Máster en Técnicas Estadísticas

DATOS IDENTIFICATIVOS DE LA MATERIA

Nombre de la materia: Modelos de regresión

Número de créditos ECTS: 5

Curso académico: 2019/2020

Profesorado:

Paula Saavedra Nieves (coordinadora, 3.5 créditos)

Luis Alberto Ramil Novo (1.5 créditos)

OBJETIVOS DE LA MATERIA

En esta materia se pretende familiarizar al alumnado con los modelos de regresión. Los objetivos a alcanzar como resultado del aprendizaje son:

• Conocer en profundidad los aspectos teóricos del análisis de regresión lineal y en concreto, del modelo lineal general.

• Saber aplicar los métodos de regresión lineal en el análisis de datos reales de naturaleza compleja.

• Saber comunicar los resultados obtenidos con las técnicas de regresión lineal en entornos multidisciplinares.

• Conocer las potencialidades y limitaciones del análisis de regresión lineal.

CONTENIDOS DE LA MATERIA

Tema 1. **Modelo de regresión lineal simple.**

Elementos de un modelo de regresión: el modelo lineal. Estimación de los parámetros por mínimos cuadrados. Propiedades de los estimadores. Inferencia sobre los parámetros. Descomposición de la variabilidad. El test F. Predicción.

Tema 2. **Validación de un modelo de regresión.**
El coeficiente de determinación. Diagnosis del modelo. Transformaciones previas a la regresión.

Tema 3. **El modelo lineal general: regresión múltiple.**
El modelo de regresión lineal múltiple y el modelo lineal general. Estimación de los parámetros. Interpretación de los parámetros: regresión particionada y regresión parcial. Coeficientes de correlación simple, múltiple y parcial. Propiedades de los estimadores. Inferencia sobre los parámetros. Descomposición de la variabilidad. El test F. Predicción.

Tema 4. **Diagnosis de observaciones atípicas o influyentes.**
Introducción a las observaciones atípicas e influyentes. Los apalancamientos en regresión simple y en regresión múltiple. Detección del carácter atípico: estandarización de los residuos. Diagnosis de la normalidad. Detección del carácter influyente: medidas de influencia. Pautas de actuación ante datos atípicos o influyentes.

Tema 5. **Construcción de un modelo de regresión.**
Regresión polinómica. Interacciones. Modelos linealizables. Validación de un modelo de regresión múltiple. Colinealidad. Métodos de selección de variables.

Tema 6. **Análisis de la varianza.**
El modelo de análisis de la varianza. Parametrización de una variable explicativa discreta. Descomposición de la variabilidad. El test F. Comparaciones múltiples. Contraste de igualdad de varianzas.

Tema 7. **Análisis de la covarianza.**
Modelo con una variable explicativa discreta y otra continua, sin interacción y con interacción. Contraste de los efectos principales y contraste de la interacción. Modelos de regresión con varias variables explicativas discretas y continuas.

Tema 8. **Regresión logística.**
El modelo de regresión logística: la odds y la odds-ratio. Estimación de los parámetros por máxima verosimilitud. Algoritmos de estimación. Inferencia sobre los parámetros. Contraste de modelos mediante la deviance.

BIBLIOGRAFÍA BÁSICA Y COMPLEMENTARIA

Básica

Faraway, J.J. (2015). Linear models with R (2nd edition). Chapman and Hall.
Faraway, J.J. (2006). Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models. Chapman and Hall.
Montgomery, D. C.; Peck, E. A. y Vining, G. G. (2012). Introduction to linear regression analysis (5th ed). Wiley
Ritz, C. y Streibig, J.C. (2008). Nonlinear regression with R. Springer.
Sheather, S.J. (2009). A modern approach to regression with R. Springer.

Complementaria

Agresti, A. (1996). An introduction to categorical data analysis. Wiley.
Fox, J. y Weisberg, S. (2011). An R companion to applied regression. SAGE Publications.
Greene, W.H. (1999). Análisis econométrico. Prentice Hall.
Hosmer, D. W., Lemeshow, S. y Sturdivant, R. X. (2013). Applied logistic regression (3rd edition). John Wiley & Sons.
Huet, S., Bouvier, A., Gruet, M.A. y Jolivet, E. (1996). Statistical tools for nonlinear regression (A practical guide with S-Plus examples). Springer.
Peña, D. (2010). Regresión y diseño de experimentos. Alianza Editorial.
Venables, W.N. y Ripley, B.D. (2010). Modern applied statistics with S (4th edition). Springer.

COMPETENCIAS

Se indican a continuación cuáles son las competencias específicas (E), 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.
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.

En esta materia se trabajarán las competencias básicas (CB6-CB10), generales (CG1-CG5) y transversales (CT1-CT5) recogidas en la memoria del título, incidiendo en las que se indican a continuación:

CB7 - Saber aplicar los conocimientos avanzados adquiridos, integrándolos en la resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.
CB9 - 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.

CG3 - Desarrollar la capacidad para realizar estudios y tareas de investigación y transmitir los resultados a públicos especializados, académicos y generalistas.
CG4 - Integrar conocimientos avanzados y enfrentarse a la toma de decisiones a partir de información científica y técnica.
CG5 - Desarrollar la capacidad de aplicación de algoritmos y técnicas de resolución de problemas complejos en el ámbito de la estadística y la investigación operativa, manejando el software especializado adecuado.

CT3 - Ser capaz de resolver problemas complejos en entornos nuevos mediante la aplicación integrada de los conocimientos.

METODOLOGÍA DOCENTE

La enseñanza constará de clases expositivas e interactivas, así como de la tutorización del aprendizaje y de las tareas encomendadas al alumnado.

En las clases expositivas e interactivas se resolverán ejemplos mediante el software R, por lo que es necesario que el alumnado disponga en el aula 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.

Se proporcionarán los apuntes de la materia, así como otro material orientativo del aprendizaje del software. Los apuntes y otros instrumentos didácticos estarán disponibles a través de alguna herramienta de acceso por vía web.

CRITERIOS Y MÉTODOS DE EVALUACIÓN

Evaluación continua (30%): 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 CB7 e CB9 y generales CG3, CG4 y CG5. También se evaluará el nivel alcanzado en las competencias específicas E2, E5 y E8. Así mismo, se tendrá en cuenta en la evaluación el nivel alcanzado en la competencia transversal CT3.

Examen final (70%): el examen 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 evaluarán las competencias específicas: E1, E2, E4, E5, E6 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.

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 actividades presenciales: 20 horas de sesiones expositivas, 15 horas de sesiones interactivas (seminarios, laboratorios en aula de informática y presentación de trabajos) y 3 horas para realización de exámenes.

Se estima que el alumno debe dedicar 87 horas a las actividades de tipo no presencial, entre las que se incluyen la resolución de ejercicios, la resolución de casos prácticos; actividades de análisis de datos y modelos; la elaboración de trabajos y tiempo de estudio personal.

En total, resultan 25 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.

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 de la asignatura.

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. En este sentido, si bien los contenidos de la materia se centran únicamente en modelos de regresión lineal (con respuesta continua y binaria – regresión logística - y con variables explicativas continuas pero también categóricas), estos se estudiarán de una manera exhaustiva, presentando todas las fases del proceso de modelado de manera rigurosa: formulación del modelo, estimación, validación y diagnosis. Se discutirán, además, los errores que se pueden cometer al tomar decisiones basadas en modelos con problemas de especificación (modelos que no cumplen las hipótesis bajo las que se formula la inferencia, o modelos que directamente, no se ajustan bien a las observaciones).