



UNIVERSIDADE DA CORUÑA *Universidade de Vigo*

Máster en Técnicas Estadísticas

DATOS IDENTIFICATIVOS DE LA MATERIA

Código de la materia: P1061211

Nombre de la materia: Análisis Multivariante

Número de créditos ECTS: 5

Curso académico: 2022/2023

Profesorado:

Beatriz Pateiro López (coordinadora, 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. 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 4. 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 5. 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 6. Análisis Discriminante.

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

Tema 7. 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

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.

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

En esta materia se trabajarán las competencias básicas, generales y transversales recogidas en la memoria del título. Se indican a continuación cuáles son las competencias específicas, que se potenciarán en esta materia:

Competencias específicas:

E1 - Conocer, identificar, modelar, estudiar y resolver problemas complejos de estadística e investigación operativa, en un contexto científico, tecnológico o profesional, surgidos en aplicaciones reales.

E2 - Desarrollar autonomía para la resolución práctica de problemas complejos surgidos en aplicaciones reales y para la interpretación de los resultados de cara a la ayuda en la toma de decisiones.

E3 - Adquirir conocimientos avanzados de los fundamentos teóricos subyacentes a las distintas metodologías de la estadística y la investigación operativa, que permitan su desarrollo profesional especializado.

E4 - Adquirir las destrezas necesarias en el manejo teórico-práctico de la teoría de la probabilidad y las variables aleatorias que permitan su desarrollo profesional en el ámbito científico/académico, tecnológico o profesional especializado y multidisciplinar.

E5 - Profundizar en los conocimientos en los fundamentos teórico-prácticos especializados del modelado y estudio de distintos tipos de relaciones de dependencia entre variables estadísticas.

E6 - Adquirir conocimientos teórico-prácticos avanzados de distintas técnicas matemáticas, orientadas específicamente a la ayuda en la toma de decisiones, y desarrollar la capacidad de reflexión para evaluar y decidir entre distintas perspectivas en contextos complejos.

E8 - Adquirir conocimientos teórico-prácticos avanzados de las técnicas destinadas a la realización de inferencias y contrastes relativos a variables y parámetros de un modelo estadístico, y saber aplicarlos con autonomía suficiente un contexto científico, tecnológico o profesional.

E9 - Conocer y saber aplicar con autonomía en contextos científicos, tecnológicos o profesionales, técnicas de aprendizaje automático y técnicas de análisis de datos de alta dimensión (big data).

E10 - Adquirir conocimientos avanzados sobre metodologías para la obtención y el tratamiento de datos desde distintas fuentes, como encuestas, internet, o entornos "en la nube".

METODOLOGÍA DOCENTE

La actividad presencial del alumnado será de 35 horas contando la asistencia y participación en clases expositivas y sesiones interactivas. 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 (plataformas de comunicación o correo electrónico) como de forma presencial en grupos reducidos.

CRITERIOS Y MÉTODOS DE EVALUACIÓN

La evaluación de la materia se realiza mediante evaluación continua y la realización del examen final.

La evaluación continua tendrá una ponderación del 30% y tendrá en cuenta los ejercicios/trabajos realizados a lo largo del curso. La evaluación continua permite valorar la adquisición de las competencias básicas CB7-CB9 y generales CG1-CG5. Se tendrá en cuenta el nivel alcanzado en las competencias transversales CT1-CT5. También se evaluará el nivel alcanzado en las competencias específicas CE1, CE2, CE5, CE6 y CE9.

El examen final, de problemas y cuestiones, tendrá una ponderación del 70%. En el examen, se evaluarán las competencias específicas CE1-CE6.

La calificación obtenida en la evaluación continua se conservará en las dos oportunidades de evaluación de la convocatoria de cada curso. En la segunda oportunidad de evaluación, se realizará un examen y la nota final será el máximo entre: i) la nota de la primera oportunidad, ii) la nota del nuevo examen y iii) la media ponderada del nuevo examen (70%) y la evaluación continua (30%).

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 125 horas repartidas como sigue:

- 1) Actividad presencial (38 horas): 35 horas (expositiva-interactiva)+3 horas (examen).
- 2) Actividad no presencial (87 horas): Se estima 1 hora de trabajo personal por cada hora de actividad presencial (sin incluir el examen), además del tiempo para resolución de ejercicios y casos prácticos, actividades de análisis de datos y modelos y la elaboración de trabajos.

RECOMENDACIONES PARA EL ESTUDIO DE LA MATERIA

Conviene acudir a esta materia con conocimientos básicos de álgebra lineal y geometría métrica, así como 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

El desarrollo de los contenidos de la materia se realizará teniendo en cuenta que las competencias a adquirir por el alumnado deben cumplir con el nivel MECES3. Se incidirá en los fundamentos técnicos

de las herramientas multivariantes que se estudien, y se aplicarán en distintos ejemplos prácticos, de manera que el alumnado conozca tanto las potencialidades como las posibles limitaciones de los métodos.

Para los casos de realización fraudulenta de ejercicios o pruebas, será de aplicación lo recogido en las respectivas normativas de las universidades participantes en el Máster en Técnicas Estadísticas.

Esta guía y los criterios y metodologías en ella descritos están sujetos a las modificaciones que se deriven de normativas y directrices de las universidades participantes en el Máster en Técnicas Estadísticas.