

Lectures for chemists on statistics. I. Belief, probability, frequency, and statistics: decision making in a floating world

Günther Meinrath

Received: 20 July 2007 / Accepted: 9 November 2007 / Published online: 30 November 2007
© Springer-Verlag 2007

Abstract A brief account on the development of probability theory and statistics is given. These subjects may help the chemist to interpret the information content of his experimental data. Their source and validity, however, is often unclear. A series of papers, of which this is the first, intends to present a clearer view by illustrating the origin of the data-analysis tools. The major figures in this field are mentioned, their achievements presented and the mutual links outlined. The intention is the provision of a common basis for the discussion of statistical aspects relevant to metrology in chemistry. The chemist should see probability theory and statistics as tools to assist in the daily practice of data interpretation to arrive at an objective conclusion.

Keywords Probability theory · Statistics · Determinism · Pearson · Neyman · Fisher

Introduction

“I often say that when you can measure what you are speaking about and express it in numbers, you know something about it. And when you cannot measure it, when you cannot express it in numbers, your knowledge is meager and of unsatisfactory kind. It may be the beginning of knowledge, but you have scarcely, in your thought, advanced to the stage of science, whatever the

matter be” [1]. This famous statement of the physicist Lord Kelvin certainly ignores some achievements of biology and geology but illustrates strikingly the focus on numbers and mathematical formulations in the sciences. These imply precision, clarity, and unambiguity. Galilei’s statement “if I were again beginning my studies, I would follow the advice of Plato and start with mathematics” combines two major names in history of human intellect in their appreciation of mathematics. Mathematics imply purity, symmetry, harmony, and truth. These elements describe what human lives lack. Human life is unstable and, for sure, unpredictable. Despite expressing itself in the language of mathematics with numbers and formulas, statistics does not have a reputation similar to that of, e.g., geometry and arithmetic, or physics. “Statistics—the curse of the analytical classes” is the title of an article on the use and abuse of statistics in papers devoted to analytical chemistry [2]. Science (from the Latin “scire”: to know) is devoted to knowledge. Statistics is devoted to doubt and uncertainty. Hence the subjects seem to have conflicting (if not mutually exclusive) objectives. There are numerous jokes (and less joking quotes) about statistics and statisticians. Isn’t it foolish to study uncertainty and doubt when all the scientific world craves for certainty? Knowledge and uncertainty are opposite faces of the same medal. One cannot go without the other. Objectivity in proposing a scientific hypothesis also requires a statement on the relevant aspects the hypothesis cannot cover.

“It is the primary task of statistics intended for scientific use to convince oneself and others that the data have been interpreted fairly” [3]. No learned person would reject statistics as a useless or even stupid subject. Nevertheless, when his arguments are commented on with reference to statistics, the predominant reaction is depreciation. Or

G. Meinrath
RER Consultants Passau, Schiessstattweg 3a,
94032 Passau, Germany

G. Meinrath (✉)
Technische Universität Bergakademie Freiberg,
Leipziger Str. 27, 09599 Freiberg, Germany
e-mail: rer@panet.de

helplessness when advice on certain aspects of a data set is given in terms of statistics. The reasons for these reactions lie, in part, in the lack of formal education of most chemists in statistics. In many cases statistical knowledge reduces to mean value and standard deviation. Few are able to explain, e.g., the meaning of degrees of freedom. The meaning of the correlation coefficient in linear regression [4–7] is another practical example of the role of statistical properties in chemistry. Despite its rather poor performance, the correlation coefficient is regularly found in chemical communications for the purpose of underscoring the quality of the data presented. The other situation is the frustration in the eyes of a chemist seeking advice on his/her data interpretation. Modern statistical techniques, e.g. computer-intensive statistics, non-parametric statistics, and classification techniques including factor analysis, are widely unknown. It is not the primary objective of this paper, the first in a series, to remedy the situation. However, metrologists have the task of communicating the magnitude of doubt associated with a set of measurement values. And doubt is one subject where statistics can make a helpful contribution—if some basic understanding of the history and methods of this important field of data analysis are available. No science can escape the stamp of time which coins its progress and its language. Therefore this paper deals mainly with history. To focus the vast subject, the development of statistics through the centuries will be described on the basis of eminent contributions. The presentation is necessarily subjective.

A highly abbreviated history on probability and statistics

Arguably, statistics does not have its basis in pondering doubt and probability, but in data collection. The words “statistics” and “state” do have the same linguistic roots. The oldest scriptures are telling not about gods but about goods. The administrators of the first complex human social structures had to document tax laws and their implementation by citizens. Thus the libraries of cuneiform writings in Khorsabhad and Ninive (established in Mesopotamia about 1000–500 BC) did not predominantly forward artistic literature like the Gilgamesh Epic but heaps of clay panels holding the number of crop bushels delivered to the ruler’s stores and the tributes received from conquered neighbors. The initially pictorial symbols used to denote these quantities eventually developed into script. In a world where life was short and instability prevailed, human concepts about the surrounding world were handed down orally as myths. Homer’s Iliad and Odyssey are prominent examples from the Achaian period of Greek history (eighth century BC). A major achievement of

pre-Socratic philosophers in Greece (fifth century BC) was to separate rational reasoning from myth. Prior to this, the gods and their deeds were as real as are electrons and molecules in our times. Logical reflection about this world (as perceived by the five senses) provided evidence that some events can be predicted while in other cases the event cannot be known beforehand. In order to know, one must try. In Latin, to try is “probare” and from this the word “probabilitas” is derived. The earliest documentation of this word comes from Cicero [8] and connects what usually happens with what is ordinarily believed [9]. From the subjective interpretation of probability to a mathematically founded probability theory was a long journey, and one which had many sideways dead-ends. Thus, it is easier to say where mathematical probabilities did not come from than the opposite [10]. Where chance rules, determinism was sought—and found in the seemingly eternal stability of the stars. Here was predictability, stability, and harmony. At least, if the outcome of an event is uncertainty, one wants to know what the likely outcomes might be or what could be expected. In fact, for lasting years during early modern times chance was discussed usually in connection with expectation E :

$$pV = E \quad (1)$$

where p is the probability of the event, V its outcome value, and E the expectation (note: this is a historic definition of expectation). If the probability p of obtaining an outcome of $V = 100 \text{ €}$ is 0.1, then the expectation E is 10 €. In relationship expressed by Eq. (1) probability has changed from a term expressing personal belief into a (dimensionless) quantity. It may, therefore, not be surprising that the first text offering a mathematical formulation for probability was in terms of what players in a game should get. Huygens’ “De ratiociniis in ludo aleae” [11] was published briefly after Blaise Pascal and Pierre de Fermat exchanged letters over a question nowadays known as the “problem of points”. The question was posed by A. Gombaud Chevalier de Méré: two players P1 and P2 agree to play a series of fair games until one of them has won a specified number N of games. The play is suddenly interrupted. P1 has won N_1 games and P2 N_2 games. How should the stakes be divided? A series of publications on probabilities was initiated, from Graunt’s “Observations on the Bills of Mortality” [12] to Bernoulli’s “Ars Conjectandi” [13]. Graunt’s work on mortality tables, together with Halley’s similar work using data of the City of Breslau, reminds us that probability was seen as a subject of science but not primarily as a tool for science. The “Ars Conjectandi” was published posthumously in 1713 and contains, in addition to an annotated version of Huygens’ treatise, a proof of Newton’s binomial theorem,

$$(1+x)^n = \sum_{i=0}^n \binom{n}{i} x^i, \quad (2)$$

an application of Eq. (2) to games with dice and an outline of possible uses in civilian, moral, and economic matters (developing the (weak) law of large numbers). Experimental science was just at its beginning. The first natural law (on acceleration) was formulated by Galilei in approximately 1610. The great period of systematic experimentation was just to start. Nevertheless, Galilei emphasized in the “Discorsi” [14] that measurement results cannot be arbitrarily exact

From the data collected during the eighteenth century on such diverse subjects as birth, death, crime, and the movement of stars it became evident that chance follows rules. De Moivre even mentioned the error law Eq. (3)

$$N = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (3)$$

in 1733 even though its basis remained enigmatic. Bayes’ posthumously published treatise “An essay towards solving a problem in the doctrine of chances” [15] laid the foundations of what is known today as conditional probabilities. In its simple form Bayes’ theorem is given by Eq. (4)

$$P(A|B) = \frac{P(A) P(B|A)}{P(A) P(B|A) + P(A') P(B|A')}. \quad (4)$$

where events A and B have probabilities of $P(A)$ and $P(B)$, respectively. A' is the complementary event of A. The symbol $P(A|B)$ denotes a conditional probability, that is a probability that event A occurs provided event B occurred. The importance of the theorem can be appreciated from both the terms $P(A|B)$ and $P(B|A)$. If the conditional probability $P(B|A)$ is known but $P(A|B)$ is required, Eq. (4) provides a way to obtain it. Bayes’ theorem, despite its importance, often remains obscure to people for several reasons. First the concept of “conditional probability” is unknown to most chemists. Further, it is clear that in case of independent (unconditional) events the probability of observing the two events A with probability $P(A)$ and B with probability $P(B)$ is $P(A)P(B)$ but “conditional probability” is less clear. Therefore, a brief example may illustrate its importance.

Consider an immunoassay. The immunoassay will be positive in presence of a doping substance in 99% of cases. Only in 1% of the cases where no drug is present the immunoassay will be positive (so-called false-positive result). The percentage of samples to the laboratory actually contaminated with the drug is, say, 1%. The question of interest is: if the immunoassay indicates contamination with doping drug, what will be the probability that the sample is actually contaminated? This is a simple example

because only two possibilities exist. The probability $P(A)$ that a sample holds the doping drug is 0.01. The probability that the immunoassay is positive if the sample is contaminated is 0.99. The probability $P(A')$ that the sample is not contaminated is $(1 - P(A)) = 0.99$. The probability $P(B|A')$ that the immunoassay is positive despite the fact the sample is clean is 0.01. Hence we have

$$\begin{aligned} P(A|B) &= 0.01 \cdot 0.99 / ((0.01 \cdot 0.99) \\ &\quad + (0.99 \cdot 0.01)) \\ &= 0.5. \end{aligned} \quad (5)$$

The result can be presented in another way. Of 10,000 samples, only 100 are contaminated (and 99 of these are detected by the immunoassay). 9,900 samples are not contaminated. Of these 9,900 samples, 1% are falsely detected positive. i.e. 99 samples. Hence, if the immunoassay forwards a positive result, it is correct in only 50% of the cases. This example may also illustrate the importance of the presumption of innocence. It is also important to recognize that Bayes’ theorem is valid irrespective of whether frequencies or subjective probabilities are used. Bayes’ theorem plays an important role if the probabilities are of a subjective nature. If there are assumptions about the distribution of the outcome A (say a measurand in a measurement), these assumptions can be expressed by a probability function $f(A)$. This function is the prior. If some measurements on A are performed, new data D are available and the distribution $L(A|D)$ needs to be evaluated (the likelihood function). Using Eq. (4) with $P(A) = f(A)$ and $P(B|A) = L(A|D)$, the posterior function $P(A|D)$ can be obtained. It represents the new information about the distribution of the measured values given the new data D. The numerator is usually chosen to ensure that the area under the distribution $P(A|D)$ is unity. Calculating with distributions is usually not simple and, in its time, Bayes’ theorem was mostly ignored. Nowadays, with computers at hand for numerical integration, Bayes’ theorem is receiving increasing attention [16–18].

Probability theory around 1750 concentrated on mortality, life insurance, and annuities. Statistics contributed to the social field—including gambling. Leibnitz considered probability theory as a mathematical version of legal reasoning. It is, at this stage, important to understand that probability and chance only referred to limited human knowledge. The world itself was deterministic. Given the initial constraints were known, all future development was thought to be predictable. On this basis, Laplace analyzed the seemingly odd behavior of the planets Jupiter and Saturn [19], thereby not only illustrating the stability of the solar system but also introducing new mathematical concepts (e.g. the term “potential” and the Laplace operator). His treatise on probability [20] became a cornerstone of

probability theory. Before Laplace, probability theory was solely concerned with mathematical analysis of games of chance. Laplace applied probabilistic ideas to many scientific and practical problems. The theory of errors, actuarial mathematics, and statistical mechanics are examples of some of the important applications of probability theory developed in the nineteenth century.

For metrologists, the period between 1760 and 1810 is of special significance. Laplace, like Lavoisier, was a member of the Metrological Commission founded in 1791 to establish unified standards for weight and measures. The foundations of what today has developed into the SI were laid in these times. From these times also came the recognition that chance follows laws. For one, the mortality tables proved the constancy of female and male birth rates and Gauss [21] gave an elegant proof of Eq. (3), nowadays known as the Gaussian distribution. In between, the expeditions sent out to measure the Paris meridian [22] accomplished their goal [23] and, eventually, established internationally recognized references like the meter, the kilogram, and the second. Even though probability theory and science went separately, the laws of chance and probability were developed in parallel to the extremely deterministic natural laws of Galilei and Newton. In the nineteenth century scientists were searching for laws, causality, and prediction. Probability theory, however, did not provide a sufficient basis for any conclusion about individual events. While one faction of the 19th centuries probabilists focused on the mean values of their observations (e.g. crime, marriages, suicides, birth, death) thereby discovering surprising stability, the other faction studied variation revealing fluctuation incompatible with pure chance. Still, measurement data did not play a role in probability theory. The leaders in probability theory in the nineteenth century concentrated on eugenics [24]. Eugenics deals with interference to improve the human gene pool. Galton invented the term “eugenics” and wrote the first book on the subject [25]. Galton also invented linear regression and pioneered the use of the Gaussian distribution, Eq. (3), to interpret histograms of collected data. Galton’s successor was Karl Pearson. Pearson is a co-founder of psychology and biometry. His contributions on statistics should be known to almost all scientists—he is the originator of the correlation coefficient [26], the χ^2 goodness-of-fit test [27], and the method of moments [28]. The institution created by Galton at University College, London, became an Institute of Applied Statistics in 1911. The institute attracted further excellence, e.g. Pearson’s son Egon, Jerzy Neyman, John Wishart, and Ronald Fisher. The often recipe-like application of numerical operations to derive certain figures of merit (mean value, standard deviation, confidence limits, significance tests) is based on the work of statisticians from Galton to Fisher. A

considerable part of Fisher’s scientific work is accessible at the University of Adelaide Digital Library [29].

Pearson, Neyman and Fisher

Egon Pearson, Jerzy Neyman, and Ronald Fisher shaped what is known to many chemists as “the statistics”. The terms “statistical model”, “hypothesis testing” [30, 31], “Null hypothesis”, “Type I error”, “Type II error”, and “significance test” are some of those introduced by the London group. Unfortunately, despite (or because of) their close work on the same subject, the personal relationship among them was rather poor. A consequence is that a real synopsis of the different views (Neyman/Pearson, Fisher, Bayes) and their specific developments is not available. While statistical methods gradually entered many scientific fields, the in part controversial views of the originators of the methods were ignored or, even worse, mishmashed. It is not uncommon that some of the present readers are reading the names of Pearson, Neyman, and perhaps even Fisher here for the first time. Original literature, where important concepts are developed and not seldom controversially discussed, is rarely referred to. Instead the various approaches are often mixed and the underlying concepts almost ignored. The Neyman/Pearson school compares hypotheses [32] while the Fisherian school uses the significance test [33]. The implications are subtle and often confusing. Fisher significance tests are widely required in medical and other literature and are often applied mechanically [34]. The significance test seems to be favored over hypothesis testing, because only one hypothesis (the Null hypothesis stating that the observed effect is due to random variation) is tested against the data. Here, the only possible error is a Type I error: The effect is accepted as significant despite being random. This risk can be reduced by choosing a small test criterion p . The hypothesis test compares the Null hypothesis against its alternative. Two types of error are possible—a Type I error that the effect is accepted despite being random must be balanced against the Type II error to reject an effect as random despite being significant. R. Fisher moved to Rothamstead Agricultural Station in 1919. Most of the popularization of his statistical methods was done during this time, especially via his textbooks [35–37].

In Rothamstead a large number of data awaited analysis. Typical studies asked for the effect of some substance on biomass production. Which fertilizer is better? What crop-rotation system works better for what kind of soil? Such questions were highly relevant economically. But how to address them? Fisher’s answer was randomization, analysis of variance (ANOVA), maximum likelihood, and the theoretical concepts of efficiency and sufficiency. Here is no

space to fully appreciate the contribution of Fisher who, in addition to his statistical work, continued to contribute to eugenics and genetics. It is, however, almost impossible to overestimate Fisher's impact: "Fisher was a genius who almost single-handedly created the foundations for modern statistical science" [38].

Kolmogorov and the theoretical foundation of probability

This sweep through the concepts of probability would be incomplete without a reference to A.N. Kolmogorov. The name Kolmogorov is, today, mainly associated with the Kolmogorov–Smirnov test [39]. However, Kolmogorov provided an axiomatic basis of the properties the concept of probability should have [40]. Probability is a very old concept but, as will be summarized below, a concept which has multiple meanings. To give an example, classical probability calculus introduced by, e.g., Bernoulli [13] and de Moivre [41], gives the probability of an event as a ratio of equally likely cases that favor it to the total number of equally likely cases possible under the circumstances. The "equally likely case" was important. In play, "equally likely" also means fair. Introducing probability in terms of urn models (as usually practiced in school education) is also implicitly based on the assumptions that any draw from the urn is equally likely. It should be noted in passing that this classical approach led to some paradoxes, e.g. the Bertrand paradoxes [42]. Kolmogorov based probability on set theory. An event is a set from a measurable space (note: "measurable set" is a general mathematical concept implying that certain properties of the set can be quantified). Measure theory assigns to each set of a field of sets a number between 0 and 1, its probability. This setting seems simple but develops into a complex mathematical structure. Striking analogies exist, e.g., between expectation and the integral over the set or orthogonality of functions and independence of random variables. Two principles were offered to connect probabilities $P(X)$ with reality. First, if an experiment E is repeated a large number of times, the relative frequency of event X will differ very slightly from $P(X)$. Second, if $P(X)$ is very small, one can be practically certain that when E is carried out only once, the event X will not occur at all [43]. Thus, the primitive notion of the probability was related to first principles.

Frequentists, Bayesians and the different meanings of probability

All the previous discussions would be of limited importance to metrologists if the situation concerning the

fundamentals of science had not changed radically at the beginning of the twentieth century. One reason for that is quantum theory. Kolmogorov, like Fisher, was a frequentist. To a frequentist the probability of an event is equal to its relative occurrence in a larger number of repetitions of an experiment. Thus, the repeatability of an experiment is an essential element in the application of statistical tools. As co-worker at the agricultural station at Rothamstead and in his eugenic studies, Fisher had access to a large number of data. It was clear, especially after the contributions of Gosset [44], that the law of large numbers required a larger number of repetitions of an experiment if the power of a significance test was not to be compromised. In the social sciences, also, the sample size was usually larger than about 30. Fisher insisted on the use of frequencies because these were based on actually observed/measured data. These data formed the closest connection with reality available. And reality was, still, deterministic. Thus, compared with the subjective probabilities of Laplace, the frequentist approach appeared to be a major progress in the direction of objectivity. During the 19th century, Laplace was the authority on the use of probability: "It is remarkable that a science which began with the consideration of games of chance has become the most important object of human knowledge" [20]. He suggested a subjective interpretation of probability along Bayes' theorem. His famous quote: "Probability theory is nothing but common sense reduced to calculation" is an account to this position. It is also worth noting that the major statistical tool in the 19th century, the least-squares method, was not a subject of dispute, because it can be derived by sampling theory approaches, on the basis of maximum likelihood, and by Bayesian inference with the same final result. The dispute between subjectivists and frequentists is, nevertheless, still vivid [45].

Quantum mechanics gradually brought probability theory into physics and chemistry [46]. Determinism faded and with it, the close relationship between reality and measurement data. Heisenberg's Uncertainty Principle introduced chance to the so-called exact sciences and the Copenhagen interpretation of quantum chemistry made clear that experimenter and experiment are never independent of each other [47]. Probability theory is still expanding, e.g. in biology, genetics, politics, economics, physiology, and psychology. So, it may be expected that at some future date a unified interpretation of probability will become available. However, there are at least three different interpretations of probability. To Laplace and his contemporaries probability was a state of mind. Another view defines probability as an essentially unanalyzable but intuitively understandable logical relationship between propositions. There must be a logical intuition of the probable relationships between evidence and conclusions,

the latter becomes a subject of “rational belief” [48]. The third view rests on the concept of frequency which emerged during the 19th century.

Variation, uncertainty, and metrology

It seems possible that the three different views turn out as different facets of the same gem. Probability is a rational concept to deal with doubt. If the doubt can be quantified on a rational basis, better (whatever the criterion for “better” may be) decisions can be made. If the likely outcomes of an experiment are pondered, the information available on the experiment and the conditions under which the experiment is performed play a role in assessing, e.g., the risks of the experiment. The fact that circumstances, even at the highest level of control, cannot be predictable to an arbitrary degree, gives weight to the aspects of variation and associated uncertainty. What is the likely outcome if the experiment is repeated? If it is repeated under different circumstances? What can be concluded if many independent outcomes from an experiment are available, i.e. from round-robin tests? It is common to search for answers to these questions with statistics. The Gaussian distribution (Eq. 3) is the almost undisputed (and unchecked) model for data variance [49–51]. It is, however, also evident that the data interpreted by metrologists, the data forwarded by metrologists, and the quality assurance of the data forwarded by metrologists, are not Gaussian or, at least, should not be expected a priori to be a sample from a Gaussian distribution. Metrologists care about the quality of data which are measured. The uncritical interpretation of measurement values in terms of Gaussian distributions (which is already implied by the evaluation of a measurement uncertainty, u_c , which is symmetric with the reported measurement value, and the use of $k = 2$ as expansion factor) does have its own risks. It is important to understand the rationale in the use of the Gaussian distribution and its limitations in order not to compromise the metrological intentions, e.g. to warrant traceability. There are alternatives. Simulation, e.g. by Monte Carlo methods, is one possibility, rarely encountered in metrological approaches in chemical measurement. Another possibility is the use of robust estimators [52] or non-parametric statistics.

Conclusions

Reporting measurement results without a statement about their likely variability is an anachronism. For some time the complexity of chemical measurements could be used as an excuse. Nowadays, the computer is a valuable tool in

generating uncertainty estimates. How to do that (the procedures) must be a field of active research when taking a metrological approach to chemical measurement. In practice the willingness to request and to provide a statement of measurement uncertainty by the stakeholders in the field is currently rather low. This reluctance may be explained by inertia [53]. From observation during the past 15 years of performing research on statistical/metrological properties of complex chemical measurements [54] it must be concluded that the fear of having to deal with details of statistics is at least an additional obstacle.

This very brief summary intends to provide a “golden thread” through the development of probability theory and statistics. The reader is encouraged to use these comments and the information provided by the references as a starting point for an independent journey through statistics. Statistics is an active field of research. The existence of a mathematical basis for the interpretation of probability (a basis for theoretical statistics) does not preclude the existence of three different interpretations of probability. It should be emphasized that the concept of probability, in its beginnings, was purely heuristic. The latin word “probare” means “to try” and therefore carries an essentially empirical notion. Thus it does not offer a consistent axiomatic building free of contradictions. The tools predominantly used by chemists are founded in the theory of errors established during nineteenth century. To appreciate the differences between the various schools, the reader may study the controversy between the Fisherian and the Pearson/Neyman approach to statistical inference as an example. In general, study of the original papers is highly encouraged.

Further papers will deal with the Gaussian distribution and its properties and with least-squares regression, correlation, and robust statistics. With a basic understanding of the rationale behind expansion factors and standard deviations, the discussion of practical aspects, e.g. the combination of measurement values obtained from proficiency tests, will hopefully become more efficient.

References

1. Thompson W (1883) In: Popular lectures and addresses. Macmillan and Co, London
2. Thompson M (1993) *Analyst* 119:127N
3. Efron B (1983) *Am Stat* 37:36–48
4. Huber W (2004) *Accred Qual Assur* 9:726
5. Hibbert DB (2005) *Accred Qual Assur* 10:300–301
6. Ellison SLR (2006) *Accred Qual Assur* 11:146–152
7. Gonzales AG, Herrador MA, Asuero AG, Sayago (2006) *Accred Qual Assur* 11:256–258
8. Cicero MT (83 BC) *De Inventione*. <http://www.volutabrum.de>. Accessed July 2007
9. Schüßler R (2003) *Moral im Zweifel* Mentis Paderborn/FRG

10. Gigerenzer G, Swijtink Z, Porter Th, Daston L, Beatty J, Krüger L (1989) *The Empire of Chance*. Cambridge University Press, Cambridge
11. Huygens CH (1657) In: Hald A (ed) (1990) *A history of probability and statistics and their applications before 1750*. Wiley, NY
12. Graunt J (1662) Natural and political observations mentioned in a following index and made upon the bills of mortality (available from <http://www.ac.wvu.edu/~stephan/Graunt/bills.html>). Accessed June 2007
13. van der Waerden BL (1975) *Die Werke von Jacob Bernoulli Vol 3*. Birkhäuser Verlag, Basel/CH
14. Galilei G (1638) *Discorsi et demonstrationi matematiche intorno a due nuove scienze attenenti alle mechanic & i movimenti locali*. Elsevier Leyden/NL
15. Bayes T (1764) *Phil Trans Roy Soc London* 53: 370–418, (<http://www.stat.ucla.edu/history/essay.pdf>). Accessed June 2007
16. Box GEP, Tiao GC (1973) *Bayesian inference in statistical analysis*. Wiley, NY
17. Königsberger E, Gamsjäger H (1990) *Monatshefte Chemie* 121:119–127
18. Olbricht W, Chatterjee N, Miller K (1994) *Phys Chem Minerals* 21:36–49
19. Laplace PS (1799) *Traité de mécanique céleste*, Rééditions Jacques Gabay (1990) Paris/F
20. Laplace PS (1812) *Théorie analytique des probabilités*, Tome VII des œuvres complètes (Paris-3e édition-1820). Réédition Jacques Gabay (1995) Paris/F
21. Gauss CF (1809) *Theoria motus corporum celestium*. Perthers, Hamburg/FRG
22. Guedj D (2000) *Le mètre du monde*. Editions de Seuil, Paris/F
23. Delambre JBJ (1808–1810) *Base du système métrique décimal, ou mesure de l'arc du méridien compris entre les parallèles de Dunkerque et Barcelone, exécutée en 1792 et années suivantes*, par MM Méchain et Delambre Baudouin Paris/F
24. MacKenzie D (1981) *Statistics in Britain 1865–1930: the social construction of scientific knowledge*. Edinburgh University Press, Edinburgh/UK
25. Galton F (1869) *Hereditary genius*. Macmillan, London (<http://www.galton.org/books/hereditary-genius/>). Accessed June 2007
26. Pearson K (1896) *Phil Trans Roy Soc A* 187:253–318
27. Pearson K (1900) *Phil Mag* 50:157–175
28. Pearson K (1894) *Phil Trans Roy Soc A* 185:71–110
29. <http://www.digital.library.adelaide.edu.au/coll/special/fisher/index.html>. Accessed June 2007
30. Neyman J, Pearson ES (1928) *Biometrika* 20:174–240
31. Neyman J, Pearson ES (1928) *Biometrika* 20:263–294
32. Neyman J, Pearson E (1933) *Phil Trans Roy Soc A* 231:289–337
33. Fisher RA (1926) “The arrangement of field experiments”. *J Ministry Agric Great Britain* 33:503–513
34. Salsburg DS (1985) *Am Stat* 39:220–223
35. Fisher RA (1925) *Statistical methods for research workers*. Oliver & Boyd, Edinburgh (<http://www.psychclassics.yorku.ca/Fisher/Methods/>). Accessed June 2007
36. Fisher RA (1971) *The design of experiments*. 9th edn. Macmillan, NY
37. Fisher RA (1951) *Statistical methods and scientific inference*. Oliver & Boyd, Edinburgh
38. Hald A (1998) *A history of mathematical statistics*. Macmillan, New York
39. Meinrath G, Ekberg C, Landgren A, Liljenzin O (2000) *Talanta* 51:231–246
40. Kolmogorov NA (1933) *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin
41. de Moivre A (1718) *The doctrine of chances*. Millar, London
42. Bertrand J (1888) *Calcul des probabilités*. Gauthier-Villars, Paris
43. Vovk V, Shafer G (2003) *The game-theoretic probability and finance project working paper #6*. (<http://www.probabilityandfinance.com>). Accessed Jan 2007
44. “Student” (Gosset W) (1908) *Biometrika* 6:1–25
45. Efron B (1986) *Am Stat* 40:1–11
46. Jaynes ET (2003) *Probability theory*. Cambridge University Press, Cambridge
47. Heisenberg W (1976) *Physik und Philosophie*. Ullstein, Frankfurt
48. Nagel E (1939) *Principles of theory and probability*. International encyclopedia of unified science, vol 1. Chicago
49. Desimoni E, Mannino S, Brunetti B (2001) *Accred Qual Assur* 11:452–458
50. Christensen JM, Holst E, Olsen E, Wilrich PT (2002) *Accred Qual Assur* 7:28–35
51. Araujo P, Froyland L (2005) *Accred Qual Assur* 10:185–189
52. Wilrich PT (2007) *Accred Qual Assur* 12:231–240
53. Bich W (2006) *Accred Qual Assur* 11:638–639
54. Meinrath G, Schneider P (2007) *Quality assurance for chemistry and environmental science*, Springer, Berlin, 326 p