

Guia docent de l'assignatura "Anàlisi multivariant"

2011/2012

Codi: 100122

Crèdits ECTS: 6

Titulació	Pla	Tipus	Curs	Semestre
2501919 Estadística Aplicada	973 Graduat en Estadística Aplicada	OB	3	1

Contacte

Nom : Mercè Farré Cervelló

Email : Merce.Farre@uab.cat

Utilització d'idiomes

Llengua vehicular majoritària: català (cat)

Algun grup íntegre en anglès: No

Algun grup íntegre en català: Sí

Algun grup íntegre en espanyol: No

Prerequisits

És imprescindible un curs d'àlgebra lineal així com probabilitats i inferència estadística.

Objectius i contextualització

La necessitat de processar gran quantitat de dades amb moltes variables de distinta naturalesa, i d'obtenir resultats de relació i classificació de casos (gens, píxels, productes, clients, països, etc.), han fet desenvolupar recentment un gran nombre de procediments que se situen en el camp multivariant.

L'objecte d'estudi d'aquesta assignatura són les tècniques i mètodes multivariants clàssics: factorials, discriminants, *clusters*, *MANOVA*, etc. Algunes d'aquestes eines es poden presentar des del punt de vista de l'aprenentatge automàtic i la mineria de dades (*data mining*) on conviuen amb arbres de regressió, xarxes neuronals, per exemple, tòpics que l'estudiant del Grau d'Estadística Aplicada veurà en altres assignatures.

L'enfoc de l'assignatura és eminentment aplicat. Les classes de teoria, en les quals els mètodes es presenten amb la major generalitat i rigor, es combinen amb sessions i treballs pràctics, on es treballa amb bases de dades d'interès.

La majoria del mètodes multivariants tenen una vessant descriptiva o exploratòria, per la qual no es requereixen hipòtesis sobre el model, i una part confirmatòria amb supòsits més restrictius. Es procurarà distingir sempre ambdues vessants.

Competències i resultats d'aprenentatge

2273:E04 - Resumir i descobrir patrons de comportament en l'exploració de les dades.

2273:E04.29 - Identificar les etapes en els problemes que requereixen tecnologies avançades.

2273:E04.30 - Emprar índexs de resum de dades multivariants, sèries temporals i totes les altres tècniques avançades.

2273:E04.31 - Emprar gràfics de resum de dades multivariades i d'evolució temporal.

2273:E04.33 - Caracteritzar grups homogenis d'individus mitjançant l'anàlisi multivariant.

2273:E05 - Dissenyar un estudi estadístic o de recerca operativa per a la resolució d'un problema real.

2273:E05.13 - Projectar un estudi a partir de metodologies multivariants i/o una mineria de dades per

resoldre un problema contextualitzat en la realitat experimental.

2273:E09 - Identificar els models estadístics i de recerca operativa més adequats per a cada context i que permetin la presa de decisions.

2273:E09.13 - Reconèixer la necessitat d'emprar mètodes multivariants en lloc de bivariants.

2273:E11 - Utilitzar correctament una bona part del programari estadístic i de recerca operativa existent, escollir el més apropiat per a cada anàlisi estadística i ser capaç d'adaptar-lo a les noves necessitats.

2273:E11.22 - Utilitzar programari estadístic per als diversos mètodes d'anàlisi multivariant.

2273:T07 - Demostrar que es té un pensament lògic, un raonament estructurat i capacitat de síntesi.

2273:T07.00 - Demostrar que es té un pensament lògic, un raonament estructurat i capacitat de síntesi.

Continguts

CONTINGUTS

Introducció

- Diverses classificacions dels mètodes multivariants.
- Presentació de mètodes multivariants en base a exemples d'aplicació.
- Descripció dels objectius i diverses classificacions dels mètodes.
- Elements bàsics de càlcul matricial. Les descomposicions espectral i en valors singulars (SVD).
- Vectors aleatoris. La normal multidimensional. Mostres multivariants.
- Inferència: alguns estadístics per al vector de mitjanes i les matrius de covariàncies i de correlacions.

Els mètodes factorials I: anàlisi de components principals (ACP).

- Objectiu: reducció de la dimensió i representació canònica de les dades.
- L'Anàlisi de components principals: el concepte.
- El resultat fonamental. Matriu de covariàncies o de correlacions.
- Representació canònica de les variables. Estandarditzacions.
- Aplicacions i relació amb altres mètodes.

Els mètodes factorials II: anàlisi factorial (AF).

- El model: *comunalitats* i *especificitats*.
- AF exploratòria i confirmatòria.
- Rotacions.
- Interpretació dels resultats. Aplicacions.

Els mètodes factorials III: anàlisi de correspondències .

- El cas simple: perfils i inèrcies.
- Distància *khi-quadrat*.
- Descomposicions de la inèrcia.
- Representacions gràfiques (*biplots*): canònica i simultània.
- L'Anàlisi de correspondències múltiple. Condensació de la informació.
- Anàlisi de la taula disjuntiva completa. Anàlisi d'una taula de *Burt*.
- Discussió.

Els L'anàlisi de conglomerats, l'escalament multidimensional i altres eines de classificació no-supervisada.

- Proximitats i disimilaritats: entre objectes i entre grups.
- Classificació jeràrquica.
- Classificació no-jeràrquica: el mètode de les *K-mitjanes*.
- Altres mètodes de classificació.

- Escalament multidimensional.

L'anàlisi de dades agrupades: Anàlisi Multivariant de la Variància (MANOVA) i Anàlisi Discriminant.

- El MANOVA com a una generalització del ANOVA.
- Discriminació quan les lleis són conegudes. Criteri del màxim de versemblança i criteri de Bayes.
- Discriminació en un model normal.
- Discriminació no-paramètrica.

Complements: breu presentació.

- Regressió múltiple i correlació canònica.
- Arbres de decisió.
- L'anàlisi conjunta.
- L'anàlisi estructural.
- Relacions entre les diverses metodologies multivariants.

Metodologia

Classes de teoria, on s'exposaran detalladament els continguts sempre mostrant exemples d'aplicació convenientment interpretats. Els resums dels apunts es penjaran al campus virtual.

Classes pràctiques amb software estadístic, R preferentment. Els guions de les pràctiques i altre material auxiliar es penjarà al Campus Virtual. A les classes pràctiques es proposaran exercicis d'ampliació que caldrà resoldre autònomament, per equips.

A les tutories s'assessorarà els alumnes en: resolució de dubtes, recerca d'informació i de dades i plantejament i seguiment del treball.

En el treball autònom, l'alumne haurà de: estudiar la matèria, realitzar les tasques anteriorment descrites, relacionades amb el treball i els exercicis d'ampliació i també preparar els exàmens i altres activitats d'avaluació.

Activitats formatives

Activitat	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Classes de pràctiques	28	1.12	2273:E04.30 , 2273:E04.31 , 2273:E04.33 , 2273:E11.22 , 2273:E09.13
Classes de teoria	28	1.12	2273:E04.29 , 2273:E04.31 , 2273:T07.00 , 2273:E04.33 , 2273:E04.30
Tutories	10	0.4	2273:E04.29 , 2273:E11.22
Tipus: Supervisades			
Exercicis d'ampliació	8	0.32	2273:E04.31 , 2273:E11.22
Tipus: Autònomes			
Estudi	25	1.0	2273:E04.29 , 2273:E04.30 , 2273:E04.33 , 2273:E11.22 , 2273:E09.13 , 2273:E04.31
Treball escrit amb presentació oral	30	1.2	2273:E04.29 , 2273:E04.33 , 2273:E09.13 , 2273:T07.00 , 2273:E11.22 , 2273:E05.13 , 2273:E04.31 , 2273:E04.30

Avaluació

Les 10 hores d' "avaluació del treball escrit i presentació oral" són per a la redacció final del treball escrit i per a la preparació i realització de la presentació oral. La realització en sí del treball ja computa a l'apartat de metodologia.

Les 5 hores d' "Assistència a pràctiques i exercicis d'aplicació" no es refereixen a l'assistència a pràctiques (ja computada a l'apartat de metodologia) sinó a la presentació escrita dels exercicis complementaris.

Activitats d'avaluació

Activitat	Pes	Hores	ECTS	Resultats d'aprenentatge
Assistència a pràctiques i exercicis d'aplicació	15	5	0.2	2273:E04.30 , 2273:E04.31 , 2273:E04.33 , 2273:E11.22
Examen de teoria i de pràctiques	60%	6	0.24	2273:E04.30 , 2273:E04.31 , 2273:E04.33 , 2273:E11.22 , 2273:T07.00 , 2273:E09.13
Treball escrit i presentació oral	25%	10	0.4	2273:E04.29 , 2273:E05.13 , 2273:E11.22 , 2273:T07.00 , 2273:E09.13 , 2273:E04.33 , 2273:E04.30 , 2273:E04.31

Bibliografia

BIBLIOGRAFIA Bàsica

- Apunts de l'assignatura: resum de la teoria. Es penjaran al Campus Virtual, amb decalatge, un cop impartits els temes a classe.
- B. Everitt, *An R and S-PLUS Companion to Multivariate Analysis*. Springer. Llibre específic per a les pràctiques que conté breus resums de teoria.
- J. Friedman, T. Hastie, R. Tibshirani, *The elements of Statistical Learning*. 2nd edition, 2008.
- M. Greenacre, *La practica del análisis de correspondencias*. Fundación BBVA. 2008.
- J. D. Jobson, *Applied Multivariate Analysis*, Vols I i II. Springer, 1992.

BIBLIOGRAFIA COMPLEMENTÀRIA

- C. Chatfield, A.J. Collins (1980), *Introduction to Multivariate Analysis*. Chapman-Hall.
- W. J. Krzanowski (1988), *Principles of Multivariate Analysis*. Oxford.
- K. V. Mardia, J. T. Kent, J. M. Bibby (1979), *Multivariate Analysis*. Acad. Press.
- Alvin C. Rencher, *Methods of Multivariate Analysis*. Wiley Series in Probability and Mathematical Statistics.