

Análisis Exploratorio de Datos

Código: 104853
Créditos ECTS: 6

2025/2026

Titulación	Tipo	Curso
Estadística Aplicada	FB	1

Contacto

Nombre: Maria Rosa Camps Camprubi

Correo electrónico: rosa.camps@uab.cat

Equipo docente

Montserrat Ferre Delgado

Idiomas de los grupos

Puede consultar esta información al [final](#) del documento.

Prerrequisitos

Conocimientos estándar de bachillerato.

Objetivos y contextualización

El objetivo de este curso es, en primer lugar, el conocimiento de las técnicas descriptivas y exploratorias aplicadas a resumir la información contenida en los conjuntos de datos experimentales.

Se empieza con el estudio de una variable, y seguidamente se inicia el estudio comparativo de dos o más variables.

Se dedican apartados especiales al estudio de la evolución temporales de las variables: series cronológicas y números índice.

El segundo objetivo es que los estudiantes adquieran soltura en el uso de software estadístico para manipular datos, realizar análisis descriptivo y gráficas.

Un objetivo paralelo, tanto o más importante, es la interpretación de los resultados y de las gráficas en el contexto de los datos.

Resultados de aprendizaje

1. CM05 (Competencia) Interpretar situaciones basadas en conjuntos de datos, representaciones gráficas y resúmenes estadísticos.
2. CM06 (Competencia) Aplicar los conocimientos adquiridos para organizar datos, construir y mostrar tablas y trabajar con diferentes representaciones de datos.
3. KM07 (Conocimiento) Identificar patrones de comportamiento en la exploración de datos.
4. SM08 (Habilidad) Analizar resultados de encuestas.

Contenido

1. Preliminares

- 1.1. Objetivo del análisis exploratorio de datos o estadística descriptiva.
- 1.2. Tipos de variables y escalas de medida.
- 1.3. Redondeos y notación científica.

2. Resumen de datos estadísticos.

- 2.1. Distribuciones de frecuencias: tablas.
- 2.2. Agrupamiento en intervalos.
- 2.3. Representaciones gráficas.

3. Características numéricas de una variable.

- 3.1. Características de posición central: media, mediana, moda.
- 3.2. Otras características de posición: cuartiles, deciles y percentiles.
- 3.3. Características de dispersión: varianza y desviación típica (muestrales y poblacionales), rango, rango inter-cuartil.
- 3.4. Características de dispersión relativa.
- 3.5. Puntuaciones tipificadas.
- 3.6. Características de forma: simetría y curtosis.

4. Complementos en el estudio de una variable.

- 4.1. Análisis exploratorio: diagrama de caja y otros gráficos.
- 4.2. Transformación de variables.
- 4.3. Otras medias: geométrica, armónica, cuadrática.
- 4.4. La desigualdad de Chebyshev.

5. Comparación de una variable en dos o más grupos: Análisis exploratorio

- 5.1. Situación de muestras independientes.
- 5.2. Situación de muestras pareadas

6. Tabulación y representación de la distribución conjunta de los valores de dos variables categóricas.

- 6.1. Tablas de contingencia (distribuciones de frecuencias conjunta, marginales y condicionadas).
- 6.2. Análisis descriptivo de la dependencia entre dos variables categóricas.

7. Descripción numérica de la distribución conjunta de dos variables estadísticas.

- 7.1. Características marginales y condicionadas.
- 7.2. Curvas de regresión y coeficiente de correlación.
- 7.3. Ajuste lineal y predicción.

8. Introducción a las series temporales.

8.1. La descomposición clásica.

Actividades formativas y Metodología

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Clases de problemas	8	0,32	
Clases teóricas	18	0,72	
Estudiar conceptos de teoría, resolver problemas a mano y con R	84	3,36	
Prácticas con ordenador	30	1,2	

El trabajo de aula, teoría y problemas, se complementará con prácticas de ordenador donde se utilizará el paquete estadístico R.

En el Moodle del curso los alumnos encontrarán la planificación de la asignatura, las listas de problemas y las prácticas, así como eventuales cambios de aula, horarios, etc.

Es importante tener en cuenta que el CampusVirtual no es una web estática sino que se irá actualizando a lo largo del curso.

Se procurará introducir en la parte más práctica de la asignatura, el análisis y comparación de datos estadísticos por sexo, comentando, en su caso, en el aula las causas y los mecanismos sociales y culturales que pueden sustentar las desigualdades observadas.

*La metodología docente propuesta puede experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Evaluación

Actividades de evaluación continuada

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Nota de las classes de prácticas	10%	2	0,08	CM05, CM06
Primer parcial de teoría i problemas	25%	2	0,08	CM05, KM07, SM08
Primera prueba de resolución de problemas con ordinador	15%	2	0,08	

Segunda prueba de resolución de problemas con ordenador	25%	2	0,08
Segundo parcial de teoría i problemas	25%	2	0,08

La nota final de la asignatura F se obtendrá a partir de:

- 1) Las notas de los dos exámenes parciales de teoría y problemas, TP1 y TP2, con pesos respectivos 25% y 25%.
- 2) Las notas de las dos pruebas con ordenador, O1 y O2, con pesos respectivos 15% y 25%.
- 3) En las sesiones prácticas semanales se propondrán entregas. Después, cada dos o tres semanas, se evaluarán con pequeños cuestionarios, dando la nota PC, con un peso total del 10%. Esta parte no es recuperable. En caso de dudas en la nota final se revisarán las entregas realizadas por el estudiante a lo largo del curso.

La nota final de la asignatura se obtiene haciendo la media ponderada

$$F = 0,25 \text{ TP1} + 0,15 \text{ O1} + 0,25 \text{ TP2} + 0,25 \text{ O2} + 0,10 \text{ PC}$$

Requisito para superar la asignatura con la media anterior, las notas TP1, TP2, O1, y O2 deben ser mayores o iguales a 4.

Habrá dos pruebas de síntesis para aquellos estudiantes que no aprueben la asignatura ($F < 5$) o bien no cumplan los requisitos mencionados ($O2, TP1, TP2 \geq 4$ y $O1 \geq 3,5$):

- Una prueba de síntesis STP de todo el temario de teoría y problemas, para los estudiantes que tengan alguna de las notas TP1 o TP2 de menos de 4 o suspendan la asignatura a causa de teoría y problemas.
- Una prueba de síntesis SO de todas las prácticas, para aquellps estudiantes que tengan menos de 4 en O2 o menos de 3.5 en O1 o bien suspendan la asignatura a causa de las prácticas.

Su nota final será $F=0,50 \text{ STP} + 0,40 \text{ SO} + 0,10 \text{ PC}$

(en caso que sólo necesiten presentarse a una de las dos pruebas de síntesis, se substituirá la nota del otro examen de síntesi porr la media ponderada de los dos parciales ya aprobados).

Después de las segundas pruebas parciales se otorgarán las matrícululas de honor que se consideren claras. Estas matrícululas serán ya definitivas. Si el número máximo de matrículas permitido no se ha alcanzado, se reconsiderará la posibilidad de otorgar más después del examen de recuperación.

Si un estudiante no se presenta en ningún parcial ni a el aprobado de síntesis, se calificará como "No Evaluable".

Sin perjuicio de otras medidas disciplinarias que se estimen oportunas, y de acuerdo con la normativa Acadèmica vigente, se calificarán con un cero las irregularidades cometidas por el estudiante que puedan conducir a una variación de la calificación de un acto de evaluación. Por lo tanto, plagiar, copiar o dejar copiar una práctica o cualquier otra actividad de evaluación implicará suspender con un cero y no se podrá recuperar en el mismo curso académico. Si esta actividad tiene una nota mínima asociada, entonces la asignatura quedará suspendida.

El uso de la inteligencia artificial (IA) está fuertemente desaconsejado cuando se hace por parte de personas no expertas en la materia que se consulta, como es el caso de los estudiantes de la asignatura. En esta asignatura, no se permite el uso de tecnologías de Inteligencia Artificial (IA) para generar las entregas de prácticas. Por lo tanto cualquier trabajo que incluya fragmentos generados con IA será considerado una falta de honestidad académica y puede comportar una penalización parcial o total en la nota de PC. Las sospechas

se aclararán con una entrevista con las profesoras. Evidentemente, la consulta de IA durante los exámenes será considerada como una copia e implicará suspender la asignatura con un cero.

El alumnado que se haya acogido a la modalidad de evaluación única deberá realizar dos exámenes globales de la asignatura: una prueba de teoría y de problemas y otra prueba con ordenador, utilizando el software R. Estos dos exámenes se realizarán el mismo día y hora que los correspondientes al segundo parcial de la asignatura. Ese día el estudiante podrá entregar los ejercicios de las entregas de prácticas que se hayan programado a lo largo del curso, cuyo enunciado encontrará en Moodle. La ponderación para la nota final será de 50% del examen de teoría y problemas, 40% del examen de prácticas y 10% de las entregas. El estudiante podrá ser convocado posteriormente a una revisión oral de sus exámenes y entregas con las profesoras de la asignatura. Si se obtiene una nota inferior a 5, podrá recuperarse el mismo día, hora y lugar que se realice la recuperación del resto de alumnos del curso. Las ponderaciones serán las mismas.

Bibliografía

Manual de teoría

X. BARDINA, M. FARRÉ, Estadística descriptiva, Manuals, 54 Servei de Publicacions, UAB

Bibliografía:

A.J.B. ANDERSON, Interpreting Data. A first cours in Statistics, Ed Chapman and Hall, 1989.

R Tutorial. An R introduction to statistics. (2016). www.r-tutor.com

E. CASA ARUTA, Problemas de Estadística Descriptiva, Ed. Vicens Vives.

R. JOHNSON, P. KUBY, Estadística elemental: Lo esencial, Ed Thomson, 1999.

B. PY, Statistique Descriptive, Ed Económica, 1988.

M. SPIEGEL, Estadística, Teoría y 875 problemas resueltos, Schaum-McGraw-Hill, 1990.

V. ZAIATS, M.L. CALLE i R. PRESAS, Probabilitat i Estadística. Exercicis I, Eumo Ed, 1998.

Bibliografía complementaria.

G. CALOT, Curso de Estadística Descriptiva. Ed Paraninfo, 1988.

FERNÁNDEZ, J.M. CORDERO, A. C'ORDOBA, Estadística Descriptiva, ed ESIC 1996.

L.C HAMMILTON, Modern Data Analysis, Brooks/Cole Publishing Company, 1990.

P.G. HOEL i R.J. JESSEN, Estadística básica para negocios y economía, Compañía Editorial Continental,Mexico, 1993.

R.K. PEARSON, Exploratory Data Analysis using R. Data Mining and Knowledge Discovery Series, Chapman & Hall/CRC, 2018.

D. PEÑA SÁNCHEZ DE RIVERA, Estadística. Modelos y métodos. 1. Fundamentos i 2. Modelos lineales yseries temporales, Alianza Editorial 1995. (2 volúmenes)

Software

R i RStudio

Grupos e idiomas de la asignatura

La información proporcionada es provisional hasta el 30 de noviembre de 2025. A partir de esta fecha, podrá consultar el idioma de cada grupo a través de este [enlace](#). Para acceder a la información, será necesario introducir el CÓDIGO de la asignatura

Nombre	Grupo	Idioma	Semestre	Turno
(PLAB) Prácticas de laboratorio	1	Catalán	primer cuatrimestre	tarde
(PLAB) Prácticas de laboratorio	2	Catalán	primer cuatrimestre	tarde
(TE) Teoría	1	Catalán	primer cuatrimestre	tarde