

MÁSTER INTERUNIVERSITARIO EN SALUD PÚBLICA

Guía docente de la asignatura: Estadística II

**Departamento de Medicina y Ciencias de la vida
UNIVERSITAT POMPEU FABRA - UNIVERSITAT AUTÒNOMA DE
BARCELONA**

DATOS DESCRIPTIVOS DE LA ASIGNATURA

Curso académico: 2024 - 2025

Nombre de la asignatura: Estadística II

Tipo de asignatura: Obligatoria

Número de créditos: 3 ECTS **Horas dedicación:** horas

Trimestre: Segundo

Idioma: Castellano

Requisitos previos: Haber cursado Estadística I y Epidemiología I.
Conocimientos básicos de R o RStudio.

Profesorado: Xavier Basagaña y Mireia Utzet

Coordinación: Mireia Utzet

PROFESORES

Profesor. Xavier Basagaña es Diplomado en estadística (1999) y Licenciado en Ciencias y Técnicas Estadísticas (2002) por la Universitat Politècnica de Catalunya (UPC), y Doctor en Bioestadística por la universidad de Harvard (2008). Actualmente es investigador del Instituto de Salud Global de Barcelona (ISGlobal). Sus líneas de investigación incluyen los efectos de la contaminación atmosférica y las temperaturas extremas en la salud y, en el apartado metodológico, los efectos de los datos faltantes y el error de medida en la inferencia estadística. Su actividad docente se centra en el máster de Salud Pública de la Universitat Pompeu Fabra (UPF) y UAB.

Email: xavier.basagana@isglobal.org

Profesora. Mireia Utzet es Licenciada en Sociología (2005) por la Universidad de Barcelona con máster en Estadística e Investigación Operativa (2009) por la Universidad Politècnica de Catalunya, y Doctora en Salud Pública (2014) por la Universitat Autònoma de Barcelona. Actualmente trabaja como investigadora en el Centro de Investigación en Salud Laboral (CiSAL) de la UPF. Sus líneas de investigación incluyen el análisis de las desigualdades sociales en salud, con especial atención en el impacto de las condiciones de Trabajo y empleo, y en la aplicación de metodología cuantitativa específica en la investigación en Salud Pública. Su actividad docente se centra en la enseñanza de epidemiología y estadística en el máster en Seguridad e Higiene en el Trabajo: Prevención de Riesgos Laborales (UB, UPC, UPF) y en el máster de Salud Pública (UPF-UAB).

Email: mireia.utzet@upf.edu

PRESENTACIÓN

Las poblaciones humanas son heterogéneas respecto a ciertas características que las pueden predisponer a desarrollar una determinada enfermedad. El estudio de esta variabilidad mediante modelos de regresión se ha convertido en una herramienta de gran utilidad para estudiar la relación entre la enfermedad y características de la población. El propósito de esta asignatura es presentar los modelos de regresión más utilizados en la investigación epidemiológica: regresión lineal múltiple, regresión logística y regresión de Cox.

Objetivos docentes

1. Presentar los modelos de regresión más útiles según el objetivo del estudio y los datos disponibles, utilizando un enfoque esencialmente aplicado y basado en ejemplos prácticos, evitando tanto como sea posible las formulaciones o demostraciones matemáticas.
2. Estimar, construir y validar los modelos de regresión con programas informáticos.
3. Interpretar los resultados de los modelos de regresión proporcionados por los programas informáticos.
4. Fomentar la lectura crítica de literatura científica en la que se apliquen modelos de regresión.

COMPETENCIAS A ALCANZAR

Competencias Generales

Instrumentales

1. Habilidad de gestión de información (recoger y analizar).
2. Competencias elementales en informática.
3. Capacidad de análisis y síntesis.
4. Comunicación oral y escrita.
5. Resolución de problemas.
6. Toma de decisiones.

Interpersonales

1. Habilidades críticas y autocríticas.
2. Trabajar en equipo.
3. Habilidad de trabajar en un equipo interdisciplinar.
4. Habilidad para comunicarse con expertos de otros campos.
5. Habilidad de trabajar en un contexto internacional.
6. Apreciación de la diversidad y multiculturalidad.
7. Compromiso ético.

Sistémicas

1. Capacidad para aplicar el conocimiento a la práctica.
2. Habilidades investigación.
3. Capacidad de aprender.
4. Capacidad de generar nuevas ideas (creatividad).
5. Habilidad de trabajo autónomo.
6. Gestión y diseño de proyectos.
7. Iniciativa y espíritu emprendedor.
8. Preocupación por la calidad.
9. Voluntad de tener éxito.

Competencias específicas

1. Justificar, argumentar y diseñar la mejor manera de recoger, procesar y almacenar la información necesaria para producir resultados y conclusiones válidos en el ámbito de la epidemiología y la salud pública.
2. Generar hipótesis específicas de investigación o de trabajo, y contrastarlas mediante el diseño de estudios y el manejo de las técnicas epidemiológicas y estadísticas útiles en cada caso, e incluso en entornos complejos y nuevos.
3. Utilizar metodología epidemiológica y estadística avanzada, útil en cada ocasión.
4. Emitir valoraciones y juicios críticos sobre la salud y sus determinantes.
5. Establecer un contacto real, como profesional competente, en el ámbito de la salud pública, bien sea en el ámbito de la investigación o de las instituciones regionales, estatales o internacionales responsables.

RESULTADOS DE APRENDIZAJE

1. Saber la formulación de los modelos de regresión lineal múltiple, regresión logística y regresión de Cox.
2. Saber seleccionar el modelo de regresión según el objetivo del estudio y los datos disponibles.
3. Saber estimar, construir y validar los modelos de regresión utilizando paquetes estadísticos
4. Saber interpretar y comunicar los resultados de los modelos de regresión.

CONTENIDO

Tema/ Módulo	Título
1	Introducción al Modelo de Regresión Lineal
2	Regresión logística
3	Análisis de supervivencia. Kaplan-Meier, log-rank, regression de Cox

METODOLOGÍA

Se realizarán actividades en el aula y fuera del aula. En el aula se impartirán clases teóricas y se realizarán ejercicios prácticos tutorizados. Fuera del aula se realizarán trabajos individuales y en grupo. Para conocer con detalle el tipo de actividad se puede consultar el apartado de programación

EVALUACIÓN

Criterios generales de evaluación de la asignatura:

Se realizará una evaluación continuada del alumno basada en los siguientes criterios:

- Participación activa en las sesiones teóricas y prácticas
- Nivel de implicación en los trabajos de grupo
- Evolución de su aprendizaje
- Capacidad de comprensión y resolución de los problemas
- Redacción e interpretación de los ejercicios propuestos
- Examen

Criterios para evaluar los entregables:

- Corrección de los conceptos preguntados
- Formato
- Ortografía y gramática
- Organización
- Claridad
- Brevedad

Porcentaje en la evaluación de entregables y examen:

Porcentaje nota final	Descripción (ponderación)
24%	Entregables E1 (50%) + E2 (50%)
24%	Entregables E3 (50%) + E4 (50%)
12%	Entregable E5 (100%)
40%	Examen (100%)

- Todos los miembros de un grupo tendrán la misma puntuación en la práctica de grupo. No se evaluarán las prácticas copiadas. Es condición necesaria entregar los entregables en el plazo establecido para poder realizar el examen.
- Para superar la asignatura es necesario obtener una nota superior a 4,9 en el examen y una nota media superior a 4,9 en la parte de entregables.
- Solo podrán realizar recuperación los estudiantes, que habiendo participado la evaluación continuada y habiéndose presentado al examen final de la asignatura, hayan obtenido la calificación final de suspenso. Los estudiantes recuperarán aquellas partes que hayan obtenido una calificación inferior a 5.
- La evaluación continua del alumno finalizará como muy tarde tres días hábiles antes de la firma de actas.

**Entregables:
Sesiones 1 y 2**

Código	Descripción	Formato	Tipo¹	Para cuándo	Tipo de corrección/fecha
E1	Correlación lineal y construcción del modelo de regresión lineal: estudio sobre madres gestantes con un embarazo único siguiendo fertilización in vitro	Electrónico (formato Word)	I	Límite 15:00 horas del 20/01/2025	Para cada grupo. (corregido antes del examen)
E2	Construcción de un modelo de regresión lineal múltiple en un estudio sobre tabaco y presión arterial.	Electrónico (formato Word)	I	Límite 15:00 horas del 27/01/2025	Para cada grupo. (corregido antes del examen)

Sesiones 3 y 4

Código	Descripción	Formato	Tipo¹	Para cuándo	Tipo de corrección/fecha
E3	Definición y estimación del modelo de regresión logística en un estudio sobre síndrome del túnel carpiano en personas trabajadoras de un hipermercado	Electrónico (formato Word)	G	Límite 15:00 horas del 3/02/2025	Para cada grupo. (corregido antes del examen)
E4	Cortès I, Artazcoz L, Rodríguez-Sanz M, Borrell C. Desigualdades en la salud mental de la población ocupada. Gac Sanit 2004; 18 (5):351-9	Electrónico (formato Word)	G	Límite 15:00 horas del 17/02/2025	Para cada grupo. (corregido antes del examen).

Sesión 5

Código	Descripción	Formato	Tipo¹	Para cuándo	Tipo de corrección/fecha
E5	Ejercicio: Análisis de supervivencia un estudio de cáncer de pulmón.	Electrónico (formato Word)	G	Límite 15:00 horas del 24/02/2025	Para cada grupo. Entregado corregido antes del examen.

¹ I: individual, G: grupal

FUENTES DE INFORMACIÓN Y RECURSOS DIDÁCTICOS

Referencias Generales:

1. Martínez-González MA, Sánchez-Villegas, Faulín Fajardos FJ. Bioestadística amigable. 2ª Edición. Ediciones Díaz de Santos, 2006.
2. Katz MH. Multivariable Analysis: A Practical Guide for Clinicians. Cambridge University Press, 2006.
3. Szklo M, Nieto J (traducción Luis Carlos Silva y Rosa Jiménez). Epidemiología intermedia. Diaz de Santos. 2003.
4. Alan Grafen, Rosie Hails. Modern statistics for the life sciences. Oxford University Press, 2002.
5. Hosmer, DW, Lemeshow S. Applied logistic regression. New York. John Wiley & Sons. 2000.
6. Silva Ayçaguer, Luis Carlos. [Excursión a la regresión logística en ciencias de la salud](#). Madrid : Díaz de Santos, cop. 1995
7. Hosmer, DW, Lemeshow S. Applied survival analysis: regression modeling of time to event data. New York. John Wiley & Sons. 1999.
8. Dalgaard, P. Introductory Statistics with R. New York: Springer. 2008.

Artículos recomendados:

Disponibles en

<https://guiesbibtic.upf.edu/br-bcum/31822>

Lecturas sesiones:

Sesiones 1 y 2

Código	Descripción	Formato	Fichero/web
S1L1	Taylor R. Interpretation of the correlation coefficient: a basic review. J Diagn Med Sonogr 1990; 1:35-39.	Electrónico	http://www.sagepub.com/salkind2study/articles/05Article01.pdf
S1L2	The Anscombe datasets	Electrónico	http://data.princeton.edu/wws509/stata/anscombe.html
S1L3	Hart J. Association between heart rate variability and manual pulse rate. J Can Chiropr Assoc. 2013 Sep;57(3):243-50.	Electrónico	http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3743650/
S1L4	Perrier E et al. Relation between urinary hydration biomarkers and total fluid intake in healthy adults. Eur J Clin Nutr. 2013 Sep;67(9):939-43.	Electrónico	http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3778844/
S2L1	Silva LC, Barroso IM. Selección algorítmica de modelos en las aplicaciones biomédicas de la regresión múltiple. Med Clin (Barc) 2001; 116: 741-745.	Electrónico	http://zl.elsevier.es/es/revista/medicina-clinica-2/articulo/seleccion-algoritmica-modelos-las-aplicaciones-13014323
S2L2	Marconi VC, et al. Cumulative viral load and virologic decay patterns after antiretroviral therapy in HIV-infected subjects influence CD4 recovery and AIDS. PLoS One. 2011;6(5):e17956	Electrónico	<i>Artículo recomendado*</i>

* Disponible en

<https://guiesbibtic.upf.edu/br-bcum/31822>

Sesiones 3 y 4

Código	Descripción	Formato	Fichero/web
S3L1	Anexo 1. Funciones exponencial y logarítmica. En: Silva LC. Excursión a la regresión logística en ciencias de la salud.	Electrónico	<i>Bibliografía recomendada</i>
S3L2	Introducción: regresión logística. Definición modelo de regresión logística. Estimación modelo de regresión logística.	Electrónico	S4L2.pdf
S3L3	Bland JM, Altman DG. The odds ratio. BMJ 2000; 320: 1468.	Electrónico	<i>Artículo recomendado*</i>
S4L1	Comparación de modelos.	Electrónico	S5L1.pdf
S4L2	Interpretación del modelo regresión logística.	Electrónico	S6L1.pdf
S4L3	Assessing the fit of the model. En: Applied logistic regression. New York. John Wiley & Sons, 2000	Capítulo de libro	<i>Bibliografía recomendada</i>
S4L4	Construcción de modelos multivariados	Electrónico	S6L2.pdf
S4L5	Cortès I, et al. Desigualdades en la salud mental de la población ocupada. Gac Sanit 2004; 18 (5):351-9.	Electrónico	<i>Artículo recomendado*</i>
S4L6	Técnicas de regresión múltiple para el ajuste (apartado 7.4). En: Szklo M, Nieto J. Epidemiología intermedia. Diaz de Santos. 2003.	Capítulo de libro	<i>Bibliografía recomendada</i>
S4L7	Silva LC. Una ceremonia estadística para identificar factores de riesgo. Salud Colectiva, Buenos Aires, 1(3): 309-322, Septiembre-Diciembre, 2005.	Electrónico	<i>Artículo recomendado*</i>

* Disponible en

<https://guiesbibtic.upf.edu/br-bcum/31822>

Sesión 5

Código	Descripción	Formato	Fichero/web
S5L1	Survival Analysis Part I: Basic concepts and first analyses. <i>British Journal of Cancer</i> (2003) 89, 232 – 238	Electrónico	<i>Artículo recomendado*</i>
S5L2	Statistics notes: Survival probabilities (the Kaplan-Meier method. <i>British Medical Journal</i> ; Dec 5, 1998; 317, 7172	Electrónico	<i>Artículo recomendado*</i>
S5L3	The logrank test. <i>British Medical Journal</i> , 2004;328;1073	Electrónico	<i>Artículo recomendado*</i>
S5L4	Pulmonary resection for metastases from colorectal cancer: prognostic factors and survival. <i>Int J Colorectal Dis</i> , 2006	Electrónico	<i>Artículo recomendado*</i>
S5L5	Holmes MD, et al. Physical Activity and Survival After Breast Cancer Diagnosis. <i>JAMA</i> , 2005	Electrónico	<i>Artículo recomendado*</i>

* Disponible en <https://guiesbibtic.upf.edu/br-bcum/31822>

PROGRAMACIÓN

SESIONES

Asignatura Estadística II		
Profesora Mireia Utzet		
Sesión 1 (13/01/2025)	El Modelo de Regresión Lineal (I)	Aula global (Disponible en la web)
Contenidos	<ul style="list-style-type: none"> • Introducir el contenido del curso. • Revisar el concepto de correlación lineal explicado en la asignatura de Estadística I. • Formular el modelo de regresión lineal simple. • Estimar e interpretar los parámetros del modelo. • ANOVA de la regresión • Medidas de bondad de ajuste del modelo. • Aplicar la regresión lineal a datos biomédicos utilizando el software RStudio 	
Seminario en aula	- Clase práctica con Ejercicio Entregable de la sesión 1 E1: Correlación lineal y construcción del modelo de regresión lineal: estudio sobre madres gestantes con un embarazo único siguiendo fertilización in vitro	E1.pdf
Lecturas obligatorias	<ul style="list-style-type: none"> - Lectura S1L1: Taylor R. Interpretation of the correlation coefficient: a basic review. J Diagn Med Sonogr 1990; 1:35-39. - Lectura S1L2: The Anscombe Datasets. 	<i>Artículos recomendados</i> *
Lecturas recomendadas	<ul style="list-style-type: none"> - Lectura S1L3: Hart J. Association between heart rate variability and manual pulse rate. J Can Chiropr Assoc. 2013 Sep;57(3):243-50. - Lectura S1L4: Perrier E et al. Relation between urinary hydration biomarkers and total fluid intake in healthy adults. Eur J Clin Nutr. 2013 Sep;67(9):939-43. 	<i>Artículos recomendados</i> *
Trabajo fuera del aula	<ul style="list-style-type: none"> - Lecturas obligatorias S1L1, S1L2 - Lecturas recomendadas S1L3, S1L4 - Ejercicio Entregable E1 	<i>Artículos recomendados</i> * E1.pdf

* Disponibles en

<https://guiesbibtic.upf.edu/br-bcum/31822>

Sesión 2 (20/01/2025)	El Modelo de Regresión Lineal (II)	Aula global (Disponible en la web)
Contenidos	<ul style="list-style-type: none"> • Formular el Modelo de Regresión Múltiple • Conocer las estrategias de construcción de los Modelos de Regresión Múltiple. • Interpretar resultados de la Regresión Múltiple en artículos científicos de aplicaciones biomédicas • Realizar el diagnóstico de los residuos del Modelo de Regresión Múltiple. 	
Seminario en aula	- Clase práctica con Ejercicio Entregable de la sesión 2 E2: regresión lineal múltiple en un estudio sobre tabaco y presión arterial	<p style="text-align: center;">E2.pdf</p> <p style="text-align: center;"><i>Artículos recomendados*</i></p>
Lecturas recomendadas	<ul style="list-style-type: none"> - Lectura S2L1: Silva LC, Barroso IM. Selección algorítmica de modelos en las aplicaciones biomédicas de la regresión múltiple. Med Clin (Barc) 2001; 116: 741-745. - Lectura S2L2: Marconi VC, et al. Cumulative viral load and virologic decay patterns after antiretroviral therapy in HIV-infected subjects influence CD4 recovery and AIDS. PLoS One. 2011;6(5):e17956 	<p style="text-align: center;"><i>Artículos recomendados*</i></p>
<i>Trabajo fuera de aula</i>	<ul style="list-style-type: none"> - Lecturas obligatorias S2L1 - Lecturas recomendadas S2L2 - Ejercicio Entregable E2 	<p style="text-align: center;"><i>Artículos recomendados*</i></p> <p style="text-align: center;">E2.pdf</p>

* Disponibles en

<https://guiesbibtic.upf.edu/br-bcum/31822>

Sesión 3 (27/01/2025)	Regresión logística (I)	Aula global (Disponible en la web)
Contenidos	<ul style="list-style-type: none"> • Introducción: regresión logística • Definición del modelo de regresión logística • Estimación del modelo de regresión logística 	S3L2.pdf
Seminario en aula	- Seminario S3 . <i>Definición y estimación del modelo de regresión logística: estudio de los factores asociados al bajo peso en recién nacidos (I)*</i>	S3.pdf E3.pdf
Lecturas obligatorias	<ul style="list-style-type: none"> - Lectura S3L1. Anexo 1. Funciones exponencial y logarítmica. En: Silva LC. Excursión a la regresión logística en ciencias de la salud. Madrid, Ediciones Díaz de Santos, 1995. - Lectura S3L2. Introducción: regresión logística. Definición del modelo de regresión logística. Estimación del modelo de regresión logística. 	<i>Bibliografía recomendada</i> S3L2.pdf
Lecturas recomendadas	- Lectura S3L3 . Bland JM, Altman DG. The odds ratio. BMJ 2000; 320: 1468.	<i>Artículo recomendado*</i>
<i>Trabajo fuera del aula</i>	<ul style="list-style-type: none"> - Lecturas obligatorias - Lecturas recomendadas - Entregable E3. <i>Definición y estimación del modelo de regresión logística: predicción de la probabilidad de tener síndrome del túnel carpiano en personas trabajadoras de un hipermercado*</i> 	S3L1 S3L2.pdf S3L3 <i>(Artículo recomendado*)</i> E3.pdf

***NOTAS:**

1. El ejercicio S3 (*Definición y estimación del modelo de regresión logística: estudio de los factores asociados al bajo peso en recién nacidos (I)*) se realizará y corregirá en la sesión 3 (día 27/01/2025). Este ejercicio no se debe entregar.

2. El entregable E3 (*Definición y estimación del modelo de regresión logística: predicción de la probabilidad de tener síndrome del túnel carpiano en personas trabajadoras de un hipermercado*) se comenzará y avanzará en su realización durante la sesión 3.

3. El artículo recomendado S3L3 está disponible en:

<https://guiesbibtic.upf.edu/br-bcum/31822>

Sesión 4 (10/02/2025)	Regresión logística (II)	Aula global (Disponible en la web)
Contenidos	<ul style="list-style-type: none"> • Comparación de modelos • Interpretación del modelo de regresión logística*. • Construcción de modelos multivariantes 	
Seminario en aula	- Seminario 4 (S4). <i>Comparación de modelos: estudio de los factores asociados al bajo peso en recién nacidos (II) *</i> .	S4.pdf
Lecturas obligatorias	<ul style="list-style-type: none"> - Lectura S4L1. Comparación de modelos - Lectura S4L2. Interpretación - Lectura S4L4. Interpretación del modelo de regresión logística. - Lectura S4L5. Construcción de modelos multivariantes. - Lectura S4L6. Cortès I, et al. Desigualdades en la salud mental de la población ocupada. Gac Sanit 2004; 18 (5):351-9 	S4L1.pdf S4L2.pdf
Lecturas recomendadas	- Lectura S4L3 . Assessing the fit of the model. En: Applied logistic regression. New York. John Wiley & Sons, 2000*	<i>Bibliografía recomendada</i>
<i>Trabajo fuera del aula</i>	<ul style="list-style-type: none"> - Lecturas obligatorias - Entregable E4. Preguntas sobre el artículo: Cortès I, et al. Desigualdades en la salud mental de la población ocupada. Gac Sanit 2004; 18 (5):351-9 	S4L1.pdf S4L2.pdf E4.pdf

***NOTAS:**

1. El seminario S4 (*Comparación de modelos: estudio de los factores asociados al bajo peso en recién nacidos (II)*) se realizará y corregirá en la sesión 4 (día 03/02/2025). Este ejercicio no se debe entregar.

2. El entregable E4 (*Preguntas sobre el artículo: Cortès I, et al. Desigualdades en la salud mental de la población ocupada. Gac Sanit 2004; 18 (5):351-9*) se comenzará y avanzará en su realización durante la sesión 4.

Asignatura Estadística II

Profesor Xavier Basagaña

Sesión 5 (17/02/2025)	Análisis de supervivencia	Aula global (Disponible en la web)
Contenidos	<ul style="list-style-type: none">• Introducción al análisis de supervivencia• Curvas de supervivencia y alternativas• Censura• Estimación de la función de supervivencia (método de Kaplan-Meier)• Comparación de grupos.• Modelo de Cox	
Seminario en aula	Seminario 5. Tiempo hasta recaída en pacientes drogodependientes (I)	S5.pdf
Lecturas obligatorias	-	
Lecturas recomendadas	<ul style="list-style-type: none">- Lectura S5L1: Introducción al análisis de supervivencia (leer sólo 3 primeras páginas)- Lectura S5L2: Estimador de Kaplan-Meier- Lectura S5L3: El test de logrank.- Lectura S5L4. Pulmonary resection for metastases from colorectal cancer: prognostic factors and survival. Int J Colorectal Dis, 2006- Lectura S5L5. Holmes MD, et al. Physical Activity and Survival After Breast Cancer Diagnosis. JAMA, 2005	Artículos recomendados*
Trabajo fuera del aula	<ul style="list-style-type: none">- Lecturas recomendadas posteriores a la sesión.- Entregable E5: Análisis de supervivencia en un estudio de cáncer de pulmón.	Artículo recomendados*

* Disponibles en

<https://guiesbibtic.upf.edu/br-bcum/31822>