

GUIA DOCENTE

Nombre de la asignatura (*) Análisis e interpretación de datos ambientales (AIDA)
Centro Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos de Barcelona
Departamento Matemática Aplicada III
Curso 2013-2014
Titulación Máster en Ingeniería Ambiental

Créditos ECTS(*):

5

Idiomas(*):

Español

Inglés

Catalán

Profesores (*)

Responsable

Juan José Egozcue Rubí

Otros

Maribel Ortego Martínez

Horario de atención

Horario

Miércoles de 15.30 a 17h. Otros a convenir.

Objetivos(*)

Aprender a extraer conclusiones fundadas mediante el estudio de datos representativos de un problema medioambiental.

Objetivos específicos

Conocimientos

Elementos básicos del análisis estadístico multivariante aplicados a datos ambientales.

Habilidades

Desarrollar un conocimiento crítico de los métodos y modelos cuantitativos multivariantes, y en especial, atender a sus limitaciones. Aprender a obtener e interpretar modelos estadísticos multivariantes aplicados a problemas medioambientales.

C.Transversales

Búsqueda de documentación científica
Capacidad de comunicación oral y escrita de resultados científicos
Práctica de varias lenguas

Contenidos(*)

1: Análisis exploratorio y descriptivo

Descripción:

Análisis crítico de una base de datos. Descriptores y resumen de la información. Selección de la escala de los datos (análisis de escala real, positiva, direccional o composicional). Análisis de componentes principales y biplot.

Descripción laboratorio

Introducción al programa R y al análisis de datos con hojas de cálculo. Técnicas de simulación de variables

Desglose del contenido:

Laboratorio: 2h lectivas; 0.6 ects
Problemas: 0.2 ects
Teoría: 6h lectivas; 0.8 ects

Objetivos específicos:

Conocimientos: conceptos de

(*) Campos obligatorios

aleatorias. Análisis descriptivo univariante y bivalente. Análisis de datos composicionales. Componentes principales y biplot de datos multivariantes.	análisis de datos composicionales, análisis de componentes principales, técnicas de representación de datos. Habilidades: programación básica; exploración y representación de datos. Interpretación y síntesis de los datos según sus características
--	---

2: Modelos lineales (regresión y ANOVA) Descripción Variables respuesta y explicativas. Transformaciones de escala de variables. Modelos lineales para las variables explicativas. Regresión y contrastes de significación. ANOVA (comparación de medias) y contrastes de significación. Gráficos de diagnóstico y usos predictivos. Descripción laboratorio: Transformaciones logarítmicas y logísticas para datos positivos y composicionales. Cálculo efectivo de modelos de regresión. Inclusión de factores como variables explicativas. Cálculo de modelos ANOVA. Contrastes y diagnósticos.	Desglose del contenido: Laboratorio: 6h lectivas, 1.0 ects Problemas: 0.2 ects Teoría: 4h lectivas, 0.8 ects Objetivos específicos: Conocimientos: contrastes de hipótesis y su uso. Concepto de p-valor y su interpretación. Formulación e interpretación de modelos lineales Habilidades: modelización estadística y conclusiones que se derivan
--	---

3: Clasificación y discriminación Descripción: Análisis discriminante (clasificación supervisada): variables respuesta discretas y observadas. Regresión logística. Otros análisis discriminantes. Contrastes de hipótesis sobre los modelos. Análisis de conglomerados (clasificación no supervisada; "cluster"): no existe clasificación previa. Métodos de análisis. Discusión de distancias y disimilaridades. Descripción laboratorio Práctica de análisis discriminante (regresión logística, de Poisson y/o discriminante lineal). Práctica de análisis de conglomerados (revista de varios métodos).	Desglose del contenido: Laboratorio: 6h lectivas, 1.0 ects Problemas: 0.2 ects Teoría: 2h lectivas, 0.2 ects Objetivos específicos: Conocimientos: análisis discriminante y de conglomerados. Interpretación Habilidades: Reconocer los modelos o técnicas adecuados a cada conjunto de datos.
--	--

Desglose de contenidos:

Laboratorio: 2.6 ects
Problemas: 0.6 ects
Teoría: 1.8 ects

Tiempo total: 5 ects

Planificación de actividades

Título actividad 1: Exposición lectiva	Dedicación total: 12 h
Descripción: asistencia a las clases de la asignatura	
Material: diapositivas y resúmenes de clase	
Entregable: NO	
Objetivos: transferir conocimientos básicos de las técnicas estudiadas, generar criterios de evaluación de su utilidad	

Título actividades 2: Tareas autónomas	Dedicación total: 40 h
Descripción: ejercicios prácticos con hoja de cálculo y R. Búsqueda de documentación de los ejercicios en plataforma Atenea.	
Material: documentación de los ejercicios en plataforma Atenea.	
Entregable: Sí	
Objetivos: profundizar en el conocimiento de los métodos, sus limitaciones y usos	

Título actividad 3: Resolución de problemas	Dedicación total: 10 h
--	---

(*) Campos obligatorios

Descripción: problemas de tipo teórico-práctico	
Material: documentación de los ejercicios en plataforma Atenea.	
Entregable: No. Corrección personalizada	
Objetivos: consolidar en el conocimiento de los modelos teóricos de probabilidad y los usos de la estadística elemental	
Título actividades 4:	Dedicación total:
Desarrollo de prácticas	14 h
Descripción: ejercicios prácticos con hoja de cálculo y R bajo supervisión	
Material: documentación de los ejercicios en plataforma Atenea y dado en clases prácticas	
Entregable: Sí. A través de plataforma Atenea	
Objetivos: Contactos con el software. Introducción a los métodos de análisis a utilizar en ejercicios autónomos.	

Título actividades 5:	Dedicación total:
Exposición oral	16 h
Descripción: elaboración autónoma y exposición oral en clase de un caso de análisis de datos ambientales	
Material: Guía de presentación y rúbrica de criterios de evaluación	
Entregable: Sí, en forma oral y material de la presentación	
Objetivos: aprender a desarrollar un análisis de datos, desde el planteamiento de la pregunta a la respuesta razonada, soportada por un análisis estadístico	

Sistema de calificación^(*)

- 1) tareas entregadas a través de Atenea (40%)
- 2) exposición oral: contenido y forma (60%)

Normas de realización de las pruebas^(*)

- 1) las tareas deben presentarse en la forma, formato, fecha y hora definidas
- 2) es obligatorio asistir al menos a 5 exposiciones orales de los compañeros
- 3) la exposición oral debe hacerse en 10 minutos + 5 minutos de preguntas. La presentación en formato electrónico es entregable y obligatoria

Capacidades previas

Conocimientos previos de álgebra lineal, cálculo, probabilidad y estadística elementales. Capacidad de abstracción. Capacidad de estudio individual. Conocimiento del acceso a la información en internet y documentación bibliográfica. Comprensión de problemas ambientales, incluyendo sus aspectos bio-geo-químicos.

Requisitos

Un curso de estadística, incluyendo modelos de regresión

Metodología docente

Combina las clases expositivas con las sesiones de discusión y las prácticas supervisadas. Se complementa con ejercicios autónomos y un trabajo en solitario, que equivalen a la mitad del esfuerzo en la asignatura

Bibliografía^(*)

Básica

Krzanowski, W. J. (1988): *Principles of Multivariate Analysis*, Clarendon Press, Oxford (UK), 563 p
Peña, D. (2002): *Análisis de datos multivariantes*, Mc Graw Hill, Aravaca (Madrid), 539p.
Peña, D. (1991): *Estadística – Modelos y Métodos. I. Fundamentos y II. Modelos lineales y series temporales*, Alianza Editorial, Madrid (E), 1093 p
Mac Berthouex, Paul, and Brown, Linfiels C. (1994): *Statistics for Environmental Engineers*, Lewis Publishers, Boca Ratón, NY (USA), 335 p

Complementaria

Pawlowsky-Glahn, V., J. J. Egozcue and R. Tolosana-Delgado (2007): *Lecture Notes on Compositional Data Analysis*, <http://hdl.handle.net/10256/297>, (Universitat de Girona), 2007.
Pawlowsky-Glahn, V. and Buccianti, A. : *Compositional data analysis: theory and applications*, Wiley, Chichester UK, 2011.
ISBN-10: 0-470-71135-3

R Development Core Team (2009). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

Otros recursos

Recursos no tabulados Disponibles en plataforma Atenea

^(*) Campos obligatorios

Material audiovisual

Material informático

**Open software: R, Open office, CoDa-Pack.
Otros: MS-Excel**
