

Titulación	Tipo	Curso
2503852 Estadística Aplicada	OT	4

Contacto

Nombre: Juan Ramon Gonzalez Ruiz

Correo electrónico: juanramon.gonzalez@uab.cat

Idiomas de los grupos

Puede consultar esta información al [final](#) del documento.

Prerrequisitos

Es recomendable, pero no imprescindible porque se realizará una clase para homogeneizar el nivel de los alumnos, conocer los modelos lineales generalizados y el modelo de Cox para el análisis de supervivencia. La asignatura sobre estudios observacionales contiene una introducción complementaria al tema que versará sobre el análisis de la evolución temporal de tasas de incidencia y mortalidad en el que se introducirá el concepto de tasa.

Es recomendable pero no imprescindible conocer el paquete estadístico R

Objetivos y contextualización

Los objetivos principales de esta asignatura son:

- Conocer los modelos estadísticos para el análisis de datos longitudinales (información que se obtiene a partir de mediciones efectuadas a lo largo del tiempo) que suelen aparecer frecuentemente en ciencias de la salud (biología, medicina, farmacología, toxicología, química y/o ingeniería)
- Conocer los modelos estadísticos para analizar la evolución temporal de las tasas de incidencia y mortalidad de una enfermedad para detectar cambios temporales y porqué son debidos
- Conocer los modelos estadísticos para analizar el tiempo hasta la ocurrencia de un evento de interés que aparece de forma recurrente (recaídas tumorales, migraña, infartos, ...) teniendo en cuenta el efecto de covariables, el efecto de la intervención y/o el efecto de observar varios eventos con anterioridad
- Conocer los modelos estadísticos para analizar datos obtenidos a partir de medidas repetidas a lo largo del tiempo utilizando modelos lineales (rehospitalizaciones, recaída de una enfermedad, ...)
- Conocer los modelos estadísticos para analizar datos obtenidos a partir de medidas repetidas a lo largo del tiempo utilizando modelos no lineales (crecimiento tumoral en ratas, evolución del peso de los niños tras nacer, ...)
- Ser capaces de leer de forma crítica un artículo científico en el que se plantee el análisis de un estudio en el que se disponga de información recogida a lo largo del tiempo.
- Ser capaces de identificar el modelo estadístico necesario para analizar un conjunto de datos que se presentarán en ejercicios prácticos y que pertenecen a estudios reales.

- Saber cómo realizar todos estos análisis utilizando R mediante las librerías adecuadas.

Resultados de aprendizaje

1. CM14 (Competencia) Proponer el modelo estadístico necesario para analizar conjuntos de datos pertenecientes a estudios reales.
2. KM17 (Conocimiento) Reconocer los modelos estadísticos para el análisis de datos con distintas estructuras y complejidad que aparecen frecuentemente en distintos ámbitos de aplicación.
3. KM18 (Conocimiento) Reconocer el lenguaje propio de las aplicaciones de economía y finanzas, ciencias biomédicas e ingeniería, aportado por la investigación y la innovación en el ámbito de la estadística.
4. KM18 (Conocimiento) Reconocer el lenguaje propio de las aplicaciones de economía y finanzas, ciencias biomédicas e ingeniería, aportado por la investigación y la innovación en el ámbito de la estadística.
5. SM16 (Habilidad) Seleccionar las fuentes de información adecuadas para el trabajo estadístico
6. SM17 (Habilidad) Debatir artículos científicos en que se plantee el análisis de un estudio propio de las diferentes áreas de aplicación.
7. SM18 (Habilidad) Depurar la información disponible para su posterior tratamiento estadístico.
8. SM19 (Habilidad) Analizar datos de estructuras complejas, ya sea por su naturaleza o por su dimensión.

Contenido

Estos son los contenidos propuestos*:

1. Introducción a la asignatura
2. Modelos lineales para datos longitudinales continuos: estructura de datos, visualización y modelos con respuesta normal
3. Modelos longitudinales para fecha continuos: modelos mixtos
4. Modelos datos longitudinales con respuesta no normal
5. Regresión 'picewise' y segmentada
6. Análisis de supervivencia con variables dependientes del tiempo
7. Análisis de supervivencia con eventos recurrentes
 - 7.1. Extensión del modelo de Cox
 - 7.2. Frailty modelos
 - 7.3. Modelo general para eventos recurrentes
 - 7.4. Modelos con evento terminal
8. Cálculo de tamaño muestral para datos longitudinales
9. Joint Models (Survival + Longitudinal markers)

*A menos que las restricciones impuestas por las autoridades sanitarias obliguen a una priorización o reducción de estos contenidos.

Actividades formativas y Metodología

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Sesiones presenciales	150	6	

Sesiones teóricas presenciales:

En estas sesiones se presentarán los principales conceptos de cada tema, así como el análisis de datos con ejemplos reales en los que se mostrará el código necesario de R para llevar a cabo dicha tarea. Las diapositivas (creadas con R Markdown - que garantiza reproducibilidad de los resultados) incluirán los conceptos teóricos, el análisis de datos y la interpretación de resultados y las conclusiones que se extraen a partir de ellos.

Sesiones prácticas presenciales:

En estas sesiones se plantearán unos ejercicios guiados que el alumno deberá de resolver de forma individual. Cada alumno dispondrá de un conjunto de datos individual para el mismo problema. Se generará una base de datos aleatoria para el mismo problema real de forma independiente para cada alumno. Con esta metodología se pretende que el alumno investigue y aprenda cómo analizar un conjunto de datos reales sin importar que pregunte a un compañero cómo se debe llevar a cabo, ya que cada alumno tendrá que analizar sus datos y obtener sus conclusiones a partir de sus propios resultados.

Asistencia a seminarios

Excepcionalmente, si se dan ciertas coincidencias horarias y si los alumnos ven factible poder realizarlo, los alumnos, junto con el profesor, asistirán a algún seminario organizado por el Servei d'Estadística Aplicada de la UAB o por algún otro centro de investigación cercano a la Universidad. Esta asistencia no será obligatoria, pero será altamente recomendada por el profesor, ya que los alumnos podrían ver cómo la metodología que están aprendiendo se utiliza en estudios reales y podrían ver cómo el trabajo de un estadístico tiene una implicación crucial en la finalización de dichas investigaciones.

Trabajo individualizado

Las soluciones de todas las prácticas que se realicen de forma presencial deberán ser entregadas al profesor. Como ya se ha comentado estas prácticas se realizarán de forma individualizada, pues cada alumno dispondrá de una base de datos personalizada sobre el mismo problema a tratar. Además de estas prácticas presenciales, el alumno tendrá que resolver otras cuatro prácticas en casa y entregar la solución numérica, así como el código de R utilizado para obtener dichos resultados. Tanto las prácticas presenciales como las que se realicen en casa formarán parte de su evaluación continuada.

El alumno tendrá acceso a todo el material didáctico a partir de una web docente en la que también existirá un foro en el que plantear dudas que idealmente podrían ser resueltas por sus compañeros y que el profesor supervisará y/o solucionará, en caso de ser necesario.

NOTA: La metodología docente propuesta puede experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Evaluación

Actividades de evaluación continuada

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Actividades individualizadas (prácticas y autoevaluación)	50%	0	0	CM14, KM17, KM18, SM16, SM17, SM18, SM19
Examen presencial	50%	0	0	CM14, KM17, KM18

Preguntas de autoevaluación:

Tras cada sesión teórica se propondrán una serie de preguntas generales para evaluar si el alumno ha aprendido los conceptos básicos del tema tratado durante esa sesión.

Entrega de prácticas:

Durante el curso el alumno deberá solucionar varias prácticas en las que el profesor podrá guiar ya que se harán de forma presencial. Se deberá enviar la solución numérica así como el código de R utilizado para resolverlas.

Examen:

Los alumnos realizarán un examen presencial tipo test para evaluar si ha adquirido los conceptos teóricos y prácticos mínimos sobre la asignatura. Esta prueba contendrá preguntas conceptuales sobre los modelos tratados en clase y salidas de R similares a las obtenidas en los análisis que se han realizado durante el curso, sobre el que se realizarán preguntas de interpretación de resultados.

La nota final de la asignatura se será:

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

donde NE corresponde a la nota del examen y NPA a la nota de prácticas y preguntas de autoevaluación. Para poder hacer media se necesita que $NE \geq 5$.

Evaluación única (opcional):

Se llevará a cabo un examen (4 horas) que permitirá evaluar de manera integral los conocimientos y habilidades adquiridos durante el curso. Este examen estará diseñado para evaluar la capacidad del alumno para aplicar los análisis estadísticos aprendidos y su comprensión de los conceptos teóricos.

El examen constará de dos partes principales: análisis estadístico y preguntas teóricas. En la sección de análisis estadístico, se proporcionarán datos relevantes que requerirán del alumno la aplicación de las técnicas y herramientas estadísticas aprendidas durante el curso. Se espera que el alumno realice los siguientes pasos:

1. Identificación del problema: El alumno deberá comprender la naturaleza de los datos y los objetivos de análisis.
2. Selección y aplicación de técnicas: El alumno utilizará los conocimientos adquiridos para seleccionar y aplicar las técnicas estadísticas apropiadas para analizar los datos. Esto puede incluir la determinación de medidas de tendencia central, dispersión, correlación, regresión, pruebas de hipótesis, entre otros.
3. Interpretación de los resultados: Una vez realizados los análisis, el alumno deberá interpretar los resultados de manera adecuada, explicando su significado en el contexto del problema planteado.

La segunda parte del examen consistirá en preguntas teóricas que requerirán respuestas escritas. Estas preguntas estarán relacionadas con los conceptos fundamentales de la estadística, su aplicabilidad en diferentes situaciones y su importancia en la toma de decisiones. El alumno deberá demostrar su comprensión de los conceptos y su capacidad para explicarlos de manera clara y coherente.

La evaluación de este examen se realizará teniendo en cuenta varios criterios:

1. Precisión y corrección en los análisis: Se evaluará la capacidad del alumno para realizar los análisis estadísticos de manera precisa y correcta, seleccionando las técnicas adecuadas y utilizando los procedimientos correctos.
2. Interpretación de resultados: Se valorará la capacidad del alumno para interpretar y explicar de manera coherente los resultados obtenidos en los análisis estadísticos realizados.
3. Completitud de respuestas teóricas: Se evaluará la capacidad del alumno para proporcionar respuestas claras y completas a las preguntas teóricas, demostrando un dominio de los conceptos y su aplicación.
4. Organización y claridad en la presentación: Se tendrá en cuenta la organización general del examen, la claridad de las respuestas escritas y la calidad de la presentación de los resultados estadísticos.

NOTA: La evaluación propuesta puede experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.

Bibliografía

Material principal:

Todos los modelos y conceptos que se presentan en las sesiones teóricas se dispondrán en diapositivas.

El código de R para analizar los datos de estas sesiones teóricas también estará disponible para los alumnos.

Todo este material se podrá bajar de este repositorio:
https://github.com/isglobal-brge/TeachingMaterials/tree/master/Longitudinal_data_analysis

Bibliografía complementaria :

Artículos (accesibles en la web de la asignatura)

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

Libros

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

Software

Tanto las clases teóricas como las prácticas se llevarán a cabo con R.

Lista de idiomas

Nombre	Grupo	Idioma	Semestre	Turno
(PLAB) Prácticas de laboratorio	1	Catalán	segundo cuatrimestre	tarde
(TE) Teoría	1	Catalán	segundo cuatrimestre	tarde