

Máster en Técnicas Estadísticas

DATOS IDENTIFICATIVOS DE LA MATERIA

Código de la materia: P1062205

Nombre de la materia: Regresión no paramétrica y semiparamétrica

Número de créditos ECTS: 5

Curso académico: 2026/2027

Profesorado:

María Alonso Pena (coordinadora, 3 créditos)

Jose Ameijeiras Alonso (2 créditos)

OBJETIVOS DE LA MATERIA

El objetivo de esta materia es que el estudiantado se familiarice con las técnicas de regresión no paramétrica y semiparamétrica, con especial énfasis en su aplicación práctica. Estos objetivos se concretan en los siguientes resultados del aprendizaje:

- Conocer las principales técnicas no paramétricas y semiparamétricas de estimación de la función de regresión.
- Saber escoger el modelo de regresión no paramétrico o semiparamétrico adecuado para analizar la dependencia existente en datos complejos procedentes de situaciones reales.
- Conocer las limitaciones de las técnicas no paramétricas en el análisis de situaciones reales con un alto número de variables.
- Ser autónomo en el análisis de datos en entornos aplicados multidisciplinares utilizando técnicas no paramétricas y semiparamétricas.

CONTENIDOS DE LA MATERIA

Tema 1. Estimación de la regresión mediante splines.

Mínimos cuadrados penalizados. Spline de interpolación: spline cúbico natural y spline de suavización. Spline de regresión: base de splines y regresión spline penalizada. Generalización multivariante thin-plate spline y tensor producto spline.

Tema 2. Modelos generalizados.

Presentación de los modelos generalizados. Modelo logístico y de conteo suavizados.

Extensión de los algoritmos de estimación para los modelos generalizados. Modelo generalizado penalizado.

Tema 3. Introducción a la estimación tipo núcleo.

Estimación tipo núcleo de la función de densidad. Selección del parámetro de suavizado. Modificaciones del estimador tipo núcleo de la densidad. Estimación de las derivadas de la función de densidad. Análisis exploratorio basado en el estimador núcleo.

Tema 4. Regresión tipo núcleo.

El regresograma. El estimador de Nadaraya-Watson. El estimador polinómico local. Selección del parámetro de suavizado: plug-in y validación cruzada. Estimación de las derivadas de la función de regresión. Estimación no paramétrica de la regresión multivariante. Verosimilitud local: regresión logística tipo núcleo.

Tema 5. Estimación de la regresión por vecinos más próximos.

El estimador clásico por los k vecinos más próximos. Variantes del estimador por vecinos más próximos. El estimador Loess. Comparación con la suavización ordinaria.

Tema 6. Modelos parcialmente lineales y modelos aditivos.

Presentación de los modelos parcialmente lineales y de los modelos aditivos. El problema de identificación y algoritmos de estimación.

Tema 7. Modelos single-index.

Presentación del modelo single-index. El problema de identificación y estimación del modelo. Extensión de los modelos single-index.

BIBLIOGRAFÍA BÁSICA Y COMPLEMENTARIA

Básica

- [1] Bowman, A.W. y Azzalini, A. (1997). Applied smoothing techniques for data analysis. Oxford University Press.
- [2] Härdle, W. (1990). Applied nonparametric regression. Econometric Society Monographs, Cambridge University Press.
- [3] Härdle, W., Müller, M., Sperlich, S. y Werwatz, A. (2004). Nonparametric and Semiparametric Models. Springer.
- [4] Wand, M.P. y Jones, M.C. (1995). Kernel Smoothing. Chapman Hall.
- [5] Wasserman, L. (2005). All of Nonparametric Statistics. Springer.
- [6] Wood, S.N. (2006). Generalized Additive Models. An Introduction with R. Chapman and Hall.

Complementaria

- [1] Fan, J. y Gijbels, I. (1996). Local Polynomial Modelling and Its Applications. Chapman and Hall.
- [2] Green, P.J. y Silverman, B.W. (1994). Nonparametric regression and generalized linear models: A roughness penalty approach. Chapman and Hall.
- [3] Hastie, T. y Tibshirani, R. (1990). Generalized Additive Models. Chapman and Hall.
- [4] Scott, D.W. (1992). Multivariate Density Estimation: Theory, Practice, and Visualization. John Wiley and Sons.
- [5] Silverman, B.W. (1986). Density Estimation for Statistics and Data Analysis. Chapman and Hall.
- [6] Simonoff, J.S. (1996). Smoothing Methods in Statistics. Springer.
- [7] Wahba, G. (1990). Spline Models for Observation Data. Society for Industrial and Applied Mathematics.

COMPETENCIAS

En esta materia se trabajarán las competencias básicas, generales y transversales recogidas en la memoria del título. Se indican a continuación cuáles son las competencias específicas, que se potenciarán en esta materia:

E1 - Conocer, identificar, modelar, estudiar y resolver problemas complejos de estadística e investigación operativa, en un contexto científico, tecnológico o profesional, surgidos en aplicaciones reales.

E2 - Desarrollar autonomía para la resolución práctica de problemas complejos surgidos en aplicaciones reales y para la interpretación de los resultados de cara a la ayuda en la toma de decisiones.

E3 - Adquirir conocimientos avanzados de los fundamentos teóricos subyacentes a las distintas metodologías de la estadística y la investigación operativa, que permitan su desarrollo profesional especializado.

E4 - Adquirir las destrezas necesarias en el manejo teórico-práctico de la teoría de la probabilidad y las variables aleatorias que permitan su desarrollo profesional en el ámbito científico/académico, tecnológico o profesional especializado y multidisciplinar.

E5 - Profundizar en los conocimientos en los fundamentos teórico-prácticos especializados del modelado y estudio de distintos tipos de relaciones de dependencia entre variables estadísticas.

E6 - Adquirir conocimientos teórico-prácticos avanzados de distintas técnicas matemáticas, orientadas específicamente a la ayuda en la toma de decisiones, y desarrollar capacidad de reflexión para evaluar y decidir entre distintas perspectivas en contextos complejos.

E8 - Adquirir conocimientos teórico-prácticos avanzados de las técnicas destinadas a la realización de inferencias y contrastes relativos a variables y parámetros de un modelo estadístico, y saber aplicarlos con autonomía suficiente un contexto científico, tecnológico o profesional.

E9 - Conocer y saber aplicar con autonomía en contextos científicos, tecnológicos o profesionales, técnicas de aprendizaje automático y técnicas de análisis de datos de alta dimensión (big data).

E10 - Adquirir conocimientos avanzados sobre metodologías para la obtención y el tratamiento de datos desde distintas fuentes, como encuestas, internet, o entornos "en la nube".

METODOLOGÍA DOCENTE

La enseñanza constará de clases expositivas e interactivas, así como de la tutorización del aprendizaje y de los trabajos encomendados al estudiantado. Se proporcionará material para el seguimiento del curso, así como otro material orientativo del aprendizaje del software. En las clases expositivas e interactivas se resolverán ejemplos mediante el software R, por lo que es necesario que el estudiantado disponga en el aula de un ordenador.

La actividad presencial, junto con el correspondiente y necesario trabajo personal del estudiantado para su preparación, es valorada con tres créditos ECTS. Esta carga de trabajo incluye el examen final. Se considera que es suficiente una hora y media de trabajo personal para la preparación de cada sesión presencial de tipo teórico-práctico. Los otros dos créditos ECTS de la materia corresponden a trabajos prácticos que el estudiantado tendrá que elaborar a lo largo del curso.

A continuación, se presenta una aproximación de las horas que se dedicarán a cada tema.

TEMA 1. SPLINES (6h expositivas, 5h de laboratorio)

TEMA 2. MOD GENER (2h expositiva, 1h de laboratorio)

TEMA 3. INTRO NÚCLEO (1h expositiva, 2h de laboratorio)

TEMA 4. REG NÚCLEO (5h expositivas, 3h de laboratorio)

TEMA 5. KNN (2h expositivas, 1h de laboratorio)

TEMA 6. PLM Y ADITIVOS (3h expositivas, 2h de laboratorio)

TEMA 7. SINGLE-INDEX (1h expositivas, 1h de laboratorio)

En el caso de que alguna de las actividades docentes deba realizarse en remoto, se utilizará la plataforma MS Teams.

CRITERIOS Y MÉTODOS DE EVALUACIÓN

Evaluación continua (60%): la evaluación continua consistirá en la preparación de un documento técnico sobre alguna temática relacionada con la materia. El estudiantado revisará literatura estadística y, en función de la temática, realizará análisis de datos reales y/o estudios de simulación. Los trabajos serán presentados oralmente. Además, se incluirá en el proceso una evaluación por pares, tanto del documento técnico como de la presentación. Si el trabajo se realiza en grupo, el profesorado considerará estrategias de evaluación que permitan discernir la nota de cada componente del equipo. Para el cálculo de la calificación se tendrá en cuenta el documento técnico, la presentación y la evaluación de los pares. La calificación se conservará entre las oportunidades (ordinaria y extraordinaria) dentro de la convocatoria de cada curso. Con las distintas actividades que se propondrán a lo largo del curso, se valorará el nivel de adquisición de las competencias básicas y generales, CB6-CB10 y CG1-CG5. También se evaluará el nivel alcanzado en las competencias transversales CT1-CT5 y de las competencias específicas E2, E6, E9 y E10

Prueba final (40%): la prueba final constará de varias cuestiones teórico-prácticas sobre los contenidos de la materia, dentro de las que se podrá incluir la interpretación de resultados obtenidos con el programa utilizado en la docencia interactiva. En el examen se evaluará la adquisición de las competencias específicas E1, E3, E4, E5 y E8.

Presentación a la evaluación: se considera que un/a estudiante concurre a una convocatoria cuando participa en alguna actividad de evaluación, bien sea de evaluación continua o de asistencia al examen. El peso de la evaluación continua en la oportunidad extraordinaria de recuperación (pruebas de julio) será el mismo que en la evaluación ordinaria. En la segunda oportunidad de evaluación (recuperación), se realizará un examen y la nota final será el máximo de tres cantidades: la nota de la evaluación ordinaria, la nota del nuevo examen y la media ponderada del nuevo examen y la evaluación continua.

Debe notarse que, para los casos de realización fraudulenta de ejercicios o pruebas, será de aplicación lo recogido en la “Normativa de evaluación del rendimiento académico de los estudiantes y de revisión de calificaciones”.

TIEMPO DE ESTUDIO Y DE TRABAJO PERSONAL QUE DEBE DEDICAR UN ESTUDIANTE PARA SUPERAR LA MATERIA

Cada crédito ECTS se traduce en 7 horas de clase de tipo presencial. Se estima que el estudiantado necesitará una hora para preparar el material correspondiente a cada hora presencial, previa a la propia clase. Posteriormente, precisará de hora y media para la comprensión global de los contenidos, incluyendo las actividades asociadas a ejercicios y otras tareas. En total resultarán 24.5 horas por crédito ECTS.

RECOMENDACIONES PARA EL ESTUDIO DE LA MATERIA

Para superar con éxito la materia es aconsejable la asistencia a las clases, siendo fundamental el seguimiento diario del trabajo realizado en el aula. También es recomendable que el estudiantado sea capaz de utilizar el software estadístico R para explorar las posibilidades de las diversas técnicas no paramétricas explicadas a lo largo del curso. Además, para un mejor aprendizaje de la materia, es conveniente tener presente el sentido práctico de los métodos introducidos a lo largo del curso.

En el caso de que la docencia se desarrolle de manera presencial, el profesorado informará en clase de cuáles son los objetivos del aprendizaje a alcanzar y los contenidos a trabajar durante la semana que corresponda. En caso de que la docencia se desarrolle parcial o totalmente de manera no presencial, se elaborará el “plan de trabajo semanal” que se facilitará al estudiantado al comienzo de cada semana. En este plan se darán una serie de pautas para ayudar al estudiantado en su aprendizaje autónomo, y donde se especificarán los contenidos a trabajar y las actividades recomendadas.

RECURSOS PARA EL APRENDIZAJE

El estudiantado dispondrá de notas elaboradas por el profesorado de la materia, así como de la bibliografía recomendada en este programa. Para la difusión del material propio de la asignatura, se hará uso de la web del máster.

OBSERVACIONES

El desarrollo de los contenidos de la materia se realizará teniendo en cuenta que las competencias a adquirir por el estudiantado deben cumplir con el nivel MECES3. Los contenidos que se incluyen en esta materia son novedosos y altamente especializados. Se trabajará sobre la correcta formulación de modelos, la construcción de estimadores y la validación y el análisis de las distintas propuestas estudiadas.

Para los casos de realización fraudulenta de ejercicios o pruebas, será de aplicación lo recogido en las respectivas normativas de las universidades participantes en el Máster en Técnicas Estadísticas.

Esta guía y los criterios y metodologías en ella descritos están sujetos a las modificaciones que se deriven de normativas y directrices de las universidades participantes en el Máster en Técnicas Estadísticas.