

Mathematical Statistics: Fringe or Frontier?

R.D. Gill

1. AN OVERVIEW

1.1. Introduction

This essay contains a personal view of the place of mathematical statistics in mathematics and science, with an eye to the future. Statisticians are sometimes paid to make predictions, but they are trained to be careful and hedge their bets by giving an indication of its uncertainty. I shall be careful not to be specific, as far as the future is concerned. Most of the essay is concerned with the recent past (which surely contains the seeds of the future) and is built around some anecdotal case-studies of fascinating but in various senses paradoxical developments in the field.

59

1.2. What is mathematical statistics?

Statistics is concerned with analysing data, especially in the presence of randomness, whether due to measurement errors, deliberate sampling from a larger population, biological or behavioural variability, or whatever (more on this later). *Mathematical statistics* is the mathematical theory of how to do this. It codifies and organizes strategies for learning from data and for drawing conclusions about the real world phenomena which generated them. It is deeply connected with, and partly grown out of, probability theory, concerned with how to calculate probabilities of outcomes in random structures. Since mathematicians tend to become fascinated with the

abstract structure in the subject before them for its own sake, modern mathematical statistics contains many deep and beautiful mathematical results whose connection with real life data analysis can be tenuous. ‘Applications’ may only turn up later, or the theory may help in retrospect to explain why certain practical procedures work as well as they do.

1.3. Position within mathematics

As a relatively young branch of mathematics, uncertain of its independence, its practitioners often have ambivalent feelings about their relationship with mathematics proper. And real mathematicians may not always consider statistics as ‘within the fold’. A large wall poster still adorns the corridor of many a German university’s mathematics department, giving a well known publishing house’s schematic plan of the whole of mathematics. It took me some time to locate my own discipline in this grand scheme of things. In the middle of the picture were boxes labelled with names built of permutations of the words algebra (or algebraic), analysis (analytic), geometry (geometric), number, and theory. These boxes were connected with a dense network of arrows indicating connections in all directions.

As one moved away from the centre the connecting arrows became less dense, the names became more varied, less ‘pure’, less abstract. A distant corner contained the box ‘mathematical statistics’, connected by a single arrow *from* probability theory. This in its turn was only linked to the rest of mathematics by arrows from ‘measure theory’ and ‘potential theory’. Only after tracing back through many links did one arrive back at the centre of mathematics.

This picture is a caricature to be sure, but it reflects a common view of statistics held both by mathematicians and by other scientists. Statistics is a bit dirty and messy, a necessary evil perhaps; but from the point of view of those ‘at the heart’ it is a fringe event. Renowned mathematicians have occasionally called for abolishment of the whole discipline. The inventor of the Kalman filter—a piece of mathematics without which man would not have set foot on the moon—argued that ‘chance’ does not exist and therefore statistics is meaningless. Outsiders delight in the sometimes bitter controversies between different schools of statistics: subjectivists versus objectivists, exploratory data analysts versus decision theorists, and so on. So is statistics a marginal activity?

A rather different picture is given by a table published in a book by N.J. Higham [1]. The table gives the six most often cited papers in mathematics and computer science, statistics included. Four of the six are actually papers in statistics, published in more or less theoretical journals (Journal of the American Statistical Association, Journal of the Royal Statistical Society); the other two are papers on numerical mathematics. (One of these two—introducing the fast Fourier transform—by an author who later went on

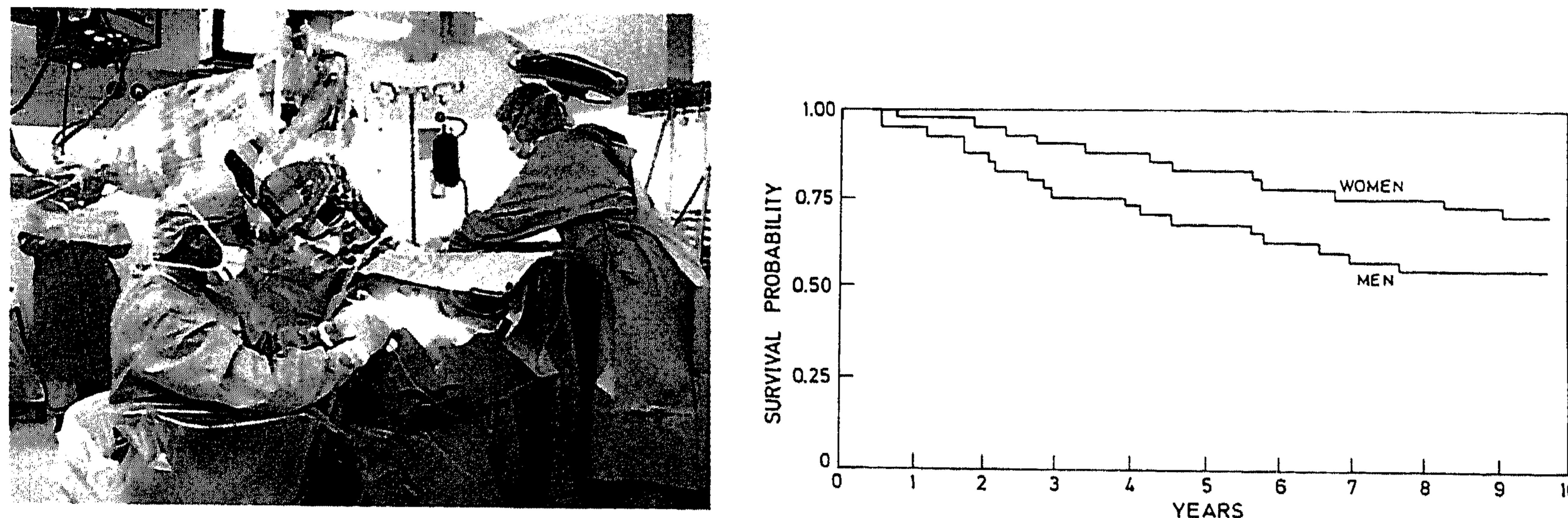


Figure 1. An important aspect of medical treatment of severe diseases like cancer is to observe survival times after treatment and draw statistical inferences from these observations. The curves drawn here (Kaplan-Meier estimates), show the survival probabilities after operation for malignant melanoma for 205 patients, stratified by sex. (Photo: courtesy Academisch Medisch Centrum Amsterdam.)

to be the founder of exploratory data analysis). Perhaps I should mention that though the cited papers were taken from mathematics and computer science, the ‘citers’ could have been in any discipline.

1.4. Survival analysis

Let me focus on two of these most cited papers [2, 3]. The second most often cited paper is by E.L. Kaplan and P. Meier (1958) and number four, by D.R. Cox (1972); both are concerned with survival analysis: the branch of statistics devoted to the special problems of analysing life-times, times till events, for instance the length of the disease-free period after cancer treatment in the life of cancer patients. These papers proposed new statistical techniques which became part of the standard repertoire of a huge army of cancer researchers.

But it was not just medical researchers who used these results and cited them alongside standard laboratory methods, as is their tradition. The papers just mentioned are cited enormously often by mathematical statisticians. The new techniques, developed to take account of a rather common feature of survival data, namely that many observations are *censored* (in other words, only known to exceed some value determined by the closing date of the study), relied on flashes of insight on the part of their inventors which could not be supported by then available mathematics. The special features of this kind of medical statistics has been an inspiration and a challenge to mathematicians since those key papers. Even now remarkable (for the insiders: amazing) mathematical properties are being discovered about the Kaplan-Meier survival curve estimator though it has become such a commonplace item in the statistician’s toolbox that its picture has been seen on the front page of major Dutch newspapers. The work of the mathe-

maticians has also been directly used by practitioners, and survival analysis is now a large established and rich area of statistics.

The development of survival analysis did not rely solely on the then available resources of mathematical statistics. A rather abstract, recently developed and then still incomplete part of probability theory came to play a major role in understanding and developing the analysis of censored survival data: namely continuous time martingale theory and the theory of stochastic integration. I won't start to explain these terms but let me emphasize that this was pure, pure mathematics, initially as unintelligible to mathematical statisticians (let alone, applied statisticians) as it will be to most readers now.

1.5. Applications and new theory

Those landmark contributions in survival analysis also contained the seeds of major new developments in statistics going far beyond the original medical setting. The theory and practice of 'semiparametric models' had its roots there. This way of doing statistics was taken up by applied researchers in econometric modelling and in psychometrics; it has fueled intense theoretical research on building a large-sample theory for statistical inference in infinite-dimensional sample spaces.

This story shows that from an unsuspected corner of applied statistics an impulse can come which leads to redrawing the map of theoretical statistics and revitalising its connections with other areas of mathematics. Is that a

once-off event? By now ancient history (the success story of the long-gone seventies and eighties)?

The answer is no. More recently—some would say, this is the success story of the present decade—financial mathematics developed deep connections with probability theory; indeed, the part of probability theory I just mentioned, the Itô stochastic calculus. The famous Black-Scholes paper [4] on how to price options was not only part of a new financial business but also part of the discovery of new deep results in probability theory. (Actually one may wonder how much recent financial catastrophies have been caused by use or mis-use of these fundamental advances.) It is easy to list many



Figure 2. The day after Black Friday, October 13, 1989. (Photo: N. Tully-Sigma. Courtesy ABC Press.)

other areas where new developments in science, technology and society have catalysed (and been effected by) revitalising activity in statistics and probability.

I would like to concentrate on two such stories. In each case the paradoxical fact is that the initial event was inimical to classical statistical theory (the abolishment crowd would lick their lips). It says something for the vitality and intrinsic need for a science of how to analyse random data that in neither case was this the end of statistics.

2. FIRST STORY: THE BOOTSTRAP

In 1979, B. Efron introduced into statistics his new ‘bootstrap method’ [5]. The essential idea of the method is to use computer simulation (in fact, a Monte-Carlo experiment, using computer generated randomness) to evaluate the accuracy of a statistical estimation procedure. ‘The real world’, from which random data has been obtained in order to learn about it, is replaced by an artificial ‘bootstrap world’ on the computer, totally under the control of the experimenter. In the bootstrap world random samples are repeatedly drawn; the variability in the estimate which is repeatedly evaluated from each new artificial dataset (an estimate of a known bootstrap world quantity) is a guide to the variability of the statistician’s actual estimate which she calculated from the actually available real data in the real world.

There is a little snag in this description: the bootstrap world has to be a faithful copy of the real world in order to make the simulation experiment appropriate, but the real world is not known: to find out about it was precisely the whole purpose of the exercise. No matter, we are statisticians, so we just use our data to estimate it. Efron’s audacious proposal was to do this in the most primitive way available: simply use the data points as they are as if only these values, and in precisely these proportions (each value equally likely), existed. The method then reduces to taking a sample of the same size as the original data-set *from* the original dataset (a random sample ‘with replacement’, so any particular data value can reappear a number of times in the new sample), recomputing the statistic of interest, repeating this procedure time and time again, and extrapolating the observed variability in the outcomes to the real world.

The method caught on like wild-fire. As the ambitions and sophistication of statisticians and the speed of their computers had increased, more and more complicated things were being done with data, and it was becoming less and less easy to use traditional means (analytic calculations in special models, or large-sample approximations) to judge the reliability of the results. Now all one had to do was leave the computer to repeat the original calculation of interest a thousand times on easily made artificial data sets, and you are done. One is then also liberated from only using methods for which explicit analytic calculations are feasible. Efron’s grand

claim was that his method abolished the need for mathematical analysis in statistics, replacing it by brute force computing power, and freeing the users of statistics from the traditionally available recipes, which were becoming a straight-jacket. Now anyone can creatively decide what he or she wants to do with the data, and leave the bootstrap to furnish the standard errors, confidence intervals, significance levels or whatever.

The method was so audacious that many people a little further from statistics refused to take it seriously. How could you get more information from data by throwing in extra random variation generated by yourself on your own computer? A senior research manager at the mathematics section of a well known Dutch multinational which at the time did not employ statisticians once told me that the reason for this was that nothing new had happened in statistics for the last forty years. I mentioned the bootstrap as a counter-example: but that was such a stupid idea it only confirmed his belief. Now the same company has belatedly caught on to the fact that a lot has happened, a lot of great value in an industrial research environment, and is rapidly building up its own sizeable internal statistical consultation unit.

Interestingly the bootstrap revolution did not put mathematical statisticians out of work after all. Along with practical success stories, came case-studies where the bootstrap gave stupid answers. In any case, why should it work at all? How well does it work? Can one make it work better? An explosion of activity took place with all kinds of variants being proposed of the original easy to understand methodology, in order to make the method more reliable, more flexible, more accurate, less computer-intensive (!), and so on. Some of the mathematical tools needed to really understand why or how the bootstrap works turned out to be linked to the traditional central activities of pure probabilists, nowadays somewhat looked down on: higher order corrections to the famous central limit theorem, which says that sample averages are approximately normally distributed. Other mathematical tools were connected to fundamental advances in pure mathematics, connected to the very abstract topic of 'probability in Banach spaces'. Yet other tools were needed from the theory of asymptotic statistics.

In a sense the bootstrap liberated applied statisticians from making tedious analytic calculations but it required a deeper and more creative level of mathematical activity, namely to understand and categorise the fundamental structure of diverse statistical methods, and their relation to fundamental probabilistic limit theorems.

3. SECOND STORY: QUANTUM STATISTICS

My other story is a story perhaps just starting, namely a new involvement of statistics in quantum physics. Let me begin this story on what may seem a philosophical issue, namely the question of whether randomness actually

exists. This is indeed a debatable point! A tossed coin or dice, for probability theory the archetypal random experiment, is just a simple dynamical system. Small differences in the initial vertical speed and angular momentum of the coin or dice are exponentially quickly magnified to large differences in its final position. Nothing random happens at all. The randomness in the initial condi-



Figure 3.

tions is presumably completely deterministically explainable in similar terms. The randomness of the bootstrap samples on the computer are also not random at all: a computer slavishly carries out its instructions; a random number generator is a sophisticated but completely deterministic way of ‘mixing up’ initial conditions so that what comes out looks random. (Interestingly enough, the modern theory of random number generation at the same time links fundamental ideas from statistics, from number theory—the heart of pure mathematics—and from the theory of computational complexity—the heart of theoretical computer science.) Now there is one place and I believe one place only where randomness really happens in the real world, and that is at the quantum level. Quantum physics describes the completely deterministic and continuous evolution (according to Schrödinger’s equation) of the state of quantum systems (systems of fundamental particles, photons, etc.). From the state at a given time can be calculated the *probability distribution* of the results of measurement of the system. To give some examples: a particular photon either does or does not pass a given polarization filter; an electron is either found or not found in a given region of space. Quantum theory tells us what the statistics would be like of many repetitions of these experiments: in a certain percentage of times, a photon passes the filter; in a certain percentage of times an electron is registered in a particular region of space. Radioactivity as measured by a Geiger-counter, showing a seemingly random series of time-points of emissions of individual particles, is another classic and nowadays even familiar example; an example where the randomness is not averaged out into the statistics of many particles but where it is still present at the macroscopic level.

Now a little thought shows up a huge paradox in the theory. A measurement device (e.g., a photo-multiplier set up to allow one to decide if a single photon does or does not go through a polarization filter) is itself just a large collection of elementary particles. The device together with the photon being measured together form a single quantum system, which develops deterministically and completely smoothly according to a huge com-

plicated Schrödinger's equation. Nothing actually 'happens', and certainly nothing happens or does not happen by chance. Similarly, a radioactive substance and a Geiger-counter together are also just one large assemblage of fundamental particles whose joint state is evolving deterministically and continuously according to the appropriate Schrödinger equation.

A traditional way out of the paradox is to suppose that things only 'happen' when a conscious observer looks at the system. Everything now becomes subjective, or circular, or an infinite regress threatens. What happens when an observer observes another observer? This is the famous Schrödinger's cat paradox, which hinges around the question of whether a cat, which is killed if and only if a certain radioactive decay takes place within a certain period of time, actually does die or not at the moment of the radioactive decay, or if this only happens when a human observer looks into the cage to see what has happened. This paradox forms part of R. Penrose's thesis (developed in his books 'The Emperor's New Mind' [6] and its sequel [7]) that human consciousness is essentially a quantum physical phenomenon and therefore artificial intelligence based on classical models of computing is impossible!

This state of affairs causes no difficulties in practice: the theory makes predictions which so far have agreed with all empirical findings. And theory which has consequences for laboratory experiments also has consequences for everyday technology. Technological advances are leading to experiments and experiments are leading to technological advances in which truly quantum phenomena have an impact on our everyday world. 'Quantum cryptography' is a practical way of transmitting messages safe from eavesdroppers which depends on paradoxical quantum phenomena (so-called entangled states) which have fascinated philosophers and visionaries and cranks for years; right now programs for quantum computers are being designed which for instance will factor large numbers in polynomial time by sheer brute force, simultaneously trying out all the possible factorizations, coded as a quantum superposition of states. Physicists predict that within five years the first real quantum computer will have been built; it will be able to successfully factor the number 'fifteen'. This is not a joke: the first digital computer was also not of much practical use. (When the first computer in The Netherlands was demonstrated to the Minister of Science, the program it ran was a program to produce a random number. That way the minister would not notice if the computer had actually worked properly or not!)

So nowadays quantum physicists are manipulating systems of a really small number of fundamental particles. These systems exhibit random behaviour of a truly fundamental character: this is randomness which cannot be explained by recourse to hidden but deterministic variation at a lower level (this is the content of 'no-go hidden variables theorems', connected to Bell's inequalities, the Einstein-Podolsky-Rosen paradox, and so on).

One would expect that a physical theory so closely involved with probability would be well-known and well-studied by probabilists and statisticians; and that these experts would even have a big contribution to make. Strangely that is not the case. Actually just before Kolmogorov in 1933 successfully axiomatized classical probability theory (the probability theory of dice and coin tosses, of insurance companies and casinos), John von Neumann axiomatized quantum probability. The mathematical structure was more general, and much more abstract looking than ordinary probability. Physicists (for instance R. Feynman in the famous Feynman lectures, see also [8]) and later on mathematicians have continually claimed that ‘quantum probability is a different kind of probability’. The field seems far from ordinary probability and statistics, and a huge barrier has been set up between them. The physicists perhaps did not really understand the mathematical modelling involved in ordinary probability; and pure mathematicians who were able to get a feeling for the mathematical structure of quantum mechanics—and who in recent years have developed an imposing theoretical structure called quantum probability theory—probably had so little feeling for physics that they accepted the statements of famous physicists like Feynman without questioning. The many paradoxes of the field anyway are enough to make any mathematician shy of saying anything about the practical side of the subject: he or she will just prove theorems in the abstract mathematical playground which physics provides.

In my opinion quantum probability is not a different kind of probability at all. My personal opinion is that quantum *reality* is a rather different kind of reality to ordinary reality (Einstein has said: reality is weirder than we imagine; weirder than we *can* imagine), the challenge is to classical deterministic thinking rather than to probability. Moreover the field is ripe for a new involvement of statistics. Already physicists are discussing ways of learning about the state of a quantum system from the (random) results of measurements which can be made on it. A couple of books and a fast growing number of papers exists on the topic (see, e.g., [9, 10, 11]). So far it is developing independently of modern statistics. Physicists are busy reinventing classical ideas from statistics: this is not a bad thing in itself; the bad thing is that they are unaware of the tremendous advances which that science has made in the last half century.

As long as both physicists and statisticians and pure mathematicians believe that quantum probability is ‘a different kind of probability’ this bad state of affairs will persist. But I think there are a lot of signs that this accepted wisdom is about to be thrown aside and the result will be a tremendous enriching both of mathematical statistics and of quantum technology.

4. CONCLUSION

The stories I have sketched above, and many others I could tell, show the young science of statistics vigorously growing at the interface between mathematics and society. New developments in technology and society immediately set huge challenges to applied and to theoretical statisticians: how to analyse data of growing complexity and how to answer the increasingly complex questions which society poses. These challenges reverberate into the heart of mathematics and sometimes answers are found using tools developed long ago in seemingly unrelated parts of mathematics, sometimes the challenges stimulate new fundamental advances.

It should be obvious now whether I think of statistics as a fringe or a frontier to mathematics. In my opinion it is part of the living frontier of mathematics; intensely alive; intensely unpredictable.

REFERENCES

1. N.J. HIGHAM (1993). *Handbook of Writing for the Mathematical Sciences*, SIAM, Philadelphia.
2. E.L. KAPLAN, P. MEIER (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, Volume 53, 457-481.
3. D.R. COX (1972). Regression models and life-tables. *Journal of the Royal Statistical Society*, Series B, Volume 34, 187-220.
4. F. BLACK, M. SCHOLES (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 639-654.
5. P. DIACONIS, B. EFRON (1983). Computer-intensive methods in statistics. *Scientific American*, 96-108.
6. R. PENROSE (1989). *The Emperor's New Mind*, Oxford University Press.
7. R. PENROSE (1994). *Shadows of the Mind*, Oxford University Press.
8. R.P. FEYNMAN (1951). The concept of probability in quantum mechanics. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 533-541.
9. J.D. MALLEY, J. HORNSTEIN (1993). Quantum statistical inference. *Statistical Science* 8, 433-457.
10. A. HOLEVO (1982). Probabilistic and statistical aspects of quantum theory. North-Holland, Amsterdam.
11. C.W. HELSTROM (1976). Quantum detection and estimation theory. Academic Press, New York.