

Modelos Lineales 2

Código: 104861
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
2503852 Estadística Aplicada	OB	3	1

La metodología docente y la evaluación propuestas en la guía pueden experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.

Contacto

Nombre: Natalia Isabel Vilor Tejedor
Correo electrónico: NataliaIsabel.Vilor@uab.cat

Uso de idiomas

Lengua vehicular mayoritaria: catalán (cat)
Algún grupo íntegramente en inglés: No
Algún grupo íntegramente en catalán: Sí
Algún grupo íntegramente en español: No

Prerequisitos

Conocimientos básicos de estadística descriptiva e inferencial. Es imprescindible un curso previo de Modelos Lineales.

Objetivos y contextualización

El objetivo del curso es el aprendizaje supervisado con un enfoque en los métodos de regresión y clasificación. El plan de estudios incluye la revisión de conceptos de regresión lineal y polinómica, regresión logística y análisis discriminante lineal; selección de modelos y métodos de regularización (Ridge y LASSO); modelos no lineales tales como: splines y modelos aditivos generalizados, entre otros.

Competencias

- Analizar datos mediante la aplicación de métodos y técnicas estadísticas, trabajando con datos de diversas tipologías.
- Aplicar el espíritu crítico y el rigor para validar o refutar argumentos tanto propios como de otras personas.
- Diseñar un estudio estadístico o de investigación operativa para la resolución de un problema real.
- Evaluar de manera crítica y con criterios de calidad el trabajo realizado.
- Formular hipótesis estadísticas y desarrollar estrategias para confirmarlas o refutarlas.
- Interpretar resultados, extraer conclusiones y elaborar informes técnicos en el campo de la estadística.
- Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
- 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.
- 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.
- Resumir y descubrir patrones de comportamiento en la exploración de los datos.
- Seleccionar los modelos o técnicas estadísticas para aplicarlos a estudios y problemas reales, así como conocer las herramientas de validación de los mismos.

- Seleccionar y aplicar procedimientos más apropiados para la modelización estadística y el análisis de datos complejos.
- Utilizar correctamente un amplio espectro del software y lenguajes de programación estadísticos, escogiendo el más apropiado para cada análisis y ser capaz de adaptarlo a nuevas necesidades.
- Utilizar eficazmente la bibliografía y los recursos electrónicos para obtener información.

Resultados de aprendizaje

1. Analizar datos mediante el modelo de regresión lineal.
2. Analizar datos mediante el modelo lineal generalizado.
3. Analizar datos mediante técnicas de inferencia usando software estadístico.
4. Analizar los residuos de un modelo estadístico.
5. Aplicar el espíritu crítico y el rigor para validar o refutar argumentos tanto propios como de otros.
6. Comparar el grado de ajuste entre varios modelos estadísticos.
7. Detectar y contemplar interacciones entre variables explicativas.
8. Detectar y tratar la colinealidad entre variables explicativas.
9. Emplear gráficos de visualización del ajuste y de la adecuación del modelo.
10. Establecer las hipótesis experimentales de la modelización.
11. Evaluar de manera crítica y con criterios de calidad el trabajo realizado.
12. Extraer conclusiones de la adecuación de los modelos con la utilización e interpretación correcta de indicadores y gráficos.
13. Identificar distribuciones de las respuestas con el análisis de residuos.
14. Identificar fuentes de sesgo en la obtención de la información.
15. Identificar las etapas en los problemas de modelización.
16. Identificar las suposiciones estadísticas asociadas a cada procedimiento.
17. Identificar las variables respuesta, explicativas y de control.
18. Medir el grado de ajuste de un modelo estadístico.
19. Modificar ligeramente el software existente si el modelo estadístico propuesto lo requiere.
20. Predecir respuestas, comparar grupos (valor causal) e identificar factores relevantes.
21. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
22. 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.
23. 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.
24. Seleccionar las variables explicativas relevantes.
25. Sintetizar e interpretar los resultados de los modelos lineales clásicos, generalizados y no lineales en función del objetivo del estudio.
26. Usar regresión logística para resolver problemas de clasificación.
27. Utilizar diversidad de software estadístico para ajustar y validar modelos lineales y sus generalizaciones.
28. Utilizar eficazmente bibliografía y recursos electrónicos para obtener información.
29. Validar los modelos utilizados mediante técnicas de inferencia adecuadas.

Contenido

1. Regresión lineal

Regresión lineal simple

Regresión lineal múltiple

Ampliación de los modelos lineales.

2. Métodos de Clasificación

Introducción de los métodos de clasificación.

Regresión logística. El modelo logístico. Estimación de los coeficientes de regresión. Predicciones.

Regresión logística múltiple

Análisis discriminante lineal.
Análisis discriminante cuadrático.

3. Selección de modelos y Regularización.

Selección de subconjuntos: Selección por pasos, y selección de modelo óptima.

Métodos de Regularización: Ridge y regresión LASSO. Selección del parámetro de afinación de los métodos de regularización.

Métodos de reducción de dimensionalidad: Análisis de componentes principales (PCA) y mínimos cuadrados parciales (PLS).

4. Más allá de la linealidad

Regresión polinomial

Regresión Step-wise

Splines

Modelos aditivos generalizados (GAM)

****A menos que las restricciones impuestas por las autoridades sanitarias obliguen a una priorización o reducción de estos contenidos.***

Metodología

La asignatura tiene, a lo largo del cuatrimestre, dos horas semanales de teoría y problemas más dos horas semanales de prácticas. El material del curso (apuntes de teoría, listas de problemas y enunciados de prácticas) estará disponible en el campus virtual, de manera progresiva a lo largo del curso.

****La metodología docente propuesta puede experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.***

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Prácticas en ordenador	50	2	1, 2, 3, 5, 11, 10, 9, 15, 16, 19, 23, 25, 28, 26
Teoría	50	2	1, 2, 3, 4, 5, 11, 6, 7, 8, 10, 12, 9, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 28, 26, 27, 29
Tipo: Supervisadas			
problemas/ejercicios para resolver	16	0,64	3, 5, 11, 12, 23, 25, 28, 27
Tipo: Autónomas			
Preparación Exámenes	10	0,4	3, 5, 11, 16, 21, 25

Evaluación

PR: Entrega de las prácticas que se indiquen lo largo del curso. Valoración de PR: 4 puntos sobre 10. Esta parte NO es recuperable.

P1: Prueba parcial 1 (teoría, problemas y prácticas). Valoración de P1: 2 puntos sobre 10.

P2: Prueba parcial 2 (teoría, problemas y prácticas). Valoración de P2: 4 puntos sobre 10.

Para que un examen se tenga en cuenta a la hora de calcular la nota final es necesario obtener una nota mínima de 3.5 puntos. La Nota final será: $\text{Nota Final} = \text{PR} + \text{P1} + \text{P2}$. En enero habrá día asignado para hacer una prueba final, PF, que permite la recuperación de P1 y P2 (6 puntos sobre 10) para los alumnos que no hayan aprobado el curso o se quiera mejorar la nota. En caso de que se haga el examen final, la nota final será: $\text{Nota Final} = \text{PR} + \text{PF}$.

****La evaluación propuesta puede experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.***

Actividades de evaluación

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Parcial 1	20%	4	0,16	1, 2, 3, 4, 5, 11, 6, 7, 8, 10, 12, 9, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 28, 26, 27, 29
Parcial 2	40%	4	0,16	1, 2, 3, 4, 5, 11, 6, 7, 8, 10, 12, 9, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 28, 26, 27, 29
Prácticas	40%	16	0,64	3, 5, 11, 10, 9, 15, 16, 19, 21, 22, 23, 25, 28, 27

Bibliografía

Montgomery, D. Peck, A. Vining, G.; *Introduction to Linear Regression Analysis*. Wiley. 2001.

Christopher Hay-Jahans; *An R Companion to Linear Statistical Models*. Chapman and Hall, 2012.

John Fox and Sandord Weisberg; *An R Companion to Applied Regression*, 2nd edition, Sage Publications, 2011.

Daniel Peña; *Regresión y diseño de Experimentos*, Alianza Editorial (Manuales de Ciencias Sociales), 2002.

Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani; *An Introduction to Statistical Learning*, Springer texts in Statistics, 2013.

Software

Lenguaje de programación R