"UNDERSTANDING STATISTICS - WHAT DO ALL THESE NUMBERS MEAN?"

The Appendix to this paper contains copies of the transparencies used at the conference. Copies of the Appendix may be obtained from the author.

Some lecture-notes prepared for a talk given in Canterbury by John Bibby, The Open University, April 1981

[These notes have been prepared in response to requests from a variety of sources. They do not correspond precisely to the paper as given. Comments would be most welcome. J.B.]

1. The institutionalisation of statistical knowledge

1.1 The quantitative revolution: why?

Recent years have seen a phenomenal growth in the use of statistics. If it were not for this "quantitative revolution", as it as been termed, then many of us here would not be statisticians, and none of us would be at this conference here today. To what then can we attribute this quantitative revolution, and what are its effects?

The answer to this question is extraordinarily complex, requiring perceptions gathered from a wide variety of disciplines. We require ideas from relatively esoteric areas such as political economy, the sociology of knowledge and of the professions, theories of managerial capitalism and the state, as well as other more everyday subject-areas such as history , economics, technology, and more recently micro-electronics and computer programming. Only a fool would attempt to reconcile all these subject-areas in a brief talk such as this. I shall make no such attempt, being more concerned with posing questions than with attempting definitive answers.

In discussing historical phenomena such as this quantitative revolution, it is difficult or even impossible to distinguish cause from effect in historical analysis. Indeed one sometimes wonders whether these terms have any meaning. To distinguish cause from effect required tortuous philosophising at times virtually indistinguishable from theology. Studying historical ideas makes me question the general truth of the maxim "correlation is not causation", which we love to teach our students. Thus if I use the terminology of causation now, it is to be interpreted in a loose and quasi-journalistic sense, rather than one of rigid scientific analysis. However, journalism should not be disdained. After all a good journalist is a far more effective educator than the best of academics.

1.2 Statistics as science, and bureaucratisation

What then are the causes or correlates of this quantitative revolution? One answer would focus on the role of statistics in sustaining an "objective" , scientific approach to world affairs. This is what I call the "statistics as science" explanation, where science is perceived as a hard, objective methodology. According to this viewpoint (which is reflected in the RSS motto) statistics is facts, and "facts is facts is facts". The role of the statistician is to collect them. This may be linked with the "rationalising " aspect of bureaucratisation, foreseen by Max Weber and the desire for uncontentious "objective" answers.

Such a portrayal of statistics is not without merit, as long as we are careful to interpret "rational" and "objective" in a strictly Weberian sense i.e. as providing commonly accepted criteria for making decisions. In passing I note here that English may be a peculiarly value-loaded language. While German distinguishes between Zweckrational and Wertrational, in English the word "rational" still carries connotations derived from its latin root i.e. a rational approach is derived from pure reason and is in some sense more "reasonable" than other approaches, and is to be preferred. This ambiguity can be played (and has been) to advantage by the statistical community and gives an exceptionally rosy-eyed view of the subject. We statisticians have not done much to counter this generally flattering viewpoint ,

except in footnotes, small print, and internal memoranda. After all, why should we? It would hardly be the imperial tailor who first spoke the truth concerning the emperor's clothes!

Unfortunately, I feel that our inability or unwillingness to counter this rosy-eyed "scientific" view of the subject is beginning to backfire on us, especially in the social sciences. In recent years the dominant view of statistics in sociology has shifted from the hard-nosed empiricist "facts is facts is facts" point of view towards various brands of relativism and nihilism. We see this in the oft-quoted dictum "There are lies, damned lies, and statistics". The same idea appears in its academic reincarnations as extreme wings of the Althusserian (Hindess) and ethnomethodological schools of sociological analysis (Cicourel). These schools display a distinctly revolutionary aspect. To adapt a dictum of Regis Debray, the quantitative revolution has revolutionized the anti-quantitative counter-revolution!!

Martin Bulmer(1980) has discussed the implications of all this for the use which sociologists make of official statistics. As he points out, the critiques by Hindess(1973) and others are themselves overly deterministic and pessimistic. They also overplay the role of the rational - see especially the Willer's book on "Systematic Empiricism" (Willer & Willer 1973). Some reconciliation is needed between the empirical and the rational. I shall take this up later when I discuss the relationship between statistics and the sociology of knowledge.

1.3 The emergence of the professions

The growth in statistics also reflects the recent burgeoning of professions in general. Thus the "lies" seen in statistics may merely be a numerical counterpart of the "trahison des clercs" which some commentators have noted.

The emergence of the professions itself reflects a restructuring of class relations. This accompanied changes in the division of labour and an increasing concentration of capital (Johnson 1977, p.94). As another commentator has put it, the professions are "a group of newly-created roles, carrying out novel and most rapidly expanding social functions"(Ben-David 1963-4,p.297). (Note the emphasis on the novelty of the function, and on the rapid rate of expansion, both undoubtedly true of the statistical profession in recent years.) The same author goes on to say that the professions have in several respects "taken...the place of the self-made entrepreneur...as the occupational ideal of western societies"(Ben-David 1963-4,p.297)

However, this identification with the entrepreneur does not ring all-together true. In any case, latter-day capitalism has little place for the small-scale innovator. What is required are agents of global capital (Carchedi 1975). Like accountants, although perhaps to a somewhat lesser extent, the statistician is characterised by corporate patronage, whether the corporation be a public company or the nation state(see Johnson 1967, p.106 for a discussion of accountancy in similar vein). The statistical profession results from a fragmentation of labour. This specialisation is one forced on us by the concentration of capital, and it is against this that we must fight when we plead for a "return to our roots" i.e. to the application-based developments of the statistical pioneers.

1.4 Quantification as alchemy

Statisticians have argued whether their subject is a science or an art. It undoubtedly possesses elements of both. Michael Healy has suggested that statistics is in fact better regarded as a technology. That may be so. However, I believe that "craft" is a better description - and not only because one needs to be "crafty" in order to practice the discipline! In fact our subject has many of the traits of the best of medieval crafts - a lot of our learning is done "on the job", and the best statisticians are life-long apprentices.

However, statistics also contains elements of alchemy. Our task is often seen as one of converting base metal into gold. The DIGO principle is "Dross in, Gold out"!

A further undoubted reason for the growth in statistics is its capacity to confuse even the simplest question. Statistics "mystifies" in the sense of obscuring

value judgements and cloaking conflicts where they exist. Cost-benefit analysis is a case in point. Another example is the TRRL car-ownership forecasts, which tended to obscure the necessary assumptions concerning government policy, and failed to consider the possibility of conflicting interests between private and public transport (see Adams 1980).

Having once confused the issue it is an easy step to say that now the question is so difficult that it must be taken out of the public realm - for who can now understand it? - and delivered into the hands of the experts. They will do with it, like Solomon in his wisdom, as they themselves decide.

Thus a real-life political question has been magically transformed into a technical one, and a potentially democratic means of arriving at a solution has been replaced by a technocratic one.

However the strength of our subject, and the magnificent completion of the mystification is that it now can provide a technical apparently objective solution to the re-defined problem . We are the Midas's of the modern world. No matter how dirty the data, we statisticians possess the philosopher's stone which will turn dross into gold . Garbage in, gold out. The alchemist of ancient times was philosopher,medic,mystic, and quack - plus ca change?!

This viewpoint is echoed by many contributors to Irvine et al's (1979) recent book on "Demystifying Social Statistics". Facts, like theory, are value-loaded - some more so, some less so - but even the most innocent looking piece of "data" is highly problematic and should by no means be regarded as "given".

To say this is easy, but in many ways I hope I am preaching to the converted. A far more difficult problem is how to teach this point of view to our students, and how to balance it with more conventional definitions of the subject. At the Open University we are trying to introduce these constructively critical ideas using a variety of viewpoints on what for convenience we term "statistics as a social product". I shall return to this point later in my talk.

1.5 Quantification as an administrative tool

A further reason for the growth of quantification has to do with the ascendancy of the corporate state. In capitalist and communist systems there has been a tendency toward the concentration of capital in larger institutions - governmental and inter-governmental , national and multi-national. The bureaucratic apparatus of these megalithic institutions demands new types of information and control procedures. The data must be uniform in style and easy to file. The decision must be routinised and decisive. There is no room for fuzzy sets in the corporate state!

Linked with this has been the growth of the Government Statistical Service (GSS) and other organisations which are specialised information producers and retrievers. It is important to emphasise here that statistics are produced and not just collected (see Irvine et al. 1977), but the implications of this are not always clear to the users.

In Britain the GSS has seen considerable changes in the past two decades. It has grown substantially and the structure has changed. When the history of this period comes to be written it will not I trust ignore the links between the concentration of capital and the concentration of information. Nor I imagine will it ignore the very real influence of personalities upon history - in this particular case the interaction between two former presidents of the Royal Statistical Society.

1.6 Quantification as technology

The growth of statistics has also been aided by technological developments since the Second World War. Computers have made number-crunching a feasible activity. It has also been extremely profitable - until recently , when microcomputers introduced structural changes into the market.

Along with this has gone a technological determinism - because the facility for complex calculations now exists, simple methods are no longer used. The means has become the end. Perhaps also there has been a loss of confidence amongst

mathematicians, who now find a niche in the evergreen glades of computer-based multivariate techniques.

An interesting question concerns recent developments such as computer graphics and micro-computers - what effect will these have upon the long-run development of statistical techniques - How will the medium affect the message? We may see a resurgence of pattern-seeking methods using compute r graphics. We may witness an "alternative technology " movement based on micros. (The impact upon developing countries will be particularly interesting). However, we shall have to just wait and see while hopefully prodding developments in the direction of more flexible and simpler statistical methods, and a greater dissemination of basic statistical skills. (After all, the numerical counterpart of basic literacy is reading statistical tables, not arithmetic dexterity.)

1.7 Statistics as Cookery

I should also like to mention one other craft which is rarely associated with teaching, but which might be. This is the wondrous skill of cookery.

What connection does cookery have with teaching, you may well ask? Surely he's not just referring to the oft-derided "cookery-book texts"? Well yes, in a way he is.

We have all come across the following sort of recipe : "Take the contingency table in your left hand; shake vigorously. Boil up a table of critical values; add 1% or 5% significance levels to taste. Then pour into the original mixture, regardless of the poisonous look of the brew; garnish with one or two stars until it looks like an entry in the Good Food Guide. Then give your product a foreign, preferably double-barrelled name such as Watney-Mann or Mann-Whitney. Thus you will be virtually assured of a place at the Table d'hôte de Maison."

Of course, none of us here today would dream of cooking like that, let alone cooking statistics! We are I trust more of the Mrs Beeton variety, tentative and non-committal in the extreme - "First catch your hare." In other words we are and should be carefully conscious of any assumptions implicit in the recipes we use.

However, I would wish to take this analogy further, since I believe that there is a sort of cook-book which could provide a highly desirable model for our statistical teaching. What I have in mind is a text that presents cookery not as pure technique, but as an art-form, coupled with notions of experimental science. (Suck it and see!) We need a publication that also takes on the chemical, physical and nutritional implications of cookery (there's a lot of science in domestic science!), and that analyses the relationship between everyday domestic routines and the social environment within which they exist. THAT surely is the book we all are waiting for, and in statistics as in cookery it has yet to be written.

However, let us now move away from gastronomic metaphors, by-passing the question of whether statistics is truly a science, an art, a craft, a technology, or what.

1.8 Institutionalisation of the profession

The emphasis which has been given above on the quantitative growth of statistics, and on its increased legitimacy and prestige, should not blind us to the qualitative evolution which has taken place. Central among these is the emergence of the statistical profession - perhaps the last of the professions (except the oldest !) to be grouping itself in a professional organisation.

The role of "statistician" is something new. How many of the great names - Fisher, Pearson, Yule, Galton - saw themselves first and foremost as "statistician"? Yet now we have two large professional bodies, a specialist branch of the civil service , and every tinpot company has its tame statistical practitioners. This year, I hear, even Wimbledon will have its "statistician", to count false serves, aces, and calculate averages.

The ISI you will be pleased to note, have anticipated this development. Their 15-category "taxonomy of statisticians", ranges from "Academic " and "Business" statisticians through "Biometric", "Medical", and "Social", and ends up with "Survey (including pollster)" and "Other (including sports, numerologist)" (Duncan and Durbin 1980, p. 147). Thus the Wimbledon bailsman is included as a pukka statistician within

the taxonomy of the ISI!

I tend to be very skeptical of any attempts to "harmonise" the profession - it runs great risks of several damaging tendencies. Professionalism tends to institutionalise knowledge - it makes it into a commodity, produced and traded amongst the experts. Moreover, as Ivan Illich (p. 100) has said "it makes people dependent on having their knowledge produced for them. It leads to a paralysis of the moral and political imagination".

It is in the nature of professions that firstly, they erect barriers to entry; secondly, they develop an internal language which marks them off from the man in the street; and thirdly that they attempt to define a professional "harmony" or orthodoxy. "Not so much a profession," as G.B.Shaw put it, " more a conspiracy" - the doctor's dilemma, but also a very topical one for the statistician.

I do not wish to encourage any of these tendencies, still less to vote for the fissiparous tendencies of the profession to divide and subdivide. Rather, I would prefer to make every man his own statistician. Or failing that, perhaps each family or block should have its own "barefoot statistician" who would act as a source of countervailing power against those larger and richer institutions who can afford to employ fully-fledged professionals. I do hope that the current discussions between the RSS and the Institute can take these considerations into account.

2. STATISTICAL "BLINKERS"

2.0 Introduction

Any mode of analysis has areas of strength and areas of weakness - potentials and limitations - and statistics is no exception. This section will look at the limitations. It is written not so much in the spirit of "putting our own house in order" as of asking "What sort of house have we got?" and "Can it be put in order, or do we have to demolish it first?"

I take it as axiomatic that statistical viewpoints can be insightful in certain situations. If that were not the case then I would not be here today. Out of the thousands of papers presented at statistical conferences, 99 per cent examine one insight or another.

My task now is diametrically opposite - I wish to examine the limitations of the statistical viewpoint , or viewpoints What assumptions does it make? What questions does it ask? What questions does it leave unanswered? What sorts of social change does it encourage? Who gains from the use of statistics, and who loses? Let us look at these questions one by one.

2.1 What assumptions does statistics make?

I shall be concerned here only with intrinsic assumptions of statistics, not with extrinsic and relatively contingent assumptions such as normality, independence , linearity, etc.. Nor shall I be concerned with specific techniques e.g. regression analysis. Some of these do have social product-type implications (see e.g. McKenzie 1979 or the use of regression analysis in assessing the Rate Support Grant), but I shall not be concerned with them here.

I shall be concerned with more fundamental issues, e.g. the assumption that we deal with numbers representing scores for a set of cases on a set of variables, and moreover that we manipulate these scores to form summaries such as averages and so on. (Some analyses such as multidimensional scaling have slightly different structures, but the basic points remain the same).

A further point concerns the use of models which aim to capture in some sense the essence of a situation. Some models aim to say how the real world really works. Others are more metaphorical; they do not portray a "truth" about the world, nevertheless, the world acts "as if" the metaphor were true. Friedman ( ) calls these "as if" models. Most statistical models are, I suggest, "as if" models - used for prediction rather than strictly causal inference . Box and Jenkins ( ) have provided one of the better known "as if" models in recent years.

A well-chosen model has a great creative potential, just like a well-chosen phrase or metaphor.

Koestler (1964, p. 322) has described how this might happen - in poetry as well as in science. However, while the essence of poetry may be linked with the existence of emotive metaphors and analogies, the same cannot be said of science. Poetry may well be defined as "that which is lost in translation". But science must be different - or else what meaning could attach to the universality of scientific laws?

Thus metaphors undoubtedly have a creative potential. What is the "Eureka!" process if not the sudden recognition of a link which had not been perceived before? However, one should not overlook the equally important restrictive potential of a metaphor. I will cite two examples from the literature of social stratification, and will return to the notion of metaphor below.

First there is the notion of social "strata", a concept which has become so reified by frequent use that one tends to forget its reference to and undoubted origins in geologically stratified formations. From certain perspectives the use of this geological metaphor in social contexts may be seen as extraordinarily well-chosen, especially if one views society as consisting of relatively unchanging hierarchically ordered groups. But what this metaphor lacks is any notion of change or of conflict. Thus it is essentially a static, functional, consensual metaphor, not easily reconciled for instance with ideas of dialectical progress.

A second metaphor, which may appear to be more dynamic in character, relates to the notion of social mobility. This, following Sorokin (1927) is often seen as happening in "social space". Such a geometrical metaphor seems hardly worthy of comment, until one realises that it immediately makes certain questions appear sensible, and simultaneously quietly rules others out of court. For instance, questions concerning the "dimensionality" of social space, and the distinction between "vertical" and "horizontal" mobility, appear to be meaningful if the geometrical analogy is employed. Otherwise they would not have any clear meaning. On the other hand, questions concerning the "conflict" between various parts of the "space" do not seem so sensible, although they undoubtedly would be if the metaphor were not used. Again, the metaphor used has an essentially anti-dialectical nature.

I do not wish to argue that the metaphors of dialectical materialism are any less likely to blinker than the metaphors of functionalism. I simply wish to point out that in a society where the functionalist ethos prevails, the use of consensual metaphors can blind us to the possibilities of conflict and the processes of change. Hence perhaps the "surprise" that has been evinced worldwide by unforeseen structural changes in society which have taken place in the last decade. To return to statistics. Don't believe metaphor too well. Thus terminology, language and notation inevitably bias the nature of the converstaion, restricting some parts and encouraging others.

2.2 The need for measurement

Statistics assumes quantification, quantification assumes measurement, and measurement assumes that there is something meaningful there to measure in the first place.

With simple measuring instruments this is no problem - if a ruler is measuring nothing, then that is usually perfectly apparent.

More complicated measuring instruments lack this advantage. For instance, recent research into ESP by a Physics professor in London revealed microscopic changes in magnetic field. However, whether they were measuring anything meaningful is highly doubtful.

In social statistics our measuring instruments often go one stage further - they actually CREATE the thing they measure. Whatever our views of the Prime Minister, we are unlikely to think of her on a 5-point scale until the interviewer is standing there in front of us and has posed the question.

Moreover, some people may not even HAVE an opinion, yet feel pressurised to give some intelligent response. One axiom of opinion polling is that people actually have an "opinion" on the thing being polled (Marsh 1979). Yet that may not be true. In such cases the "response" to a questionnaire may merely reflect a (random or

structured) attempt by the respondent to cover up his ignorance or disinterest. Opinion polling is so much part of our culture that poll results cannot be taken at face value. For instance, over 30 years ago Payne(1951) reported that more than 2/3 of his sample ventured an opinion on the non-existent "Metallic Metals Act". More recently, Bishop (1980) found over 30% of his sample were prepared to volunteer an opinion on the "Repeal of the 1975 Public Affairs Act", although this was a totally fictitious issue - no such Act existed, nor was it to be repealed!

In short, then, statisticians measure but they may also be creating that which they are measuring.

Alternatively, as Saint Augustine is rumoured to have put it: "For so it is, O Lord my God: I measure it, but what it is that I measure, I do not know."

The fact that we measure and express our data in the form of numbers limits us to what Sayer(1978, p.6), following Georgescu-Roegen, has called "arithmomorphic" categories. That is, it excludes relational characteristics of society. The philosopher Hegel thought of this a long time ago, and put it as follows: "Number is just that quantity which removes all relational aspects". See also Coser(1978).

2.3 Language, metaphor, and statistical perception.

The relationship between language and perception has become a commonplace in philosophical journals no less that in senior common room chat. The words we use affect the way we think. "Wortschatz heisst Weltanschauung", as the German philosopher might have put it. Thus some liberated women make a big thing about the sexist use of the pronoun "he". A much bigger thing could be made concerning the extreme racist overtones implicit in much of the English language. I only have to say "bete noire" in order to illustrate this point! I recently had to smile at the irony of a recent BBC news item on sporting links with South Africa which used both "black-list" and "blackmail" within a sentence or so of each other. However weeping would have been more apposite if the irony was unintended.

Statistics too uses a peculiarly loaded terminology. William Kruskal (1980) has blamed R.A.Fisher for much of this, and he certainly was no mean innovator in this regard. However, each one of us continues and reinforces the doubtful etymological tradition whenever we use terms such as "unbiased", "optimal", "efficient", etc.. We trade on the ambiguities of the words we use. Would IQ ever have achieved such prominence were it not for the seductive appeal of the phrase "simple structure"?

Of course, we experts know the limitations of the terms we use, but do others? I suspect not. I sometimes wonder whether the careless use of ambiguous language is taking us towards the Humpty-Dumpty situation described by Poincare in connection with the normal distribution: everyone believes statistics, the mathematicians because they think it has been justified experimentally, and the experimentalists because they think it has been justified mathematically.

This confusing mystification reflects the alienation of knowledge - even experts account for their own beliefs by reference to the expertise of others. However, the very use of the word "normal" surely gives the concept a greater apparent legitimacy than if the adjective "Gauss-de Moivre-Laplacian" had become internationally acceptable. (Incidentally, what is the origin of the word "normal"? Pearson claimed it: Galton used it: and William Kruskal wishes to trace it all back to the use of the word in "Ecole normale" or "Normalschule", which goes back to Vienna in the 1830's and an obscure psychologist called Mesmer - unfortunately for the story this was not the great hypnotist, but any ideas on this track would be most welcome. Was it more like Paris in 1790?)

Thus terminology must be a telescope, but it also blinkers - in focussing on particular areas it necessarily obscures others. The same point applies to statistical notation, which carries a whole load of hidden assumotions. The very use of terms like p(i,j) in Markov chain models of social mobility has already assumed away substantive questions such as (a) whether classes are continuous or discrete, and (b) whether transition proportions change over time. Alan Coddington has made this point very effectively in his book on Wage Bargaining.

However, it is not a question of "mere" terminology or of "mere" notation that I wish to raise here, not just a question of the medium and the message. It is also

the question of the whole STATISTICAL METAPHOR, the metaphor of chance, the metaphor of measurement, the metaphor of cases and variables, and the various metaphors of statistical analysis.

Analysis of variance, for instance, is based upon the notion of "decomposition" - 30% of the variance is explained by Factor X, for instance, and 40% by Factor Y. In some situations this decomposition might make sense. But what does it mean in the IQ debate, for instance? Does the metaphor mirror reality? Does ANOVAR make sense? Of course Anovar makes certain technical assumptions, but that is not the root of the problem. The crux of the matter lies in the assumption that decomposition between heredity and environment has some meaning. One might as well say, ASSUMING the moon is made of cheese, what proportion is Camembert? The IQ debate, like many other statistical debates, is woefully reminiscent of medieval theologians, all great men and true, disputing the number of angels on the head of a pin.

There may well be useful insights waiting to be gleaned from a consideration of the relationship between statistical language and statistical perception, but I do not wish to go further into that here. Rather I wish to look at the technical and conceptual ASSUMPTIONS os statistical and quantitative methods, and ask what sort of views of the world they facilitate, and what views they inhibit.

2.4 Abstraction and levels of measurement.

One essence of quantification is its need for ABSTRACTION. From any real-life situation we ABSTRACT certain features, which are then expressed as numbers. And "Numbers is numbers is numbers", as Tom Sawyer might have said, even if we be not concerned whether the numbers relate to balls in Bombay, conferences in Canterbury, or dice in Dar es Salaam. However, once wwe have abstracted the data in the form of numbers then certain operations APPEAR to be relevant. This is partly the levels of measurement problem, but it is more subtle than that, and in any case the classical dicussion of levels of measurement (due to S.S.Stevens) strikes me as unduly simplistic.

Nevertheless, even within the constraints of Stevens' categorization - nominal, ordinal, interval and ratio - I would like to propose a possibly heretical viewpoint: THERE IS NO SUCH THING AS THE LEVEL OF MEASUREMENT OF A VARIABLE. In order to say whether "distance" is nominal, ordinal, interval, or ratio, one has to EMBED THE VARIABLE IN A MODEL, and only then can one begin to ask sensible questions concerning levels of measurement. (And such questions are best posed in terms of invariance e.g. "What tranformations of the variable OR VARIABLES are irrelevant to the substantive conclusion?)

For example, consider the question of the relationship between the location of British regions and their influence on governmental decision- making. Then it may be reasonable to argue that location is

(a) nominal, if we wish to distinguish between England, Scotland, Ireland and Wales, regardless of distance from London,
or (b) ordinal, if influence is reckoned to decline with distance,
or (c) interval, if the rate of decline is relatively constant,
or (d) ratio, if doubling the distance changes influence in a fixed proportion.

The above example is intended to make the point that Stevens' taxonomy is unduly atomistic, and has the effect of diverting our view from regarding the set of variables as a SYSTEM.

2.5 Cases and variables

In my discussion I have already referred to the notion of "variable", which is imbibed by statisticians along with their mother's milk, and is impossible to question. Another key ingredient is the notion of "case" or "individual". The two concepts taken together provide a neat and apparently unquestionable rationale for the concept of "data matrix", and from there it is a turnpike to eigenvalues, eigenvectors, Householder transformations, and all the n-dimensional gyrations loved by advocates of the art of multivariate analysis.

However, whatever the mathematical formulation might suggest, there is no neat

duality between the rows and columns of the data matrix. For instance, how were they selected?

The rows of the data matrix (the "cases") may be selected according to some random sampling strategy. However this is not the case with the columns.

In order to use statistical methods, cases and variables must be defined in a precise, unequivocal, and qualitatively invariant manner. Pigs is pigs and cows is cows, and cannot become pork, hame, beef or cheese, except of course via an arithmomorphic transformation. Thus whether or not the categories have any meaning, our cases and variable ACQUIRE meaning by the use of the statistical formulation. They become reified and thus lead to problems such as the "fallacy of misplaced concreteness" (Sayer 1978).

In selecting which variables to measure, certain arbitrary decisions must be made. Even if some "random" decision rule could be used, there is no analogue to sampling theory which could assess the implications of this rule (Lord and Novick 19??). In truth however, the decision is far from random, and must be based upon some criterion of importance or simplicity. In addition to the theory-based criteria for choice, there is strong pressure towards pragmatism in selection of variables - one measures what can be easily measured.

A measure of the cost of living must focus either on prices or on quality.

??use of numeraire - money - commodification

I have already referred to the "levels of measurement" issue, and the way in which it effectively atomises our viewpoint on a problem. Something similar also happens to the ROWS of the data matrix when we think of them as "cases". this is because the choice of "cases" necessarily obscures the structure of relations within the chosen unit. Sex war, class war, war of generations - all these are muted by the statistical sieve. Professor Galbraith has pointed out how the choice of "household" as the unit of analysis in the census tacitly conceals any conflict of choice between husband and wife or between parents and children (see Galbraith 1980, and also articles in Radical Statistics 20 (1981)).

One can easily escape from a particular LEVEL of atomisation by merely moving up or down the ladder of aggregation. For instance, models of educational attainment can be formulated at the level of the individual child, the class, the school, or the education authority (Bibby and Evans 1978). At each level, different variables become relevant e.g. allocation of resources. However, any particular level completely ignores the dialectical interrelationships at that level and also between the different levels on the ladder of aggregation.

It is perhaps significant that while we have a whole industry devoted to multiVARIATE analysis, the is no such sophistication for multiCASE analysis - all we do is add them up and divide by N! This reflects and also encourages a highly simplistic view of the relationship between individuals.

The different multivariate techniques refect different inter-relationships between the variables, symmetric and a-symmetric relationships. But the relationships between people are far more complex, yet thes are ignored by statistical analysis. Indeed they are even DESIGNED to be ignored, because the whole process of random sampling has the effect of destroying any social grouping that actually exist. For social groups we get social congeries, if I may use Sorokin's distinction.

I might illustrate the sort of relationships I have in mind in educational testing,by focussing on which children play together, work together, who learns, who teaches - who helps whom?

This point was taken up in a geographical context by Sayer (1978, p.4) who complained of the "neglect of people as conscious social individuals". The fact that we use models of "atomistic association" MAY be attributed to cultural hegemony as Sayer claims. More realistically however they are equally due to the cultural hegemony of atomistic ideologies WITHIN SOCIAL SCIENCE.

Social scientists and social statisticians cannot continue to lay all their ills at the feet of some supposedly maleficent tradition stemming from positivistic natural science.(In any case, atomistic theories have now fallen into disrepute even in that area.)

No, there is a very close match between the atomistic ideologies of social science and the assumptions of statistical models. The responsibility lies with us to

expose that relationship, and to change it if necessary.

George Simmel is credited by Ginsberg (1933, p.105) with having transformed German sociology by placing the notion of Beziehung (relation) opposite that of Gegenstand (object). Thus for instance, social groupings became a dynamic complex of inter-linking relationships rather than static entities. Yet despite statisticians' oft-voiced disdain for sociology, it is questionable whether our subject has yet reached that stage in its development.

3. Statistics and the Sociology of Knowledge.

3.1 Popper and Kuhn.

Karl Popper has suggested an iterative loop which could describe how data and theory interact to produce scientific "progress". He calls this the interaction between 'conjecture' and 'refutation'.The statistical counterpart of this loop comes easily out of Tukey's "DATA equals FIT plus RESIDUAL" decomposition. We must listen to data as well as hypothesis - "Facts do not speak for themselves, but neither does theory speak for itself" (Bulmer 1980, p.505).

However, in the sociology of knowledge, Popper has been upstaged by Thomas Kuhn whose discussion of "scientific revolutions" has delineated the phenomenon of "paradigm shift". In such a shift,"normal" science moves through a "revolutionary" period, and eventually into a new period of normality.In this period however the perspectives and paradigms are different from those prevailing in the former period of normal science. Einstein's general theory of relativity is often cited as a case in.point. So also is quantum theory.

We may have been seeing something similar in statistics over the last few years. For several decades the statistical profession has been dominated by notions of mathematical models containing stochastic components, and by "true" parameters having unknown values. The statistician's purpose in life has then been defined as the elucidation of these unkown parameters using hypothesis tests and confidence intervals.

For some years there has been considerable dissatisfaction with this definition of the subject, although this dissatisfaction has undoubtedly been muted by the profession's apparent success at increasing its strength and legitimacy. The dissatisfaction may stem in part from some of the considerations voiced earlier in the talk. However, even within the sacred portals of statistical orthodoxy, voices have been raised concerning the universality of randomness.Some years ago David Finney referred to the problem of assessing doctor's reports on the side-effects of a new drug, given that no "scientific"survey had been carried out i.e. the data was haphazard rather than random. For instance, if 100 reports of undesirable side-effects are received this year compared with only 10 in each of two preceding year, what inference could be made? Statisticians are impotent in such situations, since no stochastic model can be identified. Thus a purist response would say that there is no statistical answer to the problem. However,this is hoisting ourselves by our own petard. There may indeed be a statistical answer - albeit not a definitive one - if we are prepared to define our subject slightly more broadly. Similar points can be made concerning (a) meaningful statements where data is available for the whole population, and (b) procedures such as cluster analysis which have no underlying statistical model. "C'est magnifique," one might say, "mais ca n'est pas la statistique".

Such purism has, I feel been overtaken by events. Whether we like it or not, answers are being demanded for questions which fall outside the traditional statistical field of action. It is up to us to find answers to the given questions, and not to twist the question into something which will allow us to give a purist, but irrelevant answer.

Our failure to give convincing answers to questions such as these has lead to an uneasy silence, a vacuum waiting to be filled. This may be characteristic of the atmosphere that immediately precedes a Kuhnian "parameter shift".

3.2. Tukey

In recent years the vacuum has been gently disturbed by a still small voice, calling plaintively over the waters from Bell Telephone Laboratories and Princeton University. John W. Tukey has the word, and his gospel is EDA (Tukey 1977?). Many disciples have assembled, and the jihad has now begun. Would-be adherents are now waiting in all parts of the kingdom, hoping that soon the holy scriptures will be translated into a language that everyone can understand.

Like most holy books, EDA contains a lot of sound common sense along with a considerable amount of gunge. It also is not free of mystification. Tukey's approach to data analysis reminds one of the American policy for promoting science. In attempting to catch up with the Russians the notion developed that if you throw enough dollars at a question, then some answers are likely to bounce back, or at least some useful ideas. Only the richest nation on earth could afford such a strategy. In Tukey's case it is slightly different - if you throw off enough ideas, like a Roman candle spitting flame, then some of the ideas are likely to be good ones, and the fire will spread.

In short then, with EDA, statistics may be stepping from an era of "normal science" into a "scientific revolution", in which fundamental beliefs and value-orientations will be re-examined. Eventually doubtless we shall return to a new era of "normal science", overseen by a new orthodoxy, a new paradigm. Hopefully this paradigm will be neither EDA nor the current orthodoxy. One great fear I have is that EDA will itself become the new orthodoxy. The band-wagon effect is already evident, and I am reminded of T. H. Huxley's warning that "new truths start life as heresies and end life as superstitions". I wonder whether the heretical notions of EDA will become rapidly internalised, routinised, and computerized, thereby providing examination fodder no less mechanical and uncritical than that so prevalent today.

3.3 What is distinctive about EDA?

But what is there, if anything, about EDA that can truly claim to be revolutionary? There seem to be several candidates for this accolade, which I will now discuss (see Appendix).

As is well know, Tukey's analysis is based on order statistics rather than means and standard deviations. This has theoretical advantages of robustness - order statistics tend to be "sturdy" in a way that moments are not. It also has two practical advantages. Firstly, order statistics are simple - they are simple to teach and, if well taught, are simple to use. I calculate that once one has a back-to-back stem-and-leaf, deriving a confidence interval for the difference between medians takes less than one-tenth of the time needed for the standard methods based on the difference between means. A third advantage of EDA is that it encourages good habits - it keeps one's nose in the data, and forces one to continually question and possible amend one's assumption.

Of course, the emphasis on order statistics is not new. We could well have had EDA in the last century if F. R. Weldon had achieved ascendency over Galton. A decade or more before Galton stumbled upon the "Pearsonian" correlation coefficient - in the grounds of Naworth Castle, as his Memories so vividly describe - F. R. Weldon was already using a measure of correlation based on medians, conditional medians, and quartiles.

However, as history would have it, F. R. Weldon's measure was soon delegated to the dustbin by Karl Pearson, on the grounds that the median and quartile are "too inexact for accurate statistics". (Too easy?)

At The Open University we have also been developing a measure of association which is compatible with EDA. We call this measure PSE - the Proportion of Spread Explained. PSE is a "proportional reduction in error" measure (Costner, American Sociological Review 1965), and is defined as

$$1 - dQ(\text{residuals}) \,/\, dQ(\text{data}).$$

Here "dQ" represents the interquartile range, and PSE has the advantage of being applicable to a vast variety of situations - median polishing, line fitting, curve fitting, indeed any situation where one can use the Tukey fundamental equation

DATA = FIT+RESIDUAL.

There are also other distinctive features of EDA which are referred to in the Appendix, and in the talk.

4 CONCLUSION

The house of statistics has many rooms, and many windows. Our view need not be restricted to just one. Eclecticism is important. this has been true throughout the history of statistics, and we must ensure that it remains true for the future.

SELECTED REFERENCES

Blumen(1956) Social analysis and the variable. Am.Sociol.Rev. December, pp.683-

Milton Friedman( ) Essays in Positive Economics.

Morris Ginsberg (1933) Recent tendencies in sociology. Reprinted in Essays in Sociology and Social Philosophy, vol.2.

William Kruskal(1980) The significance of Fisher. J.Am.Statist.Assoc. 75(372), December, pp.1019-1030.

Koestler, Arthur (1964) The Act of Creation. Hutchinson.

Terence Johnson(1967) The professions in the class structure. Pp.93-110 in Richard Scase, ed. Industrial Society:Class, Cleavage and Control. George Allen & Unwin.

J. Ben-David (1963-4) Professions in the class system of present-day societies. Current Sociology 12(3).

G. Carchedi(1975) On the economic identification of the new middle class. Economy and Society 4(1)

George F.Bishop et al.(1980) Pseudo-opinions on public affairs. Public Opinion Quarterly 44(2), p.198-209. See New Society 19.2.81, p.333.

Stanley L. Payne (1951) The Art of Asking Questions. Princeton University Press.

"LIES, DAMNED LIES AND STATISTICS"

Does anybody know the origin of this quote? I have seen it attributed to Mark Twain, Winston Churchill, and Benjamin Disraeli. Mark Twain certainly blames it on Benjamin Disraeli in his (Mark Twain's) autobiography. But where did Disraeli use it? Any suggestions would be most welcome.

On a more general note, I am compiling a file of apt/witty/readable quotes/poems/stories about statistics. Please send me your favourite(s), and then if I ever get round to compiling a consolidated publication I promise to send you a set.

Please send to John Bibby, Maths Faculty, The Open University, Milton Keynes. MK7 6AA