

Anàlisi de Dades Transversals

Codi: 104878

Crèdits: 6

Titulació	Tipus	Curs	Semestre
2503852 Estadística Aplicada	OT	4	1

Professor/a de contacte

Nom: Jose Barrera Gomez

Correu electrònic: jose.barrera@uab.cat

Utilització d'idiomes a l'assignatura

Llengua vehicular majoritària: espanyol (spa)

Grup íntegre en anglès: No

Grup íntegre en català: No

Grup íntegre en espanyol: No

Prerequisits

S'assumeix que l'alumnat ha cursat prèviament l'assignatura "Estadística a les Ciències de la Salut".

Objectius

Els objectius principals del curs són:

- Conèixer les principals característiques d'un estudi epidemiològic transversal.
- Aprendre a dissenyar un qüestionari de salut.
- Aprendre a crear, netejar i validar un conjunt de dades a partir de la informació continguda en un qüestionari de salut.
- Aprendre a modelar l'associació entre un indicador de salut i una exposició potencialment relacionada, en presència de possibles problemes de confusió.
- Aprendre a modelar prevalences i taxes mitjançant models lineals generalitzats en una única població o en diferents subpoblacions.
- Utilitzar R per gestionar i modelar dades transversals.
- Ser capaç d'escriure informes estadístics reproductibles mitjançant LaTeX i el paquet knitr d'R.

Competències

- Analitzar dades mitjançant l'aplicació de mètodes i tècniques estadístiques, treballant amb dades de diverses tipologies.
- Aplicar l'esperit crític i el rigor per validar o refutar arguments tant propis com d'altres persones.
- Avaluar de manera crítica i amb criteris de qualitat el treball realitzat.
- Formular hipòtesis estadístiques i desenvolupar estratègies per confirmar-les o refutar-les.
- Identificar la utilitat i la potencialitat de l'estadística en les diferents àrees de coneixement i saber aplicar-la adequadament per extreure'n conclusions rellevants.
- Interpretar resultats, extreure conclusions i elaborar informes tècnics en el camp de l'estadística.

- Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.
- 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'índole social, científica o ètica.
- Seleccionar els models o tècniques estadístiques per aplicar-los en estudis i problemes reals, així com conèixer-ne les eines de validació.
- Seleccionar i aplicar procediments més apropiats per a la modelització estadística i l'anàlisi de dades complexes.
- Seleccionar les fonts i tècniques d'adquisició i gestió de dades adequades per a fer-ne un tractament estadístic.
- Treballar cooperativament en un context multidisciplinari assumint i respectant el rol dels diferents membres de l'equip.
- Utilitzar correctament un ampli espectre del programari i llenguatges de programació estadístiques, escollint el més apropiat per a cada anàlisi i ser capaç d'adaptar-lo a noves necessitats.
- Utilitzar eficaçment la bibliografia i els recursos electrònics per obtenir informació.

Resultats d'aprenentatge

1. Analitzar dades corresponents a estudis epidemiològics o assajos clínics.
2. Aplicar l'esperit crític i el rigor per validar o refutar arguments, tant propis com d'altres.
3. Avaluar de manera crítica i amb criteris de qualitat la feina feta.
4. Dissenyar i dur a terme tests d'hipòtesi en els diferents camps d'aplicació estudiats.
5. Elaborar informes tècnics que expressin clarament els resultats i les conclusions de l'estudi utilitzant vocabulari propi de l'àmbit d'aplicació.
6. Extreure conclusions coherents amb el context experimental propi de la disciplina a partir dels resultats obtinguts.
7. Fer el mostreig més adequat per a estudis epidemiològics.
8. Identificar les tècniques d'inferència estadística més utilitzades en estudis d'epidemiologia.
9. Interpretar els resultats estadístics en contextos aplicats.
10. Justificar l'elecció de cada mètode particular dins del context en què s'aplica.
11. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.
12. 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.
13. 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'índole social, científica o ètica.
14. Reconèixer els avantatges i els inconvenients de les diferents metodologies estadístiques quan s'estudien dades procedents de diverses disciplines.
15. Reconèixer la importància dels mètodes estadístics estudiats dins de cada aplicació particular.
16. Treballar cooperativament en un context multidisciplinari assumint i respectant el rol dels diferents membres de l'equip.
17. Utilitzar diferents programes (tant lliures com comercials) associats a les diferents branques aplicades.
18. Utilitzar eficaçment bibliografia i recursos electrònics per obtenir informació.

Continguts

(*)

1. Introducció als continguts. Introducció a la investigació reproducible utilitzant el paquet knitr de R.
2. Dades transversals.

- (a) Dades de tall transversal.
 - (b) Fonts d'informació: informació reportada, informació mida.
 - (c) Aspectes a considerar durant el disseny d'una enquesta de salut.
 - (d) El llibre de codis
3. Estudis basats en la població: estudis transversals.
- (a) Característiques
 - (b) Avantatges
 - (c) Desavantatges
 - (d) Comparació amb altres dissenys d'estudis epidemiològics.
4. El mesurament de la presència de la malaltia en estudis transversals: la prevalença
5. Exposició binària i malaltia: la taula de contingència 2 x 2.
- (a) Proves d'independència entre exposició i malaltia: aproximació asimptòtica: prova de chi-quadrat, prova de Fisher: inconvenients, disseny i implementació d'una prova exacta sota disseny de secció transversal
6. Revisió del GLM.
- (a) Especificació del model
 - (b) Estimació de màxima versemblança dels paràmetres del model.
 - (c) Proves d'hipòtesis per als paràmetres del model: prova de Wald i prova de raó de versemblances.
 - (d) Interpretació dels paràmetres del model.
 - (e) Tractar amb confusores
 - (f) Considerant interaccions
 - (g) Validació.
7. Modelat de prevalences amb el GLM.
- (a) Modelat d'OR amb la regressió logística.
 - (b) Modelat de PR amb la regressió log-binomial.
 - (c) Modelat de PD amb la regressió lineal.
 - (d) Bondat d'ajustament.
8. Modelat de recomptes i taxes amb el GLM.
- (a) Regressió de Poisson.
 - (b) Regressió binomial-negativa.
 - (c) Models per excés de zeros.
9. Introducció als models de regressió per a resposta politòmica.
10. El Model lineal generalitzat mixt per modelar prevalences i taxes en dades agrupades.

*Llevat que les restriccions imposades per les autoritats sanitàries obliguin a una prioritització o reducció d'aquests continguts.

Metodologia

(*)

- Sessions teòriques: en aquestes sessions, es presenten els diferents conceptes del tema, així com exemples il·lustratius. A més, es proposa resoldre alguns exercicis (que generalment requereixen l'ús de R). La metodologia es basa en la presentació i discussió de diapositives, així com en la presentació d'alguns materials addicionals (principalment notícies publicades en mitjans en línia i articles científics buscats a PubMed).

- Sessions pràctiques: En aquestes sessions, es proposaran diversos exemples pràctics i exercicis. Es desenvoluparan activitats relacionades amb l'ús de R, cerca a PubMed, lectura d'articles i anàlisis estadístiques. Alguns dels exercicis proposats seran de lliurament obligatori.

- Assistència a seminaris: el Departament de Matemàtiques i el Servei d'Estadística de la UAB organitzen seminaris d'estadística. Els alumnes i el professor assistiran a alguns d'ells, segons el tema i l'horari.

*La metodologia docent proposada pot experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

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			
Sessions de teoria	14	0,56	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18
Tipus: Supervisades			
Sessions de pràctiques	28	1,12	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18
Tipus: Autònomes			
Treball personal	108	4,32	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18

Avaluació

(*)

- Treballs en grup durant el curs. El professor podrà avaluar la contribució individual mitjançant qüestions orals individuals.

- Examen presencial

- Examen compensatori opcional presencial. Si l'estudiant assisteix a l'examen compensatori, la seva qualificació substituirà la puntuació en l'examen ordinari anterior, independentment de la puntuació obtinguda als dos exàmens.

*L'avaluació proposada pot experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Examen (o examen compensatori)	50%	0	0	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18
Exercicis en grup	20%	0	0	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18
Treballs en grup	30%	0	0	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18

Bibliografia

Bàsica: tots els conceptes desenvolupats en les sessions de classe es publicaran en Moodle, incloses les diapositives que es tractaran en les sessions de teoria.

Altres lectures: els estudiants interessats en anar més lluny poden explorar els següents escrits:

- Agresti, Alan. Categorical Data Analysis. Wiley, 3rd Edition, 2013.
- Breslow, N., N. Day. Statistical methods in cancer research. International Agency for Research on Cancer, 1980.
- Christensen, R. Log-Linear Models and Logistic Regression. Springer, 2nd Edition, 1990.
- Clayton D., Hills, M. Statistical models in epidemiology. Oxford University Press, 1993.
- Dalgaard, P. Introductory Statistics with R. Springer, 3rd Edition, 2002.
- dos Santos, I. Cancer epidemiology: principles and methods. International Agency for Research on Cancer, 1999.
- Gordis, L. Epidemiology. W.B. Saunders, 2004.
- Hosmer, D.W., Lemeshow, S. Applied Logistic Regression. Wiley, 2nd Edition, 2000.
- Kleinbaum, D.G. y Klein, M. Logistic Regression. A Self-Learning Text. Springer, 2002.
- Lachin, J.M. Biostatistical Methods: The Assessment of Relative Risks. Wiley, 2000.
- Motulsky, H.J. Intuitive Biostatistics. Oxford University Press, 1995.
- McCullagh, P., Nelder, J.A. Generalized Linear Models. Chapman and Hall, 1983.
- Rothman, K., Greenland, S. Modern epidemiology. Lippincott Williams & Wilkins, 1998.
- Rothman, K. Epidemiology: an introduction. Oxford University Press, 2002.
- Wassertheil-Smoller, S. Biostatistics and epidemiology: a primer for health and biomedical professionals. Springer, 3rd Edition, 2004

Programari

- R
- RStudio
- LaTeX