

**Proyecto/Guía docente de la asignatura**

| | | | |
|--|--|----------------------|-------|
| Asignatura | Análisis Multivariante | | |
| Materia | 2- Técnicas Estadísticas | | |
| Módulo | | | |
| Titulación | Grado en Estadística | | |
| Plan | 549 | Código | 47097 |
| Periodo de impartición | Cuatrimestre 2º | Tipo/Carácter | OB |
| Nivel/Ciclo | Grado | Curso | 3 |
| Créditos ECTS | 6 | | |
| Lengua en que se imparte | Castellano. Se manejarán materiales teóricos y programas informáticos en inglés. | | |
| Profesor/es responsable/s | Luis Ángel García y Miguel Alejandro Fernández | | |
| Datos de contacto (E-mail, teléfono...) | lagarcia@eio.uva.es, miguelaf@eio.uva.es | | |
| Departamento | Estadística e Investigación Operativa | | |

#1# ver la memoria de verificación del grado en <http://www.eio.uva.es/docencia/grado/memoria.pdf>

1. Situación / Sentido de la Asignatura #1#**1.1 Contextualización #1#**

Esta asignatura obligatoria de Análisis Multivariante de Tercer Curso del Grado en Estadística constituye el primer contacto del estudiante con problemas de inferencia multivariante así como una continuación del estudio de métodos multivariantes iniciado con la asignatura de análisis de datos del primer cuatrimestre. Con ella se busca ampliar los conocimientos del estudiante en lo relativo a los problemas multivariantes y sus técnicas.

Tiene un doble enfoque teórico y práctico que proporciona la base para poder abordar el muestreo de la normal multivariante y modelos relacionados.

1.2 Relación con otras materias #1#

Aunque la amplitud de la Estadística obliga a distribuir en asignaturas específicas el estudio de diferentes métodos, existe una gran interconexión entre ellas y esta asignatura, en particular, ayudará a proporcionar al estudiante una visión más homogénea y global, contribuyendo a adquirir una parte del bagaje de "pensamiento estadístico" que un buen profesional debe utilizar. En una pequeña medida, esta asignatura amplía la formación probabilística que el alumno ha seguido en las asignaturas de probabilidad de primer y segundo curso. Por otro lado, extiende sus conocimientos de inferencia, abordando problemas más complejos que tornan ahora hacia enfoques multivariantes. En varios momentos de su desarrollo, también se relaciona esta asignatura con Regresión y Anova y Modelo Lineal de tercer curso, a las que aporta una perspectiva geométrica y matemática más evolucionada que será útil al año siguiente para abordar los Modelos Lineales de respuesta multivariante en Modelos Estadísticos Multivariantes de Cuarto Curso.



También se potencia el uso de R y SAS, y se introduce el módulo SAS/IML, que está diseñado para programar con matrices de forma rápida. SAS/IML supone una ampliación y complemento de los módulos SAS/Base y SAS/Stat que se introdujeron en la Asignatura de Computación Estadística de segundo curso y que se utilizan, por ejemplo, en asignaturas como Regresión y Anova, Modelo Lineal o Modelos Estadísticos Multivariantes.

1.3 Prerrequisitos #1#

Es conveniente por parte del alumno un manejo desenvuelto a nivel conceptual y práctico, del Cálculo de Probabilidades y de Estadística Matemática para seguir la asignatura con aprovechamiento. También se precisan conocimientos de Álgebra y Geometría, de Cálculo Infinitesimal y de los Lenguajes de Programación, todo ello al nivel que se imparte en los cursos 1º y 2º del Grado. Deberá leer inglés técnico y manejar programas estadísticos en inglés.

2. Competencias #1#

Se incluyen a continuación las competencias de la asignatura, según aparecen en la memoria de verificación del Grado:

Transversales Instrumentales

- I1. Capacidad de análisis y síntesis
- I2. Capacidad de gestión de la información
- I3. Capacidad de organización y planificación
- I4. Conocimientos de informática relativos al ámbito de estudio
- I5. Resolución de problemas
- I6. Comunicación oral y escrita en lengua nativa
- I7. Conocimiento de lenguas extranjeras
- I8. Toma de decisiones

Personales

- P2. Razonamiento crítico
- P4. Compromiso ético

Sistémicas

- S1. Aprendizaje autónomo
- S2. Adaptación a nuevas situaciones
- S3. Motivación por el trabajo bien hecho
- S4. Iniciativa y espíritu emprendedor
- S5. Creatividad

2.1 Generales #1#

- G1. Capacidad para la gestión de la información: Incluye la capacidad para la búsqueda, manejo y exposición de información relevante de diversas fuentes, así como el manejo de las herramientas TIC necesarias.
- G2. Capacidad para la abstracción y el razonamiento crítico: El modelado y análisis de datos de problemas reales exige una elevada capacidad de abstracción, y el razonamiento crítico es necesario para hacer interpretaciones y establecer conclusiones y soluciones con rigor científico.



G3. Capacidad para la puesta al día y el auto-aprendizaje: Incluye la capacidad para la búsqueda de recursos que permitan la solución de nuevos problemas o de nuevas técnicas, en un medio científico y tecnológico en continua evolución.

2.2 Específicas #1#

E1. Recogida y tratamiento de datos: Incluye la capacidad para decidir sobre el diseño del procedimiento de obtención de datos. Capacidad para la búsqueda de información de fuentes diversas y para la elaboración de cuestionarios. Dominio en el manejo de bases de datos y en el tratamiento y depuración de los mismos.

E2. Descripción y síntesis de datos: Esta competencia es la que permitirá describir numéricamente y mediante gráficos diferentes características de interés de variables e individuos de los datos objeto de estudio, localización, escala, diferentes tipos de asociación, outliers... con empleo y adaptación de la herramienta informática apropiada o programación específica.

E3. Ajuste de modelos estadísticos y de investigación operativa: Incluye la capacidad para crear o reconocer un modelo adecuado al problema objeto de estudio. Capacidad para el ajuste del modelo mediante las herramientas estadísticas e informáticas adecuadas.

E4. Análisis de resultados, interpretación y validación de modelos: Incluye la capacidad para la selección y validación de un modelo. Capacidad para la interpretación de los modelos ajustados y las diferencias entre ellos. Empleo de herramientas informáticas específicas.

E5. Extracción de conclusiones: Incluye la capacidad para interpretar los resultados del ajuste de un modelo seleccionado en términos del problema objeto de estudio, evaluando su utilidad y/o proponiendo la necesidad de otras orientaciones del estudio.

E6. Presentación y comunicación de resultados: Incluye la capacidad para presentar los resultados de los análisis realizados, junto a las posibles soluciones a los problemas planteados por los demandantes del estudio en contextos diversos.

E7. Dominio de conceptos y herramientas informáticas a nivel de usuario avanzado: programación, sistemas operativos, algoritmos, computación, inteligencia artificial, aprendizaje automático, almacenes y minería de datos, etc.

3. Objetivos #1#

Objetivos generales prioritarios:

- Conocimiento teórico y práctico de conceptos y técnicas básicas de análisis multivariante.
- Aprendizaje de programación en R y SAS/IML dentro del contexto del análisis Multivariante, generando datos, programando técnicas estudiadas y resolviendo casos.

4. Bloques temáticos

Bloque 1: ANÁLISIS MULTIVARIANTE

Carga de trabajo en créditos ECTS:

a. Contextualización y justificación #1#

Esta asignatura obligatoria de Análisis Multivariante de Tercer Curso del Grado en Estadística constituye una continuación del estudio de métodos multivariantes iniciado con la asignatura de análisis de datos del primer



cuatrimestre. Con ella se busca ampliar los conocimientos del estudiante en lo relativo a los problemas multivariantes y sus técnicas.

Tiene un enfoque eminentemente práctico, pero siempre tratando de comprender bien las bases conceptuales de cada técnica y evitando presentar al alumno un catálogo exhaustivo de procedimientos para que los utilice mediante "recetas" mecánicamente aplicadas.

Se tratará de dotar al estudiante de la herramienta estadística y geométrica formal mínima, más a nivel de ideas que de desarrollos teóricos, para que comprenda el funcionamiento de las técnicas estadísticas que maneja y sea capaz de plantearlas correctamente y realizar interpretaciones precisas.

b. Objetivos de aprendizaje #1#

- Que el estudiante aprenda a plantear y reconocer problemas de Análisis Multivariante localizando fuentes de datos en su caso, a elegir las técnicas adecuadas para su resolución y a plantearlas adecuadamente.
- Que el estudiante aprenda a manejar paquetes de programas estadísticos, y se inicie a programar en lenguajes matriciales, como SAS/IML o R. También manejará herramienta informática complementaria que le permitirá resolver problemas de análisis multivariante y elaborar informes.
- Que el estudiante sepa cómo interpretar con precisión los resultados de las diferentes técnicas multivariantes en problemas aplicados, dando respuesta a problemas concretos.
- Que el estudiante avance en la definición y seguimiento de los diferentes pasos del proceso que va desde la formulación de un problema real, hasta la solución estadística y su comunicación.
- Que el estudiante potencie el desarrollo de varias competencias genéricas demandadas en el ámbito profesional, como son el trabajo en equipo, la presentación de informes, la expresión oral y escrita, la elección de la solución más apropiada, la capacidad de iniciativa y el sentido crítico.

c. Contenidos #1#

- La normal multivariante y distribuciones asociadas. Inferencias sobre la normal multivariante.
- Análisis de correlaciones canónicas.
- Multidimensional scaling.
- Análisis factorial.
- Discriminación logística. Introducción a las redes neuronales.
- Support vector machines.
- Árboles de clasificación. Random forests.
- Ampliación de métodos de clasificación no supervisada.

d. Métodos docentes #1#

Se combinarán clases magistrales y prácticas en salas informáticas con elaboración, exposición y discusión de trabajos individuales y en grupo.

e. Plan de trabajo #1#

Se abordarán diversos problemas completos desde su planteamiento en la vida real, pasando por la manipulación de datos, su análisis y la aplicación práctica de los procedimientos estadísticos multivariantes más apropiados, hasta la elección de una solución final satisfactoria y su presentación mediante la elaboración



de informes y su exposición oral. Los estudiantes tratarán problemas complejos que deberán estructurar en tareas, fraccionándolos en pasos o etapas encadenadas y relacionadas. Se potenciará el uso de diversos programas informáticos estadísticos (SAS, R...) y otros auxiliares que empleará como herramienta intermedia para preparación, lectura y depuración de los datos iniciales, y para elaborar los informes finales. Deberá también utilizar técnicas estadísticas diversas aprendidas en otras asignaturas del Grado, como Regresión y Anova, Modelos Lineales o Inferencia Estadística. Con ello se persigue que el alumno progrese en su proceso de maduración dentro de su formación estadística integral.

f. Evaluación #1#

La evaluación de esta asignatura, tanto en la convocatoria ordinaria como en la extraordinaria, se efectuará a partir de las siguientes actividades.

| ACTIVIDAD | PESO EN LA NOTA FINAL | OBSERVACIONES |
|---|-----------------------|--|
| Evaluación continua (podrá incluir prácticas con ordenador evaluadas y trabajos individuales o en grupo, entre otros) | 30% | Las notas obtenidas en evaluación continua son válidas para la convocatoria ordinaria. |
| Examen en la convocatoria ordinaria | 70% | |
| Examen en la convocatoria extraordinaria | 100% | Este examen podrá incluir programación con ordenador |

g. Bibliografía básica #1#

- HASTIE, T. TIBISHIRANI, R., FRIEDMAN J. (2001), The Elements of Statistical Learning, Ed. Springer.
- SEBER, G.A.F. (1984), Multivariate Observations, Ed: Wiley.
- SRIVASTAVA, M.S. (2002) Methods of Multivariate Statistics. Ed. Wiley

h. Bibliografía complementaria

- AFIFI, A. A. Y CLARK, V. (1990), Computer-Aided Multivariate Analysis (second edition), Ed: Van Nostrand Reinhold.
- CHATFIELD, C. Y COLLINS, A.J. (1980), Introduction to Multivariate Analysis, Ed: Chapman and Hall.
- EVERITT, B. (1993), Cluster Analysis (third edition), Ed: Edward Arnold.
- EVERITT B. S. Y DUNN G. (1991), Applied Multivariate Data Analysis, Ed: Edward Arnold, London.
- FLURY BERNARD (1997). A first Course in Multivariate Statistics. Ed: Springer
- JOBSON, D.V. (1992). Applied Multivariate Data Analysis. Volume II: Categorical and Multivariate Methods. Ed: Springer-Verlag.
- KRZANOWSKI, W. J. (1988), Principles of Multivariate Analysis (a user's perspective), Ed: Oxford Science Publications.
- PEÑA D. (2002) Análisis de Datos Multivariantes. Ed.: Mc Graw Hill.

i. Recursos necesarios #1#

Software estadístico específico, disponible en las salas de informática de la Facultad de Ciencias y el Departamento de Estadística e Investigación Operativa.



j. Temporalización #1#

| BLOQUE TEMÁTICO | CARGA ECTS | PERIODO PREVISTO DE DESARROLLO |
|------------------------|------------|--------------------------------|
| ANÁLISIS MULTIVARIANTE | 6 | Segundo cuatrimestre |

5. Métodos docentes y principios metodológicos #1#

| MÉTODOS DOCENTES | OBSERVACIONES |
|-----------------------------------|---------------|
| Clases expositivas y de problemas | 30 horas |
| Laboratorios | 30 horas |
| Tutorías | |

La modalidad docente prevista es la de PRESENCIALIDAD SEGURA. Es decir, si la disponibilidad de espacios en el centro y la organización global prevista en el mismo lo permite, la docencia se desarrollará de forma presencial. En otro caso, la docencia será bimodal con webcam en aulas para que los alumnos que no puedan seguir la docencia de forma presencial la sigan en streaming preferentemente desde un espacio contiguo.

6. Tabla de dedicación del estudiante a la asignatura #1#

| ACTIVIDADES PRESENCIALES | HORAS | ACTIVIDADES NO PRESENCIALES | HORAS |
|-----------------------------------|-----------|-----------------------------------|-----------|
| Clases en aula | 30 | Estudio y trabajo autónomo | 30 |
| Laboratorios/Clases con ordenador | 30 | Elaboración de trabajos | 20 |
| | | Trabajo personal en laboratorio | 25 |
| | | Preparación de presentaciones | 10 |
| | | Otras actividades complementarias | 5 |
| Total presencial | 60 | Total no presencial | 90 |

7.

Sistema

La calificación de esta asignatura, tanto en la convocatoria ordinaria como en la extraordinaria, se efectuará a partir de los siguientes instrumentos/procedimientos.

| INSTRUMENTO/PROCEDIMIENTO | PESO EN LA NOTA FINAL | OBSERVACIONES |
|---|-----------------------|--|
| Evaluación continua (podrá incluir prácticas con ordenador evaluadas y trabajos individuales o en grupo, entre otros) | 30% | Las notas obtenidas en evaluación continua son válidas para la convocatoria ordinaria. |
| Examen escrito en la convocatoria ordinaria | 70% | |
| Examen en la convocatoria extraordinaria | 100% | Este examen podrá incluir programación con ordenador |