

Curso Académico: (2022 / 2023)

Fecha de revisión: 16-05-2022

Departamento asignado a la asignatura: Departamento de Estadística

Coordinador/a: GARCIA PORTUGUES, EDUARDO

Tipo: Optativa Créditos ECTS : 3.0

Curso : 1 Cuatrimestre : 2

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

Programación en R
Programación Avanzada

OBJETIVOS*** Competencias básicas**

- CB6: Poseer y comprender los conocimientos que proporcionan una base u oportunidad para ser original en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
- CB7: Saber aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.
- CB8: Integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.
- CB9: Comunicar conclusiones, así como el conocimiento y las razones últimas que las sustentan, a audiencias especializadas y no especializadas de una manera clara e inequívoca.
- CB10: Desarrollar las habilidades de aprendizaje que permitan continuar estudiando de manera que se sea en gran medida autodirigido o autónomo.

*** Competencias generales**

- CG1: Aplicar las técnicas de análisis y representación de la información, para adaptarla a problemas reales.
- CG2: Identificar el modelo estadístico más adecuado para cada problema real y saberlo aplicar para su análisis, diseño y solución.
- CG3: Obtener soluciones científicamente viables para problemas estadísticos reales complejos, tanto de manera individual como en equipo.
- CG4: Sintetizar las conclusiones obtenidas del análisis de datos y presentarlas de forma clara y convincente en un entorno bilingüe (español e inglés), tanto escrito como oral.
- CG5: Generar nuevas ideas (creatividad) y anticiparse a nuevas situaciones, en los contextos de análisis de datos y de toma de decisiones.
- CG6: Aplicar habilidades sociales para el trabajo en equipo y para relacionarse con los demás de forma autónoma.

*** Competencias específicas**

- CE1: Aplicar conocimientos avanzados de inferencia estadística en el desarrollo de métodos de análisis de problemas reales.
- CE2: Usar software libre como R y Python para la implementación de análisis estadísticos.
- CE5: Aplicar los fundamentos estadísticos avanzados para el desarrollo y análisis de problemas reales que implican la predicción de una respuesta variable.
- CE6: Aplicar modelos no paramétricos para la interpretación y predicción de fenómenos aleatorios.
- CE8: Aplicar y desarrollar técnicas de visualización de muestras recogidas con software de libre distribución como R y Python.
- CE9: Identificar correctamente el tipo de análisis estadístico correspondiente a unos objetivos y datos determinados.
- CE10: Aplicar la modelización estadística en el tratamiento de problemas relevantes en el campo científico.
- CE11: Formalizar fenómenos aleatorios y modelizarlos por medio de modelos probabilísticos.

- CE12: Aplicar modelos para el aprendizaje supervisado y no supervisado.
- CE13: Modelizar datos complejos con dependencia estocástica.
- CE14: Aplicar conocimientos y capacidades avanzadas de consultoría estadística.

* Resultados del aprendizaje

Adquisición de conocimientos sobre: 1) habilidades útiles en un servicio de consultoría estadística; 2) técnicas de presentación automática de resultados en informes; 3) desarrollo de aplicaciones Shiny; 4) el entorno tidyverse; 5) el entorno tidymodels.

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

Este curso cubre varias herramientas para agilizar el proceso de consultoría en R: desde la organización y el manejo de datos hasta la presentación de resultados, pasando por un rápido modelado estadístico. El énfasis se encuentra en ver las principales características de muchos paquetes y soluciones diferentes.

1. R Markdown avanzado para la creación de informes

- 1.1. Temas avanzados en R Markdown
- 1.2. Escribiendo buenos informes
- 1.3. Presentaciones personalizadas
- 1.4. Otros documentos y temas

2. Aplicaciones Shiny

- 2.1. Paradigma principal
- 2.2. Examples of simple applications
- 2.3. Reactions and appearance
- 2.4. More advanced applications
- 2.5. flexdashboard
- 2.6. Otros temas

3. Manejo de datos dentro del tidyverse I

- 3.1. dplyr
- 3.2. tidyr
- 3.3. readr
- 3.4. tibble
- 3.5. Otros paquetes

4. Manejo de datos dentro del tidyverse II

- 4.1. stringr
- 4.2. forcats
- 4.3. lubridate y hms
- 4.4. glue
- 4.5. purrr
- 4.6. Otros paquetes

5. Modelización rápida usando AutoML

- 5.1. Introducción al AutoML
- 5.2. Explicabilidad
- 5.3. Ejemplos en regresión
- 5.4. Ejemplos en clasificación binaria
- 5.5. Ejemplos en clasificación multiclase

6. Modelización rápida con tidymodels I

- 6.1. broom
- 6.2. rsample
- 6.3. parsnip
- 6.4. yardstick
- 6.5. Otros paquetes

7. Modelización rápida con tidymodels II

- 7.1. recipes
- 7.2. workflows
- 7.3. tune
- 7.4. infer
- 7.5. Otros paquetes

El programa está sujeto a modificaciones menores debido al desarrollo del curso y/o al calendario académico.

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Las clases consisten en una mezcla de exposiciones referentes al software visto y del uso práctico del mismo. Se emplea el lenguaje estadístico R. Se espera que los estudiantes traigan sus propios portátiles para experimentar con el código durante las clases.

* Actividades formativas

- AF1: Clase teórica.
- AF2: Clase práctica.
- AF5: Tutorías.
- AF6: Trabajo en grupo.
- AF7: Trabajo individual.
- AF8: Pruebas de evaluación presenciales.

* Metodologías docentes

- MD1: Exposiciones en clase del profesor con soporte de medios informáticos y audiovisuales, en las que se desarrollan los conceptos principales de la materia y se proporciona la bibliografía para complementar el aprendizaje de los alumnos.
- MD2: Lectura crítica de textos recomendados por el profesor de la asignatura: Artículos de prensa, informes, manuales y/o artículos académicos, bien para su posterior discusión en clase, bien para ampliar y consolidar los conocimientos de la asignatura.
- MD3: Resolución de casos prácticos, problemas, etc. planteados por el profesor de manera individual o en grupo.
- MD4: Exposición y discusión en clase, bajo la moderación del profesor de temas relacionados con el contenido de la materia, así como de casos prácticos.
- MD5: Elaboración de trabajos e informes de manera individual o en grupo.

SISTEMA DE EVALUACIÓN

La evaluación ordinaria se realiza enteramente mediante evaluación continua. Esta se hace con una mezcla de:

- (a) un conjunto de ejercicios prácticos;
- (b) participación activa en las clases.

La calificación de la evaluación continua (en la escala 0-10) es

$$\min(A + 0.10 * B, 10),$$

donde

- A (en la escala 0-10) es la calificación de los ejercicios prácticos;
- B (en la escala 0-10) es la calificación correspondiente a (b).

Los estudiantes que no hayan seguido la evaluación continua podrán realizar un examen final con un valor del 60% de la calificación final.

La calificación en la convocatoria extraordinaria se establece por medio de la entrega de un conjunto de ejercicios prácticos.

Se proporcionan más detalles en Aula Global. La evaluación está sujeta a modificaciones debido al desarrollo del curso y/o al calendario académico.

Peso porcentual del Examen Final:	0
Peso porcentual del resto de la evaluación:	100

BIBLIOGRAFÍA BÁSICA

- Hadley, W. y Grolemund, G. R for Data Science, O'Reilly, 2017
- Xie, Y., Allaire, J.J. y Grolemund, G. R Markdown, CRC Press/Chapman & Hall, 2019

RECURSOS ELECTRÓNICOS BÁSICOS

- Kuhn, M. y Silge, J. . Tidy Modeling with R: <https://www.tmrw.org/>

