

## Ejercicio Técnicas Estadísticas

El ejercicio debe realizarse individualmente. Se permite la consulta a compañeros o internet en caso de dudas puntuales, pero se vigilará y penalizará el plagio. El objetivo es demostrar (a vosotros mismos y a los profesores) que sois capaces de realizar estos análisis.

Podéis utilizar la aproximación seguida en clase en cuanto a paquetes, funciones, etc., pero se valorará el esfuerzo extra por vuestra parte, incluyendo la presentación esmerada del código y los resultados (por ejemplo, código ordenado, eficiente y bien comentado, tablas informativas y bien formateadas, y figuras aptas para su inmediata publicación).

### Instrucciones

En el fichero `datasets.zip` hay un dataset diferente para cada alumno. Utiliza el que te corresponda y realiza un análisis completo del problema que se plantea más abajo.

El ejercicio deberá presentarse como un documento Word o PDF (por ejemplo, generado mediante *Rmarkdown*) que incluya el código ejecutado y los resultados (tablas, figuras, discusión de resultados). El código debe estar completo (esto es, debe ser capaz de regenerar todos los resultados).

**Envía estos documentos por correo electrónico a `david.garcia.callejas@gmail.com`.**

**Fecha límite de envío: 15/02**

### Problema

Se ha detectado una nueva enfermedad parasitaria en peces de río de la provincia de Huelva. La Consejería de Medio Ambiente te ha encargado un estudio para evaluar la incidencia de esta enfermedad en distintos ríos de la provincia. Se ha realizado un muestreo de 900 peces en distintos ríos, anotando su peso (en gramos) y comprobando si estaban afectados por la enfermedad.

Debes realizar un informe completo que describa y analice los datos, discuta los resultados e incluya abundante material gráfico, incluyendo:

- descripción verbal de los datos
- análisis exploratorio de los datos (visualización)
- ajuste de modelos
- descripción y visualización de los resultados del modelo
- comprobación de la adecuación del modelo (residuos, etc.)

Además debe contestarse explícitamente a las siguientes preguntas:

1. Cuántos peces se han muestreado en cada río.
2. Cuántos peces enfermos se han encontrado en cada río.
3. ¿Cómo varía la incidencia de la enfermedad entre ríos.Cuál es la probabilidad de estar enfermo en cada río?
4. ¿Existe relación entre el peso y la enfermedad? La probabilidad de estar enfermo está asociada al tamaño del pez?
5. ¿Cuál es la probabilidad media de padecer la enfermedad en peces de 40 gramos de peso para cada río?
6. Incluye una evaluación de los residuos del modelo (por ejemplo, una figura). ¿Es adecuado el modelo para este análisis?