

Guía docente de la asignatura

Fecha de aprobación por la Comisión Académica: 13/07/2022

Análisis de Datos Funcionales
(M42/56/2/32)**Máster**

Máster Universitario en Estadística Aplicada

MÓDULO

Módulo II: Formación para la Investigación

RAMA

Ciencias

CENTRO RESPONSABLE DEL TÍTULO

Escuela Internacional de Posgrado

Semestre

Segundo

Créditos

4

Tipo

Optativa

Tipo de enseñanza

Enseñanza Virtual

PRERREQUISITOS Y/O RECOMENDACIONES

Recomendaciones: Para realizar con éxito este curso convendría que el alumno tenga una formación básica en análisis matemático, aproximación de funciones, procesos estocásticos, análisis multivariante y modelos de regresión. Esta formación está garantizada para los alumnos que hayan realizado el Grado en Estadística, el Grado en Matemáticas o las Diplomaturas y Licenciaturas equivalentes. Para los alumnos que no tengan esta formación se recomienda realizar el curso del máster “Técnicas estadísticas multivariantes y aplicaciones”.

BREVE DESCRIPCIÓN DE CONTENIDOS (Según memoria de verificación del Máster)

Las variables funcionales se caracterizan por la evolución de una variable a lo largo del tiempo (proceso estocástico), de modo que los valores que toman son, en general, funciones de uno o varios argumentos en lugar de vectores como en análisis multivariante clásico. La imposibilidad de medir la mayoría de estas variables de forma continua, unida a la complejidad teórica de muchos de los métodos estadísticos disponibles para su análisis llevan a que se manejen resúmenes periódicos que constituyen las series temporales contenidas normalmente en los anuarios estadísticos. Aunque existen muchas técnicas para la modelización y predicción de datos temporales discretos, la mayoría de ellas como por ejemplo la teoría clásica de Box-Jenkins, imponen que se verifiquen hipótesis bastante restrictivas como estacionariedad, observaciones igualmente espaciadas o pertenencia a una clase de procesos específica.

Estos problemas han planteado la necesidad de desarrollar metodologías potentes que permitan, en la práctica, la modelización y predicción a partir de datos funcionales dando lugar a una



especialidad estadística muy reciente conocida con el nombre de Análisis de Datos Funcionales (FDA), que está siendo objeto de un estudio intensivo por parte de los investigadores en los últimos 30 años. Los primeros trabajos sobre FDA estuvieron dedicados a la generalización de la técnica de reducción de dimensión Análisis en Componentes Principales (ACP) al caso funcional. Posteriormente las investigaciones en FDA se han centrado en el desarrollo de modelos de regresión funcional y sus aplicaciones en diferentes campos como la economía, el medioambiente o las ciencias de la salud, así como la extensión de todo tipo de metodologías estadísticas para el análisis de este tipo de datos.

Objetivos particulares

- Conocer las nociones básicas sobre variables funcionales (procesos estocásticos) de segundo orden.
- Manejar los métodos matemáticos de aproximación de funciones de cuadrado integrable (interpolación, aproximación mínimo cuadrática, etc) a partir de bases de funciones (trigonométricas, splines, wavelets, etc).
- Estudiar la técnica de reducción de dimensión Análisis en Componentes Principales Funcional (ACPF) y métodos de estimación a partir de observaciones en tiempo discreto de las funciones muestrales.
- Conocer la formulación, estimación muestral e implementación computacional, así como la aplicación con datos reales e interpretación de resultados, de modelos de predicción en componentes principales (PCP) de una variable funcional. Relacionar los modelos PCP con los modelos más generales de regresión lineal funcional.
- Conocer la formulación, estimación muestral e implementación computacional, así como la aplicación con datos reales e interpretación de resultados, de modelos de regresión logística para estimar una variable binaria a partir de un predictor funcional relacionado.

COMPETENCIAS

COMPETENCIAS BÁSICAS

- CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
- CB7 - Que los estudiantes sepan 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 - Que los estudiantes sean capaces de 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 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.
- CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

COMPETENCIAS GENERALES



- CG01 - Los titulados han de 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.
- CG02 - Los titulados han de ser capaces de 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.
- CG03 - Los titulados han de saber comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades
- CG04 - Los titulados deben poseer las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.
- CG05 - Los titulados han de demostrar una comprensión sistemática del campo de estudio y el dominio de las habilidades y métodos de investigación relacionados con dicho campo.
- CG06 - Los titulados deben demostrar la capacidad de concebir, diseñar, poner en práctica y adoptar un proceso sustancial de investigación con seriedad académica.
- CG07 - Los titulados han de realizar una contribución a través de una investigación original que amplíe las fronteras del conocimiento desarrollando un corpus sustancial, del que parte merezca la publicación referenciada a nivel nacional o internacional.
- CG08 - Los titulados deben ser críticos en el análisis, evaluación y síntesis de ideas nuevas y complejas.
- CG09 - Los titulados deben saber comunicarse con sus colegas, con la comunidad académica en su conjunto y con la sociedad en general acerca de sus áreas de conocimiento.
- CG10 - Los titulados han de ser capaces de fomentar, en contextos académicos y profesionales, el avance tecnológico, social o cultural dentro de una sociedad basada en el conocimiento.

COMPETENCIAS ESPECÍFICAS

- CE01 - Conocer métodos para el Análisis de Datos
- CE02 - Conocer diferentes técnicas de Muestreo
- CE03 - Adquirir conocimientos avanzados en Probabilidad y Procesos Estocásticos
- CE04 - Profundizar en las técnicas de Modelización Estocástica
- CE05 - Adquirir conocimientos avanzados en Inferencia Estadística
- CE07 - Saber identificar y aplicar diferentes Modelos Económicos
- CE08 - Conocer técnicas de teoría de Fiabilidad
- CE10 - Dominar el uso de diferentes entornos de Computación Estadística
- CE12 - Ser capaz de resolver problemas a través de técnicas de Simulación Estocástica
- CE13 - Saber llevar a cabo el diseño, programación e implantación programas de computación estadística
- CE15 - Ser capaz de identificar la información relevante para resolver un problema
- CE16 - Utilizar correcta y racionalmente programas de ordenador de tipo estadístico
- CE17 - Adquirir capacidades de elaboración y construcción de modelos y su validación
- CE18 - Ser capaz de realizar un análisis de datos
- CE19 - Saber gestionar bases de datos
- CE20 - Ser capaz de realizar una correcta representación gráfica de datos
- CE21 - Conocer, identificar y seleccionar fuentes estadísticas
- CE22 - Ser capaz de interpretar resultados a partir de modelos estadísticos
- CE23 - Adquirir capacidad para elaborar previsiones y escenarios
- CE24 - Ser capaz de extraer conclusiones y redactar informes



COMPETENCIAS TRANSVERSALES

- CT01 - Mostrar interés por la calidad y la excelencia en la realización de diferentes tareas
- CT02 - Comprender y defender la importancia que la diversidad de culturas y costumbres tienen en la investigación o práctica profesional
- CT03 - Tener un compromiso ético y social en la aplicación de los conocimientos adquiridos
- CT04 - Ser capaz de trabajar en equipos interdisciplinarios para alcanzar objetivos comunes desde campos expertos diferenciados.
- CT05 - Incorporar los principios del Diseño Universal en el desempeño de su profesión

RESULTADOS DE APRENDIZAJE (Objetivos)

El alumno sabrá/comprenderá:

- Las nociones básicas sobre variables funcionales (procesos estocásticos) de segundo orden.
- Los métodos matemáticos de aproximación de funciones de cuadrado integrable (interpolación, aproximación mínimo cuadrática, etc) a partir de bases de funciones (trigonométricas, splines, wavelets, etc)
- La técnica de reducción de dimensión Análisis en Componentes Principales Funcional (ACPF) y métodos de estimación a partir de observaciones en tiempo discreto de las funciones muestrales.
- La formulación, estimación muestral e implementación computacional, así como la aplicación con datos reales e interpretación de resultados, de modelos de predicción en componentes principales (PCP) de una variable funcional.
- Relacionar los modelos PCP con los modelos más generales de regresión lineal funcional.
- La formulación, estimación muestral e implementación computacional, así como la aplicación con datos reales e interpretación de resultados, de modelos de regresión logística para estimar una variable binaria a partir de un predictor funcional relacionado.

El alumno será capaz:

- Desarrollar aplicaciones con datos funcionales reales, eligiendo el modelo más adecuado de los estudiados para alcanzar el objetivo planteado, así como extraer conclusiones e investigar la posibilidad de usar modelos funcionales más complejos en caso necesario.

PROGRAMA DE CONTENIDOS TEÓRICOS Y PRÁCTICOS

TEÓRICO

- Tema 1: Introducción al Análisis de Datos Funcionales.
- Tema 2: Representación de datos funcionales mediante bases de funciones.
- Tema 3: Análisis en componentes principales funcional: formulación, estimación y aproximación.
- Tema 4: Modelos de regresión lineal funcional: casos de respuesta escalar y respuesta funcional. Relación con los modelos PCP de predicción en componentes principales.
- Tema 5: Modelos de regresión logística funcional para respuesta binaria. Aplicaciones en Medioambiente y en Ciencias de la Salud.



PRÁCTICO

El temario práctico consistirá en aplicar los modelos teóricos estudiados con el paquete fda del software estadístico R y con la aplicación web StatFda (/www.statfda.com).

BIBLIOGRAFÍA

BIBLIOGRAFÍA FUNDAMENTAL

- Aguilera, A.M., Escabias, M. and Valderrama, M.J. (2008): Discussion of different logistic models with functional data. Application to Systemic Lupus Erythematosus. *Computational Statistics and Data Analysis*, 53(1), 151-163.
- Escabias, M., Aguilera, A.M. and Valderrama, M.J. (2004): Principal component estimation of functional logistic regression: discussion of two different. *Journal of Nonparametric Statistics*, 16 (3-4), 365-38.
- Escabias, M., Aguilera, A.M. and Valderrama, M.J. (2005): Modeling environmental data by functional principal component logistic regression. *Environmetrics*, 16 (1), 95-107.
- Escabias, M., Aguilera, A.M. and Aguilera-Morillo, M.J. (2014): Functional PCA and base-line logit models. *Journal of classification*, 31, 296-324.
- Ferraty, F. And Vieu, P. (2006): *Nonparametric Functional Data Analysis. Theory and Practice*. Springer-Verlag.
- Horváth, L. and Kokoszka, P. (2012): *Inference for functional data with applications*. Springer-Verlag.
- Ocaña, F.A., Aguilera, A.M. and Escabias, M. (2007): Computational considerations in functional principal component analysis. *Computational Statistics*, 22, 449-466.
- Ramsay J.O. y Silverman B.W. (1997, 2005): *Functional Data Analysis. (First and Second editions)*. Springer-Verlag.
- Ramsay J.O. y Silverman B.W. (2002): *Applied Functional Data Analysis*. Springer-Verlag.
- Ramsay J.O., Hooker, G. and Graves, S. (2009). *Functional data analysis with R and MATLAB*. Springer-Verlag..
- Valderrama M.J., Aguilera A.M. y Ocaña F.A. (2000): *Predicción Dinámica mediante Análisis de Datos Funcionales*. La Muralla, Madrid.

BIBLIOGRAFÍA COMPLEMENTARIA

- Aguilera, A.M., Aguilera-Morillo, M.C. and Preda, C. (2016): Penalized versions of functional PLS regression. *Chemometrics and intelligent laboratory systems*, 154, 80-92.
- Aguilera, A.M. and Aguilera-Morillo, M.C. (2013): Penalized PCA approaches for B-spline expansions of smooth functional data. *Applied Mathematics and Computation*, 219, 7805-7819.
- Aguilera, A.M., Escabias, M., Preda, C. and Saporta, G. (2010): Estimating PLS regression with functional data by using basis expansion methods. *Chemometrics and intelligent laboratory systems*, 104(2), 289-305.
- Aguilera-Morillo, M.C., Aguilera, A.M., Escabias, M. and Valderrama, M.J. (2013): Penalized spline approaches for functional logit regression. *Test*, 22(2), 251-277..
- Ratcliffe, S. J., Heller, G. Z. y Leader, L. (2002b): Functional data analysis with application to periodically stimulated foetal heart rate data. II: Functional logistic regression. *Statistics in Medicine*, 21, 1115-1127.
- Shi, J.Q. and Choi, T. (2011): *Gaussian process regression analysis for functional data*. CRC Press.
- Zhang, J-T. (2014): *Analysis of Variance for functional data*. CRC Press.



ENLACES RECOMENDADOS

- Wel del Prof. J. Ramsay, creador de la terminología FDA: <https://www.psych.mcgill.ca/misc/fda/>
- Curso breve sobre FDA: <http://faculty.bscb.cornell.edu/~hooker/ShortCourseHandout.pdf>
- Revisión sobre FDA: <http://anson.ucdavis.edu/~mueller/Review151106.pdf>

METODOLOGÍA DOCENTE

- MD01 Lección magistral/expositiva
- MD02 Sesiones de discusión y debate
- MD03 Resolución de problemas y estudio de casos prácticos
- MD04 Prácticas de laboratorio o clínicas
- MD05 Seminarios
- MD06 Ejercicios de simulación
- MD07 Análisis de fuentes y documentos
- MD08 Realización de trabajos en grupo
- MD09 Realización de trabajos individuales

EVALUACIÓN (instrumentos de evaluación, criterios de evaluación y porcentaje sobre la calificación final)

EVALUACIÓN ORDINARIA

El artículo 17 de la Normativa de Evaluación y Calificación de los Estudiantes de la Universidad de Granada establece que la convocatoria ordinaria estará basada preferentemente en la evaluación continua del estudiante, excepto para quienes se les haya reconocido el derecho a la evaluación única final. Como se ha indicado en la metodología docente, la asignatura se configura en cinco bloques o temas que se evalúan de manera independiente. Para cada bloque se propondrá una actividad de evaluación que podrá consistir en una de las siguientes alternativas:

- La realización de un trabajo en el que se describa resumidamente uno de los modelos teóricos estudiados en el bloque y se aplique a un conjunto de datos reales o simulados, realizando un informe en el que se explique tanto la obtención de resultados con R o StatFda como la interpretación de los mismos.
- La respuesta a un cuestionario online basado en la resolución de un ejercicio o caso práctico con datos reales o simulados para el ajuste de los modelos teóricos explicados en la teoría del bloque.

La evaluación del primer bloque supondría un 12% y cada uno de los otros 4 bloques un 22% de la calificación final de la asignatura. En la calificación final se tendrá en cuenta la realización de la actividad propuesta (90%) y la participación en foros y tutorías (10%).

EVALUACIÓN EXTRAORDINARIA

El artículo 19 de la Normativa de Evaluación y Calificación de los Estudiantes de la Universidad de Granada establece que los estudiantes que no hayan superado la asignatura en la convocatoria ordinaria dispondrán de una convocatoria extraordinaria. A ella podrán concurrir todos los



estudiantes, con independencia de haber seguido o no un proceso de evaluación continua. De esta forma, el estudiante que no haya realizado la evaluación continua tendrá la posibilidad de obtener el 100% de la calificación mediante la realización de una prueba y/o trabajo.

Para la convocatoria extraordinaria se propondrán el mismo tipo de actividades propuestas en la evaluación ordinaria pero en el periodo establecido para ello.

EVALUACIÓN ÚNICA FINAL

El artículo 8 de la Normativa de Evaluación y Calificación de los Estudiantes de la Universidad de Granada establece que podrán acogerse a la evaluación única final, el estudiante que no pueda cumplir con el método de evaluación continua por causas justificadas.

Para acogerse a la evaluación única final, el estudiante, en las dos primeras semanas de impartición de la asignatura o en las dos semanas siguientes a su matriculación si ésta se ha producido con posterioridad al inicio de las clases, lo solicitará, a través del procedimiento electrónico, a la Coordinación del Máster, quien dará traslado al profesorado correspondiente, alegando y acreditando las razones que le asisten para no poder seguir el sistema de evaluación continua.

Para la evaluación única final se propondrán el mismo tipo de actividades propuestas en la evaluación ordinaria.

INFORMACIÓN ADICIONAL

Metodología docente

Por tratarse de una asignatura que se imparte de manera 100% virtual, la metodología docente tiene como principal eje de actuación el aprendizaje autónomo del alumno a partir de una temporización y secuenciación de contenidos que planifica el profesorado. Para lograr el aprendizaje autónomo en cada bloque de contenido, el profesorado facilita al alumnado el siguiente material:

- Guía de aprendizaje autónomo con la explicación del modo de trabajo que el alumnado debe seguir para conseguir un aprendizaje autónomo eficaz.
- Material docente teórico con las metodologías de ADF que el alumnado deberá estudiar apoyado de ejemplos prácticos que ilustran los conceptos teóricos.
- Material docente práctico que aborda la resolución de casos prácticos con datos reales de los métodos introducidos en el material teórico. En este material se ilustra el uso del software libre R y/o la aplicación StatFda (www.statfda.com) para llevar a cabo los ajustes de los modelos teóricos y los cálculos necesarios que involucran dichos ajustes.
- Foro de cada tema en el que el alumnado podrá discutir las dificultades y avances de su aprendizaje, y cuya participación será tenida en cuenta para la evaluación final de la materia.
- Actividad de evaluación del bloque.

Este material docente podrá estar apoyado con material audiovisual creado al efecto. Todo el material anteriormente descrito se facilita al alumnado a través de la plataforma Moodle PRADO que la Universidad de Granada tiene habilitada para la docencia virtual.

La materia se imparte durante un periodo de unas 10 semanas aproximadamente. Dado que la



materia comprende 5 bloques, la metodología docente asigna 2 semanas a cada bloque. La secuenciación de cada bloque se organiza del siguiente modo:

- El primer día de comienzo de un bloque se facilita al alumnado la guía de autoaprendizaje, el material teórico y el material práctico.
- El alumno trabajará los contenidos del bloque durante dos semanas.
- Al menos una vez por semana se planificarán tutorías virtuales a través de la plataforma de videoconferencia que facilite la Universidad de Granada.
- Tras dos semanas de autoaprendizaje se facilita al alumnado la actividad de evaluación que tendrá que entregar o resolver en los 3-4 últimos días asignados al bloque.

