Probability And Statistics Ideas In The Classroom – Lessons From History

D.R. Bellhouse
Department of Statistical and Actuarial Sciences
University of Western Ontario
London, Ontario,Canada,N6A 5B7

Abstract

When examining how the history of probability and statistics can be useful in the classroom, it is first useful to examine the styles in which the history of the subjects is written. These styles may be divided into internalist (those working within the area) and externalist (those outside the area) approaches. It is natural for teachers of probability and statistics to follow an internalist approach for classroom discussion. In order to discover what principles apply in transferring the lessons of history to the classroom, the work of William Sealy Gosset (Student) is discussed as a case study. What follows from this case study is that the most important historical lesson to convey is the motivation for an individual's work. This lesson is illustrated further in discussions of the solution to the problem of points or division of stakes and of the Fisher-Neyman dispute over their approaches to statistical inference.

1. Introduction

Almost any introductory statistics textbook is a compendium of the history of elementary probability since the Middle Ages and statistical methods since the seventeenth century. Of course, more modern developments obtained throughout the twentieth century are also included in these texts. Dicing probabilities, sometimes given as problems to solve in these introductory texts, first appeared in a manuscript written in the thirteenth century. Kolmogorov's axioms of probability from the 1930s are usually given as the basic rules of probability. The now standard technique of inference about the mean when the variance is unknown follows from results that were first obtained in the early twentieth century. With some exceptions, the methods and techniques are given without the historical references or context. Instead, the focus is on "relevant" modern applications of the material presented.

How should one approach history in the classroom? One temptation is to apply historical examples directly to the material. One could solve the dicing problems of the thirteenth century in class or present the original data that were used to demonstrate inferences about the mean when the variance is unknown. The problem is that many students are not interested in historical examples, which by their very nature are outdated – they want something more "relevant". Another approach is to give biographical sketches of some leading probabilists and statisticians. The straight biographical approach can be sterile unless the information that is presented is relevant to the more technical material given in the textbook.

A statistics textbook, even an elementary one, is a summary of knowledge, new and old, about the subject. I would put forward the view that in using history in probability and statistics the important questions to address are: how and why was this new knowledge created? There are a number of other questions that follow from this first one. When new knowledge is created, is there a clash between the old and the new knowledge? What is the nature of the clash? What happens when two strands of new knowledge compete for prominence? What is the social background of the new knowledge creator and what is its relevance to the knowledge created? In answering these questions, we often discover the motivation for the development of a new statistical technique, which deepens our understanding of it.

Much has been written on the history of probability and statistics. Before trying to decide what part of this history is useful in the teaching probability and statistics, it is helpful to look at the approaches to history that have been taken by the historians of probability and statistics over the past 140 years or so.

2. Historical Models for the History of Probability and Statistics

Historians of probability and statistics might be divided into two groups: internalists and externalists. Internalists are those who were trained in the subject, like myself; and externalists are those whose formal training comes from outside probability and statistics. Each can bring important insights to the history. Separately, each provides an incomplete picture of the history. Internalists are highly knowledgeable in the technical aspects of the subject and externalist have much greater knowledge of the social, economic and political forces that may have impact on the subject.

For historians of probability the standard early work is Todhunter (1865) who provided a list history devoted entirely to probability theory. All the major results of probability theory to the time of Laplace are listed and described in some mathematical detail. Todhunter's work is a major secondary source for early history of probability. Other list histories from the nineteenth century have been more general, massive tomes devoted to broad areas of mathematics while describing results in probability very briefly. These include, for example, Libri (1838) and Cantor (1880 – 1908). The second volume of Cantor's four-volume work lists some of the early results in probability that do not appear in Todhunter's work; it has become the second major secondary source for material in the history of probability, used most recently by Hald (1990).
There are similarities among all the analyses done by these nineteenth century historians. Their common approach to the history of probability comes from the fact that probability is a branch of mathematics and that the dominant philosophy of mathematics is Platonism or Neoplatonism.

Hersh (1997) has described three basic schools of the philosophy of mathematics, including the commonly held philosophy of Platonism, or Neoplatonism. Hersh calls the other two schools formalism and constructivism. The Neoplatonic school views all mathematical objects or results as eternally existing. Some of these objects have been discovered already, but the infinite remainder is yet to be discovered. A mathematician's approach to the history of mathematics, and consequently a probabilist's approach to the history of probability, has often been to answer the questions of who discovered what and when and who had priority for the discovery of a mathematical result. The natural way to write this history is to produce a list history. A history of probability such as Todhunter (1965) was directly influenced by this philosophy.

While probability can definitely be viewed as a branch of mathematics, there can be some debate about whether statistics can be similarly viewed. Most early and mid-nineteenth century statisticians in the Statistical Society of London and the American Statistical Association were numerate but not very mathematical. All would probably agree that today statistics is a discipline with a considerable amount of mathematical activity in it. In that vein, similar list histories to what occurred in probability and other branches of mathematics were produced describing statistical activity. Koren (1918), which is a collection of articles on the history of statistics in various countries, is one such example. It should be noted that Koren (1918) is a history of statistics written in the nineteenth century common connotation of the word. It is a description of data collection in various the states rather than a history of the development of statistical methods.

Hersh (1997) has rejected the three philosophies of mathematics as unsatisfactory in describing mathematical activity and has put forward in the preface to his book what he calls a humanist approach in which,

> "… mathematics must be understood as a human activity, a social phenomenon, part of human culture, historically evolved, and intelligible only in a social context."

Hersh's position is not new and may be compared, for example, to Karl Pearson's lecture notes on the history of statistics given at University College, London during the 1920s and 30s. Pearson began to move away from the list history approach, stating (Pearson, 1978):

> "… it is impossible to understand a man's work unless you understand something of his character and unless you understand something of his environment. And his environment means the state of affairs social and political of his own age."

F.N. David was Karl Pearson's research assistant in the 1930s and probably attended Pearson's lectures on the history of statistics. No doubt it was Pearson's philosophy that inspired her to deviate from the list history approach. Her book (David, 1962) on the early history of probability, *Gods, Games and Gambling*, contains biographical material and historical background, as well as technical analyses. Stigler (1986) in his *The History of Statistics* has taken this approach to its ultimate conclusion. As well as the biographical material, Stigler has provided a wealth of historical and scientific background so that the motivation is given in most cases for the technical developments that were achieved.

Internalist historians of probability and statistics have moved substantially in the direction of responding to the original question that I posed: how and why was new knowledge, particularly in probability and statistics, created? The how is the discovery of new tools and techniques and their influence on the subject's development. The why is the motivation to developing a new technique or result. Early historians such as Todhunter answered the how question listing what result was obtained, when, where and by whom.

In the past two or three decades, externalists, among them professional historians, sociologists and philosophers, have become interested in the history of probability and statistics. Their approach reflects their backgrounds and training. Typically the emphasis is on the social and political background to discover what social forces encouraged certain developments. Very little of this type of history deals with the technical development of the subject. For example, in the development of statistical methodology in Britain from Galton to Fisher, MacKenzie (1981), a sociologist traces this development in a non-technical way through the eugenics movement in Britain and its ties to the interests of the British professional middle class. Porter (2004) has described how Karl Pearson's intellectual pursuits eventually developed into work on fundamental problems in statistical analysis. In probability, Daston (1988), a historian of science, again in a non-technical way shows the connections between Enlightenment thought and the development of theory of probability and its application from its accepted initial development in the mid-seventeenth century through the mid-nineteenth century.

3. Some of Approaches to Using History in the Classroom

In view of the fact that we are statisticians, it is natural to follow what the internalists have done when looking to see how history can be used in the classroom. Mostly, I have followed the internalist approach in my own teaching and research work. It is easier for me, and for the students listening, to cover some technical detail and follow it with some relevant historical sidelight. Taking a note from the externalist approach, many years ago I once introduced a course that I taught on the mathematics of finance (interest calculations, annuities, etc.) with part of a lecture on usury, touching the religious and legal aspects of it since the Middle Ages. The greatest impact of this lecture was to increase my reputation for eccentricity among the students. In statistics courses I have used history in the classroom more positively in at least three ways: historical problems, historical personages and historical data.

The most typical use of historical problems in the classroom is through probability problems, and the most typical problem is the problem of the Chevalier de Méré: why does it pay to bet on seeing at least one six in four rolls of a single die and not on seeing at least one double six in twenty-four rolls of two dice? Some texts, Wild and Seber (2000) for example, give the problem as an exercise and then go on to give a brief historical description of the problem and how it led in part to the development of the probability calculus by Blaise Pascal and Pierre de Fermat. There are good and bad points in the use of this example. On the positive side it is a good exercise in a simple probability calculation and it introduces students to some historical characters. On the negative side, the problem as stated is a gambling problem and many students are either not interested in gambling or have a negative disposition to it. It can also give a wrong impression of the entire field of statistics if the course starts with probability and several calculations are made relating to dicing, cards and lotteries. Some textbooks, unnamed here, compound the problem by describing de Méré as a gambler and possibly an inveterate one at that. This actually may be historically inaccurate; Ore (1960) quotes one of de Méré's negative pronouncements on gambling. Although technically a gambling problem, de Méré's problem to him at the time may have been little more than an intellectual exercise set in a familiar courtly surrounding.

One of the increasingly popular personages to appear in biographical vignettes in statistics textbooks is Florence Nightingale. She is often the lone female in the gallery of male statisticians presented in these texts. One day in class I decided to describe another female statistician, a woman whom I had interviewed personally and believe to be the first woman to work professionally as a statistician in Canada, beginning in about 1940. She was the original quality control statistician at a company known as Northern Electric, now operating under the name Nortel. One mistake I made was to wait until near the end of class to talk about this remarkable woman. Now Canadians have this reputation of being polite people. It was not evident in this class. Binders began snapping and people began to leave while I was talking. One obvious lesson, of course, is not to present such things at the end of class. The other, subtler, lesson is that most students today are not interested in history. They want something that they think is immediately relevant to their studies, or more particularly to the exam, and to their future careers. "Think" is the operative word; the understanding of history can be highly relevant both to career and to study.

Whether using historical or modern data in the classroom, the same issue is present. Students respond most positively to any data presentation when the scientific background to the data is given and when some of the scientific points made in the introduction to the data are illustrated in the analysis. The issue in this case is not history and how to use it. Instead it is being familiar with the data, knowing the setting in which the data occur and being interested in the setting so that the instructor's enthusiasm for the problem is passed on to the student.

My own experience with using history in the classroom has been mixed. In learning from this experience, I believe that there are some underlying principles that would help to blend history into the classroom in a positive way. In order to discover this positive way, it is useful to look at a case study. In the next section I use William Sealy Gosset as my case study.

4. William Sealy Gosset: a Case Study

There are historical references, and especially to Gosset, in several introductory textbooks in probability and statistics. I examined not a random sample of these texts, but a dozen that happened to have recently crossed my desk. As expected, since statisticians wrote these books, all the uses of historical examples in them fall somewhere along the Todhunter – Pearson – David – Stigler spectrum.

Here is a brief biography of Gosset taken from the *Dictionary of National Biography*, written by Gosset's friend and associate E.S. Pearson (Pearson, 1996). Additional biographical information may be obtained from Pearson (1990). Gosset was born in 1876 and died in 1937. He studied at Oxford where he obtained a first class in mathematics in 1897 and another first class in chemistry in 1899. Shortly after graduation Gosset took a position at Guinness Breweries in Dublin where he eventually rose to the position of Chief Brewer. Soon after joining Guinness, Gosset found himself among a mass of data that had been collected relating to the whole brewing process from the cultivation of the ingredients to the finished product. In 1905, during a holiday in England, Gosset briefly met Karl Pearson so that he could discuss his statistical problems with Pearson. The following year, with Guinness's approval, Gosset went to London to work at Pearson's Biometric Laboratory for a couple of terms during that academic year. Gosset returned to Dublin where he was put in charge of the company's Experimental Brewery, a position that also put Gosset into contact with more data. Pearson had been highly impressed with Gosset and tried to convince him to take an academic position. By this time Gosset was married and had a child. His current salary at Guinness was £800 per year; the average academic salary for a professor at the time was £600 (*plus* *ça change* – only the amounts are different today). Gosset wrote his first paper while at Pearson's Biometric Laboratory. Guinness agreed to let Gosset publish his statistical research provided that he used a pseudonym (he used "A Student") and that none of the company's data appeared in the publication. The paper for which he is most famous was written the following year (Student, 1908). This is the paper in which the Student *t* distribution for small samples was obtained. Later Gosset corresponded with Fisher and maintained good relationships with both Fisher and Karl Pearson despite the animosity between the two.

It is of interest to see how the textbooks deal with Gosset and his statistical result. Some introductory textbooks contain no historical references to Gosset, or to anyone else (Mendenhall, Beaver and Beaver, 2003; Sanders and Smidt, 2000). Others make very few direct historical references and mention Gosset in passing when introducing the Student *t* distribution (Freund, 2001; Woodbury, 2002; McClave and Sinich, 2000; Wild and Seber, 2000). At the next level some texts contain historical vignettes of a few sentences, including one for Gosset, in sidebars or footnotes on an appropriate page. For Gosset the appropriate page is one by the discussion of the *t* distribution (Bluman, 2001; Moore and McCabe, 1998). Then the historical detail increases substantially. A number of texts provide biographies of various probabilists and statisticians, often at the beginning or end of a chapter (Johnson and Kuby, 2000; Moore, 2000; Weiss, 1999). At the extreme end of the scale Larsen and Marx (2001) give early histories, one of probability and one of statistics, at the beginning of the book. Then some biographical vignettes on a variety of probabilists and statisticians are given at the beginning of each chapter. Some of these more detailed biographies contain some additional information to what I have given. For example, Gosset married Marjory Surtees Phillpotts in 1906.

Many of the authors of the introductory texts discussed above have pointed out the requirement for Gosset to write under a pseudonym, often as an interesting historical tidbit without further explanation (as I also have done). What all but one of the biographies omit (and I have also purposely omitted it from my own vignette) is the motivation for Gosset's research into the *t* distribution. To me these two things are the most important pieces of information (the further explanation for the pseudonym and the motivation for the research) that a student could obtain from the whole biography.

The motivation for Gosset's research and subsequent discovery of the *t* distribution can be found in some of his published letters. In a letter dated May 12,1907 to a colleague at Guinness in Dublin, Gosset, who was at Pearson's Biometric Laboratory in London at the time, wrote of his working day:

> "… and on other days work at small numbers; a greater toil than I had expected, but I think absolutely necessary if the Brewery is to get all possible benefit from statistical processes."
>
> (in McMullen and Pearson, 1939)

This quotation shows that Gosset thought that small sample inference should be of interest to Guinness so that the 1908 paper, which makes no reference to the work Guinness Breweries or the brewery's data, was the direct motivation for Gosset's work. The letter does not show the full extent of Gosset's motivation for this work. This appears in a later letter dated September 15, 1915 to R.A. Fisher:

> "… and the Experimental Brewery which concerns such things as the connection between the analysis of malt or hops, and the behaviour of beer, and which takes a day to each unit of the experiment, thus limiting the numbers …"
>
> (in Pearson, 1990)

Obtaining a single observation took an entire day and was therefore an expensive thing to do. Often in most treatments of small sample inference in any textbook the cost factor is ignored and yet it remains the prime factor leading to small samples. Moore (2000), the one exception among the introductory textbooks, mentions the problem that field experiments run by Gosset resulted in small numbers of observations.

Karl Pearson was able to obtain scads of data for his Biometric Laboratory and so could never fully understand why Gosset concerned himself so much with small sample inference. Moreover, it was not until Fisher followed up on Gosset's research that small sample inference became more widely used beyond Guinness Breweries. This is one example of new knowledge not being adopted quickly after its discovery. The statistical needs of Karl Pearson, the leader in the field at that time, had not changed. Fisher's needs, based on designed experiments with relatively few observations compared to Pearson, were quite different.

Guinness's requirement for secrecy about data from the brewery and about what their employees were doing scientifically can seem foreign to an academic, although some academics are very secretive about their own work until it is in print. Consequently, the presentation of Guinness's secrecy requirement is usually made as a statement without further comment. Guinness's policy could easily be put into context today giving students insight into some of the statistical practices prevalent in industry. In today's world when academics are carrying out consulting work, there can be nondisclosure agreements attached to a consulting contract. My own experience with a brewery did not involve such an agreement, but it was certainly in the spirit of what Gosset agreed to. I was asked to give a one-day workshop on experimental design for employees in the scientific research section of Labatt Breweries. In order to make my presentation meaningful I asked for and received data from an experiment run as a 2^5 factorial design with two replicates. I wanted to demonstrate the use of a fractional factorial design and to compare the results to a full factorial; getting the most out of small samples is a continuing concern. With respect to secrecy, the catch was that I was not told what the response variable was other than it had something to do with the head on the beer. Nor was I told what each of the factors were. As I presented my analysis of the data some of the workshop members nodded in agreement over the statistical significance that I found for some factors (it made scientific sense to them – I was completely in the dark) and they could easily explain the presence of an unusual residual that I found. To this day, I have no idea what the data were, other than data on beer; Labatt Breweries wanted to keep their secrets a secret so that competitors would not have any idea of what they were doing.

5. Motivation, Motivation, Motivation

When it comes to real estate the standard phrase is "location, location, location." For the use of history for a class in probability and statistics, the length of the phrase is the same but the word is "motivation." To my mind the best use of history in class is to discover and describe what motivated people to work on various problems. And it will often turn out that what motivated our statistical forbears to come up with certain techniques or theory is the same as the motivation for using the results and techniques today. Likewise, when we examine the motivation behind certain probability problems, it turns out to be different from what we expect (the best gambling strategy, for example) and gives us further insight into these problems.

The strive to find motivation falls directly in line with what professional historians see as the reason for studying history. Stearns (2003) provides two major reasons for the study of history: (1) history helps us to understand people and societies; and (2) history helps us to understand change and how the society we live in came to be. Along the lines of the first reason, we understand better why we use a certain technique if we understand its original motivation for development. In concert with the second reason, we understand the need for further technical research or change in standard techniques used, if we know the motivation behind the change.

Within the mathematical sciences there are two distinct sources of motivation for new developments. The first we have seen through Gosset. His research was motivated through a practical problem. The second source of motivation is intellectual exercise; new results are added to an intellectual structure because the structure is there and the structure is interesting or intellectually challenging to the researcher.

I would put forward that the Pascal-Fermat problems in probability were actually more intellectual exercises than practical gambling problems. I have already alluded to the problem of the Chevalier de Méré as a probable intellectual exercise. The other problem that Pascal and Fermat worked on was known as the problem of points. This problem can be stated as: in a series of games or a tournament in which the final winner is the one to win in total a specified number of games, how should the stakes be divided if the series or tournament is concluded early? The problem first shows up in Italian arithmetic or abbaco books, the earliest in manuscript form in about 1400 prior to the invention of printing. Some famous Italian Renaissance mathematicians, Luca Pacioli, Girolamo Cardano and Niccolo Tartaglia all worked on this problem and all published incorrect solutions to it. One very important thing to note is that all attempted solutions to the problem of points were in abbaco books. This includes the French commercial arithmetic books, the likely source of the problem for Pascal and Fermat. The French books can be described essentially as technology transfer of the abbaco books from Italy. The Italian abbaco books were written typically as reference manuals for the teachers or for merchants rather than as texts for the students. Van Egmond (1981) has a description of these abbaco books and an extensive list of manuscript and printed abbaco books to 1600. The need for abbaco books began in the thirteenth century as the Italian city-states became increasingly involved in trade with the Arab world on the other side of the Mediterranean. The books typically contain treatments of the basic arithmetical operations of addition, subtraction, multiplication and division, as well as discussion of fractions and the extraction of square and cube roots. Many of these books go well beyond these basic arithmetic operations by including business problems, recreational mathematics problems, discussions of elementary geometry and algebra, and miscellaneous material such as calendars and astrology. Abbaco books contain no mathematical proofs, but rather are descriptions of mathematical problem solving techniques with many examples. When the problem of points is examined in the context of the genre of the books in which it appears it turns out, at least in Pacioli (1494), to be an application of, and a twist on, the standard business problem of the divisions of profits in a partnership given in the books. It is a practice calculation of a problem put in a slightly different way in order to maintain the interest of the reader. By the time that Pascal and Fermat attacked the problem, it had been an unsolved mathematical puzzle for over 250 years that had philosophical implications. As Daston (1988) has described the issue of probability calculations, beginning in the time of Pascal and Fermat, became one trying to answer to the question of how to behave rationally under uncertainty. This interpretation changes how the problem is presented in the classroom or as an exercise perhaps making it more palatable for those whose interests do not run to gambling.

Statistics is not purely a mathematical exercise. There are differences in the approach to statistical inference and differences in the interpretation of what the data say. The depth and significance of the differences can lead to heated exchanges between those who hold opposing opinions. And there is some beauty to these heated exchanges for teaching purposes. Many are fascinated by a good fight and so students' interests are piqued. More importantly, positions and issues are often made very clear by both sides so that the motivation behind an approach is also clarified.

Here is one such example. It begins with a quotation taken from a 1935 discussion to an article (Neyman, 1935) in *Journal of the Royal Statistical Society*.

> "Professor R.A. Fisher, in opening the discussion, said he had hoped that Dr. Neyman's paper would be on a subject with which the author was fully acquainted, and on which he could speak with authority, as in the case of his address to the Society delivered last summer. Since seeing the paper, he had come to the conclusion that Dr. Neyman had been somewhat unwise in his choice of topics."

Thus began a dispute between two giants of statistics that lasted until Fisher's death in 1962. When examined closely the dispute was about the nature of statistical inference. Neyman on the one side put forward confidence intervals as well as an approach to hypothesis testing with Egon Pearson that takes into account the errors in the possible decisions to be made. Fisher instead propounded fiducial intervals and the concept of p-values in significance testing. Neyman recognized the scientific value of the dispute. After twenty-five years of disagreement, Neyman (1961) wrote:

> "In general, scientific disputes are useful even if, at times, they are somewhat bitter. For example, the exchange of opinions and the studies surrounding the definition of probability by Richard von Mises, clarified the thinking considerably. On the one hand, this dispute brought out the superiority of Kolmogoroff's axiomization of the theory. On the other hand, the same dispute established firmly von Mises' philosophical outlook on 'frequentist' probability as a useful tool in indeterministic studies of phenomena. There are many similar examples in the history of science."

An excellent non-technical description of the issues in significance or hypothesis testing that have been clarified as a result of this heated (at least on Fisher's side) dispute can be found in Salsburg (2001).

The impact of Fisher's dispute with Neyman was felt for years after Fisher's death. Here is one small example that shows up in textbook writing. The first edition of Paul Hoel's *Introduction to Mathematical Statistics* (Hoel, 1947), for example, contains only the Neyman-Pearson approach to hypothesis testing. At the time the book was written, Hoel was at the University of California, Los Angeles. Neyman, who had a great influence over the development of statistics not only worldwide but also more locally, was at the University of California, Berkeley. On the other side of the textbook exposition of hypothesis testing is the Canadian statistician, Cyril Goulden whose first edition of a book on experimental design (Goulden, 1939) contains only a discussion of p-values. Goulden had intensively studied Fisher's work and had gone to England to spend some time learning directly from Fisher. North America tended strongly to follow the Neyman-Pearson approach to hypothesis testing. I saw this as a student in the 1960s and later as a young faculty member; the Neyman-Pearson approach was standard fare in all the introductory statistics textbooks that I saw during that time. It has only been in the last decade or so that the Fisherian approach with p-values began to appear in North Anerican statistics textbooks.

A look at my dozen non-randomly chosen textbooks shows the lingering effect of this dispute. The whole dozen now cover p-values, but in slightly different ways. Two-thirds of these books treat the Neyman-Pearson theory first in detail and then follow it with a briefer discussion of p-values. Sometimes the p-value discussion in these books appears to be an add-on. There is little or no discussion of the philosophical differences between the two approaches. Freund (2001), for example merely states that p-values are appropriate when you "cannot, or do not want to, specify a level of significance." The remaining four textbooks (Johnson and Kuby, 2000; Moore, 2000; Moore and McCabe, 1998; and Wild and Seber, 2000) begin with p-values and then move to the selection-rejection criteria under the Neyman-Pearson approach. Interestingly, Johnson and Kuby (2000) refer to the Neyman-Pearson approach as the "classical approach." This is a bit of a misnomer since Fisher's approach to hypothesis testing using p-values predates the Neyman-Pearson theory. Wild and Seber (2000) and Moore (2000), as well as Moore and McCabe (1998), are the best in describing the difference between the two approaches, the former referring to the Neyman-Pearson approach as hypothesis testing for decision making and the latter as tests with a fixed level of significance.

What are the differences between the two approaches since both approaches use the same test statistic? In the Neyman-Pearson setup a risk of error α is set so that if the null hypothesis is true, in repeated testing the hypothesis is rejected in error $100\alpha\%$ of the time. The size of the risk can be set in advance and related to the cost of making a wrong decision. The motivation here is decision making and controlling cost. The Fisherian approach comes out of the evaluation of scientific evidence and is a summary of the evidence against the null hypothesis. The motivation has changed to scientific inference. Fisher's procedure is a natural departure away from a method of deductive inference. A hypothesis implies data of a particular type or value. If the data obtained are in contradiction to the hypothesis, or could not have been obtained if the hypothesis were true, then we can conclude with certainty that the hypothesis is false. This is basic deductive logic using proof by contradiction. In the usual statistical setting, rather than contradictory data we may obtain data that are unusual or unlikely to have occurred if the hypothesis were true. The p-value is a measure of the "unlikeliness" of the data under the null hypothesis so that the p-value is an evaluation of the data rather than a formal decision. Now if we want to find the probability that the null hypothesis is true, then we need to take one more step and go the Bayesian route, which is another academic fight altogether.

What a little study of history has done so far is to put some of the textbook writing into context. It could also provide teachers in the classroom with a deeper discussion of the issues surrounding our approaches to hypothesis testing.

6. Conclusion

My conclusions are brief: motivation, motivation, motivation. History can be both useful and interesting in a classroom setting dealing with statistical methods. What the study of history does is to provide the motivation and context for methods that are covered in the class. By providing the motivation and context, it also provides deeper insight into the methodology covered.

REFERENCES

Bluman, A.B. (2001). *Elementary Statistics: A Step by Step Approach*. 4th ed. Boston: McGraw-Hill.

Cantor, M. (1880 – 1908). *Vorlesungen über Geschichte der Mathematik*. Reprinted by Johnson Reprint Corporation, New York, 1965.

Daston, L. (1988). *Classical Probability in the Enlightenment*. Princeton: Princeton University Press.

David, F.N. (1962). *Gods, Games and Gambling*. London: Griffin.

Freund, J.E. (2001). *Modern Elementary Statistics*. 10th ed. Upper Saddle River: Prentice Hall.

Goulden, C.H. (1937). *Methods of Statistical Analysis*. New York: Wiley.

Hald, A. (1990). *A History of Probability and Statistics and Their Applications before 1750*. Wiley: New York.

Hersh, Reuben (1997). *What is Mathematics, Really?* Oxford University Press: New York.

Johnson, R. and Kuby, P. (2000). *Elementary Statistics*. 8th ed. Pacific Grove: Duxbury.

Hoel, P.G. (1947). *Introduction to Mathematical Statistics*. New York: Wiley.

Koren, J. (1918). *The History of Statistics: Their Development and Progress in Many Countries*. Reprinted Burt Franklin, New York, 1970.

Larsen, R.J. and Marx, M.L. (2001). *An Introduction to Mathematical Statistics and Its Applications*. 3rd ed. Upper Saddle River: Prentice Hall.

Libri, G. (1838). *Histoire des Sciences Mathématiques en Italie depuis la renaissance de letters jusqu'à lad fin de 17e siècle*. Paris. Reprinted by Georg Olms Verlagsbuchhandlung, Hildesheim, 1967.

MacKenzie, D.A. (1981). *Statistics in Britain 1865 – 1930*. Edinburgh: Edinburgh University Press.

McClave, J.T. and Sinich, T. (2000). *Statistics.* 8th ed. Upper Saddle River: Prentice Hall.

McMullen, L. and Pearson, E.S. (1939). William Sealy Gosset, 1876 – 1937. *Biometrika* **30**: 205 – 250.

Mendenhall, W., Beaver, R.J. and Beaver, B.M. (2003). *Introduction to Probability and Statistics*. 11th ed. Pacific Grove: Brooks/Cole – Thomson.

Moore, D.S. (2000). *The Basic Practice of Statistics*. 2nd ed. New York: Freeman.

Moore, D.S. and McCabe, G.P. (1998). *Introduction to the Practice of Statistics*. 3rd ed. New York: Freeman.

Neyman, J. (1935). Statistical problems in agricultural experimentation. *Supplement to the Journal of the Royal Statistical Society* **2**: 107 – 180.

Neyman, J. (1961). Silver jubilee of my dispute with Fisher. *Journal of the Operations Research Society of Japan* **3**: 145 – 154.
Ore, O. (1960). Pascal and the invention of probability theory. *American Mathematical Monthly 67*: 409 – 419.
Pacioli, L. (1494). *Summa de arithmetica, geometria, proportioni et proportionalita*. Venezia.
Pearson, E.S. (1990). *'Student': A Statistical Biography of William Sealy Gosset*. R.L. Plackett and G.A. Barnard (eds.) Oxford: Clarendon Press.
Pearson, E.S. (1996). William Sealy Gosset. In *The Dictionary of National Biography 1931 –1940*. L.G. Wickham Legg (ed.). London: Oxford University Press, pp. 353 – 354.
Pearson, K. (1978). *The History of Statistics in the 17th and 18th Centuries against the changing background of intellectual, scientific and religious thought*, E.S. Pearson (editor). London: Charles Griffin and Company.
Porter, T.M. (2004). *Karl Pearson: the Scientific Life in a Statistical Age*. Princeton: Princeton University Press.
Salsburg, D. (2001). *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century*. New York: Freeman.
Sanders, D.H. and Smidt, R.K. (2000). *Statistics: A First Course.* 6th ed. Boston: McGraw-Hill.
Stearns, P.N. (2003). *Why Study History?* American Historical Association.
Stigler, S.M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, Massachusetts: Belknap Press.
Student (W.S. Gosset) (1908). The probable error of a mean. *Biometrika* **6**: 1 – 25.
Todhunter, I. (1865). *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace*. Cambridge University Press. Reprinted 1965, Chelsea Publishing, New York.
Van Egmond, W. (1981). *Practical Mathematics in the Italian Renaissance: a Catalog of Italian Abbacus Manuscripts and Printed Books to 1600*. Firenze: Editrice Giunti Barbèra.
Weiss, N.A. (1999). *Elementary Statistics*. 4th ed. Reading Massachusetts: Addison-Wesley.
Wild, C.J. and Seber, G.A.F. (2000). *Chance Encounters: A First Course in Data Analysis and Inference*. New York: Wiley.
Woodbury, G. (2002). *An Introduction to Statistics*. Pacific Grove: Duxbury.