Modern statistics: the myth and the magic

David J. Hand

*Imperial College London, UK*

*[The address of the President, delivered to The Royal Statistical Society on Wednesday, December 10th, 2008]*

**Summary.** The paper is a personal exploration of the puzzling contradiction between the fundamental excitement of statistics and its poor public image. It begins with the historical foundations and proceeds through the role of applications and the dramatic impact of the computer in shaping the discipline. The mismatch between the reality of statistics and its public perception arises from a number of dichotomies, some of which are explored. In particular, although statistics is perhaps typically seen as an impersonal discipline, in some sense it is very personal, and many of its applications are aimed at providing unique benefit to individuals. This benefit depends on the creation of detailed data sets describing individuals, but the contrary view is that this represents an invasion of privacy. Some observations on statistical education are made, and issues which will affect the future health of the discipline are examined.

Keywords: Applications; Computer; Education; Empirical models; Greater statistics; Privacy; Public perception

‘The Reader may here observe the Force of Numbers, which can be successfully applied, even to those things, which one would imagine are subject to no Rules. There are very few things which we know, which are not capable of being reduc'd to a Mathematical Reasoning; and when they cannot, it's a sign our Knowledge of them is very small and confused.’

John Arbuthnott, 1692

1. Introduction

It is a great honour, and also a very considerable challenge, to be chosen to serve as President of our Society. Not only does one have an opportunity to steer things in the directions that one considers important, but one also has to represent the Society in many different situations. I think this means that all Presidents climb a steep learning curve when they assume office. Although this is probably true for the officers of *any* Learned Society and professional body, I think that it is particularly true for the officers of the Royal Statistical Society. The reason for this is simply that statistics is so ubiquitous. There is no aspect of modern life upon which statistics does not impinge. Sometimes this impact is obvious—as in the drivers of government policy or the development of new medicines. But often the role of statistics is latent, hidden in algorithms which steer the direction of modern life—in biometrics systems, in control mechanisms for complex machines, in transport planning, in crime detection, etc.

In addition to the extraordinary breadth of relevance of its subject-matter, the Royal Statistical Society, in particular, is also unusual in its historical depth. Relatively few Learned Societies have a history stretching back as far as our Society’s—next month brings us into the

*Address for correspondence:* David J. Hand, Department of Mathematics, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK.

E-mail: d.j.hand@imperial.ac.uk

© 2009 Royal Statistical Society

0964–1998/09/172287
175th-anniversary year. As we might expect, over this 175-year period the nature and role of both statistics and the Society have changed.

The Presidential address is an opportunity for a personal view, so it gives me an excuse to discourse on certain issues which are of interest to me. To do so, I shall begin by glancing back at the past, to see how statistics has developed, and this will enable me to put its current situation, and in particular how others view it, into perspective. The ‘myth’ in my title refers to this general public perception. The ‘magic’ refers to the reality of the modern discipline. From here I shall then try to identify some of the issues which are important in ensuring the future health of the discipline.

I thought that a good place to begin my talk would be to try to convey some of the ubiquity of statistics by describing just a few widely distinct areas in which statistics is fundamental. So here are just a few examples of particular issues in which, whether it has been explicitly stated or not, statistics serves a central role. I could have chosen from an infinite number of others.

(a) In the UK, there is, of course, an Office for National Statistics. All developed countries have such an office, monitoring the changing economic, social and cultural aspects of society so that effective government is possible. I note parenthetically that few other disciplines have their own national office: few other disciplines are so central to effective government.

(b) Awareness of the risks that are associated with climate change is of growing importance. But detecting the change required careful statistical analysis of data describing the past. Determining the future risk requires sophisticated statistical extrapolation and forecasting techniques.

(c) We are currently experiencing a bioinformatic revolution, in which new data capture technologies, such as microarrays, are providing a previously unheard of wealth of data, probing how biological organisms act and function. Inferring biological mechanisms from these data is a fundamentally statistical challenge. It promises a revolution in prevention and treatment of disease.

(d) Large modern retailers, such as supermarket chains, rely heavily on careful data analysis of massive data sets describing customer purchases, to decide how best to meet customers’ wishes and demands.

As John Tukey is reported as having said:

‘The best thing about being a statistician is that you get to play in everyone’s back yard’.

The reasons for the ubiquity of statistics will be well known to most people in this audience. Statistics is both the science of uncertainty and the technology of extracting information from data. Indeed, it is a great pity that the snappy phrase *information technology* has been appropriated by the computer scientists. The truth is that computer scientists are really *data* technologists, concerned with storing and manipulating data. But it is statisticians for whom the *raison d’être* is the extraction of meaning from data—whose job is to transform data into information. This relationship between statistics and computer science is something to which I shall return below.

2. Inside statistics

2.1. The foundation of statistics

There are many different definitions of statistics. Depending on the precise nature of the work that a statistician does, they tend to put different degrees of emphasis on terms such as data,
uncertainty, information, analysis, understanding, prediction, inference, summarization, and so on. The range of definitions arises from the diversity of kinds of statistical exercises, something which was also noted by David Bartholomew in his Presidential address (Bartholomew, 1995). In that address, he identified four different types of statistics, ranging from concern with collecting and presenting data to reveal their main features, through the formal inferential procedures which tend to be taught on statistics courses, via attempts to understand horrendously complex stochastic systems, to issues of interacting in higher level discussions about domains with large amounts of intrinsic uncertainty.

I think, however, that the diversity is also a consequence of how the discipline of statistics developed. Statistics as we now know it is the confluence of several strands of thought and has evolved through several phases. Each of these strands, and each of these phases, has left its mark, and each continues to be of interest.

The important strands include the following.

(a) *The theory of probability*: the beginnings of the development of a formal calculus of probability is often dated from around 1650, when mathematicians focused their attention on some issues arising in games of chance. Kolmogorov’s axioms, which were formulated around 1930, are now the basis of modern statistical treatments of probability, whichever interpretation one puts on that word.

(b) *Surveys of people for governmental and economic purposes, as well as the collection of administrative data*: indeed, the word ‘statistics’ derives from the Italian *stato*, meaning state, and with this meaning dates back at least as far as the 16th century. This work is related to the construction of life tables for insurance purposes.

(c) The third strand is *the development of arithmetic tools for coping with observational and measurement error*, especially in disciplines such as astronomy and mechanics.

### 2.2. The development of statistics

The notion of statistics as a primarily mathematical discipline really developed during the 20th century, perhaps up to around 1970, during which period the foundations of modern statistical inference were laid. However, during more recent decades the effect of the computer has been felt, so there is nowadays as much justification for regarding the discipline as a computational one as a mathematical one. Certainly, the computer has revolutionized statistics and its practice. The computer and the particular requirements of the different domains to which statistics has been applied have been the two primary factors in influencing the direction in which it has developed in the past few decades.

One of the consequences of the ubiquity of statistics is that it has a tendency to diversify and specialize into different application areas. In some sense, this is a good thing. Such growth and change characterize a dynamic discipline, evolving to fill new data analytic niches. A discipline which failed to change and adjust to new challenges would rapidly die. However, it also has consequences which may not be beneficial. Just as evolution produces new species, which cannot interbreed with their cousins, so data analytic evolution has generated subdisciplines such as econometrics, chemometrics and environmetrics. Not all practitioners of these disciplines will define themselves as statisticians, perhaps arguing that they have particular expertise that is not possessed by statisticians. To some extent that argument is always valid—the very breadth of statistics ensures that different practitioners, in different application areas, are likely to have rather different skills. It is a well-known fact that statisticians working within any particular application area do develop an understanding of and expertise within that area, which enables them to make significant, sometimes fundamental, contributions to it. But I think that this
has also served to fracture the idea that they have a common statistical core. I think that the discipline as a whole suffers when this common core is not acknowledged.

If, in some cases, subdisciplines have spun off from statistics, and also adopted ideas from other sources, in other cases statistical ideas have been rediscovered. Typically, of course, it then takes some time before it is recognized that the ideas are the same, merely with different names. Such duplication of research effort is unfortunate. It does not seem to be restricted to any particular application domain but occurs in both the natural and the social sciences, and in engineering and commercial applications. At the least it can lead to cases of reinvention of the wheel and failure to take advantage of advances. I believe that this is a consequence of lack of awareness of the common foundation, and that it arises from too narrow a view of what statistics is. The problem is illustrated by an examination of US patents and patent applications. Unlike in the UK and the European Union, in the USA it is possible to patent ‘business processes’, and statistical techniques often fall into this category. Such an examination shows many instances of apparent rediscovery of well-known and well-understood statistical techniques which have been around for years.

I have wondered whether the tendency to invent new names for statistical methods has gathered pace in recent years. At a higher level, in my work with the business community I have sometimes found it difficult to keep up with the fact that I was a ‘techie’, or a ‘quant’, or worked in ‘analytics’ or in ‘business intelligence’, when all I thought I was doing was statistics. Although these business terms might simply be part of the need for new packaging to attract attention (of senior management as much as of customers) I think that failure to recognize that it is really statistics is detrimental to our discipline. At the least, we do not benefit from the fact that it is our tools, methods and ideas which are driving a company to success. I should say that there are exceptions, and that sometimes the fact that it is statistics is explicitly acknowledged: as Sam Alkhalaf, of Mastercard, put it in an interview, ‘The actual magic comes from our statistical analysis team’ (Higgins, 2003).

I have already commented that the application of statistics to different areas has been one of the primary drivers behind the development of the discipline. We are probably all familiar with early agricultural work being the genesis of formal experimental design, of factor analysis and item response theory originating in the behavioural sciences, of survival analysis being developed for medical applications, of ordination and multidimensional scaling being driven by work in ecology and of survey sampling being developed in the social sciences, although, of course, sometimes parallel developments of similar ideas occur. In general, however, once a new tool has been developed for one application domain, it often readily finds applications in others. Experimental design is widely used in most quantitative areas nowadays, from medicine through manufacturing to social policy evaluation. Factor analysis is used in areas including chemistry, ecology, geology and finance: and so on.

This role of application domains being one of the prime motivators for the development of new statistical ideas is continuing—and there is no reason to expect it to stop in the future. New domains will present new problems, and new solutions will be required to those problems. Current examples of this are the ‘large p, small n’ problems in bioinformatics, and streaming data problems arising in many applications as a result of dynamic data capture. There are also new challenges arising from different kinds of data: text data, clickstream data, image data, metadata, and so on. The World Wide Web, in particular, has led to the development of many very new and exciting statistical methods, characterized by the large size of the data sets and their dynamic nature.

Of course, these are non-traditional statistical problems. But, if statistics is to continue to be a healthy discipline, it must continue to grow and develop, to cope with new kinds of data
analytic problems. Despite my earlier comments, I believe that it has not always been up to this challenge, and that that is another reason that its full contribution and potential have not been properly appreciated. Statisticians have often been hesitant about tackling new kinds of data, in new areas. In fact, this is the source of one caricature of the statistician: that she requires the assumption of a particular model (say, a normal distribution) and bases all the analysis on that assumption. Of course, we all know that this is untrue: that models are chosen because they reflect the structure of the data or a theory describing the genesis of the data, and not simply because they are mathematically attractive; indeed that there are whole swathes of statistical theory and methods concerned with non-parametric or essentially model-free methods. However, as with all good caricatures, there is an element of truth in it. For example, much statistical theory is based on the notion that the data have been randomly drawn from a population about which one wishes to make an inference. That is fine in carefully controlled experimental or survey situations, when the population on which the observations are made is well defined. But in observational situations it may not be realistic: there may be unknown and undetectable factors influencing the subject-matter and altering the probability that certain values will be observed. Dust clouds absorbing light from distant stars might make some objects less likely to be observed than others. Patients may not turn up for appointments because they felt unwell. If the dust clouds are more likely to occur in regions where new young stars are forming, and if the patients feel unwell because the treatment is not working, any simplistic conclusions that are based on the observed data could be highly misleading. It is for reasons such as this that statisticians, quite properly, have exercised caution.

Caution is good, up to a point. It decreases the chance of making errors and poor decisions. However, it also decreases the chance of making major discoveries and advances. Explorers who sailed across uncharted oceans on voyages of discovery took great risks and occasionally reaped great rewards. Perhaps I can illustrate this by continuing to caricature: in statistics the approach has been to understand a method before applying it, but in some other disciplines the reverse seems to have been the case. In the 1980s, for example, I followed the development of neural networks within the computer science community with interest. This class of tools had its origins in the 1960s, with work on the so-called ‘perceptron’—essentially a linear classifier, like linear discriminant analysis, but based on optimizing a different performance measure and placing emphasis on sequential algorithms. Advanced versions, so-called phi-machines, introduced non-linearities into the models, so producing what were really the first neural networks. However, progress was then delayed because of the computational requirements in fitting such models. But, as hardware capabilities advanced, so, in the 1980s, effective fitting and estimation algorithms were produced, and the area took off, accompanied by a real buzz of excitement and enthusiasm. There were also claims that neural networks could outclass statistical tools with the same objective. But such claims sometimes went too far: I can recall attending meetings where the proponents of these new tools described how they could solve discrimination problems to produce zero error rate. In general, for most real problems, this is impossible: to claim such a thing is to invite ridicule. The occurrence of such a result should have aroused suspicion. In fact it arises from ‘overfitting’, a phenomenon that is well understood by statisticians. In this case, the adventurous spirit and lack of caution led to mistakes.

Things have moved on since those days. The overfitting problem is well recognized within the neural network community. Standard, well-understood statistical procedures are used to estimate the parameters of neural networks. And a variety of other tools have been developed, by statisticians, with the same power as neural networks.

Perhaps the ideal would be to have a single discipline which was both cautious and adventurous, but maybe that is not possible—although it depends, of course, on what is meant by
a ‘discipline’. In this context, Chambers (1993) coined the useful phrase ‘greater statistics’. He defined greater statistics as

‘everything related to learning from data, from the first planning or collection to the last presentation or report’.

In contrast, he defined lesser statistics as

‘the body of specifically statistical methodology that has evolved within the profession—roughly, statistics as defined by texts, journals and doctoral dissertations’

and he went on to say

‘Greater statistics tends to be inclusive, eclectic with respect to methodology, closely associated with other disciplines, and practiced by many outside of academia and often outside professional statistics. Lesser statistics tends to be exclusive, oriented to mathematical techniques, less frequently collaborative with other disciplines, and primarily practiced by members of university departments of statistics.’

Leo Breiman drew attention to the limited domain of lesser statistics in his comment on a paper entitled ‘Statistical fraud detection: a review’ (Bolton and Hand, 2002). He wrote (Breiman, 2002):

‘The authors titled their paper “Statistical fraud detection,” implying that this area is within the realm of statistics—would that it were—but the number of statisticians involved is small. The authors write that they are covering a few areas “in which statistical methods can be applied.” The list of statistical methods that I extracted from the article are

- Neural nets
- Rule-based methods
- Tree-based algorithms
- Genetic algorithms
- Fuzzy logic
- Mixture models
- Bayesian networks
- Meta-learning

These were developed in machine learning, not statistics (with the exception of mixture models)....’

I am arguing, following Chambers’s ‘greater statistics’, that we should regard these as part of statistics.

I started this discussion of developments in statistical methodology by highlighting the role of applications. However, I have segued into the other factor which has provided a huge impact over recent decades. This is progress in computer technology. I think that none of the areas that are given in Breiman’s list would be possible without computers, and there are many other examples (even within ‘lesser statistics’) for which this is also so. Indeed, even at the most basic level of data analysis the effect of computer technology has been vast. To calculate a mean or standard deviation nowadays requires no arithmetic skill. The application of standard techniques, which were developed in the early days of statistics, merely requires an ability to request that a computer does the calculation—the press of a button. This, by itself, radically alters the way that statisticians function. No longer do they need to be concerned with minitiae, and obsessive accuracy, but instead they can take a higher level view: they can concern themselves with understanding, rather than with mechanical manipulation. This changes the complexion of the discipline. But, beyond its effect on the basic tools, the development of computers has led to entirely new classes of tools, of which those mentioned by Leo Breiman are examples, and which were literally inconceivable before modern computer power became available. In fact, it is this change which has led me to describe statistics as ‘the most exciting of disciplines’ because
‘today’s statisticians use these [advanced software] tools to probe data in the search for structures and patterns, [to] peel back the layers of mystification and obscurity, revealing the truths beneath’

(Hand, 2007).

One qualification is appropriate, however, especially in the context of the enthusiasm of the preceding paragraph. This is that being a statistician still requires numeracy in the sense of understanding the numbers. It still requires, for example, the ability to recognize when a data entry error has produced an absurd result. It still, perhaps even more than ever, requires a feel for data.

2.3. Empirical models

In the preceding sections I have attempted to describe the breadth of modern statistics, and I briefly noted the primary drivers in forming it. In this section, almost as a parenthetical remark, I want to focus on a particular development, which has attracted considerable attention in recent years, and which is not going to go away. Whereas enthusiasm for new classes of data analytic approaches tends to follow a rising curve, which peaks and declines as their limitations become understood, plateauing at some level when their range of applicability is known, the area of which I speak is continuing, and I predict will continue, to grow in importance. This is what has become known as ‘data mining’.

Data mining is not a new class of data analytic approaches but rather is a term that is used to describe the application of a collection of existing techniques, along with a limited number of intrinsically non-statistical techniques, to the analysis of large data sets:

‘data mining is the analysis of (often large) observational data sets to find unsuspected relationships and to summarise the data in novel ways that are both understandable and useful to the data owner’

(Hand et al., 2001). As such, it falls squarely within Chambers’s ‘greater statistics’. And the reason that I believe that interest in the area will not decline is simply that we are being presented with increasingly more and increasingly larger data sets. That is, the demand for the solutions that are provided by data mining is continuing and will continue to increase.

To some statisticians, it seems unnecessary to have coined a new phrase to describe the sort of activities which, they might argue, they already undertook. This is a perfectly reasonable perspective. However, it is clear from the attention that data mining has attracted that statisticians were not meeting all the needs. I think that this was partly a hesitancy on the part of statisticians to become involved in analysing data sets of dubious provenance: another example of their intrinsic caution. Many such large data sets are not in any sense ‘random samples’ and many are not entire populations. The risk of drawing misleading conclusions is large. There are also more general problems arising from poor quality data. Such problems are difficult enough to tackle at the best of times, but they are far more difficult when there are millions or even billions of observations, or perhaps tens of thousands of variables. Manual inspection is out of the question, so automated procedures are essential. I also think that another reason that data mining did not originate in the statisticians’ camp is that it found a natural home as a spin-off of database technology and expertise in storing and manipulating large quantities of data. If one wants to make a statement about the items that were sold by a supermarket last week, one processes the database describing last week’s purchases. And database technology resides in the camp of the computer scientist.

However, that last example is revealing—and shows why data mining arguably belongs more naturally within the purview of the statistician. Although it is true that one will sometimes want to ‘make a statement about the items that were sold by a supermarket last week’, one will also, perhaps even more often, want to make a statement about future sales, i.e. the observed data
are to be used as the basis for inference to the future, or to other circumstances. And inference is what statisticians do. The danger is that, if we let our natural and justifiable caution arising from concern about the quality of the data override our willingness to be involved, then we shall lose ground. As John Chambers put it

‘If statisticians remain aloof, others will act. Statistics will lose; in addition, I believe science and society will lose also, because the statistician’s mental attitude at its best provides qualities likely to be missing elsewhere.’

The size of many modern data sets leads to new problems: not simply low level logistical ones of how to examine each of a billion data points, but also more subtle ones concerning the nature of inference. For example, with a billion data points, minuscule effects will be statistically significant. However, a more radical development arises from the impetus that data mining has given to the development of ‘empirical’ models.

It is convenient to divide statistical models into two classes (Box and Hunter, 1965; Lehmann, 1990; Cox, 1990; Hand, 1985, 2008a) which go under various names. Here I shall call them ‘empirical’ and ‘iconic’. Iconic models are based on some theory concerning the underlying mechanism which has generated the data (‘iconic’ because they represent an image of how one believes reality works). Empirical models are devoid of causal or other kinds of theory but merely summarize the data in convenient ways. Clearly, empirical models will be particularly important in areas where theory is weak, or potentially vastly complex. Such areas are illustrated by the social sciences and commercial applications. For example, in trying to predict whether a customer will buy a particular item next week, one does not base one’s prediction on a set of differential equations (an iconic model), but rather on a (probably fairly simple) descriptive model (an empirical model) relating past purchases to the characteristics of the customers making them. Contrast this with trying to predict where a ballistic shell will fall. The distinction between empirical and iconic models is an old one, and one which is related to, but is not identical to, the distinction between data models and algorithmic models that was described by Breiman (2001). Most data mining work is based on empirical models.

One reason why empirical models have really come into their own with data mining is that very large data sets permit small features of the underlying data-generating process to be modelled. Whereas a small data set allows us to fit a linear regression model adequately, fitting a high order polynomial regression requires a larger data set—and fitting the higher order terms leads to greater predictive accuracy. (At the risk of demonstrating the cautious approach that is implied by my statistical background, I feel bound to remark that such apparent improved predictive power is not without its risks—see Hand (2006).)

Examples of such empirical models that are produced by data mining activities are the predictive models built by hedge funds, the extensive experimentation and model search carried out by retail financial services organizations such as credit card companies, identifying relevant genes and gene combinations in bioinformatics and the model fitting exercises of large supermarket chains.

3. External challenges

So far I have focused on introspection, looking at what statistics is, or perhaps at what I think it should be—the magic if you like. I would now like to complement that perspective by looking at some aspects of how statistics is perceived, the myth, because the way that it is perceived will have a major influence on its future. I shall do this in two parts. The first looks at the specific issue of the relationship between statistics and the individual, and the second at a wide range of issues which impact on the popular perception of statistics.
3.1. Statistics and the individual

A traditional, and entirely natural, perspective on statistics is that it is concerned with the mass of observations rather than the individual. This perspective is understandable because that is perhaps the aim of most statistical activity, and that is certainly how statistics is explained in introductory courses. David Bartholomew, in his Presidential address to the Society (Bartholomew, 1995), quoted Egon Pearson as introducing statistics as ‘the study of the collective characters of populations’ and Maurice Kendall as defining it as ‘the science of collectives and group properties’ (Kendall, 1950). Maurice Kendall, in his Presidential address (Kendall, 1961) quoted a passage defining the interests of the Statistical Section of the British Association, when it was established in 1832, as being concerned with

‘facts relating to communities of men which are capable of being expressed by numbers, and which promise when sufficiently multiplied to indicate general laws’.

However, there is a complementary perspective, and one which is becoming increasingly important in social contexts as a direct consequence of automatic data capture via modern electronics and computer systems. This is that statistics is about the identification of the unique, characterizing the behaviour and needs of the individual members of a population. In many ways this is an attractive notion, because it personalizes statistics; it brings home the fact that, although statistics does indeed enable us to make broad summary statements about large groups of people, it also enables the unique, idiosyncratic, and indeed special nature of people to be recognized and taken into account.

Let me give you some of the context for this perspective, at least as I want to develop it today. Not so long ago, to collect data about an individual it was necessary to speak to them, to correspond with them, or to observe them in some way, and then to transcribe the results to a written record. The effort that is involved in this activity limited the size of data sets. Nowadays, however, in many contexts data capture has become automated. Electronic devices note and record the details, requiring no human intervention, and furthermore they do this accurately, tirelessly and vastly quicker than could a human. Since computer storage facilities have grown at a comparable pace, often there is no need to reduce the amount of data stored—it can all be retained. A typical example is that of credit card transactions. Each such transaction leads to the automatic recording of between 70 and 80 items of information, and several billion such transactions are carried out annually in the UK. Some data capture is explicit and overt—such as examination scores, medical test results and data from application forms. We all know that this information is being recorded and we collude with its collection. Increasingly often, however, data capture is implicit and covert—such as the details of the items that we bought at a supermarket, the details of our credit card transactions, of our train and plane tickets and indeed journeys (think of Oyster cards—prepaid cards permitting travel on London’s public transport network, which are scanned at the start and finish of each journey), of the phone calls and e-mails that we made and received (perhaps not the content, but where, when and to whom), and radio frequency identification scans revealing not only our location but also the location of goods that we bought as we carry them home. In this general context, I note that the UK has the world’s highest per capita rate of closed-circuit television cameras. Increasingly, also, we are seeing the fusion of databases describing different aspects of behaviour. Overall, data miners say that we each cast a long data shadow.

In many situations, the primary aim of all this data capture, whether overt or covert, is immediate: to calculate how much to bill a customer, to print out a travel itinerary, to match a car with a congestion charge or a road toll, and so on. Increasingly, however, once the data have been recorded, they are stored for later analysis. And this is where the statistician comes in. The statistician can analyse the data to find the structures, patterns and relationships within them.
Summarizing the mass of records of the behaviour of people allows organizations (governments, corporations, educational institutions, etc.) to provide a better service. Knowing the proportion of the population within each age group allows a local authority to decide how best to allocate its resources to services. Knowing what sort of items people prefer to buy enables retailers to make sure that those items are available. Of course, there are economic restrictions on how well organizations can react in meeting requirements, as anyone who has been forced to stand because of overcrowding on a long train journey will know.

But we can turn this view on its head. If we have built a model summarizing the relationships between the variables in a population, then we can apply that model to the unique characteristics of an individual to make predictions about the specific behaviour, needs and wishes of that individual, and then choose the most appropriate action for that individual. We condense the information about a mass of individuals, anonymizing it in the process, and reducing it to a comparatively simple set of relationships (simple, at least compared with the possibly many billions of values that are stored in the database). Applying this model to what information we have about any particular individual (regardless of whether or not they were included in the data that were used for building the model) then allows us to make further statements about that individual, describing things which were not latent in the personal characteristics of the individual alone.

Here are just three examples.

(a) The aim of a clinical trial is not really to work out whether drug A is superior to drug B ‘on average’, but to enable a decision to be made about which drug to prescribe for the next patient who walks through the door, i.e. for the individual.

(b) The purpose of building a credit scoring model is not the modelling in itself, but rather the application of that model to the characteristics of an individual so that a statement can be made about the riskiness of that particular person.

(c) The analysis that is implicit in the Amazon book recommender system is not to achieve any kind of global statement about what books people buy, but rather it is to enable the prior behaviour of an (any) individual to be matched to books in which they, as individuals, are most likely to be interested.

Each of these examples has two components:

(i) a large mass of data which is reduced, by a statistical modelling process, to a relatively simple model;

(ii) the particular characteristics of an individual.

Putting the two together enables us to deduce further information about the individual. I cannot tell from a person’s medical records and tests alone whether they are likely to react well to a particular treatment. But I can tell this if I have analysed the characteristics and reaction of a large mass of people who have previously been exposed to the treatment.

Although the modelling process depersonalizes and anonymizes the individuals whose behaviour is modelled, that is not the aim of the exercise. The aim is to be able to make personal and particular statements about individuals, and the analysis enables that. From this perspective, we see that it is wrong to regard statistics as a depersonalizing science, focusing attention on Quetelet’s *homme moyen*, average man. Rather, it is the science of the individual and the specific. From this perspective, what statistics does is enable knowledge of the many to be focused down to improve the lot of the one.

A rather interesting example of this way in which the traditional view of statistics is turned on its head is the following. As I write this talk, we are in the throes of a US Presidential election,
with all its attendant publicity. (In fact, now that I think of it, this seems often to be so.) What is not apparent from all this publicity is that electioneering, at least on a large scale, has changed dramatically over the past couple of decades—and it is the inversion of the view of statistics which has enabled this change.

Before and during the 19th century, an electoral candidate would distribute his manifesto and make public speeches. The speeches would be directly heard by just a few tens of listeners, although the content would be conveyed more broadly by newspapers. The key thing would be for the candidate to communicate to the electorate what he stood for, so that the voters could make an informed decision about which candidate they should support. Clearly I am taking an idealized perspective here, and glossing over any notion that the candidate might try to avoid an explicit statement of his position, to encourage more people to vote for him. In any case, as technology advanced into the 20th century, so first radio and then television changed this. Now politicians could speak directly to millions. This broadcasting changed the way that candidates interacted with the electorate, but it was still a one-way process: the aim was still to communicate to the electorate what the candidate’s position was.

Towards the end of the 20th century, however, another technological innovation occurred which stood all this on its head. This was, again, the advent of the computer, and with it the collation of vast databases describing voters and their attitudes and opinions. Piggybacking on this, we had the statistical technology that was described above, namely the ability to analyse these large databases to determine what sort of person voted how, what issues were of most concern to them, whether they were likely to be swayed by particular arguments, and so on. Particular individuals could then be targeted, so that a candidate’s resources were used most effectively in the campaign. Floating voters and those in marginal seats were clearly of more relevance than those whose minds were made up, or who would have little influence on the outcome whichever way they voted. Whereas previously what mattered was what the voter knew about the candidate, what mattered now was what the candidate (or, at least, his supporting analytics team) knew about the voter. In place of broadcasting we had ‘narrowcasting’, ‘political sharpshooting’ or ‘microtargeting’.

This is just one example of the use of statistics to make statements which are of direct and particular personal relevance to individuals. My examples of clinical trials, credit scoring and recommender systems are three more. There are innumerable others: some which clearly provide unqualified benefits; others which exist in more murky ethical domains. But I believe that we should educate the public to this role of statistics: although statistics is certainly about summarizing the mass of observations, it also makes highly particular statements about the individual. Raising public awareness of the way that statistics focuses down to benefit individuals must reflect well on the discipline.

Now I have deliberately focused on the positive in the above discussion. But, as we are all aware, although sophisticated technologies can be used beneficially, they can also be misused—and unfortunately it is not possible to have the former without the risk of the latter. There are deep issues, of fundamental importance, relating to personal privacy, which we, as statisticians, need to consider. Things are complicated by the fact that, as statisticians know better than perhaps any other group, databases are seldom perfect—data errors do arise. Moreover, as the media keeps reminding us, computer records do become lost. I think that such issues pose vital questions for society in the coming years. And statisticians need to play a central role in this debate. I also think that they also pose significant challenges to our legislators, and that the increasingly dramatic rate of change in automatic capture and the use of data means that legislation is lagging increasingly further behind. Piecemeal legislation, such as the exclusion of particular characteristics from credit scorecards or constraints on disclosure control in
official statistics, is pecking around the edges. More coherent high level strategic thought is required.

3.2. Statistics and the public perception
I think that everyone in this room is aware of the poor public image from which the discipline of statistics suffers. There are many quips and quotations ridiculing the subject. Admitting to being a statistician in casual conversation seldom stimulates an effusion of interest and enthusiasm.

To anyone with an appreciation of the role that statistics plays in modern society, this is rather puzzling. In particular, to statisticians themselves, who use their skills and tools to dig deep into data, gaining insights and understanding, and unearthing previously undiscovered and possibly unsuspected structures and patterns, it is particularly mystifying. How could such a modern voyage of discovery be anything but exciting? Of course, not everyone suffers from such misunderstandings: a geologist friend once remarked to me that he envied the statisticians, since they were always in at the kill—they were inevitably there when the data gave up their secrets and the discoveries were actually made. But such enlightened individuals are too few and far between among the non-statistical masses.

If we are to tackle this misperception of statistics, we need to understand why it occurs. It seems to me that there are several reasons, and I want to explore some of them, in the hope that pinning them down may help us to formulate strategies to counteract them.

The first is what I might call the ‘historical baggage of tedium’. This is the notion that statistics is primarily concerned with the manual manipulation of masses of figures, so that a facility with mechanical arithmetic is necessary to be a statistician. Now this audience, above all, is one which does not need to be told how false this perception is. It is a hangover from the dark ages, before we had built machines which could take over the mechanical burden of juggling numbers. But it lingers on: perhaps partly because of the way that statistics is taught; perhaps partly because of a widespread innumeracy in the general population. I believe that, for the non-mathematical lay public, ways of teaching the discipline which focus on the properties of the statistics, rather than their derivation, could help to eliminate this false perception. When reading, in a newspaper, that the mean value of one group of people is higher than that of another, the arithmetic details of how these values were arrived at is unimportant. What matters are the properties of the statistics. I like to think that things will improve with increasing familiarity and ready use of tools which permit elementary calculations at the touch of a button, like Excel (notwithstanding any issues of the accuracy of such programs). These entirely remove the need for arithmetic ability and allow the user to concentrate on the meaning.

The second issue is a narrow public perception of what statistics is, which restricts awareness of its extraordinary ubiquity in modern life. The public will be aware of sports statistics and of government statistics, but perhaps of little else. They may not recognize that the debate around climate change and global warming was primarily a statistical one, that the hedge fund industry is based on statistical models and analysis, that statistics plays a central role in modern drug development, that epidemiological issues such as severe acute respiratory syndrome, bovine spongiform encephalopathy, acquired immune deficiency syndrome and avian flu are fundamentally statistical, and so on. The question for us is how to raise public awareness of the key role of statistics in all these and other activities. Although I hesitate to use the word, what is lacking is a proper ‘branding’ of statistics, so that when people read about these issues the role of statistics is explicitly stated and recognized.

A perennial problem is the notion that one can prove anything by judicious manipulation of statistics. There are several contributory factors here. One is that different analyses may lead to different conclusions. This is true at an elementary level—a classic example being the different
conclusions which can be reached by using the mean instead of the median to summarize data. But, of course, it is even more true at a higher level when complex models are fitted. The low level issues may be eased by education—such as teaching an understanding of the properties of different elementary summary statistics—but when conflicting conclusions are drawn there is inevitably a tendency for the non-expert to throw up their hands and to say ‘a plague on all your houses’.

Things are further complicated by the knowing propagation of incorrect statistics, or the selective choice or concealment of statistics to ‘prove’ a case. It is obvious that, in such cases, the discipline of statistics should not be blamed, and perhaps one way to reinforce this principle is for statisticians to be more outspoken when such incidents occur. This is perhaps most pertinent in the political arena, and the newly formed UK Statistics Authority has been tasked with exactly this role of standing up and loudly drawing attention to such abuses of official statistics. Taking this further, perhaps the reputation of the discipline would be enhanced if more noise was made about the forensic role of statistics in detecting abuses in other areas, such as banking fraud and money laundering, fabricated election results and insurance fraud, as well as in areas such as school examination grade inflation.

Sometimes the misuse of statistics is more subtle, and perhaps statisticians are partly to blame. I am thinking of situations which are fundamentally statistical, and in which statisticians are involved, but where perhaps insufficient attention is drawn to the central role of statistics. League tables, key performance indicators and the UK research assessment exercise are examples of the sorts of things that I have in mind: statisticians are certainly involved in such areas, but perhaps they are insufficiently outspoken. The development of any such measure should be led by statisticians, and the first thought of anyone who wanted to develop such a system should be that they should seek statistical advice. I might even assert that it is too dangerous to leave such things in the hands of non-statisticians. Examples of the complications which can arise, and with which statisticians are familiar, are Goodhart’s law, the Hawthorne effect and the placebo effect. Related to this is the concept of measurement error, an aspect of uncertainty: in an idealized lay world, measurements are always completely accurate but, in the real world, this is seldom true. The concept of measurement error, and how to handle it, is well understood by statisticians but is clearly something which others sometimes fail to grasp. There have been recent high profile cases of misdiagnosis of child abuse, based on lack of awareness of the fallibility of diagnostic measurements.

Another factor contributing to the notion that one can prove anything by statistics is a deeper misunderstanding of the nature of science, of evidence and of justifiable belief. The failure is a lack of recognition that science is contingent: that theories are constructed to explain the evidence, and that they can and should be expected to change as more data become available and the evidence changes. I am reminded of the remark by John Maynard Keynes, who, when criticized for having changed his position on monetary policy, replied ‘When the facts change, I change my mind. What do you do, sir?’. This issue is also illustrated by the retrospective updating of official statistics time series as more data become available. It is necessary to raise awareness of the fact that the figures are always estimates, and estimates can be improved as more data become available.

Although I am entirely sympathetic to the unease that uncertainty inspires, the solution to dispelling that unease is not to hide from the uncertainty by criticizing those who draw attention to it, but to try to diminish or dispell it by collecting more evidence, more data, so that understanding grows. The proper response to the fact that scientists report conflicting results about whether particular foods are beneficial or harmful is, not to categorize them as making errors, but to recognize that the world is a complex place and that considerable data and investigation
are needed to understand it. This perhaps applies even more forcefully in the domain of social policy. Although the importance of evidence-based social policy is increasingly acknowledged, I still find myself frustrated by the frequent unsupported assertions in this realm.

Yet another contributor to the notion that one can prove anything with statistics is that it involves profound ideas, which may not be easy to grasp in an instant—the reactions of students when first faced with notions of $p$-values or subjective probabilities illustrate this. Sometimes, as we might expect would be the case for a discipline with some depth, the ideas may even be counterintuitive. Those with the time and inclination will explore these ideas, and come to terms with them, but it would be unrealistic to expect most people to do this. These profound ideas are associated with their own technical language and notation—without which it would be extraordinarily laborious, if not impossible, to discuss them. This then presents a dilemma. Why should we expect that a non-expert, who does not have the time to explore things properly, will be able to distinguish the fact that the group of symbols in a statistics paper (a configuration such as

\[(n_m\alpha^d)^{-1} \sum_{i=1}^{n_m} \prod_{j=1}^{d} K_0\{(y_j - x_{ij})/\alpha\},\]

for example) represents a procedure from a powerful technology which can be used to model reality accurately whereas the combination of symbols in an astrology chart (such as the symbols $\zeta \varphi \gamma \delta \sigma \epsilon \phi \eta \pi \rho \alpha \beta \gamma$, for example) represent nothing more than fantasy? To the untutored eye, they all look like gobbledygook. An appeal to the evidence, the demonstrable power of the first in predicting phenomena in the real world and the lack of such clear demonstrations for the latter, is all very well, but it is only relevant if one has the time and inclination to spend on it. Sometimes we just have to accept what we are told, based on how convincing are the arguments presented by those doing the telling. Once again I have returned to the importance of promoting statistics—its ‘brand’ if you like.

4. **Statistical education: the interface between internal and external**

Given that the discipline of statistics has changed dramatically over the past few decades, and is continuing to change, and given issues of poor public perception of statistics, I thought it appropriate to say a few words about statistical education. Particularly because of the ubiquity of statistics, this is a broad topic, at one end of which we have the training of the next generation of statisticians, and at the other end we have the education of the lay public in the importance of statistics and in elementary statistical concepts. This makes it rather different from many other disciplines. In neurosurgery, for example, although we might consider how best to train the next generation of neurosurgeons, we may not concern ourselves with educating the lay public about the rudiments of neurosurgery.

The identification of the spectrum of statistical education means that, in discussing it, it is necessary to consider whom we are teaching, what we are teaching and indeed why we are teaching it. I do not, in the limited space that I have available, intend to try to present a comprehensive discussion of these various issues, but merely to raise some important aspects.

I think that one of the most important aspects—one of the ways of enhancing brand awareness—is to drive home the central role of statistics when interacting with students from other disciplines. To illustrate that aspect, perhaps I can share with you something that I sometimes say to undergraduate engineering students. Such students are often studying the statistics course in their degree because it is mandatory—they need to pass this course to obtain their degree. But by choice they would probably not be there, and they certainly do not regard it as the high point of
their undergraduate studies. However, the truth is, as I explain to them, that the statistics course is in fact probably the single course which will have most influence on their careers. And the reason is that the skills and understanding that they acquire on that course will be used throughout their lives. After all, when they graduate and begin their engineering work, they will use statistics in designing bridges, constructing aircraft wings, building power stations or whatever else it is that they specialize in. But, after a while, as their careers progress, they will find themselves moving back from the front line into managerial roles. Gradually, if they continue in that profession, they will work their way up the chain, dealing with increasingly higher level issues. And at every level, from the basic application of their engineering knowledge to the higher level decisions about staff numbers, market opportunities and corporate direction, they will need to understand, manipulate and base decisions on data, i.e. they will need to use statistics.

That example was from engineering, but I could have picked any other discipline. For many years I collaborated with psychologists. My experience is that fresh undergraduate psychologists are often shocked by the level of numeracy and quantitative skills that they are expected to have, and in particular by the statistical understanding that they are expected to acquire. But observation also shows that, at least at the more scientific end of psychology, such skills are essential to a successful career and are heavily used. The same is true of medicine.

Engineering, psychology and medical undergraduates are all very well. We might perhaps expect them not to appreciate the excitement, breadth and value of modern statistics (or, presumably, they would have chosen to study statistics instead of engineering, psychology or medicine!). But unfortunately I think it is also true that most mathematics students are also not aware of the true nature of the discipline; and mathematics is still the initial training arena for most of those who actually go on to become statisticians. In particular, fresh graduate statisticians are typically heavily steeped in the mathematical formalisms of inference and are less aware of the wealth of contexts in which they will apply their skills. In this vein, I think that there is much to be said for undergraduate statistics degrees, even if they are taught within a mathematics department, including a course which indicates the breadth of applicability of statistics. It could tell the students something about the use of statistics in medical research, in the physical sciences, in social policy, in economics, in manufacturing, in the service industries, in banking, etc. I am almost tempted to go so far as to claim that to do otherwise is to do a disservice both to the students, who will graduate blind to the multitude of opportunities which are open to those with statistical qualifications, and to the discipline, as the failure serves to promote a narrow view of statistics.

A related point is that statistical competence requires some maturity in understanding the world. Whereas we have mathematics prodigies, we do not have statistics prodigies and I think that a primary reason for this is that statistics is not merely a question of working within a well-defined world of axioms and operations but is fundamentally about relating such a system to the real world. Poincaré wrote:

"Mathematicians do not study objects, but relations between objects. Thus, they are free to replace some objects by others so long as the relations remain unchanged. Content to them is irrelevant: they are interested in form only."

Although statisticians also study relationships, to them content is fundamental: their aim is to make statements about the objects which are the subject of the relationships. The distinction is also driven home by aspects of real data, such as their quality: as John Nelder put it in his Presidential address to the Society,

"data are untidy; they do not satisfy exact mathematical relationships, and so do not lend themselves to deductive reasoning"
It seems to me that failure to drive home this fundamental distinction between mathematics and statistics when teaching the pool from which the next generation of statisticians will be drawn is a lost opportunity.

Since statistics is fundamental to so many different areas, we might have expected to find huge numbers of statisticians being trained. Indeed, we might have expected to see the number increasing, as the world comes increasingly to depend on quantitative and especially data analytic skills. But it is not obvious that they are. In fact, the Royal Statistical Society has expressed concern that the numbers of statistical staff in universities may be declining, with an obvious potential risk for statistics teaching (see Smith and Staetsky (2007))—although, of course, things are complicated by recruitment from overseas (at one point I had five nationalities represented in my own teaching staff of 13 people).

Things are also complicated by statistics being taught to, and statistical analyses being carried out by, people who do not call themselves statisticians but who nevertheless work within the realm of greater statistics as defined above—people such as researchers in pattern recognition, data mining and neural networks. The question is, does it matter whether the statistics is being done (assuming, of course, that it is done correctly) if it is not done by statisticians? The answer is that, although it might not matter to the immediate ‘client’, it could matter to the discipline of statistics. It could also matter to subsequent clients because of a loss of insight consequent on the narrower perspective that is implied by the more restricted objectives of these particular areas. For example, I know of a major corporation which uses a fraud detection system based on a neural network which may be relatively effective, but where the tools that are used for evaluating the system lead to biased, and hence potentially misleading, conclusions.

5. The future

I have stressed the ubiquity of statistics. The modern world relies increasingly on the analysis of data, either overtly or via automatic systems. Modern commerce would be impossible without data analysis. The cutting edge Web-based corporations, supermarkets, the financial sector, all hinge on data. Furthermore, science is increasingly interdisciplinary, with data, and hence statistics, lying at the root of all of science: one might say that statistics bridges the silos of science. The truth is that statistical analysis lies all around us. And yet the general public is unaware of this foundational role that statistics plays. We need to drive home this fundamental truth.

I think that I might see signs of a glimmer of hope. Although there have always been popular statistics books, such as How to Lie with Statistics (Huff, 1954), Facts from Figures (Moroney, 1951) and Use and Abuse of Statistics (Reichman, 1961), this subgenre of popular science is continuing and, it seems to me, growing, with the appearance of such works as Damned Lies and Statistics (Best, 2001), The Lady Tasting Tea (Salsburg, 2002), The Tiger that Isn’t (Blastland and Dilnot, 2007), Dicing with Death (Senn, 2003) and Statistics—a Very Short Introduction (Hand, 2008b). In parallel with these, some books have recently appeared which describe the excitement and opportunity arising from the large data sets which are now accumulating, such as The Numerati (Baker, 2008), Super Crunchers (Ayres, 2007), Competing on Analytics (Davenport and Harris, 2007), Scoring Points (Humby et al., 2007) and my own Information Generation (Hand, 2007). Unfortunately, not all of these present statistics as the core underlying discipline—again a failure in brand promotion.

I have also described greater statistics as the overarching discipline concerned with ‘everything related to learning from data’, to use John Chambers’s phrase. It is very clear that a narrow perspective on statistics has risked marginalizing the discipline, allowing the intellectual proprietorship (along with big research grants) to go elsewhere. The fact that distinct names have been
coined for areas such as pattern recognition, machine learning, data mining, neural networks and expert systems is a sign of lost opportunities. Breiman (2002) wrote:

‘we are ceding some of the most interesting of current statistical problems to computer scientists and engineers allied to the machine learning area’.

Perhaps what is also needed is a little more pride, and a recognition by others that being a statistician justifies pride. I am reminded of Fred Smith’s comment, in his Presidential address (Smith, 1993):

‘Whenever I visit Canada and say that I am a statistician, instead of quoting the wisecracks that greet that admission in this country, Canadians often quote statistics at me and talk with pride about Statistics Canada’.

In the UK we still have a long way to go.

Statistics needs to become more outward looking. With some hesitation, I have used the term ‘brand’ to describe what we have, but it seems an apt term. We need to promote the greater statistics brand. I opened my book *Statistics—a Very Short Introduction* with an assertion which I believe to be true, and which I believe we need to communicate to the wider public. This is that statistics is the most exciting of disciplines.

References


May I begin by congratulating and thanking David for his service to the Society as our President? Although he states in his address that the very ubiquity of statistics poses each President with the daunting task of trying to ‘come up to speed’ with the breadth of the role, he himself has a breadth of interests and experience through statistical methodology and applications in engineering, financial services, psychology, health and social sciences that equip him as well as anyone.

To be accurate I should limit my congratulations and thanks to the first year of David’s presidency. It is unusual in recent times for Presidents to deliver their address before the end of their term of office. David has broken with this practice and, as someone immune from the consequences, I can cheerfully commend the innovation to David’s successors.

This address focuses on the puzzling contrast between the excitement and importance of statistics as an activity, and the dreary public image which the term engenders. Within this there are thoughtful and incisive comments on the very nature of statistics, its ‘brand image’ and our unwillingness as statistical professionals to push out the envelope of what the term might cover to embrace important developments in fields such as computer science. David’s view is that our reticence to be involved with what might be termed empirical as opposed to theory-driven analyses of partial or incomplete data sets, however large, has ceded areas that are essentially statistical to other disciplines.

The President rightly refers to the profound effect of modern computing power and storage capacity on our discipline and the effect of this on our role as statisticians. Many of the computer-intensive methods that are now in common use were beyond the wildest dreams of analysts 30 years ago. As well as permitting new methods, the sheer size and complexity of models that can now be fitted has increased hugely. There is now much less emphasis for statisticians in carrying out detailed precise calculations and an opportunity to take a higher level view: a focus on understanding rather than mechanical manipulation. This is absolutely right but, as David says, modern developments place an even higher premium on ‘a feel for the data’.

This raises an important issue of how we, as teachers, promote and nurture a feel for data in the next generation. This feel for data is really a feel for the interaction between the data and the model being fitted. Errant data points or the ‘shape’ of the data may have no effect on some analyses and distort others greatly. In simple analyses we try to nurture and develop this feel for data in our students. In the vast data arrays and highly complex models that computing power has opened up for us, do we yet understand how to nurture this feel for data or even through what processes we, ourselves, have acquired it? I would add this to the President’s points on statistical education.

Two of the features of modern administrative data sets are their partial or selective nature with respect to any population of interest and their dynamic nature—with new cases, new variables and new values being added to the data set continuously. If we are to use these data for inference purposes, and I am sure that we should, then there is a need to keep in mind some very traditional issues. These are the relationship between the data set and the target population of interest and a clear statement of the targets of inference, particularly in dynamic situations, and how these may be affected by the partial or dynamic nature of the data. Emphasis on these old issues is at a premium when using modern administrative data sets.

The President turns the traditional idea of statistics as summaries or generalizations to large groups on its head and stresses that many analyses of large modern data sets are concerned with predicting individual characteristics, e.g. to predict my personal car or home insurance premium or my personal credit rating. He is surely right and cites the effect on marketing and political strategies. But, as he recognizes, this raises ethical issues about the use of data, particularly the merger of different data sets into ever larger agglomerations. I am sure he is right that high level strategic thought is required to help to inform legislators. I sense that the issue is not one of data design and structures or of methods of analysis because all of these can bring both social benefits as well as privacy dangers in different contexts. I suspect that the focus of any high level thought must be on the intended purpose. Perhaps this is an issue that the Society could give some attention to, in conjunction with other relevant bodies, with a view to producing a public report.

The President regrets the lack of public understanding of scientific processes, the nature of evidence and its relationship to scientific theory. I am sure that we all recognize this and it is a concern which we can share with other sciences. This reflects a real failure in education and communication since for many
people there appears to be no distinction between scientific evidence, well based on a body of empirical and theoretical evidence, and sets of ideas based on homespun ideas. There are many examples where the two are treated as competing sets of ideas that may be accepted or not on an equal basis. This must be a cause for concern for the whole scientific community.

There is much in this address which will demand further thought and I commend it to you. I have the greatest pleasure in proposing a vote of thanks to Professor David Hand for his stimulating Presidential address.

Andrew P. Grieve (King’s College London)
I would like to begin by joining with Tim Holt in congratulating and thanking David Hand for his services to our Society and to wish him well for the second half of his term of office. I have spent the great majority of my career involving myself in the application of statistics to the development of new drugs, albeit not solely in the clinical trials phase. It is consequently humbling to set my own relative narrowness against the breadth and depth of the President’s interests and experience. That personal history can only have been a help to him in scaling that steep learning curve that all Presidents of the Society, as he himself acknowledges, need to scale.

The President’s canvas is vast. It stretches from the beginnings of probability theory in 17th-century France. It passes by the use of administrative data to further the aims of the state and the development of techniques for dealing with observational and measurement error in the 19th and 20th centuries, until at the end it reaches today, a time when techniques, explicitly or implicitly, statistical are used to bring order out of the seeming chaos of the vast terabytes and possibly yottabytes of data that scientists, social scientists and commercial enterprises have the capability of collecting and require interpreting.

It is an exciting prospect but as the President argues it raises profound administrative, social, ethical and educational questions. Administratively, the size of data depositories raises many practical and security issues. We heard this year that Google is considering, and has already patented, a floating data centre that could be powered and cooled by the ocean. These offshore data centres could sit 3–7 miles offshore and reside in about 50–70 m of water. Who today cannot be concerned about the loss of data? Socially there will be public concern about the consequences of commercial and governmental organizations controlling access to such data and the use to which they are put. Ethically, there is the need to address issues of consent to the use of personal identifiable data for secondary analyses of benefit to the wider society, and here I am thinking particularly of epidemiological investigations, and the Society has a role in defining the appropriate boundaries for such use. Finally education will be important because we need to ensure that the excitement that the President has highlighted is instilled in succeeding generations. Part of the solution here may be to revisit the suggestion in ‘Making mathematics count’, the report of Professor Adrian Smith’s inquiry into ‘Post-14 mathematics education’, that ‘data analysis’ be taught in the curriculum as part of other subjects such as physics and geography. This week, hot off the press, comes a proposal by Sir Jim Rose, former Ofsted Chief Inspector and the government’s advisor on primary schools, that lessons such as history and geography should be cut from the curriculum to allow teachers greater flexibility to teach fewer subjects in greater detail. Instead of teaching subjects individually, Rose proposes that the curriculum should focus on cross-curricular studies encompassing a range of subjects and ideas. I see an opportunity here for data analysis to become a cross-cutting, cross-curriculum theme providing support to other subjects while introducing young pupils to the ‘joys’ of data.

With one caveat I support the notion that much of modern statistics has as its focus the individual as well as the collective. The caveat is clinical trials. I would suggest that many clinical trials are designed primarily to support a decision about whether drug A should be licensed, or be recommended as part of a clinical guideline, or reimbursed. It is true that subsequently the results of the trial may be used, together with results from other trials, in a systematic review to answer the treating physician’s justifiable question: ‘Is there evidence to suggest that the probability of benefit to the patient in front of me now could be very different from the aggregate probability of benefit that was reported in the literature?’ A recent Lancet publication edited by Peter Rothwell addresses precisely such objectives. Such analyses are post hoc and I tend to lean towards Stephen Senn’s view that the genomic revolution has yet to prove the contention of many, including former GlaxoSmithKline Chief Executive Officer Sir Richard Sykes, that prospective clinical trials will provide a platform where ‘individual patients will be targeted with specific treatment and personalised dosing regimens to maximise efficacy . . .’.

Finally I offer an additional ‘popular’ book to the President’s list: Nassim Nicholas Taleb’s The Black Swan—the Impact of the Highly Improbable. It is not entirely comfortable reading for a statistical audience.
and yet it shows quite clearly how important it is for statisticians to break out into the wider world of application, to their own benefit but also to the benefit of the wider society.

The President has delivered an address that covers many issues of great importance to statistics and statisticians and it gives me great pleasure to second the vote of thanks.

The vote of thanks was passed by acclamation.