

**Análisis de Big Data en Bioinformática**

Código: 104886  
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: Natalia Isabel Vilor Tejedor  
Correo electrónico: [Nataliabel.Vilor@uab.cat](mailto:Nataliabel.Vilor@uab.cat)

### Uso de idiomas

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

### Equipo docente

Natalia Isabel Vilor Tejedor  
Arnau Cordomi Montoya  
Angel González Wong

### Prerequisitos

Ninguna. Recomendable haber cursado la asignatura de Bioinformática.

### Objetivos y contextualización

La asignatura pretende dar una visión sobre las posibilidades del análisis de big data en bioinformática. La asignatura constado de dos bloques: 1) metodologías computacionales aplicadas al descubrimiento de fármacos y 2) análisis de datos ómicos. La asignatura forma parte de la Mención en Estadística para las Ciencias de la Salud.

### 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 los modelos o técnicas estadísticas para aplicarlos a estudios y problemas reales, así como conocer las herramientas de validación de los mismos.
- 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. Aplicar el espíritu crítico y el rigor para validar o refutar argumentos tanto propios como de otros.
2. Aplicar métodos estadísticos al análisis de datos de expresión génica.
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. Interpretar los resultados estadísticos en contextos aplicados.
8. Justificar la elección de cada método particular dentro del contexto en que se aplica.
9. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
10. 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.
11. 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.
12. Reconocer los métodos de inferencia estadística más utilizados en bioinformática.
13. Reconocer la importancia de los métodos estadísticos estudiados dentro de cada aplicación particular.
14. Trabajar cooperativamente en un contexto multidisciplinar asumiendo y respetando el rol de los diferentes miembros del equipo.
15. Utilizar distintos programas (tanto libres como comerciales) asociados a las distintas ramas aplicadas.
16. Utilizar eficazmente bibliografía y recursos electrónicos para obtener información.

## Contenido

### BLOQUE 1. Big Data en Diseño de Fármacos

1. Introducción al big data en diseño de fármacos.
2. Estructura de proteínas y espacio químico en moléculas pequeñas (tipo fármaco).
3. Interacciones proteína-fármaco.
4. Cribado virtual.
5. Dinámica molecular.

### BLOQUE 2. Big Data en Análisis de Datos Ómicos

1. Introducción a Bioconductor y herramientas bioinformáticas para el análisis de datos ómicos.
2. Estudios de asociación genética y GWAS (Estudios de genoma completo).
3. Métodos Multivariantes para la Integración de Datos Ómicos y Big Data.ç

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

## Metodología

La asignatura está organizada en sesiones de 3 horas. Cada sesión consta de una parte teórica (aulas de teoría) donde s'introducirá el temario nuevo seguida de una parte práctica (aulas de informática) donde se trabajará la aplicación de los conceptos explicados en la parte teórica. En cada sesión el profesor indicará a los estudiantes algunas tareas a realizar de manera autónoma, como lectura de artículos o elaboración de informes de prácticas. El material utilizado por los profesores estará disponible en el Campus Virtual de la asignatura.

**\*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			
Clases prácticas	21	0,84	2, 5, 3, 6, 7, 15, 16
Clases teóricas	21	0,84	9, 11, 12, 13
Presentación Proyecto de Investigación	3	0,12	9, 10
Tipo: Supervisadas			
Tutorizaciones	10	0,4	6
Tipo: Autónomas			
Estudio	70	2,8	10
Preparación Proyecto de Investigación	20	0,8	1, 3, 14

## Evaluación

BLOQUE 1. Big Data en Diseño de Fármacos (50%):

- ejercicios de prácticas (10%)
- prueba teórico-práctica (20%)
- presentación trabajo bioinformático ante una comisión (20%)

BLOQUE 2. Big Data en Análisis de Datos Ómicas (50%):

- ejercicios de prácticas (30%)
- prueba teórico-práctica (20%)

La calificación mínima global necesaria para superar la asignatura será de 5 puntos. Para hacer media es necesario que la nota mínima de cada una de las actividades evaluables sea igual o superior a 4 puntos. Los estudiantes que tengan alguna de las pruebas suspendidas o no presentadas podrán hacer el examen de recuperación donde se podrán examinar del bloque suspendido.

***\*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
Exámenes teórico-prácticos	40	4	0,16	2, 5, 3, 6, 7, 8, 9, 10, 11, 12, 13, 15
Preparación de informes de prácticas	40	0,5	0,02	4
Presentación Proyecto de Investigación	20	0,5	0,02	1, 4, 9, 14, 16

### Bibliografía

- Lesk A.M. *Introduction to Bioinformatics*. Oxford University Press 2005.
- Attwood, T.K., Parry-Smith, D.J., *Introducción a la Bioinformática*. Pearson Education, 2002.
- Foulkes A.S. *Applied Statistical Genetics with R. For Population-based Association Studies*. Springer Dordrecht Heidelberg London New York. ISBN 978-0-387-89553-6
- Gonzalez JR, Cáceres A. *Omic association studies with R and Bioconductor*. Chapman and Hall/CRC, ISBN 9781138340565, 2019.
- <https://www.bioconductor.org/>