

Marina Fragalá

Universidad Nacional de General Sarmiento, Argentina

mfragala@campus.ungs.edu.ar

El problema del análisis estadístico de datos composicionales sigue siendo una fuente de preocupación desde que en 1897 Karl Pearson pusiera de manifiesto la inadecuación de los métodos estadísticos clásicos para el estudio de los mismos. Los datos composicionales son realizaciones de vectores aleatorios positivos de suma constante. Suelen darse en forma de proporciones, porcentajes o concentraciones. Son habituales en ciencias aplicadas como biología, química, geología, economía, medicina, sociología, etc. Por eso es tan imprescindible disponer de herramientas adecuadas para su análisis.

Una posible distribución para las composiciones es la Dirichlet. Como los modelos de Dirichlet no siempre ajustan bien, Monique Graf (2020) propuso una generalización de dicha distribución, denominada distribución Beta Generalizada Simplicial (SGB). Esta distribución es lo suficientemente flexible como para adaptarse a muchas situaciones prácticas. La estimación por máxima verosimilitud y los modelos de regresión SGB fueron desarrollados por la misma autora.

En esta charla propondremos generalizaciones robustas con buenas propiedades asintóticas. Analizaremos cómo se comportan estos estimadores en escenarios de simulación con outliers y lo compararemos con el estimador clásico de Graf.

Trabajo en conjunto con Marina Valdora (Instituto de Cálculo, Universidad de Buenos Aires - Conicet) y Alfio Marazzi (Facultad de Biología y Medicina, Universidad de Lausanne, Suiza).

Referencias

- [1] Aitchison J. (1986). The statistical analysis of compositional data. Monographs on statistics and applied probability. Chapman and Hall Ltd (reprinted 2003 with additional material by the Blackburn Press, London (UK).
- [2] Aitchison J. (2003). The Statistical Analysis of Compositional Data. The Blackburn Press, Caldwell, NJ.
- [3] García Ben M., Martínez E., Yohai V.J. (2006). Robust estimation for the multivariate linear model based on a Tau-scale. Journal of Multivariate Analysis 97, 1600 - 1622.
- [4] Graf M. (2020). Regression for compositions based on a generalization of the Dirichlet distribution. Statistical Methods Applications.
- [5] Graf M. (2020). SGB: Simplicial Generalized Beta Regression, R package.
- [6] Marazzi A., Valdora M., Yohai V.J., Amiguet M. (2019). A robust conditional maximum likelihood estimator for generalized linear models with a dispersion parameter. Test. 28(1), 223-241.