

| Titulación | Tipo | Curso |
|---------------------|------|-------|
| Ingeniería de Datos | OB | 3 |

Contacto

Nombre: Sandra Cobo Ollero

Correo electrónico: sandra.cobo@uab.cat

Idiomas de los grupos

Puede consultar esta información al [final](#) del documento.

Prerrequisitos

Se recomienda haber cursado la asignatura de descripciones probabilísticas y estadísticas. Nociones de álgebra y cálculo.

Objetivos y contextualización

El curso introduce al estudiante en el aprendizaje y el análisis estadístico utilizando las siguientes técnicas:

1. Temas básicos de inferencia estadística y análisis de varianza de una vía.
2. Elementos básicos que intervienen en el análisis estadístico de modelos de regresión simple y multilíneal.
3. Herramientas implicadas en el análisis estadístico (estimación de parámetros, descomposición de la varianza, tabla ANOVA, contrastes) de diseños de bloques aleatorios completos y no completos.

El curso también pretende familiarizar a los estudiantes con el uso del software R.

Competencias

- Analizar los datos de forma eficiente para el desarrollo de sistemas inteligentes con capacidad de aprendizaje autónomo y/o para la minería de datos.
- Buscar, seleccionar y gestionar de manera responsable la información y el conocimiento.
- Desarrollar un pensamiento y un razonamiento crítico y saber comunicarlo de manera efectiva, tanto en las lenguas propias como en inglés.
- 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.
- Utilizar las técnicas propias de la probabilidad y estadística para analizar y modelar fenómenos complejos, y para resolver problemas de optimización.

Resultados de aprendizaje

1. Buscar, seleccionar y gestionar de manera responsable la información y el conocimiento.
2. Desarrollar un pensamiento y un razonamiento crítico y saber comunicarlo de manera efectiva, tanto en las lenguas propias como en inglés.
3. Diseñar e implementar una estrategia integrada de técnicas estadísticas y de inteligencia artificial para el desarrollo de sistemas descriptivos y predictivos.
4. Interpretar correctamente el resultado de un test o modelo estadístico para el análisis poblacional de datos experimentales o la validación de un algoritmo.
5. 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.

Contenido

1. Elementos de inferencia estadística (Repaso)

- 1.1 Población y muestra, parámetros y estadísticos
- 1.2 Distribuciones de probabilidad e inferencia
- 1.3 Estimación puntual e intervalos de confianza
- 1.4 Conceptos básicos de contraste de hipótesis
 - 1.4.1 Planteamiento de una hipótesis estadística
 - 1.4.2 Contraste para la media
 - 1.4.3 Contraste para la varianza
 - 1.4.4 Tres criterios equivalentes de rechazo o aceptación
 - 1.4.5 Hipótesis para dos medias: comparación de dos tratamientos
 - 1.4.6 Contraste para la igualdad de varianzas
 - 1.4.7 Poblaciones apareadas (comparación de dos medias con muestras dependientes)

2. Principios del diseño de experimentos

- 2.1 El diseño de experimentos hoy
- 2.2 Definiciones básicas en el diseño de experimentos
- 2.3 Etapas en el diseño de experimentos
- 2.4 Consideraciones prácticas sobre el uso de métodos estadísticos
- 2.5 Clasificación y selección de los diseños experimentales

3. Análisis de Varianza (ANOVA)

- 3.1 Diseño completamente al azar y ANOVA
- 3.2 Comparaciones múltiples o pruebas de rango
- 3.3 Verificación de los supuestos del modelo
- 3.4 Elección del tamaño muestral
- 3.5 Métodos no paramétricos en el análisis de varianza
 - 3.5.1 Prueba de Kruskal-Wallis

4. Regresión lineal simple y múltiple

- 4.1 Regresión lineal simple
 - 4.1.1 Contrastes de hipótesis en regresión lineal simple
 - 4.1.2 Calidad del ajuste en regresión lineal simple
 - 4.1.3 Estimación y predicción por intervalo en regresión simple (Método de mínimos cuadrados. Estimadores de máxima verosimilitud)
- 4.2 Regresión lineal múltiple
 - 4.2.1 Contrastes de hipótesis en regresión lineal múltiple
 - 4.2.2 Intervalos de confianza y predicción en regresión múltiple
 - 4.2.3 Expresión matricial del modelo y estimadores de los coeficientes. Interpretación de los coeficientes en el modelo múltiple
 - 4.2.4 Prueba de "ligaduras" para resolver restricciones lineales sobre los coeficientes

- 5. Modelos lineales generalizados y mixtos
 - 5.1 Regresión logística: variables respuesta dicotómicas
 - 5.2 Estimación de parámetros e interpretación de coeficientes
 - 5.3 Evaluación del ajuste del modelo
 - 5.4 Aplicaciones y ejemplos en clasificación binaria
- 6. Análisis de la supervivencia
 - 6.1 Conceptos básicos: tiempo hasta el evento, censura
 - 6.2 Funciones de supervivencia y de riesgo
 - 6.3 Estimación de la función de supervivencia: método de Kaplan-Meier
 - 6.4 Comparación de funciones de supervivencia: test log-rank
 - 6.5 Modelos de regresión para datos de supervivencia: modelo de Cox
- 7. Análisis multivariante (introducción)
 - 7.1 Análisis de componentes principales (PCA): reducción de dimensionalidad
 - 7.2 Interpretación de los componentes y varianza explicada
 - 7.3 Representación gráfica y aplicaciones prácticas

Actividades formativas y Metodología

| Título | Horas | ECTS | Resultados de aprendizaje |
|---------------------|-------|------|---------------------------|
| Tipo: Dirigidas | | | |
| Clases de Practicas | 14 | 0,56 | 1, 2, 3, 4, 5 |
| Clases de Problemas | 14 | 0,56 | 1, 2, 3, 4, 5 |
| Clases de Teoría | 26 | 1,04 | 1, 2, 3, 4, 5 |
| Tipo: Autónomas | | | |
| Estudio | 60 | 2,4 | 1, 2, 3, 4, 5 |

La asignatura se estructura a partir de clases de teoría, problemas y prácticas. A las clases de teoría se les dará introducción a los conceptos y técnicas que describen el programa del curso. se puede seguir a través de la bibliografía básica recomendada. Las clases de problemas tienen como objetivo trabajar y entender los conceptos estadísticos. En el Campus Virtual se escriben las listas de problemas y, cuando se resuelven en clase, también las soluciones. El objetivo de las prácticas es la utilización del programa estadístico R, para obtener y aclarar los resultados de los procedimientos que se han introducido en las clases de teoría y problemas. En el Campus Virtual se leerá el enunciado de cada práctica con antelación.

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.

Evaluación

Actividades de evaluación continuada

| Título | Peso | Horas | ECTS | Resultados de aprendizaje |
|--------|------|-------|------|---------------------------|
|--------|------|-------|------|---------------------------|

| | | | | |
|---------------------------------|-----|----|------|---------------|
| Evaluación continua y prácticas | 30% | 32 | 1,28 | 1, 2, 3, 4, 5 |
| Examen final | 40% | 2 | 0,08 | 1, 2, 3, 4, 5 |
| Examen Parcial | 30% | 2 | 0,08 | 1, 2, 3, 4, 5 |

La evaluación de la asignatura consistirá en:

1. Examen parcial (30%)
2. Examen final (40%)
3. Evaluación continua y práctica (30%)

Ninguna de las actividades de Evaluación elimina la materia para el examen final. La nota final será la media ponderada de las actividades. No se establece ninguna política de nota mínima por actividad. Si aplicando los pesos mencionados anteriormente la calificación del alumno es 5 o superior, se considera superada la asignatura y ésta no podrá ser objeto de una nueva Evaluación. Un alumno se considera que está "No evaluado" en la asignatura siempre y que no ha participado de ninguna de las actividades de evaluación. Por lo tanto, se considera que un estudiante que realiza algún componente de Evaluación continuada ya no puede optar a un "No evaluado".

Proceso de Recuperación "Para participar en el Proceso de Recuperación el Alumno tiene que haber sido evaluado previamente en un conjunto de actividades que representen un mínimo de dos terceras partes de la calificación total de la asignatura o módulo." Apartado 3 del artículo 112 ter. La recuperación (Normativa Académica UAB). Los y las estudiantes deben haber obtenido una calificación media de la asignatura entre 4,0 y 4,9. Los datos de esta prueba estarán programados en el calendario de exámenes de la Facultad. El estudiante que esté presente y la supere aprobará la asignatura con una nota de 5. En caso contrario mantendrá la misma nota.

Irregularidades en el Acto de Evaluación

Sin perjuicio de otras medidas disciplinarias que se consideren oportunas, y de acuerdo con la normativa académica vigente, "en el caso de que el alumno presente alguna Irregularidad que pueda suponer una variación significativa en la calificación de un acta de evaluación, dicha acta de evaluación será calificada con un 0, con independencia del proceso disciplinario que se pueda incoar. en el caso de que se produzcan varias irregularidades en las actas de evaluación de una misma asignatura, la calificación final de la misma será 0". Apartado 10 del artículo 116. Resultados de la evaluación. (Normativa Académica UAB) La Propuesta de Evaluación puede sufrir alguna modificación en función de las restricciones impuestas por las autoridades sanitarias.

Bibliografía

1. Brenton, R. C. (2008). Linear models: the theory and applications of analysis of variance. ISBN: 978-0-470-0566-6.
2. Arnold, E. y Davis, Ch. S. (2002). Statistical methods for the analysis of repeated measurements. Springer.
3. Dobson, A.J. y Barnett, A.G. (2008). An introduction to generalized linear models. Series: Chapman & Hall/CRC texts in Statistical Science.
4. Fisher, R.A. (2003). Statistical methods, experimental design, and scientific inference. ISBN: 978-0-19-852229-4.
5. Gutiérrez P.H. (2003). Análisis y diseño de experimentos. McGraw-Hill.

6. Hocking, R. R. (2003). Methods and applications of linear models: regression and the analysis of variance. Wiley Series in Probability and Statistics. ISBN: 978-0-471- 23222-3.
7. Kish, L. (2004). Statistical design for research. Wiley Interscience.
8. Lindman, H. R. (1992). Analysis of variance in experimental design. Springer-Verlag.
9. Kuehl, R. O. (2001). Diseño de experimentos. Principios estadísticos del diseño y análisis de investigación. Thomson Learning.
10. Peña, D. (2002). Regresión y diseño de experimentos. Alianza.
11. Montgomery, D. C. (2002). Diseño y análisis de experimentos. Limusa-Wiley.
12. Scheiner, S.M. (2001). Design and analysis of ecological experiments. Oxford University Press.
13. Toutenburg, H. (2002). Statistical analysis of designed experiments. Springer.

Software

Se dará al inicio de la asignatura

Grupos e idiomas de la asignatura

La información proporcionada es provisional hasta el 30 de noviembre de 2025. A partir de esta fecha, podrá consultar el idioma de cada grupo a través de este [enlace](#). Para acceder a la información, será necesario introducir el CÓDIGO de la asignatura

| Nombre | Grupo | Idioma | Semestre | Turno |
|--------------------------|-------|-----------------|---------------------|-------|
| (PAUL) Prácticas de aula | 81 | Catalán/Español | primer cuatrimestre | tarde |