SIGNIFYING NOTHING: TRADITIONAL HISTORY, LOCAL HISTORY, STATISTICS AND COMPUTING

PETER LASLETT

The pursuit of history in all its modes, political, economic, religious, social, literary, aesthetic, scientific and so forth, is concerned above all with causal analysis and probability. It has to be known with the greatest possible degree of certainty what and why things happened as they did, why the condition of politics or of the arts or of literature or of society was as it was, what is the likelihood that things might have been different, the likelihood of a particular explanation being true or false. Yet in my experience traditional historians proceed as if causality and probability were topics irrelevant to them, with no place in the reconstruction of what happened in the past. They know next to nothing about causal analysis, or the theory of probability and their practical applications. Correlation and regression are closed books.

Traditional historians are rightly concerned with significance. They are anxious that their own statements shall be accepted as significant, and not regarded as intuitive guesswork, such as a novelist or a journalist might produce. Journalists of course have always been the bugbear of established historical criticism. Many traditional historians, perhaps most of them indeed, are alive to the necessity that the evidence which they use should be truly typical, a favourite word, if that evidence is to be accepted as significant. Yet they have no truck with the study of significance itself, with its theory and with the practical applications of that theory, significance testing that is to say.

By traditional historians is intended here those who still rely exclusively on the methods of collecting and analyzing evidence which were in use before the arrival of automatic computation. Such scholars are common amongst political historians, constitutional historians, ecclesiastical or military historians, indeed those who profess subjects which are in themselves recognized as traditional. But some of those who are concerned with the newer lines of study, such as economic history or social history in its various branches can still be called

traditional in the sense laid down. Writers of history in this established, conventional mode are perpetually proclaiming that such and such an individual was typical, such and such an event or a trait was characteristic. But they do so without apparently recognizing that to decide on typicality requires that the whole class of the phenomena under examination has to be surveyed and that all possible individual events or all possible traits must be taken into account.

Most of them seem to be uneasily aware that small samples may be deceptive, and often apologize for inferences based on minimal data. This reveals some recognition of the presence of that randomness which makes any individual case, or any restricted sample, restricted in relation to the whole, so unreliable. Yet the notion of randomness itself seems not yet to have entered the intellectual purview of such historians. When they do try to allow for the effect of exiguous numbers, they are prone to betray their ignorance by supposing that their being able to generalize on one or a few instances can actually be the *cause* of an unexpected or unwelcome outcome, rather than an instance of a value happening to lie at the extreme end of a distribution.

They are in a muddle in fact, because the established laws of sample-size in relation to population at risk and its distribution, as well as well-tried sampling procedures, are not recognized as relevant in the practice of traditional history. We may notice here that this is not entirely the case with politicians or even with journalists themselves, for they cannot always evade the requirement that randomness be taken into account when it comes to such things as the outcomes of opinion polls. These people have been forced to acknowledge that the computer analysis of opinion surveys, and surveys of other kinds too, is a control upon their own activity, and that it won't do to neglect the margin of error derived from the laws of probability.

No such salience, unfortunately, attaches to computer analysis in the everyday conduct of traditional historians, which is one element in their regrettable isolation from their own time. One may wonder nevertheless why the word processors on which so many of them produce their texts, the machines they encounter at the supermarket check-out, or their own sententious clichés about how we all now live in an age of computers, have not brought home to them that computers may have something for their own intellectual pursuit.

In spite of their resistance to sampling, however, they have always been prepared to recognize that the phenomena with which they are dealing varies from area to area, from place to place. This, along with the dilemma of everything being much too large to handle, often compels traditional historians to embark upon local studies, acceptable in their canon always providing that such work is not undertaken by people calling themselves local historians. Without an acquaintance with statistics and statistical theory, however, and it is these highly developed branches of knowledge which have been at issue in

all the statements made here so far, going local in this way does little to get the traditional historian off the hook.

For one thing the question of typicality comes up again in an acute form. If a set of influences can only be handled effectively by the selection of a particularly well-documented county or a village in a study, how can it be known whether the results will be valid for the whole country, the whole economy, or for any of the wholes which historians of the traditional kind so often have to deal with? Or, if, conscientiously, they take two units rather than one to form their sample, and reckon the differences between them, how can they tell whether within-sample differences are greater or less than the differences between that sample and the whole which really interests them? Only statistical reasoning and recommendations can answer such questions as these.

These examples may seem rather remote from the direct concerns of those using computation in historical analysis. But everyone at all familiar with these very general issues will be aware that nowadays reckoning significance, providing for sampling vagaries, working out correlations, regressions, degrees of significance and so on will be done by machine, and practically always by the use of packaged programs. Everyone in the know to this extent will also recognize these further circumstances. One is that the existence of computing at its present state of development (extremely advanced of course compared with what it was only a few years ago, but doubtless due to be rapidly overtaken) makes it possible to undertake tasks of a scale and complexity never before envisaged. In this straightforward way computation has enormously enlarged the field of historical enquiry, or could have done so if the great body of historians had been willing to pursue its possibilities. Enlarged, but also changed, or even transformed, as computing has certainly done with economics.

The computer has in fact now begun to reveal itself in its proper guise, not so much as a sophisticated machine useful for doing difficult sums, but as an all-purpose instrument of intellectual exploration.[1] But a further circumstance has to be taken into account, in that this instrument can be a substitute for thought, as well as an intellectual disciplinarian and a stimulus to imaginative exploration. Packaged programs may stand in the way of the unaware, the unwilling or the plain lazy ever understanding how statistical operations like multiple regressions actually work. What is more, the fascination of the machine itself could give rise amongst those already engaged in historical computation, to a society of computer buffs, who are in it for rewards clearly removed from the substance of historical research and analysis.

I am aware that this last is a distant danger and that I go rather far in my criticisms of established historians in relation to computers. Many writers of history of the traditional kind, even amongst the political historians who are usually the most backward-looking, are quite familiar with statistics and

computer techniques. There is even increasing pressure to require all historians in training, that is to say all graduate students, to be instructed in computing usage, though less disposition, which I myself deplore, to require that they be made to familiarize themselves with elementary statistics and statistical reasoning. The situation has certainly changed, and changed for the better from our point of view, since the first movements towards the introduction of these advances into historical analysis were made in the 1960s and 1970s.

But everyone who is far enough along the road to want to read this journal will recognize that there have been so many ignorant, though sometimes apparently wilful misunderstandings, lost opportunities and unnecessary controversies that these words of critical defiance are still worth uttering again. It is only necessary to recall what happened when counterfactual analysis, analysis which inevitably draws upon statistical reasoning and upon automatic computation, was first introduced into the traditional historical arena by Robert Fogel in his splendid works on American rail-roads, and American slavery.[2] We ourselves had a similar experience at the Cambridge Group for the History of Population and Social Structure when we ventured to use an extremely elementary statistical argument in the book which has been said to have initiated the whole study of family and household in the past.[3] We have also witnessed the almost entire failure of the only full-length attempt to employ and to justify, in theoretical as well as substantive terms, the use of sophisticated statistical arguments and demanding computer theory, practice and processes, in the historical study of social structure. It went practically unnoticed by those to whom it was directed, it was extensively misunderstood by such reviewers as it attracted and seems to have been forgotten by all but the historical demographers.[4]

The studies in this book are for the most part of a more limited character, and their use of computers closer to the ordinary run than this last named volume. That books of the present kind are being written, read and used: that meetings should take place of the character of that to which the present writing was delivered as an opening address: that computers are now doing such splendid things, not only in the hands of non-traditional historians in official places, but of researchers labouring away in restricted areas and calling themselves local historians, and especially of scholars working with such enthusiasm in their spare time: that all this should be true must gladden the heart of a veteran.

Perhaps after all a cohort of historians is in fact coming into being, which will make it unnecessary ever to fill out the Shakespearean quotation which I have taken as a title, so that history to be composed from now on will have nothing whatever of the flavour of

> A tale told by an idiot, full of sound and fury
> Signifying nothing.

ENDNOTES

[1] This capital point cannot be pursued here, even at an elementary level. It is expounded at that level by a book of the broadcasts, J. Palfreman and D. Swade, *The dream machine* (London, 1991).

[2] R. W. Fogel, *Railroads and American economic growth: essays in econometric history*, (Baltimore, 1964); R. W. Fogel and S. Engerman, *Time on the cross*, 2 vols (Boston, 1974).

[3] P. Laslett with the assistance of Richard Wall, eds, *Household and family in past time* (Cambridge, 1972). For the unfortunate results, see P. Laslett, 'The character of familial history, its limitations and the conditions for its proper pursuit', *Journal of family history*, 12 (1987), 263–284.

[4] K. W. Wachter, E. A. Hammel and P. Laslett, *Statistical studies of historical social structure* (New York, 1978).