



Guía docente				
Datos Identificativos				2021/22
Asignatura (*)	Modelos de Regresión	Código	614G02012	
Titulación	Grao en Ciencia e Enxeñaría de Datos			
Descritores				
Ciclo	Periodo	Curso	Tipo	Créditos
Grado	1º cuatrimestre	Segundo	Obligatoria	6
Idioma	Castellano			
Modalidad docente	Presencial			
Prerrequisitos				
Departamento	Matemáticas			
Coordinador/a	Vilar Fernandez, Jose Antonio	Correo electrónico	jose.vilarf@udc.es	
Profesorado	Vilar Fernandez, Jose Antonio	Correo electrónico	jose.vilarf@udc.es	
	Vilar Fernandez, Juan Manuel		juan.vilar@udc.es	
Web				
Descripción general	Estudio de modelos de regresión paramétricos y no paramétricos. En el contexto paramétrico, se abordan procedimientos de inferencia, diagnóstico y predicción para el modelo lineal simple, el modelo lineal múltiple y el modelo logístico. La modelización no paramétrica abarcará regresión tipo kernel y por splines, además de una introducción a modelos de regresión parcialmente lineales y a modelos aditivos.			
Plan de contingencia	<p>1. Modificaciones en los contenidos Ninguna.</p> <p>2. Metodologías Las metodologías docentes de esta materia pueden llevarse a cabo con independencia del grado de presencialidad en la docencia, empleando recursos telemáticos.</p> <p>3. Mecanismos de atención personalizada al alumnado Correo electrónico y uso de recursos telemáticos (Teams)</p> <p>4. Modificacines en la evaluación El método de evaluación no precisa ninguna adaptación. En el caso de que el examen no se pueda realizar de forma presencial, se realizará virtualmente a través de herramientas telemáticas.</p> <p>5. Modificaciones de la bibliografía o webgrafía Ninguna</p>			

Competencias del título	
Código	Competencias del título
A17	CE17 - Capacidad para la construcción, validación y aplicación de un modelo estocástico de un sistema real a partir de los datos observados y el análisis crítico de los resultados obtenidos.
A18	CE18 - Capacidad para el análisis de un conjunto de datos y la selección y aplicación de las técnicas de inferencia estadística y de regresión más adecuadas para la adquisición de conocimiento para la toma de decisiones.
A20	CE20 - Conocimiento de las herramientas informáticas en el campo del análisis de los datos y modelización estadística, y capacidad para seleccionar las más adecuadas para la resolución de problemas.
B2	CB2 - Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio
B3	CB3 - Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética



B7	CG2 - Elaborar adecuadamente y con cierta originalidad composiciones escritas o argumentos motivados, redactar planes, proyectos de trabajo, artículos científicos y formular hipótesis razonables.
B8	CG3 - Ser capaz de mantener y extender planteamientos teóricos fundados para permitir la introducción y explotación de tecnologías nuevas y avanzadas en el campo.
B9	CG4 - Capacidad para abordar con éxito todas las etapas de un proyecto de análisis de datos: exploración previa de los datos, preprocesado, análisis, visualización y comunicación de resultados.
B10	CG5 - Ser capaz de trabajar en equipo, especialmente de carácter multidisciplinar, y ser hábiles en la gestión del tiempo, personas y toma de decisiones.
C1	CT1 - Utilizar las herramientas básicas de las tecnologías de la información y las comunicaciones (TIC) necesarias para el ejercicio de su profesión y para el aprendizaje a lo largo de su vida.

Resultados de aprendizaje			
Resultados de aprendizaje	Competencias del título		
Conocer los conceptos generales de la regresión.		B2 B3 B7 B8 B9 B10	
Conocer técnicas de regresión no paramétricas.	A17 A18 A20	B2 B8 B9	C1
Entender la importancia de llevar a cabo una diagnosis de un modelo de regresión.	A17 A18 A20	B2 B3 B8 B9	C1
Saber estimar los parámetros de modelos de regresión lineal simple, múltiple y logístico	A17 A18 A20		C1
Ser capaz de interpretar los resultados y hacer predicciones utilizando modelos de regresión.	A17 A18	B2 B3 B7 B8 B9 B10	
Ser capaz de aplicar las principales técnicas de regresión a conjuntos de datos reales o simulados.	A17 A20	B2 B3 B8 B9	C1
Saber manejar con soltura programas informáticos avanzados de análisis estadístico.	A20	B2 B3 B7 B8 B9 B10	C1

Contenidos	
Tema	Subtema



1. Regresión Lineal Simple (RLS). Metodoloxía e Inferencia	<p>1.1 Conceptos xerais de regresión</p> <p>1.2 O modelo RLS. Estimación dos parámetros. Propiedades</p> <p>1.3 O coeficiente de correlación</p>
2. Regresión Lineal Simple. Diagnóstico e Predición.	<p>2.1 Diagnóstico do modelo RLS. Análise de residuos: linealidade, homocedasticidade, normalidade e independencia</p> <p>2.2 Observacións atípicas e influentes no modelo RLS</p> <p>2.3 Transformacións para conseguir linealidade</p> <p>2.4 Predición co modelo RLS</p>
3. Regresión Lineal Múltiple (RLM). Metodoloxía e Inferencia	<p>3.1 O modelo RLM. Hipóteses básicas do modelo</p> <p>3.2 Estimación dos parámetros. Propiedades dos estimadores</p> <p>3.3 Tabla ANOVA. Contraste de la F</p> <p>3.4 Correlación en regresión múltiple</p>
4. Regresión Lineal Múltiple. Diagnóstico e Predición	<p>4.1 O problema da multicolinealidade. Definición, efectos, identificación e tratamento</p> <p>4.2 Diagnóstico do modelo RLM. Análise de residuos: erro de especificación, homocedasticidade, normalidade e independencia</p> <p>4.3 Robustez do modelo. Observacións atípicas e influentes no modelo RLM</p> <p>4.4 Predición co modelo RLM</p> <p>4.5 Selección do modelo de regresión. Regresión paso a paso</p> <p>4.6 Regresión polinómica</p> <p>4.7 Regresión con variables cualitativas. Variables ficticias</p>
5. Modelo de Regresión con resposta cualitativa	<p>5.1 Formulación de modelos con resposta cualitativa</p> <p>5.2 O modelo logístico</p> <p>5.3 Estimación e contrastes no modelo logístico</p> <p>5.4 Diagnóstico do modelo</p>
6. Regresión non paramétrica	<p>6.1 Introducción aos métodos de suavización en regresión</p> <p>6.2 Estimación núcleo, o estimador de Nadaraya-Watson</p> <p>6.3 O estimador polinómico local</p> <p>6.4 Selección do parámetro de suavizado</p> <p>6.5 Estimadores de tipo spline</p> <p>6.6 Extensiones. O modelo semiparamétrico (modelo parcialmente lineal). O modelo aditivo</p>

Planificación				
Metodoloxías / probas	Competencias	Horas presenciales	Horas no presenciales / traballo autónomo	Horas totales
Sesión magistral	A17 A18 B3 B8 B9 B10	30	30	60
Prácticas a través de TIC	A17 A18 A20 B2 B3 B8 B9 C1	20	20	40
Seminario	A20 A18 B2 B3 B8 C1	10	10	20
Solución de problemas	A17 A18 A20 B2 B7 B9 C1	0	20	20
Proba obxectiva	A17 A18 A20 B2 B9 C1	6	0	6
Atención personalizada		4	0	4
(*) Los datos que aparecen en la tabla de planificación són de carácter orientativo, considerando la heterogeneidad de los alumnos				

Metodoloxías
--------------



Metodologías	Descrición
Sesión magistral	Se introducirán los fundamentos teóricos de la materia y las principales técnicas para su aplicación en la práctica. Se pondrá particular atención a las condiciones estructurales requeridas para su aplicación y a la diagnosis y evaluación de los modelos ajustados. Su utilidad se ilustrará con ejemplos específicos de diferentes áreas de conocimiento, enfatizando sus ventajas y limitaciones. Se promoverá la participación de los estudiantes.
Prácticas a través de TIC	Se presentarán problemas específicos y los procedimientos de resolución poniendo en práctica conceptos y algoritmos expuestos en las sesiones magistrales. Se tratará de sesiones interactivas donde los problemas propuestos se resolverán con ayuda de scripts con código libre del software R. El docente apoyará y supervisará la puesta en práctica de los conocimientos adquiridos por los estudiantes.
Seminario	Se propondrán cuestiones y ejercicios para discusión y realización conjuntamente con los estudiantes.
Solución de problemas	Se propondrán ejercicios para la realización por parte de los estudiantes de forma individual.
Prueba objetiva	Pruebas objetivas de evaluación de conocimientos.

### Atención personalizada

Metodologías	Descrición
Seminario Prácticas a través de TIC	Las prácticas a través de TIC están pensadas para resolver ejercicios usando código en R. Los estudiantes deben familiarizarse con el código y las librerías de R que se usen, comprender y saber interpretar las salidas que se generen y ser capaces de emplearlo individualmente en la realización de ejercicios similares. Para alcanzar estos objetivos lo más deprisa posible es deseable proporcionar atención personalizada, lo que se realizará durante el desarrollo de las sesiones presenciales y en tutorías individualizadas.

### Evaluación

Metodologías	Competencias	Descrición	Calificación
Prueba objetiva	A17 A18 A20 B2 B9 C1	La prueba de conocimientos final será un examen escrito que constará de varias cuestiones teórico-prácticas sobre los contenidos de la materia. Será obligatorio y supondrá el 80% de la calificación.	80
Solución de problemas	A17 A18 A20 B2 B7 B9 C1	La resolución y entrega de ejercicios propuestos a lo largo del curso formará parte de la evaluación continua. Computará hasta el 20% de la calificación.	20

### Observaciones evaluación

Presentación y evaluación: Se considerará que un estudiante concurre a una convocatoria cuando participa en actividades que le permiten obtener cuando menos un 50% de la evaluación final. La calificación final obtenida se conservará entre las oportunidades (ordinaria y extraordinaria) dentro de la convocatoria de cada curso. Segunda oportunidad (prueba de julio) El peso de la evaluación continua en la segunda oportunidad (prueba de julio) será el mismo que en la evaluación ordinaria, un 20%. En la segunda oportunidad de evaluación se realizará un examen y la nota final será el máximo de tres cantidades: la nota de la evaluación de la primera oportunidad, la nota del nuevo examen y la media ponderada del nuevo examen (80%) y la evaluación continua (20%).
---

### Fuentes de información



<b>Básica</b>	<ul style="list-style-type: none"> <li>- Vilar Fernández, J.M. (2006). Modelos estadísticos aplicados. Universidade da Coruña, Servizo de Publicacións</li> <li>- Faraway, J.J. (2015). Linear models with R . Chapman and Hall</li> <li>- Faraway, J.J. (2006). Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models. Chapman and Hall</li> <li>- Peña, D. (2002). Regresión y diseño de experimentos. Alianza Editorial</li> <li>- Montgomery, D.C., Peck, E.A. y Vining, G.G. (2012). Introduction to linear regression analysis . Wiley</li> <li>- Ritz, C. y Streibig, J.C. (2008). Nonlinear regression with R. Springer</li> <li>- Wand M.P. y Jones M.C. (1995). Kernel Smoothing. Chapman and Hall/CRC</li> <li>- Hosmer, D.W., Lemeshow, S. y Sturdivant, R.X. (2013). Applied logistic regression . John Wiley &amp; Sons</li> </ul>
<b>Complementaría</b>	<ul style="list-style-type: none"> <li>- Bowman A.W. y Azzalini A. (1997). Applied Smoothing Techniques for Data Analysis. Oxford University Press</li> <li>- Fox, J. y Weisberg, S. (2011). An R companion to applied regression. SAGE Publications</li> <li>- Agresti, A. (1996). An introduction to categorical data analysis. Wiley</li> <li>- Fan J. y Gijbels I. (1996). Local polynomial modelling and its applications. Chapman and Hall/CRC</li> <li>- Sheather, S.J. (2009). A modern approach to regression with R. Springer</li> <li>- Venables, W.N. y Ripley, B.D. (2010). Modern applied statistics with S . Springer</li> <li>- Wood, S.N. (2006). Generalized Additive Models: An introduction with R. Chapman and Hall/CRC</li> </ul>

## Recomendaciones

### Asignaturas que se recomienda haber cursado previamente

Álgebra Lineal/614G02001

Cálculo Multivariable/614G02006

Inferencia Estadística/614G02007

Probabilidad y Estadística Básica/614G02003

### Asignaturas que se recomienda cursar simultáneamente

Modelización Estadística de Datos de Alta Dimensión/614G02013

### Asignaturas que continúan el temario

Técnicas de Simulación y Remuestreo/614G02036

Análisis Estadístico de Datos Complejos/614G02031

Optimización Matemática/614G02020

Análisis Estadístico de Datos con Dependencia/614G02022

### Otros comentarios

(\*) La Guía Docente es el documento donde se visualiza la propuesta académica de la UDC. Este documento es público y no se puede modificar, salvo cosas excepcionales bajo la revisión del órgano competente de acuerdo a la normativa vigente que establece el proceso de elaboración de guías