

Titulación	Tipo	Curso
2503852 Estadística Aplicada	OB	2

## Contacto

Nombre: María Merce Farre Cervello

Correo electrónico: merce.farre@uab.cat

## Idiomas de los grupos

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

## Prerrequisitos

Fundamentos de estadística descriptiva e inferencial y de probabilidades, así como conocer los rudimentos de programación con el lenguaje R.

## Objetivos y contextualización

El objetivo de la asignatura es el estudio de los Modelos Lineales, así como las aplicaciones en distintos ámbitos. Los métodos y técnicas se introducen a partir de ejemplos y se trabajan a partir de la solución de los problemas propuestos y prácticas informáticas diseñadas para su ejecución con el lenguaje R. En primer lugar, se presenta el modelo de regresión simple porque tiene numerosas aplicaciones y porque es una buena introducción a la comprensión del modelo múltiple. El modelo de regresión múltiple, expresado matricialmente y que incluye algunas variantes (polinomial, con regresoras ficticias, interacciones, etc.), constituye la segunda parte del curso. Utilizando el análisis de los residuos como herramienta principal, se analiza el ajuste del modelo y la especificación correcta, el desempeño de las hipótesis y la detección de datos "especiales" (anómalas y/o influyentes). Por último, se abordan temas de especial relevancia, como la multicolinealidad y la selección de variables.

## Resultados de aprendizaje

1. CM09 (Competencia) Valorar la adecuación de los modelos con la utilización e interpretación correcta de indicadores y gráficos.
2. CM09 (Competencia) Valorar la adecuación de los modelos con la utilización e interpretación correcta de indicadores y gráficos.
3. CM10 (Competencia) Modificar el software existente si el modelo estadístico propuesto lo requiere, o crear nuevo software, si fuera necesario.
4. KM12 (Conocimiento) Proporcionar las hipótesis experimentales de la modelización, teniendo en cuenta las implicaciones técnicas y éticas relacionadas.
5. KM12 (Conocimiento) Proporcionar las hipótesis experimentales de la modelización, teniendo en cuenta las implicaciones técnicas y éticas relacionadas.
6. KM13 (Conocimiento) Detectar interacciones, colinealidad e importancia entre variables explicativas.
7. SM11 (Habilidad) Analizar los residuos de un modelo estadístico.

8. SM12 (Habilidad) Interpretar los resultados obtenidos para formular conclusiones respecto a las hipótesis experimentales
9. SM13 (Habilidad) Comparar el grado de ajuste entre diversos modelos estadísticos.
10. SM14 (Habilidad) Emplear gráficos de visualización del ajuste y de la adecuación del modelo.

## Contenido

### 1. El modelo de Regresión lineal simple

- Introducción a los modelos de regresión y pasos previos en la regresión simple: Exploración de los datos.
- La regresión lineal simple: Modelo, hipótesis, parámetros.
- Estimación puntual de los parámetros del modelo: Método de mínimos cuadrados. Los estimadores de máxima verosimilitud.
- Inferencia sobre los parámetros del modelo bajo las hipótesis de Gauss-Markov: Intervalos y tests.
- Intervalo de confianza para la respuesta media e intervalo de predicción de nuevas observaciones. Inferencias simultáneas en la regresión simple. Bandas de confianza y de predicción.
- Análisis de la varianza (ANOVA) del modelo de regresión simple.
- Diagnósticos del modelo: Evaluación gráfica de la linealidad y verificación de las hipótesis mediante el análisis de los residuos. Test de falta de ajuste lineal.
- Datos anómalos o influyentes.

### 2. El modelo de regresión lineal múltiple

- Pasos previos en la regresión múltiple: Exploración de los datos con herramientas de visualización multidimensional.
- Expresión matricial del modelo y los estimadores de los coeficientes. Interpretación de los coeficientes del modelo múltiple.
- Leyes de los estimadores de los coeficientes, de las predicciones y de los residuos: Aplicación de las propiedades de las matrices idempotentes.
- Inferencia en el modelo lineal múltiple. Anova del modelo.
- Test de "ligaduras" para resolver restricciones lineales sobre los coeficientes: El principio de la variabilidad incremental.
- Discusión de las hipótesis del modelo lineal: Análisis de los residuos. Transformaciones de Box-Cox.
- El problema de la multi-colinealidad entre variables regresoras: Detección y soluciones.
- Variables ficticias en regresión (*dummies*): Interpretación de los coeficientes y aplicaciones.
- Selección de variables en un modelo lineal: El estadístico  $C_p$  de Mallows, la validación cruzada de modelos y la selección automática por pasos.

## Actividades formativas y Metodología

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Clases de teoría	26	1,04	
Prácticas tutorizadas	26	1,04	
Tipo: Autónomas			
Estudio y consultas	36	1,44	
Resolución de ejercicios con R	32	1,28	
Resolución de problemas	18	0,72	

La asignatura consta de teoría, problemas y prácticas. En la teoría se presentan y motivan las herramientas y métodos de modelos lineales y se hace un trabajo analítico. Se proporcionarán listas de problemas a lo largo del curso. Además, hay prácticas para analizar datos con el lenguaje de programación R. Se propondrán algunas tareas (ejercicios teóricos y con R) para entregar. Además de las entregas, el estudiante también realizará otro trabajo autónomo consistente en investigación bibliográfica y preparación de exámenes.

El material del curso (notas de teoría, listas de problemas y tareas de ordenador) estará disponible en el aula *moodle*.

La perspectiva de género va más allá de los contenidos de los cursos, ya que implica también una revisión de las metodologías y las interacciones entre los estudiantes y los profesores, tanto dentro como fuera del aula. En este sentido, las metodologías participativas de enseñanza que dan lugar a un entorno de igualdad, menos jerárquicas en el aula, evitando ejemplos estereotipados en el género y el vocabulario sexista, suelen ser más favorables a la plena integración y participación de las alumnas. Por ello, se hará esta aplicación efectiva durante el curso.

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
Entrega de tareas (problemas y prácticas resueltas)	20%	0	0	CM09, CM10, KM12, KM13, SM11, SM12, SM13, SM14
Examen final	80% (recupera los dos parciales)	4	0,16	CM09, KM13, SM11, SM12, SM13, SM14
Primer parcial	30%	4	0,16	CM09, KM13, SM11, SM12, SM13, SM14
Segundo parcial	50%	4	0,16	

---

PR: Entrega de los ejercicios teóricos y prácticos (con R). Valoración máxima de PR: 2 puntos. Esta parte no es recuperable.

P1: Prueba parcial de regresión simple (teoría, ejercicios, y prácticas). Valoración máxima de P1: 3 puntos.

P2: Prueba parcial de regresión múltiple (teoría, ejercicios y prácticas). Valoración máxima de P2: 5 puntos.

La nota de curso se calculará:  $NC = PR + P1 + P2$ . El aprobado por curso requiere que NC sea igual o mayor que 5 y que las notas de cada parcial sean mayores o iguales que 3.5 (sobre 10).

Al final del semestre se realizará un examen de recuperación que será una prueba de síntesis, PS, (teoría, ejercicios y prácticas) de los contenidos de todo el curso con una puntuación máxima de 8 puntos, por los alumnos que no hayan aprobado por curso o quieran mejorar la nota. Sólo se podrán presentar a la prueba de síntesis los estudiantes que hayan participado en 2/3 de las actividades de evaluación.

La nota final de los presentados a la prueba de síntesis se calculará:  $NF = PR + \max(PS, P1 + P2)$ .

Las matrículas de honor que eventualmente se concedan a partir de la NC no se retirarán incluso si otro estudiante obtiene una nota superior después de la PS.

#### Evaluación única

La evaluación única será una prueba de síntesis de las competencias de ambos parciales, en base a: (1) Un examen con cuestiones de teoría y de prácticas (peso: 50%). (2) Una prueba de prácticas delante del ordenador (peso: 40%). (3) La entrega de las tareas programadas que se indiquen, con la posibilidad de que el profesorado pida que el estudiante explique detalles de estas entregas (peso: 10%).

Atención: "Sin perjuicio de otras medidas disciplinarias que se estimen oportunas, y de acuerdo con la normativa académica vigente, se calificarán con un cero las irregularidades cometidas por el estudiante que puedan conducir a una variación de la calificación de un acto de evaluación. Por lo tanto, plagiar, copiar o dejar copiar una práctica o cualquier otra actividad de evaluación implicará suspender con un cero y no se podrá recuperar en el mismo curso académico. Si esta actividad tiene una nota mínima asociada, entonces la asignatura quedará suspendida. "

#### Bibliografía

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

Clarke, B.R.; Linear Models: The Theory and Applications of Analysis of variance. Wiley, 2008.

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

Fox, J. and Weisberg, S.; An R Companion to Applied Regression. Sage Publications, 2nd edition, 2011.

N. R. Mohan Madhyastha; S. Ravi; A. S. Praveena. A First Course in Linear Models and Design of Experiments. 2020.

<https://link-springer-com.are.uab.cat/content/pdf/10.1007%2F978-981-15-8659-0.pdf>

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

#### Bibliografía complementaria:

Sen, A., Srivastava, M.; Regression Analysis: Theory, Methods and Applications. Springer, 1990.

Neter, M. H. Kutner, C. J. Nachtsheim, W. Wasserman; .Applied Linear Models. Irwin (4th edition), 1996.

Faraway, J.; Linear Models with R. Chapman&Hall/CRC (2nd ed), 2014.

Rao, C. R., Toutenburg, H., Shalabh, Heumann, C; Linear Models and generalizations. Springer, 2008.

## Software

Software libre con R en el entorno RStudio.

## Lista de idiomas

Nombre	Grupo	Idioma	Semestre	Turno
(PLAB) Prácticas de laboratorio	1	Catalán	segundo cuatrimestre	tarde
(PLAB) Prácticas de laboratorio	2	Catalán	segundo cuatrimestre	tarde
(TE) Teoría	1	Catalán	segundo cuatrimestre	tarde