

Tema 4: Análisis de datos. Ejercicios resueltos.

1. Se está poniendo a prueba un proceso que en fotobiología se denomina abscisión, con la esperanza de aumentar la cosecha de fruta en los naranjos. El proceso implica exponer los árboles a luz coloreada durante quince minutos cada noche. Se recolectó fruta de 10 árboles experimentales, bajo condiciones normales primero y luego después del nuevo tratamiento. Se obtuvieron las siguientes observaciones:

Conds. normales	100	50	98	26	65	95	86	100	108	60	175
Nuevo tratamiento	129	60	141	56	150	100	102	126	111	59	179

El promotor del nuevo proceso pretende comprobar que éste incrementa la recolección. Realiza el análisis estadístico oportuno para averiguar si estos datos respaldan la pretensión del promotor.

Solución: tenemos un problema de comparación de la media de dos muestras emparejadas unilateral (porque nos piden que un grupo tenga la media mayor que el otro).

```
nor <-c(100,50,98,26,65,95,86,100,108,60,175)
trat<-c(129,60,141,56,150,100,102,126,111,59,179)
```

Como son muestras emparejadas comprobamos si la variable diferencia es normal:

```
dif<-nor-trat
shapiro.test(dif); ks.test(dif,"pnorm",mean(dif),sd(dif))
```

Como podemos observar un test nos dice que no hay normalidad mientras que el otro nos dice que sí. Para hacer un estudio más detallado se puede observar los gráficos qq:

```
qqnorm(dif); qqline(dif)
```

en los que vemos que efectivamente las diferencias no son normales.

Utilizamos pues una prueba no paramétrica ya que tenemos pocos datos:

```
wilcox.test(nor, trat, paired = TRUE, alternative = "less")
```

Como el p-valor 0.0009766 es menor que 0.05 se rechaza la hipótesis de que la mediana de la recolección en condiciones normales es mayor que con el nuevo tratamiento y por lo tanto pensamos que el nuevo método incrementa la recolección.

Si pensáramos que sí que hay normalidad entonces haríamos un test t:

```
t.test(nor, trat, paired=T, alternative="less")
```

En ese caso el p-valor 0.006411 también es menor que 0.05 y la conclusión es similar salvo que podemos decir que el incremento medio es mayor en el nuevo método.

2. Las manadas de lobos son territoriales, con territorios de 130 km² o más. Se piensa que los aullidos de los lobos, que comunican información tanto de la situación como de la composición de la manada, están relacionados con la territorialidad. Se obtuvieron los siguientes valores de la variable X , duración en minutos de una sesión de aullidos de una determinada manada sometida a estudio: 1.0, 1.8, 1.6, 1.5, 2.0, 1.8, 1.2, 1.9, 1.7, 1.6, 1.6, 1.7, 1.5, 1.4, 1.4 y 1.4 ¿Confirman estos datos que la duración media es superior a 1.5 minutos?

Solución: tenemos un problema de una muestra en el que queremos analizar si la media es superior a un valor. Contraste unilateral. Tras comprobar que los datos son normales hacemos el test t habitual.

```
X<-c(1.0,1.8,1.6,1.5,2.0,1.8,1.2,1.9,1.7,1.6,1.6,1.7,1.5,1.4,1.4,1.4)
shapiro.test(X); ks.test(X,"pnorm",mean(X),sd(X))
t.test(X, mu=1.5, alternative = "greater")
```

Como el p-valor 0.15 es mayor que 0.05, no tenemos evidencia estadística suficiente para rechazar la hipótesis nula.

3. Normalmente las hojas de la Mimosa púdica son horizontales. Si se toca ligeramente una de ellas, las hojas se pliegan. Se afirma que el tiempo medio desde el contacto hasta el cierre completo es de 2.5 segundos. En un experimento dirigido a comprobar dicha hipótesis se han obtenido las siguientes observaciones: 3.0, 2.9, 2.8, 2.7, 2.6, 2.4, 2.5, 2.4, 2.6 y 2.7:

- Discutir la normalidad de estos datos.
- Plantear y resolver un contraste de hipótesis adecuado para comprobar si el valor 2.5 es en verdad el tiempo medio hasta el cierre o si por el contrario es diferente.
- Extraer las conclusiones oportunas.

Ayuda: `ks.test(mimosa); t.test(mimosa, mu=2.5)`