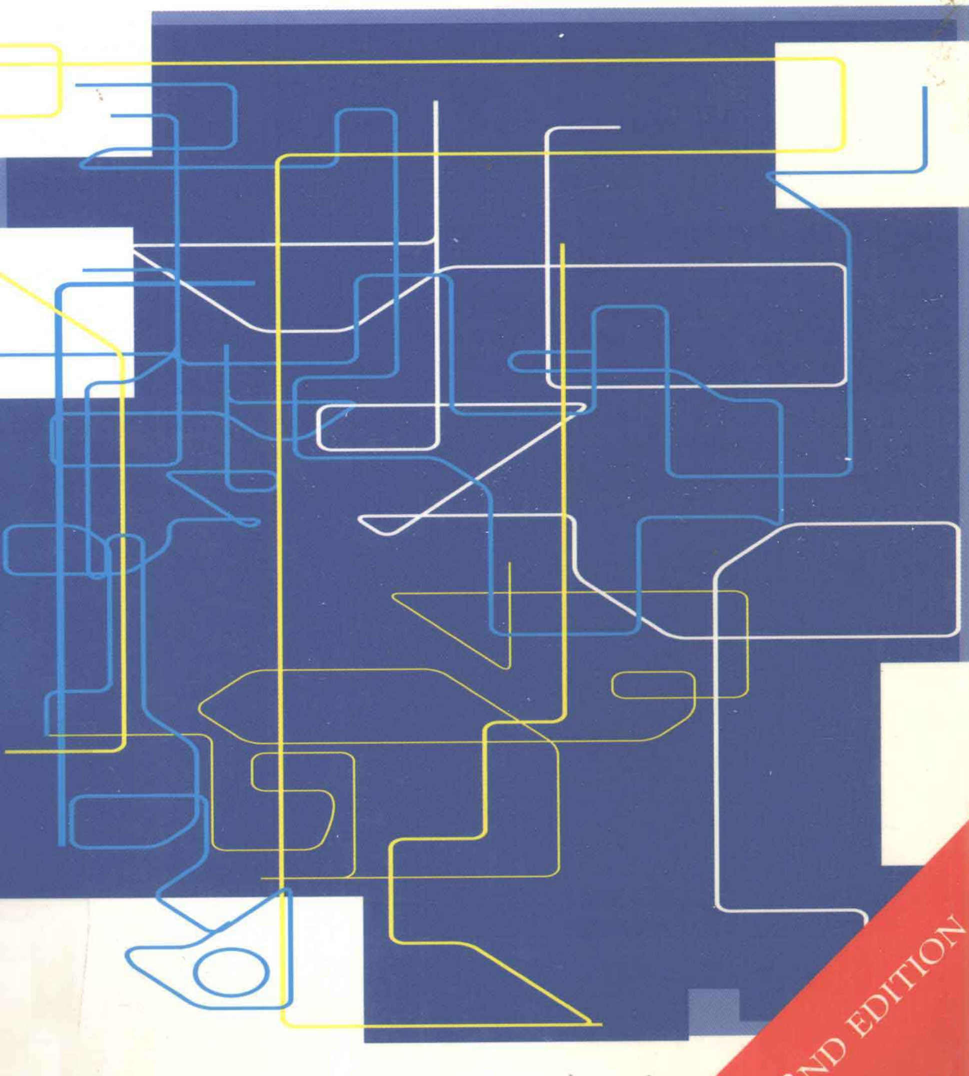
Quantifying Archaeology

STEPHEN SHENNAN



2ND EDITION

Quantifying ARCHAEOLOGY

Stephen Shennan

Second Edition

University of Iowa Press, Iowa City

For Sue

Published in the United States by University of Iowa Press,
Iowa City 52242

Published in the United Kingdom by Edinburgh University Press,
22 George Square, Edinburgh

First published 1988
Second edition 1997

Typeset in Times and Helvetica
by Pioneer Associates, Perthshire
Printed and bound in Great Britain

International Standard Book Number 0–87745–598–8

Library of Congress Catalog Card Number 97–60561

97 98 99 00 01 02 P 5 4 3 2 1

ACKNOWLEDGEMENTS

I would like to thank Mike Baxter, Clive Orton and Richard Wright for helpful comments and suggestions. Needless to say, they are not responsible for any errors.

The following material is reproduced by permission of the organisation specified: tables A, B and E, Cambridge University Press, from *New Cambridge Statistical Tables* (2nd edn) by D. V. Lindley and W. F. Scott, tables 12, 18 and 21; figure for exercise 12.2 and figure 11.10, Society for American Archaeology, respectively from *American Antiquity* 60 (4), 1995, and *American Antiquity* 61 (4), 1996; figure for exercise 13.3, the American Anthropological Association, from *American Anthropologist* 94 (3), September 1992; figures 14.5–14.6, John Wiley and Sons Inc, from *Sampling* by S. K. Thompson, 1992 and the American Statistical Association, from *Journal of the American Statistical Association*, vol. 85, 1990; figure 3.9, Chapman and Hall Publishers, from M. Wand and M. Jones, 1995, *Kernel Smoothing*, figure 2.3; figures 14.7–14.10, University of Southampton and English Heritage, from *Planning for the Past*, vol. 3, by T. Champion, S. J. Shennan and P. Cuming, 1995, figures 1 and 8–10.

I would also like to thank the following for permission to use their published material: the Biometrika Trustees for the tables of the χ^2 and t distributions; Stanley Thornes (Publishers) Ltd for the table of the standardised normal distribution from J. White, A. Yeats and G. Skipworth, *Tables for Statisticians*.

Every effort has been made to trace the copyright holders but, if any have been inadvertently overlooked, the publisher will be pleased to make the necessary acknowledgement at the first opportunity.

CONTENTS

One

INTRODUCTION

The days of the innumerate are numbered
(Colin Renfrew)

The aim of this book is to make students of archaeology familiar with some of the basic quantitative methods used within the discipline as well as some of the more advanced ones which are widely used. The techniques are not specific to archaeology, they are used in a great variety of fields, but experience has shown that archaeologists do not gain a great deal from attending statistics classes for sociologists or biologists. Although the statistical theory and method are the same, the examples tend to be either boring or incomprehensible or both. This situation is particularly unsatisfactory for archaeology students because by and large they are not mathematically inclined, so if alien mathematical concepts are to be understood it has to be from a base in relevant subject matter, preferably involving worked examples.

It is hoped that by the end of the book students will themselves be able to use the simple techniques described and to understand the more advanced ones. But in many ways specific skills are less important than some more general attitudes the book aims to put across. The first of these is a knowledgeably sceptical attitude to the results of quantitative analyses rather than a 'knee-jerk' acceptance or rejection on the basis of uninformed prejudice. The second is a feel for the way in which archaeological questions can be translated into quantitative terms. The third is a basis of knowledge for talking to statisticians about data analysis problems. If you turn to a statistician for help and neither of you knows what the other is talking about, you will probably end up with the wrong answer to the wrong question.

The book assumes very little in the way of prior knowledge. Only the most basic mathematical operations of addition, subtraction, multiplication and division are required, together with a vague memory of roots and powers.

WHY USE QUANTITATIVE METHODS?

The key argument here is that quantitative reasoning is central to archaeology and that a better grasp of its implications is likely to improve our work as archaeologists. Clive Orton's book *Mathematics in Archaeology* (1980) provides an excellent demonstration of why this is the case by taking some of the standard questions which archaeologists ask, such as 'What is it?', 'How old is it?', 'Where does it come from?' and 'What was it for?', and showing how a quantitative approach can help to provide the answers. It follows, therefore, that quantitative methods should be seen, not as a distinct scientific specialism within archaeology, like artefact characterisation techniques, for example, but as part of every archaeologist's mental toolkit. Statistical, mathematical and computer specialists may often be required to cope with particular problems, but archaeologists must have sufficient quantitative awareness to recognise when problems arise which can be helpfully tackled in a quantitative fashion. No one else can do this for them.

Given that this is the case, it remains to be specified exactly where the mathematics and the archaeology come together. Part of the answer is in the simple description of the archaeological record: *counts* of potsherds or lithics of different types, *sizes* of pits, and so on. Such quantitative information is an essential part of all modern archaeological reports, and simple quantitative description is the first topic we will consider, in the next chapter. Methodologically, it is very straightforward; conceptually, it raises important issues which tend not to get the attention they deserve. The results of such quantitative summaries are tables of data and it is on the basis of these that archaeologists build their arguments and draw their inferences. The process usually involves the claim that some sort of patterning exists in the data being considered. One way of doing this is simply to look at the table of data and on the basis of this point out what appears to be important and significant within it. As Wright (1989) emphasises, this is unlikely to be very satisfactory. Mathematically-based techniques can help us to recognise patterning in archaeological data and to specify its nature. The area where mathematics meets the messier parts of the real world is usually statistics. It is precisely this fact that makes statistics in many ways a tricky subject, because mathematical and factual considerations are both involved, and because the patterns are only rarely very clear cut. Nevertheless, inasmuch as all interpretation of the archaeological record is concerned with identifying patterning, it is capable of benefiting from a quantitative approach. The point that, within certain constraints, we are *identifying* patterning rather than creating it is an important one to which we will have to return later. Without such an assumption archaeological evidence would not tell us anything, but one of the virtues of the quantitative approach is that it can tell us in particular cases what a lack of patterning actually looks like.

In this context it is unfortunate that the emergence of 'post-processual' archaeology in the 1980s has led to a reaction against the use of quantitative methods, perceived as associated with the processual approaches which have been rejected. While it is certainly the case that such techniques have been used by archaeologists of the processual school more than anyone else, and some at least of these entertained the over-optimistic view that quantitative data analysis could somehow provide direct insights into the past denied to more traditional approaches, the definition of patterning in data remains fundamental to the archaeological enterprise, whether demonstrating associations between rock art motifs or showing the existence of 'structured deposition' in the archaeological record, and quantitative methods have a vital role to play in this, not least as an antidote to our ever-present weakness for self-deception. In recent years there has been a tendency for archaeology to split between the retrieval and description of data, on the one hand, and discussions of high-level theory with little empirical grounding on the other. This weakness will persist until the zone in between is occupied by the rigorous analysis and interpretation of archaeological data patterning.

THE PLACE OF QUANTITATIVE METHODS IN ARCHAEOLOGICAL RESEARCH

Before turning to the techniques themselves it is appropriate to say something about the place of quantitative methods in the research process. The analysis itself generally comes at a very late stage in the sequence, immediately before interpretation and conclusions, but it is not a good idea to leave it until then before thinking about appropriate techniques of analysis for a particular study (cf. Fieller, 1993). At the research design stage the investigator should be deciding not just what to do but how to do it, including appropriate forms of analysis. Once these decisions are made they define the conduct of the research and nowhere is this more important than in ensuring that the data collected and the method of their collection correspond to the requirements of the techniques it is proposed to use, including the theoretical assumptions the techniques presuppose. Discovering the problems at the analysis stage is too late. Research is not a linear process, of course; it is a loop, because the conclusions will (or should) send you or somebody else back to the first stage again, to design a new investigation.

Two

QUANTIFYING DESCRIPTION

Collections of archaeological material do not speak for themselves; it is necessary for archaeologists to specify aspects of the material which interest them, and these will be determined by their aims (or, very often, by what has become traditional within the discipline). The process of going from aims to relevant aspects of one's material is by no means straightforward. Some archaeologists would say that it has rarely been done successfully and that consequently many if not most archaeological (re)constructions of the past are little more than fictions.

Let us consider an example. Suppose one is interested in studying social stratification through time in a given area. The next step might be to look at the archaeological record of that area and to decide that the best aspect for giving us an indication of changing social stratification would be the variation, through time, in the quantity of metal grave goods deposited in the richest graves in the area. A diachronic picture showing the changing quantities of metal could then be drawn. However, if the quantities of metal deposited related not to the social power of the individuals buried but, for example, to changes in mining technology or in the trade contacts of the area, then the picture would not reflect changing social stratification, but something else. If, after we had mistakenly argued that metal deposition related to social stratification, we then went on to try and explain the reasons for growing social stratification, we would be making matters even worse, because we would be trying to understand a process that never occurred! Presented in this form, the pitfalls seem obvious enough, but they are very easy to fall into in practice, and much recent work has been devoted to improving our understanding of the enormous variety of processes which produce the archaeological record.

For the purposes of this text we will have to skirt round this problem most of the time and to assume that we have selected for investigation an aspect of our material which is appropriate to our interests. In practice, particularly at the level of describing material for an excavation report, for example, there is broad agreement about what categories of information should be recorded and presented, so that we do not have to agonise too much. But we can rightly raise the question whether what has become traditional in such matters is always what we want.

Once we have defined the aspects of our material in which we are interested, it is necessary to prepare a record of them ready for analysis. When data are being collected, the process of assigning a value or score to some aspect of the material in which we are interested constitutes the process of measurement. This is a much more general definition than simply measuring things with a set of calipers or weighing them on a pair of scales – measurement can be of many different kinds. If we are studying a collection of pottery, for example, there are many aspects in which we could be interested: the height or volume of the vessels, the decorative motifs used on them, the fabrics of which they are made, or their shapes. For each vessel in our collection we need to record the information in which we are interested. The result of this work will be a large table of scores and values for each aspect of interest to us (e.g. Table 2.1). The aspects of our material in which we are interested in a given study are usually referred to as the *variables* of interest. Each item that we are studying, whether the items are sites, regions, ceramics, lithics or whatever, will have a specific *value* for each variable.

TABLE 2.1. Example of the information recorded for a group of ceramic vessels.

	Height (mm)	*Rim diameter (mm)*	*Fabric type*	*Rim type*	*Motif in position 1*	*Motif in position 2*	. . .
Vessel 1	139	114	1	1	16	11	. . .
Vessel 2	143	125	2	1	12	9	. . .
⋮							
Vessel *n*	154	121	4	3	21	15	. . .

The process of measurement, especially the description of complex items such as pottery or graves, is by no means straightforward and requires a lot of clear thinking (see for example Gardin, 1980; Richards and Ryan, 1985). It is not, or should not be, simply a matter of sitting in front of a database program filling in the names of fields.

In the past there was little choice about ways of entering data: numbers had to be entered directly into the rather clumsy data entry modules of large statistics packages. These days it is much more usual to enter data into a database or spreadsheet program unless the data set is a trivial one. Spreadsheets, of course, have extensive statistical functions and if data can be easily structured for spreadsheet entry then they are in a form amenable to quantitative analysis. More complex data need to be organised in a relational database and there is an extensive literature on the way to construct these successfully in a way which captures the characteristics of interest and enables information to be retrieved in ways which are flexible, accurate and consistent (e.g. Elmasri and Navathe, 1989); unfortunately, most archaeologists seem unaware of it! An analysis of burials, for example, may involve information about the grave itself, the skeleton(s), possibly individual bones, the positions of the grave goods, their number and their attributes, such as detailed descriptions of pottery. These sets of information are likely to be best stored in separate tables within the database, following the formal rules for relational structure, and must be correctly linked together. Such a structure provides the flexibility to examine, for example, the relationship between the sex or age of a buried individual and the decorative motifs on pottery grave goods buried with them. On the other hand, the rules of good relational structure which make this possible may mean that outputting the data in a form suitable for statistical analysis can be quite complex.

Software does not remove the substantive problems of data description, although it may make it easier to make a good descriptive scheme work; it certainly does not save you from making mistakes. It is obviously vital to use terms and codes consistently and without ambiguity and to avoid logical inconsistencies between different parts of the descriptive system. Systematically describing pottery decoration for computer input can be especially difficult since it can involve making decisions about what are the basic units of the decorative scheme, what are simply variations within the basic structure, and many others (see Plog, 1980, for a good discussion of this).

A general question which often arises is what to include and what to omit from the description, even when you know what your aims are. For example, if you are studying a cemetery of inhumation burials with a view to understanding patterns of burial ritual and social structure, do you include information on the position of each grave good in the grave? Perhaps the exact position of the limbs of the skeleton is significant in some way? The usual answer is to err on the side of inclusion rather than omission, but in a very large study this may involve an enormous amount of work which may not prove relevant and which, if it involves fieldwork, is likely to cost a great deal of money as well as time. It may also produce a dataset which is simply too unwieldy to analyse (cf. Fieller, 1993).

The best way to sort out all the problems which may arise is to carry out a pilot study – a preliminary analysis of a small part of the data using the proposed descriptive system. The importance of this cannot be urged too strongly. It is no exaggeration to say that decisions taken at the coding/description stage will have a major effect on the outcome of the subsequent analyses and that time spent getting it right will more than repay itself.

It might be thought that there is an exception to the above comments: increasingly data are being captured by various kinds of automatic data logging techniques, perhaps in particular the use of video cameras to capture images which can then be manipulated using image analysis techniques. Even here, however, choices and decisions cannot be avoided prior to analysis (see Durham et al., 1994); if it is an image of an object, for example, we have to define what parts of the image will be analysed: the shape only? internal detail? the texture? a segment of the shape? Furthermore, as with all the more laborious descriptive methods, we end up with a table of numbers which we need to do something with.

LEVELS OF MEASUREMENT

Once we have produced a table or tables of data then all the information is there but it is not yet very accessible to us. We are not usually interested in the characteristics of each individual item, but in the assemblage of material as a whole. When we ask questions like 'How common are the different pottery fabrics?', 'Are the vessels a standard size?', answers are not immediately available from the table. We need to summarise our data (the values of our variables) in some way, whether by means of diagrams or summary numbers. Whatever form of summary we choose, however, we first need to consider the measurement characteristics of our variables, or what are known as *levels of measurement*. What are these levels or scales? They are, in order of their mathematical power from lowest to highest, the *nominal*, *ordinal*, *interval* and *ratio* scales.

The nominal scale is so-called because it involves no more than giving names to the different categories within it. You might not think of this as measurement at all, but as the process of classification: placing things in groups or categories, a basic first step in virtually any investigation. Suppose we were studying British Bronze Age funerary pottery and we categorised our pots, following a long-standing classification, as collared urns, globular urns, barrel urns and bucket urns. This would represent a nominal scale, appropriate for this particular set of pots, in which there were four categories. In this case the process of measurement would consist of assigning one of these categories or values to each of our pots. There is no inherent ordering among the pots implied by categorising them in this way. We could assign numbers to the categories, e.g.:

1 = collared urn
2 = globular urn
3 = barrel urn
4 = bucket urn

If we did this we would be using the numbers merely as symbols that are convenient to us for some reason – perhaps as a shorthand notation. It would be meaningless to add or multiply these numbers together.

If it is possible to give a rank order to all of the categories according to some criterion, then the ordinal level of measurement has been achieved. Thus if we categorised the sherds in a pottery assemblage as fine ware, everyday ware and coarse ware, we could say that this was an ordinal scale with respect to some notion of quality. We could rank the fine wares as 1, domestic wares as 2, and coarse wares as 3. Similarly, the well-known and much maligned classification of societies into bands, tribes, chiefdoms and states (Service, 1962) is a rank-ordering of societies with respect to an idea of complexity of organisation. Each category has a unique position relative to the others. Thus, if we know that chiefdom is higher than tribe and that state is higher than chiefdom, this automatically tells us that state is higher than tribe. On the other hand, we do not know *how much* lower chiefdom is than state, or tribe than chiefdom, we simply know the order – it is lower. It is this property of ordering which is the sole mathematical property of the ordinal scale.

In contrast to the ordinal scale, where only the ordering of categories is defined, in interval and ratio scales the distances between the categories are defined in terms of fixed and equal units. The difference between these two, however, is rather less obvious than the others and is best illustrated by an example. Is the measurement of time in terms of AD or BC on an interval or ratio scale? It is certainly more than an ordinal scale because time is divided into fixed and equal units – years. The distinction between the two depends on the definition of the zero point – whether it is arbitrary or not. Defining chronology in terms of years AD or BC is an arbitrary convention. Other different but equally valid systems exist, with different starting points, for example the Jewish or Islamic systems. If, on the other hand, we consider physical measurements, such as distances, volumes or weights, then the zero point is not arbitrary. For example, if we measure distance, whatever units of measurement we use, a zero distance is naturally defined: it is the absence of distance between two points; and the ratio of 100 mm to 200 mm is the same as that between the equivalent in inches, 3.94 and 7.88, i.e. 1:2. This is not true of our chronological systems: the ratio of AD 1000 to 2000 (1,000 years) is 1:2, but if we take the corresponding years in the Islamic chronology, 378 and 1378 (also 1,000 years), the ratio is 1:3.65. Chronology then is an example of an interval scale but physical measurements are examples of ratio scales. In practice, once we get beyond the ordinal scale, it is usually ratio scale variables that we are dealing with in archaeology – physical measurements of the various types referred to above, and counts of numbers of items.

The reason for knowing about these distinctions is that they affect the statistical techniques which we can use in any particular case, whether we are using complex methods of multivariate analysis or merely drawing diagrams. In the chapters which follow, as the different techniques are presented, one of the first considerations will always be the level of measurement of the data for which the methods are appropriate. It is particularly easy to slip into using inappropriate methods these days when the work is always done by computer rather than by hand calculation, since the program will take the numbers you give it at face value and not question whether they are suitable for the operations being carried out.

The discussion so far has emphasised the distinctions between the various levels of measurement but it is worth noting that the scale of measurement for a particular property of a set of data is not necessarily immutable and indeed to some extent is a matter of choice.

Let us return to our example of dividing a pottery assemblage into fine ware, everyday ware and coarse ware, an ordinal scale based on an idea of fineness or quality. In principle, there is no reason why we should not quantify the fineness of the pottery fabric, for example in terms of the mean grain size of the tempering material, or the ratio of inclusions to clay. We would then have a ratio scale measure of fineness and we could place each sherd or vessel on the line from fine to coarse, measured in terms of fixed and equal units. Clearly such a ratio scale contains more information about the property in question than the ordinal scale of fine, medium and coarse and in that sense it might be regarded as preferable.

There is, of course, no reason in principle why we cannot reverse the process. Starting with measurements of grain sizes in our pottery fabrics, for example, we could then categorise them as fine, everyday and coarse. If we do this, however, we are neglecting information, which is generally not a good thing to do. Nevertheless, the argument is not completely straightforward and controversies have raged in the archaeological literature about when and whether it is appropriate to categorise ratio scale variables (see the contributions to Whallon and Brown, 1982, particularly those of Hodson and Spaulding).

The best guide is to make use of the level of measurement that will provide an answer to the question being investigated for the least cost. To refer again to the pottery example, if our investigation requires no more than a distinction between fine ware, everyday ware and coarse ware, it is a waste of time and money to produce a detailed quantified description of every vessel's fabric. However, we may want to analyse a few examples of each fabric type to demonstrate that our distinctions between the fabrics are not totally subjective.

EXERCISES

2.1 Look at the series of German neolithic ceramic vessels in Figure 2.1, p. 14 (after Schoknecht, 1980), and devise a set of variables and values that you think provides the basis for a systematic description of them suitable for entry into a database or a statistics program. Apply your system to the vessels and produce table(s) of values of your variables for each vessel. What problems arose, if any? (Scale: 3:16.)

2.2 Try the same exercise with the set of illustrations of grave plans and their contents from a late neolithic cemetery in the Czech Republic which appear in Figures 2.2 to 2.7, pp. 15–20 (after Buchvaldek and Koutecky, 1970). The contents of the graves are also listed below since the nature of the objects is not always clear from the drawings and not all of them are illustrated. (Scale: plans 1:27, pottery and grindstone 1:4, other items 1:2.)

GRAVE 1
1. Amphora
2. Decorated beaker
3. Flat axe
4. Flint blade
5. Grindstone

GRAVE 2
1. Base sherds of beaker
2. Decorated beaker

GRAVE 3
1. Decorated beaker with handle
2. Decorated amphora
3. Flint blade
4. Piece of copper spiral