

Análisis de Big Data en Bioinformática

Código: 104886

Créditos ECTS: 6

2025/2026

| Titulación | Tipo | Curso |
|----------------------|------|-------|
| Estadística Aplicada | OP | 4 |

Contacto

Nombre: Angel Gonzalez Wong

Correo electrónico: angel.gonzalez@uab.cat

Equipo docente

Gianluigi Caltabiano

Angel Gonzalez Wong

Juan Ramon Gonzalez Ruiz

Carolina Soriano Tarraga

Idiomas de los grupos

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

Prerrequisitos

Conocimientos básicos del idioma inglés, ya que buena parte de los artículos, tutoriales y paquetes de software están redactados en esta lengua.

Recomendable haber cursado la asignatura de Bioinformática o tener un dominio equivalente de:

- Fundamentos básicos de Biología Molecular y Genómica.
- Programación básica en R.

Objetivos y contextualización

La asignatura pretende ofrecer una visión sobre las posibilidades del análisis de Big Data enfocado en Biomedicina y Bioinformática.

La asignatura consta de dos bloques temáticos:

1. Metodologías Computacionales Aplicadas al Descubrimiento de Fármacos
2. Análisis de Datos Ómicos

Al finalizar la asignatura, el estudiante será capaz de:

- Caracterizar y gestionar datos masivos de la investigación biomédica.
- Aplicar algoritmos estadísticos, bioinformáticos y de aprendizaje automático a datos biológicos y biomédicos.
- Integrar y visualizar diversas capas de datos para interpretar hipótesis biológicas.
- Evaluar críticamente resultados e investigaciones de la especialidad, integrar conocimientos y comunicar la información de forma clara y adecuada al contexto disciplinario.

La asignatura forma parte de la Mención en Estadística para las Ciencias de la Salud.

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. SM16 (Habilidad) Seleccionar las fuentes de información adecuadas para el trabajo estadístico
5. SM17 (Habilidad) Debatir artículos científicos en que se plantea el análisis de un estudio propio de las diferentes áreas de aplicación.
6. SM18 (Habilidad) Depurar la información disponible para su posterior tratamiento estadístico.
7. SM19 (Habilidad) Analizar datos de estructuras complejas, ya sea por su naturaleza o por su dimensión.

Contenido

BLOQUE 1. Big Data en el Descubrimiento de Fármacos

- Introducción al Big Data en Biociencias, Bioconductor y el ecosistema R
- Bases de datos y representación de componentes biológicos y compuestos químicos.
- Análisis, agrupamiento y visualización de sustancias químicas y farmacológicas.
- Cribado virtual en el descubrimiento de fármacos.

BLOQUE 2. Big Data en el Análisis de Datos Ómicos

- Introducción a Bioconductor y herramientas bioinformáticas para el análisis de datos ómicos.
- Estudios de asociación genética y GWAS (Estudios de genoma completo).
- Métodos Multivariantes para la Integración de datos ómicos y big data.

Actividades formativas y Metodología

| Título | Horas | ECTS | Resultados de aprendizaje |
|------------------|-------|------|---------------------------|
| Tipo: Dirigidas | | | |
| Clases prácticas | 21 | 0,84 | |

| | | |
|--|----|------|
| Clases teóricas | 21 | 0,84 |
| Presentación Proyecto de Investigación | 3 | 0,12 |
| Tipo: Supervisadas | | |
| Tutorizaciones | 10 | 0,4 |
| Tipo: Autónomas | | |
| Estudio | 70 | 2,8 |
| Preparación Proyecto de Investigación | 20 | 0,8 |

La asignatura está organizada en sesiones de 3 horas. Cada sesión consta de una parte teórica (aulas de teoría) donde se 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, resolución de ejercicios de clase o elaboración de informes de prácticas. El material utilizado por los profesores estará disponible en el Campus Virtual de la asignatura.

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 |
|--|------|-------|------|------------------------------------|
| Exámen Teórico-Práctico | 20 | 2 | 0,08 | CM14, KM17, KM18, SM19 |
| Preparación de Informes de Prácticas | 30 | 0,5 | 0,02 | CM14, KM17, KM18, SM16, SM18, SM19 |
| Presentación ejercicios clase | 30 | 0,5 | 0,02 | CM14, KM17, KM18, SM18, SM19 |
| Presentación Proyecto de Investigación | 20 | 2 | 0,08 | KM18, SM16, SM17, SM18, SM19 |

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

- Presentación ejercicios clase (15%)
- Preparación Informes de Prácticas (15%)
- Presentación Trabajo Bioinformático ante una comisión (20%)

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

- Presentación ejercicios clase (15%)
- Preparación Informes de Prácticas (15%)
- 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 3,5 puntos.

El alumnado, para poder optar a la recuperación, debe haber sido previamente evaluado en un conjunto de actividades cuyo peso equivalga a un mínimo de dos terceras partes de la calificación total de la asignatura. Los estudiantes que tengan alguna de las pruebas suspendidas o no presentadas podrán hacer el examen de recuperación correspondiente al bloque suspendido. En caso de no superar el umbral establecido en alguno de los bloques en la recuperación, la nota final del curso será la nota mínima de los bloques.

Esta asignatura no contempla el sistema de evaluación única.

Bibliografía

- 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
- Buffalo, V. *Bioinformatics Data Skills*. O'Reilly Media, 2015.
- Lesk, A. M. *Introduction to Bioinformatics*. Oxford University Press, 2019.
- González, J. R., Cáceres, A. *Omic Association Studies with R and Bioconductor*. Chapman and Hall/CRC, ISBN 9781138340565, 2019.
- Lecturas y artículos especializados disponibles en el campus virtual de la asignatura
- <https://www.bioconductor.org/>

Software

R: <https://www.r-project.org/>

Rstudio: <https://www.rstudio.com/>

Grupos e idiomas de la asignatura

La información proporcionada es provisional hasta el 30 de noviembre de 2025. A partir de esta fecha, podrá consultar el idioma de cada grupo a través de este [enlace](#). Para acceder a la información, será necesario introducir el CÓDIGO de la asignatura

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