

Anàlisi multivariant**2013/2014**

Codi: 100122

Crèdits: 6

Titulació	Tipus	Curs	Semestre
2500149 Matemàtiques	OT	4	0

Professor de contacte

Nom: Josep Solé Clivillés

Correu electrònic: JosepLluis.Sole@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

La majoria de les dades recollides en diferents disciplines són multivariants, és a dir per a la mateixa unitat experimental observem simultaneament els valors de varies variables. L'Anàlisi Multivariant estudia els mètodes per a descriure i explorar aquestes dades vectorials, així com per fer-hi una inferència estadística.

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**Matemàtiques**

- Demostrar de forma activa una elevada preocupació per la qualitat en el moment d'argumentar o exposar les conclusions dels seus treballs
- Formular hipòtesis i imaginar estratègies per confirmar-les o refutar-les.
- Hauran desenvolupat aquelles habilitats d'aprenentatge necessàries per emprendre estudis posteriors amb un alt grau d'autonomia
- Seran capaços de transmetre coneixements, procediments, resultats i idees matemàtiques.
- Utilitzar aplicacions informàtiques d'anàlisi estadística, càlcul numèric i simbòlic, visualització gràfica,

- optimització o altres per experimentar en Matemàtiques i resoldre problemes
- Utilitzar eficaçment bibliografia i recursos electrònics per obtenir informació

Resultats d'aprenentatge

1. Análisis de datos.
2. Capacidad de elaboración y construcción de modelos y su validación.
3. Capacidad para detectar y modelizar el azar en problemas reales.
4. Diseño, programación e implantación de paquetes estadísticos.
5. Extracción de conclusiones y redacción de informes.
6. Identificación de la información relevante para resolver un problema.
7. Identificación de relaciones o asociaciones.
8. Interpretación de resultados a partir de modelos estadísticos.
9. Pensamiento y razonamiento cuantitativo.
10. Representación gráfica de datos.
11. Utilización correcta y racional del software.

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.

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

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Classes de pràctiques	28	1,12	1, 6, 10, 11
Classes de teoria	28	1,12	1, 2, 3, 6, 9
Tutories	10	0,4	2, 6, 7, 9
Tipus: Supervisades			
Exercicis d'ampliació	8	0,32	1, 4, 10, 11
Tipus: Autònomes			
Estudi	25	1	1, 2, 3, 6, 7
Treball escrit amb presentació oral	30	1,2	1, 2, 3, 4, 6, 7, 10, 11

Avaluació

La nota de curs (NC) serà la mitjana ponderada de la nota de presentació del treball (NT), la de les entregues de pràctiques i exercicis (NPE) i la dels diferents examens parcials (NP). Els pesos respectius seran 0.2, 0.15 i 0.65.

Si la nota de curs és superior a 5.5 es considerarà com la nota final.

L'examen final (NEF) s'adreça a aquells alumnes que tenint una nota de curs superior a 5.5 la vulguin millorar, i a aquells que no arribin a 5.5. Per tant es considera un examen de recuperació. La nota final serà el màxim entre els valors $0.2NT+0.8NEF$ i $0.2NT+0.4NC+0.4NEF$

La presentació del treball és obligatoria. L'examen final, que tindrà una part pràctica, és també obligatori, per a validar les seves entregues, per als estudiants no presencials.

Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Entrega de pràctiques i exercicis d'aplicació	15%	5	0,2	4, 5, 7, 8, 10
Examens de teoria i de pràctiques	65%	6	0,24	1, 2, 3, 7, 9, 10, 11
Treball escrit i presentació oral	20%	10	0,4	1, 3, 4, 6, 7, 9, 10, 11

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, T. Hothorn, An Introduction to Applied Multivariate Analysis with R. Springer. 2011. 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 práctica del análisis de correspondencias. Fundación BBVA. 2008.
- J. D. Jobson, Applied Multivariate Analysis, Vols I i II. Springer, 1992.

BIBLIOGRAFIA COMPLEMENTÀRIA

- R. A. Johnson, D. W. Wichern, Applied Multivariate Statistical Analysis. Pearson Education International. 2007
- W. Härdle, L. Simar, Applied Multivariate Statistical Analysis. Springer. 2007
- K. V. Mardia, J. T. Kent, J. M. Bibby, Multivariate Analysis. Acad. Press. 2003.
- A. C. Rencher, Methods of Multivariate Analysis. Wiley Series in Probability and Mathematical Statistics. 2002.