

Análisis de Datos Transversales

Código: 104878
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
2503852 Estadística Aplicada	OT	4	0

La metodología docente y la evaluación propuestas en la guía pueden experimentar alguna modificación en función de las restricciones a la presencialidad que impongan las autoridades sanitarias.

Contacto

Nombre: Jose Barrera Gomez
Correo electrónico: Jose.Barrera@uab.cat

Uso de idiomas

Lengua vehicular mayoritaria: español (spa)
Algún grupo íntegramente en inglés: No
Algún grupo íntegramente en catalán: No
Algún grupo íntegramente en español: No

Prerequisitos

El alumnado que asista a esta asignatura debería haber cursado previamente la asignatura "Estadística en Ciencias de la Salud".

Objetivos y contextualización

Los objetivos principales del curso son:

- Conocer las principales características de un estudio epidemiológico transversal.
- Aprender a diseñar un cuestionario de salud.
- Aprender a crear, limpiar y validar un conjunto de datos a partir de la información contenida en un cuestionario de salud.
- Aprender a modelar la asociación entre un indicador de salud y una exposición potencialmente relacionada, en presencia de posibles problemas de confusión.
- Aprender a modelar prevalencias y tasas mediante modelos lineales generalizados en una única población o en diferentes subpoblaciones.
- Utilizar R para gestionar y modelar datos transversales.
- Ser capaz de escribir informes estadísticos reproducibles mediante LaTeX y el paquete knitr de R.

Competencias

- Analizar datos mediante la aplicación de métodos y técnicas estadísticas, trabajando con datos de diversas tipologías.
- Aplicar el espíritu crítico y el rigor para validar o refutar argumentos tanto propios como de otras personas.
- Evaluar de manera crítica y con criterios de calidad el trabajo realizado.
- Formular hipótesis estadísticas y desarrollar estrategias para confirmarlas o refutarlas.

- Identificar la utilidad y la potencialidad de la estadística en las distintas áreas de conocimiento y saber aplicarla adecuadamente para extraer conclusiones relevantes.
- Interpretar resultados, extraer conclusiones y elaborar informes técnicos en el campo de la estadística.
- Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
- Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.
- Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.
- Seleccionar las fuentes y técnicas de adquisición y gestión de datos adecuadas para su tratamiento estadístico.
- Seleccionar los modelos o técnicas estadísticas para aplicarlos a estudios y problemas reales, así como conocer las herramientas de validación de los mismos.
- Seleccionar y aplicar procedimientos más apropiados para la modelización estadística y el análisis de datos complejos.
- Trabajar cooperativamente en un contexto multidisciplinar asumiendo y respetando el rol de los diferentes miembros del equipo.
- Utilizar correctamente un amplio espectro del software y lenguajes de programación estadísticos, escogiendo el más apropiado para cada análisis y ser capaz de adaptarlo a nuevas necesidades.
- Utilizar eficazmente la bibliografía y los recursos electrónicos para obtener información.

Resultados de aprendizaje

1. Analizar datos correspondientes a estudios epidemiológicos o ensayos clínicos.
2. Aplicar el espíritu crítico y el rigor para validar o refutar argumentos tanto propios como de otros.
3. Diseñar y llevar a cabo tests de hipótesis en los diferentes campos de aplicación estudiados.
4. Elaborar informes técnicos que expresen claramente los resultados y las conclusiones del estudio utilizando vocabulario propio del ámbito de aplicación.
5. Evaluar de manera crítica y con criterios de calidad el trabajo realizado.
6. Extraer conclusiones coherentes con el contexto experimental propio de la disciplina, a partir de los resultados obtenidos.
7. Identificar las técnicas de inferencia estadística más utilizadas en estudios de epidemiología.
8. Interpretar los resultados estadísticos en contextos aplicados.
9. Justificar la elección de cada método particular dentro del contexto en que se aplica.
10. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
11. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.
12. Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética.
13. Realizar el muestreo más adecuado para estudios epidemiológicos.
14. Reconocer la importancia de los métodos estadísticos estudiados dentro de cada aplicación particular.
15. Reconocer las ventajas e inconvenientes de las distintas metodologías estadísticas cuando se estudian datos procedentes de diversas disciplinas.
16. Trabajar cooperativamente en un contexto multidisciplinar asumiendo y respetando el rol de los diferentes miembros del equipo.
17. Utilizar distintos programas (tanto libres como comerciales) asociados a las distintas ramas aplicadas.
18. Utilizar eficazmente bibliografía y recursos electrónicos para obtener información.

Contenido

(*)

1. Introducción a los contenidos. Introducción a la investigación reproducible utilizando el paquete knitr de R.
2. Datos transversales.
 - (a) Datos de corte transversal.
 - (b) Fuentes de información: información reportada, información medida.
 - (c) Aspectos a considerar durante el diseño de una encuesta de salud.
 - (d) El libro de códigos
3. Estudios basados en la población: estudios transversales.
 - (a) Características
 - (b) Ventajas
 - (c) Desventajas
 - (d) Comparación con otros diseños de estudios epidemiológicos.
4. La medición de la presencia de la enfermedad en estudios transversales: la prevalencia
5. Medir la asociación entre exposición y enfermedad.
 - (a) Riesgo relativo
 - (b) Odds ratio
 - (c) Riesgo atribuible.
6. Exposición binaria y enfermedad: la tabla de contingencia 2 x 2.
 - (a) Pruebas de independencia entre exposición y enfermedad: aproximación asintótica: prueba de chi-cuadrado, prueba de Fisher: inconvenientes, diseño e implementación de una prueba exacta bajo diseño de sección transversal
7. El modelo de regresión logística.
 - (a) Presentando el modelo
 - (b) Estimación de máxima verosimilitud de los parámetros del modelo.
 - (c) Pruebas de hipótesis para los parámetros del modelo: prueba de Wald y prueba de razón de verosimilitudes.
 - (d) Interpretación de los parámetros del modelo.
 - (e) Tratar con confusoras
 - (f) Considerando interacciones
 - (g) Selección de modelos. Significación, bondad de ajuste, sensibilidad, especificidad, ROC.
8. El modelo de regresión de Poisson.
 - (a) Presentando el modelo
 - (b) Estimación de máxima verosimilitud de los parámetros del modelo.

(c) Pruebas de hipótesis para los parámetros del modelo: prueba de Wald y prueba de razón de verosimilitudes.

(d) Interpretación de los parámetros del modelo.

(e) Tratar con sobredispersión

(f) Modelado de tasas

9. Introducción a los modelos de regresión para respuesta politómica.

10. El Modelo lineal generalizado mixto para modelar prevalencias y tasas en datos agrupados.

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

Metodología

(*)

- Sesiones teóricas: en estas sesiones, se presentan los diferentes conceptos del tema, así como ejemplos ilustrativos. Además, se propone resolver algunos ejercicios (que generalmente requieren el uso de R). La metodología se basa en la presentación y discusión de diapositivas, así como en la presentación de algunos materiales adicionales (principalmente noticias publicadas en medios en línea y artículos científicos buscados en PubMed).

- Sesiones prácticas: En estas sesiones, se propondrán varios ejemplos prácticos y ejercicios. Se desarrollarán actividades relacionadas con el uso de R, búsqueda en PubMed, lectura de artículos y análisis estadísticos. Algunos de los ejercicios propuestos serán de entrega obligatoria.

- Asistencia a seminarios: el Departamento de Matemáticas y el Servicio de Estadística de la UAB organizan seminarios de estadística. Los alumnos y el profesor asistirían a algunos de ellos, según el tema y el horario.

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

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Sesiones de teoría	14	0,56	1, 2, 5, 3, 4, 6, 13, 7, 8, 9, 10, 11, 12, 14, 16, 17, 18
Tipo: Supervisadas			
Sesiones de prácticas	28	1,12	1, 2, 5, 3, 4, 6, 13, 7, 8, 9, 10, 11, 12, 14, 16, 17, 18
Tipo: Autónomas			
Trabajo personal	108	4,32	1, 2, 5, 3, 4, 6, 13, 7, 8, 9, 10, 11, 12, 14, 16, 17, 18

Evaluación

(*)

- Trabajos individuales durante el curso

- Examen presencial

- Examen compensatorio opcional presencial. Si el estudiante asiste al examen compensatorio, su calificación sustituirá la puntuación en el examen ordinario anterior, independientemente de la puntuación obtenida en el examen compensatorio.

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

Actividades de evaluación

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Ejercicios individuales	30%	0	0	1, 2, 5, 3, 4, 6, 13, 7, 8, 9, 10, 11, 12, 15, 14, 16, 17, 18
Examen (o examen compensatorio)	40%	0	0	1, 2, 5, 3, 4, 6, 13, 7, 8, 9, 10, 11, 12, 15, 14, 16, 17, 18
Trabajos personales	30%	0	0	1, 2, 5, 3, 4, 6, 13, 7, 8, 9, 10, 11, 12, 15, 14, 16, 17, 18

Bibliografía

Básica: todos los conceptos desarrollados en las sesiones de clase se publicarán en Moodle, incluidas las diapositivas que se tratarán en las sesiones de teoría.

Otras lecturas: los estudiantes interesados en ir más lejos pueden explorar los siguientes escritos:

- 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

