26

Error structures of ceramic assemblages

Clive Orton*
Paul Tyers†

26.1 Background

It has long been appreciated that useful archaeological information can be obtained by a quantitative approach to ceramics, and in particular to the comparison of assemblages. Such information has proved valuable in chronological studies, especially seriation (e.g. Millett 1979a), spatial analysis (e.g. Fulford & Hodder 1974), and functional/social analysis (e.g. Redman 1979). Related statistics have proved useful in the study of site formation processes (Schiffer 1987, p. 282–5).

However, there is no consensus as to how ceramic assemblages should be quantified. Over the years, four main contenders for the role of the variable by which to quantify pottery (here called the *measure* of quantity) have emerged. They are:

1. sherd count,
2. weight (or a closely-related measure, such as surface area or displacement volume),
3. vessels represented (i.e. how many vessels are there, sherds of which are in the assemblage?),
4. vessel-equivalents (i.e. counting each sherd as a fraction of its parent vessel, what is the total of all the fractions?).

Measures (i) and (ii) can be counted/measured directly, but (iii) and (iv) must usually be estimated. The abbreviations for these estimates are *evreps* and *eves*; the latter may be modified by specifying which part of the vessel is used to generate the estimate. The most common is *rim-eves*, since rim sherds can often be measured as a fraction of a complete rim.

There is a long history of attempts to compare the relative merits of different measures (e.g. Solheim 1960; Bloice 1971; Glover 1972; Egloff 1973; Evans 1973; Hulthén 1974; Hinton 1977; Vince 1977; Millett 1979b). The main problem has been to devise criteria by which different measures can be assessed. Our view is that this can best be done by formal criteria, e.g. bias and standard errors of estimates. A first attempt, using sampling theory (Orton 1975) gave some results on bias. A second attempt, using simulation (Orton 1982) confirmed the results of the earlier work, and suggested that evreps had lower standard errors than other measures, but failed to make significant progress.

Fundamental questions remained unanswered: it was still impossible to attach confidence intervals to the proportions of different types in an assemblage, or to

* Institute of Archaeology
University College London,
31–34 Gordon Square
London WC1H OPY

† Institute of Archaeology
University College London,
31–34 Gordon Square
London WC1H OPY

assess the statistical significance of differences between the compositions of two or more assemblages.

In October 1988 a research project, 'Statistical analysis of ceramic assemblages', funded by SERC-SBAC, started at the Institute of Archaeology, with the overall aim of producing a computer package which would enable archaeologists to answer such questions routinely. This paper gives the first outcome of the project; as it is still in its early stages, we shall concentrate on the theory, but shall present a few experimental results.

26.2 Statistical theory

26.2.1 Aims and notation

There are two main theoretical aims of the project:

1. to be able to set confidence limits on the proportions of a ceramic assemblage that belong to different types,
2. to be able to compare the compositions of two or more assemblages in terms of the proportion of each type present in each assemblage, and to assess the statistical significance of the differences between them.

The practical aim of the project is to apply the theory to assemblages from a wide range of sites, of different types and periods, to assist in their interpretation, and hence the interpretation of the sites themselves. We expect that the work will also lead to recommendations about the recording of ceramic assemblages.

The term *type* is used in a perfectly general sense, to mean a categorical variable that takes a value on all the pottery from an assemblage. In practice, we use type to mean either fabric, form, or the combination fabric-by-form.

We assume that the pottery is catalogued as *records*, each of which contains (at least) the type of the pottery and a value of its measure (which may be zero); they may also contain other information. A record relates to pottery which is all from the same assemblage and is all of the same type; the totality of the pottery of a particular type in an assemblage may be represented by one or more records.

The number of assemblages making up a dataset is denoted by A, and the number of types by T.

The numbers of records of the jth type in an assemblage is denoted by $m_j, (j = 1, \ldots, T)$, and the total number of records by m. The measure of the ith record of the assemblage is denoted by $w_i, (i = 1, \ldots, m)$. The total measure of a type is denoted by $W_j (j = 1, \ldots, T)$, and the overall total by W. The sum of squares $\sum_j w_i^2$ is denoted by S_j^2.

The symbol $\sim j$ refers to all types *except* the jth, and $\sum_j$ means summation over the jth type.

The approach is to treat each assemblage as a sample from a different population of vessels. Our task then becomes

1. to make point and interval estimates of the proportions of different types in each population, and
2. to test the significance of the differences between the estimates of proportions obtained from the different samples. This can also be seen as testing whether the assemblages could reasonably have come from the same population.

We note that for standard statistical theory to be applicable, the 'observations' (in our notation, records) must be independent of each other. This implies that all sherds with non-zero measure from the same vessel in the same assemblage must be included in the same record. If two sherds of non-zero measure from the same vessel form part of two different records, then those records are correlated and standard statistical theory cannot be used. A record may represent sherds from more than one vessel. The minimum unit of record is the set of all sherds in an assemblage that are from the same vessel (the *sherd family*).

26.2.2 Estimates of proportions in a single assemblage

26.2.2.1 Proportions

The proportion p_j is estimated by

$$\hat{p}_j = W_j/W, \text{ for } j = 1, \ldots, T.$$

26.2.2.2 Variance and covariance

By defining $\hat{p}_j$ as the ratio of two dummy variables $x_i(j)$ and $y_i(j)$, where

$$\begin{aligned} x_i(j) &= w_i \\ y_i(j) &= w_i \text{ if the ith record relates to the jth type,} \\ &= 0 \text{ otherwise,} \end{aligned}$$

we can write $\hat{p}_j = W_j/W = \sum y_i(j)/\sum x_i(j)$, a *ratio estimate.*

Cochran (Cochran 1963, p. 30–1) gives a formula for the variance of a ratio estimate, leading to

$$var(\hat{p}_j) \approx (m/(m-1)W^4)\{W_{\sim j}^2 S_j^2 + W_j^2 S_{\sim j}^2\} \tag{26.1}$$

A similar argument lead to

$$cov(\hat{p}_j, \hat{p}_k) \approx -(m/(m-1)W^4)\{WW_j S_k^2 + WW_k S_j^2 - W_j W_k S^2\} \tag{26.2}$$

This section enables us, for the first time, to estimate the variances and covariances of the proportions of different types in a single assemblage. We can therefore attach confidence intervals to the estimate of the proportion of a type or any combination of types.

26.2.3 Comparing proportions in two or more assemblages

26.2.3.1 Equivalent sample size in a binomial model

Before we can tackle this problem, we need to develop some preliminary theory. Given any type j, we can compare $var(\hat{p}_j)$ with the variance of an estimate based on a binomial model, i.e. on an assemblage of complete vessels. In the latter case, the formula is $var\prime(\hat{p}_j) = \hat{p}_j\hat{q}_j/n$, for a population of size n, where $\hat{q}_j = 1 - \hat{p}_j$. So the variances would be the same if $var(\hat{p}_j) = \hat{p}_j\hat{q}_j/n$. We can turn this round and define $n_j = \hat{p}_j\hat{q}_j/var(\hat{p}_j)$, so that n_j is the number of whole vessels that would give the same value of $var(\hat{p}_j)$ as our sample of m measurable records. The full formula is

$$n_j = ((m-1)/m)W_j W_{\sim j} W^2/\{W_{\sim j}^2 S_j^2 + W_j^2 S_{\sim j}^2\} \tag{26.3}$$

It is important to note that n_j is just a number which, when applied to the binomial formula for variance, give the same result as equation (26.1). It is *not*, for example, an estimate of the number of vessels in the original population.

It has been found that $n_j = n_k$ when either

1. there are only two types, or

2. $S_j^2/W_j = c$ for all j; this is the weakest condition so far found. It is satisfied if all types have the same mean and variance of w; it can be satisfied under other, weaker, conditions. These however seem to be mathematical oddities, and not likely to be encountered in practice.

26.2.3.2 Pooled estimate n of the equivalent sample size

Suppose now that all types have the same mean and variance of w. Recall that

$$var(\hat{p}_j) = (m/(m-1)W^4)W_{\sim j}^2 S_j^2 + W_j S_{simj}^2) \quad (1.1)$$

and that $n_j = \hat{p}_j \hat{q}_j / var(\hat{p}_j)$,

and $\hat{p}_j = W_j/W$

We pool our estimates of the mean and sum of squares of w, obtaining W/m and S^2 respectively, and replace W_j by $W(m_j/m)$, S_j^2 by $S^2(m_j/m)$, leading to

$$n_j = ((m-1)/m)W^2/S^2 \quad (26.4)$$

It follows that

$$var(\hat{p}_j) \approx (m/(m-1))(S^2/W^2)(m_j/m)(m_{\sim j}/m) \quad (26.5)$$

and

$$cov(\hat{p}_j, \hat{p}_k) \approx -(m/(m-1))(S^2/W^2)(m_j/m)(m_k/m) \quad (26.6)$$

26.2.3.3 Homogeneity

We define an assemblage as *statistically homogeneous* if the observed measures for all types could reasonably have come from the same frequency distribution. It follows from this definition that the mean, variance and n-value of each type should not differ significantly from those of all other types.

This definition will prove to be inadequate for the complexities of archaeological data (see section 1.2.4). Nevertheless, it is a useful starting point for the development of a theory for the handling of such data.

26.2.3.4 Tests of homogeneity

Two aspects need to be examined:

1. whether the types as a whole are homogeneous,
2. if not, which types are the cause of departure from homogeneity.

In this section we look at the former; the latter will be dealt with in section 1.2.4.

The weakest current condition for homogeneity is that

$$S_j^2/W_j = c \text{ for all types } j,$$

i.e. $((m_j - 1)/m_j)var_j(w_i)/\bar{w}_j + \bar{w}_j = c$ for all j. For the sorts of values we are dealing with, $var_j(w_i) < \bar{w}_j$, so the major component of variation in n is likely to come from variation in $\bar{w}_j$, rather than in $var_j(w_i)$.

This being so, we can reasonably test for homogeneity in the n-values by testing for homogeneity of the means $\bar{w}_j$. The obvious test is analysis of variance for the equality of means (Kendall & Stuart 1973, p.522). The statistic

$$F = (1/(T-1)) \sum^T m_j(\bar{w}_j - \bar{w})^2 / \{(1/(m-L)) \sum^T_j \sum (w_i - \bar{w}_j)^2\}$$

is distributed as $F_{T-1,m-T}$. This has the drawback that it assumes normality, although it is said to be very robust. Fieller (*pers comm*) has suggested on theoretical grounds that a lognormal distribution (i.e. a logarithmic transformation) might be appropriate. This view is supported by data.

None of the alternative tests e.g. those due to Fisz 1963, p. 407–10, Birnbaum & Hall 1960 and Kruskal & Wallis 1952 are appropriate. The F-test is therefore used.

26.2.3.5 Several assemblages

Now we at last have the background theory we need to look at the comparison of several assemblages, say A of them. We have:

$$\begin{array}{rl} \text{vectors of measures} & \{W_{r1},\ldots,W_{rT}\}, \\ \text{of numbers of observations} & \{m_{r1},\ldots,m_{rT}\}, \\ \text{of sums of squares} & \{S^2_{r1},\ldots,S^2_{rT}\}, \\ \text{and estimates of proportions} & \{\hat{p}_{r1},\ldots\hat{p}_{rT}\}, \\ \text{and variance-covariance matrices} & \|cov(\hat{p}_{rj},\hat{p}_{rk})\|, \end{array}$$

all for $1 \le r \le A$.

We want to compare the vectors of estimated proportions, e.g. to test a hypothesis H_o: all assemblages are 'the same', i.e. can be thought of as samples from the same parent population.

We assume that each assemblage is homogeneous and so has a single n-value, which we call n_r, for $1 \le r \le A$.

We replace each W_{rj} by $n_r(W_{rj}/W_r)$, for $j = 1,\ldots,T$ and $r = 1,\ldots,A$.

Calling the new numbers $W\prime_{rj} = (n_r/W_r)W_{rj}$, we have $W\prime_r = n_r$ for all assemblages r.

The estimates of proportions are unchanged:

$$\hat{p}\prime_{rj} = W\prime_{rj}/W\prime_r = W_{rj}/W_r = \hat{p}_{rj},$$

and so are their variances and covariances:

$$\begin{array}{rcl} var(\hat{p}'_{rj}) & = & (m/(m-1))(S\prime^2_r/W\prime^2_r)(m_{rj}/m_r)(m_{r\sim j}/m_r) \\ & = & (m/(m-1))(S^2_r/W^2_r)(m_{rj}/m_r)(m_{r\sim j}/m_r) = var(\hat{p}_{rj}), \\ \text{since } S'^2_r & = & (n_r/W_r)^2 S^2_r \text{ and } W'^2_r = (n_r/W_r)^2 W^2_r. \\ \text{And } cov(\hat{p}'_{rj},\hat{p}'_{rk}) & = & (m/(m-1))(S'^2_r/W'^2_r)(m_{rj}/m_r)(m_{rk}/m_r) \\ & = & (m/(m-1))(S^2_r/W^2_r)(m_{rj}/m_r)(m_{rk}/m_r) = cov(\hat{p}_{rj},\hat{p}_{rk}) \end{array}$$

for the same reason. Recalling (26.4), that $n_r = ((m-1)/m)(W^2_r/S^2_r)$, and writing $m_{rj}/m_r \approx \hat{p}_{rj}$, etc., since the assemblages are homogeneous, we have

$$var(\hat{p}_{rj}) \approx \hat{p}_{rj}\hat{q}_{rj}/n_r$$

and

$$cov(\hat{p}_{rj},\hat{p}_{rk}) \approx -\hat{p}_{rj}\hat{p}_{rk}/n_r.$$

But these are exactly the same as the variance and covariances we would obtain from a multinomial distribution with parameter p and sample size n.

This is a very important result. It means that, as a large-sample approximation, we can treat the transformed data as a series of samples from multinomial distributions. We can therefore treat them collectively as a contingency table, and use any of the theory appropriate to contingency tables (e.g. log-linear models, see section 26.2.5; correspondence analysis, see section 26.4).

For the first time, this approach enables us to make proper statistical comparisons between the proportions of different types in different assemblages.

As a point of notation, we refer to the transformed values W'_{rj} as *pseudo-counts*. They are not usually integers, but can be treated for statistical purposes as if they were.

26.2.4 The quantum effect

When the above theory was applied to data, a snag immediately appeared. Certain types (e.g. flagons) consistently had higher values of $\bar{w}$ than other types when the rim-eves measure was used. This is because they have small rims which break into fewer fragments. The rims are more 'chunky' than the vessels as a whole, so that using rim-eves overstates the mean completeness. Such types are called *chunky* types, and the effect is called the *quantum effect.*

The problem can be overcome by detecting such types (a multiple t-test on the $\bar{w}$ values was found to be suitable), and omitting them from the estimation of n. If the first m_j records are chunky and the other $m_{\sim j}$ are ordinary, equation (26.4) becomes

$$\hat{n} = (1 - W_{\sim j}/m_{\sim j}W)WW_{\sim j}/S^2_{\sim j} \qquad (26.7)$$

This equation enables us to accommodate the quantum effect within our definition of a statistically homogeneous assemblage, and still use the transformation to pseudo-counts and contingency table theory.

26.2.5 Log-linear models

Suppose we have a table of pseudo-counts (see section 26.2.3.5) n, which may be two-way (e.g. context-by-fabric; context-by-form) or three-way (context-by-fabric-by-form).

26.2.5.1 Notation

To follow the standard notation (e.g. Fienberg 1977) we replace n by x, with subscripts i, j and k for three variables. The expected counts in the cells, under various models, are denoted by $\hat{m}_{ij}, \hat{m}_{ijk}$ (two- and three-way respectively). The marginal totals are denoted by 'dots', e.g. $\sum_j x_{ij} = x_{i\cdot}$; $\sum_k x_{ijk} = x_{ij\cdot}$; $\sum_{i,j} x_{ijk} = x_{\cdot\cdot k}$ etc.

26.2.5.2 Two-way tables

We can test for the independence of the two variables as follows:

$$\text{estimate } \hat{m}_{ij} = (x_{i\cdot}x_{\cdot j})/x_{\cdot\cdot} \text{ for } i = 1, \ldots I \text{ and } j = 1, \ldots, J.$$

We test for differences between the 'observed' (x_{ij}) and the 'expected' ($\hat{m}_{ij}$) by

$$\text{chi-squared } = \sum_i \sum_j (O-E)^2/E = \sum_i \sum_j (x_{ij} - x_{i\cdot}x_{\cdot j}/x_{\cdot\cdot})^2/(x_{i\cdot}x_{\cdot j}/x_{\cdot\cdot})$$

which has $(I-1)(J-1)$ degrees of freedom.

26.2.5.3 Three-way tables

We choose the subscripts i, j, and k to refer to context, fabric and form respectively. Context is treated as an explanatory variable, and fabric and form as response variables.

We can construct a set of nested models of increasing complexity and archaeological reality.

Model 1: $log\hat{m}_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)}$, "complete independence",

Model 2: $log\hat{m}_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{23(jk)}$, "fabric-by-form interaction only",

Model 3: $log\hat{m}_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{23(jk)} + u_{31(ki)}$, "fabric-by-form and context-by-form interactions",

Model 4: $log\hat{m}_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{23(jk)} + u_{31(ki)} + u_{12(ij)}$, "all pairwise interactions",

Model 5: $log\hat{m}_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{23(jk)} + u_{31(ki)} + u_{12(ij)} + u_{123(ijk)}$, "all interactions", the saturated model.

Other routes from Model 1 (complete independence) to Model 5 (saturated) are possible, and can be created by bringing in the two-variable interactions in a different order. The order chosen here is the one that seems to be the most reasonable archaeologically.

Within each model, we can calculate the estimates $\hat{m}_{ijk}$ and carry out a goodness-of-fit test, by calculating

$$X^2 = \sum(\text{observed} - \text{expected})^2 / \text{expected}$$

and

$$G^2 = 2\sum(\text{observed}) \log(\text{observed} / \text{expected}),$$

both of which have approximately chi-squared distributions.

We use a method due to Lancaster 1951 to partition the overall chi-squared statistic, thus enabling us to find the simplest model, out of a chosen hierarchy of models, that fits the data reasonably.

The simplest statistically is complete independence of fabric, form and context (model 1), i.e. the proportions of the different forms are the same in all fabrics, and the proportions of both are the same in all contexts. It is archaeologically incredible.

The next simplest (model 2) is that different forms occur in different proportions in different fabrics, but that the proportions of fabrics and forms are the same in all contexts. This may occur if all the contexts are 'similar', or the assemblages are so small that differences between them are not significant.

Model 3 introduces the possibility that the proportions of forms may vary from context to context; so may the proportions of fabrics, but only as a side-effect of variations in forms and the preference of certain forms for certain fabrics.

Model 4 allows proportions of fabrics and of forms to vary independently of each other from context to context. This would allow for functional variability (forms) as well as chronological variability (fabrics and forms) and geographical variability (fabrics).

Model 5 is the most complicated, and one hopes it would not be needed as it would be difficult to interpret. It is here as a 'backstop' should all other models fail to fit the data.

We have found it necessary to use quasi-log-linear models (Bishop *et al.* 1975, p. 177–228) because of the incomplete nature of the data.

26.3 Data

So far we have concentrated on data catalogued as one sherd-family per record (see section 26.2.1), as this format is the most suited to our approach. Other formats will be used later. Four datasets in this format have been acquired to date: Lime Street, City of London (Richardson 1985, p. 49) and Silchester phases 1 to 3 (Frere 1987, p. 348). Artificial datasets have been used to test aspects of the programs.

The Lime Street catalogue relates to Roman pottery dating from about AD 70 to AD 160. It has been classified by context, fabric and form, and quantified by rim-eves, using the standard Museum of London, Department of Urban Archaeology, recording system. Because the amounts of pottery in each context are generally small, the analysis below is based on a grouping into six phases.

26.4 Results

We here present the results of a correspondence analysis carried out on the data described in section 26.3, after transformation to pseudo-counts (see section 26.2.3.5) but without any special treatment, e.g. no types were omitted from the calculation of the pseudo-totals. The program used is part of the iastats package (Duncan *et al.* 1988), based on one published by Greenacre (1984). Two analyses were made: forms by phase and fabrics by phase.

26.4.1 Forms by phase

The 1st principal axis (50% of total inertia) is dominated by FINE BOWL (90%) and phase 6 (97%). The 2nd principal axis shows a contrast between FLASK (6%), BEAKER (32%) and perhaps BOWL (11%) against AMPHora (13%) and FINE CUP (13%), matched by a contrast between phase 3 (50%) and phase 4 (46%). These hint at possible functional differences which need further investigation. One would not expect functional differences to appear clearly, if at all, at the level of phase-assemblages.

26.4.2 Fabrics by phase

The 1st principal axis (35% of total inertia) is dominated by fabrics SHEL (56%) and SAND (20%), and by a contrast between phases 1 and 2 (77% and 9% respectively) and phase 4 (11%). On the 2nd principal axis, fabrics BB2 (54%) and KOLN (4%) stand out, as does the contrast between phases 5 (70%) and 3 (23%).

The picture is clearer when both are seen together (Fig. 26.1). Here we can see the characteristic parabola shape of a chronological sequence (Madsen 1988, p. 24). Only phase 6 is out of order; its data point is of low quality (lies 'off the plot') and we have noted a possible functional difference (section 26.4.1). An 'early' fabric (SHEL) is at the beginning of the sequence and three 'late' ones (for this site)—BB1, BB2 and KOLN—are at the end. MORT occupies a central position; it is a 'rag-bag' type comprising a variety of rare and unidentified fabrics.

26.5 Discussion

As well as solving two long-standing theoretical problems, this work shows great potential for the interpretation of ceramic assemblages. It has implications for the way in which pottery is catalogued. It is likely that different sorts of interpretation (functional, chronological, distributional) will be possible at different levels of grouping (context, phase and site assemblages).

Acknowledgements

This work has been supported by SERC grant GR/E 95873. We are grateful to the Museum of London Department of Urban Archaeology for permission to use and publish data from Lime Street, and to Jane Timby and Michael Fulford for allowing us to use data from the recent Silchester excavations. We thank Nick Fieller for valuable theoretical advice.

Bibliography

BIRNBAUM, Z. W. & R. A. HALL 1960. "Small sample distributions for multi-sample statistics of the Smirnov type", *Annals of Mathematical Statistics*, 31: 710–20.

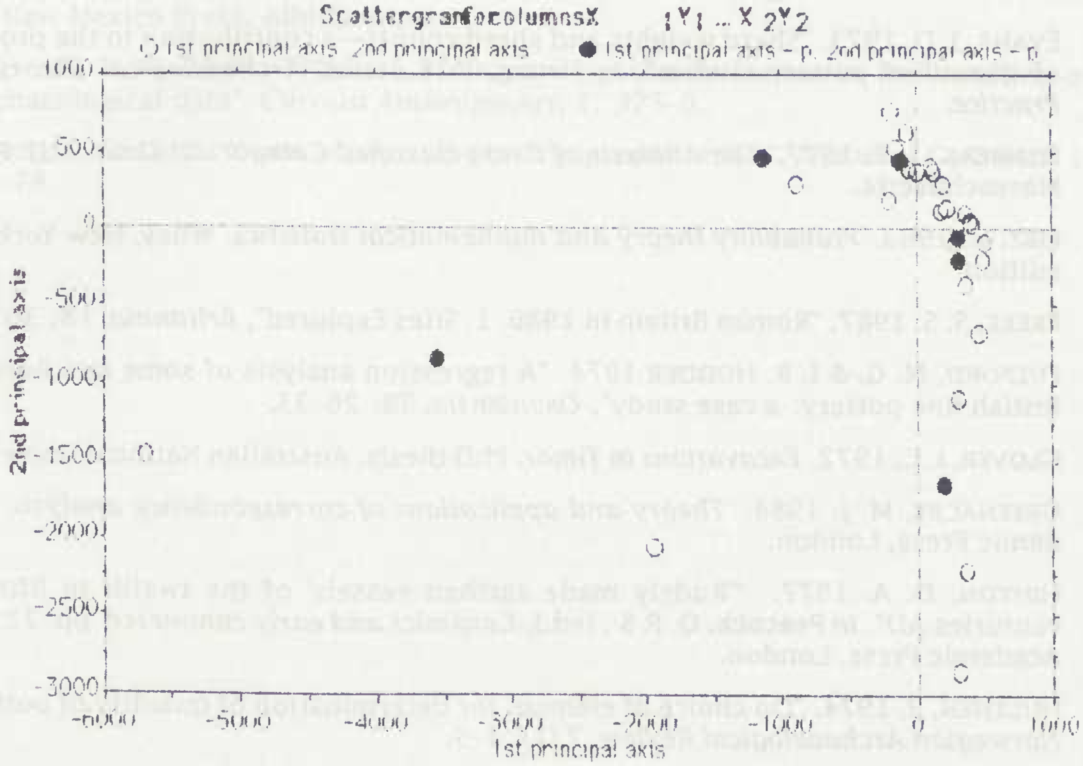
Scattergram for columns X 1Y1 ... X 2Y2
1st principal axis 2nd principal axis
1st principal axis – p, 2nd principal axis – p
1000
500
0
-500
-1000
-1500
-2000
-2500
-3000
-6000
-5000
-4000
-3000
-2000
-1000
0
1000
1st principal axis
2nd principal axis

Figure 26.1:

BISHOP, Y. M. M., S. E. FIENBERG, & P. W. HOLLAND 1975. *Discrete Multivariate Analysis.*

BLOICE, B. J. 1971. "'Note' *in* G. Dawson (ed) Montague Close Part 2", *London Archaeologist*, 1 (11): 250–1.

COCHRAN, W. G. 1963. *Sampling Techniques.* Wiley, New York, 2nd edition.

DUNCAN, R. J., F. R. HODSON, C. R. ORTON, P. A. TYERS, & A. VEKARIA 1988. "Data analysis for archaeologists: the Institute of Archaeology programs". MSS.

EGLOFF, B. J. 1973. "A method for counting ceramic rim sherds", *American Antiquity*, 38 (3): 351–3.

EVANS, J. D. 1973. "Sherd weights and sherd counts—a contribution to the problem of quantified pottery studies". *in* Strong, D. E., (ed.), *Archaeological Theory and Practice.*

FIENBERG, S. E. 1977. *The Analysis of Cross-classified Categorical Data.* MIT Press, Massachusetts.

FISZ, M. 1963. *Probability theory and mathematical statistics.* Wiley, New York, 3rd edition.

FRERE, S. S. 1987. "Roman Britain in 1986. 1. Sites Explored", *Britannia*, 18: 301–59.

FULFORD, M. G. & I. R. HODDER 1974. "A regression analysis of some late Romano-British fine pottery: a case study", *Oxoniensia*, 39: 26–33.

GLOVER, I. C. 1972. *Excavations in Timor.* PhD thesis, Australian National University.

GREENACRE, M. J. 1984. *Theory and applications of correspondence analysis.* Academic Press, London.

HINTON, D. A. 1977. "'Rudely made earthen vessels' of the twelth to fifteenth centuries AD". *in* Peacock, D. P. S , (ed.), *Ceramics and early commerce*, pp. 221–38. Academic Press, London.

HULTHÉN, B. 1974. "On choice of element for determination of quantity of pottery", *Norwegian Archaeological Review*, 7 (1): 1–5.

KENDALL, M. G. & A. STUART 1973. *The Advanced Theory of Statistics, Vol. 2.* 3rd edition.

KRUSKAL, W. H. & W. A. WALLIS 1952. "Use of ranks in one-criterion variance analysis", *Journal of the American Statistical Association*, 47: 582–640.

LANCASTER, H. O. 1951. "Complex contingency tables treated by the partition of chi-square", *Journal of the Royal Statistical Society, Series B*, 13: 242–9.

MADSEN, T., (ed.) 1988. *Multivariate Archaeology, Numerical Approaches to Scandinavian Archaeology.* Jutland Archaeological Society Publications XXI. Aarhus University Press, Aarhus.

MILLETT, M. 1979a. "The dating of Farnham (Alice Holt) pottery", *Britannia*, 10: 121–77.

MILLETT, M. 1979b. "How much pottery?". Occasional Publication 4, pp. 77–80. London University Institute of Archaeology.

ORTON, C. R. 1975. "Quantitative pottery studies: some progress, problems and prospects", *Science and Archaeology*, 16: 30–5.

ORTON, C. R. 1982. "Computer simulation experiments to assess the performance of measures of quantity of pottery", *World Archaeology*, 14 no. 1: 1–20.

REDMAN, C. L. 1979. "Description and inference with the late medieval pottery from Qsar es-Seghir, Morocco", *Medieval Ceramics*, 3: 63–79.

RICHARDSON, B. 1985. "Excavation Round-up 1984, Part 1", *London Archaeologist*, 5 (2): 47–52.

SCHIFFER, M. B. 1987. *Formation processes of the Archaeological Record*. University of New Mexico Press, Albuquerque.

SOLHEIM, W. G. 1960. "The use of sherd weights and counts in the handling of archaeological data", *Current Anthropology*, 1: 325–9.

VINCE, A. G. 1977. "Some aspects of pottery quantification", *Medieval Ceramics*, 1: 63–74.