

# Diseño y Análisis de Experimentos - enunciado PEC6

## Máster de Bioinformática y Bioestadística

**NOTA: Todos los problemas tienen la misma importancia. Se entregarán, un único script Rmd, separando de manera clara cada problema, y, el pdf o html asociados.**

### Problema 1. Comparación de espectrómetros para la medida de la proporción de 14N a 15N en dos tipos de suelos.

Se realizó un experimento para cuantificar las proporciones de 14N a 15N en dos tipos de suelos (urbano, rural) con tres espectrómetros distintos. De cada suelo se eligieron 18 muestras. A cada una de las 18 muestras del suelo urbano se la asignó uno de los tres espectrómetros al azar de tal forma que con cada espectrómetro se analizaron 6 de ellas, y, de la misma forma, se asignaron las muestras del suelo rural. Por restricciones operativas para cada combinación de espectrómetro y suelo se eligieron 3 días al azar para medir en cada fecha la proporción de 14N a 15N.

En este diseño intervienen el factor **suelo** y el factor **espectrometro**. Además, como el **suelo** se analiza en los tres **espectrometros**, estos dos factores estarán cruzados y podemos resolver en el modelo la interacción. De hecho interesa comprobar si la interacción es significativa, y si fuera este el caso, la diferencia entre los tres espectrometros dependería del tipo de suelo. Por otra parte, al tomar fechas al azar para cuantificar la proporción hay que tener en cuenta el factor **día**. Este factor es de efectos aleatorios, al contrario que los factores **suelo** y **espectrometro** que son de efectos fijos. El factor **día** está anidado a la combinación de **suelo** y **espectrometro**.

Los datos (`dades_problema1_pec6.csv`) son las proporciones 14N a 15N (multiplicadas por 1000).

Pedimos:

1. Realizar un análisis descriptivo inicial de los datos. Comentar los resultados.
2. Formular detalladamente un modelo lineal para el experimento.
3. Escribir las hipótesis de interés correspondientes a los parámetros del modelo lineal.
4. ¿Hay diferencias significativas entre espectrómetros? ¿y entre suelos? ¿Es significativa la interacción?
5. En su caso, efectuar las comparaciones múltiples.

### Problema 2. Complementos vitamínicos.

Un investigador realizó un experimento para evaluar los efectos de cuatro complementos vitamínicos en el aumento de peso de animales de laboratorio. El experimento se realizó con un diseño totalmente aleatorizado, con 20 animales, en jaulas separadas, para cada tratamiento. Como el consumo de calorías difiere entre los animales e influye en el aumento de peso, el investigador lo midió en cada animal. Los datos del incremento de peso (gramos), la covariable de consumo de calorías (calorías/10) y la dieta según el complemento vitamínico se encuentran en el fichero `dades_problema2_pec6.csv`.

Pedimos:

1. Realizar un análisis descriptivo para visualizar la relación entre consumo calórico y el incremento de peso. Comenta los resultados.
2. Formular el modelo lineal para modelar el efecto de la dieta en el incremento de peso, ajustando por el consumo calórico y sin ajustar.
3. Estimar el efecto de la dieta en el incremento de peso, con los dos modelos anteriores. Comparar las estimaciones según los dos modelos de las dietas 1 y 2.
4. Realizar la diagnosis de las suposiciones de modelización.
5. Determinar entre que dietas hay diferencias significativas.

### Problema 3. Tratamiento con estrógenos para la depresión postparto.

Los datos surgen de un estudio clínico doble ciego controlado con placebo para el ensayo de un tratamiento de estrógenos para la depresión postparto. La variable principal fue una medida de depresión, registrada una ocasión antes de la aleatorización al tratamiento y en seis visitas bimensuales después de la aleatorización. A todas las mujeres del ensayo se les registró la variable depresión en las siete visitas programadas. Utiliza modelos de efectos aleatorios apropiados para investigar estos datos, en particular para determinar si el tratamiento activo con estrógenos ayuda a reducir la depresión postparto. Los datos están en el fichero `dades_problema3_pec6.csv`.

Pedimos:

1. Realizar un análisis descriptivo para evidenciar si hay diferencias a nivel de intercepciones y pendientes en la relación entre la variable de depresión y el tiempo, en los pacientes. Discute los resultados.
2. Formular el modelo lineal mixto para contrastar el efecto del tratamiento en la depresión. Plantea el modelo de intercepciones y pendientes aleatorias. Describe la matriz de regresión de los efectos fijos y la de los efectos aleatorios.
3. Implementar los modelos lineales mixtos de intercepciones aleatorias y de intercepciones y pendientes aleatorias. ¿Qué modelo resulta más adecuado?
4. Con el modelo elegido en 3), ¿Qué factores han resultado significativos?
5. Con el modelo elegido en 3), proporcionar el modelo estimado. En particular, interpreta los modelos obtenidos en los pacientes 1 y 26.