

## Máster en Técnicas Estadísticas

### **DATOS IDENTIFICATIVOS DE LA MATERIA**

**Código de la materia:** P1062111

**Nombre de la materia:** Regresión generalizada y modelos mixtos

**Número de créditos ECTS:** 5

**Curso académico:** 2022/2023

**Profesorado:**

Rosa M. Crujeiras Casais (3 créditos)

Mercedes Conde Amboage (coordinadora, 2 créditos)

### **OBJETIVOS DE LA MATERIA**

En esta materia se pretende familiarizar al alumnado con los modelos de regresión, tratando de obtener los siguientes resultados del aprendizaje:

- Conocer los modelos de regresión lineal generalizados.
- Saber usar de forma autónoma las técnicas de regresión avanzada (regresión generalizada y modelos mixtos) para la toma de decisiones en contextos multidisciplinares.
- Saber formular y aplicar el modelo adecuado para estudiar la dependencia entre una variable respuesta y un conjunto de variables explicativas.
- Conocer diferentes extensiones de la regresión lineal, identificando los factores diferenciales de cada una.

### **CONTENIDOS DE LA MATERIA**

#### **Tema 1. Introducción a los modelos lineales generalizados.**

Introducción a los modelos lineales generalizados. El modelo de regresión logística. El modelo de Poisson para datos de conteo. Estimación e inferencia sobre los parámetros. Contraste de modelos mediante la deviance. Sobre-dispersión en el modelo de Poisson. Formulación y análisis de modelos lineales generalizados.

### **Tema 2. Regresión no lineal.**

Ejemplos notables de modelos no lineales de regresión. Estimación de los parámetros por mínimos cuadrados. Algoritmos de estimación. Inferencia sobre los parámetros en base a la distribución asintótica y mediante el perfil de RSS. El test F.

### **Tema 3. Regresión cuantil.**

Introducción: la mediana, los cuantiles, la desviación absoluta y la función de Koenker. La función de regresión cuantil. El modelo de regresión cuantil lineal. Algoritmos de estimación. Propiedades del estimador cuantil. Inferencia sobre los parámetros. Regresión cuantil no lineal. Regresión cuantil no paramétrica.

### **Tema 4. Análisis de la varianza con efectos aleatorios.**

Revisión de los modelos ANOVA y ANCOVA. Introducción a los datos multinivel. Modelo de análisis de la varianza con efectos aleatorios. Estimación del modelo. Predicción dos efectos aleatorios.

### **Tema 5. Modelos multinivel con respuesta continua.**

Modelos con variable explicativa asociada al nivel inferior: efectos aleatorios en la tendencia e en la ordenada en el origen. Modelos con variable explicativa asociada al grupo. Correlaciones intra-grupos, entre-grupos y efecto contextual. Estimación por máxima verosimilitud restringida. Predicción dos efectos aleatorios.

### **Tema 6. Modelos multinivel con respuesta binaria.**

Formulación o do modelo de regresión logística con efectos aleatorios. Métodos de estimación. Interpretación de los elementos del modelo en el caso logístico.

## ***BIBLIOGRAFÍA BÁSICA Y COMPLEMENTARIA***

### **Básica**

- [1] Faraway, J.J. (2006). Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models. Chapman and Hall.
- [2] Koenker, R. (2005). Quantile regression. Cambridge University Press.
- [3] McCulloch, C.E. e Neuhaus, J.M. (2005). Generalized linear mixed models. Encyclopedia of Biostatistics, 4.
- [4] Ritz, C. y Streibig, J.C. (2008). Nonlinear regression with R. Springer.
- [5] Sheather, S.J. (2009). A modern approach to regression with R. Springer.
- [6] West, B. T., Welch, K. B. e Galecki, A. T. (2014). Linear mixed models: a practical guide using statistical software. Chapman and Hall/CRC.

### **Complementaria**

- [1] Furno, M. y Vistocco, D. (2018). Quantile regression: estimation and simulation (Vol. 216). John Wiley & Sons. Disponible en: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118863718>.
- [2] Galecki, A. y Burzykowski, T. (2013). Linear mixed-effects models using R: A step-

bystep approach. Springer Science & Business Media.

[3] Huet, S., Bouvier, A., Gruet, M.A. y Jolivet, E. (1996). Statistical tools for nonlinear regression (A practical guide with S-Plus examples). Springer.

### **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.

### **METODOLOGÍA DOCENTE**

La enseñanza constará de clases expositivas e interactivas, así como de la tutorización del aprendizaje y de las tareas encomendadas al alumnado. Se proporcionarán los apuntes de la materia, 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 alumnado disponga de un ordenador. Se propondrán actividades para el alumnado, que consistirán en la resolución de cuestiones, ejercicios y ejemplos relacionados con los modelos de regresión generalizados y los modelos mixtos.

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

- TEMA 1. INTRODUCCIÓN GLM (4h expositivas, 3h de laboratorio)
- TEMA 2. NO LINEAL (2h expositivas, 2h de laboratorio)
- TEMA 3. CUANTIL (2h expositivas, 3h de laboratorio)
- TEMA 4. ANOVA ALEATORIO (2h expositivas, 4h de laboratorio)
- TEMA 5. MULTINIVEL CONTINUA (3h expositivas, 5h de laboratorio)
- TEMA 6. MULTINIVEL BINARIA (2h expositivas, 3h de laboratorio)

En el caso de que alguna de las actividades docentes deba realizarse en remoto, se utilizará la plataforma MS Teams, institucional para la USC y la UDC, con acceso como invitados/as para el alumnado matriculado en la UVigo.

### ***CRITERIOS Y MÉTODOS DE EVALUACIÓN***

**Evaluación continua (40%):** la evaluación continua se realizará en base a la resolución de problemas, presenciales o no presenciales, por parte del alumnado. En estos problemas, el alumnado utilizará el programa R y redactará las conclusiones extraídas. La calificación obtenida 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 y E6.

**Prueba final (60%):** 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 lenguaje estadístico utilizado en la docencia interactiva (R). 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 alumno/a concurre a una convocatoria cuando participa en actividades que le permiten obtener al menos un 50% de la evaluación final. 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 e 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 alumnado 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***

Es conveniente que el alumnado posea conocimientos básicos de cálculo de probabilidades y estadística. También es recomendable disponer de unas habilidades medias en el manejo de ordenadores, y en concreto de software estadístico. Para un mejor aprendizaje de la materia, conviene tener presente el sentido práctico de los métodos que se están conociendo.

En el caso de que la docencia se desarrolle de manera presencial, el profesorado informará en clase del plan semanal, indicando 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 alumnado al comienzo de cada semana, y donde se especificarán los contenidos a trabajar y las actividades recomendadas.

### ***RECURSOS PARA EL APRENDIZAJE***

Bibliografía y apuntes. Uso del campus virtual de la USC y del sitio web del Máster en Técnicas Estadísticas como soporte para el material del programa.

### ***OBSERVACIONES***

El desarrollo de los contenidos de la materia se realizará teniendo en cuenta que las competencias a adquirir por el alumnado deben cumplir con el nivel MECES3. Los contenidos que se incluyen en esta materia son técnicamente complejos (como los considerados en los modelos lineales generalizados) y/o altamente especializados y novedosos (por ejemplo, los relativos a los modelos multinivel), y su estudio se acompañará de implementaciones prácticas, empleando software específico.