



Guia docent

270210 - PIE2 - Probabilitat i Estadística 2

Última modificació: 19/07/2023

Unitat responsable: Facultat d'Informàtica de Barcelona
Unitat que imparteix: 715 - EIO - Departament d'Estadística i Investigació Operativa.
Titulació: GRAU EN CIÈNCIA I ENGINYERIA DE DADES (Pla 2017). (Assignatura obligatòria).
Curs: 2023 **Crèdits ECTS:** 6.0 **Idiomes:** Català

PROFESSORAT

Professorat responsable: MARTA PÉREZ CASANY
Altres: Primer quadrimestre:
VÍCTOR PEÑA PIZARRO - 11, 13
MARTA PÉREZ CASANY - 11, 12, 13

CAPACITATS PRÈVIES

L'estudiant ha d'haver assolit de forma satisfactòria els continguts d'el'assignatura de PIE1 i de Càlcul.

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

CE3. Analitzar fenòmens complexos mitjançant la probabilitat i l'estadística, i plantejar models d'aquests tipus en situacions concretes. Formular i resoldre problemes d'optimització matemàtica.

Genèriques:

CG1. Concebre sistemes computacionals que integren dades de procedències i formes molt diverses, construeixen amb ells models matemàtics, raonen sobre aquests models i actuen en conseqüència, aprenent de l'experiència.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

Transversals:

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

Bàsiques:

CB1. Que els estudiants hagin demostrat posseir i comprendre coneixements en una àrea d'estudi que parteix de la base de l'educació secundària general, i se sol trobar a un nivell que, si bé es recolza en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda del seu camp d'estudi.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB5. Que els estudiants hagin desenvolupat aquelles habilitats d'aprenentatge necessàries per emprendre estudis posteriors amb un alt grau d'autonomia

METODOLOGIES DOCENTS

La meitat de les classes seran de teoria, expositives i participatives, i l'altra meitat seran pràctiques i es faran a un aula informàtica o bien en una aula normal amb el propi portàtil de l'estudiant. A l'acabar les classes pràctiques, es proposaran exercicis d'anàlisi de dades que els estudiants haurà de treballar de forma autònoma i es discutiran a la sessió següent.

OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

1. Aprendre a construir models estadístics per sintetitzar informació, explicar una variable resposta i fer prediccions.
2. Entendre els conceptes bàsics de l'estadística Bayesiana
3. Aprendre a utilitzar software estadístic per a la resolució de problemes d'anàlisi de dades
4. Aprendre a validar models estadístics, i a millorar els models quan aquests siguin inadequats per al problema
5. Aprendre a redactar informes presentant els resultats d'una anàlisi de dades
6. Entendre la diferència entre l'aproximació Bayesiana a l'estadística, i l'aproximació no Bayesiana
7. Saber identificar el model estadístic més adequat per a cada problema
8. Aprendre a interpretar un model ajustat
9. Entendre el concepte de la validació creuada, i el problema del sobre-ajust i el sota-ajust
10. Aprendre a fer servir un model per a fer prediccions
11. Entendre la diferència entre paràmetre i estimador, i a resoldre problemes d'inferència en el context dels models lineals i lineals generalitzats
12. Aprendre a incorporar variables explicatives categòriques als models lineals i models lineals generalitzats
13. Analitzar de forma autònoma i amb sentit crític dades de temes rellevants per la societat
14. Saber estimar paràmetres per intervals de confiança i interpretar-los.
15. Entendre la importància dels test d'hipòtesis. Conèixer com fer les comparacions més habituals i tècniques per a fer-ne de noves que se'ls hi puguin presentar.

HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup gran	30,0	20.00
Hores grup petit	30,0	20.00
Hores aprenentatge autònom	90,0	60.00

Dedicació total: 150 h

CONTINGUTS

Distribucions relacionades amb la Normal. Estimació per Intervals de confiança.

Descripció:

Distribucions Chiquadrat, t-d'Student i F-Fisher-Snedecor. Definició d'interval de confiança. IC per un valor esperat, per una variància, per una probabilitat i per la diferència de dos valors esperats i dues probabilitats. Quantitats pivotals.

Tests d'hipòtesis

Descripció:

Conceptes generals en l'entorn dels test d'hipòtesis. Comparacions d'una esperança i una variància amb un valor concret. Comparació de dos valors esperats, comparació de dues variàncies. Comparació d'una probabilitat amb un valor concret. Comparació de dues probabilitats.

Model lineal

Descripció:

Definició de model lineal normal. Estimació dels paràmetres Taula ANOVA i mesures de bondat d'ajust. Inferència sobre els paràmetres. Predicció. Validació del model. Selecció del model. Interpretació del model; Biaix, colinealitat i causalitat. Us de variables explicatives categòriques.



Model lineal generalitzat

Descripció:

Definició de model lineal generalitzat. Models per a resposta binària. Estimació dels paràmetres. Inferència sobre els paràmetres. Validació del model. Selecció del model. Predicció. Interpretació del model.

Introducció a l'estadística Bayessiana

Descripció:

Teorema de Bayes. Model Bayessià. Distribució predictiva a priori i a posteriori. Elecció de la distribució a priori.

ACTIVITATS

Models lineals

Descripció:

Definició de model lineal normal. Estimació. Inferència. Predicció. Validació. Selecció de model. Interpretació. Ús de variables explicatives categòriques. Model no lineal normal.

Objectius específics:

1, 3, 4, 5, 7, 8, 10, 11, 12, 13

Competències relacionades:

CG1. Concebre sistemes computacionals que integren dades de procedències i formes molt diverses, construeixen amb ells models matemàtics, raonen sobre aquests models i actuen en conseqüència, aprenent de l'experiència.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CE3. Analitzar fenòmens complexos mitjançant la probabilitat i l'estadística, i plantejar models d'aquests tipus en situacions concretes. Formular i resoldre problemes d'optimització matemàtica.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB1. Que els estudiants hagin demostrat posseir i comprendre coneixements en una àrea d'estudi que parteix de la base de l'educació secundària general, i se sol trobar a un nivell que, si bé es recolza en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda del seu camp d'estudi.

CB5. Que els estudiants hagin desenvolupat aquelles habilitats d'aprenentatge necessàries per emprendre estudis posteriors amb un alt grau d'autonomia

Dedicació: 59h

Grup gran/Teoria: 12h

Grup petit/Laboratori: 12h

Aprenentatge autònom: 35h



Model lineal generalitzat

Descripció:

Definició del model. Model per resposta binària. Estimació. Inferència. Validació. Predicció. Selecció del Model. Interpretació.

Objectius específics:

1, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13

Competències relacionades:

CG1. Concebre sistemes computacionals que integren dades de procedències i formes molt diverses, construeixen amb ells models matemàtics, raonen sobre aquests models i actuen en conseqüència, aprenent de l'experiència.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CE3. Analitzar fenòmens complexos mitjançant la probabilitat i l'estadística, i plantejar models d'aquests tipus en situacions concretes. Formular i resoldre problemes d'optimització matemàtica.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB1. Que els estudiants hagin demostrat posseir i comprendre coneixements en una àrea d'estudi que parteix de la base de l'educació secundària general, i se sol trobar a un nivell que, si bé es recolza en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda del seu camp d'estudi.

CB5. Que els estudiants hagin desenvolupat aquelles habilitats d'aprenentatge necessàries per emprendre estudis posteriors amb un alt grau d'autonomia

Dedicació: 27h

Grup gran/Teoria: 6h

Grup petit/Laboratori: 6h

Aprenentatge autònom: 15h

Estadística Bayesiana

Descripció:

Model estadístic. Inferència basada en versemblança. Model Bayesia. Distribució a posteriori. Distribució predictiva a priori i a posteriori. Elecció de priori. Inferència Bayesiana. Validació del model.

Objectius específics:

2, 6

Competències relacionades:

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CE3. Analitzar fenòmens complexos mitjançant la probabilitat i l'estadística, i plantejar models d'aquests tipus en situacions concretes. Formular i resoldre problemes d'optimització matemàtica.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CB1. Que els estudiants hagin demostrat posseir i comprendre coneixements en una àrea d'estudi que parteix de la base de l'educació secundària general, i se sol trobar a un nivell que, si bé es recolza en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda del seu camp d'estudi.

CB5. Que els estudiants hagin desenvolupat aquelles habilitats d'aprenentatge necessàries per emprendre estudis posteriors amb un alt grau d'autonomia

Dedicació: 14h

Grup gran/Teoria: 2h

Grup petit/Laboratori: 2h

Aprenentatge autònom: 10h

Distribucions relacionades amb la Normal. Estimació per Interval de confiança.

Descripció:

Es presenten les distribucions chi-quadrat. T-d'student i F-Fisher-Snedecor. Es defineix interval de confiança i s'ensenya com calcular-los a partir de quantitats pivotals. Es calculen IC per un valor esperat, per una variància, per una probabilitat i per la diferència de dos valors esprats i dues probabilitats.

Objectius específics:

13, 14

Competències relacionades:

CG1. Concebre sistemes computacionals que integren dades de procedències i formes molt diverses, construeixen amb ells models matemàtics, raonen sobre aquests models i actuen en conseqüència, aprenent de l'experiència.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB5. Que els estudiants hagin desenvolupat aquelles habilitats d'aprenentatge necessàries per emprendre estudis posteriors amb un alt grau d'autonomia

Dedicació: 23h

Grup gran/Teoria: 4h

Grup petit/Laboratori: 4h

Aprenentatge autònom: 15h

Test d'hipòtesis

Descripció:

Es presenten els conceptes bàsics a l'entorn del test d'hipòtesis. comparació d'una esperança i una variància amb un valor concret. Comparació d'un probabilitat amb un valor concret. Comparació de dos valors esperats. Comparació de dues variàncies. Comparació de dues probabilitats.

Objectius específics:

5, 13, 15

Competències relacionades:

CG1. Concebre sistemes computacionals que integren dades de procedències i formes molt diverses, construeixen amb ells models matemàtics, raonen sobre aquests models i actuen en conseqüència, aprenent de l'experiència.

CG2. Elegir i aplicar els mètodes i tècniques més adequats a un problema definit per dades que representin un repte pel seu volum, velocitat, varietat o heterogeneïtat, inclosos mètodes informàtics, matemàtics, estadístics i de processament del senyal.

CE3. Analitzar fenòmens complexos mitjançant la probabilitat i l'estadística, i plantejar models d'aquests tipus en situacions concretes. Formular i resoldre problemes d'optimització matemàtica.

CT6. Aprenentatge autònom. Detectar deficiències en el propi coneixement i superar-les mitjançant la reflexió crítica i l'elecció de la millor actuació per ampliar aquest coneixement.

CT5. Ús solvent dels recursos d'informació. Gestionar l'adquisició, l'estructuració, l'anàlisi i la visualització de dades i informació en l'àmbit de l'especialitat i valorar de forma crítica els resultats d'aquesta gestió.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB5. Que els estudiants hagin desenvolupat aquelles habilitats d'aprenentatge necessàries per emprendre estudis posteriors amb un alt grau d'autonomia

Dedicació: 27h

Grup gran/Teoria: 6h

Grup petit/Laboratori: 6h

Aprenentatge autònom: 15h

SISTEMA DE QUALIFICACIÓ

Hi haurà un examen parcial i un examen final, a més a més d'exercicis d'anàlisi de dades que es realitzaran al llarg del curs

El parcial correspondrà als dos primers temes.

El Final correspondrà a la resta de continguts de l'assignatura

La nota de curs serà la mitjana aritmètica dels exercicis puntuables realitzats al llarg del curs.

La nota final de l'assignatura es calcularà de la forma següent:.

$$\text{Nota Assignatura} = 0,25 \cdot \text{NCurs} + 0,25 \cdot \text{Parcial} + 0,5 \cdot \text{Final}$$

En el cas dels estudiants que vagin a la reavaluació, la nota de l'assignatura es calcularà de la forma següent:

$$\text{Nota Assignatura} = \max(\text{NRevaluació}, 0,25 \cdot \text{NCurs} + 0,75 \cdot \text{NRevaluació})$$

BIBLIOGRAFIA

Bàsica:

- James, G.; Witten, D.; Hastie, T.; Tibshirani, R. An introduction to statistical learning. Springer, 2013. ISBN 97-1461471370.
- Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari, A.; Rubin, D.. Bayesian data analysis. 3rd ed. Chapman & Hall, 2014. ISBN 978-1439840955.
- Weisberg, S. Applied linear regression. 4th ed. John Wiley and Sons, 2014. ISBN 9780471704096.
- Dobson, A.J.; Barnett, A.G. An introduction to generalized linear models. 4th ed. Chapman & Hall, 2018. ISBN 978-1138741515.
- Dobrow, R.P. Introduction to stochastic processes with R. John Wiley and Sons, 2016. ISBN 978-1118740651.