of chance (*La Géométrie du hasard*). However, he did not mention any other applications of the new ideas and methods than the problem of points.

Huygens prophetically remarked that the study of games of chance lays the *fondements d'une spéculation fort intéressante et profonde* (1657), but he was unable to provide a definition. Nevertheless his correspondence and manuscripts, published during 1888 – 1920, included discussions of games of chance and dealt with important problems in the statistics of mortality.

Montmort, before reading Jakob Bemoulli's *Ars Conjectandi* (1713), mentioned games of chance and *les autres choses de la vie* and indirectly formulated his aims as offering *des règles infaillibles pour calculer les différences qui se trouvent entre diverses probabilités* (1713, p. ix). In addition he noted that [mathematical] analysis is used only to discover *des rapports constans & immuables entre des nombres & des figures*, whereas he will apply this discipline

Pour découvrir des rapports de probabilité entre des choses incertaines & qui n'ont rien de fixe, ce qui semble fort opposé à l'esprit de la Géomètrie (p. x).

Jakob Bernoulli expressly defined the *Ars Conjectandi sive stochastice* as the

Kunst so genau als möglich die Wahrscheinlichkeiten der Dinge zu messen und zwar zu dem Zwecke, dass wir bei unseren Urteilen und Handlungen stets das auswählen und befolgen können was uns besser, trefflicher, sicherer oder rathsamer erscheint.

Darin allein beruht die ganze Weisheit des Philosophen und die ganze Klugheit des Staatsmannes (1713, Ch. 2 of Part 4, p. 75).

Bernoulli hardly thought of calculating the probabilities of events by directly enumerating the appropriate favourable and unfavourable cases, or of equating, on the strength of his law of large numbers, the theoretical probability of an event to its statistical counterpart. Yet we may say that he was feeling his way to our modern definition (cf. § 2.4). His statement about the application of stochastic reasoning was also in keeping with modern ideas.[1]

If Bemoulli had any thought of applying the *Ars Conjectandi* to the natural sciences, we have no documentation of that. Nevertheless the above passage is important since it was the first definition of the theory of probability.[2] In agreement with this formulation Bemoulli intended to apply the new theory to *bürgerliche, sittliche und wirthschaftliche Verhältnisse* (see the title of the fourth Part of his treatise), but he had no time. Some of his preliminary deliberations on the issue can be found in his diary (1975)[3].

2.2. Chance and Design. De Moivre called his main book on probability *The Doctrine of Chances* (1718). The aims of this doctrine, he stated, were to serve, *in Conjunction with the other parts of the Mathematics, as a fit Introduction to the Art of Reasoning* (p. ii); *to help to cure a kind of Superstition, viz, that there is in Play such a thing as Luck, good or bad* (p. iii); but mainly to *establish a due comparison between Chance and Design* (p. v).

In 1718 he emphasized the last aim in his *Dedication* of the first edition of his book to Newton. Although he did not consider any other applications of probability than to games of chance and population statistics (mortality and sex ratio at birth), De Moivre concluded his Dedication by expressing the hope that the isolation of chance from design will

Excite in others a desire of prosecuting these studies, and of learning from your [Newton's] *Philosophy how to collect, by a just Calculation, the Evidences of exquisite Wisdom and Design, which appear in the Phenomena of Nature throughout the Universe*. (1756, p. 329).

Yes, for De Moivre the *Doctrine of Chances* was indeed a tool for distinguishing between chance and design by a *just Calculation* (and comparison of the appropriate probabilities). Quite a few scholars can be mentioned in this connection (e. g.. Derham, or Süssmilch) and Pearson (1926, p. 552) correctly noted that

De Moivre expanded the Newtonian theology and directed statistics into the new channel down which it flowed for nearly a century.

Unlike Montmort before him (§ 2.1), De Moivre, like Bayes after him, considered probability as a part of pure mathematics, about which his version of the 'De Moivre – Laplace' limit theorem clearly testifies.

2.3. The Modern Term. Laplace dwelt on the use of probability in his earlier memoirs stating that *le plus grand nombre des phénomènes* can only be studied stochastically.

Dans l'impossibilité de les connaitre leurs différents degrés de vraisemblance can be determined, *en sorte que nous devons a la faiblesse de l'esprit humain une des théories les plus délicates et les plus ingénieuses des Mathématiques, savoir la science des hasards ou des probabilités* (1776, pp. 144 – 145, cf. 1786, p. 296).

It would be more natural to attribute the origin of the theory of probability to positive rather than to negative circumstances, viz. to the existence of stochastic laws. It was the impossibility of studying most natural phenomena deterministically (as mentioned by Laplace himself), and not the feebleness of the mind, that led to the introduction of logical and subjective probability.

Like De Moivre (§ 2.2), Laplace had to isolate chance from design; more precisely, from order (he never mentioned divine will or intervention). He had to study small effects to decide, for instance, whether or not a certain observed magnitude significantly differed from zero, i.e. whether it really existed (Laplace, 1812, p. 361; Sheynin 1976, Epigraph).

Laplace coined the term *Théorie des probabilités*, for instance in the title of his treatise (1812)[4], but he did not define it; his celebrated sentence, *la théorie des probabilités n'est, au fond, que la bon sens réduit au calcul* (1814, p. CLIII), could just as well have been formulated with respect to mathematics of his time in general.

Laplace explained that the theory gave rise to new analytical methods (largely invented by himself), demanded *fine et délicate* logic and was extremely useful for studying the *Philosophie naturelle et des Sciences morales* (Ibidem). And, of course, he actually proved the last point by making important contributions to natural science (astronomy in the first place) as well as, to a lesser extent, to demography.

Laplace's theory of probability belonged to applied mathematics, although, as is the case with De Moivre (§ 2.2), some of his results are of utmost importance for pure science. First, he apparently considered himself an applied mathematician; thus, while introducing integrals of complex functions, he expressed the hope that the *géomètres* (1812/1886, p. 304) will become interested in this topic. Similar passages occur elsewhere (1774, p. 62, 1812/1886, p. 365) and Poisson justly remarked that for Laplace *l'analyse mathématique était un instrument qu'il pliait aux applications les plus variées* (1827, p. 20). Second, Laplace's theory of probability was insufficiently abstract. He did not introduce, even on a heuristic level, any notion of a random variable; hence he was unable to study densities or characteristic functions per se, so that his results did not admit of development and the theory had to be created anew.

By 1850 it had become clear that the contemporary theory of probability did not belong to pure mathematics (Öttinger 1852 , pp. iii and iv).[5] Poisson proved most important theorems non-rigorously, just as Laplace, but at the same time he (1812, p. 161) called Laplace's *Théorie analytique* (1812)

Un Traité complet de la théorie des hazards dans lequel on trouvera des méthodes uniformes et générales,

and even 25 years later he stated that the theory became *une des principales branches des mathématiques* (1837, p. 1). With regard to Laplace's classic I note that, except for the central limit theorem, the *méthodes* indeed pertained to mathematics in general.

2.4. The modern definition. According to a modern definition (Prokhorov & Sevastianov 1971, p. 540) the theory of probability is a mathematical science that enables one to determine the probabilities of random events if they are in certain ways connected with other events whose probabilities are known. This is exactly what Chebyshev (1845, p. 29) and Boole (1851, p. 251) stated more than a century earlier, although Boole discussed propositions rather than events.[6]

Von Mises (1928/1972, p. 36), the eminent critic of probability theory, was of a somewhat different (and, regrettably, forgotten) opinion: his definition mentioned collectives rather than events. At present a collective is a random sequence, and I think that a modern definition of probability theory should state that it studies random variables, their sequences (e. g. Markov chains) and systems (e. g. random vectors). After all, a random event is a particular case of a random variable.

Here I am going further than I went previously (1994, p. 337), I am now stating in essence that from Chebyshev's time mathematicians have been developing the theory by ever more fully using the power (Kolmogorov's expression) of the concept of random variable.

2.5. Pure or applied mathematics? Von Mises' example is all the more interesting since he insisted that the theory is *eine Naturwissenschaft gleicher Art wie die Geometrie oder die theoretische Mechanik* (1919b, p.58).[7] He did have a point with respect to probability, whose axiomatic construction occurred more than two decades later, but his attitude towards geometry, which by that time was axiomatized, was hardly sound. I also note that pre-axiomatic probability theory was extremely useful for natural sciences, civil life and demography.

However, it belonged to applied mathematics. Chebyshev's and Boole's definition of probability (§ 2.4) was not sufficient for elevating it to the realm of pure mathematics, but its very appearance, many decades before mathematics acquired its present more abstract nature, is an interesting fact.

3. Statistics

Poisson et al. (1835, p. 174) were the first to connect statistics explicitly with probability by maintaining that the former was *the functioning mechanism of the calculus of probability*. Quetelet, who was the most influential statistician of his time, paid lip service to probability theory; while advocating its use, he hardly ever applied it. What is more, after his death (1874) German statisticians began anathematizing him and declared that probability was not needed at all. It was Bortkiewicz (1904) who opposed this line. It is true that direct definitions of statistics did not (and do not) any longer connect it with probability, but the link is still felt. Thus, Pearson (1978, p. 3) stated that statistics is *the application of mathematical theory to the interpretation of mass observations*.

The relations between the two mathematical disciplines are fuzzy. Von Mises (1964a, p. l) included statistics within probability theory:

Certain classes of probability problems … are customarily designated as <u>theory of statistics</u> or <u>mathematical statistics</u>.

Neyman (1950, p. 4) was of the same opinion, whereas Kolnogorov (1948, p. 216) held that the theory of probability *must be considered its* [he meant statistics] *structural part*. He (p. 218) added, however, that

The system of the main concepts of theoretical statistics … is still in the making. Statistics only gradually ceases to be the <u>applied theory of probability</u>.

It is difficult to combine these two statements, but at in any case Kolmogorov's note (1948) was only an abstract of his report, and he hardly ever repeated them. Drawing on my conclusion in § 2.4, I may also venture to express Kolmogorov's first statement in another way: the study of random variables and their systems must be considered the structural part of the interpretation of observations (of realizations of the corresponding random variables).

Nevertheless it is generally thought best to consider the two disciplines as somehow separate, and m any case it is hardly natural to unite entities differing from each other with respect to their dependence or independence from induction or deduction. Indeed, probability theory is based on deduction, but statistics, considered in its entirety, is not altogether deductive since, for example, the estimation of such magnitudes as statistical probabilities is inductive.

No similar problem existed two hundred years ago when Laplace, while actually discussing the statistical approach (but never using the terms *statistics* or *statistical*), called it *un nouvelle branche de la théorie des probabilités* (1781, p. 383). He repeated Lagrange's statement (letter to Laplace of 13.1.1775; *Oeuvr.*, t. 14. Paris, 1892, p. 58).

4. The stochastic theory of errors

The aims of the theory of errors include the choice of optimal methods and circumstances of observations, the design of instruments that enable us to use such methods, etc.; in other words, the theory, in addition to its much better known stochastic branch, has a deterministic branch as well. Below I discuss only the stochastic theory of errors; I believe that the other branch belongs to experimental design (understood in its wider sense).

The stochastic theory of errors originated in the 18th century largely at the hands of Simpson and Lambert. Simpson proved that, for several laws of distribution, the arithmetic mean was preferable to a single observation, and Lambert introduced the principle of maximum likelihood. Thus, according to our modern point of view, the stochastic error theory belonged, from its very beginning, both to the theory of probability (Simpson) and to statistics (Lambert).

The same statement is valid for the work of Laplace and Gauss as well. Laplace applied a few different assumptions for deriving optimal values of the constants sought. His main condition was that the absolute expectation of the error should be minimal,

$$\mathrm{E}\,|\xi| = \int_{-\infty}^{\infty} |x|\varphi(x) = \min.$$

Here $\varphi(x)$ was the density of the errors of observation, normal on the strength of the central limit theorem whose several versions he extremely non-rigorously proved. In 1809, after introducing statistical considerations (the principle of maximum likelihood and the postulate of the arithmetic mean), he derived both the normal distribution and the principle of least squares. In 1823, he applied an alternative assumption (the condition of least variance) and again arrived at the same principle (this time, the *method* of least squares).

Thus, the stochastic theory of errors made use of its own theorems and assumptions, but the theorems now belong to probability theory and the assumptions were appropriated by statistics. My conclusion that error theory (or treatment of observations) is a separate independent discipline contradicts the opinion of several Russian authors, e. g., Romanovsky (1955)[8] or Bolshev (1989) who maintain that the stochastic theory of errors belong to mathematical statistics[9].

Three circumstances strengthen my conclusion. First, the theory of errors makes use of the notion of real value whereas statistics applies it only occasionally. Second, contrary to the opinion of those authors error theory has to study systematic errors.[11] Statisticians attempt to isolate them during exploratory data analysis which belongs to theoretical but not to mathematical statistics. Bolshev not only excluded this topic from error theory, but attributed it to data processing which seems to be a stillborn counterpart of exploratory analysis. Third, during the last few decades the theory of errors has not developed as a chapter of statistics.

I mention two contributions. Kemnitz (1957) noticed that many actual series of geodetic observations possessed negative excesses whereas Eddington (1933) had shown that the excess of a mixture of normal laws was positive. Kemnitz explained the situation by indicating that, complying with official manuals, practitioners had to reject outlying observations and thus to make use of truncated normal laws. This is an important conclusion although hardly interesting from the theoretical viewpoint.

The other writing is a survey Markuse (1985) of contributions on the treatment of geodetic observations published in 1976 – 1984. It convincingly proves that the theory of errors though certainly important, is an applied discipline.

Having originated as a chapter of probability theory, and remaining extremely important for that discipline until the 1920s[12], the theory of errors does not belong to it anymore.

Later note

I have mentioned Fourier's (indirect) introduction of the notion of *true value*. Now, I (2007) say that his innovation was left unnoticed but that many authors independently from him and from one another followed him. A corollary was also put on record: the unavoidable residual systematic error is necessarily included in the true value.

Second point. Pearson (1892, p. 15) stated that *the unity of science* (I would say, of a definite science) *consists alone in its method*. For statistics, it means that it is an independent science in spite of its lacking a certain subject.

Then, medical statistics is the application of the statistical method to medicine, and the theory of errors, its application to the treatment of observations (so that statisticians ought to be familiar with that theory).

I have not found any definition of *statistical method* and I think that it amounts to the arrangement and ordering of the statistical data and that therefore that method is inseparably linked with statistical theory, with mathematical statistics.

Notes

1. Newton's approach was much the same (Schell, 1960). '

2. I am using this term because Bernoulli's classical work embodied the first elementary theory of probability.

3. Recall also Leibniz' celebrated opinion (1765, Bd. 2, p. 515):

Ich habe schon mehr als einmal gesagt, dass man eine neue Art Logik braucht, die die Grade der Wahrscheinlichkeit behandelt.

4. He made use of other expressions as well. In the second *livre* of his book, devoted to probability proper, the terms 'Calcul', and 'Analyse des probabilités' are found as often as *Théorie*. In the *Essai* (1814) *Calcul des probabilités* occurs more often than *Théorie* and *Analyse* taken together, whereas in the (later) Supplements (1818c, 1819) to the *Théorie analitique* (1812) *Calcul* is already used almost exclusively. On rare occasions (although not in these Supplements), *Théorie*, or *Analyse*, or *Science des hasards*, and *Science des probabilités,* are also found.

During 1888 – 1925 Bertrand, Poincaré and Lévy called their treatises *Calcul des probabilités*, whereas its Russian equivalent, *Ischislenie veroiatnostei*, was the title of Markov's book (four editions, 1900 – 1924).

5. Anticipating Hilbert by about 50 years, Boole argued that in order to "rank among the pure sciences" probability theory should be "founded on principles of axiomatic nature" (1854, p. 288). In 1880 Chebyshev (§ 2.4, note 6) made a small heuristic step in the direction of an axiomatic theory.

6 For that matter, Chebyshev himself (1880/1936, p. 148), perhaps in a somewhat simplifying manner, argued that

The aim of the theory of probability is to determine the chances of the occurrence of a certain event; the word 'event' means anything whose probability is being determined.

He continued: *Thus, in mathematics, the term probability serves to denote some magnitude that is to be measured*. Note that he defined both terms, event and probability, in an abstract manner.

7. Hilda Geiringer apparently did not convincingly describe Mises' attitude; she stated that, for von Mises, there *were never two different theories* [of probability], *one 'pure', the other 'applied', but one theory only, a frequency theory, mathematically rigorous and guided by an operational approach.* (von Mises 1964a, p. v).

8. A prominent statistician and mathematician, creator of the Tashkent statistical school.

9. In a previous note Romanovsky (1939, p. 726) argued, however, that the theory of errors was *a most important field of applying probability theory* and did not mention statistics at all. Western authors, whether statisticians or natural scientists, hardly discussed this topic. They were apparently not interested in delimiting the related scientific disciplines from each other or in defining their goals. Thus, treatises on the adjustment of geodetic measurements published a few decades ago (Rainsford 1957; Grossmann 1961) are highly disappointing in that they do not dwell on the classical theory of errors.

10. Beginning with Fourier (1826, p. 564) whose definition was forgotten many authors including von Mises (1919a, pp. 40 and 46) have stated independently that this is the limit of the appropriate arithmetic mean. See Sheynin (2007).

11. It was Romanovsky's considered view (1939b) that the error theory did not study systematic errors.

12. Without the theory of errors Lévy's main contribution on stable laws (1925), as he himself had seen fit to remark, would have lost its *raison d'être* (Sheynin 1995, p. 104).

Bibliography

BSE = *Bolshaya Sov. Enz*. Third edition translated into Englich as *Great Sov. Enc.*, 1973 – 1983. New York – London.

OC = *Oeuvr. Compl.* Paris.

Bernoulli J. (1713), *Ars Conjectandi*. German translation: *Wahrscheinlichkeitsrechnung*, 1899.

--- (1975), Aus den *Meditationes*. *Werke*, Bd. 3. Basel, pp. 21 – 89. Written in 1684 – 1689.

Bolshev L. N. (1975), Errors, theory of. BSE, vol. 19, pp. 44 – 45 of GSE.

Boole G. (1851), On the theory of probabilities. In author's *Studies in Logic and Prob*., vol. 1, pp. 247 – 259. London, 1952.

--- (1854), On the conditions by which the solution of questions in the theory of probabilities are limited. Ibidem, pp. 280 – 288.

Bortkiewicz L. (1904), Anwendung der Wahrscheinlichkeitsrechnung auf Statistik. *Enc. math. Wiss.*, Bd. 1. Leipzig, 822 – 851.

Chebyshev P. L. (1845), *Opyt elementarnogo analysa teorii veroiatnostei*. (Essay on an Elementary Study of the Theory of Prob.). *Poln. sobr. soch.* (Complete Works), vol. 5. M. – L., 1951, pp. 26 – 87.

--- (1936), *Teoriya veroyatnostei* (Theory of Prob.). Lectures read in 1880. Published from notes taken by A. M. Liapunov. M. – L. Defective edition with great many essential mistakes.

De Moivre A. (1756), *Doctrine of Chances*. New York, 1967. Previous editions 1718, 1738.

Eddington A. S. (1933), Notes on the method of least squares. *Proc. Phys. Soc.*, vol. 45, pp. 271 – 287.

Fourier J. B. J. (1826), Sur les résultats moyens. *Oeuvr*., t. 2, pp. 525 – 545. Paris, 1890. **S, G,** 72.

Grossmann W. (1961), *Grundzüge der Ausgleichungsrechnung*. Berlin.

Huygens C. (1657), Letter to van Schooten prefixed to his memoir *De ratiociniis* … OC, t. 14. La Haye, pp. 49 – 91. In Dutch and French. The *Letter* in French is on pp. 57 – 58.

Kemnitz Yu. V. (1948), On the density law of the errors of observation. *Geod. i Kartograph.*, No. 10, pp. 21 – 29. (R)

Kolmogorov A. N. (1948), The main problems of theoretical statistics. *Soveshchanie* (1948, pp. 216 – 220).

Laplace P. S. (1774), Sur la probabilité des causes par les événements. OC, t. 8, pp. 27 – 65.

--- (1776), Recherches sur l'intégration des equations différentielles aux différences finies. OC, t. 9, 1893, pp. 69 – 197.

--- (1781), Sur les probabilités. OC, t. 9, 1893, pp. 383 – 485.

--- (1786), Sur les approximations des formules etc. Suite. OC, t. 10, 1894, pp. 295 – 338.

--- (1812), *Théorie analytique des probabilités*. OC, t. 7, No. 2, 1886, pp. 181 – 496.

--- (1814, French), *Philosophical Essay on Probabilities*. New York, 1995. Translator and Editor A. I. Dale.

--- (1816, 1818, ca. 1819), Supplément 1, 2, 3 to the *Théor. anal.* OC, t. 7, No. 2, 1886, pp. 497 – 530; 531 – 580; 581 – 616.

--- (1827), Sur le flux et reflux lunaire atmosphèrique. OC, t.13, 1904, pp. 342 – 358.

Leibniz G. W. (1765), *Neue Abhandlungen über den menschlichen Verstand*. Hamburg, 1996.

Lévy P. (1925), *Calcul des probabilités*. Paris.

Markuse Yu. I. (1985), *Matematicheskaya obrabotka resultatov geodesicheskikh izmereniy* (Math. Treatment of Geod. Obs.). *Itogi Nauki i Teckniki. Geod. i Aeros'emka* ser., vol. 23. The whole volume. Moscow.

Mises R. von (1919a), Fundamentalsätze der Wahrscheinlichkeitsrechnung. *Math. Z.*, Bd. 4. In author's (1964b, pp. 35 – 56).

--- (1919b), Grundlagen der Wahrscheinlichkeitsrechnung. *Math. Z.*, Bd. 5. In author's (1964b, pp. 57 – 105).

--- (1928), *Wahrscheinlichkeit, Statistik und Wahrheit*. Wien – New York, 1972. English translation: New York, 1981.

--- (1964a), *Mathematical Theory of Probability and Statistics*. Edited and complemented by Hilda Geiringer. New York – London. Based on lectures (1939 – 1952) and other materials. Author's Harvard lectures: see (1964b, p. 563).

--- (1964b), *Selected Papers*, vol. 2. Providence, RI.

Montmort P. R. (1708, 1713), *Essay d'analyse sur les jeux de hazard*. New York, 1980.

Neyman J. (1950), *First Course in Probability and Statistics*. London.

Öttinger L. (1852), *Wahrscheinlichkeitsrechnung*. Berlin.

Pascal B. (1963), *Oeuvres complètes*. Paris. In one volume.

Pearson K. (1892), *Grammar of Science*. Many later editions and translations into many languages.

--- (1926), De Moivre. *Nature*, vol. 117, pp. 551 – 552.

--- (1978), *History of Statistics in the 17th and 18th Centuries* etc. Lectures of 1921 – 1933. Editor, E. S. Pearson. London – High Wycombe.

Poisson S.-D. (1812), Review of Laplace (1812). *Nouv. Bull. Sci. Soc. Philom.*, 5e ann., t. 3, pp. 160 – 163.

--- (1827), Discours prononcé aux obsèques de … Laplace. *Conn. des tems* pour 1830, Additions, pp. 19 – 22.

--- (1837), *Recherches sur la probabilité des jugements* etc. Paris, 2003. **S, G,** 53.

Poisson S.-D, Dulong P. L., Larrey F. H., Double F. J. [Report on a manuscript by J.] Civiale, Recherches de statistique … *C. r. Acad. Sci. Paris*, t. 1, pp. 167 – 177.

Prokhorov Yu. V., Sevastianov B. A. (1971), Probability, theory of. BSE, vol. 4, pp. 425 – 429 of English edition of that, third edition of the BSE.

Rainford H. F. (1957), *Survey Adjustment and Least Squares*. London.

Romanovsky V. I. (1939, 1955), Errors, theory of. BSE, 1st edition, vol. 43, pp. 725 – 727. 2nd edition, vol. 31, pp. 500 – 501. (R)

Schell E. D. (1960), S. Pepys, I. Newton and probability. *Amer. Statistician*, vol. 14, No. 4, pp. 27 – 30.

Sheynin O. B. (1976), Laplaces's work on probability. *Arch. Hist. Ex. Sci.*, vol. 16, pp. 137 – 187.

--- (1994), Chebyshev's lectures on the theory of probability. Ibidem, vol. 46, pp. 321 – 340.

--- (1995), Density curves in the theory of errors. Ibidem, vol. 49, pp. 163 – 196.

--- (2007), The true value of a measured constant and the theory of errors. *Historia Scientiarum*, vol. 17, pp. 38 – 48.

Soveshchanie (1948), *Vtoroe vsesoiuznoe soveshchanie po matematicheskoi statistike* (Second All-Union Conf. Math. Stat.). Tashkent.