



1. DATOS BÁSICOS DEL TFG:

Título: Técnicas de análisis multivariante en muestreos con diseños complejos.

Descripción general (resumen y metodología):

Un diseño muestral se considera complejo cuando las probabilidades de obtener cada muestra no son uniformes o cuando el espacio muestral no abarca todas las muestras de un determinado tamaño que se podrían obtener de la población. En este tipo de diseños, surge la importancia de utilizar estimadores que tengan en cuenta la complejidad del diseño para obtener estimaciones lo más precisas posible. Es habitual utilizar estimadores que ponderan cada unidad a través de los pesos de diseño para así obtener estimaciones de parámetros puntuales (media, total, proporción). Sin embargo, su aplicación en las técnicas de análisis multivariante, como los modelos predictivos, los métodos de reducción de la dimensionalidad o los métodos de clasificación no supervisada no ha sido tan estudiada, a pesar de que reflejar la complejidad del muestreo también podría favorecer una mejor representación de las relaciones existentes entre individuos.

En este trabajo se abordará la incorporación de información sobre el diseño muestral a los principales métodos de análisis multivariante. Para ello, se realizará una encuesta probabilística en el contexto de una institución de enseñanza superior, de cara a obtener información sobre posibles similitudes entre grupos de individuos o la predicción de diversas variables de interés.

Tipología: Estudio de casos, teóricos o prácticos, relacionados con la temática del Grado.

Objetivos planteados:

El objetivo principal de este trabajo es el estudio y la aplicación de técnicas de análisis multivariante adaptadas a diseños muestrales complejos. Los objetivos específicos son los siguientes:

OE1. Estudio de la literatura referente a las técnicas de análisis multivariante y la incorporación de información sobre el diseño muestral: modelos predictivos ponderados, técnicas de reducción de dimensiones y de agrupamiento ponderadas.

OE2. Realización de una encuesta probabilística en el contexto de una institución educativa: diseño del cuestionario, pilotaje, trabajo de campo, depuración, ajustes y análisis final.

OE3. Estudio de relaciones entre individuos, grupos de individuos y predicción de variables en función de la información auxiliar disponible.

Bibliografía básica:

- Peña, D. (2013). Análisis de datos multivariantes. Cambridge: McGraw-Hill España.
- Hashemi, M., & Karimi, H. (2018). Weighted machine learning. Statistics, Optimization and Information Computing, 6(4), 497-525.
- Delchambre, L. (2015). Weighted principal component analysis: a weighted covariance eigendecomposition approach. Monthly Notices of the Royal Astronomical Society, 446(4), 3545-3555.
- Lohr, S. L. (2021). Sampling: design and analysis. Chapman and Hall/CRC.
- Ackerman, M., Ben-David, S., Brânzei, S., & Loker, D. (2012). Weighted clustering. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 26, No. 1, pp. 858-863).

Recomendaciones y orientaciones para el estudiante:

Se recomienda que el estudiante haya cursado las materias Muestreo Estadístico, Diseño de Encuestas, Diseño de Experimentos y Modelos Lineales del Grado en Estadística.

Plazas: 1

2. DATOS DEL TUTOR/A:

Nombre y apellidos: RAMÓN FERRI GARCÍA

Ámbito de conocimiento/Departamento: ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

Correo electrónico: rferri@ugr.es

3. COTUTOR/A DE LA UGR (en su caso):

Nombre y apellidos: Jorge Luis Rueda

Ámbito de conocimiento/Departamento: ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

Correo electrónico: jorgerueda@ugr.es

4. COTUTOR/A EXTERNO/A (en su caso):

Nombre y apellidos:

Correo electrónico:

Nombre de la empresa o institución:

Dirección postal:

Puesto del tutor en la empresa o institución:

Centro de convenio Externo:

5. DATOS DEL ESTUDIANTE:

Nombre y apellidos: ROCIO ALABARCE PEREZ

Correo electrónico: rocioalabarce@correo.ugr.es