

Modelización e Inferencia

Código: 104392
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
2503740 Matemática Computacional y Analítica de Datos	OB	2	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: Amanda Fernandez Fontelo
Correo electrónico: Amanda.Fernandez@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

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.

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.

Competencias

- Calcular y reproducir determinadas rutinas y procesos matemáticos con agilidad.
- Evaluar de manera crítica y con criterios de calidad el trabajo realizado.
- Formular hipótesis e imaginar estrategias para confirmarlas o refutarlas.
- Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio.
- 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.
- Relacionar objetos matemáticos nuevos con otros conocidos y deducir sus propiedades.
- Utilizar aplicaciones informáticas de análisis estadístico, cálculo numérico y simbólico, visualización gráfica, optimización u otras para experimentar y resolver problemas.

Resultados de aprendizaje

1. Analizar datos mediante técnicas de inferencia para una o dos muestras.
2. Conocer los diferentes métodos de recogida de datos.
3. Describir las propiedades básicas de los estimadores puntuales y por intervalo.
4. Escoger el software estadístico apropiado para analizar los datos mediante técnicas de inferencia.
5. Evaluar de manera crítica y con criterios de calidad el trabajo realizado.
6. Identificar distribuciones estadísticas.
7. Identificar la inferencia estadística como instrumento de pronóstico y predicción.
8. Interpretar los resultados obtenidos y concluir respecto a la hipótesis experimental.
9. Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio.
10. 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.
11. 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.
12. Reconocer la utilidad de los métodos bayesianos y aplicarlos oportunamente.
13. Utilizar las propiedades de las funciones de densidad.
14. Utilizar las propiedades de las funciones de distribución.
15. Utilizar software estadístico para gestionar bases de datos.
16. Utilizar software estadístico para obtener índices de resumen de las variables del estudio.
17. Validar y gestionar la información para su tratamiento estadístico.

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.

1.5. Estadísticos suficientes.

2. Métodos para obtener estimadores.

2.1. Método de los momentos.

2.2. Método de la máxima verosimilitud (EMV).

2.2.1. Invarianza de los EMV.

2.2.2. Función Score e información de Fisher.

2.2.3. Cota de Crámer-Rao.

2.2.4. Propiedades de los EMV.

2.2.5. Método Delta.

2.2.6. Procedimientos numéricos para determinar los 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. Para la varianza de una población Normal con media conocida.

3.4. Intervalos de confianza asintóticos: Wald, Score and LRT.

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.

5. Técnicas bootstrap.

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: Wals, 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.

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. También se procurará introducir en la parte más práctica de la asignatura, el análisis y comparación de datos estadísticos por sexo, comentando a la hora de la clase las causas y los mecanismos sociales y culturales que pueden sustentar las desigualdades observadas.

Metodología

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 y, cuando ya se hayan resuelto en clase, también las soluciones.

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 en género 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.

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Clase de problemas	18	0,72	5, 2, 3, 6, 7, 8, 9, 10, 11, 13, 14
Clases de prácticas	12	0,48	1, 5, 4, 7, 8, 9, 10, 11, 15, 16, 17
Clases de teoría	30	1,2	1, 5, 2, 3, 6, 7, 8, 9, 10, 11, 13, 14
Tipo: Autónomas			
Exámenes	15	0,6	1, 5, 2, 3, 6, 7, 8, 9, 10, 11, 13, 14
Resolución de problemas	25	1	1, 5, 2, 3, 6, 7, 8, 9, 10, 11, 13, 14
Resolución de prácticas	20	0,8	1, 5, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17

Evaluación

La nota de evaluación continua se obtendrá a partir de un control de los problemas que dará una nota C, y a partir de un control de las prácticas de la asignatura que dará una nota P. La nota C tiene un peso del 20% y la nota P un peso del 30%. La nota del examen final E1 vale un 50% de la nota final. Con las notas C, P y E1 se obtiene la nota de la asignatura, N, de la siguiente forma:

$$N = 0.50 \times E1 + 0.20 \times C + 0.30 \times P$$

Recuperación y / o mejora de la nota de examen:

El alumno supera la asignatura si N es mayor o igual que 5. En caso contrario, o bien si el alumno quiere mejorar nota, hay una posibilidad de mejorar la parte de la nota del examen E1 mediante un examen de recuperación, la nota será E2. Así, a partir de esta nota de recuperación se obtiene la nota final de la asignatura:

$$NF = 0.50 \times \max(E1, E2) + 0.20 \times C + 0.30 \times P$$

Observación 1: Las notas C y P de evaluación continua no son recuperables.

Observación 2: Se considera que el alumno se ha presentado a la convocatoria de la asignatura si se presenta cualquiera de los dos exámenes que dan lugar a las notas E1 o E2. en caso contrario, será un No Presentado, aunque tenga alguna nota de evaluación continuada (C y / o P).

Actividades de evaluación

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Entrega de problemas (C)	0,20	8	0,32	1, 5, 2, 3, 6, 7, 8, 9, 10, 11, 12, 13, 14
Examen de prácticas (P)	0,30	12	0,48	1, 5, 2, 3, 4, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17
Examen final / Recuperación (E)	0,50	10	0,4	1, 5, 2, 3, 6, 7, 8, 9, 10, 11, 13, 14

Bibliografía

Berger, R.L., Casella, G.: Statistical Inference. Duxury Advanced Series. 2002.

Daalgard, P.: Introductory Statistics with R. Springer. 2008.

Daniel, W.W.: Biostatistics. Wiley. 1974.

DeGroot, M. H.: Schervish, M.J. Probability and Statistics. Pearson Academic. 2010.

Peña, D.: Estadística. Fundamentos de estadística. Alianza Universidad. 2001.

R Tutorial. An introduction to Statistics. <https://cran.r-project.org/manuals.html>. juny 2019.

Silvey, S.D.: Statistical Inference. Chapman&Hall. 1975.

Held, Sabanes Bove (2013): Applied Statistical Inference: Likelihood and Bayes. Springer

Pawitan (2013): In all Likelihood: Statistical Modelling and Inference using Likelihood. Oxford University Press

Young, Smith (2005): Essentials of Statistical Inference. Cambridge University Press

Cox, D.R. and Hinkley, D.V. (1979). Theoretical Statistics. 1st Edition, Chapman and Hall/CRC

Software

<https://www.r-project.org/>