

Modelització i Inferència

Codi: 104392
Crèdits: 6

Titulació	Tipus	Curs	Semestre
2503740 Matemàtica Computacional i Analítica de Dades	OB	2	1

La metodologia docent i l'avaluació proposades a la guia poden experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

Professor/a de contacte

Nom: Amanda Fernandez Fontelo
Correu electrònic: Amanda.Fernandez@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

Es considera molt important un bon coneixement dels continguts de les assignatures cursades durant el primer curs, especialment les de Probabilitat i Càlcul.

Objectius

Aquesta assignatura és la primera del Grau dedicada a la Inferència Estadística, que és la part de l'Estadística que permet obtenir, de manera controlada, informació sobre una població a partir de les dades d'una mostra "representativa". L'assignatura té un caràcter central dins dels estudis, ja que s'hi presenten conceptes i tècniques que seran emprades en moltes de les matèries que es cursaran a partir d'ara. Concretament, es començarà fent una introducció a l'Estadística, i després es tractarà l'estimació de paràmetres, tant puntual com per intervals de confiança, així com els tests d'hipòtesis paramètrics clàssics per a una i dues poblacions normals i dicotòmiques, i contrastos d'independència. Finalment, s'introduirà el model de regressió lineal simple.

La presencialitat de la docència i de les activitats avaluable s'adaptarà seguint les recomanacions de les autoritats sanitàries, per tal de garantir la seguretat de totes les persones.

Competències

- Avaluar de manera crítica i amb criteris qualitat el treball realitzat.
- Calcular i reproduir determinades rutines i processos matemàtics amb agilitat.
- Formular hipòtesis i imaginar estratègies per confirmar-les o refutar-les.
- Que els estudiants hagin demostrat que comprenen i tenen 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 basa en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda d'aquell camp d'estudi.
- Que els estudiants sàpiguin aplicar els coneixements propis a la seva feina o vocació d'una manera professional i tinguin les competències que se solen demostrar per mitjà de l'elaboració i la defensa d'arguments i la resolució de problemes dins de la seva àrea d'estudi.

- Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins de la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes destacats d'indole social, científica o ètica.
- Relacionar objectes matemàtics nous amb altres de coneguts i deduir-ne les propietats.
- Utilitzar aplicacions informàtiques d'anàlisi estadística, càlcul numèric i simbòlic, visualització gràfica, optimització o altres per experimentar i resoldre problemes.

Resultats d'aprenentatge

1. Analitzar dades mitjançant tècniques d'inferència per a una o dues mostres.
2. Avaluar de manera crítica i amb criteris de qualitat el treball desenvolupat.
3. Conèixer els diferents mètodes de recollida de dades.
4. Descriure les propietats bàsiques dels estimadors puntuals i per interval.
5. Escollir el programari estadístic apropiat per analitzar les dades mitjançant tècniques d'inferència.
6. Identificar distribucions estadístiques.
7. Identificar la inferència estadística com a instrument de pronòstic i predicció.
8. Interpretar els resultats obtinguts i concloure respecte a la hipòtesi experimental.
9. Que els estudiants hagin demostrat que comprenen i tenen 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 basa en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda d'aquell camp d'estudi.
10. Que els estudiants sàpiguin aplicar els coneixements propis a la seva feina o vocació d'una manera professional i tinguin les competències que se solen demostrar per mitjà de l'elaboració i la defensa d'arguments i la resolució de problemes dins de la seva àrea d'estudi.
11. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins de la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes destacats d'indole social, científica o ètica.
12. Reconèixer la utilitat dels mètodes bayesians i aplicar-los oportunament.
13. Utilitzar les propietats de les funcions de densitat.
14. Utilitzar les propietats de les funcions de distribució.
15. Utilitzar programari estadístic per gestionar bases de dades.
16. Utilitzar programari estadístic per obtenir índexs de resum de les variables de l'estudi.
17. Validar i gestionar la informació per fer-ne el tractament estadístic.

Continguts

Preliminars de Probabilitat (recordatori): Probabilitat i variables aleatòries. Concepte de llei. Distribucions discretes. Funció de densitat i de probabilitat. Esperança i variància. Funció generatriu de moments. Exemples.

Tema 1. Introducció a la Estadística.

1. Estadística descriptiva i estadística inferencial.

1.1. Conceptes bàsics en inferència: població estadística i mostra; paràmetres, estadístics i estimadors.

1.2. Models estadístics: paramètrics i no paramètrics.

2. Estadístics més usats: els moments mostrals. Els estadístics d'ordre.

3. Distribució d'alguns estadístics.

3.1. D'una mostra d'una població Normal: Teorema de Fisher.

3.2. El Teorema Central del Límit: normalitat asimptòtica dels moments mostrals i de la proporció

Tema 2: Estimació puntual.

1. Estimadors puntuals: definició i propietats.

1.1. Biaix.

1.2. Comparació d'estimadors sense biaix. Eficiència relativa.

1.3. Comparació d'estimadors amb biaix: l'error quadràtic mitjà.

1.4. Consistència d'un estimador.

1.5. Estadístic suficients.

2. Mètodes per a obtenir estimadors.

2.1. Mètode dels moments.

2.2. Mètode de la Màxima Versemblança (EMV)

2.2.1. Invariància dels EMV.

2.2.2. Funció Score i Informació de Fisher.

2.2.3. Cota de de Cramér-Rao.

2.2.4. Propietats dels EMV.

2.2.5. Mètode Delta.

2.2.6. Procediments numèrics per determinar els EMV.

Tema 3. Estimació per intervals de confiança.

1. Concepte de regió i interval de confiança.

2. El mètode del "pivot" per a la construcció d'intervals de confiança.

3. Intervals de confiança per als paràmetres d'una població.

3.1. Per a la mitjana d'una població Normal amb desviació coneguda i desconeguda.

3.2. Per a la variància d'una població Normal amb mitjana desconeguda i coneguda.

3.3. Intervals de confiança asimptòtics: Wald, Score i LRT.

4. Intervals de confiança per als paràmetres de dues poblacions.

4.1. Intervals de confiança amb mostres independents.

4.2. Intervals de confiança per a la diferència de mitjanes de dues poblacions Normals amb dades aparellades.

5. Tècniques Bootstrap.

Tema 4: Tests d'hipòtesis.

1. Introducció.

1.1. Errors tipus I i II.

1.2. Funció potència.

1.3. Consistència dels tests.

1.4. p-valors.

1.5. Dualitat entre intervals de confiança i tests d'hipòtesis.

2. Tests per als paràmetres d'una població.

2.1. Per a la mitjana d'una població Normal amb desviació coneguda i desconeguda.

2.2. Tests asimptòtics per a la mitjana d'una població quan la mostra és gran.

2.3. Per a la variància d'una població Normal.

2.4. Tests asimptòtics: Wald, Score i LRT.

3. Tests per als paràmetres de dues poblacions.

3.1. Tests d'hipòtesis amb mostres independents.

3.2. Tests d'hipòtesis amb dades aparellades.

Tema 5. Regressió lineal simple.

1. Objectius del model.

2. Estimadors de mínims quadrats ordinaris (MQO).

3. Inferència amb el model de regressió simple.

4. Prediccions.

IMPORTANT: En la docència, la perspectiva de gènere implica revisar els biaixos androcèntrics i qüestionar els supòsits i estereotips de gènere ocults. Aquesta revisió comporta incloure als continguts de l'assignatura el coneixement produït per les dones científiques, sovint oblidades, procurant el reconeixement de lesseves aportacions, així com el de les seves obres a les referències bibliogràfiques. També es procurarà introduir a la part més pràctica de l'assignatura, l'anàlisi i comparació de dades estadístiques per sexe, comentant a l'aula les causes i els mecanismes socials i culturals que poden sustentar les desigualtats observades.

Metodologia

L'assignatura s'estructura a partir de classes de teoria, problemes i pràctiques.

A les classes de teoria anirem introduint els conceptes i tècniques que descriu el programa del curs. Donat el seu contingut estàndard d'un primer curs d'inferència estadística, es pot seguir fent us de la bibliografia bàsica recomanada. També s'anirà penjant al Campus Virtual el material corresponent a cada tema explicat a les classes presencials.

Les classes de problemes tenen per objectiu treballar i entendre els conceptes estadístics. Al Campus Virtual es penjaran les llistes de problemes i, quan ja s'hagin resolt a classe, també les solucions.

L'objectiu de les pràctiques és la utilització de programari estadístic R, per a obtenir i aclarar els resultats dels procediments que s'han introduït a les classes de teoria i problemes. Al Campus Virtual es penjarà l'enunciat de cada pràctica amb antelació.

IMPORTANT: Per a treballar més còmodament amb R, es recomana fer servir l'interface RStudio: és lliure, "open source" i funciona amb Windows, Mac i Linux. <https://www.rstudio.com/>

OBSERVACIÓ: La perspectiva de gènere en la docència va més enllà dels continguts de les assignatures, ja que també implica una revisió de les metodologies docents i de les interaccions entre l'alumnat i el professorat, tant a l'aula com a fora. En aquest sentit, les metodologies docents participatives, on es genera un entorn igualitari, menys jeràrquic a l'aula, evitant exemples estereotipats en gènere i vocabulari sexista, amb

l'objectiu de desenvolupar el raonament crític el respecte a la diversitat i pluralitat d'idees, persones i situacions, solen ser més favorables a la integració i plena participació de les alumnes a l'aula, i per això es procurarà la seva implementació efectiva en aquesta assignatura.

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, per a la complementació per part de l'alumnat de les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura/mòdul.

Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Classes de problemes	18	0,72	2, 3, 4, 6, 7, 8, 9, 10, 11, 13, 14
Classes de pràctiques	12	0,48	1, 2, 5, 7, 8, 9, 10, 11, 15, 16, 17
Classes de teoria	30	1,2	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 13, 14
Tipus: Autònomes			
Exàmens	15	0,6	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 13, 14
Resolució de problemes	25	1	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 13, 14
Resolució de pràctiques	20	0,8	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17

Avaluació

La nota d'avaluació continuada s'obtindrà a partir d'un control dels problemes que donarà una nota C, i a partir d'un control de les pràctiques de l'assignatura que donarà una nota P. La nota C té un pes del 20% i la nota P un pes del 30%. La nota de l'examen final E1 val un 50% de la nota final. Amb les notes C, P i E1 s'obté la nota de l'assignatura, N, de la manera següent:

$$N = 0.50 \times E1 + 0.20 \times C + 0.30 \times P$$

Recuperació i/o millora de la nota d'examen:

L'alumne supera l'assignatura si N és més gran o igual que 5 . En cas contrari, o bé si l'alumne vol millorar nota, hi ha una possibilitat de millorar la part de la nota de l'examen E1 mitjançant un examen de recuperació, la nota del qual serà E2. Així, a partir d'aquesta nota de recuperació s'obté la nota final de l'assignatura:

$$NF = 0.50 \times \max(E1, E2) + 0.20 \times C + 0.30 \times P$$

Observació 1: Les notes C i P d'avaluació continuada no son recuperables.

Observació 2: Es considera que l'alumne s'ha presentat a la convocatòria de l'assignatura si es presenta a qualsevol dels dos exàmens que donen lloc a les notes E1 o E2. En cas contrari, serà un No Presentat, encara que tingui alguna nota d'avaluació continuada (C i/o P) .

Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Entrega de problemes (C)	0,20	8	0,32	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14

Examen de pràctiques (P)	0,30	12	0,48	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17
Examen final / Recuperació (E)	0,50	10	0,4	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 13, 14

Bibliografia

Berger, R.L., Casella, G.: Statistical Inference. Duxury Advanced Series. 2002.

Daalgard, P.: Introductory Statistics with R. Springer. 2008.

Daniel, W.W.: Biostatistics. Wiley. 1974.

DeGroot, M. H.: Schervish, M.J. Probability and Statistics. Pearson Academic. 2010.

Peña, D.: Estadística. Fundamentos de estadística. Alianza Universidad. 2001.

R Tutorial. An introduction to Statistics. <https://cran.r-project.org/manuals.html>. juny 2019.

Silvey, S.D.: Statistical Inference. Chapman&Hall. 1975.

Held, Sabanes Bove (2013): Applied Statistical Inference: Likelihood and Bayes. Springer

Pawitan (2013): In all Likelihood: Statistical Modelling and Inference using Likelihood. Oxford University Press

Young, Smith (2005): Essentials of Statistical Inference. Cambridge University Press

Cox, D.R. and Hinkley, D.V. (1979). Theoretical Statistics. 1st Edition, Chapman and Hall/CRC

Programari

The R Project for Statistical Computing

<https://www.r-project.org/>