

Mixture method clustering of mixed three-mode three-way data with missing information

Lyn Hunt¹ and Kaye Basford²

¹Department of Statistics,
University of Waikato,
Private Bag 3105,
Hamilton, New Zealand
e-mail: lah@stats.waikato.ac.nz

²School of Land and Food Sciences,
The University of Queensland,
Brisbane Qld 4072, Australia
e-mail: k.e.basford@uq.edu.au

Keywords: Categorical and continuous data, missing values

Abstract

This abstract provides One difficulty with all classification studies is the unobserved or missing observations that occur in data sets. Hunt and Jorgensen (2003) implemented the mixture likelihood approach for clustering mixed two-mode two-way data where data are missing at random in the sense of Little and Rubin (1987). Hunt and Basford (2001) specified methodology that enabled the clustering of three-mode three-way data to include situations where not all attributes are observed for all individuals and where the component distributions are assumed to be multivariate normal. In this talk, we show that how this approach to clustering can be extended to incomplete mixed three-mode three-way data. We illustrate this approach by clustering the genotypes in a three-way data set where various attributes were measured on genotypes grown in several environments and where there is a moderate amount of missing data.

References

- Hunt, L.A. and Basford, K.E. (2001). Fitting a mixture model to three-mode three-way data with missing information. *Journal of Classification*, 18:209-226.
- Hunt, L.A. and Jorgensen, M.A. (2003). Mixture model clustering for mixed data with missing information. *Computational Statistics and Data Analysis*, 41:429-440.
- Little, R.J.A. and Rubin, D.B. (1987). *Statistical Analysis with Missing Data*. J. Wiley & Sons Ltd, New York.