



MÉTODOS ESTADÍSTICOS AVANZADOS

PRESENTACIÓN

En esta asignatura se recorrerá el flujo completo de un análisis de datos utilizando el lenguaje de programación estadística R, desde la lectura, limpieza y análisis exploratorio de los datos hasta la modelización estadística utilizando tanto técnicas paramétricas clásicas como técnicas computacionalmente intensivas basadas en remuestreo y simulación. Se hará especial énfasis en el aspecto gráfico del análisis de datos, tanto en el aspecto exploratorio como expositivo.

- **Titulación:** MÁSTER UNIVERSITARIO EN CIENCIA DE DATOS PARA CIENCIAS EXPERIMENTALES
- **Módulo/Materia:** Módulo 2 General / Materia 2.3. Fundamentos de aprendizaje automático
- **ECTS:** 5
- **Curso, semestre:** 1.º curso y primer semestre
- **Carácter:** Obligatoria
- **Profesor responsable:** Sergio Ardanza-Trevijano
- **Idioma:** Castellano
- **Aula y horario:** Consultar el calendario del Máster

RESULTADOS DE APRENDIZAJE DE LA ASIGNATURA

RA1 Aplicar soluciones computacionales para problemas científicos, usando una amplia gama de plataformas de análisis de datos

RA7 Implementar algoritmos eficientes para la resolución de problemas científicos utilizando programación estructurada y pseudocódigo

RA12 Manejar con soltura el lenguaje de programación R para la manipulación y análisis de datos científicos.

RA13 Implementar las nociones teóricas de la ciencia de datos en un entorno computacional utilizando el lenguaje de programación R.

RA14 Manipular y procesar datos científicos de diversas fuentes (archivos de texto, bases de datos, etc.) utilizando técnicas de data wrangling en diferentes entornos de programación

RA15 Generar e interpretar representaciones gráficas adecuadas para comunicar resultados científicos de manera clara y efectiva en diferentes entornos de programación

RA16 Aplicar modelos probabilísticos para estudiar fenómenos aleatorios en contextos científicos.

RA17 Implementar técnicas de inferencia estadística para analizar datos científicos, extraer conclusiones y tomar decisiones informadas.

PROGRAMA

1. Introducción al análisis estadístico de datos y programación en R. Programación funcional con -apply.
2. Limpieza y manejo de datos con el tidyverse.
3. Análisis exploratorio de datos. Principios de la visualización y paquetes de visualización en R
4. Modelos probabilísticos y simulaciones en R. Cadenas de Markov.



Universidad
de Navarra

Facultad de Ciencias

5. Introducción a la inferencia. Estimación de máxima verosimilitud analítica y numérica. Intervalos de confianza y tests de hipótesis.
6. Métodos de Resampling. Bootstrap y permutaciones.
7. Álgebra lineal para análisis de datos.
8. Análisis de componentes principales.
9. Regresión lineal simple y múltiple. Introducción al diseño.
10. Modelos lineales generalizados. Regresión logística y de Poisson.

ACTIVIDADES FORMATIVAS DE LA ASIGNATURA

Clases expositivas: 25h

Clases prácticas: 25 h

Estudio personal: 60h

Trabajos dirigidos/problemas: 10h

Evaluación: 3h

Tutorías: 2h

SISTEMAS DE EVALUACIÓN

Quizzes 25% Resolución de problemas

Casos prácticos 25%

Examen final 50%

HORARIOS DE ATENCIÓN

Para horario de consultas por favor concertar una cita escribiendo un correo a sardanza@unav.es

BIBLIOGRAFÍA

G. Golemund, H. Wickham *R for data science*. Disponible online en <https://r4ds.hadley.nz/>

P. Dalgaard, *Introductory Statistics with R* (Springer) Disponible en PDF para alumnos y empleados de la Universidad

Julian Faraway *Linear Models with R*. CRC press.

E. Demidenko, *Advanced Statistics with Applications in R* (Wiley)

Kieran Healy *Data visualization, a practical introduction* Disponible online en socviz.co

A. Garcimartín y S. Ardanza-Trevijano *Introducción a la estadística con R*