

**Anàlisi Multivariant****2015/2016**

Codi: 100122

Crèdits: 6

| Titulació                    | Tipus | Curs | Semestre |
|------------------------------|-------|------|----------|
| 2500149 Matemàtiques         | OT    | 4    | 0        |
| 2501919 Estadística Aplicada | OB    | 3    | 1        |

**Professor de contacte**

Nom: Josep Solé Clivillés

Correu electrònic: JosepLluis.Sole@uab.cat

**Utilització de llengües**

Llengua vehicular majoritària: català (cat)

**Prerequisites**

És imprescindible un curs previ d'àlgebra lineal, així com el de probabilitats i inferència estadística.

**Objectius**

La majoria de dades recollides són multivariants, és a dir per a la mateixa unitat experimental observem de manera simultània els valors de varies variables. L'Anàlisi Multivariant estudia els mètodes apropiats per a descriure i explorar aquestes dades vectorials, així com per a aplicar-hi la inferència estadística.

La necessitat de processar gran quantitat de dades amb moltes variables de distinta naturalesa i d'obtenir relacions entre elles, i la necessitat de la classificació de casos han fet desenvolupar recentment un gran nombre de procediments que se situen en l'escenari multivariant.

L'objectiu de l'assignatura és un primer contacte de l'estudiant amb aquest camp, i fer-li veure la potència i aplicabilitat d'alguns mètodes, els quals a vegades es basen en idees molt simples i naturals. Per a això l'enfoc de l'assignatura és eminentment aplicat, encara que a teoria els estudiarem amb el rigor i generalitat adequats.

**Competències****Matemàtiques**

- Davant de situacions reals amb un nivell mig de complexitat, demanar i analitzar dades i informació rellevants, proposar i validar models utilitzant eines matemàtiques adequades per a, finalment, obtenir conclusions
- Formular hipòtesis i imaginar estratègies per confirmar-les o refutar-les.
- Utilitzar aplicacions informàtiques d'anàlisi estadística, càlcul numèric i simbòlic, visualització gràfica, optimització o altres per experimentar en Matemàtiques i resoldre problemes

**Estadística Aplicada**

- Demostrar que es té un pensament lògic, un raonament estructurat i capacitat de síntesi.
- Dissenyar un estudi estadístic o de recerca operativa per a la resolució d'un problema real.
- Expressar i aplicar rigorosament els coneixements adquirits en la resolució de problemes
- Identificar els models estadístics i de recerca operativa més adequats per a cada context i que permetin la presa de decisions.
- Resumir i descobrir patrons de comportament en l'exploració de les dades.
- 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.

## Resultats d'aprenentatge

1. Caracteritzar grups homogenis d'individus mitjançant l'anàlisi multivariant.
2. Demostrar que es té un pensament lògic, un raonament estructurat i capacitat de síntesi.
3. Emprar gràfics de resum de dades multivariades i d'evolució temporal.
4. Emprar índexs de resum de dades multivariants, sèries temporals i totes les altres tècniques avançades.
5. Expressar i aplicar rigorosament els coneixements adquirits en la resolució de problemes.
6. Identificar les etapes en els problemes que requereixen tecnologies avançades.
7. Projectar un estudi a partir de metodologies multivariants i/o una mineria de dades per resoldre un problema contextualitzat en la realitat experimental.
8. Reconèixer la necessitat d'emprar mètodes multivariants en lloc de bivariants.
9. Utilitzar programari estadístic per als diversos mètodes d'anàlisi multivariant.

## Continguts

### Continguts

#### El model aleatori i les dades multivariants.

- Classificació dels mètodes multivariants. Exemples.
- Vectors aleatoris. El vector d'esperances i les matrius de covariàncies i correlacions. Propietats.
- Dades multivariants. El vector de mitjanes i les matrius de covariància i correlació mostrals. Les seves propietats com a estimadors del vector d'esperances i de la matriu de covariàncies.
- Elements bàsics del càlcul matricial. Les descomposicions espectral i en valors singulars.

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

- Introducció a l'ACP.
- El resultat fonamental. Diferències entre el tractament de la matriu de correlacions i la de covariàncies.
- Mètodes de decisió de les components a retenir.
- Representacions de les variables i de les dades. Estandarditzacions. Interpretació geomètrica de les CP.

#### Els mètodes factorials II: anàlisi factorial (AF).

- El model factorial: comunalitats i especificitats.
- Rotacions.
- Interpretació dels resultats.

#### Els mètodes factorials III: anàlisi de correspondències.

- Perfils i inèrcies.
- Distància khi-quadrat.
- Descomposicions de la inèrcia.
- Representacions gràfiques (biplots).

#### Elements d'inferència multivariant

- Alguns contrastos sobre la matriu de covariàncies.
- Contrastos sobre el vector d'esperances.

#### Anàlisi de conglomerats

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

## Anàlisi Multivariant de la Variància (MANOVA) i Anàlisi Discriminant (AD).

- El MANOVA com una generalització de l'ANOVA.
- AD en un model normal.
- AD no paramètric.
- Anàlisi de correlació canònica.

## 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. Els guions de les pràctiques i l'altre material auxiliar es penjaran al Campus Virtual. A les classes pràctiques es proposaran exercicis d'ampliació que caldrà resoldre autònomament, per equips.

## Activitats formatives

| Títol                                            | Hores | ECTS | Resultats d'aprenentatge  |
|--------------------------------------------------|-------|------|---------------------------|
| <b>Tipus: Dirigides</b>                          |       |      |                           |
| Classes de pràctiques                            | 28    | 1,12 | 1, 2, 3, 4, 6, 7, 8, 9    |
| Classes de teoria                                | 28    | 1,12 | 3, 4, 6, 8                |
| <b>Tipus: Supervisades</b>                       |       |      |                           |
| Estudi de casos                                  | 8     | 0,32 | 2, 3, 6, 8, 9             |
| <b>Tipus: Autònomes</b>                          |       |      |                           |
| Estudi                                           | 38    | 1,52 | 3, 4, 6, 8, 9             |
| Presentació d'informes de pràctiques i problemes | 30    | 1,2  | 1, 2, 3, 4, 5, 6, 7, 8, 9 |

## Avaluació

### Mètode d'avaluació.

La nota de curs (NC) es calcularà a partir de les entregues d'exercicis i pràctiques (10%), i de les qualificacions de proves parcials, tant de teoria (70%) com de pràctiques (20%). Excepte pels alumnes virtuals i d'altres amb causa justificada, l'assistència a pràctiques és obligatòria per a ser qualificat de l'assignatura.

Si NC és igual o superior a 5, no cal anar a l'examen final. La nota del curs serà NC.

L'examen final (que es pot considerar una activitat de recuperació) tindrà una part teòrica i una de pràctica. La nota de l'examen (F) serà la mitjana ponderada de les dues parts, amb pesos respectius 0.6 i 0.4. La nota final (NF) es calcularà amb la següent fórmula:  $NF = \text{Max}(F, 0.5NC + 0.5F)$ .

L'examen final és obligatori per a aquells alumnes virtuals pels quals no hagi tingut suficient contacte presencial per a veure com treballen.

### Activitats d'avaluació

| Títol                                           | Pes | Hores | ECTS | Resultats d'aprenentatge  |
|-------------------------------------------------|-----|-------|------|---------------------------|
| Entregues d'informes de pràctiques i problemes. | 10% | 10    | 0,4  | 1, 2, 3, 4, 5, 6, 7, 8, 9 |
| Exàmens parcials teòrics                        | 70% | 5     | 0,2  | 1, 2, 3, 4, 5, 8, 9       |
| Proves pràctiques                               | 20% | 3     | 0,12 | 1, 2, 3, 4, 6, 7, 8, 9    |

## Bibliografia

### Bibliografia bàsica

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).

W. Härdle, L. Simer. Applied Multivariate Statistical Analysis. Springer. 2007.

D. Peña. Análisis de Datos Multivariantes. McGraw Hill. 2002.

A. Rencher. Methods of Multivariate Analysis. Wiley Series in Probability and Mathematical Statistics. 2002.

Little book of R for Multivariate Analysis. <https://little-book-of-r-for-multivariate-analysis.r...> 2015. (web específica per a les pràctiques)

Notes de l'assignatura penjades al campus virtual.

### Bibliografia complementària.

B. Flury. A first course in Multivariate Analysis. Springer 1997.

J.D. Jobson. Applied Multivariate Analysis. Vol I i II. Springer 1992.

R. Johnson. D.W. Wichern. Applied Multivariate Statistical Analysis. Pearson Education International. 2007.

K.V. Mardia, J.T. Kent, J.M. Bibby. Multivariate Analysis. Academic Press. 2003.

D. Skillicorn. Understanding Complex Data. Data Mining with Matrix Decomposition. Chapman&Hall. 2007.