2470

Ancient Monuments Laboratory
Report 165/87

ON THE RELATIONSHIP BETWEEN THE NUMBER OF TAXA AND THE NUMBER OF SPECIMENS IDENTIFIED IN ARCHAEOLOGICAL BONE ASSEMBLAGES.

T P O'Connor

Ancient Monuments Laboratory Report 165/87

ON THE RELATIONSHIP BETWEEN THE NUMBER OF TAXA AND THE NUMBER OF SPECIMENS IDENTIFIED IN ARCHAEOLOGICAL BONE ASSEMBLAGES.

T P O'Connor

Summary

Results are presented of a series of analyses investigating the relationship between sample size and taxon yield in four hand-collected and sieved bone assemblages from archaeological sites in York. The results confirm Casteel's identification of a significant non-linear relationship, and the implications of this relationship for sampling and analytical strategies are briefly discussed.

Author's address :-

Environmental Archaeology Unit
University of York
Heslington
York
YO1 5DD

0904 430000 x5849

On the relationship between the number of taxa and the number of specimens identified in archaeological bone assemblages.

Introduction

The increasing use of sieving to recover bones from soil samples, coupled with a greater need to justify the extent of studies of archaeological bone assemblages, has prompted closer investigation of the quantity of material which has to be examined in order to obtain that improbable thing, the 'representative sample'. Because the examination of a sample of bones leads to the accumulation of several different datasets, with any one bone contributing to one or to several of these datasets, it is more useful to consider the problem of what constitutes a representative sample of bones for each different category of information which is sought. Blanket use of sampling concepts irrespective of the actual question being asked is at best unhelpful and at worst thoroughly misleading.

This short paper thus considers just one topic; the relationship between the numbers of specimens identified in a sample, and the number of taxa which might typically be expected to be identified amongst that number of bones. Although the mere recording of more and more taxa is not the main purpose of studying ancient bones, there is much information in the areas of wild resource exploitation, seasonal occupation, and vermin commensalism which can only be investigated if the limits to the recovery of infrequent taxa are properly understood. Comparisons of 'diverse' and 'restricted' assemblages cannot be made unless sample size can be excluded as a factor restricting diversity in the identified assemblage. Without over-stating the 'stamp-collecting' of rare taxa, then, there are important reasons for being able to predict, if only approximately, what size of sample would impose what limits on the recovery of the taxon diversity of the original sampled population.

Method

The mathematical relationship between the number of taxa identified (henceforth S) and the number of identified specimens in the sample (N) was examined by Casteel (1979). Working from first principles, Casteel showed that the two variables should be linked by an equation of the general form $S = a.N^b$, where a would approach unity and b would lie between 0 and 1. He then used data obtained from results published by Parmalee and Shane (1970) to derive a predictive equation

$$S = 1.08.N^{0.54} \qquad \text{--- (1)}$$

which described a statistically significant relationship between S and N for one particular assemblage. It is not stated in Casteel's work (indeed it does not seem to have been considered relevant) whether the assemblage was recovered by hand-collection or sieving, or to what level taxonomic identification proceeded. Obviously, the potential diversity of the identified sample depends in part upon the diversity of the sampled population, and in part on whether, for example, a number of rodent and insectivore bones are recorded as one taxon (small mammal), as three taxa (mice; voles; shrews), or as seven different species.

Following on from Casteel's work, it was proposed to construct general equations linking S and N for several different bone assemblages from York sites. The aim was to see how far the values obtained for a and b resembled those obtained by Casteel, and to what extent such equations could be used as an aid to developing sampling strategies on future excavations. The assemblages which were chosen were:

1/ Hand-collected bones from Roman deposits at 24-30 Tanner Row (see O'Connor 1985). This assemblage was well-preserved, but recovered under difficult conditions (i.e. waterlogged organic sediments, winter light and weather).
2/ Hand-collected bones from Roman deposits at the City Garage, Blake Street. Though not so obviously well-preserved as the bones from Tanner Row, recovery was, judging by the incidence of small bones in the recovered assemblages, quite good.
3/ Hand-collected bones from Area II of the Bedern excavation (Scott 1985). These bones were well-preserved, and the frequent records of deer and hare bones were expected to give the characteristics of a diverse assemblage.
4/ Bones from Roman deposits at 24-30 Tanner Row recovered by wet-sieving large soil samples to 1mm, then dry-sieving the residue to 2mm and recovering bones from the 2+mm retent (Kenward, Hall and Jones 1980). Because sieving will have permitted the recovery of a wider range of taxa, this assemblage should show the characteristics of a much higher potential diversity than the hand-collected groups.

The three hand-collected assemblages were used in this analysis at the same level of taxonomic precision, viz. all fish bones were attributed to one taxon (fish) and bird bones to three taxa (fowl; goose; others). The sieved assemblage was identified to species, but all fish bones were excluded.

For each assemblage, a record was made of the number of bones identified and the number of taxa identified in each sampling unit. For the hand-collected assemblages, the sampling unit was the excavated context; thus the record comprised values of S and N for each of dozens of contexts. For the sieved assemblage, the sampling unit was the soil sample, giving values of S and N for each of 110 samples.

The data were analysed by means of the statistical package Microtab, using a BBC model B microcomputer. The required equations were obtained by performing linear regression analyses on log-10 transformed data. Goodness-of-fit was generally poor but statistically acceptable.

Results

The equations and correlations obtained from the four assemblages were as follows:

Tanner Row hand-collected

$$S = 1.24.N^{0.36} \qquad \text{--- (2)}$$

Correlation logS : logN = 0.815, n=99

Blake Street hand-collected

$$S = 1.11.N^{0.40} \qquad \text{--- (3)}$$

Correlation logS : logN = 0.808, n=79

Bedern hand-collected

$$S = 1.20.N^{0.43} \qquad \text{--- (4)}$$

Correlation logS : logN = 0.869, n=86

Tanner Row sieved

$$S = 1.18.N^{0.54} \qquad --- (5)$$

Correlation logS : logN = 0.894, n=110

It can be seen that the values obtained for the constant and coefficient are close to those in equation 1. The coefficient serves as a measure of potential diversity (i.e. the higher the coefficient, the more taxa predicted per given number of specimens), and in this respect the best match with Casteel's American data is the sieved assemblage from Tanner Row. Fig. 1 plots the curves to these equations.

Discussion

The results confirm Casteel's observation that there is a statistically significant non-linear relationship between S and N. However, no single curve can be used to predict an optimum value for N from the required value of S. Instead, a range of coefficients must be used, selecting a value of N from the range thus obtained according to *a priori* assessment of the likely diversity of the sampled population. Taking equations 2 and 4 as examples of the range of curves likely to be encountered in British urban hand-collected material, it is clear that beyond a mean sample size of 50 identified bones per context, further taxa are added but slowly. At N = 50, S ranges from 5 to 6.5. At N = 100, S = increases to 6.5 to 8.5, and at N = 200, S rises to 8.5 to 11.5. Thus quadrupling the sample size will increase the average number of taxa by a factor of about 1.7; hardly a greatly increased information return for the increased investment of time and effort. Increasing the mean sample size to N = 500 gives a range of values of S from 11.6 to 17.4. The predicted number of taxa has more than doubled, but in return for ten times the amount of work.

Turning to equation 5, the graph may be used to explore optimum soil sample size, although this will be to pile averages upon averages. The mean concentration of identified bone in soil samples from Roman deposits at Tanner Row was 0.698 bones/kg, equivalent to 1 identified bone per 1.432kg. It can thus be predicted that a series of samples of mean weight 30kg would yield a mean value of N = 20.9, and mean S = 6.1. Doubling the mean sample weight would be predicted to increase mean S to 8.9. It is important, therefore, that in any consideration of the diversity of small mammal or bird taxa in assemblages obtained by sieving, the mean sample weights of each group of samples in the comparison should not differ by more than a few kilograms, the critical margin for error obviously increasing with the mean sample weight. In real life, the factors which determine mean sample weight are likely to have as much to do with available man-power on site and the volume of sample containers as with mathematically-predicted yield. A knowledge of the relationship between sample size and taxon yield can at least add a degree of objectivity to an assessment of such a pragmatically constrained sampling strategy.

References

Casteel, R.W. (1979). Taxonomic abundance and sample size in archaeological faunal assemblages. Pp.129-136 in M. Kubasiewicz (ed.) Archaeozoology. Agricultural Academy, Szczecin.

Kenward, H.K., Hall, A.R. and Jones, A.K.G. (1980). A tested set of techniques for the extraction of plant and animal macrofossils from waterlogged archaeological deposits. Science and Archaeology 22, 3-15.

O'Connor, T. P. (1985). Hand-collected bones from Roman to medieval deposits at the General Accident Extension site, York (1983-4.32). Ancient Monuments Laboratory Report 4723.

Parmalee, P. and Shane, O.C. (1970). The Blain site vertebrate fauna. Pp. 185-206 in Prufer, O.H. and Shane, O.C. (eds.) Blain village and the Fort Ancient tradition in Ohio. Kent State University Press.

Scott, S.A. (1985). Bones from the Bedern, York. Ancient Monuments Laboratory Report 4821.

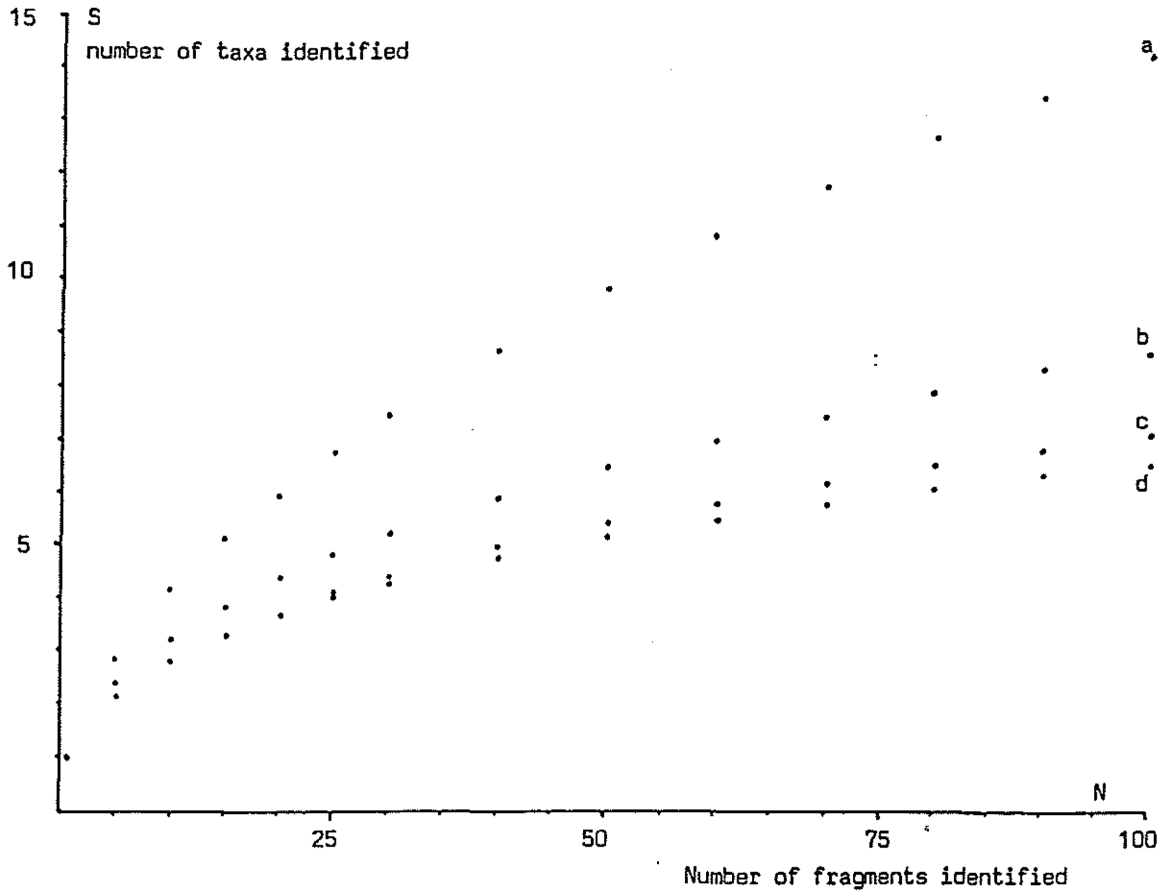

Fig. 1. Graphs plotted to equations 2 to 5.

a/ Tanner Row sieved bones	$S = 1.18\ N^{0.54}$
b/ Bedern hand-collected bones	$S = 1.20\ N^{0.43}$
c/ Blake Street hand-collected bones	$S = 1.11\ N^{0.40}$
d/ Tanner Row hand-collected bones	$S = 1.24\ N^{0.36}$