HISTORIA MATHEMATICA **23** (1996), 54–67
ARTICLE NO. 0004

Mendeleev and the Mathematical Treatment of Observations in Natural Science

OSCAR SHEYNIN

D. I. Mendeleev (1834–1907), the eminent chemist, rejected doubtful experiments and spoke out against amassing observations. He gave thought to eliminating systematic errors and offered a simple test of the "harmony" of observations. Modern statistics has recognized harmony as symmetry of the appropriate density function and has independently quantified asymmetry in accordance with Mendeleev's idea. Mendeleev made mistakes in estimating the plausibility of his data, and he hardly knew Gauss's second formulation of the method of least squares. An analysis of his work sheds light on the level of statistical knowledge in the natural sciences beyond astronomy and geodesy in the late 19th century. © 1996 Academic Press, Inc.

Der berühmte Chemiker D. I. Mendeleev (1834–1907) warf zweifelhafte Beobachtungen weg und trat gegen die Anhäufung von Beobachtungen auf. Er hat nach Eliminierung systematischer Fehler gestrebt und ein einfaches Kriterium für die "Harmonie" der Beobachtungen vorgeschlagen. Die moderne Statistik hat Harmonie als Symmetrie der entsprechenden Dichtefunktion anerkannt und hat selbst ein quantitatives Maß der Asymmetrie, das Mendeleev's Idee entspricht, eingeführt. Mendeleev irrte sich bei der Abschätzung der Sicherheit seiner Beobachtungen und war kaum mit der zweiten Gauss'schen Begründung der Methode der kleinsten Quadrate bekannt. Eine Analyse seines Werkes bringt Licht in den Stand der statistischen Kenntnis der Naturwissenschaft des späten 19. Jahrhundert außerhalb der Astronomie und Geodäsie. © 1996 Academic Press, Inc.

Известный химик Д.И.Менделеев (1834 — 1907) отбрасывал сомнительные эксперименты и высказывался против чрезмерного накопления наблюдений. Он стремился исключать систематические ошибки и предложил простой критерий "стройности" наблюдений. Современная статистика признала стройность как симметрию соответствующей функции плотности и независимо ввела количественную меру асимметрии, отвечающую идее Менделеева. Менделеев ошибался при оценке надежности своих наблюдений и вряд ли был знаком со вторым гауссовским обоснованием метода наименьших квадратов. Исследование его работ проливает свет на уровень статистических знаний в естествознании второй половины 19=го века за пределами астрономии и геодезии. © 1996 Academic Press, Inc.

AMS 1991 subject classifications: 01A55, 62-07.
KEY WORDS: Gauss, Mendeleev, symmetric densities, treatment of observations in natural science.

0315-0860/96 $18.00

1. INTRODUCTION

The noted chemist, Dmitrii Ivanovich Mendeleev (1834–1907), was also a prominent metrologist and from 1893 to 1907 the director of Russia's Main Board of Weights and Measures. In this paper, I describe his thoughts on the planning, selection, and treatment of observations, i.e., on the problems which inevitably presented themselves in connection with both chemistry and metrology.[1]

For the benefit of those not conversant with it, I set the stage in Section 2 by examining the Gaussian theory of errors. Section 3 sketches the general situation in the natural sciences of the late 19th century relative to how physicists and chemists contemporaneous with Mendeleev treated their observations. In Sections 4–7, I analyze some of the problems pertaining to the stochastic branch of error theory which Mendeleev had to solve, problems which today would belong to the appropriate chapter of mathematical statistics. Section 8 concludes.

Mendeleev also considered the determinate branch of the theory of errors and, more precisely, the planning of measurements and the preliminary study of their results. I discuss his ideas on this topic in Sections 4–5 and note that it now falls within the theory of experimental design and exploratory data analysis (both of which also involve stochastic considerations).

Mendeleev was an exceptionally versatile scholar, who studied demography and industrial statistics in addition to chemistry. These varied interests gave him an interesting perspective on statistics as well as on its role in meteorology. On statistics in general, he viewed the subject's possibilities as essentially limitless, writing that "[s]ome day they [the poets] will clothe the coarse statistical prose in verses because the numbers reveal force, power, men's weaknesses, the paths of history and many other ... aspects of the world ..." [25, 54]. Yet, relative to meteorology, he felt that the prevailing school tended to amass observations, seemingly needing nothing but "numbers and numbers" [20, 267]. As a result, its practitioners got no further than plotting lines of equal values of meteorological elements. Mendeleev, however, envisioned a new meteorology (also based on statistical data and using contour lines) which "little by little" would "master, synthesize, forecast" [23, 527]. Clearly, Mendeleev did not perceive that the introduction of contour lines into meteorology would be an early masterpiece of exploratory data analysis,[2] but he nevertheless had keen insight into the overall direction meteorology should take.

2. THE THEORY OF ERRORS: THE MAIN PROBLEM

The main problem which Mendeleev had to solve involved adjusting direct observations, i.e., choosing some final value (or estimator in modern terms) for the unknown constant x given observations $x_1, x_2, \ldots, x_n$ and estimating the plausibility of this final value. In a more general setting, the problem was to adjust indirect

[1] I may also add physics. Indeed, some of Mendeleev's works, e.g., [15], at least bordered on this science.

[2] This was achieved in 1817 by Alexander Humboldt, who thus separated climatology from meteorology. In general, he conditioned the study of natural phenomena by discovering pertinent mean values (or states) [31, Sect. 4.1].

observations, i.e., to deduce some final values for the unknown constants x, y, z, ... from a redundant system of equations

$$a_i x + b_i y + c_i z + \cdots + l_i = 0, \quad i = 1, 2, \ldots, n, \tag{1}$$

with known coefficients a_i, b_i, c_i, ... and measured free terms l_i, and to estimate the plausibility of these values. The linearity of Eq. (1) was not restrictive since the approximate values of $x, y, z, \ldots$ were always known so that only small corrections were actually needed. For "physically" independent free terms l_i, Eq. (1) became inconsistent, and any set $(\hat{x}, \hat{y}, \hat{z}, \ldots)$ leading to reasonably small values of residual free terms v_i had to be admitted as a solution. The method of least squares, made known by Adrien-Marie Legendre in 1805 [13] but used by Carl Friedrich Gauss as early as 1794 or 1795, was no exception. It demanded that

$$\sum v_i^2 \equiv [vv] = \min \tag{2}$$

from among all possible sets $(\hat{x}, \hat{y}, \hat{z}, \ldots)$. In what follows, I use the apt notation $[vv]$, as well as the general symbol

$$[ab] = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n,$$

both of which are due to Gauss.

Legendre had come to the method of least squares via qualitative considerations, whereas Gauss had offered a mathematical substantiation in 1809 and published his alternative and definitive thoughts on the subject in 1823. On the other hand, Pierre Simon de Laplace developed the method for the case of a large number of observations by nonrigorously proving quite a few versions of what is now called the central limit theorem; the first of his several pertinent contributions appeared in 1810 [12]. Although extremely important from a theoretical point of view, Laplace's work had no direct bearing on the practitioner. Since Gauss's ideas had the greatest immediate impact, I restrict my attention to his findings.

In 1809, Gauss assumed that the errors of observation possess a unimodal (with a single peak) and symmetric density function $\varphi(x)$, and that the arithmetic mean $\bar{x}$ of n observations

$$x_1, x_2, \ldots, x_n \tag{3}$$

of an unknown constant should be chosen as its final value (as the "best" estimator of x). Finally, Gauss required that the joint probability of obtaining observations (3) or, rather, errors $(x - x_i)$ should be maximal if $\hat{x} = \bar{x}$:

$$\varphi(\bar{x} - x_1)\, \varphi(\bar{x} - x_2) \cdots \varphi(\bar{x} - x_n) = \max. \tag{4}$$

(Here, the differences $(x - x_i)$ were the true errors of observation and $(\bar{x} - x_i)$ the apparent errors.)

This condition, (4), the principle of maximum likelihood, which had been stated by Johannes Heinrich Lambert in 1760 and used by Daniel Bernoulli in 1778, led Gauss to the "normal" distribution (a later term)

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

Applying (4) once more, Gauss easily derived condition (2). For the case of one unknown (of direct observations), he naturally returned to the arithmetic mean, $\hat{x} = \bar{x}$, which thus became the most probable estimator of the true value sought.

The assumptions made weighed heavily on Gauss. Indeed, the postulate of the arithmetic mean failed to seem self-evident, and the principle of maximum likelihood did not ensure a sufficiently high joint probability (4). In addition, the universality of the normal law, although justified by the central limit theorem, conflicted with common sense (and was eventually abandoned).

Gauss's later substantiation of the method of least squares was founded on the principle of maximal weight (or minimal variance) rather than on condition (4), which he abandoned. The variance of a random variable[3] ξ with density $\psi(x)$ given on the interval $(-\infty, +\infty)$ is, by definition,

$$m^2 = \int_{-\infty}^{\infty} x^2\psi(x)\, dx.$$

Note that $\psi(x)$ is not necessarily the normal distribution. If ξ is a random error with values $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, then the "sample variance" of ξ is $[\varepsilon\varepsilon]/n$. Suppose, however, that the apparent errors v_i are substituted instead of the unknown ε_i's; then the sample variance is $[vv]/(n-1)$ and the principle of least variance leads to the method of least squares. The root of the sample variance is called the mean square error, or the standard deviation, σ. Offering this new justification, Gauss changed his terminology as well: understandably, the estimators $\hat{x}, \hat{y}, \hat{z}, \ldots$ were no longer *most probable,* but *most plausible.*

Both the sample variance and the mean square error are distribution-free estimators of the precision of observations. This is extremely important since the definitive theory of errors does not depend on the occurrence of the normal law. Nevertheless, another estimator of precision which lacks this property, the so-called probable error, also gained currency, apparently because of its intuitive appeal, and Gauss himself used it on occasion in his letters [33, 261].

By definition, the probable value of a random variable (of a random error) ξ with density $\psi(x)$ is the root ρ of the equation

$$\int_{-\rho}^{\rho} \psi(x)\, dx = 1/2.$$

For the normal distribution, $\rho = 0.477/h$. The sample estimator of h is [7, Sect. 3]

$$\hat{h} = ((n-1)/2\,[vv])^{1/2},$$

and r, the sample estimator of ρ or the probable error, will therefore be

[3] Both "variance" and "random variable" are later terms. Gauss considered random errors (rather than variables in general), calling them "irregulares seu fortuiti."

$$r = 0.477\ (2[vv]/(n - 1))^{1/2} = 0.675\sigma$$

for the normal law.

By the middle of the 19th century the concept of the mean square error and its bastard substitute, the probable error, became standard and accepted tools in astronomy and geodesy. That they were not widely known or used in physics or chemistry is a sad fact. Mendeleev, however, was an exception.

3. A PROPER PERSPECTIVE

To place Mendeleev's work in perspective, I must consider the (insufficiently studied) situation in contemporaneous physics and chemistry. Specialists in these fields did not master, or did not even know, Gauss's definitive error theory (see Sections 3.2–3.3 below). Moreover, in treating observations, they adhered to a nonstochastic approach (Sect. 3.1), whereas Mendeleev's reasoning about harmonious series of observations (Sect. 6) links him with later probabilistic ideas.

3.1. "Best" observations versus stochastic reasoning. Beginning with Ptolemy, scientists regularly neglected to describe how they adjusted their observations. Since this task cannot be accomplished formally, moreover, the approach is necessarily subjective. In fact, problems in natural science are not infrequently solved in a nonstandard and possibly subjective manner. Johannes Kepler, for example, adjusted Tycho Brahe's observations, correcting them by small arbitrary quantities which remained "within the limits of observational precision" [10, 334]. (This suggests an early use of what is now termed statistical simulation.)

Ancient astronomers as well as modern physicists (and perhaps chemists) sometimes selected "best" observations using the rest to make a rough check; in other words, they based their arguments at least partially on subjective appraisals. Witness Robert Boyle: "... experiments ought to be estimated by their value, not their number; ... a single experiment ... may as well deserve an entire treatise ... [a]s 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 ..." [3, 376]. Indeed, his estimation of the "value" of his experiments likely required some subjective reasoning.

In 1756 Thomas Simpson broke away from this tradition by proving that, for the discrete uniform and the discrete triangular distributions, the arithmetic mean was stochastically preferrable to a single observation.[4] As he put it, he aimed to refute "some persons, of considerable note, [who] have been of [the] 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]" [34, 82].

The rule of choosing the "best" observation nevertheless persisted in some branches of natural science, if not in astronomy. James Prescott Joule [9] determined five values for the mechanical equivalent of heat and rejected four of them. The selected value satisfied two conditions at once: it was derived from the experiment consisting of the maximal number of observations; and the coefficient of the un-

[4] In 1757, he also studied the case of the continuous uniform distribution.

known was maximal among the five of them so that the error of measurement was divided by the largest factor.[5]

The choice of a "best" measurement can mean the denial of the stochastic approach, a position actually defended by a recent author. Eric Mendoza has maintained that

> The original applications of error theory [in astronomy and geodesy] were concerned with many repeated measurements of the same simple quantity, a scale reading of the angle of a telescope or theodolite, and this had no obvious relevance to a small number of complex measurements of atomic weight or specific heat. Chemists were preoccupied with arguments about techniques, purity of samples, and so on, and defended their results against all others. Physicists were sometimes more concerned with limits of accuracy,[6] but used a whole variety of methods to estimate them. [27, 283]

In fact, the difference between the two pairs of branches of natural science, namely, astronomy–geodesy and chemistry–physics, was much more subtle than Mendoza contended.

In astronomy and geodesy, methods of observation were devised and developed, instruments required adjustment, and the "simple quantities" to which Mendoza referred had to receive several corrections. In spite of these precautions, systematic errors, such as those occasioned by lateral (horizontal) refraction, corrupted observations and were no less harmful in astronomical and geodetic work than was the impurity of samples in chemistry.[7] Moreover, what Mendoza termed "complex measurements" in physics and chemistry (to distinguish them from the alleged "simple quantities" measured in astronomy and geodesy), likely involved the carrying out of a series of "simple" observations, as was the case with Joule.

Note, however, that physical, chemical, or astronomical constants are determined in at least several places (laboratories, observatories), possibly over a substantial period of time, so that the deduction of their "final" values is really complicated. On the other hand, the angles of a chain of triangulation are measured only once (in one series of observations).

Whether to take a stochastic approach or to opt for selecting the "best" observations is indeed a difficult choice. The practitioner—whether astronomer, geodesist, chemist or physicist—should choose the latter possibility if, but only if, the presence of considerable systematic influences is suspected in the data and cannot be excluded.

[5] It seems possible that Joule selected the "best" experiment beforehand and chose the number of observations accordingly. I have calculated the mechanical equivalent of heat, taking into account all five of his experiments, but did not thus change his value significantly.

[6] "Accuracy" is a term connected with the action of systematic errors whereas random errors determine "precision." Judging by his discussion of physical literature (Sect. 3.2), the author was thinking of the second case.

[7] At the same time, geodesists enjoy the possibility of averaging out some systematic influences by measuring angles under differing meteorological conditions. In addition, the existence of redundant observations in a chain of triangulation (e.g., of three rather than two measured angles in each triangle) makes it possible to reveal and, to some extent, to eliminate residual systematic errors.

3.2. Estimating the plausibility of observations. As noted in Section 2, the plausibility of observations was measured by two estimators, the mean square error and (in spite of its dependence on the law of error involved) the probable error. However, physicists and chemists have often fared worse by making use of doubtful estimators, *viz.*, of

$$(x_n - x_1), \quad (x_n - \bar{x}), \quad \text{or} \quad (\bar{x} - x_1), \tag{5}$$

where x_1 and x_n were the extreme observations ($x_n > x_1$) and $\bar{x}$ was the mean of all n observations, or of

$$\frac{(x_n - x_1)}{\bar{x}}, \quad \frac{(x_n - \bar{x})}{\bar{x}}, \quad \text{or} \quad \frac{(\bar{x} - x_1)}{\bar{x}}. \tag{6}$$

Thus, Henry Cavendish [4, 284] used both the differences (5) and the quotients (6) as early as 1798, whereas James Ivory [32, 179] preferred the quotients in his work of 1830. In [27, 292], however, Mendoza cited other relevant physical memoirs published in 1842 and 1877. In addition, he quoted John William Strutt Rayleigh's paper of 1883 ("The degree of accordance in the numbers ... shows the success of the observations ..." [27, 294]) as an instance of the use of estimators different from (5) and (6). Nevertheless, the accordance (or diversity) between the numbers, to which Rayleigh referred, seems to be measured by the same estimators.

The use of estimators (5) or (6) is hardly justified since they depend on whether or not the extreme observations are rejected and since both the differences and the corresponding quotients tend to increase with the number of observations.

3.3. The normal law. Many scholars, who should have known the classical results achieved by Gauss, were hardly conversant with them, and their ignorance was not restricted to estimating the plausibility of observations. Thus, in 1826–1828 Ivory published several papers on the adjustment of pendulum observations, only gradually mastering the method of least squares [32]. Generally speaking, however, the second most widely misunderstood statistical concept among the physical scientists was the law of error.

The normal law continued to be regarded as the universal law of error for several reasons. First, textbook writers (mostly astronomers and geodesists) usually ignored Gauss's much more complicated mature approach. Second, the exponential law was easy to use and more or less reasonably described the scatter of observational errors. Third, in natural science the normal law became especially popular after 1860, when James Clerk Maxwell nonrigorously proved that the distribution of molecular velocities, appropriate to a gas in equilibrium, is normal.

One immediate consequence of the mistaken belief that the normal law was the universal law of error was the implicit, or perhaps even instinctive, assumption that the formula for calculating the probable error (Section 2) is universally valid.

The above discussion of "best" observations, plausibility of observations, and the normal law sketch the contours of 19th-century physical scientists' understanding of statistical methods. It serves as a backdrop to the work of the chemist Mendeleev, to which I now turn.

4. THE ARITHMETIC MEAN AND THE MEDIAN

Two simple cases provide keen insight into Mendeleev's treatment of direct observations: his use of the arithmetic mean and his reference to the median. While estimating the volume of oil export from the United States in the years 1870–1874, he stated that he "had only one way to judge the extent of error in the published data [on the statistics of petroleum]: to check the maximum possible number of independently obtained figures against one another. If there are no grounds for preferring one set of figures to another ... the average of all figures should be taken ... as the true value, but it should not be assumed error-free ..." [21, 156]. Mendeleev [14, 181] had already formulated the same recommendation even more resolutely, in 1856, hinting that the arithmetic mean should be chosen if the separate results were of unknown plausibility, and he persisted in this conviction when, in 1895, he declared that "[o]nly when the relative merit of various determinations is either absolutely unknown or not clearly determined by anything, is it possible, and sometimes even necessary, to choose their mean" [26, 159].[8]

In [14, 26], Mendeleev was concerned with physical measurements (not with oil export), and his insistence on the mean seems all the more strange. However, at least once, without elaborating, he [19, 209] favorably referred to a work of J. E. Estienne (apparently [6]), in which Estienne had attempted to prove that the median was always preferrable to the mean. (Still, Mendeleev should have qualified his reference to the over-enthusiastic French author: for the case of near-normal observations, it is, in fact, better to use the arithmetic mean.)

Elsewhere, studying the influence of fertilization on yields, Mendeleev stated that "[t]he mean number [the arithmetic mean] calculated from data obtained under differing conditions, by different methods and observers, says little and is always less probable than results achieved by precise methods and experienced persons" [16, 101].[9] Mendeleev did not elaborate on the phrase "different methods." Apparently, he had in mind insufficiently studied methods which led to doubtful results.

5. SELECTION OF OBSERVATIONS AND DESIGN OF EXPERIMENTS

Mendeleev categorically rejected the use of doubtful data, denying them even the slightest consideration: "When, however, one of the numbers gives demonstrably better assurance of precision than the others, it alone should be taken into account, ignoring the numbers that either certainly represent worse experimental or observational conditions or give any cause for doubt. ... To consider worse numbers taking them with some [even small] 'weight' is tantamount to deliberately corrupting the best number" [26, 159]. In one case, however, Mendeleev did assign smaller weight

[8] According to modern beliefs, in such cases one should choose the median and, for example, Andrei Nikolaevich Kolmogorov [11] expressly advised doing so.

[9] While this is certainly the case, the terminology "less plausible" would have been better than "less probable" (cf. Gauss's mature justification of the method of least squares in Sect. 2). Note also that Mendeleev's memoir [16] contained recommendations (although they were not methodically presented) which belong to the prehistory of experimental design.

to some observations because of their lesser "harmony" [24, 458] (see Sect. 6 below). Yet, even in this paper, he made clear his opinion of doubtful data:

> The analysis of data and the selection therefrom of reliable figures demand so much work and time that it is more advantageous to make new observations instead. But, once the new set of numbers has been introduced, the similar, more or less random observations not only give nothing new but they simply complicate the search for a realistic result. Until the disadvantageous [unfavorable] data have been eliminated by a clear critical appraisal, there is no hope of achieving a realistic result. [24, 82][10]

Note that Mendeleev rejected here observations which either "behaved irregularly," or were not precisely described, or were unreliable.

Rejection of outlying observations is an extremely delicate procedure. In a letter to Wilhelm Olbers dated May 3, 1827, Gauss [8, 152–153] stated that "without an extensive knowledge of the subject [of astronomy], rejection is always questionable [*misslich*], especially if the number of observations is not very large." Many statistical criteria for rejection have appeared since Gauss's time, and a few of these could have been known to Mendeleev. Still, as Vic Barnett and Toby Lewis admit in their 1984 book on the subject, "the major problem ... remains the one that faced the very earliest workers in the subject—what is an outlier and how should we deal with it?" [1, 360].

Since Mendeleev demanded a precise description of data and their analysis, he evidently thought about the most difficult problem of uniting the results obtained by several observers.[11] No wonder that he therefore preferred to collect data according to the principle "better less but better." This attitude is clearly reflected in his 1872 statement concerning a determination of the empirical relation between the density of a gas and its pressure, i.e., a refinement of the Boyle—Mariotte law:

> In order to avoid, if possible, the application of the method of least squares, I prefer to make a few but precise and repeated measurements at several significantly different pressures ... Amassing observations made at various closely spaced pressures not only presents many difficulties for analysis, but also increases the error in drawing inferences.
>
> Out of the many now existing observations of the compressibility of gases it is yet impossible to determine precise numbers with known limits of error, and the most basic reason for this [state of affairs] is the insufficient knowledge of the degree of precision of the observations. [17, 144][12]

In the first half of this passage, Mendeleev alludes to the fact that, with a large number of observations, the application of the method of least squares to their adjustment would have required considerable effort. Mendeleev's main point, however, is his recommendation about differing pressures. Apparently, he was con-

[10] Mendeleev formulated a similar thought somewhat earlier [22, 362] in terms of the expression "worthless data."

[11] *Cf.* Mendoza. The comparison of data collected in different laboratories began [in earnest] in the 1880s [27, 290]; yet systematic errors in physics greatly impeded such activities "during much" of the 19th century [27, 291].

[12] Mendeleev effectively repeated this same reasoning in [18, 256]. Note that the limit of error was then and is now usually determined by the appropriate variance.

cerned with the influence of systematic errors which might have depended on the values of the pressure p. Indeed, suppose that a systematic error ε is a function of p: $\varepsilon = \varepsilon(p)$. In this case, observations made at significantly different values of p will result in $\varepsilon(p)$ essentially changing its magnitude. On the other hand, observations made at closely spaced pressures will be corrupted by almost the same systematic error: this alternative is worthless.

An interesting case in point is the adjustment of pendulum observations (which Mendeleev did not mention). Jean Baptiste Biot [2, 16–17] remarked that, before pendulum stations are adjusted, those having almost the same latitude may be easily replaced by one fictitious mean station. Since the period of vibration of a pendulum of a given length, or its length when the period is the same, depend only on the latitude of the station, Biot was correct. However, if the replaced stations were actually close to one another (i.e., if, in addition, their longitudes did not differ much), then all the observations could be corrupted by almost the same local gravimetric anomaly. In this case, the weight of the mean station should not be increased as compared to that of a real station, an observation Biot did not make.

6. HARMONIOUS OBSERVATIONS

Mendeleev [19, 209] called a series of observations "harmonious" if its median coincided with its arithmetic mean, adding, however, that he preferred another definition of this concept, to wit, the coincidence of the mean of its middlemost third ($\bar{x}_2$) with the mean of the means ($\bar{x}_1$ and $\bar{x}_3$) of its extreme thirds. He also remarked that Russia's Main Board of Weights and Measures kept to the "usual Gaussian methods," that is, it obviously did not make any use of the median.

Nevertheless, Mendeleev himself hardly ever used this estimator (see Sect. 4). Neither did he explain what practitioners should do if their observations were not harmonious. He may (sometimes?) have checked whether the median of his observations coincided with their mean, and if it did not, this may have cast some doubt in his mind. In one such case (see beginning of Sect. 5), at least, Mendeleev did assign his observations a small weight. In another instance, he expressed satisfaction at obtaining harmonious observations (but made theoretical mistakes and committed a few elementary errors in his pertinent calculations): "The probable conclusion ... here agrees perfectly with the arithmetic mean, and this indicates that the errors [of observations] follow a definite law assumed by the Gaussian theory of probability [of errors!], i.e., that the observations do not contain large random deviations, but are subject to unavoidable observational errors ..." [19, 209].

Several problems appear in this passage relative to Mendeleev's understanding of this sort of analysis. First, "probable conclusion" is an imprecise and, therefore, unfortunate substitute for "median." Second, Gauss, in his earlier justification of the method of least squares, arrived at "a definite" (the normal) law by assuming that its mode (not median!) coincided with the mean of an observational series (see, however, a qualifying remark below). Third, the normal law allows large errors to exist, although with a low probability. Fourth and last, the implicit assumption that the case of normally distributed errors of observation is the best possible one is correct. The latter opinion is accepted by modern statisticians in spite (or, rather, irrespective) of Gauss's second substantiation of the method of least squares.

It is extremely interesting that Mendeleev's notion of harmony is recognized in modern statistics: observations are called harmonious in his sense if their distribution is symmetrical. Indeed, in [35, 161] one of the measures of asymmetry is

$$\text{Skewness of distribution} = 3\ (\text{Mean} - \text{Median})/\text{Standard deviation}.$$[13]

Here, it is important to note that Mendeleev was not altogether wrong in stating that the coincidence of the median with the mean is characteristic of the normal law. The relation

$$\text{Mode} = \text{Mean} - 3\ (\text{Mean} - \text{Median})$$

holds "with a surprising closeness for moderately asymmetrical distributions ..." [35, 117].

7. MEASURES OF PRECISION OF OBSERVATIONS

At least twice Mendeleev made use of probable error in order to determine the permissible deviation of observations from their mean, or the permissible difference between two means. In [15, 46], his value of the density of a certain substance differed from the value derived by another experimenter by more than the sum of the probable errors of the respective means. Consequently, he decided that the substance analyzed by his predecessor was of lesser purity.

In fact, Mendeleev assumed that observations should not deviate from their mean by more than its probable error. He used the same reasoning in [18, 312, note 2]. Let $\bar{y}$ be the arithmetic mean of n observations y_i, $i = 1, 2, \dots, n$, and denote $z_i = \bar{y} - y_i$. Then the probable error of the mean will be

$$r = 0.675\ [(z_1^2 + z_2^2 + \cdots + z_n^2)/n(n-1)]^{1/2},$$

and the "true value" of the constant sought will be located in the interval $[\bar{y} - r; \bar{y} + r]$. However, the above expression for r holds only for the normal distribution and, furthermore, the permissible error of $\bar{y}$ was usually taken, at least in astronomy, as $3r/0.675$ (the "three sigma rule").

Mendeleev adduced, without explanation, an approximate expression for the probable error,

$$\begin{aligned} r &= 1.196\ (2n - 0.86)^{-1/2} \sum |\varepsilon_i|/n \qquad (7) \\ &= (1.196/\sqrt{2})\ (n - 0.43)^{-1/2} \sum |\varepsilon_i|/n. \end{aligned}$$

He thus made use of the so-called mean error of observations,

[13] Skewness can force the practitioner to reject a large proportion of extreme observations. According to one meteorologist, since the densities in meteorology are asymmetric, the theory of errors in this branch of science is "inadmissible in principle [principiell unzulässig]" [28, 32].

$$\theta = \Sigma |\varepsilon_i|/n = \Sigma |y - y_i|/n,$$

where y is the true value sought. For normally distributed errors

$$\theta = \sigma\sqrt{2}/\sqrt{\pi},$$

where σ is the mean square error and [7, Sects. 4 and 6]

$$r = 0.84535 \ldots \theta. \tag{8}$$

Although $1.196/\sqrt{2} \approx 0.845$ and although the Gaussian relation (8) should be transformed by taking into account the apparent deviations, $|\bar{y} - y_i|$, rather than the "true" ones, $|y - y_i|$ [29], formula (7) is wrong even for the normal distribution, and its origin in Mendeleev's work is a mystery.

There is one curious fact which exonerates Mendeleev with regard to his use of "twice the probable error" as a criterion of permissible empirical differences. Even in 1903 Andrei Andreevich Markov, referring to an oral remark made by the eminent astronomer, Fedor Aleksandrovich Bredikhin, stated: "I like very much Bredikhin's rule, that 'if a computed quantity is to be accepted as realistic, it must be at least twice as large as its probable error.' Only I do not know who established this rule or whether it is recognized by all experienced calculators" [30, 21].

8. CONCLUSION

Neither physicists nor chemists of the late 19th century knew the mature Gaussian theory of errors.[14] This fact can be explained by their attitude of selecting the "best" experiment rather than of adhering to the stochastic approach; in turn, the reliance on the "best" was likely occasioned by the presence of undetectable systematic errors in their experiments.

In adjusting direct observations, Mendeleev placed too much trust in the arithmetic mean at the expense of the median, and in estimating the precision of data, he sometimes made incorrect or at least doubtful use of the probable error. On the other hand, Mendeleev expressed reasonable ideas on the need to base all inferences on plausible data and to avoid unnecessary and even disadvantageous amassing of observations. Instead, he planned his own experiments and demanded a precise description and a preliminary study of all the data available. Most interesting are Mendeleev's thoughts on the need to obtain observations whose errors have a symmetric density. In this respect, he anticipated some modern statistical ideas, although, perhaps because they were not encountered in physics or chemistry, he did not indicate that asymmetric distributions have to be tackled as well.

[14] Even a recent author testified that the existence of the second formulation of the method of least squares "seems to be virtually unknown to almost all American users of least squares except students of advanced mathematical statistics" [5, 24].

ACKNOWLEDGMENTS

This paper represents part of a research program undertaken at the Mathematical Institute of the University of Cologne (Professor J. Pfanzagl) with the support of the Axel-Springer Stiftung. I am grateful to the referees, who justly demanded that the first version of this paper be rewritten. Professor Roger Cooke gave me sound historical and linguistic advice, and Professor Karen Parshall helped me to strengthen my arguments. A preliminary version of this paper appeared in *Istoriko-matematicheskie issledovania* **35** (1994), 56–64.

REFERENCES

A note about the contributions of Mendeleev. All of Mendeleev's contributions to which I refer originally appeared and were reprinted in Russian. The reprints were published between 1934 and 1952 in his *Sochinenia* [Works] (25 vols.) in Leningrad by the Ob'edinennoe Nauchno-Tekhnicheskoe Izdatel'stvo (vols. 1–4) or in Leningrad and Moscow by the Akademia Nauk SSSR (the other volumes). The editors were V. E. Tishchenko (vols. 1, 4, 6); V. G. Khlopin (vols. 5, 16–20, 22); L. M. Khlopin (vols. 7–15); S. E. Frish (vols. 21, 23, 25); A. N. Bakh, B. N. Vyropaev, I. A. Kablukov, N. S. Kurnakov, and V. E. Tishchenko (vols. 2 and 3); and G. S. Vasetsky (vol. 24). Original publication dates are given in [brackets].

1. Vic Barnett and Toby Lewis, *Outliers in Statistical Data,* Chichester: Wiley, 1984.
2. Jean Baptiste Biot, Mémoire sur la figure de la terre, *Mémoires de l'Académie des sciences Paris* **8** (1829), 1–56.
3. Robert Boyle, A Physico-chymical Essay. *Works,* 6 vols., London: Rivington, 1772, **1:** 359–376; reprint ed., Hildesheim: Olms, 1965–1966.
4. Henry Cavendish, To Determine the Density of the Earth (1798), *Scientific Papers,* ed. J. C. Maxwell reviewed by Sir J. Larmor (vol. 1) and Sir E. Thorpe (vol. 2), 2 vols., Cambridge, UK. Cambridge Univ. Press, 1921, **2:** 249–286.
5. Churchill Eisenhart, The Meaning of "Least" in Least Squares, *Journal of the Washington Academy of Sciences* **54** (1964), 24–33.
6. J.-E. Estienne, Étude sur les erreurs d'observation, *Revue d'artillerie* **36** (1890), 235–259.
7. Carl Friedrich Gauss, *Bestimmung der Genauigkeit der Beobachtungen,* 1816, or *Werke,* 12 vols., Göttingen: Königlichen Gesellschaft der Wissenschaften, 1870–1929, **4:** 109–117.
8. Carl Friedrich Gauss, *Werke,* 12 vols., Göttingen: Königlichen Gesellschaft der Wissenschaften, 1900, **8**.
9. James Prescott Joule, On the Mechanical Equivalent of Heat, *Philosophical Transactions of the Royal Society of London* (1849 for 1850), 61–82.
10. Johannes Kepler, *New Astronomy,* trans. W. H. Donahue, Cambridge, UK: Cambridge Univ. Press, 1992. [originally published in Latin in 1609].
11. Andrei Nikolaevich Komogorov, Applying the Median in the Theory of Errors, *Matematichesky Sbornik* **38,** nos. 3–4 (1931), 47–49. [In Russian].
12. Pierre Simon de Laplace, Sur les approximations des formules qui sont fonctions de très grands nombres, 1810, or in *Oeuvres complètes,* publiées sous les auspices de l'Académie des sciences par MM. Secrétaires perpétuels, 14 vols., Paris: Gauthier–Villars, 1878–1912, **12:** 301–353.
13. Adrien-Marie Legendre, *Nouvelles méthodes pour la détermination des orbites des comètes,* Paris: Courcier, 1805.
14. Dmitrii Ivanovich Mendeleev, Specific Volumes, *Sochinenia* **1:** 139–311 [1856].
15. Dmitrii Ivanovich Mendeleev, On the Cohesion of Some Liquids, *Sochinenia* **5:** 40–55 [1860].
16. Dmitrii Ivanovich Mendeleev, Report on the Experiments of 1867 and 1869, *Sochinenia* **16:** 99–113 [1872].
17. Dmitrii Ivanovich Mendeleev, On the Compressibility of Gases, *Sochinenia* **6:** 128–171 [1872].

18. Dmitrii Ivanovich Mendeleev, On the Elasticity of Gases, 1, *Sochinenia* **6:** 221–589 [1875].
19. Dmitrii Ivanovich Mendeleev, Progress of Work on the Restoration of the Prototypes of Measures of Length and Weight, *Sochinenia* **22:** 175–213 [1875].
20. Dmitrii Ivanovich Mendeleev, On the Temperatures of the Atmospheric Layers, *Sochinenia* **7:** 241–269 [1876].
21. Dmitrii Ivanovich Mendeleev, The Oil Industry in Pennsylvania and in the Caucasus, *Sochinenia* **10:** 17–244 [1877].
22. Dmitrii Ivanovich Mendeleev, The Dependence of the Specific Weights of Solutions on Their Composition and Temperature, *Sochinenia* **4:** 279–383 [1884].
23. Dmitrii Ivanovich Mendeleev, A Note on the Scientific Work of A. I. Voeikov, *Sochinenia* **25:** 526–531 [1885].
24. Dmitrii Ivanovich Mendeleev, An Investigation of the Specific Weight of Aqueous Solutions, *Sochinenia* **3:** 3–468 [1887].
25. Dmitrii Ivanovich Mendeleev, The Power for the Future Now Resting Idly on the Banks of the Donets, *Sochinenia* **11:** 53–207 [1888].
26. Dmitrii Ivanovich Mendeleev, On the Weight of a Definite Volume of Water, *Sochinenia* **22:** 105–171 [1895].
27. Eric Mendoza, Physics, Chemistry and the Theory of Errors, *Archives internationales d'histoire des sciences* **41** (1991), 282–306.
28. Hugo Meyer, *Anleitung zur Bearbeitung meteorologischer Beobachtungen,* Berlin: Springer-Verlag, 1891.
29. Christian August Friedrich Peters, Über die Bestimmung des wahrscheinlichen Fehlers einer Beobachtung, *Astronomische Nachrichten* **44** (1856), 29–32.
30. *Prenia mezdu Akademikami v Zasedaniakh 1-go Otdelenia Akademii Nauk* [Debates between Academicians at Meetings of the 1st Department of the Academy of Sciences (in connection with B. B. Golitzin's nomination for effective membership in the Academy of Sciences)], St. Petersburg: Academy of Sciences, 1903. [In Russian].
31. Oscar Sheynin, On the History of the Statistical Method in Meteorology, *Archive for History of Exact Sciences* **31** (1984), 53–95.
32. Oscar Sheynin, Ivory's Treatment of Pendulum Observations, *Historia Mathematica* **21** (1994), 174–184.
33. Oscar Sheynin, C. F. Gauss and Geodetic Observations, *Archive for History of Exact Sciences* **46** (1994), 253–283.
34. Thomas Simpson, On the Advantage of Taking the Mean of a Number of Observations, *Philosophical Transactions of the Royal Society* **49** (1756), 82–93.
35. George Udny Yule and Maurice George Kendall, *Introduction to the Theory of Statistics,* 14th ed., London: Griffin, 1958.