

Anàlisi de Dades Longitudinals

Codi: 104879

Crèdits: 6

Titulació	Tipus	Curs	Semestre
2503852 Estadística Aplicada	OT	4	2

Professor/a de contacte

Nom: Juan Ramón Gonzalez Ruiz

Correu electrònic: juanramon.gonzalez@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

- És recomanable, però no imprescindible perquè es realitzarà una classe per homogeneïtzar el nivell dels alumnes, conèixer els models lineals generalitzats i el model de Cox per a l'anàlisi de supervivència.
- L'assignatura sobre estudis observacionals conté una introducció complementària al tema que versarà sobre l'anàlisi de l'evolució temporal de taxes d'incidència i mortalitat en el qual s'introduirà el concepte de taxa.
- És recomanable però no imprescindible conèixer el paquet estadístic R

Objectius

Els objectius principals d'aquesta assignatura són:

- Conèixer els models estadístics per a l'anàlisi de dades longitudinals (informació que s'obté a partir de mesuraments efectuats al llarg del temps) que solen aparèixer freqüentment en ciències de la salut (biologia, medicina, farmacologia, toxicologia, química i / o enginyeria)
- Conèixer els models estadístics per analitzar l'evolució temporal de les taxes d'incidència i mortalitat d'una malaltia per a detectar canvis temporals i per què són deguts
- Conèixer els models estadístics per analitzar el temps fins l'ocurrència d'un esdeveniment d'interès que apareix de forma recurrent (recaigudes tumorals, migranya, infarts, ...) tenint en compte l'efecte de covariables, l'efecte de la intervenció i / o l'efecte d'observar diversos esdeveniments amb anterioritat
- Conèixer els models estadístics per analitzar dades obtingudes a partir de mesures repetides al llarg del temps utilitzant models lineals (rehospitalitzacions, recaiguda d'una malaltia, ...)
- Conèixer els models estadístics per analitzar dades obtingudes a partir de mesures repetides al llarg del temps utilitzant models no lineals (creixement tumoral en rates, evolució del pes dels nens després de néixer, ...)
- Ser capaços de llegir de forma crítica un article científic en el qual es planteja l'anàlisi d'un estudi en el qual es disposi d'informació recollida al llarg del temps.
- Ser capaços d'identificar el model estadístic necessari per analitzar un conjunt de dades que es presentaran en exercicis pràctics i que pertanyen a estudis reals.
- Saber com realitzar totes aquestes anàlisis utilitzant R mitjançant les llibreries adequades.

Competències

- 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 sàpiguen 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.
- 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. Aplicar l'esperit crític i el rigor per validar o refutar arguments, tant propis com d'altres.
2. Avaluar de manera crítica i amb criteris de qualitat la feina feta.
3. Dissenyar i dur a terme tests d'hipòtesi en els diferents camps d'aplicació estudiats.
4. Elaborar informes tècnics que expressin clarament els resultats i les conclusions de l'estudi utilitzant vocabulari propi de l'àmbit d'aplicació.
5. Extreure conclusions coherents amb el context experimental propi de la disciplina a partir dels resultats obtinguts.
6. Interpretar els resultats estadístics en contextos aplicats.
7. Justificar l'elecció de cada mètode particular dins del context en què s'aplica.
8. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.
9. Que els estudiants sàpiguen 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.
10. 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.
11. Reconèixer la importància dels mètodes estadístics estudiats dins de cada aplicació particular.
12. Treballar cooperativament en un context multidisciplinari assumint i respectant el rol dels diferents membres de l'equip.
13. Utilitzar diferents programes (tant lliures com comercials) associats a les diferents branques aplicades.
14. Utilitzar eficaçment bibliografia i recursos electrònics per obtenir informació.

Continguts

Aquests son el continguts de l'assignatura*

1. Introducció a l'assignatura
2. Models lineals per a dades longitudinals contínues: estructura de dades, visualització i models amb resposta normal
3. Models longitudinals per a data continus: models mixts
4. Models dades longitudinals amb resposta no normal

5. Regressió 'picewise' i segmentada
6. Anàlisi de supervivència amb variables dependents del temps
7. Anàlisi de supervivència amb esdeveniments recurrents
 - 7.1. Extensió del model de Cox
 - 7.2. Frailty models
 - 7.3. Model general per a esdeveniments recurrents
 - 7.4. Models amb esdeveniment terminal
8. Càlcul de mida mostral per a dades longitudinals
9. Joint Models (Survival + Longitudinal markers)

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

Metodologia

Sessions teòriques presencials:

En aquestes sessions es presentaran els principals conceptes de cada tema, així com l'anàlisi de dades amb exemples reals en els quals es mostrarà el codi necessari de R per dur a terme aquesta tasca.

Les diapositives (creades amb R Markdown - que garanteix reproductibilitat dels resultats) inclouran els conceptes teòrics, l'anàlisi de dades i la interpretació de resultats i les conclusions que s'extreuen a partir d'ells.

Sessions pràctiques presencials:

En aquestes sessions es plantejaran uns exercicis guiats que l'alumne haurà de resoldre de forma individual. Cada alumne disposarà d'un conjunt de dades individual per al mateix problema.

Es generarà una base de dades aleatòria per al mateix problema real de manera independent per a cada alumne.

Amb aquesta metodologia es pretén que l'alumne investigui i aprengui com analitzar un conjunt de dades reals sense importar que pregunti a un company com s'ha de dur a terme, ja que cada alumne haurà d'analitzar les seves dades i obtenir les seves conclusions a partir dels seus propis resultats.

Assistència a seminaris:

Excepcionalment, si es donen certes coincidències horàries i si els alumnes veuen factible poder realitzar-lo, els alumnes, juntament amb el professor, assistiran a algun seminari organitzat pel Servei d'Estadística Aplicada de la UAB o per algun altre centre d'investigació proper a la Universitat.

Aquesta assistència no és obligatòria, però serà altament recomanada pel professor, ja que els alumnes podrien veure com la metodologia que estan aprenent s'utilitza en estudis reals i podrien veure com el treball d'un estadístic té una implicació crucial en la finalització d'aquestes investigacions.

Treball individualitzat:

Les solucions de totes les pràctiques que es realitzen de forma presencial hauran de ser lliurades al professor. Com ja s'ha comentat aquestes pràctiques es realitzaran de forma individualitzada, ja que cada alumne disposarà d'una base de dades personalitzada sobre el mateix problema a tractar.

A més d'aquestes pràctiques presencials, l'alumne haurà de resoldre quatre pràctiques a casa i lliurar la solució numèrica, així com el codi de R utilitzat per obtenir aquests resultats. Tant les pràctiques presencials com les que es realitzen a casa formaran part de la seva avaluació continuada.

L'alumne tindrà accés a tot el material didàctic a partir d'una web docent en què també hi haurà un fòrum en el qual plantejar dubtes que idealment podrien ser resoltes pels seus companys i que el professor supervisarà i / o solucionarà, en cas de ser necessari.

NOTA: 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 presencials	150	6	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14

Avaluació

Preguntes d'autoavaluació:

Després de cada sessió teòrica es proposaran una sèrie de preguntes generals per avaluar si l'alumne ha après els conceptes bàsics del tema tractat durant aquesta sessió.

Lliurament de pràctiques:

Durant el curs l'alumne haurà de solucionar diverses pràctiques en què el professor podrà guiar ja que es faran de forma presencial. S'haurà d'enviar la solució numèrica així com el codi de R utilitzat per resoldre-les.

Examen:

Els alumnes realitzaran un examen presencial tipus test per avaluar si ha adquirit els conceptes teòrics i pràctics mínims sobre l'assignatura. Aquesta prova contindrà preguntes conceptuals sobre els models tractats a classe i sortides de R similars a les obtingudes en l'anàlisi que s'han realitzat durant el curs, sobre el qual es realitzaran preguntes d'interpretació de resultats.

La nota final de l'assignatura serà:

$$0,5*NE + 0,5*NPA$$

on NE correspon a la nota de l'examen i NPA a la nota de pràctiques i preguntes d'autoavaluació. Per poder fer mitja NE ≥ 5 .

NOTA: 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
Activitats individualitzades (pràctiques i autoavaluació)	50%	0	0	1, 2, 3, 4, 6, 7, 8, 9, 10, 12, 13, 14
Exàmen presencial	50%	0	0	5, 6, 7, 11

Bibliografia

Tots els models i conceptes que es presenten a les sessions teòriques es disposaran en diapositives. El codi de R per analitzar les dades d'aquestes sessions teòriques també estarà disponible per als alumnes.

Tot aquest material es podrà baixar d'aquest repositori:
https://github.com/isglobal-brge/TeachingMaterials/tree/master/Longitudinal_data_analysis

En aquesta pàgina també es podran descarregar els arxius de dades individualitzades per a la realització de les sessions pràctiques

Bibliografia complementària:

Articles (accessibles a la web de l'assignatura)

González JR, Llorca F, Moreno V. Algunos aspectos metodológicos sobre los modelos edad-periodo-cohorte. Aplicación a las tasas de mortalidad por cáncer. *Gaceta Sanitaria*, 2002;16:267-273

Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Statistics in Medicine*, 2000;19:335-51

Fernandez E, Gonzalez JR, JM Borrás, et al. Recent decline in cancer mortality in Catalonia (Spain). A Joint point regression analysis. *European Journal of Cancer*, 2001;37:2222-2228.

Gonzalez JR, Peña E, Slate E. Modelling intervention effects alter cancer relapses. *Statistics in Medicine*, 2005;24:3959-1975

V Rondeau, Gonzalez JR. Frailtypack: a computer program for the análisis of correlated failure time data using penalized likelihood estimation. *Computer Methods and Programs in Biomedicine*, 2005;80:154-164.

González JR, Peña E. Estimación no paramétrica de la función de supervivencia para datos con eventos recurrentes. *Revista Española de Salud Pública*, 2004;78:211-220

Llibres

Gonzalez JR. Modelling recurrent event data with application to cancer research. VDM Verlag, Saarbrken, Germany, 2009 (pdf del libro accesible en la web de la asignatura)

Therneau T and Grambsch P. Modeling Survival Data: Extending the Cox Model. Springer-Verlag, New York, 2000.

Duchateau, L and Janssen, P. The Frailty model. Springer-Verlag, New York 2008

Ritz C and Streibig JC. Nonlinear Regression with R. Use R! Springer, New York 2009

Programari

Tan la part pràctica com teòrica es farà amb R