

Modelo skew- t de coeficientes parcialmente variables (ST-PVCM)

Bladimir Morales

Departamento de Estadística, Universidad de Bío-Bío

Christian Caamaño

Departamento de Estadística, Universidad de Bío-Bío

Germán Ibacache-Pulgar

Instituto de Estadística, Universidad de Valparaíso

Resumen

Los modelos parciales con coeficientes variando (PVCMS) son una herramienta importante en la modelación de datos ambientales, económicos, financieros, biomédicos y otros, que tienen en su formulación un componente paramétrico y otro no paramétrico además de presentar interacción de las funciones suaves con otras covariables, que flexibiliza a los modelos de regresión clásicos, que tienen usualmente en el componente aleatorio (errores) distribución gaussiana, pero en muchas ocasiones los datos presentan asimetría y datos atípicos (colas pesadas). El objetivo de este trabajo es extender los PVCMS que permitan tener errores de distribución asimétrica t -Student denominada Skew- t presentada por Azzalini y Capitanio (2003), Gupta et al. (2002), Gupta (2003) incrementando la flexibilidad del modelo. Específicamente se desarrolló el algoritmo de Esperanza-Maximización (EM) propuesto por Dempster et al. (1977) y el análisis de diagnóstico mediante influencia local, así como el apalancamiento generalizado siguiendo a Zhu y Lee (2001). Luego para evaluar la eficiencia de la estimación del algoritmo EM se realizó un estudio de simulación con la implementación computacional en el software R. Finalmente se aplicó a un conjunto de datos del Sistema de Información Nacional de Calidad del Aire (SINCA) de Chile, específicamente a datos de la Estación de Pudahuel y Coronel Norte mostrando la utilidad del modelo ST-PVCM.

Referencias Bibliográficas

1. Azzalini, A., Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. *Journal of the Royal Statistical Society, Series B* **65**, 367-389.
2. Dempster, A., Laird N., Rubin D. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B* **39**, 1-38.
3. Ferreira, C., Paula, G. (2016). Estimation and diagnostic for skew-normal partially linear models. *Journal of Applied Statistics* **44**, 1-21.
4. Gupta A., Chang F., Huang W. (2002). Some skew-symmetric models. *Random Operators and Stochastic Equations* **10**, 133-140.
5. Gupta A. (2003). Multivariate skew t distribution. *Statistics* **37**, 359-369.
6. Ibacache-Pulgar G., Paula G. (2011). Local influence for Student- t partially linear models *Computational Statistics and Data Analysis* **55**, 1462-1478.
7. Ibacache-Pulgar G., Paula G., Galea M. (2012). Influence diagnostics for elliptical semiparametric mixed models. *Statistical Modelling* **12**, 165-193.
8. Zhu H., Lee S. (2001). Local influence for incomplete-data models. *Journal of the Royal Statistical Society, Series B* **63**, 111-126.