

# Lección 4: Estadística inferencial

## Bioestadística con R

### Ejercicios

**Ejercicio 1:** En un estudio, se midió el desarrollo del timo. Los investigadores pesaron 10 embriones de gallinas. Cinco de los embriones se incubaron durante 14 días y otros cinco se incubaron 15 días. Los pesos fueron los siguientes:

Table 1: Pesos de las gallinas (en g)

14 días	15 días
29.6	32.7
21.5	40.3
28.0	23.7
34.6	25.2
44.9	24.2

- Obtener la media ( $\bar{x}$ ) y desviación estándar ( $s$ ).
- Realizar una prueba de  $t$  para comprar las medias, con un  $\alpha = 0.10$ .
- ¿Existen diferencias estadísticamente significativas entre los dos grupos?

**Ejercicio 2:** La siguiente tabla muestra el número de colonias de bacterias presentes en distintos platos de Petri después de haber sido inoculados con *E. coli* e incubados por 24 horas. El tratamiento “jabón” son aquellos que contienen una solución basada en jabón, mientras que “control” son aquellos que contienen agua destilada.

Table 2: Colonias de *E. coli*

Control	Jabón
30	76
36	27
66	16
21	30
63	26
38	46
35	6
45	

- Realizar una prueba de  $t$  con un  $\alpha = 0.05$ .

## Soluciones

**Ejercicio 1:** Para este ejercicio utilizaremos la función `t.test()`. Primero necesitamos crear un data frame con nuestros datos.

```
days <- c(rep("14d", 5), rep("15d", 5))
peso <- c(29.6, 21.5, 28.0, 34.6, 44.9, 32.7, 40.3, 23.7, 25.2, 24.2)
ej1 <- data.frame(days, peso)
ej1
```

days	peso
14d	29.6
14d	21.5
14d	28.0
14d	34.6
14d	44.9
15d	32.7
15d	40.3
15d	23.7
15d	25.2
15d	24.2

Es *muy importante* que tengamos 2 columnas, una con nuestro grupo y otra con el peso. Posterior a esto podemos calcular la media. Para esto vamos a echarnos una manita con la librería `dplyr`. Llamaremos a nuestra variable `desc`, porque son nuestros estadísticos descriptivos.

```
library(dplyr)
desc <- ej1 %>% group_by(days) %>% summarise(media = mean(peso), sd = sd(peso))
desc
```

days	media	sd
14d	31.72	8.729089
15d	29.22	7.188672

Podemos ahora utilizar la función `t.test()` y obtener nuestro resultado. Sin embargo, como nuestro  $\alpha = 0.10$  necesitamos añadir el argumento `conf.level` para indicar que queremos un nivel de confianza del 0.9.

```
t.test(peso ~ days, ej1, conf.level = 0.9)
```

```
##
##  Welch Two Sample t-test
##
## data:  peso by days
## t = 0.49435, df = 7.7163, p-value = 0.6348
## alternative hypothesis: true difference in means between group 14d and group 15d is not equal to 0
## 90 percent confidence interval:
##  -6.949042 11.949042
## sample estimates:
## mean in group 14d mean in group 15d
##           31.72           29.22
```

Como podemos ver en este caso, no existe suficiente información para rechazar la hipótesis nula ( $H_0 : \mu_1 = \mu_2$ ), por lo que decimos que se acepta.

**Ejercicio 2:** Simplemente utilizamos la función `t.test()`, sin agregar el argumento `conf.level`.

```
tratamiento <- c(rep("control", 8), rep("jabon", 7))
colonias <- c(30, 36, 66, 21, 63, 38, 35, 45, 76, 27, 16, 30, 26, 46, 6)
ej2 <- data.frame(tratamiento, colonias)
ej2
```

tratamiento	colonias
control	30
control	36
control	66
control	21
control	63
control	38
control	35
control	45
jabon	76
jabon	27
jabon	16
jabon	30
jabon	26
jabon	46
jabon	6

Ahora que tenemos nuestros datos, simplemente hacemos una prueba de  $t$  usando la función `t.test()`.

```
t.test(colonias ~ tratamiento, ej2)
```

```
##
##  Welch Two Sample t-test
##
## data:  colonias by tratamiento
## t = 0.90954, df = 10.43, p-value = 0.3836
## alternative hypothesis: true difference in means between group control and group jabon is not equal
## 95 percent confidence interval:
##  -13.38678  32.02964
## sample estimates:
## mean in group control    mean in group jabon
##           41.75000           32.42857
```