

PLAN DE FORMACIÓN DOCENTE DE LA UNIVERSIDAD DE CÁDIZ

MODELOS AVANZADOS DE ANALISIS ESTADISTICO CON R CAMPUS DE PUERTO REAL

Dirigido a: Profesorado de la Universidad de Cádiz.

Objetivos:

El sistema estadístico R es un software de código abierto que aporta implementaciones de la casi totalidad de las técnicas conocidas en investigación. El propósito del curso presentado es, manteniendo un carácter eminentemente práctico pero genérico, mostrar al investigador la aplicación e interpretación de dichas técnicas. Para ello, se trabaja sobre datos reales, haciendo uso de un entorno gráfico desarrollado a tal efecto que simplifica notablemente la labor del usuario. En futuros cursos orientados al Análisis de Datos, la formación se centrará en las técnicas específicas empleadas por las distintas disciplinas y/o áreas de investigación en la UCA

Contenidos:

**** LECTURA Y ALMACENAMIENTO DE DATOS:**

Importación y exportación de datos, desde los formatos Rd, txt, csv, Spss, Excel.
Ejecución mediante la interfase mejorada de Rcmdr.

**** CONSTRUCCION Y ETIQUETADO DE VARIABLES.**

Transformaciones de cambio de escalas.

Etiquetado de categorías.

Remodificación, Categorización, Transformaciones estadísticas y matemáticas.

Transformaciones Box-Cox.

Ejecución mediante la interfase mejorada de Rcmdr.

**** DESCRIPCION DE VARIABLES:**

Variables Simples,

Descriptivas por Estratos.

Variables categóricas, tablas y gráficos asociados.

Correlaciones paramétricas y no paramétricas.

Ejecución mediante la interfase mejorada de Rcmdr.

**** CONTRASTES SOBRE MEDIAS Y PROPORCIONES**

Prueba t-test para una muestra simple.

Comparación de dos muestras apareadas e independientes.

Análisis de k muestras independientes.

Comparación de una y dos proporciones.

**** ANALISIS DE LA VARIANZA**

Modelos de efectos fijos y aleatorios.

Anova de 1 factor. Anova de dos factores.

Anova Multifactorial. Hipótesis previas y diagnosis del modelo.

**** MODELOS DE REGRESION Y AJUSTE.**

Modelos linealizables con una predictora. Regresión Lineal Múltiple.

Hipótesis previas y diagnosis del modelo.

Modelos polinómicos.

Modelos de ajuste no paramétricos.

Ejecución mediante la interfase mejorada de Rcmdr.

**** REDUCCION DE DIMENSION:**

Análisis de Componentes Principales.
Regresión en Componentes principales.
Análisis Factorial, Análisis de Correspondencias.
Ejecución mediante la interfase mejorada de Rcmdr.

**** CLASIFICACION NO SUPERVISADA**

Técnicas Clusters.
Modelos Jerárquicos y su representación.
Modelos no jerárquicos.
El método k-mean no supervisado.
Ejecución mediante la interfase mejorada de Rcmdr.

**** CLASIFICACION SUPERVISADA**

Clasificadores Bayesianos sencillos.
Análisis discriminante de Fisher.
Método de los k-vecinos más cercanos.
Modelo de discriminación Cuadrática.
Discriminación Logística.
Árboles de clasificación.
Ejecución mediante la interfase mejorada de Rcmdr.

Metodología:

Cada día de curso se plantea como una sesión diferente de 4 horas de trabajo.
Cada sesión consta de dos partes diferenciadas: una parte teórica donde el profesor/tutor explicará una serie de contenidos y una parte practica donde el alumno trabajando solo o por grupos resolverá una serie de casos prácticos relativos a los contenidos explicados.

La parte Teórica .-

La parte teórica a su vez consta de una exposición breve y no muy profunda de la teoría estadística que subyace en cada método, la explicación de cómo se realizaría en la practica usando el paquete estadístico R y la explicación de se obtendría los mismos resultados usando la interfase Rcmdr mejorada y desarrollada a tal efecto. De esta forma se consigue un doble objetivo, por una parte la enseñanza del R en modo comandos y como alternativa usando la interfase con el usuario.

Casos y Ejercicios.-

Cada día de curso se trabajará durante las 2 o 2.5 últimas horas en casos reales realizando para ello ejercicios orientados a poner en práctica la teoría explicada ese día y en los anteriores. El alumno podrá realizar los ejercicios solo o en grupo, con comandos R y mediante la interfase explicada. Al final del día el tutor recogerá los resultados del trabajo de cada alumno.

Material Didáctico.-

Consta de:

- * Presentación PowerPoint (con proyector) para explicar la parte conceptual.
- * Ejercicios suministrado en papel.
- * Datos obtenidos almacenados previamente en disco o www.
- * El programa R se encontrara instalado.
- * La interfase de Rcmdr ampliada se encontrara en una dirección www donde el alumno pueda bajarla e instalarla durante el curso.
- * El alumno recibirá además un manual en formato papel que le facilite el seguimiento de las técnicas explicadas en el curso.

Trabajo Final.-

El estudiante realizará un trabajo final, que constará de un estudio completo apoyado por datos reales y lecturas complementarias. Este trabajo lo realizará a modo de prueba final de conocimientos durante las horas no lectivas del curso.

Evaluación:

La reflexión sobre la propia visión docente y ejercitar el pensamiento crítico se debe plasmar en la construcción de un modelo de docencia que se ajuste a los objetivos definidos por el propio docente. Los criterios de valoración son sociales, por tanto, dichos criterios son negociables en el contrato-pedagógico. El peso fundamental de la evaluación se fundamentará en el PORTAFOLIO personal de cada partícipe, considerando los siguientes aspectos:

1. Autoevaluación del propio alumno mediante un informe final razonado sobre logro de los objetivos propuestos y la definición de su perfil docente. La autoevaluación implica, asimismo, un juicio integrador de la interacción educador-educando.
2. La evaluación de sus compañeros de grupo mediante la evaluación continua y un informe final razonando sobre el trabajo del compañero que valor3e su desempeño en el grupo en cuanto a su papel en el grupo: iniciativas y aportaciones, actitudes e interrelaciones.

Así mismo, al final del curso se planteará al alumno que evalúe tanto el curso y diversos aspectos.

Selección:

Para la selección se tendrá como preferente a los profesores inscritos, no admitidos en la edición anterior.

Como segundo criterio se atenderá al orden de inscripción.

Impartido por:

Andrés Jiménez Jiménez

Número máximo de participantes:

12

Lugar de celebración:

Aula de Informática CITI Campus de Puerto Real

Fecha de celebración:

Sesiones presenciales:

14, 15, 16, 17 y 18 de abril del 2008

Duración:

18 horas presenciales y 7 horas no presenciales

Horario: de 9:30 a 13:30h

Dirección para realizar la inscripción:

http://www-personal.uca.es/PDI/FORMACION/F_Solicitud_Curso_PDI.htm

Lugar de celebración:

Sala de informática del CITI, Campus de Puerto Real.

Fecha límite para realizar la inscripción: 7 de abril.