

**CURSO ACADÉMICO 2008 – 2009**

**TITULACIÓN: BIOLOGÍA**

**ANÁLISIS DE DATOS**

**CÓDIGO: 200810422**

**Departamento de adscripción: Estadística, Investigación Operativa y Computación**  
**Área de conocimiento: Estadística e Investigación Operativa**

**Ciclo: 2º Curso: 4º Tipo: Optativa Créditos: 4,5 (1,5T + 3P) Carácter: Cuatrimestral**  
**Periodo lectivo en que se imparte: Primer cuatrimestre**  
**Dirección web de la asignatura: [http://webpages.ull.es/users/rodorta/analisis\\_datos\\_0809/](http://webpages.ull.es/users/rodorta/analisis_datos_0809/)**

<b>HORARIO DE CLASES TEÓRICAS</b>					
<a href="http://webpages.ull.es/users/vicebiol/">http://webpages.ull.es/users/vicebiol/</a>					
<b>PRIMER CUATRIMESTRE</b>					
<b>GRUPO CT01</b>			<b>GRUPO CT02</b>		
<b>Día</b>	<b>Horario</b>	<b>Aula</b>	<b>Día</b>	<b>Horario</b>	<b>Aula</b>
Viernes	de 16:30 a 18:30 h	1			
<b>HORARIO DE CLASES PRÁCTICAS*:</b>			<b>LUGAR DE REALIZACIÓN DE LAS PRÁCTICAS:</b>		
Fecha prevista de inicio: octubre			<input type="checkbox"/> Laboratorio <input type="checkbox"/> Campo/mar <input checked="" type="checkbox"/> Aula <input checked="" type="checkbox"/> Aula de informática		
Turno: mañana					
Horario: de 08:00 a 11:00 h y de 11:00 a 14:00 h					
* para más detalles <a href="http://webpages.ull.es/users/vicebiol/">http://webpages.ull.es/users/vicebiol/</a>					

**PROFESORADO:**

**Teoría:**

Mercedes Suárez Rancel  
 Roberto Dorta Guerra

Grupo: CT01 Y CT02  
 Grupo: CT01 Y CT02

**Prácticas:**

Carlos J. Pérez González

**COORDINADOR/ES DE LA ASIGNATURA:**

Roberto Dorta Guerra

Teoría y Practicas

**LUGAR Y HORARIO DE TUTORIAS:**

**Roberto Dorta Guerra**

Atenderá a los alumnos en: Escuela Técnica de Ingeniería Informática 2ª planta  
 Miércoles de 09:30 a 13:30  
 Viernes de 09:00 a 10:00  
 Viernes de 13:00 a 14:00

**Teléfono (opcional):** 922845044    **Correo electrónico (opcional):** rodorta@ull.es

**Mercedes Suárez Rancel**

Atenderá a los alumnos en: Facultad de Matemáticas 4ª planta  
 Viernes de 09:00 a 14:00

**Teléfono (opcional):** 922318177    **Correo electrónico (opcional):** msuarez@ull.es

**Carlos J. Pérez González**

Atenderá a los alumnos en: Torre de Químicas 2ª Planta  
Viernes de 12:00 a 14:00

Teléfono (opcional): 922319982 Correo electrónico (opcional): cpgonzal@ull.es

**OBJETIVOS DE LA ASIGNATURA:**

El objetivo principal es conseguir que el alumno tome conciencia de la importancia de los métodos estadísticos aplicados en la biología. Así se introducen los métodos multivariantes como técnicas más avanzadas después de los conocimientos básicos de estadística que han sido impartidos en años anteriores. El alumno será capaz de interpretar datos así como de llevar a cabo estudios estadísticos avanzados con técnicas como el análisis Cluster, componentes principales, modelos lineales generales o ANOVA. Estas técnicas serán útiles a la hora de interpretar artículos relacionados con la biología donde se haya realizado algún tipo de estudio estadístico.

**METODOLOGÍA DOCENTE:**

- |   |   |
|---|---|
| <input checked="" type="checkbox"/> Clase magistral.          | <input type="checkbox"/> Salidas al mar.                            |
| <input type="checkbox"/> Seminarios.                          | <input type="checkbox"/> Visitas.                                   |
| <input checked="" type="checkbox"/> Prácticas de laboratorio. | <input checked="" type="checkbox"/> Trabajo, individual o en grupo. |
| <input checked="" type="checkbox"/> Prácticas en aula.        | <input type="checkbox"/> Exposición oral.                           |
| <input checked="" type="checkbox"/> Aula de informática       | <input type="checkbox"/> Docencia Virtual.                          |
| <input type="checkbox"/> Prácticas de campo.                  | <input type="checkbox"/> Otras.                                     |

**PROGRAMA DE CONTENIDOS TEÓRICOS:**

UNIDAD 0: Introducción al Análisis Multivariante.

Introducción al Análisis Multivariante. Introducción. Visión general de las técnicas Multivariantes.

UNIDAD I: Regresión Lineal Simple

Análisis Estadístico de la Regresión Lineal. Introducción. Modelo de Regresión Lineal. Estimación Puntual. Regiones e Intervalos de Confianza para los parámetros del Modelo. Test de Hipótesis en la Regresión Lineal.

UNIDAD II: Regresión Lineal Múltiple.

Análisis Estadístico de la Regresión Lineal. Introducción. Modelo de Regresión Lineal. Estimación Puntual. Regiones e Intervalos de Confianza para los parámetros del Modelo. Test de Hipótesis en la Regresión Lineal.

Eligiendo el Mejor Modelo de Regresión. Comparación de diferentes Modelos. Métodos de selección de variables hacia delante. Método de selección de variables hacia detrás. Métodos Paso a Paso.

Análisis de Sensibilidad en la Regresión Lineal. Introducción. Matriz Predicción. Álgebra de eliminar (añadir) regresores. Efecto de una(s) observación(es) en la ecuación de regresión. Validación del modelo e hipótesis asociadas.

Otros Temas:

Regresión no Lineal. Introducción. Modelo de regresión no Lineal. Estimación Mínimos Cuadrados. Métodos computacionales para mínimos cuadrados no lineales: Algoritmo Gauss-Newton y Métodos sin derivadas.

Análisis de Covarianza. Introducción. Variables Concomitantes. Modelo General. Diseño Completamente Aleatorizado con una covarianza. Diseño factorial con dos covarianzas.

UNIDAD III: Métodos de Reducción y Clasificación de Variables.

Análisis Cluster. Introducción. Medidas de Similitud. Técnicas Cluster: Jerárquicas y Particionadas. Métodos Gráficos. Verificación de las Soluciones Cluster.

Análisis de Componentes Principales. Introducción. Planteamiento geométrico del análisis de Componentes Principales. Autovalores y Autovectores de una matriz cuadrada. Diagonalización de Matrices. Test de Hipótesis asociados.

Otros Temas:

Análisis factorial. Introducción. Modelo del análisis factorial Interpretación de los factores. Rotación de los factores. Estimación de los factores. Test de Hipótesis de dimensionalidad.

Análisis de Correspondencias. Cálculo de ejes factoriales y de los factores. Contribuciones absolutas y relativas. Interpretación de resultados.

Análisis Multidimensional de Proximidades. Introducción. Representación gráfica de una matriz de desemejanzas. Métodos métricos. Solución clásica. Métodos no métricos. Algoritmos. Aplicaciones.

Diseño de Experimentos: Anova con 1 y 2 factores de Efectos Fijos. Estimación e Inferencia. Análisis de Sensibilidad. Comparaciones Múltiples.

Análisis Discriminante. Introducción. Discriminación entre dos poblaciones con una variable. Discriminación entre dos poblaciones con más de una variable. Funciones de clasificación. Probabilidades de clasificación en cada grupo.

### PROGRAMA DE CONTENIDOS PRÁCTICOS:

Se realizarán 6 prácticas de una duración de dos horas y media cada una en el aula de informática para reforzar los conocimientos adquiridos en las clases teóricas. Las prácticas son:

- Práctica 1 (Introducción al paquete STATISTICA)
- Práctica 2 (ANOVA)
- Práctica 3 (Componentes principales)
- Práctica 4 (Regresión Múltiple I)
- Práctica 5 (Regresión Múltiple II)
- Práctica 6 (Análisis Cluster)

### EVALUACIÓN:

La evaluación teórica consta de un examen escrito en cada convocatoria oficial. La evaluación práctica se realiza por medio de la asistencia obligatoria y realización de las prácticas antes descritas. Para poder presentarse al examen teórico, se deben superar como mínimo el 80% de las prácticas totales (deberán asistir como mínimo a 5 prácticas). Aquellos alumnos que no la superen deberán realizar un examen práctico en el ordenador. Dicho examen se realizará al comienzo de cada convocatoria durante 1 hora.

### CALENDARIO DE EXÁMENES (el aprobado en Junta de Facultad):

<http://webpages.ull.es/users/vicebiol/>

**Diciembre:** -

**Enero:** -

**Febrero:**

Primer llamamiento: 21/01/2009

Segundo llamamiento: 31/01/2009

**Junio:**

Primer llamamiento: 27/05/2009

Segundo llamamiento: -

**Julio:** 17/07/2009

### NORMAS DEL CURSO:

#### ENTREGA DE TRABAJOS ADICIONALES:

Los alumnos tendrán la posibilidad de entregar trabajos adicionales que serán evaluados por los profesores siguiendo los siguientes criterios:

- Como máximo se podrán entregar dos trabajos. Uno relacionado con los temas de Componentes principales o ANOVA y otro con Regresión Múltiple o Cluster.
- Cada trabajo tendrá una puntuación máxima de 0,5 puntos.
- La fecha final de entrega será el 31 de enero de 2009 durante el examen de 2ª convocatoria de febrero

- La nota del trabajo sólo se tendrá en cuenta para las convocatorias de Febrero. (Sólo se sumará la nota de los trabajos siempre y cuando se apruebe la convocatoria de Febrero. En caso de no aprobarla no se sumará la nota en posteriores convocatorias.)
- El trabajo deberá ser entregado impreso al profesor correspondiente en los plazos establecidos.

**BIBLIOGRAFÍA RECOMENDADA:**

- SUÁREZ RANCEL, MM. DORTA ROBERTO (2004). "Transparencias de Análisis de Datos". Fotocopiadora Facultad de Biología y página web.  
[http://webpages.ull.es/users/rodorta/analisis\\_datos\\_0708/](http://webpages.ull.es/users/rodorta/analisis_datos_0708/)
- Pérez Cabrera, Ana L. (2000), "Estadística Multivariante Aplicada con el Paquete Estadístico STATISTICA", Librería Campus.
- ABRAIRA SANTOS, B. Y PÉREZ DE VARGAS LUQUE A. (1996). "Métodos Multivariantes en Bioestadística". Centro de Estudios Ramón Areces. Este texto es muy didáctico, con gran cantidad de ejemplos, dirigidos a aspectos más instrumentales que formales, claramente adaptado a la Bioestadística en general. Constituye una referencia básica para el alumno.
- JOSÉ L. CARRASCO Y MIGUEL A. HERNÁN "Estadística Multivariante en las Ciencias de la Vida" Ed. Ciencia 3. Texto muy claro con poca carga matemática que toca la esencia de los métodos multivariantes desde un punto de vista de las ciencias de la vida, biología, medicina, enfermería, etc.
- FERNÁNDEZ PALACIOS, ANTONIO DE SANTOS "Ecología de las Islas Canarias".(1996). Sociedad la Cosmológica.
- STATISTICA, "Manuales del paquete estadístico STATISTICA". Guías del usuario del Paquete Estadístico Statistica. Es una guía donde los conceptos se plasman de forma muy clara y concisa. Sirve como soporte práctico y teórico.
- SUÁREZ RANCEL, M.M., (2000), "Análisis de Datos", Librería Campus. Nivel superior al exigido en esta asignatura.
- SUÁREZ RANCEL, M. MERCEDES y GLEZ. SIERRA, M.A. (1999). "Análisis de Regresión Múltiple: Teoría, Métodos y Aplicaciones". Consejería de Educación, Cultura y Deportes. Dirección General de Universidades e Investigación.
- VICTOR, J.M. (1987). "Métodos de regresión y análisis multivariantes". Uned. Resultados muy bien justificados, pero con un alto contenido matemático para el programa que aquí se propone.

**PÁGINAS WEB DE INTERÉS:**

[http://webpages.ull.es/users/rodorta/analisis\\_datos\\_0809/](http://webpages.ull.es/users/rodorta/analisis_datos_0809/)

**OBSERVACIONES:**