

How Much Information is there in an NMR Measurement?

Peter Belton

School of Chemical Sciences and Pharmacy, University of East Anglia, Norwich NR4
7TJ, UK

p.belton@uea.ac.uk

In many applications of analysis the problem is to classify samples into various sets. Typical of these types of analytical problems are authentication and metabolomics. The analytical results on the sample must therefore be compared with the members of different sets. In many cases NMR is the analytical method of choice, however often the choice of analytical method is determined by practical considerations such as equipment availability, the skills of the analyst and cost. A method is then developed which may or may not solve the problem. Instead of going through the lengthy trial and error processes is it possible to make some choices a priori about what methods are likely to be more successful on the basis of an objective analysis?

An approach to this problem is to consider the information content of the analytical method and compare with the information requirements to define the membership of a set. The amount of information in the spectrum can then be compared with that required to define membership of a set. The information theoretic concept of entropy proves to be a valuable way of comparing and classifying different measurement methods, however this needs some modification in order to better reflect the useful information content in analytical measurements. Within sets of data, two separate measures of entropy can be distinguished. The criteria for the quality of a data set can then be defined using these and an additional concept which measures the structure in the spectrum. Finally a probabilistic approach is used to define the membership of sets and to indicate the ways in which measurement methods may be chosen.