



## FICHA IDENTIFICATIVA

### DATOS DE LA ASIGNATURA

**Código:** 36434

**Nombre:** Estadística para datos ómicos

**Ciclo:** Grado

**Créditos ECTS:** 6

**Curso académico:** 2026-27

### TITULACIONES

Titulación	Centro	Curso	Periodo
1406 - Grado en Ciencia de Datos	Escola Tècnica Superior d'Enginyeria	3	Segundo cuatrimestre

### MATERIAS

Titulación	Materia	Carácter
1406 - Grado en Ciencia de Datos	Salud	OBLIGATORIA

### COORDINACIÓN

GRAU PEREZ MARIA

DZUNKOVA - MARIA

## RESUMEN

Se propone una introducción al análisis estadístico de datos ómicos. Entre otros utilizaremos datos obtenidos con microarrays y RNA-Seq. Estudiamos los conceptos biológicos básicos previos. Se comentan las técnicas ómicas, el procesado previo de la información y los problemas estadísticos. En particular, estudiamos el problema de las comparaciones múltiples. Luego mostramos su aplicación al problema de la expresión diferencial. También consideramos análisis de grupos de genes. El software utilizado es R y Bioconductor.

Las clases de teoría se impartirán en castellano y las clases prácticas y de laboratorio según consta en la ficha de la asignatura disponible en la web del grado.

## CONOCIMIENTOS PREVIOS

### RELACIÓN CON OTRAS ASIGNATURAS DE LA MISMA TITULACIÓN

No se han especificado restricciones de matrícula con otras asignaturas del plan de estudios.



## OTROS TIPOS DE REQUISITOS

Un primer curso en Estadística así como un dominio medio del lenguaje R.

## COMPETENCIAS / RESULTADOS DE APRENDIZAJE

### 1406 - Grado en Ciencia de Datos

(CB5) Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.

(CE12) Capacidad para diseñar y poner en marcha soluciones basadas en análisis de datos en el ámbito de la medicina y de los negocios, teniendo en cuenta los requisitos específicos de este tipo de casos de uso.

(CE15) Capacidad para modelizar y analizar la incertidumbre en estudios basados en datos así como saber interpretar y contextualizar los resultados obtenidos.

(CG01) Conocimiento de materias básicas y tecnologías, que le capacite para el aprendizaje de nuevos métodos y tecnologías, así como que le dote de una gran versatilidad para adaptarse a nuevas situaciones.

(CG03) Capacidad para la realización de modelos, cálculos, informes, planificación de tareas y otros trabajos análogos en el ámbito específico de la Ciencia de Datos.

(CG07) Capacidad para tomar decisiones de forma autónoma, elaborando de forma adecuada y original, argumentos razonados, pudiendo obtener así hipótesis razonables y contrastables.

(CT01) Ser capaces de acceder a herramientas de información (bibliográficas) y de utilizarlas apropiadamente en el desarrollo de sus tareas cotidianas.

(CT03) Habilidad para defender su trabajo con rigor y argumentos, exponiéndolo de forma adecuada y precisa, apoyándose en los medios necesarios.

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.

## DESCRIPCIÓN DE CONTENIDOS

### 1. Parte de biología - Generación de datos ómicos

- Introducción a las ciencias ómicas.
- Ácidos nucleicos y proteínas.
- Los genomas de los seres vivos.
- Cromosomas, mutaciones y la herencia.
- Extracción de ácidos nucleicos y la PCR.



- Preparación de los ácidos nucleicos para la secuenciación.
- Diferentes plataformas de secuenciación.
- Tipos de archivos en análisis genético.
- Expresión génica y la epigenética.
- La mutación y la variación poblacional.
- Las bases de datos generales de secuencias.
- El programa BLAST, anotación de los genes, mapeo de las secuencias.
- El cáncer: las mutaciones génicas y cromosómicas.
- Aplicaciones de la secuenciación de ADN.

## 2. Introducción a los datos ómicos

- Introducción a los datos ómicos y los retos asociados.
- Métodos clásicos de selección de variables en el contexto de los datos ómicos.
- Corrección por comparaciones múltiples.
- Bioconductor.

## 3. Datos

- Datos de variantes genéticas (SNPs) y preprocesado.
- Datos de expresión genética de microarrays y preprocesado.
- Datos de expresión genética de RNA-seq y preprocesado.

## 4. Análisis de expresión diferencial

- Expresión diferencial marginal o gen a gen.
- Datos de microarrays: método limma.
- Datos de RNA-Seq: métodos DESeq2 y edgeR.
- Representación gráfica de resultados.

## 5. Anotación genómica y análisis de enriquecimiento

- Anotación genómica.
- Análisis de enriquecimiento por sobre representación.

## VOLUMEN DE TRABAJO (HORAS)

### ACTIVIDADES PRESENCIALES

Actividad	Horas
Teoría	32,00



Prácticas en aula	8,00
Laboratorio	20,00
<b>Total horas</b>	<b>60,00</b>

## ACTIVIDADES NO PRESENCIALES

Actividad	Horas
Asistencia a otras actividades	0,00
Elaboración de trabajos individuales o en grupo	0,00
Estudio y trabajo autónomo	60,00
Preparación de clases	0,00
Preparación de actividades de evaluación	30,00
Resolución de casos prácticos	0,00
<b>Total horas</b>	<b>90,00</b>

## METODOLOGÍA DOCENTE

Se plantean clases magistrales así como clases prácticas en aula de Informática. Se proporciona el material desarrollado en un manual. Se propondrán cuestionarios online semanales.

(Evaluación de las competencias CB5, CT1)

## EVALUACIÓN

En la primera convocatoria, la asignatura se evaluará atendiendo a dos aspectos:

SE1: Prueba objetiva, que se realizará al terminar la docencia y que constará de cuestiones teórico-prácticas y problemas. Este apartado de la evaluación contará un 70% de la nota final. De este 70%, el 50% (sobre el total de la asignatura) corresponde a la parte de Estadística y el 20% restante (sobre el total de la asignatura) corresponde a la parte de Biología (evaluación de las competencias CB4, CB5, CG1, CE9, CE15). Las clases de la parte de biología son obligatorias y la asistencia forma 10% de la nota de la parte de biología.

SE2: Realización de cuestionarios durante el curso evaluando, mediante preguntas de teoría y práctica, el trabajo del estudiantado. Esta parte corresponde completamente a la parte de Estadística y supone el 30% sobre la nota final. Esta parte de la evaluación no es recuperable (evaluación de las competencias CB2, CB5, CG1, CT2, CE9, CE15).

El reparto de las valoraciones se hará de modo que el primer bloque de contenidos biológicos tendrá una valoración global del 20% y los restantes bloques del 80%.

El apartado SE2 no es recuperable en la segunda convocatoria.

En la segunda convocatoria se mantendrá el peso relativo de los contenidos de forma de modo que el primer bloque cuente un 20% y los restantes un 80%.



En cualquier caso, el sistema de evaluación se regirá por lo establecido en el Reglamento de Evaluación y Calificación de la Universidad de Valencia para Grados y Másteres (<https://webges.uv.es/uvTaeWeb/MuestraInformacionEdictoPublicoFrontAction.do?accion = inicio & idEdictoSeleccionado = 5639>).

La copia o plagio manifiesto de cualquier actividad que forma parte de la evaluación supondrá la imposibilidad de superar la asignatura, sometiéndose seguidamente a los procedimientos disciplinarios oportunos indicados en el *PROTOCOLO DE ACTUACIÓN ANTE PRÁCTICAS FRAUDULENTAS EN LA UNIVERSITAT DE VALÈNCIA* ([ACGUV 123/2020](#)).

## BIBLIOGRAFÍA

- Bioinformática Estadística. Análisis estadístico de datos ómicos. Guillermo Ayala. <https://www.uv.es/ayala/docencia/tami/tami13.pdf>
- Pevsner J. (2015) Bioinformatics and Functional Genomics, 3rd Edition Wiley-Blackwell
- Genetics Home Reference (2020). Help Me Understand Genetics (<https://ghr.nlm.nih.gov/>)