

Curso Académico: ( 2022 / 2023 )

Fecha de revisión: 16-05-2022

Departamento asignado a la asignatura: Departamento de Estadística

Coordinador/a: GARCIA PORTUGUES, EDUARDO

Tipo: Obligatoria Créditos ECTS : 3.0

Curso : 1 Cuatrimestre : 2

**REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)**

Matemáticas para Data Science  
Probabilidad  
Inferencia Estadística  
Programación en R  
Análisis Multivariante  
Modelos de Regresión  
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.
- 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.
- 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.
- CE10: Aplicar la modelización estadística en el tratamiento de problemas relevantes en el campo científico.

**\* Resultados del aprendizaje**

Adquisición de conocimientos sobre: 1) estimadores núcleo de la densidad y sus aplicaciones; 2) métodos de regresión no paramétricos basados en suavizado; 3) contrastes de hipótesis no paramétricos.

## DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

Este curso está diseñado para dar una visión panorámica de varias herramientas disponibles en la estadística no paramétrica, a un nivel intermedio-avanzado. Esta visión abarca en profundidad los principales conceptos en la estimación de las funciones de densidad y regresión mediante métodos núcleo (con sus correspondientes aplicaciones), y la descripción de varias pruebas no paramétricas comunes. Se hace hincapié en proporcionar las principales ideas sobre los fundamentos estadísticos/matemáticos de los métodos y en mostrar la aplicación efectiva de los métodos mediante el uso de software estadístico. Esto se logra mediante una mezcla de teoría y código reproducible.

### 1. Introducción

- 1.1. Repaso de probabilidad
- 1.2. Repaso de distribuciones
- 1.3. Repaso de la convergencia estocástica
- 1.4. Notación OP y oP
- 1.5. Repaso de herramientas analíticas básicas
- 1.6. ¿Por qué la estadística no paramétrica?

### 2. Estimación núcleo de la densidad I

- 2.1. Histogramas
- 2.2. Estimación núcleo de la densidad
- 2.3. Propiedades asintóticas
- 2.4. Selección del ancho de banda
- 2.5. Cuestiones prácticas
- 2.6. Estimación núcleo de la densidad con  $k_s$

### 3. Estimación núcleo de la densidad II

- 3.1. Estimación núcleo de la densidad multivariante
- 3.2. Propiedades asintóticas
- 3.3. Selección del ancho de banda
- 3.4. Aplicaciones de la estimación núcleo de la densidad

### 4. Estimación núcleo de la regresión I

- 4.1. Estimación núcleo de la regresión
- 4.2. Propiedades asintóticas
- 4.3. Selección del ancho de banda
- 4.4. Regresograma
- 4.5. Estimación núcleo de la regresión con  $n_p$

### 5. Estimación núcleo de la regresión II

- 5.1. Regresión núcleo con datos multivariantes mixtos
- 5.2. Selección del ancho de banda
- 5.3. Predicción e intervalos de confianza
- 5.4. Verosimilitud local

### 6. Tests no paramétricos

- 6.1. Tests de bondad de ajuste para distribuciones
- 6.2. Comparación de distribuciones
- 6.3. Tests de independencia
- 6.4. Otros tests

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 teoría (descripción de los métodos) y práctica (implementación y aplicación de los métodos). Se emplea el lenguaje estadístico R. Se espera que los estudiantes traigan sus propios portátiles para experimentar con el código durante algunas partes de 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.

- 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) dos cuestionarios sobre ideas clave y conceptos teóricos;

(b) un conjunto de ejercicios prácticos a realizar en grupo;

(c) participación activa en las clases y realización de ejercicios voluntarios.

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

$\min(0.5 * A + 0.5 * B + 0.10 * C, 10)$ ,

donde

- A (en la escala 0-10) es la calificación ponderada de los cuestionarios;

- B (en la escala 0-10) es la calificación de los ejercicios prácticos;

- C (en la escala 0-10) es la calificación correspondiente a (c).

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 un cuestionario y una 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.

|  |     |
|--|-----|
| <b>Peso porcentual del Examen Final:</b>           | 0   |
| <b>Peso porcentual del resto de la evaluación:</b> | 100 |

#### BIBLIOGRAFÍA BÁSICA

- Chacón, J. E. y Duong, T. Multivariate Kernel Smoothing and Its Applications, Chapman and Hall/CRC, 2018

- Fan, J. y Gijbels, I. Local polynomial modelling and its applications, Chapman & Hall, 1996

- Li, Q. y Racine, J. S. Nonparametric Econometrics, Princeton University Press, 2007

- Wand, M. P. y Jones, M. C. Kernel Smoothing, Chapman & Hall, 1995

#### BIBLIOGRAFÍA COMPLEMENTARIA

- Wasserman, L. All of Nonparametric Statistics, Springer-Verlag, 2006

- Wasserman, L. All of Statistics, Springer-Verlag, 2004