

Desafíos teórico-metodológicos en la construcción del espacio social argentino

Gabriela Heredia Baek, IDH

Resumen

En continuidad con el trabajo realizado desde el equipo Estrategias de Reproducción Social en Gran Córdoba, se propone como temática general reconstruir el espacio social de clases de Argentina a partir de los datos provenientes de la Encuesta Nacional sobre la Estructura Social (ENES).

Como antecedente se cuenta con el documento de trabajo nro. 1 (Gutiérrez y Mansilla, 2016) en el que se ha propuesto construir el espacio social de clases cordobés utilizando el análisis multidimensional de datos, en particular el análisis de correspondencia múltiple (ACM) y la Clasificación Jerárquica Ascendente (CJA) basada en los datos de la EPH los terceros trimestres de 2003 y 2011. A partir de una selección de 51 modalidades activas, se obtiene un espacio en el que los dos primeros factores representan aproximadamente un 13% de la inercia total.

En esta dirección se propone aportar metodológicamente a la tarea de agrupamiento de las unidades de análisis aprovechando la multiplicidad de datos disponibles.

Resumen ampliado

Introducción

En la presente ponencia se plantea la construcción del espacio social de Argentina a partir de los datos de la Encuesta Nacional sobre la Estructura Social (ENES) realizada en el marco del Programa de Investigación sobre la Sociedad Argentina Contemporánea (Pisac).

Esta encuesta fue llevada a cabo en los años 2014-2015 con el objeto de abordar la heterogeneidad social argentina desde una mirada estructural y multidimensional y de dar cuenta de las condiciones de vida de los hogares (Salvia y Piovani, 2018).

Sobre la noción de espacio social

La construcción multidimensional del espacio social constituye la primera parte o momento objetivista del análisis de la desigualdad social. A partir de este es posible reconstruir el espacio

de posiciones en los que se insertan las distintas clases de agentes para luego, en un segundo momento poder dar cuenta de los sentidos vividos de los agentes, de las prácticas concretas que estos desarrollan para reproducir sus posiciones (Gutiérrez et al., 2021).

La construcción del espacio de clases a partir de datos provenientes de encuestas supone una serie de decisiones a nivel teórico-metodológico. En el presente trabajo no pretendemos cubrir cada una de estas sino hacer alusión a algunos de los obstáculos que se presentan a la hora de modelar el espacio de posiciones argentino.

Sobre los hogares como unidades de análisis

En numerosos trabajos empíricos que se ocupan del estudio de la estructura social y, en particular, del análisis de las clases sociales, se consideran como unidades de análisis los hogares o unidades familiares/domésticas argentinas (Torrado, 1982; Gutiérrez y Mansilla, 2016; Maceira, 2015). En este contexto se entiende que en el seno de las familias es donde “se puede asegurar su reproducción biológica, preservar la vida y desarrollar todas aquellas prácticas económicas y no económicas, indispensables para la optimización de las condiciones materiales y no materiales de existencia de la unidad y de cada uno de sus miembros” (Torrado, 1982: 3-4). En el caso particular de la ENES se asume que la caracterización de los miembros del hogar está mediada por la posición social del jefe del hogar (Maceira, 2015). A diferencia de estas perspectivas, planteamos considerar como unidad de análisis a todo agente económicamente activo mayor de 10 años de edad. Una de las principales implicancias de esta decisión es que se hace hincapié en la participación en el mundo del trabajo –o trabajo extra-doméstico- ignorando el valor de la participación de los miembros en el trabajo dentro de los hogares..

Los valores faltantes

Uno de los problemas comunes en el procesamiento de los datos es la presencia de valores nulos/faltantes. Ante su presencia, muchas de las técnicas de análisis estadístico no pueden procesar la información o bien ven sus resultados afectados por el sesgo que estos valores pueden introducir. La presencia de no respuesta puede ser tratada de distintas maneras, desde la eliminación de las filas que presenten valores perdidos hasta la utilización de técnicas de aprendizaje automático para el reemplazo de valores faltantes (Rosati, 2021). En el caso particular de la implementación del análisis de correspondencias múltiples (ACM) en R, existen paquetes que imputan los valores sin que estos pesen en los resultados.

ACM: Variables suplementarias y activas

Para transformar los datos en un espacio de representación menor a partir del ACM, resulta imprescindible definir qué variables serán consideradas para la construcción del espacio – variables activas- y cuáles serán consideradas variables ilustrativas/suplementarias. Al mismo tiempo se deben evaluar maneras en las que se puede mejorar la inercia representada por los primeros factores para lograr una buena aproximación gráfica de los datos originales (Baranger, 1992)

Clusterización: La cantidad de grupos (clases y fragmentos de clases)

Para algunos de los algoritmos de clusterización como el K-medias, se deben conocer previamente la cantidad de clusters en los que se deben particionar los datos (Géron, 2022). Para conocer la cantidad óptima de clusters existen métodos como el método del codo (Elbow method) o el coeficiente de silueta (Sillhouette score). A pesar de que existan este tipo de herramientas, es necesario evaluar alternativas capaces de definir clusters automáticamente en función de los datos.

Bibliografía

Baranger, D. (1992). Construcción y análisis de datos. *Posadas, UnaM*

Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc."

Gutiérrez, A., & Mansilla, H. (2016). El espacio social de las clases y los instrumentos de reproducción social. *Dinámicas de mercado de trabajo, el mercado de las políticas sociales, el mercado escolar y el mercado habitacional. Gran Córdoba, 20032011*.

Gutiérrez, A., Mansilla, H. & Assusa, G. (2021) *De la grieta a las brechas: pistas para estudiar las desigualdades en nuestras sociedades contemporáneas*. Eduvim

Maceira, V. (2015). Un abordaje teórico-metodológico para la investigación de la estructura, la movilidad social y las condiciones de vida: la propuesta ENES-PISAC. *Revista Latinoamericana de Metodología de las Ciencias Sociales*, 5.

Piovani, J. I., & Salvia, A. (2018). *La Argentina en el siglo XXI: cómo somos, vivimos y convivimos en una sociedad desigual: Encuesta nacional sobre la estructura social*. Siglo Veintiuno Editores.

Rosati, G. (2021). Métodos de Machine Learning como alternativa para la imputación de datos perdidos. Un ejercicio en base a la Encuesta Permanente de Hogares. *Estudios del trabajo*, (61), 122-145.

Torrado, Susana (1982) "El enfoque de las estrategias familiares de vida en América Latina. Orientaciones teórico metodológicas". Cuaderno del CEUR no2.