



UNIVERSIDADE DA CORUÑA *Universidade de Vigo*

PROGRAMA OFICIAL DE POSTGRADO EN ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

DATOS IDENTIFICATIVOS DE LA MATERIA

Código de la materia: P1061211

Nombre de la materia: Análisis Multivariante

Número de créditos ECTS: 5

Curso académico: 2018/2019

Profesorado:

Beatriz Pateiro López (5 créditos)

OBJETIVOS DE LA MATERIA

En esta materia se pretende que el alumno aprenda las técnicas más importantes en el análisis de datos multidimensionales. Se abordan tanto los aspectos metodológicos como del uso de software que proporcionan los paquetes estadísticos en el campo del análisis multivariante.

CONTENIDOS DE LA MATERIA

Tema 1. Introducción al análisis multivariante.

Conceptos básicos de álgebra matricial para el tratamiento de datos multivariantes. Descripción de datos multivariantes: matriz de datos, vector de medias, matriz de covarianzas y matriz de correlaciones. Medidas de proximidad. Representaciones gráficas.

Tema 2. Inferencia en poblaciones normales multivariantes.

Inferencia sobre la media y la matriz de covarianzas de una población normal. Regiones de confianza y comparaciones simultáneas. Comparación de poblaciones normales multivariantes. Contraste de la normalidad multivariante.

Tema 3. El modelo lineal multivariante.

Presentación del modelo, estimación de los parámetros e inferencia sobre los parámetros. Contraste de restricciones lineales. Predicción.

Tema 4. Análisis multivariante de la varianza.

MANOVA de un factor. Presentación del modelo. La tabla de descomposición de la variabilidad, contraste de igualdad de medias, comparaciones múltiples. MANOVA de dos factores. MANOVA de dos factores con interacción.

Tema 5. Análisis de componentes principales.

Descomposición de un vector aleatorio en sus componentes principales. Propiedades e interpretación. Variabilidad explicada por las componentes principales. Criterios para la selección del número de componentes principales. El biplot.

Tema 6. Análisis de correspondencias.

Expresión de la inercia de una tabla de contingencia a través de los perfiles de fila o de columna. Extracción de componentes. Representación simultánea de filas y columnas. Interpretaciones.

Tema 7. Análisis Discriminante.

Fundamentos del análisis discriminante. Análisis discriminante en dos poblaciones, caso particular de poblaciones normales. Generalización al caso de varias poblaciones.

Tema 8. Técnicas de formación de grupos.

Técnicas de agrupamiento jerárquico. Distancias y similitudes entre individuos. Distancias entre grupos. Métodos de particionamiento: método de las k-medias. Determinación del número de grupos.

BIBLIOGRAFÍA BÁSICA Y COMPLEMENTARIA

Bibliografía básica

Everitt, B.S. (2005). An R and S-Plus companion to multivariate analysis. Springer.

Hardle, W.K. y Simar, L. (2015). Applied multivariate statistical analysis. Fourth Edition. Springer.

Johnson, R.A. y Wichern, D.W. (2007). Applied multivariate statistical analysis. Pearson Education.

Mardia, K.V., Kent, J.T. y Bibby, J.M. (1979). Multivariate analysis. Academic Press.

Bibliografía Complementaria

Everitt, B.S. y Dunn, G. (2001). Applied multivariate data analysis. Hodder Education.

Hastie, T., Tibshirani, R. y Friedman, J. (2009). The elements of statistical learning. Springer.

Koch, I. (2014). Analysis of multivariate and high-dimensional data. Cambridge.

Peña, D. (2002). Análisis de datos multivariantes. McGraw-Hill.
Pérez, C. (2004). Técnicas de análisis multivariante de datos. Pearson Educación, S.A.
Seber, G.A.F. (1984). Multivariate observations. Wiley.

COMPETENCIAS GENERALES Y ESPECÍFICAS

Competencias generales:

- [G6] Capacidad de identificar y resolver problemas
- [G3] Capacidad de trabajo en equipo y de forma autónoma
- [G13] Redacción de informes estadísticos con precisión, orden y claridad

Competencias específicas:

- [E2] La adquisición de los conocimientos de estadística e investigación operativa necesarios para la incorporación en equipos multidisciplinares.
- [E3] Capacidad para comprender, plantear, formular y resolver aquellos problemas susceptibles de ser abordados a través de métodos de análisis multivariante.
- [E84] Ser capaz de manejar diverso software (en particular R) e interpretar los resultados que proporcionan éstos en los correspondientes estudios prácticos.

METODOLOGÍA DOCENTE: ACTIVIDADES DE APRENDIZAJE Y SU VALORACIÓN EN CRÉDITOS ECTS

La actividad presencial del alumnado será de 35 horas entre docencia expositiva e interactiva, distribuidas en sesiones de dos horas. En la parte expositiva el profesorado hará uso de presentaciones multimedia, mientras que en la parte interactiva el alumnado resolverá, usando el software estadístico R, distintas cuestiones planteadas sobre los contenidos de la materia.

El alumnado dispondrá, a través del repositorio de material del que dispone la página web del propio programa, del material docente (presentaciones, apuntes, ejercicios) de la materia. A lo largo del curso se propondrán ejercicios/trabajos que los alumnos deberán resolver con la tutorización de los profesores. Esta tutorización será realizada tanto a través de medios virtuales (fundamentalmente correo electrónico) como de forma presencial en grupos reducidos

CRITERIOS Y MÉTODOS DE EVALUACIÓN

En la evaluación continua se tendrán en cuenta los trabajos realizados a lo largo del curso (30%) y el examen final, de problemas y cuestiones (70%). La resolución de los ejercicios/trabajos de evaluación continua permiten evaluar la adquisición de las competencias G3 y G13, así como las competencias específicas. Además de dichas competencias, en el examen final también se evalúa la adquisición de la competencia G13. El alumnado, si así lo manifiesta al comienzo del curso, tendrá el derecho a renunciar a la evaluación continua por lo que su calificación consistirá únicamente en un examen final teórico/práctico.

TIEMPO DE ESTUDIO Y DE TRABAJO PERSONAL QUE DEBE DEDICAR UN ESTUDIANTE PARA SUPERAR LA MATERIA

Se considera que el tiempo de trabajo personal del alumnado para superar la materia es de 115 horas repartidas como sigue:

- 1) Actividad presencial (38): 35 horas (expositiva-interactiva)+3 horas (examen)
- 2) Estudio del material (35): Se estima 1 horas por cada hora de actividad presencial (sin incluir el examen)
- 3) Trabajos de evaluación continua (42 horas)

RECOMENDACIONES PARA EL ESTUDIO DE LA MATERIA

Conviene acudir a esta materia con conocimientos previos de álgebra lineal y geometría métrica, así como los conocimientos básicos de cálculo de probabilidades y estadística. También es recomendable disponer de unas habilidades medias en el manejo de ordenadores, y en concreto de software estadístico. Para un mejor aprendizaje de la materia, conviene tener presente el sentido práctico de los métodos que se están conociendo, así como una visualización gráfica de los procedimientos que tratan datos multivariantes.

RECURSOS PARA EL APRENDIZAJE

Apuntes elaborados por el profesor, bibliografía y ordenador.

OBSERVACIONES