

Titulació	Tipus	Curs	Semestre
2501919 Estadística Aplicada	OB	2	2

Professor de contacte

Nom: Mercè Farré Cervelló

Correu electrònic: Merce.Farre@uab.cat

Utilització d'idiomes a l'assignatura

Llengua vehicular majoritària: català (cat)

Grup íntegre en anglès: No

Grup íntegre en català: Sí

Grup íntegre en espanyol: No

Prerequisits

Coneixements bàsics d'estadística Descriptiva i Inferencial.

Objectius

L'objectiu del curs és l'estudi de la modelització i anàlisi de problemes estadístics mitjançant la metodologia dels Models Lineals i les seves aplicacions a diversos àmbits (economia, salut, enginyeria, i ciències en general), basades en exemples, problemes i pràctiques amb el sistema estadístic R. En primer lloc es presenta el model lineal simple, per la seva simplicitat i importància i perquè és un bon pròleg per a la comprensió del model lineal múltiple. El model de regressió múltiple expressat matricialment, i amb les seves variants (polinomial, amb interaccions, utilitzant variables regressores fictícies,...), és la segona part del curs. Tant en la regressió simple com múltiple s'analitzen l'adequació del model, la satisfacció de les hipòtesis i la detecció de dades "especials" (anòmales i influents), i s'estudien possibles solucions quan es detecten anomalies.

Competències

- Analitzar les dades mitjançant l'aplicació de mètodes i tècniques estadístiques i treballar amb dades qualitatives i quantitatives.
- Demostrar que es té un pensament lògic, un raonament estructurat i capacitat de síntesi.
- Desenvolupar estratègies d'aprenentatge autònom.
- Desenvolupar un pensament i un raonament crítics i saber comunicar-los de manera efectiva, tant en les llengües pròpies com en una tercera llengua.
- 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.
- Identificar la utilitat i la potencialitat de l'estadística en les diferents àrees de coneixement i saber aplicar-la adequadament per a extreure'n conclusions rellevants.
- Interpretar resultats, extreure conclusions i elaborar informes tècnics.
- Reconèixer els avantatges i els inconvenients dels procediments estudiats.
- Reconèixer situacions complexes i dissenyar estratègies per a afrontar-les.
- 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. Analitzar dades mitjançant el model d'anàlisi de la covariància.
2. Analitzar dades mitjançant el model de regressió lineal.
3. Analitzar els residus d'un model estadístic.
4. Comparar el grau d'ajust entre diversos models estadístics.
5. Demostrar que es té un pensament lògic, un raonament estructurat i capacitat de síntesi.
6. Desenvolupar estratègies d'aprenentatge autònom.
7. Desenvolupar un pensament i un raonament crítics i saber comunicar-los de manera efectiva, tant en les llengües pròpies com en una tercera llengua.
8. Detectar i preveure interaccions entre variables explicatives.
9. Detectar i tractar la colinealitat entre variables explicatives.
10. Detectar i tractar la falta d'homogeneïtat en els grups d'una variable explicativa principal qualitativa.
11. Detectar i tractar les dades absents.
12. Emprar gràfics de visualització de l'ajust i de l'adequació del model.
13. Establir les hipòtesis experimentals de la modelització.
14. Expressar i aplicar rigorosament els coneixements adquirits en la resolució de problemes.
15. Extreure conclusions de l'adequació dels models amb la utilització i la interpretació correcta d'indicadors i gràfics.
16. Identificar distribucions de les respostes amb l'anàlisi de residus.
17. Identificar els biaixos més rellevants en l'obtenció de resultats d'un estudi en funció de l'àmbit d'utilització.
18. Identificar els criteris de compliment per a l'aplicació d'un procediment.
19. Identificar fonts de biaix en l'obtenció de la informació.
20. Identificar les etapes en els problemes de modelització.
21. Identificar les particularitats en la interpretació dels resultats segons els objectius de l'estudi: predir, comparar grups (valor causal), identificar factors rellevants.
22. Identificar les suposicions estadístiques associades a cada procediment.
23. Identificar les variables resposta, explicatives i de control.
24. Mesurar el grau d'ajust d'un model estadístic.
25. Mesurar el grau de compliment dels criteris d'aplicació d'un procediment.
26. Planificar l'anàlisi estadística.
27. Projectar el qüestionari de recollida de dades.
28. Reconèixer la necessitat d'utilitzar models per a errors no independents.
29. Reconèixer les tècniques estadístiques més comunes en les diferents àrees d'utilització.
30. Reconèixer situacions complexes i dissenyar estratègies per a afrontar-les.
31. Seleccionar el paquet o paquets més apropiats per a cada problema de modelització.
32. Seleccionar les variables explicatives rellevants.
33. Simular models estadístics utilitzant programari específic.
34. Sintetitzar i interpretar els resultats dels models lineals clàssics, generalitzats i no lineals en funció de l'objectiu de l'estudi.
35. Utilitzar programari estadístic divers per a la modelització lineal i no lineal.

Continguts

1. L'anàlisi de regressió. Regressió lineal simple

- Introducció als models de regressió i passos previs en la regressió simple: Exploració de les dades.
- La regressió lineal simple: Model, hipòtesis, paràmetres.
- Estimació puntual dels paràmetres del model: Mètode de mínims quadrats. Els estimadors de màxima versemblança.
- Inferència sobre els paràmetres del model sota les hipòtesis de Gauss-Markov: Interval i tests.
- Interval de confiança per a la resposta mitjana. Interval de predicció de noves observacions. Inferències simultànies en la regressió simple. Bandes de confiança i de predicció.

- Anàlisi de la variància (ANOVA) del model de regressió simple.
- Diagnòstics del model: Avaluació gràfica de la linealitat i de la verificació de les hipòtesis mitjançant l'anàlisi dels residus. Test de manca d'ajust lineal i contrastos de les hipòtesis de Gauss-Markov.
- Dades anòmales o influents.

2. Regressió lineal múltiple

- Passos previs en la regressió múltiple: Exploració de les dades amb eines de visualització multidimensional.
- Model i estimadors dels coeficients mitjançant mínims quadrats. Interpretació dels coeficients del model lineal.
- Lleis dels estimadors dels coeficients, de les prediccions i dels residus: aplicació de les propietats de les matrius idempotents.
- Inferència en el model lineal múltiple. Anova del model.
- Test de "lligadures" per resoldre restriccions lineals sobre els coeficients: El principi de la variabilitat incremental.
- Discussió de les hipòtesis del model lineal: Anàlisi dels residus. Transformacions de Box-Cox i altres.
- El problema de la multi-colinealitat entre variables regressores: Detecció i solucions.
- Variables fictícies en regressió (dummies): Interpretació dels coeficients i aplicacions.
- Selecció de variables en un model lineal: L'estadístic Cp de Mallows, la validació creuada de models i els mètodes per passos.

Metodologia

L'assignatura té, al llarg del quadrimestre, dues hores setmanals de teoria i problemes més dues hores setmanals de pràctiques.

Les llistes de problemes es penjaran al campus virtual, amb algunes indicacions.

Tot el temari és complementa amb pràctiques amb software, amb apartats més tutoritzats i altres en forma d'exercicis proposats a l'estudiant.

El material del curs (apunts de teoria, llistes de problemes i enunciats de pràctiques) està disponible al campus virtual, de manera progressiva al llarg del curs. L'alumne es manté informat setmanalment del seguiment de les classes amb notícies i avisos al campus virtual.

Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Pràctiques amb ordinador	50	2	2, 3, 8, 9, 10, 11, 12, 16, 20, 23, 24, 27, 28, 32, 33, 35
Sessions teòriques	50	2	2, 3, 8, 12, 20, 24, 32
Tipus: Supervisades			
pràctiques per resoldre	9	0,36	3, 8, 9, 10, 12, 32, 33

problemes per resoldre 7 0,28 2, 8, 14, 19, 22, 24, 28

Tipus: Autònomes

Realització de exàmens de cursos anteriors 12 0,48 3, 6, 8, 10, 11, 16, 24, 29, 33

Avaluació

PR: Lliurament de les pràctiques que s'indiquin al llarg del curs i dels informes (el previ i el definitiu) d'un treball amb dades pròpies que caldrà presentar i exposar públicament, dins dels terminis establerts. Valoració de **PR: 2 punts sobre 10**. Aquesta part NO és recuperable. La presentació del treball és imprescindible per aprovar.

P1: Prova parcial de regressió simple (teoria, problemes i pràctiques). Valoració de **P1: 3 punts sobre 10**.

P2: Prova parcial de regressió múltiple (teoria, problemes i pràctiques). Valoració de **P2: 5 punts sobre 10**.

Aleshores la Nota final serà: **Nota Final = PR+ P1 + P2**

En juny es té un dia assignat per fer una prova de síntesi, **PS**, que permet només la **recuperació de P1 i P2 (8 punts sobre 10)** pels alumnes que no hagin aprovat per curs o desitgin pujar la seva nota dels parcials.

Cas que es faci la prova de síntesi, la nota final serà: **Nota Final = PR+ PS.**

Atenció: "Sense perjudici d'altres mesures disciplinàries que s'estimin oportunes, i d'acord amb la normativa acadèmica vigent, es qualificaran amb un zero les irregularitats comeses per l'estudiant que puguin conduir a una variació de la qualificació d'un acte d'avaluació. Per tant, plagiar, copiar o deixar copiar una pràctica o qualsevol altra activitat d'avaluació implicarà suspendre-la amb un zero i no es podrà recuperar en el mateix curs acadèmic. Si aquesta activitat té una nota mínima associada, aleshores l'assignatura quedarà suspesa."

Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Lliurament d'exercicis, pràctiques i informes	0,64	16	0,64	1, 2, 3, 4, 7, 8, 9, 10, 12, 13, 14, 15, 16, 18, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 35
Prova de regressió múltiple	50%	3	0,12	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 22, 23, 24, 25, 28, 32, 33, 35
Prova de regressió simple	30%	3	0,12	2, 3, 5, 6, 10, 11, 12, 14, 16, 17, 18, 19, 20, 22, 24, 25, 28, 33, 35

Bibliografia

Bibliografia bàsica:

M. Farré; Apunts de Models Lineals, al Campus Virtual de la UAB, per parts.

Montgomery, D. Peck, A. Vining, G.; Introduction to Linear Regression Analysis. Wiley. 2001.

Clarke, B.R.; Linear Models. The Theory and Applications of Analysis of variance. Wiley.

Christopher Hay-Jahans; An R Companion to Linear Statistical Models. Chapman and Hall, 2012.

John Fox and Sandord Weisberg; An R Companion to Applied Regression, 2nd edition, Sage Publications, 2011.

Daniel Peña; Regresión y diseño de Experimentos, Alianza Editorial (Manuales de Ciencias Sociales), 2002.

Bibliografía complementaria:

A. Sen, M. Srivastava; Regression Analysis: Theory, Methods, and Applications, Ed Springer (Springer Texts in Statistics), 1990.

J. Neter, M. H. Kutner, C. J. Nachtsheim, W. Wasserman; Applied Linear Models, Ed Irwin (4th edition), 1996.

Faraway, J.; Linear Models with R. Chapman&Hall/CRC. 2005.

Rao, Toutenburg, Shalabh, Heumann; Linear Models and generalizations. Springer.