

Inferencia ecológica: descifrando el comportamiento individual de los votantes desde datos agregados

Duración: 8 horas

Formato: En persona y en línea

Instructor: Jose M. Pavía

Descripción del curso:

Politólogos, medios de comunicación y partidos políticos, entre otros agentes, están interesados en conocer el comportamiento electoral de diferentes subgrupos de población y en disponer de estimaciones de las matrices de transferencia de voto a nivel global y local. La inferencia ecológica permite dar respuesta a este problema explotando datos agregados. Este curso tiene como principal objetivo formar en el uso del paquete `lphom` de R y sus múltiples funciones como instrumento para responder las preguntas anteriores.

En este curso, ofertado por RECSM, se va a mostrar las técnicas más avanzadas para la realización de dichas inferencias, centrándose en el uso del paquete `lphom` de R. El curso constará de dos partes. En la primera, tras una breve introducción teórica y repaso de los principales paquetes disponibles en R, se detallan las ventajas de `lphom` mostrándose su sencillez de uso mediante el desarrollo de varios ejemplos con una de sus principales funciones. Se presentan los principales outputs que ofrece y se indaga en las posibilidades que ofrecen sus múltiples argumentos. En la segunda parte, se profundiza en el problema y en sus soluciones. Se presentan nuevas funciones y argumentos, con sus algoritmos subyacentes, mostrando también las mejoras de precisión que se obtienen mediante (i) la introducción de información a priori (datos de encuesta), (ii) la clusterización automática de unidades, o (iii) el uso de algoritmos de inteligencia artificial (`bagging`, `genetic boosting` y aprendizaje reforzado).

Horario:

Día 1	14.00 -16.00	Inferencia ecológica, tipos de problemas, especificación del problema, la estimación de las matrices de transferencia de voto, el paquete <code>lphom</code> , otros paquetes en R, desarrollo de un ejemplo.
	16.00 - 16.15	Pausa
	16.15 - 18.00	Pre-procesamiento de datos (externo e interno), argumentos de las funciones, estructura y contenido de los outputs.

		Fuente de datos electorales en España: GIPEyOP. Herramientas auxiliares: sc2sc.
Día 2	14.00 - 15.45	Introducción de información a priori (datos de encuesta), elecciones simultáneas, el enfoque bottom-up, clusterización, errores de estimación.
	15.45 - 16.15	Pausa
	16.15 - 18.00	Soluciones basadas en aprendizaje automático: bagging, genetic boosting, aprendizaje reforzado. Funciones <code>_gb</code> and <code>_rl</code> .

Prerequisites: Experiencia en uso de R

Software: R y RStudio

Lectures:

Pavía, JM and Romero, R "[Improving estimates accuracy of voter transitions. Two new algorithms for ecological inference based on linear programming](#)", **Sociological Methods and Research**, *online available*.

Pavía, JM and Romero, R (2023) "[Data Wrangling, Computational Burden, Automation, Robustness and Accuracy in Ecological Inference Forecasting of R×C Tables](#)", **SORT - Statistics and Operations Research Transactions**, 47(1), 151-186.

Pavía, JM (2023) "[Adjustment of initial estimates of voter transition probabilities to guarantee consistency and completeness](#)", **SN Social Sciences**, 3, 75.

Instructor short bio:



Jose M. Pavía, MSc in Maths and PhD in Economics and Business Science, is Quantitative Methods Professor in the Economics Faculty of the Universitat de Valencia, Spain. Pavía develops and applies statistical and machine learning methods in many areas of social sciences. With broad and varied research interests, his work focuses on the search for innovations that bridge the gap between theory and practical applications and include issues related to ecological inference, prediction, statistical

(machine) learning, electoral processes, survey research, sampling, public opinion, crime detection, behavioral economics, experiments, public policy evaluation, regional economy or inequality.

His main research has been published in journals like *Journal of the American Statistical Association*, *Journal of Statistical Software*, *Sociological Methods and Research*, *Environment and Planning A*, *Journal of the Royal Statistical Society*, *International Journal of Forecasting*, *European Urban and Regional Studies*, *Technological and Economic Development of Economy*, *Regional Studies*, or *Geographical Analysis*. More info at <http://go.uv.es/ZfX6E7w>.