

PROGRAMA OFICIAL DE POSTGRADO EN ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

DATOS IDENTIFICATIVOS DE LA MATERIA

Código de la materia: P1061101

Nombre de la materia: Modelos de Regresión

Número de créditos ECTS: 6

Curso académico: 2018/2019

Profesorado: Luis Alberto Ramil Novo (3 ECTS)
Paula Saavedra Nieves (1.5 ECTS)
César Andrés Sánchez Sellero (coordinador, 1.5 ECTS)

OBJETIVOS DE LA MATERIA

En esta materia se pretende familiarizar al alumnado con los modelos de regresión, con los siguientes objetivos:

- Conocer los modelos que describen la influencia de unas variables (variables explicativas) sobre otra variable (variable respuesta).
- Saber realizar las tareas de selección del modelo, y de su aplicación en objetivos de inferencia y predicción.

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. Diagnóstico del modelo. Transformaciones previas a la regresión.

Tema 3. Operaciones lineales y cuadráticas sobre vectores aleatorios.

Vectores aleatorios: vector de medias, matriz de covarianzas, transformaciones lineales y estandarización. La distribución normal multivariante. Formas cuadráticas sobre una muestra de variables normales.

Tema 4. 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 5. Diagnóstico 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. Diagnóstico de la normalidad. Detección del carácter influyente: medidas de influencia. Pautas de actuación ante datos atípicos o influyentes.

Tema 6. 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 7. 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 8. 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 9. 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 verosimilitud. Contraste de modelos: el test F.

Tema 10. 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.

Tema 11. **Regresión de Poisson y modelos lineales generalizados.**

El modelo de Poisson para datos de recuento. 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.

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 GENERALES Y ESPECÍFICAS

En esta materia se trabajarán las competencias básicas indicadas en la memoria del Máster en Técnicas Estadísticas con los códigos CB6, CB7, CB8, CB9 y CB10, así como las competencias transversales T4 y T9. Se indican a continuación cuáles son las competencias generales y específicas que se potenciarán en Modelos de Regresión.

Competencias básicas y generales:

- G3 - Capacidad de trabajo en equipo y de forma autónoma.
- G6 - Capacidad de identificar y resolver problemas.
- G13 - Redacción de informes estadísticos con precisión, orden y claridad.

Competencias específicas:

- E4 - Conocer las aplicaciones de los modelos de la estadística y la investigación operativa.
- E11 - Modelar la dependencia entre una variable respuesta y varias variables explicativas.

E12 - Realizar inferencias respecto a los parámetros que aparecen en el modelo.

E84 - Ser capaz de manejar diverso software (en particular R) e interpretar los resultados que proporcionan éstos en los correspondientes estudios prácticos.

En los apartados donde se describa la metodología y el sistema de evaluación se especificará la manera en la que las competencias (básicas, generales, específicas y transversales) a adquirir en la materia se trabajan y evalúan.

METODOLOGÍA DOCENTE: ACTIVIDADES DE APRENDIZAJE Y SU VALORACIÓN EN CRÉDITOS ECTS

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 en el aula de un ordenador. En estas sesiones se trabajará sobre las competencias específicas E4, E11, E12 y E84, y la transversal T4 (conocimiento de aplicaciones, modelado, inferencia y manejo de software), así como sobre la competencia general G6 (resolución de problemas).

Se propondrán actividades para el alumnado, que consistirán en la resolución de cuestiones, ejercicios y ejemplos relacionados con los Modelos de Regresión. Con estas actividades, se pretende que el alumnado trabaje las competencias generales G3 y G13 y la transversal T9 (trabajo en equipo y autónomo y redacción de informes, y comunicación), así como las específicas E8 y E11 (manejo de software y modelado). Con estas actividades también se potenciarán las competencias básicas CB6, CB7 y CB8 (aplicación de conocimientos e ideas en un contexto de investigación y/o interdisciplinar e integración de conocimientos) y CB9 y CB10 (comunicación y aprendizaje autónomo).

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 generales G4, G6 y G13, así como de las competencias específicas E4 y E84 y las transversales T4 y T9. Las competencias básicas se evaluarán parcialmente mediante la evaluación continua y el examen final.

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). Además de las competencias específicas E4 y E84, que se evalúan parcialmente a

través de la evaluación continua, también se evaluarán la competencia general G6 y la específica E11.

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 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.

RECURSOS PARA EL APRENDIZAJE

Bibliografía y apuntes. Uso del campus virtual de la USC, repositorio de material del programa.

OBSERVACIONES

Como materiales, se emplearán apuntes elaborados por el profesorado, bibliografía y ordenador. Los apuntes y otros instrumentos didácticos estarán disponibles a través de alguna herramienta de acceso por vía web.