

**Anàlisi de Dades Complexes**

Codi: 104399

Crèdits: 6

Titulació	Tipus	Curs	Semestre
2503740 Matemàtica Computacional i Analítica de Dades	OB	2	2

**Professor/a de contacte**

Nom: Pedro Puig Casado

Correu electrònic: pere.puig@uab.cat

**Idiomes dels grups**

Podeu accedir-hi des d'aquest [enllaç](#). Per consultar l'idioma us caldrà introduir el CODI de l'assignatura. Tingueu en compte que la informació és provisional fins a 30 de novembre de 2023.

**Equip docent extern a la UAB**

Dorota Mlynarczyk

**Prerequisits**

Es recomana un bon coneixement de l'assignatura Modelització i Inferència així com un certa practica amb el software R.

**Objectius**

El principal objectiu és proporcionar eines estadístiques per a l'anàlisi de dades, dominant les tècniques més rellevants per a poder enfrontar-se amb models complexes.

**Competències**

- Avaluar de manera crítica i amb criteris qualitat el treball realitzat.
- Calcular i reproduir determinades rutines i processos matemàtics amb agilitat.
- Formular hipòtesis i imaginar estratègies per confirmar-les o refutar-les.
- Que els estudiants hagin demostrat que comprenen i tenen coneixements en una àrea d'estudi que parteix de la base de l'educació secundària general, i se sol trobar a un nivell que, si bé es basa en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda d'aquell camp d'estudi.
- Que els estudiants sàpiguen aplicar els coneixements propis a la seva feina o vocació d'una manera professional i tinguin les competències que se solen demostrar per mitjà de l'elaboració i la defensa d'arguments i la resolució de problemes dins de la seva àrea d'estudi.

- Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins de la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes destacats d'índole social, científica o ètica.
- Treballar cooperativament en un context multidisciplinar asumido y respetando el rol de los diferentes miembros del equipo.
- Utilitzar aplicacions informàtiques d'anàlisi estadística, càlcul numèric i simbòlic, visualització gràfica, optimització o altres per experimentar i resoldre problemes.

## Resultats d'aprenentatge

1. Analitzar dades mitjançant tècniques d'inferència per a una o dues mostres.
2. Avaluar de manera crítica i amb criteris de qualitat el treball desenvolupat.
3. Conèixer els diferents mètodes de recollida de dades.
4. Escollir el programari estadístic apropiat per analitzar les dades mitjançant tècniques d'inferència.
5. Identificar la inferència estadística com a instrument de pronòstic i predicció.
6. Identificar les diferents fonts d'informació disponibles.
7. Interpretar els resultats obtinguts i concloure respecte a la hipòtesi experimental.
8. Que els estudiants hagin demostrat que comprenen i tenen coneixements en una àrea d'estudi que parteix de la base de l'educació secundària general, i se sol trobar a un nivell que, si bé es basa en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda d'aquell camp d'estudi.
9. Que els estudiants sàpiguen aplicar els coneixements propis a la seva feina o vocació d'una manera professional i tinguin les competències que se solen demostrar per mitjà de l'elaboració i la defensa d'arguments i la resolució de problemes dins de la seva àrea d'estudi.
10. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins de la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes destacats d'índole social, científica o ètica.
11. Treballar cooperativament en un context multidisciplinari assumint i respectant el rol dels diferents membres de l'equip.
12. Utilitzar les propietats de les funcions de densitat.
13. Utilitzar les propietats de les funcions de distribució.
14. Utilitzar programari estadístic per gestionar bases de dades.
15. Utilitzar programari estadístic per obtenir índexs de resum de les variables de l'estudi.
16. Validar i gestionar la informació per fer-ne el tractament estadístic.

## Continguts

1- Models lineals: regressió múltiple i ANOVA.

2- Models lineals generalitzats: regressió logística i de Poisson.

3- Mètodes de remostreig 1: tests permutacionals.

4- Mètodes de remostreig 2: bootstrap.

5- Mètodes de remostreig 3: Jackknife.

Si hi ha temps també es farà una introducció a l'Anàlisi de Components Principals.

## Metodologia

D'acord amb els objectius de l'assignatura, el desenvolupament del curs es basa en les següents activitats:

Classes de teoria: L'alumne adquireix els coneixements científic-tècnics propis de la assignatura assistint a les classes de teoria i complementant-les amb l'estudi personal dels temes explicats. Les classes de teoria són les activitats en les quals s'exigeix menys interactivitat a l'estudiant: estan concebudes com un mètode fonamentalment unidireccional de transmissió de coneixements del professor a l'alumne. Les classes es faran utilitzant un suport de diapositives PowerPoint en Anglès que es penjaran també al Campus Virtual.

Problemes i pràctiques: Els problemes i les pràctiques són sessions amb una doble missió. D'una banda es treballen els coneixements científic-tècnics exposats en les classes de teoria per a completar la seva comprensió i aprofundir en ells desenvolupant activitats diverses, des de la típica resolució de problemes fins la discussió de casos pràctics. D'altra banda, les classes de problemes són el fòrum natural en el qual discutir en comú el desenvolupament del treball pràctic, aportant els coneixements necessaris per a portar-lo endavant, o indicant on i com es poden adquirir.

El curs pràctic d'aquesta assignatura es planteja com un camí per a orientar l'estudiant en un treball de camp d'estadística en cadascuna de les seves etapes. Així realitzen pel seu compte amb el programari R uns exercicis pràctics dirigits a resoldre problemes reals concrets. Aquest plantejament està orientat a promoure un aprenentatge actiu i a desenvolupar el raonament crític i la capacitat d'anàlisi i síntesi.

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, per a la complementació per part de l'alumnat de les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura/mòdul.

## Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Classes de Teoria	24	0,96	2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 16
Classes de problemes	20	0,8	2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13
Tipus: Supervisades			
Sessions pràctiques	20	0,8	1, 2, 4, 11, 14, 15, 16
Tipus: Autònomes			
Treball personal	61	2,44	2, 4, 5, 6, 7, 9, 10, 12, 13, 14, 15

## Avaluació

L'avaluació es realitza al llarg de tot el curs. L'avaluació continuada té diversos objectius fonamentals: Monitoritzar el procés d'ensenyament i aprenentatge, permetent tant a l'alumne com al professor conèixer el grau d'assoliment de les competències i corregir, si és possible, les desviacions que es produeixin. Incentivar l'esforç continuat de l'alumne enfront del sobreesforç, freqüentment inútil, d'última hora. Verificar que l'alumne ha assolit les competències determinades en el pla d'estudis.

Per a fer aquesta avaluació es compta amb els següents instruments: Els exercicis pràctics lliurats pels estudiants (30%), un examen parcial de Teoria al mig del curs (35%), un altre examen parcial de Teoria al final del curs (35%). A la recuperació només podran anar els alumnes que tinguin un mínim de 3 en la nota final, recuperant només la part de Teoria.

L'alumnat que s'hagi acollit a la modalitat d'avaluació única haurà de realitzar una prova final que consistirà en un examen on hi podran haver qüestions de teoria i resolució de problemes, i un examen de pràctiques

d'abant l'ordinador. Aquesta prova es realitzarà el mateix dia, hora i lloc que es realitzi la prova del segon parcial. Qui no es presenti a dita prova sense causa justificada, obtindrà la qualificació de NO AVALUABLE. Si s'obté una nota inferior a 5, es podrà recuperar el mateix dia, hora i lloc que es realitzi la recuperació de la resta d'alumnes del curs.

## Activitats d'avaluació continuada

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Examen parcial 1	35	2,5	0,1	1, 3, 5, 6, 7, 8, 9, 10, 12, 13, 16
Examen parcial 2	35	2,5	0,1	1, 3, 5, 6, 7, 8, 9, 10, 12, 13, 16
Exercicis	30	20	0,8	2, 4, 11, 14, 15

## Bibliografia

- Introduction to Linear Regression Analysis. Montgomery, D. Peck, A. Vining, G., 2001.
- An R Companion to Linear Statistical Models. Christopher Hay-Jahans, 2012.
- Generalized Linear Models. McCullagh, P. and Nelder, J., 1992.
- Resampling methods: a practical guide to data Analysis. Phillip I. Good, 2006.
- The jackknife, the bootstrap and other resampling plans. Bradley Efron, 1982.
- Bootstrap methods and their application. A.C. Davison, D.V. Hinkley, 1997.

## Programari

Es farà servi el programari R.