ADDENDA TO THE PAPER *C.F. GAUSS AND THE METHOD OF LEAST SQUARES*

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 12(18)

Oscar Sheynin

ISSN 1644-6739

DOI: 10.15611/sps.2014.12.02

Addendum No. 1

ELEMENTARY EXPOSITION OF GAUSS' FINAL JUSTIFICATION OF LEAST SQUARES

Oscar Sheynin

Summary: Legendre was the first to publish the principle of least squares in 1805, this principle was known to Gauss since 1795. But it was Gauss who introduced the method of least squares. He offered its final justification based on the principle of maximum weight (minimal variance) in 1823 and 1828. I begin with a few words about Legendre and Laplace and continue with describing Gauss' final justification of least squares. It is extremely complicated, but modern authors removed this difficulty. My own exposition (§ 3) is quite elementary and, I think, methodically necessary.

Keywords: the principle of least squares, the method of least squares.

Legendre (1805) was the first to publish the principle of least squares (known to Gauss since 1795), but it was Gauss who introduced the method of least squares; he reasonably rejected his own first attempt (1809) and offered its final justification (1823b, 1828) based on the principle of maximum weight (minimal variance). I begin with a few words about Legendre and Laplace and continue with describing Gauss' final justification of least squares. It is extremely complicated, but modern authors removed this difficulty. My own exposition (§ 3) is quite elementary and, I think, methodically necessary.

1. Legendre and Laplace

1.1. Legendre

Here is his crucial statement (1805, pp. 72–73): *It is necessary that the extreme errors without regarding their signs be restricted between the shortest possible boundaries*.

His equations can be written as

$$a_i x + b_i y + ... + l_i = v_i, \quad i = 1, 2, ..., n. \tag{1}$$

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 12 (18)

The free terms l_i are the results of physically independent observations whose number, n, is larger than the number of the unknowns, k. The coefficients a_i, b_i, … are given by the appropriate theory, and the linearity is not restrictive since the approximate values of the unknowns can be calculated (for example, from any k equations). For equations appearing in practice no solution is possible and any set of $\hat{x}, \hat{y},\ldots$ leading to reasonable residual free terms v_i is assumed as the solution.

The optimal approach which he applied was to make the sum of the squares of the errors the least possible. This approach, as stated, was wrong: actually, Legendre minimized the sum of the squares of the residual free terms of his equations. His first statement implies that the principle of least squares is at the same time the minimax principle

$$|v_{\max}| = \min,$$

where the maximum allows for the appropriate magnitudes of all the equations, and the minimum is thought to cover any set of $\hat{x}, \hat{y},\ldots$ Actually, as it is easy to prove, the minimax principle is tantamount to making minimal the sum of v_i^{2n} with $n \to \infty$.

1.2. Laplace

He is known to have non-rigorously proved several versions of the central limit theorem and, accordingly, presumed that the observations were numerous and that their errors were normally distributed (a later term). Then, he based the adjustment of observations on minimal absolute expectation of error, which meant that calculations were only practically possible for the normal distribution. Each of the two assumptions made his method of adjustment barely useful and Gauss (1821) criticized it. Laplace did sometimes apply the mean square error (the root of the sample variance) as his criterion, but on the whole he led French mathematicians including Poisson away from Gauss; this was made easier by the priority strife between Legendre and Gauss.

2. Gauss

2.1. Prior to 1805

There is no direct proof that he applied the principle of least squares before 1805. Gerardy (1977, p. 19, Note 16) came close to achieving this, but regrettably he concentrated on elementary geodetic calculations. On the other hand, it is impossible to refute Gauss' claim of having applied it. First, Gauss made many mistakes in his computations

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 12 (18)

(Maennchen 1918/1930, p. 65ff); one example is in § 2.2.2-1 below. Second, he could have assigned differing weights to his observations; third, he (1809, § 185) allowed himself some deviation from strict procedure; fourth and last, he could have well mostly applied least squares for trial computations unknown to us.

Add to this that his contemporaries including Laplace (1812/1886, p. 353) believed Gauss and that he informed his friends and colleagues about his innovation. Among those were Bessel (1832, p. 27), Wolfgang Bolyai (Sartorius von Waltershausen 1856/1965, p. 43), the father of János Bolyai, one of the discoverers of the non-Euclidean geometry, and the astronomer Olbers.

There still exists a misunderstanding about the last-mentioned. The main point is this: in 1812 Olbers agreed to confirm Gauss in that he had indeed come to know the principle of least squares from Gauss before 1805, but he only publicly stated that in 1816. However, the *Catalogue of Scientific Literature* published by the Royal Society lists, in its proper volume, Olbers' contributions, and it is seen that during 1812–1815 he did not publish anything suitable for inserting such a statement.

2.2. The year 1823

2.2.1. General remarks

In § 2 (with an explanation of a term in § 1) Gauss restricted his investigation by excluding systematic errors from consideration. He repeated this point in § 17 and promised to present a new investigation of the case in which systematic errors are not totally excluded, but he never fulfilled his intention.

Then, in § 18 Gauss offered his definition, although not quite formal, of independent functions of observations: they should not have contained common observations. In § 19 he specified that those functions were linear; otherwise his statement would have contradicted the Student–Fisher theorem on the independence of the sample variance and the arithmetic mean.

Gauss (§ 6) introduced his measure of precision (the variance, as it is now called). In his letter to Bessel of 1839, he (W-8, pp. 146–147) stressed that an integral measure of precision was preferable to a local measure. In the same § 6 he indicated that the quadratic function was the simplest [from integral measures] and in his preliminary report he (1821/1887, p. 192) noted that his choice was connected with other advantages but did not elaborate. I leave it at that. At the end of § 17 Gauss somewhat elliptically explained that minimal variance was his criterion for adjusting observations.

The main body of Gauss (1823b) is extremely difficult to read, which had undoubtedly been one of the reasons for numerous textbook authors to discuss Gauss' first substantiation of least squares (1809) rather than the second one. Those wishing to acquaint themselves with that main body without leaving aside his deliberations can consult Helmert (1872). Modern exposition is provided, for example, by Kolmogorov (1946) and Hald (1998, pp. 471–475).

2.2.2. The sample variance

Then, in § 38, Gauss derived his celebrated formula for the sample variance, as it is now called:

$$\sigma = \sqrt{\frac{[vv]}{n-k}}, \tag{2}$$

where, in Gauss' notation, $[vv]$ is the sum of the squares of v_i. More precisely, Gauss calculated the expectation of σ and had to assume that σ itself was equal to it. The reader can find the derivation in many sources, for example Helmert (1872/1924, pp. 102–104) and Kolmogorov (1946).

2.2.2-1. The precision of the sample variance

Gauss (§§ 39, 40) derived the variance of σ^2. His direct approach was somewhat laborious but easy to follow and his final formula provided the boundaries for σ^2. Additionally, he remarked that for the normal distribution

$$\operatorname{var}\sigma^2 = \frac{2\sigma^4}{n-k}. \tag{3}$$

One of the boundaries was wrong; Helmert (1904) corrected that mistake and Kolmogorov et al. (1947) independently derived the same formula as Helmert did:

$$\frac{v_4 - s^4}{n-k} < \operatorname{var}\sigma^2 < \frac{v_4 - s^4}{n-k} + \frac{k}{n}\cdot\frac{3s^4 - v^4}{n-k},$$

for $v^4 - 3s^4 < 0$ with a similar formula for the alternative. Here, $s^2 = E\sigma^2$. In a companion paper, Maltzev (1947) proved that both inequalities can be understood as being conditional.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 12 (18)

2.2.2-2. Unbiasedness

At least in geodesy, the estimator of precision is σ rather than σ^2 and, unlike σ^2, it is biased. Anyway, how important is unbiasedness? It seems that bias is now somewhat tolerated (Sprott 1978, p. 194) and in any case unbiased estimates sometimes just do not exist.

An additional consideration is interesting. Czuber (1891, p. 460) discussed the problem of bias with Helmert, and they concluded that the main point was not bias itself, but the relative value of $\mathrm{var}\sigma^2/\sigma^2$.

Eddington (1933, p. 280) independently stated the same.

For a biased estimate of the sample variance, i.e., for $k = 0$ instead of $k = 1$, Cramér (1946, § 27.4) derived the formula

$$\mathrm{var}\,\sigma^2 = \frac{\mu_4 - \mu_2^2}{n} - \frac{2(\mu_4 - 2\mu_2^2)}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}$$

in terms of the central moments μ_2 and μ_4. In case of normality he (Ibidem) additionally offered the formula

$$\mathrm{var}\,\sigma^2 = \frac{2(n-1)}{n^2}\sigma^4.$$

2.2.2-3. Application of the formula

As noted in § 2.2.1, Gauss did not consider systematic errors. In particular, this meant that formula (2) was practically inadequate, and Gauss understood it perfectly well. When performing geodetic work, he measured each angle as many times as he felt necessary, see W-9, pp. 278–281 or Schreiber (1879, p. 141). In at least three letters Gauss recommended, when the number of observations was not large, to derive a single value of σ^2 for several stations. These letters were: in 1821, to Bessel, see Gauss (1880/1975, p. 382); and in 1844 and 1847, to Gerling (1927/1975, pp. 687 and 744). At least once Laplace acted the same way even earlier, see Supplement No. 3 of ca. 1819 to his treatise (1812/1886) and another author (Ku 1967/1969, p. 309) expressed the same opinion.

In spite of the above, geodesists have been applying formula (2), although only after completing work on a chain of triangulation. It is then possible to allow for the closures of the triangles, for the discrepancies between the baselines situated at the ends of the chain, and between the astronomically fixed end lines of the chain. In other words, applying that formula only after having revealed the influence of systematic errors as much as it was possible. Supplemented with baselines and astronomical observations, a chain is to the most possible extent independent (in Gauss' sense, see § 2.2.1) of the neighbouring chains.

2.2.2-4. Criticism

Bertrand translated Gauss' contributions on the theory of errors and least squares into French (Gauss 1855). Note that Gauss, at least by the end of his life, agreed to have some of his work appearing in French; previously, owing to political reasons, he refused to publish anything in that language. Gauss died the same year, 1855, and Bertrand (1855) made known that he, Gauss, had no time for really studying the prepared translation.

Many years later Bertrand (1888) criticized the Gauss formula (2). Tacitly assuming the normal distribution, he provided an example in which his own estimate of σ^2 was less than that provided by Gauss. He forgot, however, that formula (2) provided an unbiased estimate whereas his own estimate was biased. Then, he calculated σ^2 forgetting formula (3). It was this episode that led Czuber to the discussion described in § 2.2.2-2.

Later events seem to indicate that the Gaussian theory of errors remained for a long time almost forgotten. Chebyshev (1880/1936, p. 249) stated that *recently, some authors had begun to apply* formula (2). More generally, at least up to the middle of the 20^{th} century statisticians of the ordinary rank did not know Gauss' second justification of least squares (Campbell 1928; Eisenhart 1964, p. 24).

In other countries the situation had been likely about the same.

Indeed, Fisher (1925/1990, p. 260) thought that the method of least squares was a *special application of the method of maximal likelihood* which was only correct for the first justification of the method. And Poincaré (1896/1912, p. 188) stated that Gauss' rejection of his own first justification of the method was *assez étrange*.

In Russia, however, the situation was somewhat different since Markov, citing Gauss, resolutely upheld the second substantiation. At the same time he stated that the method did not possess any optimal properties and thus contradicted himself: such methods do not require any substantiation. See Sheynin (2006, pp. 80–81).

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 12 (18)

3. Conclusion: an alternative justification of the method of least squares

After proving formula (2), Kolmogorov (1946) remarked in passing that it was only a definition of σ. Yes, if the number of the degrees of freedom is correctly allowed for. As I understand it, the formula seems plausible, but the proof is still required; after that, it can be interpreted as that definition.

Many authors beginning with Gauss had provided the proof which is not difficult. The necessary restrictions are: linearity of the equations (1), independence of their free terms (the results of observation), and the unbiasedness of the estimators $\hat{x}, \hat{y}, ...$ The main point, however, is that the proof does not depend on the condition of least squares. On the contrary, this condition can now be introduced at once since it means minimum variance.

The formulas derived by Gauss for constructing and solving the normal equations and calculation of the weights of $\hat{x}, \hat{y}, ...$ and of their linear functions will still be useful. Gauss had actually provided two justifications (of which I only left the second one), but why did not he even hint at this fact? I can only quote Kronecker (1901, p. 42) and Stewart (Gauss 1823b, 1828/1995, p. 235):

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

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

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

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

Acknowledgement. A summary of this paper appeared in *Math. Scientist*, vol. 37, 2012, pp. 147–148.

References

C.F. Gauss

– 1809, in Latin. German title, *Theorie der Bewegung* [...]. German translation of relevant place: Gauss (1887, pp. 92–117). English translation: *Theory of Motion*... Boston, 1857; Mineola, 2004.

– 1821, Preliminary author's report on Gauss (1823b, pt. 1). Ibidem, pp. 190–195.

– 1823a, Preliminary author's report on Gauss (1823b, pt. 2). Ibidem, pp. 195–199.

– 1823b, in Latin. German title: *Theorie der den kleinsten Fehlern unterworfenen Combination der Beobachtungen*, pts 1–2. German translation: Ibidem, pp. 1–53.

– 1828, in Latin. Supplement to Gauss (1823b). German translation: Ibidem, pp. 54–91. English translation of both parts and Supplement by G.W. Stewart: *Theory of combination of observations*... Philadelphia, 1995.

– 1855, *Méthode des moindres carrés.* Paris.

– 1870–1929, *Werke*, Bde 1–12. Göttingen. Hildesheim, 1973–1981. Separate volumes denoted W-*i*.

– 1880–1927, Correspondence with Bessel (1880), Olbers (1900–1909) and Gerling (1927). Reprinted, respectively, in *Werke, Ergänzungsreihe*, Bde 1, 4(1), 3; Hildesheim, 1975, 1976, 1975.

– 1887, *Abhandlungen zur Methode der kleinsten Quadrate*. Hrsg. A. Börsch, P. Simon. Latest edition: Vaduz, 1998.

Other authors

Bertrand J. (1855), Sur la méthode des moindres cárres. *C. r. Acad. Sci. Paris*, t. 40, pp. 1190–1192.

– (1888), Sur les conséquences de l'égalité acceptée entre la valeur vraie d'un polynôme et sa valeur moyenne. *C. r. Acad. Sci. Paris*, t. 106, pp. 1259–1263.

Bessel F.W. (read 1832), Über den gegenwärtigen Standpunkt der Astronomie. *Populäre Vorlesungen*. Hamburg, 1848, pp. 1–33.

Campbell N.R. (1928), *Account of the Principles of Measurement and Calculations*. London.

Chebyshev P.L. (Lectures 1879/1880, 1936, in Russian), *Teoria Veroiatnostei* (Theory of probability). Moscow–Leningrad.

Cramér H. (1946), *Mathematical Methods of Statistics*. Princeton.

Czuber E. (1891), Zur Kritik einer Gauss'schen Formel. *Monatsh. Math. Phys*., Bd. 2, pp. 459–464.

Eddington A.S. (1933), Notes on the method of least squares. *Proc. Phys. Soc.*, Vol. 45, pp. 271–287.

Eisenhart C. (1946), The meaning of *least* in least squares. *J. Wash. Acad. Sci.*, Vol. 54, pp. 24–33. Also in Ku (1969, pp. 265 –274).

Fisher R.A. (1925), *Statistical Methods for Research Workers*. In author's book (1990), separate paging.

– (1990), *Statistical Methods, Experimental Design and Statistical Inference*. Oxford.

Gerardy T. (1977), Die Anfänge von Gauss' geodätische Tätigkeit. *Z. f. Vermessungswesen*, Bd. 102, pp. 1–20.

Hald A. (1998), *History of Mathematical Statistics from 1750 to 1930*. New York.

Helmert F.R. (1872), *Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate*. Leipzig. Later editions: 1907, 1924.

– (1904), Zur Ableitung der Formel von Gauss für den mittleren Beobachtungsfehler und ihrer Genauigkeit. *Sitz. Ber. Kgl. Preuss. Akad. Wiss. Berlin*, Hlbbd. 1, pp. 950–964. Reprint: *Akademie-Verträge*. Frankfurt/Main, 1993, pp. 189–208. Shorter version: *Z. f. Vermessungswesen*, Bd. 33, 1904, pp. 577–587.

Kapteyn J.C. (1912), Definition of the correlation coefficient. *Monthly Notices Roy. Astron. Soc.*, Vol. 72, pp. 518–525.

Kolmogorov A.N. (1946, in Russian), Justification of the method of least squares. *Sel. Works*, Vol. 2. Dordrecht, 1992, pp. 285–302.

Kolmogorov A.N., Petrov A.A., Smirnov Yu.M. (1947 in Russian), A formula of Gauss in the method of least squares. Ibidem, pp. 303–308.

Kronecker L. (1901), *Vorlesungen über Zahlentheorie*, Bd. 1. Leipzig.

Ku H.H. (1967), Statistical concepts in metrology. In Ku (1969, pp. 296–310).

– Editor (1969), *Precision Measurement and Calibration. Sel. Nat. Bureau Standards Stat. Concepts and Procedures*. NBS Sp. Publ. No. 300, Vol. 1. Washington.

Laplace P.-S. (1812), *Théorie analytique des probabilités*. *Oeuvr. Compl.*, t. 7. Paris, 1886.

Legendre A.M. (1805), *Nouvelles méthodes pour la détermination des orbites des comètes*. Paris.

Maennchen Ph. (1918/1930), *Gauss als Zahlenrechner*. In Gauss W-10, Tl. 2; Abt. 6. Separate paging.

Maltzev A.I. (1947 in Russian), Remark on Kolmogorov et al. (1947). *Izvestia Akademii Nauk*, ser. math., Vol. 11, pp. 567–578.

Olbers W. (1816), Über den veränderlichen Stern im Halse des Schwans. *Z. f. Astron. u. verw. Wiss.*, Bd. 2, pp. 181–198.

Poincaré H. (1896), *Calcul des probabilités*. Paris, 1912.

Sartorius von Waltershausen W. (1856), *Gauss zum Gedächtnis*. Wiesbaden, 1965.

Schreiber O. (1879), Richtungsbeobachtungen und Winkelbeobachtungen. *Z. f. Vermessungswesen*, Bd. 8, pp. 97–149.

Sheynin O. (2006), Markov's work on the treatment of observations. *Hist. Scientiarum*, Vol. 16, pp. 80–95.

Sprott D.A. (1978), Gauss' contribution to statistics. *Hist. Math.*, Vol. 5, pp. 183–203.

Stewart G.W. (1995), Translation of Gauss (1823b) with Afterword (pp. 207–241). Philadelphia.

ELEMENTARNE PRZEDSTAWIENIE OSTATECZNEGO GAUSSOWSKIEGO UZASADNIENIA NAJMNIEJSZYCH KWADRATÓW

Streszczenie: Legendre był pierwszym, który opublikował w 1805 r. zasadę najmniejszych kwadratów, znaną Gaussowi od 1795 r. Ale to Gauss wprowadził metodę najmniejszych kwadratów. Pełne jej uzasadnienie, oparte na zasadzie maksymalnych wag (minimalna wariancja) Gauss podał w pracach opublikowanych w latach 1823 i 1828. W niniejszej pracy, po prezentacji prac Legendre'a i Laplace'a, podaję elementarne wyjaśnienie tej metody.

Słowa kluczowe: zasada najmniejszych kwadratów, metoda najmniejszych kwadratów.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 12 (18)

Addendum No. 2

ANTISTIGLER

Oscar Sheynin

Summary: Stigler is the author of two books (1986; 1999) in which he dared to profane the memory of Gauss. Stigler is considered as the best historian of statistics of the 20th century. This paper contains my critical remarks on his works.

Keywords: the history of statistics, Gauss, Stigler.

Stigler is the author of two books (1986; 1999) in which he dared to profane the memory of Gauss.

I had vainly criticized the first one (1993; 1999a; 1999b), but not a single person publicly supported me, whereas several statisticians, only justifying themselves by arguments *ad hominem*, urgently asked me to drop that subject. The appearance of Stigler's second book showed that they were completely wrong but the same general attitude is persisting. One of those statisticians, apparently believing that a living dog was more valuable than a dead lion, was the President of the International Statistical Institute (2008). But to go into detail.

1) A few years ago Stigler was elected President of that same Institute (and had served in that capacity). He is now member of the Institute's committee on history to which I was also elected (chosen?) without my previous knowledge or consent. I refused to work together with him (and with Descrosières, – of all members of the Institute, see below!).

2) A periodical (*Intern. Z. f. Geschichte u. Ethik* (!) *der Naturwissenschaften, Technik u. Medizin*, NTM) refused to consider my proposed subject, – the refutation of Stigler. The Editor politely suggested I should apply to a statistical periodical.

3) The Gauss-Gesellschaft-Göttingen is silent and had not even answered my letter urging them to support me.

4) Healy (1995, p. 284) indirectly called Stigler the best historian of statistics of the 20th century, and Hald – yes, Hald (1998, p. xvi) even called Stigler's book (1986) *epochal*. Epochal, in spite of slandering Gauss, of humiliating Euler (below), and of its being an essay rather than THE HISTORY (!) of statistics, as Stigler had the cheek to name it.

So much is absent in THE HISTORY, – cf. my book Sheynin (2005/2009), – in spite of which it became the statisticians' Bible, that I shall extrapolate this phenomenon by reducing it with Lewis Carroll's help *ad absurdum*:

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 12 (18)

Other maps are such shapes, with their islands and capes:
But we've got our brave Captain to thank
(So the crew would protest) "That he's bought us the best –
A perfect and absolute blank!"

Stigler is regarded as a demigod. *Historia Mathematica* had published a review of his book (1999). Instead of providing its balanced account, the reviewer (an able statistician; H. M. vol. 33, no. 2, 2006) went out of his way to praise, *to worship* both the book and Stigler himself.

5) *Centaurus* rejected the manuscript of my paper (1999a) initially submitted to them since the anonymous reviewer, contrary to facts and common sense, did his damnedest to exonerate Stigler.In addition to my papers mentioned above, I can now add two more publications (2005; 2006, see their Indices), but I ought to add several points here.

1. Stigler (1986, p. 145): *Gauss solicited reluctant testimony from friends that he had told them of the method* [of least squares, MLSq] *before* [the appearance of the Legendre memoir in] *1805*.

And in 1999, p. 322, repeating his earlier (of 1981) statement of the same ilk: *Olbers did support Gauss's claim … but only after seven years of repeated prodding by Gauss*. Grasping at straws, Stigler adds an irrelevant reference to Plackett (1972).

So what happened with Olbers? On 4.10.1809 Gauss had asked him whether he remembered that he had heard about the MLSq from him (from Gauss) in 1803 and again in 1804. Olbers apparently did not answer (or answered through a third party). On 24.1.1812 Gauss asked even more: Was Olbers prepared to confirm publicly that fact? And Olbers answered on 10.3.1812: *gern und willig* (with pleasure), and at the first opportunity. However, during 1812–1815 Olbers had only published a few notes on the observation of comets (*Catalogue of Scientific Literature*, Roy. Soc. London), and he therefore only fulfilled Gauss' request in 1816. Much later Gauss, who became sick and tired of the whole dispute, in a letter of 3.12.1831 to Schumacher mentioned that his friend had acted in good faith, but that he was nevertheless displeased by Olbers' testimony made public.

2. Again in 1999, Stigler had deliberately omitted to mention Bessel's statement on the same subject. I discovered it while being prompted by Stigler's attitude and quoted Bessel in a paper (1993) which Stigler mentioned in 1999. Bessel's testimony, all by itself, refutes Stigler's accusation described above.

3. Stigler (1999, pp. 322–323) mentions von Zach, his periodical (*Monatl. Corr.*) and some material published there in 1806–1807 which allegedly (indirectly) proved that von Zach had not considered Gauss as the inventor of the MLSq. Stigler leaves out a review published in the same periodical in 1809 whose anonymous author (von Zach?) described the actual history of the discovery of the MLSq, see p. 191. Incidentally, I (1999a, p. 258) found von Zach's later statement in which he repeated Gauss' explanation to the effect that he, Gauss, discovered the MLSq in 1795.

4. Stigler (1986, p. 57): *It is clear* […] *that Legendre immediately realized the method's potential*. And, on p. 146: *There is no indication that [Gauss] saw its great general potential before he learned of Legendre's work*. Stigler thus denies Gauss' well-known statement that he had been applying the MLSq since 1794 or 1795, denies simply because he is inclined to dethrone Gauss and replace him by Legendre.

5. Stigler (1986, p. 143): Only Laplace saved Gauss' first justification (in 1809) of the MLSq from joining "an accumulated pile of essentially ad hoc constructions". And how about Legendre? Stigler (1986, p. 13): *For stark clarity of exposition the presentation* [by Legendre in 1805] *is unsurpassed; it must be counted as one of the clearest and most elegant introductions of a new statistical method in the history of statistics*. His work (Stigler, p. 57) revealed his "depth of understanding of his method". All this in spite of two mistakes made by Legendre and lack of any demonstration of the method. Legendre alleged that the MLSq agreed with the minimax principle, and he mentioned errors instead of residual free terms of the initial equations. And can we believe that Stigler did not know that the Gauss' proof of 1809, which allegedly almost joined "the accumulating pile" of rubbish, had been repeated in *hundreds* of books on the treatment of observations? Was it only due to Laplace?

6. Stigler (p. 146): *Although Gauss may well have been telling the truth about his prior use of the method, he was unsuccessful in whatever attempts he made to communicate it before 1805*. The first part of the phrase was appropriate in respect to a suspected rapist, but not to Gauss. As to his "attempts", Gauss had communicated his discovery to several friends and colleagues but did not proclaim it through a public crier or by a publication in a newspaper.

Other pertinent points.

7. Stigler (1986, p. 27) denounced Euler as a mathematician who did not understand statistics. After I (1993) had refuted that pernicious statement, Stigler (1999, p. 318) declared that, in another case, Euler *was acting in the grand tradition of mathematical statistics*. He did not,

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 12 (18)

however, renounce his previous opinion. More: in that second case, Euler had rejected the method of maximum likelihood, because, as he put it, the result should not change whether an outlying observation be rejected or not (read: the treatment should be such that …). Euler suggested to keep to the known and reliable method, to the mean; he had not mentioned the median although it (but not the term itself) had actually been earlier introduced by Boscovich.

8. Descrosières (1998, transl. from French) believes that Poisson had introduced the strong law of large numbers and that Gauss had derived the normal distribution as a limit of the binomial law, see my review in *Isis*, vol. 92, 2001, pp. 184–185. And Stigler (1999, p. 52)? He called Descrosières *a scholar of the first rank*!

9. There also, Stigler named another such high ranking scholar, Porter, and he (p. 3) also called Porter's book of 1986 *excellent*. I reviewed it (*Centaurus*, vol. 31, 1988, pp. 171, 172) and declared an opposite opinion. In 2004 Porter published Pearson's biography, see my review in *Hist. Scientiarum*, vol. 16, 2006, pp. 206–209. I found there such pearls of wisdom as (p. 37) *Even mathematics has aspects that cannot be proven, such as the fourth dimension*. In my opinion, that book is barely useful.

10. In 1983, issuing from a biased stochastic supposition, Stigler declared that another author rather than Bayes had actually written the Bayes memoir. In 1999, while reprinting his 1983 paper, in spite of his sensational finding being stillborn and forgotten, Stigler got rid of its criticisms in a tiny footnote (p. 391).

11. Stigler (1986) is loath to mention his predecessors. On pp. 89–90 he described the De Moivre–Simpson debate forgetting to refer to me (1973a, p. 279). And on pp. 217–218 he discussed the once topical but then completely forgotten conclusion concerning statistics of population without citing his only possible source of information, Chuprov's letter to Markov of 10.3.1916 (Ondar 1977/1981, No. 72, pp. 84–85).

Long before that Stigler (1977) dwelt on Legendre's accusation of Gauss concerning number theory without naming me (1973b, p. 124, note 83).

So why does Stigler remain so popular? Because the statistical community is crassly ignorant of the history of its own discipline; because it pays absolutely no attention to the slandering of Gauss' memory (even if realizing that fact, as the reviewer for *Hist. Math*. did, see above, – I personally informed him about it in 1991, but he had known it himself); because it possesses a narrow scientific

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 12 (18)

Weltanschauung; and because the tribe of reviewers do not feel any social responsibility for their output. And of course there is a special reason: Stigler published his book (1986) when there was hardly anything pertinent except for papers in periodicals. The same happened to a lesser extent with Maistrov's book of 1974 which is still remembered!

To end my pamphlet, I quote, first, the most eminent scholar and historian of science, the late Clifford Truesdell (1984, p. 292), whom I will never forget and whose alarm bell apparently fell on deaf ears, and, second, Einstein's letter of 1933 to Gumbel, a German and later an American statistician (Einstein Archives, Hebrew Univ. of Jerusalem, 38615, in translation):

1) *No longer is learning the objective of scholarship.* [...] *By definition, now, there is no learning, because truth is dismissed as an old-fashioned superstition*.

2) *Integrity is just as important as scientific merits*.

References

Descrosières A. (1998), *The Politic of Large Numbers.* Harvard Univ. Press, Cambridge (Mass.) – London.

Hald A. (1998), *History of Mathematical Statistics etc*. Wiley, New York.

Healy M.G.R. (1995), Yates, 1902–1994, *Intern. Stat. Rev*., Vol. 63, pp. 271–288.

Ondar Kh.O., Editor (1977), *Correspondence between A.A. Markov and A.A. Chuprov etc*. Springer, New York, 1981 [in Russian].

Plackett R.L. (1972), The discovery of the method of least squares. *Biometrika*, Vol. 59, pp. 239–251. Reprinted: Kendall M.G., Plackett R.L., Editors (1977), *Studies in the History of Statistics and Probability*, Vol. 2. Griffin, London, pp. 279–291.

Porter T.M. (2004), *Karl Pearson*. Princeton Univ. Press, Princeton–Oxford.

Sheynin O.B. (1973a), Finite random sums. *Arch. Hist. Ex. Sci*., Vol. 9, pp. 275–305.

– (1973b), Mathematical treatment of astronomical observations. Ibidem, Vol. 11, pp. 97–126.

– (1993), On the history of the principle of least squares. Ibidem, Vol. 46, pp. 39–54.

– (1999a), Discovery of the principle of least squares. *Hist. Scientiarum*, Vol. 8, pp. 249–264.

– (1999b), Gauss and the method of least squares. *Jahrbücher f. Nationalökonomie u. Statistik*, Bd. 219, pp. 458–467.

– (2005), *Theory of Probability. Historical Essay*. NG Verlag, Berlin, 2009. Also www.sheynin.de.

– (2006), *Theory of Probability and Statistics As Exemplified in Short Dictums*. NG Verlag, Berlin. Also www.sheynin.de.

Stigler S. M. (1977), An attack on Gauss published by Legendre in 1820. *Hist. Math*., Vol. 4, pp. 31–35.

– (1986), *The History of Statistics*. Harvard Univ. Press, Cambridge, Mass.

– (1999), *Statistics on the Table.* Harvard Univ. Press, Cambridge, Mass.

Truesdell C. (1984), *An Idiot's Fugitive Essays on Science*. Springer, New York.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 12 (18)

ANTISTIGLER

Streszczenie: Stigler jest autorem dwóch książek o historii statystyki (pierwsza została opublikowana w 1986 r., druga w 1999 r.), w których dopuszcza się jednak nieścisłości, profanując nawet pamięć Gaussa. Stigler jest uznawany za najlepszego historyka statystki XX w. W pracy niniejszej omówione są niedostatki jego prac.

Słowa kluczowe: historia statystyki, Gauss, Stigler.

Addendum No. 3

THEORY OF ERRORS AND STATISTICS. SOME THOUGHTS ABOUT GAUSS

Oscar Sheynin

Summary: It is explained that the theory of errors is the application of the statistical method to the entire process of measuring physical magnitudes. In particular, the present--day experimental design and exploratory data analysis can be considered as a branch of the error theory.

Keywords: the theory of errors, experimental design, exploratory data analysis

It is my understanding that the theory of errors is the application of the statistical method to the entire process of measuring physical magnitudes. In particular, the present-day experimental design and exploratory data analysis had a precursor, the forgotten determinate branch of the error theory. It aimed at uncovering structures in the data (especially systematic errors) and at establishing the best circumstances of measurements. A simplest pertinent problem: How does the form of triangle ABC with fixed points A and B influence the precision of determining point C if angles A and B are measured with a known error?

The main obstacle to applying statistical methods to the treatment of observations is caused by the presence of unavoidable systematic errors so that the notion of random variable becomes almost meaningless. In the main paper, I mentioned the ensuing difficulties of determining the necessary number of observations, of estimating actual precision, and of dealing with outliers.

At the same time, it is proper to indicate that statistics had borrowed two important principles from the theory of errors: maximum likelihood (introduced by Lambert, see § 6.3.1) and minimal variance (Gauss). Moreover, I quoted Eisenhart who had noted that the work of Gauss was essential to statistical theory at large.

Now, why did Gauss enter (and develop the only practically important) theory of errors? The main reason was (Subbotin 1956, p. 246) that

just like Newton, Gauss was not only a mathematician, but not less a natural scientist who needed to feel directly nature and physical reality, to feel astronomy and geodesy (and, later, terrestrial magnetism). The theory of errors quite naturally accompanied the two first named sciences.

In astronomy, Gauss (*Theory of Motion*) developed his methods of orbit calculations and was able to relocate the first discovered (but then lost) minor planet. In geodesy, he personally participated in triangulating a vast region of Germany and in a meridian arc measurement. As a by-product, he invented the heliotrope for reflecting the sunlight from a triangulation station and thus avoiding night observations on lamps. On his geodetic work see Gauss (1958).

And it was extremely important that geodesy led him to his fundamental work in differential geometry and conformal mapping.

Geodesists in Russia and (in spite of all the horrible campaigns against all foreign) the Soviet Union invariably considered Gauss as a demigod.

References

Gauss C.F. (1958), *Izbrannye Geodezicheskie Sochinenia*, vol. 2. *Vysshaia Geodesia*. (Sel. Geod. Writings, vol. 2. Higher Geodesy). Moscow.

Subbotin M.F. (1956), *Astronomicheskije i gieodezicheskije raboty Gaussa*, Moscow.

TEORIA BŁĘDÓW I STATYSTYKA. PEWNE PRZEMYŚLENIA GAUSSOWSKIE

Streszczenie: Wyjaśniono, że teoria błędów stanowi zastosowanie metod statystycznych do całościowego procesu pomiaru wiekości fizycznych. Dzisiejsze dziedziny planowania doświadczeń czy eksploracyjna analiza danych mogą być traktowane jako odrębne gałęzi teorii błędów.

Słowa kluczowe: teoria błędów, planowanie doświadczeń, eksploracyjna analiza danych.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 12 (18)

GAUSS THEOREM ON CONTINUED FRACTIONS –
remarks by Witold Więsław

Gauss theorem stating that the limit of the geometric means of partial quotients of a real number is an absolute constant (for almost all real numbers) was for the first time proved by P.O. Kuzmin in rather complicated way (On a Gauss problem (in Russian), *Doklady AN*, ser. (A), 1928, 375–380). The next proof was given by P. Lévy (Sur les lois de probabilité dont dépendent les quotients complets et incomplets d'une fraction continue, *Bull. Soc. Math*. 57 (1929), 178–194). Another proof can be found in A.Ya. Chinčin's book *Continued Fractions* (in Russian), for example in its fourth edition in the year 1978. C. Ryll-Nardzewski (On the ergodic theorems. II. Ergodic theory of continued fractions, *Studia Mathematica* 12 (1951), 74–79) gave a proof based on the ergodic theory. Following ideas of E. Marczewski (Szpilrajn) he constructed an invariant measure on the reals and then applied the ergodic theorem to continued fractions.