Dingwall, L., S. Exon, V. Gaffney, S. Laflin and M. van Leusen (eds.) 1999. Archaeology in the Age of the Internet. CAA97. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 25th Anniversary Conference, University of Birmingham, April 1997 (BAR International Series 750, CD-ROM). Archaeopress, Oxford.

Testing Multivariate Normality, with Applications to Lead Isotope Data Analysis in Archaeology

M.J. Baxter

Abstract

Applications of multivariate analysis in archaeology often either assume that the data have a multivariate normal distribution, or work better if this is the case. The assumption is rarely tested. In this paper a test of multivariate normality based on a multivariate kernel density estimate, developed independently by Bowman and Foster (1993) and Henze and Zirkler (1990), is applied to trivariate lead isotope ratio data. There is clear evidence of non-normality, violating an assumption that is common in the statistical analysis of such data. The paper both illustrates the utility of kernel density estimates in a manner not previously demonstrated in the archaeological literature, and challenges an assumption that is prevalent in the field of lead isotope ratio analysis.

1 Introduction

In many applications of multivariate statistics in archaeology the normality of the data is either assumed or considered to be desirable. For example, in archaeometric studies of artefacts based on their chemical compositions, elemental concentrations are often logarithmically transformed in the hope that this will make the data more normally distributed. This is done because of a belief that trace elements have a natural log-normal distribution in nature. Similarly, in studies that use trivariate lead isotope ratio data for provenancing, some of the techniques in common use assume that the data are trivariate normal (Sayre *et al*. 1992).

Given the prevalence of the normality assumption it is surprisingly hard to find papers where the assumption is tested, either formally or informally. This is particularly true for multivariate data, even though over 50 tests of multivariate normality exist (Looney 1995). The prime purpose of the present paper is to describe and illustrate a recent approach to testing for multivariate normality that seems promising.

A version of this paper was delivered at the 1997 CAA conference in a session whose theme was the use of kernel density estimates (KDEs) in archaeology. Accordingly, this paper will examine tests of normality that are directly based on KDEs. Although kernel density estimation is an established statistical technique it has had little application in archaeology. Baxter *et al*. (1997) give examples of primarily bivariate and descriptive uses. The present paper shows how KDEs may be exploited in an inferential fashion.

The next section gives details of the theory, followed in later sections by illustrative applications and a discussion.

2 Using KDEs for testing normality

The basic idea behind using KDEs for testing normality is a simple one that has been proposed independently, and in different guises, by various authors. The outline given here follows Bowman (1992) and Bowman and Foster (1993).

For exposition the univariate case is first described, specialising to the case of KDEs with normal kernels. It is assumed throughout that the data are first standardised to have zero mean and unit variance. A KDE can be obtained by centring ‘bumps’ about the data points, displayed on a line graph, and summing the heights of the bumps at each point on the line. For present purposes the bumps are taken to be normal probability distributions of the form

$$(2\pi h^2)^{-1/2} \exp\{-x^2/(2h^2)\}$$

where the standard deviation, h, is a smoothing parameter that determines the spread of the bump and the precise shape of the KDE.

The test of univariate normality described in Bowman (1992) is based on the 'distance' between the KDE and the expected normal density that would arise if the data were sampled from a normal distribution. The only slightly tricky part is that the KDE is a biased estimate of the true density, and the bias depends on h. It can be shown that the expected density at x is $N(x, 1 + h^2)$, where $N(x - \mu, \sigma^2)$is the normal density function in x with mean μ and variance σ^2 (Bowman 1992, 5). Bowman defines distance in terms of the integrated squared error (ISE) and gives an analytic expression for the ISE statistic

$$\int \{N(x,1+h^2)-\hat{f}(x)\}^2$$

where $\hat{f}(x)$ is the KDE.

Using exactly the same ideas Bowman and Foster (1993) develop the ISE statistic for the multivariate case and this results in a test statistic that has the analytic form

$$N(0,2(1+h^2)I) - 2\sum_i N(x_i, (1+2h^2)I)/n$$
$$+\ N(0,2h^2 I)/n + 2\sum_{i<j} N(x_i - x_j, 2h^2 I)/n^2$$

Where $N(a,\sum)$ is the density of the p-variate normal distribution with covariance matrix $\sum$ evaluated at the p-vector a, and I is the p by p identity matrix. In this approach the data are first 'sphered' (equivalent to standardisation in the univariate case) to give standardised data vectors x_i, and h is the single smoothing parameter of the multivariate kernel applied to the sphered data.

Though it is not immediately obvious, the statistic given by Bowman and Foster (1993) is essentially the same as that given in Henze and Zirkler (1990) who derive it as 'a weighted integral of the squared modulus of the difference between the empirical characteristic function of the scaled residuals ... and its pointwise limit' under the hypothesis of multivariate normality. Henze and Zirkler note that their statistic can be expressed in terms of KDEs. Their work is a generalisation of earlier results of Baringhaus and Henze (1988), motivated by consideration of the empirical characteristic function.

For the test statistic to be usable critical values are needed, and these depend on the choice of h. Bowman and Foster (1993) give 5% critical values for

$$h = \{4/(p+2)n\}^{1/(p+4)}$$

which is asymptotically optimal. Henze and Zirkler (1990) also give results for this (however the value they give for h in page 3600 of their paper is incorrect) and other values. The earlier Baringhaus and Henze (1988) paper gives critical values for the case $h = 1/\surd 2$.

3 Illustrative example

The Bowman and Foster (1993) and Baringhaus and Henze (1988) statistics have been used, along with several other tests of multivariate normality, by Baxter (1998), to study the normality of lead isotope ratio fields. The analysis of such data has attracted considerable recent controversy (Budd et al., 1995; Stos-Gale et al., 1997) that will not be entered into here.

Very simplistically, a sample from an ore-body can be characterised by three lead isotope ratios, $^{208}Pb/^{206}Pb$, $^{207}Pb/^{206}Pb$ and $^{206}Pb/^{204}Pb$, which define a point in three-dimensional space. Given a sufficient number of samples (20 is often stated) a cloud of points in three-dimensional space is obtained that can be used to estimate the lead isotope field for the data. Artefacts, such as Late Bronze Age oxhide ingots, can also be characterised by their lead isotope ratios and compared to the different fields that might have been a source for the ore used in their manufacture.

Some proponents of the use of lead isotope analysis (e.g. Sayre *et al.* 1992) make extensive use of the assumption that lead isotope fields have a normal distribution, either to define the true extent of a field or to assess the probability that an artefact comes from a field. The results in Baxter (1998) suggest that normality may be the exception rather than the rule, and this contradicts what is often asserted in the literature.

To give a flavour of the use of tests of normality, an analysis of one of the data sets used in Baxter (1998), but not presented in detail there, is given. These data consist of 59 observations from the Lavrion field, given in Stos-Gale *et al.* (1996). Two cases, 8 and 57, can be identified as outliers in three-dimensional space using brush-and-spin plots. Interestingly neither case (which can also be identified as outliers using more formal tests; J. Whittaker, pers. comm.) is a uni- or bi-variate outlier. These cases are omitted from the analysis here. This treatment differs from that of previous analyses, and will allow the possibility that previous indications of non-normality are attributable to the outliers to be examined.

The ISE statistic is .00292 (compared to .0033 for the full data set) and is significant (just) at the 1% level. The statistic can also be calculated for bivariate pairs of ratios and yields values of .00569, .00838 and .00619. All are significant at the 5% level, with the second pair based on the first and third ratios, the most non-normal. The univariate ISE statistics for the individual ratios are .0105, .0072 and .0123. The first and third of these are significant at 1%; the second is not quite significant at 5%. In summary, there is strong evidence to suggest that the field for Lavrion does not have a mutivariate normal distribution; it is the first and third ratios that seem to be most involved in the non-normality, of which the latter is individually and clearly non-normal.

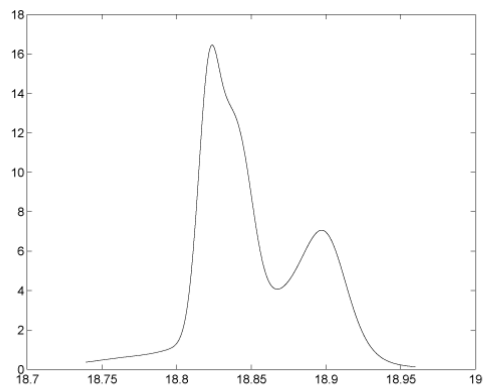

Figure 1. An adaptive kernel density estimate of the distribution of the $^{206}Pb/^{204}Pb$ ratio for the Lavrion data, after omitting two multivariate outliers.

In general, if non-normality is established, projection pursuit methods can be used to identify linear combinations of the variables that are particularly non-normal. To understand the nature of the non-normality the distribution of these linear combinations can be examined in a variety of ways, including the use of KDEs. An illustration of this general approach is given in Baxter (1998). In the present case it is simplest to illustrate the idea using the third ratio only, as this is clearly non-normal.

Figure 1 shows a KDE for the third ratio. An adaptive KDE has been used in which a pilot estimate of h, the smoothing parameter, has been chosen using the 'solve the equation' methodology given in Wand and Jones (1995). The procedure outlined in Silverman (1986), implemented in the MATLAB package using routines available in Beardah and Baxter (1996), has been followed. The particular approach used will tend to smooth the tails and highlight areas of low density in the centre of the distribution. The present example suggests that non-normality is being detected because of multi-modality, and this is characteristic of other analyses presented elsewhere (Baxter 1998).

4 Discussion

This paper has illustrated the inferential use of a statistic based on KDEs to address a specific archaeological question - concerning the normality or otherwise of lead isotope ratio fields - that has attracted occasionally contentious debate. The paper also illustrates the use of KDEs in a descriptive role, to show the nature of the non-normality detected. Other descriptive applications in archaeology are given in Baxter et al. (1997) and Beardah and Baxter (1996), and references cited there.

The use of density estimation methodology in archaeology has been relatively limited to date. Five papers, with a range of applications, were originally offered for the 1997 CAA conference session devoted to the use of KDEs in archaeology (though, in the event, only three were delivered - see this volume). It is to be hoped that this is a sign that this useful methodology is beginning to be accepted as one of the many tools that should be routinely available to quantitatively minded archaeologists and their co-workers in other disciplines.

Bibliography

Baringhaus, L, and Henze, N, 1988 A consistent test for multivariate normality based on the empirical characteristic function, *Metrika,* 35, 339-348

Baxter, M J, 1998 On the multivariate normality of lead isotope fields, *Journal of Archaeological Science*

Baxter, M J, Beardah, C C and Wright, R V S, 1997 Some archaeological applications of kernel density estimates, *Journal of Archaeological Science,* 24, 347-354

Beardah, C C, and Baxter, M J, 1996 The archaeological use of kernel density estimates, *Internet Archaeology,* 1, (http://intarch.ac.uk/journal/issue1/beardah)

Bowman, A W, 1992 Density based tests of goodness-of-fit, *Journal of Statistical Computation and Simulation,* 40, 1-13
Bowman, A W and Foster, P J, 1993 Adaptive smoothing and density-based tests of multivariate normality, *Journal of the American Statistical Association,* 88, 529-537
Budd, P, Pollard, A M, Scaife, B and Thomas, R G, 1995 Oxhide ingots, recycling and the Mediterranean metals trade, *Journal of Mediterranean Archaeology,* 8, 1-32
Henze, N, Zirkler, B, 1990 A class of invariant consistent tests for multivariate normality, *Communications in Statistics: Theory and Methods,* 19, 3595-3617
Looney, S W, 1995 How to use tests for univariate normality to assess multivariate normality, *The American Statistician,* 49, 64-70
Sayre, E V, Yener, K A, Joel, E C and Barnes, I L, 1992 Statistical evaluation of the presently accumulated lead isotope data from Anatolia and surrounding regions, *Archaeometry,* 34, 73-105
Silverman, B, 1986 *Density Estimation for Statistics and Data Analysis,* Chapman and Hall, London
Stos-Gale, Z A, Gale, N H, and Annets, N, 1996 Lead isotope data from the Isotrace laboratory, Oxford: Archaeometry data base 3, ores from the Aegean, part 1, *Archaeometry,* 38, 381-390
Stos-Gale, Z A, Maliotis, G, Gale, N H and Annets, N, 1997 Lead isotope characteristics of the Cyprus copper ore deposits applied to provenance studies of copper oxhide ingots, *Archaeometry,* 39, 83-123
Wand, M P, and Jones, M C, 1995 *Kernel Smoothing,* Chapman and Hall, London

Contact details

M. J. Baxter
Dept of Mathematics
Statistics and Operational Research
The Nottingham Trent University, Clifton Campus
Nottingham NG11 8NS
UK
e-mail: mjb@maths.ntu.ac.uk