

Anàlisi Multivariant

Dades generals

Nom de l'assignatura	Anàlisi Multivariant
Codi de l'assignatura	361232
Curs acadèmic	2025/2026
Coordinació	Sergi Ramirez Mitjans
Departament	Facultat d'Economia i Empresa
Crèdits	6
Programa únic	S

Hores estimades de dedicació

Hores totals: 150 Hores

Activitats	Tipus de formació	Hores	Observacions
Activitats presencials i/o no presencials		60	
- Teoricopràctica	Presencial	30	
- Pràctiques d'ordinadors	Presencial	30	Es fan a l'aula d'informàtica, però inclouen pràctica d'ordinadors i també desenvolupament d'un projecte llarg.
Treball tutelat/dirigit		40	
Aprenentatge autònom		50	

Recomanacions

Capacitats prèvies:

- Àlgebra lineal: espai vectorial real, mètriques, projeccions, diagonalització de matrius, etc.
- Coneixements bàsics de teoria de la probabilitat i d'estadística descriptiva i inferencial.
- Coneixements bàsics de R i algorítmia per dissenyar scripts d'anàlisi automatitzada.

Competències / Resultats d'aprenentatge que es desenvolupen

- Capacitat d'aprenentatge i responsabilitat (capacitat d'anàlisi, de síntesi, de visions globals i d'aplicació dels coneixements a la pràctica / capacitat de prendre decisions i d'adaptació a noves situacions).
- Capacitat creativa i emprenedora (capacitat de formular, dissenyar i gestionar projectes / capacitat de cercar i integrar nous coneixements i actituds).
- Capacitat de reunir i d'interpretar dades rellevants que permetin d'emetre informes raonats i obtenir conclusions en problemes científics o d'altres àmbits que requereixin eines matemàtiques.
- Saber aplicar els coneixements adquirits i la capacitat d'anàlisi a la resolució de problemes en contextos acadèmics i professionals.
- Capacitat per usar, interpretar, documentar i adaptar eines informàtiques per a l'anàlisi estadística i la gestió de bases de dades, que permeti l'ajust de models i la resolució de problemes.
- Capacitat d'ordenar, representar i resumir, amb criteris objectius, la informació proporcionada per un conjunt de dades.
- Capacitat per utilitzar el raonament lògic i els instruments matemàtics en un context .
- Capacitat per seleccionar el mètode més adequat en la realització d'un estudi estadístic, d'avaluar les possibles alternatives i, si és procedent, incloure-hi l'anàlisi de costos i de recursos disponibles.
- Capacitat de proposar, modelitzar, analitzar, validar i interpretar situacions i problemes reals, adaptant els models teòrics a les necessitats específiques de les diferents àrees d'aplicació.

Objectius d'aprenentatge

Referits a coneixements

L'objectiu de l'assignatura és presentar tècniques estadístiques d'anàlisi de grans taules per tal d'extreure de manera ràpida la informació més rellevant continguda en les dades; els problemes abordats són de diversa tipologia: des de la definició d'eixos dominants a la caracterització estadística de sub poblacions. Aquest objectiu es particularitza presentant àmpliament el punt de vista de tres grans famílies de tècniques estadístiques multivariants:

1. Tècniques multivariants de classificació automàtica orientades a establir tipologies i a caracteritzar-les; es veuen diferents famílies de mètodes, des dels més clàssics als més recents: mètodes de particions, mètodes jeràrquics, mètodes basats en densitats; s'incideix especialment en eines d'interpretació de les classes; s'estudia l'adequació dels diferents mètodes a diferents casos, en funció de l'escalabilitat, el tipus de dades, etc.

2. Tècniques multivariants enfocades a sintetitzar i a resumir la informació, estudiar relacions multidimensionals entre variables i, eventualment, definir indicadors latents; es concreta en tres tècniques fonamentals: anàlisi en components principals, anàlisi de correspondències simples i anàlisi de correspondències múltiples; es planteja l'anàlisi factorial com a marc formal general del qual es deriven les tècniques esmentades com a cas particular; es dona particular importància a l'anàlisi de resultats gràfics; s'il·lustren algunes extensions addicionals, com ara l'anàlisi textual.

3. Tècniques d'anàlisi discriminant, es tracta de tècniques multivariants enfocades a obtenir regles d'assignació; s'incideix en la seva relació amb les tècniques vistes anteriorment.

4. Tècniques d'anàlisi textual, on es treballa amb textos lliures provinents de documents, pàgines web o xarxes socials per identificar-ne els conceptes subjacents i les relacions entre ells.

L'objectiu de l'assignatura és doble des del punt de vista conceptual. D'una banda, es vol donar una base formal sòlida per a les tècniques multivariants que componen el programa. De l'altra, l'alumnat ha de desenvolupar una capacitat pràctica d'aplicació a dades reals d'aquestes tècniques. En aquest sentit, les sessions de pràctiques segueixen el temari de l'assignatura des de la perspectiva de l'aplicació, i es treballa amb dades reals. Amb aquest objectiu, cal introduir un pas previ de preprocessament de dades per preparar-les per a l'anàlisi.

Finalment, i tenint en compte que el curs no pot ser exhaustiu i que posteriorment es poden tractar altres aspectes, es presenten diverses tècniques multivariants d'una manera més introductòria i enfocant-les menys algebraicament des d'un punt de vista més algorímic.

Referits a habilitats, destreses

En aquesta assignatura es dona particular importància a l'entrenament en certes competències transversals importants en el desenvolupament professional d'un estadístic, com ara la capacitat d'anàlisi, síntesi, comunicació, integració de coneixements, redacció d'informes i sobretot el treball en equip, incloent-hi les capacitats de planificació a mitjà termini, repartiment de tasques i gestió d'incidències en el pla de treball al llarg del curs.

La pràctica s'estructura sota un esquema que permet entrenar aquestes capacitats amb els suports necessaris del professorat de l'assignatura.

Blocs temàtics

1 Introducció

1.1 Introducció i preprocessament de dades

2 Clustering

Presentació conceptual. Mètodes de particions. Mètodes jeràrquics. Interpretació de les classes. Descripció de tipologies. Aplicació a casos reals i implicacions pràctiques

3 Escalament multidimensional

Formalització general, resultats teòrics.

4 Anàlisi en components principals

Formalització, resultats teòrics, interpretació, aplicació a casos reals, implicacions pràctiques

5 Anàlisi en correspondències simples

Formalització, resultats teòrics. Interpretació. Aplicacions a casos reals. Implicacions pràctiques

6 Anàlisi en correspondències múltiples

Formalització. Resultats teòrics. Interpretació. Aplicació a casos reals i implicacions pràctiques que evidenciïn avantatges en el tractament de dades d'enquesta. Introducció a l'anàlisi factorial amb dades mixtes.

7 Anàlisis de correlacions canòniques

Formalització, resultats teòrics, interpretació, aplicació a casos reals, implicacions pràctiques

8 Anàlisi discriminant

Formalització, resultats teòrics. Relació amb l'anàlisi factorial. Interpretació bàsicament en el cas de dos grups. Aplicacions a casos reals. Implicacions pràctiques

9 Anàlisi textual

Tècniques d'anàlisi textual de dades. Anàlisi de sentiments i topic modelling

Metodologia i activitats formatives

Teoria: s'hi presenten les diferents tècniques suficientment fonamentades i s'exemplifiquen amb casos reals.

Pràctiques: s'hi fan exercicis de dimensió real amb sistemes informàtics estadístics (R) que permetin aprofundir i consolidar els conceptes vistos a teoria; s'hi plantegen també treballs d'aplicació pràctica de més llarg termini on l'estudiant pugui entrenar la seva capacitat personal de dissenyar, planificar projectes on hagi d'aplicar les tècniques estudiades per analitzar un conjunt de dades real de certa envergadura, integrar els diferents coneixements adquirits en un informe ben argumentat, i mostrar les seves capacitats de síntesi i comunicació presentant el treball.

El grup de teoria es desdobra en dos grups de pràctiques, que fan classe de pràctiques dues hores per setmana cadascun.

Avaluació acreditativa dels aprenentatges

Avaluació continuada

La nota final (N) de l'assignatura s'obté a partir d'un examen parcial (NP), un examen final (EF) i una nota de pràctiques corresponent al treball en equip desenvolupat al llarg del curs (NTE) segons l'expressió:

$$N = NP \cdot 0.30 + NTE \cdot 0.30 + EF \cdot 0.40$$

Per aprovar l'assignatura amb avaluació continuada s'ha de treure un mínim de 3.5 en la nota de l'examen parcial (NP) i haver presentat tots els lliuraments parcials del treball en equip (NTE). Si no es lliuren la part pràctica, a l'avaluació continuada hi consta una qualificació de «no presentat».

La nota de pràctiques (NTE) s'obté amb el desenvolupament d'un treball pràctic de llarga durada que s'ha de fer en grup de màxim 5 estudiants i que ha d'integrar les diferents tècniques vistes al llarg del curs. Hi ha tres lliuraments parcials, amb una puntuació específica per a cadascun: 0,1, 0,45 i 0,45. La qualificació consta d'una nota global de la pràctica comuna a tot l'equip i de bonificacions o penalitzacions individuals segons com respongui cada estudiant a les preguntes el dia de la presentació.

Avaluació única

L'avaluació única (sense continuïtat) consisteix en un examen final, que té una part teòrica (EF) i una part pràctica (PR). La part teòrica és similars als exàmens NP i EF de l'avaluació continua. En la part pràctica, l'estudiant haurà d'executar rutines en R sobre els tòpics de l'assignatura i hauran de respondre preguntes en base a les dades analitzades i facilitades en aquell moment.

La nota final obtenida per aquesta avaluació serà:

$$N = EF * 0.50 + PR * 0.50$$

L'alumnat que no superi l'assignatura té dret a una prova de reavaluació (**amb les mateixes característiques que la prova d'avaluació única**), que té lloc durant el mes de juliol en la data fixada pel Consell Docent. A aquesta prova s'hi pot presentar qualsevol estudiant, independentment del pla d'avaluació escollit durant el curs. La prova dona accés a la qualificació màxima.

Fonts d'informació bàsica

Libre

ALUJA, Tomàs, et al. *Aprender de los datos: el análisis de componentes principales: una aproximación desde el Data Mining*. Barcelona: EUB, 1999

Comentari: Bibliografia bàsica.

Enllaços: [Catàleg UB](#)

Libre

ESCOFIER, Brigitte, et al. *Análisis factoriales simples y múltiples: objetivos, métodos e interpretación*. Bilbao: Servicio Editorial. Universidad del País Vasco, 1992

Comentari: Bibliografia bàsica.

Enllaços: [Catàleg UB](#)

Libre

GREENACRE, Michael J. *Correspondence analysis in practice*. Boca Raton (Fla.) [etc.]: Chapman & Hall/CRC, 2007

Comentari: Bibliografia bàsica.

Enllaços: [Catàleg UB](#) [Edició en castellà de 2008](#)

Libre

HUSSON, François, et al. *Exploratory multivariate analysis by example using R*. Boca Raton: CRC Press, 2011

Comentari: Bibliografia bàsica.

Enllaços: [Catàleg UB](#)

Libre

JOHNSON, Richard Arnold, et al. *Applied multivariate statistical analysis*. 6th ed. Upper Saddle River, N.J.: Pearson Education, Prentice-Hall, 2007

Comentari: Bibliografia bàsica.

Enllaços: [Catàleg UB](#)

Libre

BOUROCHE, Jean-M et al. *L'analyse des données*. Paris: Presses Universitaire de France, 1980

Comentari: Bibliografia complementària.

Enllaços: [Catàleg UB](#)

Libre

JOBSON, J.D. *Applied multivariate data analysis*. Vol. I y Vol. II. New York; Barcelona [etc.]: Springer, 1992

Comentari: Bibliografia complementària.

Enllaços: [Catàleg UB](#)

Libre

LEBART, Ludovic, et al. *Tratamiento estadístico de datos: métodos y programas*. Barcelona [etc.]: Marcombo, 1985

Comentari: Bibliografia complementària.

Enllaços: [Catàleg UB](#)

Libre

SAPORTA, Gilbert. *Probabilités, analyse des données et statistique*. 3e éd. rév. Paris: Technip, 2011.

Comentari: Bibliografia complementària.

Enllaços: [Catàleg UB](#)

Libre

VOLLE, Michel. *Analyse des données*. 4e éd. Paris: Economica, 1985

Comentari: Bibliografia complementària.

Enllaços: [Catàleg UB](#)