A Generalisation of the Cressie-Read family of Statistics and Three-Way Correspondence Analysis for Archaeological Data

Lombardo R., Eric J. Beh, Gianmarco Alberti

Three-way correspondence analysis (CA) provides a quick and simple graphical summary of how categories and variables are related to one another (Carlier and Kroonenberg, 1996; Lombardo, Beh, and Kroonenberg, 2021; Lombardo, van de Velden, and Beh, 2023). Here, we present some variants of three-way CA based on a generalisation of the Cressie–Read family of divergence statistics (Cressie and Read, 1984; Pardo, 1996; Beh and Lombardo, 2024).

This family relies on a power parameter, denoted by δ; changes in δ lead to special cases within the family of divergences. For example, when $\delta = 1$, the statistic reduces to the three-way chi-squared statistic; when $\delta = -1/2$, it yields the Freeman–Tukey statistic; and when $\delta = 2/3$, it gives the Cressie–Read statistic.

In particular, our focus is on an ANOVA-like partitioning of these statistics and on comparing Pearson's three-way chi-squared statistic with the generalisations of the Freeman–Tukey and Cressie–Read statistics for three-way contingency tables.

Using archaeological data from Bellanger et al. (2021), these three-way CA variants are used to explore the distribution of artefacts in the Middle Loire Valley (France). We apply them to investigate a concrete archaeological question: did ceramic innovations diffuse at different rates across sub-regions of the middle Loire Valley, or did certain areas maintain distinct traditions while others converged over time?

References

Beh, E.J., Lombardo, R.,: Correspondence analysis and the Cressie-Read family of divergence statistics. International Statistical Review 92, 17—42 (2024)

Bellanger, L., Coulon, A., Husi, P.: Determination of cultural areas based on medieval pottery using an original divisive hierarchical clustering method with geographical constraint (MapClust). Journal of Archaeological Science 132, 105431 (2021).

Carlier, A., Kroonenberg, P. M.: Biplots and decompositions in two-way and three-way correspondence analysis. Psychometrika, 61, 355–373 (1996)

Cressie, N. A. C., Read, T. R. C.: Multinomial goodness-of-fit tests. Journal of the Royal Statistical Society, Series B 46, 440–464 (1984)

Lombardo, R, Beh, E.J., Kroonenberg, P.: Symmetrical and non-symmetrical variants of three-way correspondence analysis for ordered variables. Statistical Science, 36 (4), 542 - 561 (2021)

Lombardo, R.; van de Velden, M.; Beh, E.J. Three-way correspondence analysis in R. R Journal 15, 237–262, (2023)

Pardo, M. C.: An empirical investigation of Cressie and Read tests for the hypothesis of independence in three-way contingency tables. Kybernetika, 32, 175–183 (1996)