



Guía docente de la asignatura

Asignatura	COMPUTACIÓN ESTADÍSTICA		
Materia	Probabilidad y Estadística		
Módulo			
Titulación	GRADO EN ESTADÍSTICA (549) INdat: doble título Grado en INFORMÁTICA + Grado en ESTADÍSTICA		
Plan	549	Código	47089
Periodo de impartición	C2	Tipo/Carácter	Optativa
Nivel/Ciclo	Grado	Curso	2º
Créditos ECTS	6 ECTS		
Lengua en que se imparte	Español		
Profesor/es responsable/s	María Teresa González Arteaga		
Datos de contacto (E-mail, teléfono...)	Despacho 229 en la Facultad de Ciencias. teresag@eio.uva.es		
Horario de tutorías	Consultar la web de la UVa		
Departamento	Estadística e Investigación Operativa		



1. Situación / Sentido de la Asignatura

1.1 Contextualización

Como objetivo general pretendemos que el alumnado conozca los paquetes estadísticos R y SAS como entornos de trabajo y los maneje con la soltura suficiente como para ser capaz de plantearse utilizar nuevas funciones y nuevas librerías cuando sea necesario en el futuro.

La elección del R no es casual, sino que responde a sus características. Primeramente, debemos destacar que R es un entorno y un lenguaje para la computación estadística que proporciona una gran variedad de técnicas estadísticas y gráficas (test de hipótesis, modelos lineales y no lineales, series temporales, clasificación, etc). En segundo lugar, de gran interés resulta que este software es de distribución libre al cual se puede acceder fácilmente en la web (<http://www.cran.r-project.org/>).

Por otro lado, este entorno permite que los futuros profesionales accedan fácilmente a muchas librerías para enfrentar otros métodos estadísticos tanto ya estudiados como los que aún no conocen. Estas librerías les serán de fácil uso una vez conocido el manejo esencial del R que planteamos en esta asignatura.

El término "entorno" lo caracteriza como un sistema completamente diseñado y coherente, antes que como una agregación incremental de herramientas muy específicas e inflexibles, como ocurre frecuentemente con otros programas de análisis de datos.

La elección de SAS tampoco es casual. Este entorno es uno de los más utilizados por muchas empresas e instituciones estando entre los más importantes en el mercado. Por las mismas razones haremos una breve incursión en el paquete SPSS.

La asignatura está totalmente orientada a las aplicaciones y por ello todas las horas lectivas se impartirán en el laboratorio y casi todo el trabajo que el alumnado tendrá que realizar será de índole práctico. Además, utilizaremos la herramienta de e-learning MOODLE para que el alumnado tenga mayor facilidad con la disponibilidad de los distintos recursos que la profesora proporciona, para realizar sus tareas, así como, para relacionarse con el resto del grupo.

1.2 Relación con otras materias

En esta asignatura se introduce software estadístico que será utilizado en la mayoría de las asignaturas posteriores a la misma en el plan de estudios de Grado en Estadística y en el doble título INdat.



1.3 Prerrequisitos

Es recomendable tener conocimientos básicos de probabilidad, estadística descriptiva, álgebra y cálculo.

2. Competencias

G1 G2 G3 E1 E3 E6 I1 I2 I3 I4 I5 I6 I7 P2 P3 P4 S1 S2 S3 S4 S5

2.1 Generales

- G1. Capacidad para la gestión de la información.
- G2. Capacidad para la abstracción y el razonamiento crítico.
- G3. Capacidad para la puesta al día y el auto-aprendizaje.

2.2 Específicas

- E1. Recogida y tratamiento de datos
- E3. Ajuste de modelos estadísticos y de investigación operativa
- E6. Presentación y comunicación de resultados

3. Objetivos

Resultados de aprendizaje recogidos en la memoria del grado:

- Capacidad para manejar conjuntos de datos en el entorno de SAS y en R.
- Saber hacer análisis descriptivos, así como ajustar modelos de regresión lineal simple con SAS y R.
- Saber utilizar las herramientas de simulación de SAS y R.
- Capacidad para realizar gráficos con R.
- Capacidad para programar, a nivel básico, en SAS y en R.

4. Tabla de dedicación del estudiante a la asignatura

ACTIVIDADES PRESENCIALES	HORAS	ACTIVIDADES NO PRESENCIALES	HORAS
Clases teóricas	-	Estudio y trabajo autónomo individual	90
Clases prácticas	45	Estudio y trabajo autónomo grupal	
Laboratorios	11		
Prácticas externas, clínicas o de campo			
Seminarios			
Otras actividades	4		
Total presencial	60	Total no presencial	90



5. Bloques temáticos

Bloque 1: COMPUTACIÓN ESTADÍSTICA

Carga de trabajo en créditos ECTS:

a. Contextualización y justificación

La contextualización puede leerse en el apartado correspondiente a la asignatura completa que aparece más arriba.

b. Objetivos de aprendizaje

- Manejar con soltura el programa R como entorno de programación para Análisis de Datos y la realización de gráficos.
- Realizar el análisis estadístico básico de datos con R y desarrollar la capacidad de abordar otros más complejos de forma autónoma por el alumnado.
- Conocer e implementar algoritmos para generar muestras de distintas distribuciones aleatorias (continuas, discretas, multivariantes).
- Utilizar la simulación de distribuciones con fines específicos dentro de la Estadística como por ejemplo para aplicar procedimientos de inferencia estadística.
- Introducir el entorno de SAS para manejar conjuntos de datos y realizar análisis descriptivos.
- Introducir el software SPSS para manejar conjuntos de datos y realizar análisis descriptivos.

c. Contenidos

Temas a desarrollar en el programa de la asignatura:

1. Introducción al sistema SAS: lectura de datos y manejo de archivos. Paso data. Procedimientos básicos de SAS. Programación con SAS.
2. Introducción al paquete R: manejo de objetos en R. Manejo de archivos. Funciones básicas. Programación con R. Gráficos tradicionales con R. Librería Lattice.
3. Algunas aplicaciones estadísticas con R: Funciones lm, loess, density. Simulación de muestras para distribuciones de probabilidad, normal bidimensional. Representación de datos en mapas.
4. Introducción a SPSS.



d. Métodos docentes

La asignatura se desarrollará mediante la realización de diversas actividades, clases en el laboratorio, tanto teóricas como prácticas, entrega de ejercicios y cuestionarios, exposición oral de ejercicios resueltos, tutorías individualizadas y examen final.

En esta asignatura utilizaremos la herramienta de e-learning MOODLE que está implementada en el campus virtual de la UVa. Esta herramienta permitirá una relación más fluida en la realización de actividades y en la disposición de los distintos recursos para el alumnado.

Todas las actividades tienen como objetivo principal el de potenciar el aprendizaje de los alumnos, facilitando la adquisición de cuantos conocimientos y competencias precisas. Las diferentes actividades estarán sujetas a un proceso de evaluación continua, y algunas permitirán dar la certificación necesaria del aprendizaje. Véase el apartado dedicado a la evaluación del aprendizaje.

En esta asignatura se utilizan programas sobre el sistema operativo Windows. Estos programas son R, SAS básico y SPSS.

A continuación se detallan las diferentes actividades que se realizarán a lo largo del curso en el ámbito de esta asignatura.

Clases:

La teoría básica necesaria será expuesta en clase por la profesora de la asignatura y se ilustrará continuamente su aplicación mediante ejemplos, lo cual llevará a que no podamos diferenciar claramente entre clases de teoría y clases prácticas. Todas las horas lectivas se desarrollarán en el laboratorio.

La participación activa del alumnado será necesaria en todos los casos.

Pruebas y ejercicios:

En las fechas señaladas en el calendario de actividades se solicitará a cada alumno/a la realización de ejercicios o la respuesta a un cuestionario. Estas dos pruebas tendrán lugar dentro del horario de clases previsto. Estas pruebas se realizarán en equipos con sistema operativo Windows.

Se llevará a cabo un seguimiento de la asignatura y de la realización de ejercicios. A lo largo del curso se pedirá al alumnado que exponga oralmente algunos ejercicios de las listas de ejercicios propuestos por la profesora y/o que responda a algún cuestionario. También se podrá solicitar la grabación de la exposición de tareas.

Examen Final:

El alumnado tendrá que dar respuesta a varios ejercicios que se resolverán en el laboratorio utilizando el paquete R y/o SAS básico y/o SPSS con cierto material de apoyo que se indicará con antelación. El examen se realizará en equipos con sistema operativo Windows.

Tutorías:

Las tutorías individualizadas podrán ser atendidas en el horario que se disponga para ello al inicio del curso o a cualquier otra hora, previa cita con la profesora.

En todas las actividades realizadas se llevará un **control de asistencia**.



e. Plan de trabajo

Programa detallado de la asignatura:

1. Introducción al sistema SAS: lectura de datos y manejo de archivos. Paso data. Procedimientos básicos de SAS. Programación con SAS.

- Nociones generales sobre el entorno SAS.
- Paso data:
 - o Creación de ficheros de datos sas. Lectura de datos en formatos no SAS. Almacenamiento de datos en ficheros externos. Operaciones con ficheros sas.
 - o Sentencias en el paso data: crear nuevas variables, if, do, output, ...
- Generalidades del paso proc.
- Procedimientos básicos en SAS: Proc PRINT, FREQ, MEANS, SUMMARY, UNIVARIATE, TABULATE, GPLOT, ...
- Introducción a la realización de mapas temáticos con SAS: proc MAPS

Bibliografía:

Pérez, C. "El Sistema estadístico SAS". 2001 de. Prentice Hall

"Introducción al SAS". Joan Valls, Llorenç Badiella.

Manuales de SAS.

2. Introducción al paquete R: Objetos en R. Manejo de archivos. Funciones básicas. Programación con R. Gráficos tradicionales con R. Librería Lattice.

- Instalación. Primeros pasos.
- Objetos (vectores, matrices, arrays, marcos de datos, listas)
- Operaciones básicas con objetos. Funciones básicas.
- Manejo de archivos. Importar y exportar datos.
- Gráficos tradicionales con R.
- Programación básica con R.
- Librería lattice.
- Realización de informes con R con rmarkdown. Introducción a la realización de mapas temáticos.

Web: <http://www.cran.r-project.org/>

Bibliografía:

Dalgaard, P.

"R para principiantes". Paradis E. (Texto disponible en la web de R)

"Introducción a R" R Core Team. (Texto disponible en la web de R)

3. Algunas aplicaciones: análisis descriptivo, simulación de muestras para distribuciones de probabilidad, método de Montecarlo, remuestreo.



- Análisis descriptivo. Regresión lineal simple. Funciones lm y loess.
- Manejo de las distribuciones de probabilidad clásicas con R.
- Función density.
- Concepto de número pseudoaleatorio. Generadores de U(0, 1).
- Algunos métodos de generación de muestras.
- Método de Montecarlo para simulación de la distribución de estadísticos.

Bibliografía:

Gentle, J.E

Ríos Insua, D.

Martinez W., Martínez A. "Computational statistics handbook with MATLAB".

4. Introducción al SPSS.

f. Evaluación

Se detalla a continuación el procedimiento para asignar la calificación final.

- El 40% de la nota se obtiene con la **evaluación continua** a través de las actividades realizadas a lo largo del cuatrimestre.

Se realizarán dos pruebas con peso en la calificación final del 10% para la prueba primera y 15% para la segunda. *Las fechas provisionales de estas pruebas son el 17 de marzo y el 4 de mayo de 2016. Estas pruebas se realizan en equipos con sistema operativo Windows y sólo se permite utilizar el software indicado en el aula.*

También se hará un seguimiento del trabajo del alumnado y de la realización de ejercicios incluyendo la exposición oral de algunos de ellos o su grabación. Esto tendrá un peso del 15% de la calificación final.

Llamémos A a la puntuación obtenida en la evaluación continua.

- El 60% de la nota se obtiene con un **examen final**. Sea B la puntuación de este examen en una escala de 0 a 10.
- Para realizar la media ponderada de la puntuación en la evaluación continua y la puntuación del examen final, B deben ser mayores de 3.
- **Nota final = $0.40 * A + 0.60 * B$** si se cumple la condición anterior.
- Para aprobar B debe ser mayor de 3 y la Nota final debe superar el valor de 5.

En la convocatoria extraordinaria de julio habrá un examen único correspondiente al 100% de la nota y no se tendrá en cuenta la nota de la evaluación continua obtenida durante el cuatrimestre.

g. Bibliografía básica

- "R para principiantes". Paradis E. (Texto disponible en la web de R)
- "Introducción a R" 2001 R Core Team. (Texto disponible en la web de R)
- Dalgaard, P. "Introductory statistics with R". Springer. 2002.



- Pérez, C. "El Sistema estadístico SAS". 2001 de. Prentice Hall
- "Introducción al SAS". Joan Valls, Llorenç Badiella. Servei d'Estadística UAB

h. Bibliografía complementaria

- Gentle, J.E., "Random Number Generation and Monte Carlo Methods". Springer-Verlag, 1998.
- Ríos Insua, D., Ríos Insua, S., Martín, J, Jiménez, A. "Simulación. Métodos y aplicaciones" 2008. Ed. Rama
- Maindonald, J., Braun, J. "Data analysis and graphics using R. An example-based approach" Cambridge University press 2007.
- Martínez W., Martínez A. "Computational statistics handbook with MATLAB". Chapman & Hall 2002.

i. Recursos necesarios

Software estadístico especificado en los temas del programa.

6. Temporalización (por bloques temáticos)

BLOQUE TEMÁTICO	CARGA ECTS	PERIODO PREVISTO DE DESARROLLO
COMPUTACIÓN ESTADÍSTICA	6	C2

7. Sistema de calificaciones – Tabla resumen

INSTRUMENTO/PROCEDIMIENTO	PESO EN LA NOTA FINAL	OBSERVACIONES
Ejercicios a lo largo del curso	15,00%	La realización de una cantidad suficiente de ejercicios. Respuesta a las actividades de seguimiento.
Prueba 1	10,00%	<i>Fecha provisional: 8 de marzo</i>
Prueba 2	15,00%	<i>Fecha provisional: 27 de abril</i>
Examen final	60,00%	

En la convocatoria extraordinaria de julio se realizará un examen con una calificación sobre 10 sin tener en consideración las calificaciones parciales obtenidas a lo largo del cuatrimestre.

8. Consideraciones finales

Las clases y actividades se desarrollarán en el laboratorio.

Es recomendable la capacidad de leer inglés técnico.