

# **MÁSTER INTERUNIVERSITARIO EN SALUD PÚBLICA**

## **Guía docente de la asignatura: Estadística III**

**Departamento de Ciencias Experimentales y de la Salud**

**UNIVERSITAT POMPEU FABRA - UNIVERSITAT AUTÒNOMA DE  
BARCELONA**



## DATOS DESCRIPTIVOS DE LA ASIGNATURA

---

**Curso académico:** 2023-2024

**Asignatura:** Estadística III

**Tipo:** Optativa de itinerario

**Itinerarios:** Prevención y Servicios Sanitarios

**Número de créditos:** 3 ECTS

**Trimestre:** Tercero

**Idioma:** Castellano (materiales en inglés)

**Requisitos:** Haber cursado previamente los cursos de estadística y de epidemiología avanzada del máster, o cursos de nivel avanzado similares. En particular, poseer los conocimientos necesarios para, a partir de la pregunta científica, poder formular, estimar e interpretar el modelo de regresión lineal adecuado (p. ej. asignatura “Estadística II” de este máster).

**Coordinador:** Jose Barrera

## PROFESOR

---

**Jose Barrera** Licenciado en Física (Universitat de Barcelona), Graduado en Estadística (Universitat Autònoma de Barcelona) y Máster en Estadística (interuniversitario Universitat Politècnica de Catalunya y la Universitat de Barcelona). Trabaja como bioestadístico en el Instituto de Salud Global de Barcelona (ISGlobal) y como profesor asociado en el Departamento de Matemáticas de la Universitat Autònoma de Barcelona. [jose.barrera@isglobal.org](mailto:jose.barrera@isglobal.org). Más información en <https://sites.google.com/view/josebarrera/>.

## PRESENTACIÓN

---

Esta asignatura se centra en el estudio de indicadores epidemiológicos para cuantificar la presencia de una determinada característica de salud en la población. El interés principal se centrará en los conceptos de prevalencia e incidencia. A partir de bases de datos obtenidas en estudios epidemiológicos, se enseñará a obtener estimaciones tanto del indicador de interés como de su comparación entre diferentes subpoblaciones, obteniendo así estimaciones de medidas de asociación, en forma de diferencias o cocientes de riesgos, que facilitan resumir e interpretar tales comparaciones. Se verá cómo llevarlo a cabo teniendo en cuenta la presencia de confusoras así como complejidades en la asociación como la interacción. Adicionalmente, se enseñará cómo estimar tasas de mortalidad y cómo compararlas entre diferentes poblaciones.

## ORGANIZACIÓN Y METODOLOGÍA DEL CURSO

---

- **Número de créditos:** 3 ECTS (75 horas)
- **Organización del trabajo:**
  - **Sesiones de teoría en el aula (11 horas):** El material principal será un conjunto de diapositivas en formato PDF que se comentarán en clase y archivos con el código R que permiten reproducir todos los resultados que aparecen en las diapositivas. También se facilitará material complementario para el alumnado que desee profundizar en algunos contenidos del curso. Antes de cada clase, el material estará disponible en el Aula Global.
  - **Sesiones de prácticas en el aula (11 horas):** Se trabajará en una serie de ejercicios obligatorios (ver detalles más abajo). El profesor introducirá los ejercicios y resolverá dudas durante el trabajo del alumnado. Los materiales estarán disponibles en el Aula Global. La entrega de los informes resultado de los ejercicios se realizará a través del Aula Global.
  - **Trabajo individual fuera de clase (32 horas)**<sup>1</sup>
  - **Trabajo en grupo fuera de clase (21 horas)**<sup>2</sup>.
- **Consultas al profesor:** El horario de atención al alumnado, tanto a nivel individual como grupal (grupos de trabajo), sobre cualquier aspecto relacionado con el aprendizaje en la asignatura será cualquiera de los dos siguientes:
  - Durante las sesiones de prácticas;
  - Inmediatamente después del final de cada clase.

El alumnado que no utilice esos marcos horarios no será atendido por otras vías, salvo casos excepcionales que pueda considerar el profesor.

---

<sup>1</sup>Nótese que, por cada hora lectiva, se espera un total de 2.4 horas de trabajo fuera de clase.

<sup>2</sup>Ver nota al pie 1.

# COMPETENCIAS

---

## ■ COMPETENCIAS GENERALES

### • Instrumentales

- Comprensión básica del análisis e interpretación de la investigación epidemiológica en salud pública
- Capacidad para expresar y aplicar rigurosamente los conocimientos adquiridos en la resolución de problemas
- Capacidad para resolver problemas y comparar y contrastar hipótesis epidemiológicas
- Capacidad para tomar decisiones basadas en la evaluación crítica de la información científica
- Habilidades estadísticas e informáticas intermedias

### • Interpersonales

- Capacidad para trabajar en un equipo interdisciplinar
- Capacidad para valorar información críticamente
- Compromiso ético

### • Sistémicas

- Capacidad para aplicar los conocimientos, métodos y herramientas epidemiológicos y estadísticos a problemas concretos
- Capacidad para aprender y aplicar, de forma autónoma e interdisciplinar, nuevos conceptos y métodos relacionados con la epidemiología y la estadística
- Capacidad para analizar los datos desde las perspectivas epidemiológica y estadística
- Capacidad para expresar, de manera clara y precisa, tanto oralmente como por escrito, los resultados de los análisis
- Motivación por la calidad y la creatividad

## ■ COMPETENCIAS ESPECÍFICAS

- Entender las relaciones entre la estadística y la epidemiología a partir de la introducción explícita de ejemplos y comentarios
- Comprender los principios básicos estadísticos utilizados habitualmente en el ámbito de la salud pública para cuantificar la presencia de la enfermedad y entender las diferencias entre las medidas epidemiológicas de enfermedad más comunes
- Conocer las ventajas, inconvenientes y limitaciones de los distintos tipos de análisis de datos estudiados
- Saber elaborar y estimar modelos estadísticos adecuados a problemas reales utilizando software estadístico e interpretar los resultados
- Valorar la importancia de la estadística y de su correcta utilización para la resolución de problemas concretos del ámbito de la salud pública

# EVALUACIÓN

---

La evaluación del alumnado se basará en las actividades y criterios siguientes:

## ■ EJERCICIOS

Resolución en grupo de tres ejercicios que se irán introduciendo a lo largo del curso y que deberán ser entregados a través del Aula Global y cumpliendo de los plazos de entrega y requisitos que se establecerán al comienzo del curso. Los grupos de trabajo serán establecidos el primer día de clase. Cada uno de estos ejercicios se introducirá en una sesión de prácticas en clase, durante la cual el alumnado trabajará en su resolución con el apoyo del profesor en el aula. Los ejercicios se acabarán fuera del aula. R será el único software permitido para realizar todos los análisis de datos y para obtener todos los resultados (numéricos, tabulares y gráficos) derivados. De cada ejercicio, se entregará un único informe por grupo (cuyo formato se indicará con antelación) y la calificación del informe será común para todos los miembros del grupo. El motivo de la realización de los trabajos en grupo no es repartirse las tareas sino que todos los miembros del grupo contribuyan en todos los aspectos (análisis, interpretación y redacción) y que la interacción entre los miembros pueda servir para enriquecer el resultado final del informe. En ese sentido, el alumnado debe entender y poder explicar cualquier contenido de los trabajos presentados por su grupo. Este aspecto podrá ser evaluado en el examen.

Los criterios para evaluar los ejercicios serán:

- Claridad en la exposición de la metodología utilizada
- Corrección de los resultados
- Interpretación de los resultados
- Formato
- Ortografía y gramática
- Organización
- Claridad
- Brevedad

Los informes que sean total o parcialmente copiados, o los que no se hayan entregado dentro del plazo establecido, no serán evaluados y constarán como no presentados.

## ■ EXAMEN

Realización de un examen que evaluará tanto los conceptos y procedimientos trabajados durante las sesiones de teoría como la contribución individual a los trabajos en grupo.

No tener ningún ejercicio grupal valorado como no presentado (ver detalles en la sección EJERCICIOS) es condición necesaria para poder realizar el examen. Si el examen no se realiza por no cumplir esta condición, su calificación a efectos de computar la calificación de la asignatura será de 0.

## ■ PARTICIPACIÓN

La participación y generación de debate durante las clases, cuya evaluación supondrá un 10 % de la calificación global del curso. Cada estudiante partirá con una calificación de 0 que podrá incrementar en función de su participación.

## ■ CALIFICACIÓN FINAL DEL CURSO

La calificación final del curso se obtendrá de la siguiente manera:

- Sea:
  - $A_1$ ,  $A_2$  y  $A_3$  las calificaciones (sobre 10) en los ejercicios 1, 2 y 3, respectivamente;
  - $E$  la calificación (sobre 10) en el examen.  $E = 0$  si algún ejercicio consta como no presentado o si no se cumple la asistencia a clase (ver “ASISTENCIA A CLASE” más abajo);
  - $P$  la calificación (sobre 10) en la participación (ver “PARTICIPACIÓN” más arriba);
  - $F$  la calificación final (sobre 10) en la asignatura.
- La calificación global en los ejercicios (sobre 10) será:

$$A = 30\%A_1 + 35\%A_2 + 35\%A_3.$$

Los ejercicios valorados como no presentados (ver “EJERCICIOS” más arriba) serán calificados con 0.

- La calificación final en la asignatura será:

$$F = \begin{cases} \text{mín} \{A, E\}, & \text{si } A < 4 \text{ o } E < 4 \\ 60\%A + 30\%E + 10\%P, & \text{si } A \geq 4 \text{ y } E \geq 4 \end{cases}.$$

- Nótese que la obtención de una calificación inferior a 4 (sobre 10) en  $A$  o en  $E$  implica una calificación de suspenso en la asignatura.
- No existen actividades de recuperación.

## ■ ASISTENCIA A CLASE:

Siguiendo la información proporcionada por la coordinación del máster, la asistencia a las clases (tanto a las sesiones de teoría como a las de prácticas) es obligatoria. Para poder aprobar la asignatura, es condición necesaria pero no suficiente la asistencia a no menos del 80 % del tiempo lectivo de la asignatura.

Nótese que las consultas al profesor están programadas en horario lectivo (ver “ORGANIZACIÓN Y METODOLOGÍA DEL CURSO” más arriba).

# PROGRAMA

---

1. Introducción al curso
  - Presentación
  - Grupos de trabajo: configuración y software
2. Estandarización de tasas (2T + 2P)<sup>3</sup>
  - Método directo
  - Método indirecto
3. Introducción a los modelos lineales generalizados (1T)
4. Modelado de incidencia (4T + 5P)
  - Medidas de presencia: *incidence vs prevalence*
  - Medidas de asociación: *risk ratio* (RR) y *risk difference* (RD)
  - Medidas de impacto: *attributable fraction for the population* (AFp) y *attributable fraction among the exposed* (AFe)
  - Modelo de regresión de Poisson para incidencias
  - Modelo de regresión Binomial Negativa para incidencias con sobredispersión
  - Modelado con exceso de ceros
5. Modelado de prevalencia (4T + 4P)
  - Medidas de asociación: *odds ratio* OR, *prevalence ratio* (PR) and *prevalence difference* (PD)
  - Modelo de regresión log-binomial para PR
  - Modelo de regresión lineal para PD

---

<sup>3</sup>Horas de teoría (T) y horas de prácticas (P).

## BIBLIOGRAFÍA Y RECURSOS DE INFORMACIÓN

---

- Libros de consulta: Breslow and Day<sup>[1]</sup>, Breslow and Day<sup>[2]</sup>, Clayton and Hills<sup>[3]</sup>, Gordis<sup>[4]</sup>, Hilbe<sup>[5]</sup>, Hosmer and Lemeshow<sup>[6]</sup>, Kleinbaum and Klein<sup>[7]</sup>, McCullagh and Nelder<sup>[8]</sup>, Rothman and Lash<sup>[9]</sup>.
- Artículos científicos: Anderson and Rosenberg<sup>[10]</sup>, Loeys et al.<sup>[11]</sup>, Naimi and Whitcomb<sup>[12]</sup>, Pittman et al.<sup>[13]</sup>, Steenland and Armstrong<sup>[14]</sup>, Tamhane et al.<sup>[15]</sup>, Zeileis et al.<sup>[16]</sup>, Zou<sup>[17]</sup>.

### Referencias

- [1] NE. Breslow and NE. Day. *Statistical Methods in Cancer Research Volume I: The Analysis of Case-Control Studies*. Number 32 in IARC Scientific Publication. International Agency for Research on Cancer, Lyon, France, 1980.
- [2] NE. Breslow and NE. Day. *Statistical Methods in Cancer Research Volume II: The Design and Analysis of Cohort Studies*. Number 82 in IARC Scientific Publication. International Agency for Research on Cancer, Lyon, France, 1987.
- [3] D. Clayton and M. Hills. *Statistical Models in Epidemiology*. Oxford University Press, Oxford, 1993.
- [4] L. Gordis. *Epidemiology (5th Edition)*. Elsevier, Amsterdam, The Netherlands, 2013.
- [5] J. Hilbe. *Negative Binomial Regression (2nd Edition)*. Cambridge University Press, Cambridge, UK, 2011.
- [6] DW. Hosmer and S. Lemeshow. *Applied Logistic Regression (3rd Edition)*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2013.
- [7] DG. Kleinbaum and M. Klein. *Logistic Regression. A Self-Learning Text*. Springer, New York, 2002.
- [8] P. McCullagh and JA. Nelder. *Generalized Linear Models (2nd Edition)*. Number 37 in Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, Boca Raton, Florida, 1989.
- [9] KJ. Rothman and TL. Lash. *Modern Epidemiology (4th Edition)*. Wolters Kluwer/Lippincott Williams & Wilkins, Philadelphia, Pensilvania, 2020.
- [10] RN. Anderson and HM. Rosenberg. Age Standardization of Death Rates: Implementation of the Year 2000 Standard. *National Vital Statistical Reports*, 47(3):1–16, 1998.
- [11] T. Loeys, B. Moerkerke, O. De Smet, and A. Buysse. The analysis of zero-inflated count data: beyond zero-inflated Poisson regression. *British Journal of Mathematical and Statistical Psychology*, 65(1):163–180, 2012. URL <https://doi.org/10.1111/j.2044-8317.2011.02031.x>.
- [12] AI. Naimi and BW. Whitcomb. Estimating risk ratios and risk differences using regression. *American Journal of Epidemiology*, 189(6):508–510, 2020. URL <https://doi.org/10.1093/aje/kwaa044>.
- [13] B. Pittman, E. Buta, S. Krishnan-Sarin, SS. O'Malley, T. Liss, and R. Gueorguieva. Models for analyzing zero-inflated and overdispersed count data: An application to cigarette and marijuana use. *Nicotine & Tobacco Research*, 22(8):1390–1398, 2020. URL <https://doi.org/10.1093/ntr/nty072>.

- [14] K. Steenland and B. Armstrong. An overview of methods for calculating the burden of disease due to specific risk factors. *Epidemiology*, 17(5):512–519, 2006. URL <https://doi.org/10.1097/01.ede.0000229155.05644.43>.
- [15] AR. Tamhane, AO. Westfall, GA. Burkholder, and GR. Cutter. Prevalence odds ratio versus prevalence ratio: choice comes with consequences. *Statistics in Medicine*, 35(30):5730–5735, 2016. URL <https://doi.org/10.1002/sim.7059>.
- [16] A. Zeileis, C. Kleiber, and S. Jackman. Regression Models for Count Data in R. *Journal of Statistical Software, Articles*, 27(8):1–25, 2008. URL <https://doi.org/10.18637/jss.v027.i08>.
- [17] G. Zou. A modified poisson regression approach to prospective studies with binary data. *American Journal of Epidemiology*, 159(7):702–706, 2004. URL <https://doi.org/10.1093/aje/kwh090>.