

Titulación	Tipo	Curso
2503740 Matemática Computacional y Analítica de Datos	OB	2

## Contacto

Nombre: Rosario Delgado De la Torre

Correo electrónico: rosario.delgado@uab.cat

## Idiomas de los grupos

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

## Prerrequisitos

Se considera muy importante un buen conocimiento de los contenidos de las asignaturas cursadas durante el primer curso, especialmente las de probabilidad y cálculo.

## Objetivos y contextualización

Esta asignatura es la primera del Grado dedicada a la Inferencia Estadística, que es la parte de la Estadística que permite obtener, de manera controlada, información sobre una población a partir de los datos de una muestra "representativa". La asignatura tiene un carácter central dentro de los estudios, ya que se presentan conceptos y técnicas que serán utilizados en muchas de las materias que se cursarán a partir de ahora. Concretamente, se empezará haciendo una introducción a la Estadística, y luego se tratará la estimación de parámetros, tanto puntual como por intervalos de confianza, así como los tests de hipótesis paramétricos clásicos para una y dos poblaciones normales y dicotómicas, y contrastes de independencia. Finalmente, se introducirá el modelo de regresión lineal simple. Se hará especial énfasis en los métodos estadísticos que se pueden utilizar para comparar algoritmos de aprendizaje automático.

La presencialidad de la docencia y de las actividades evaluables se adaptará siguiendo las recomendaciones de las autoridades sanitarias, a fin de garantizar la seguridad de todas las personas.

## Resultados de aprendizaje

1. CM14 (Competencia) Implementar estrategias para confirmar o refutar hipótesis.
2. CM15 (Competencia) Gestionar la información para su validación mediante un tratamiento estadístico.
3. CM15 (Competencia) Gestionar la información para su validación mediante un tratamiento estadístico.
4. CM16 (Competencia) Evaluar, a partir de datos obtenidos, desigualdades por razón de sexo/género.
5. KM12 (Conocimiento) Identificar la inferencia estadística como instrumento de pronóstico y predicción.
6. KM12 (Conocimiento) Identificar la inferencia estadística como instrumento de pronóstico y predicción.
7. KM12 (Conocimiento) Identificar la inferencia estadística como instrumento de pronóstico y predicción.
8. KM12 (Conocimiento) Identificar la inferencia estadística como instrumento de pronóstico y predicción.
9. KM13 (Conocimiento) Describir las propiedades básicas de los estimadores puntuales y de intervalo.
10. KM13 (Conocimiento) Describir las propiedades básicas de los estimadores puntuales y de intervalo.
11. KM14 (Conocimiento) Identificar la utilidad de los métodos Bayesianos, aplicándolos oportunamente.

12. KM14 (Conocimiento) Identificar la utilidad de los métodos Bayesianos, aplicándolos oportunamente.
13. SM14 (Habilidad) Utilizar las propiedades de las funciones de densidad y de distribución.
14. SM14 (Habilidad) Utilizar las propiedades de las funciones de densidad y de distribución.
15. SM15 (Habilidad) Utilizar software estadístico apropiado para gestionar bases de datos, para obtener índices de resumen de las variables del estudio y para analizar los datos mediante técnicas de inferencia.
16. SM15 (Habilidad) Utilizar software estadístico apropiado para gestionar bases de datos, para obtener índices de resumen de las variables del estudio y para analizar los datos mediante técnicas de inferencia.
17. SM15 (Habilidad) Utilizar software estadístico apropiado para gestionar bases de datos, para obtener índices de resumen de las variables del estudio y para analizar los datos mediante técnicas de inferencia.

## Contenido

Preliminares de Probabilidad (recordatorio): Probabilidad y variables aleatorias. Concepto de ley. Distribuciones

discretas. Función de densidad y de probabilidad. Esperanza y varianza. Función generatriz de momentos. Ejemplos.

Tema 1. Introducción a la Estadística.

1. Estadística descriptiva y estadística inferencial.

1.1. Conceptos básicos en inferencia: población estadística y muestra; parámetros, estadísticos y estimadores.

1.2. Modelos estadísticos: paramétricos y no paramétricos.

2. Estadísticos más usuales: los momentos muestrales. Los estadísticos de orden.

3. Distribución de algunos estadísticos.

3.1. De una muestra de una población Normal: Teorema de Fisher.

3.2. El Teorema Central del Límite: normalidad asintótica de los momentos muestrales y de la proporción.

Tema 2: Estimación puntual.

1. Estimadores puntuales: definición y "buenas" propiedades.

1.1. Sesgo.

1.2. Comparación de estimadores sin sesgo. Eficiencia relativa.

1.3. Comparación de estimadores con sesgo: el error cuadrático medio.

1.4. Consistencia de un estimador.

2. Métodos para obtener estimadores.

2.1. Método de los momentos.

2.2. Método de la máxima verosimilitud (EMV).

Tema 3. Estimación por intervalos de confianza.

1. Concepto de región e intervalo de confianza.
2. El método del "pivote" para la construcción de intervalos de confianza.
3. Intervalos de confianza para los parámetros de una población.
  - 3.1. Para la media de una población Normal con desviación conocida y desconocida.
  - 3.2. Para la varianza de una población Normal con media conocida y desconocida.
  - 3.3. Otras aplicaciones del método del pivote.
  - 3.4. Intervalos de confianza asintóticos.
4. Intervalos de confianza para los parámetros de dos poblaciones.
  - 4.1. Intervalos de confianza con muestras independientes.
  - 4.2. Intervalos de confianza para la diferencia de medias de dos poblaciones Normales con datos aparejados.

#### Tema 4: Tests de hipótesis.

1. Introducción.
  - 1.1. Errores tipo I y II.
  - 1.2. Función potencia.
  - 1.3. Consistencia de las pruebas.
  - 1.4. p-valores.
  - 1.5. Dualidad entre los intervalos de confianza y los test de hipótesis.
2. Tests para los parámetros de una población.
  - 2.1. Para la media de una población Normal con desviación conocida y desconocida.
  - 2.2. Tests asintótico para la media de una población cuando la muestra es grande.
  - 2.3. Para la varianza de una población Normal.
  - 2.4. Tests asintóticos: Wald, Score y LRT.
3. Tests para los parámetros de dos poblaciones.
  - 3.1. Tests de hipótesis con muestras independientes.
  - 3.2. Tests de hipótesis con datos aparejados.
4. Pruebas de la ji-cuadrado.
  - 4.1. De bondad de ajuste.
  - 4.2. De independencia.
5. Tests no paramétricos para comparar algoritmos de aprendizaje automático.

#### Tema 5. Regresión lineal simple.

1. Objetivos del modelo.

2. Estimadores de mínimos cuadrados ordinarios (MCO).
3. Inferencia con el modelo de regresión simple.
4. Predicciones.

**IMPORTANTE:** En la docencia, la perspectiva de género implica revisar los sesgos androcéntricos y cuestionar los supuestos y estereotipos de género ocultos. Esta revisión conlleva incluir a los contenidos de la asignatura el conocimiento producido por las mujeres científicas, a menudo olvidadas, procurando el reconocimiento de sus aportaciones, así como el de sus obras en las referencias bibliográficas.

## Actividades formativas y Metodología

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Clase de problemas	12	0,48	
Clases de prácticas	10	0,4	
Clases de teoría	27	1,08	
Tipo: Autónomas			
Exámenes	15	0,6	
Resolución de problemas	33	1,32	
Resolución de prácticas	23	0,92	

La asignatura se estructura a partir de clases de teoría, problemas y prácticas.

En las clases de teoría iremos introduciendo los conceptos y técnicas que describe el programa del curso. Teniendo en cuenta que el contenido es esencialmente el estándar de un primer curso de inferencia estadística, se puede seguir haciendo uso de la bibliografía básica recomendada. También se irá colgando en el Campus Virtual el material correspondiente a cada tema explicado en las clases presenciales.

Las clases de problemas tienen por objetivo trabajar y entender los conceptos estadísticos. En el Campus Virtual se colgarán las listas de problemas.

El objetivo de las prácticas es la utilización de software 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 colgará el enunciado de cada práctica con antelación.

**IMPORTANTE:** Para trabajar más cómodamente con R, se recomienda utilizar la interfaz RStudio: es libre, "Open source" y funciona con Windows, Mac y Linux. <https://www.rstudio.com/>

**OBSERVACIÓN:** La perspectiva de género en la docencia va más allá de los contenidos de las asignaturas, ya que también implica una revisión de las metodologías docentes y de las interacciones entre el alumnado y el profesorado, tanto en el aula como fuera. En este sentido, las metodologías docentes participativas, donde se genera un entorno igualitario, menos jerárquico en el aula, evitando ejemplos estereotipados engéneros y vocabulario sexista, con el objetivo de desarrollar el razonamiento crítico y el respeto a la diversidad y pluralidad de ideas, personas y situaciones, suelen ser más favorables a la integración y plena participación de las alumnas en el aula, y por eso se procurará su implementación efectiva en esta asignatura.

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
Examen final	0,60	10	0,4	CM14, KM12, KM13, KM14, SM14
PEC1: Examen parcial	0,25	8	0,32	CM14, KM12, KM13, KM14, SM14
PEC2: Ejercicios evaluables	0,15	12	0,48	CM15, CM16, SM15

En las clases de teoría se introducirán los conceptos básicos de la asignatura y se presentará un amplio conjunto de ejemplos. En las clases de problemas y prácticas, se resolverán ejercicios y se realizarán prácticas con R. Se recomienda la asistencia a clase para tener una idea sobre el curso en general, así como de los ejercicios y prácticas.

Evaluación:

La calificación del estudiante será la media ponderada de las siguientes actividades:

PAC1: examen parcial, que supone el 25% de la nota.

PAC2: entregas de ejercicios relacionados con las prácticas de ordenador con R realizadas en el aula, que supone un 15% de la nota.

Examen final: que constará de algunas preguntas de tipo conceptual en forma de cuestiones cortas y de unos problemas en los que tendrá que resolver una serie de ejercicios similares a los que se han trabajado en las sesiones de clase. Esta prueba supone el 60% restante de la nota.

Importante: si la nota de alguna de estas actividades no alcanza el 3 sobre 10, contará como 0 en el cómputo de la calificación final.

Recuperación: si esta calificación no llega a 5, el estudiante tiene derecho a otra oportunidad de superar la asignatura mediante el examen de recuperación. En esta prueba podrá recuperarse el 85% de la nota correspondiente al examen final y al PAC1. La parte práctica con R (PAC2) no es recuperable. En ningún caso el examen de recuperación puede servir para subir la nota si el alumno ya ha superado la asignatura con el primer examen.

Evaluación única:

El alumnado que se haya acogido a la modalidad de evaluación única deberá realizar una prueba final que consistirá en un examen que constará de algunas preguntas de tipo conceptual en forma de cuestiones cortas y de unos problemas en los que deberá resolver una serie de ejercicios similares a los que se han trabajado en las sesiones de clase. Una vez finalizado, entregará, además del examen, los ejercicios relacionados con las prácticas de ordenador con R realizados a lo largo del curso.

La calificación del estudiante será la media ponderada de las dos actividades anteriores, donde la prueba final supondrá el 85% de la nota, y la evaluación de las hojas de respuestas de las prácticas de ordenador con R el 15% restante.

Importante: si la nota de alguna de estas actividades no alcanza el 3 sobre 10, contará como 0 en el cómputo de la calificación final.

Si esta calificación no alcanza 5, el estudiante tiene derecho a otra oportunidad de superar la asignatura mediante el examen de recuperación que se celebrará en la fecha que fije la coordinación de la titulación. En esta prueba podrá recuperarse el 85% de la nota correspondiente a la prueba final. La parte práctica con R no es recuperable. En ningún caso el examen de recuperación puede servir para subir la nota si el alumno ya ha superado la asignatura con el primer examen.

## Bibliografía

1. Daalgard, P.: Introductory Statistics with R. Springer. 2008.
2. Daniel, W.W.: Biostatistics. Wiley. 1974.
3. DeGroot, M. H.: Schervish, M.J. Probability and Statistics. Pearson Academic. 2010.
4. Delgado, R.: Probabilidad y Estadística con aplicaciones. 2018.  
<https://www.amazon.es/Probabilidad-Estad%C3%ADstica-aplicaciones-Rosario-Delgado/dp/1983376906>
5. Heumann, C., Schomaker, M., Shalbh: Introduction to Statistics and Data Analysis: With Exercises, Solutions and Applications in R. Second Edition. Springer. 2023.  
Disponible on-line a través de la UAB: <https://link.springer.com/book/10.1007/978-3-031-11833-3>
6. Plaeue, M.: Data Science: An Introduction to Statistics and Machine Learning. Springer. 2023.  
Disponible on-line a través de la UAB: <https://link.springer.com/book/10.1007/978-3-662-67882-4>
7. R Tutorial. An introduction to Statistics. <https://cran.r-project.org/manuals.html>.
8. Salsburg, D. The Lady tasting tea : how statistics revolutionized science in the twentieth century. 2002. ISBN-13: 978-0805071344
9. Silvey, S.D.: Statistical Inference. Chapman&Hall. 1975.

<https://app.jove.com/science-education/v/12796/introduction-to-statistics>

## Software

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

## Lista de idiomas

Nombre	Grupo	Idioma	Semestre	Turno
(PLAB) Prácticas de laboratorio	1	Catalán	primer cuatrimestre	mañana-mixto
(SEM) Seminarios	1	Catalán	primer cuatrimestre	mañana-mixto
(TE) Teoría	1	Catalán	primer cuatrimestre	mañana-mixto