

# HW\_03.R

jairo

2025-09-22

```
# HW_03
# Ejercicio: comparación de concentraciones de estroncio en cuerpos de agua
# Jairo Alberto Leal Gómez
# Matricula 1723093
# DCOMRN 1 semestre - Estadística en la Investigación Científica

# Descripción -----

# Un investigador midió la concentración de estroncio (mg/ml) en cinco cuerpos
# de agua para evaluar si existen diferencias significativas entre ellos.
# Cada sitio tuvo 6 réplicas (n=6).

estroncio <- read.csv("estroncio.csv", header = TRUE)
estroncio
```

```
##      graysonpond beaverlake anglercove appletreelake rockriver
## 1          28.2         39.6         46.3          41.0         56.3
## 2          33.2         40.8         42.1          44.1         54.1
## 3          36.4         37.9         43.5          46.4         59.4
## 4          34.6         37.1         48.8          40.2         62.7
## 5          29.1         43.6         43.7          38.6         60.0
## 6          31.0         42.4         40.1          36.3         57.3
```

```
estroncio2 <- read.csv("estroncio2.csv", header = TRUE)
estroncio2
```

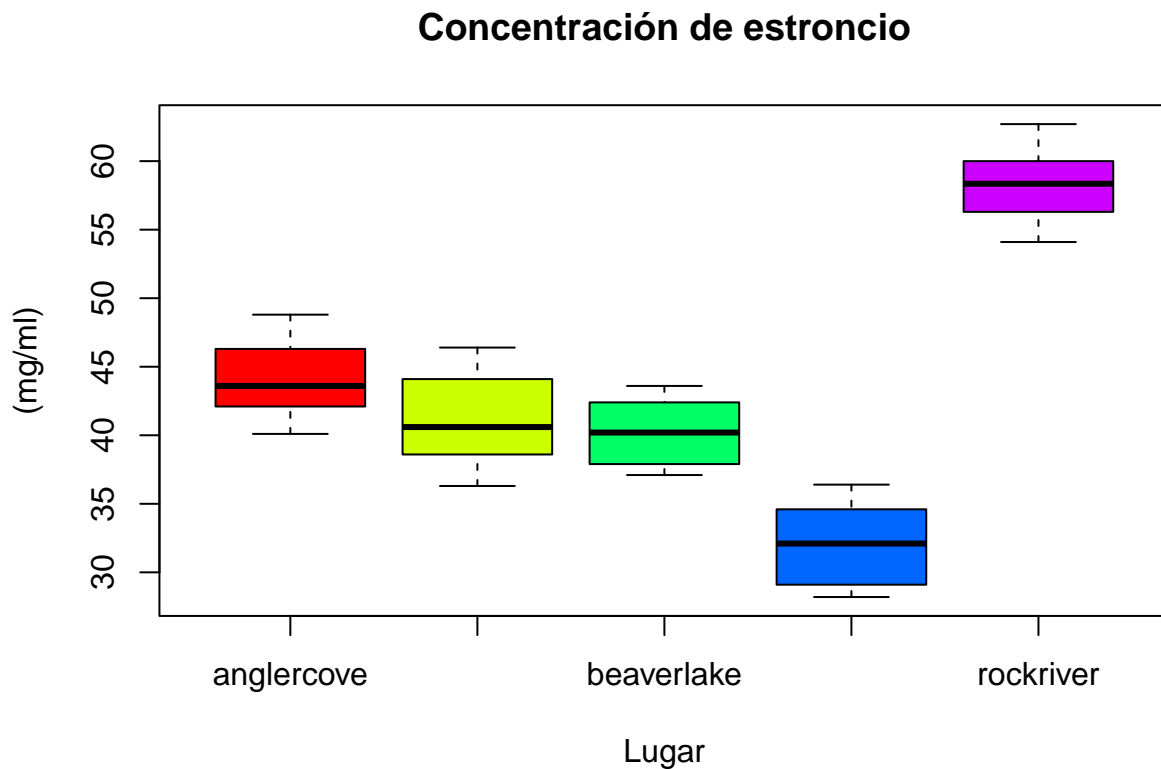
```
##      sitio concentracion
## 1  graysonpond         28.2
## 2  graysonpond         33.2
## 3  graysonpond         36.4
## 4  graysonpond         34.6
## 5  graysonpond         29.1
## 6  graysonpond         31.0
## 7   beaverlake         39.6
## 8   beaverlake         40.8
## 9   beaverlake         37.9
## 10  beaverlake         37.1
## 11  beaverlake         43.6
## 12  beaverlake         42.4
## 13  anglercove         46.3
## 14  anglercove         42.1
## 15  anglercove         43.5
## 16  anglercove         48.8
## 17  anglercove         43.7
```

```
## 18    anglercove      40.1
## 19 appletreelake     41.0
## 20 appletreelake     44.1
## 21 appletreelake     46.4
## 22 appletreelake     40.2
## 23 appletreelake     38.6
## 24 appletreelake     36.3
## 25    rockriver      56.3
## 26    rockriver      54.1
## 27    rockriver      59.4
## 28    rockriver      62.7
## 29    rockriver      60.0
## 30    rockriver      57.3
```

*# Planteamiento del problema -----*

```
estroncio_ordenado <- estroncio[, order(names(estroncio))]
```

```
boxplot(estroncio_ordenado,
        xlab = "Lugar",
        ylab = "(mg/ml)",
        main = "Concentración de estroncio",
        col = rainbow(5))
```



*# La Figura 1 muestra las concentraciones de estroncio (mg/ml) registradas  
# en cinco cuerpos de agua: Grayson's Pond, Beaver Lake, Angler's Cove,  
# Appletree Lake y Rock River.*

```

# Cada sitio contó con seis repeticiones independientes.

# Se observa que Rock River presenta las concentraciones más elevadas,
# con valores consistentemente superiores al resto de los sitios.
# En contraste, Grayson's Pond mostró los niveles más bajos, mientras que
# Beaver Lake, Angler's Cove y Appletree Lake presentaron valores intermedios
# y relativamente cercanos entre sí.

# Este patrón sugiere la existencia de diferencias significativas entre sitios,
# lo cual motiva la aplicación de un ANOVA de una vía, seguido de
# pruebas post-hoc (LSD y Tukey HSD) para identificar con precisión
# qué grupos difieren estadísticamente en sus medias.

# HIPOTESIS DE LA ANOVA

# H0 = Todas las medias de concentración de estroncio en los cinco cuerpos de agua
# son iguales. No existe diferencia significativa entre los sitios.

# H1 = Al menos un sitio presenta una media distinta de concentración de
# estroncio respecto a los demás. Existen diferencias significativas.

estroncio2

```

```

##          sitio concentracion
## 1    graysonpond      28.2
## 2    graysonpond      33.2
## 3    graysonpond      36.4
## 4    graysonpond      34.6
## 5    graysonpond      29.1
## 6    graysonpond      31.0
## 7    beaverlake      39.6
## 8    beaverlake      40.8
## 9    beaverlake      37.9
## 10   beaverlake      37.1
## 11   beaverlake      43.6
## 12   beaverlake      42.4
## 13   anglercove      46.3
## 14   anglercove      42.1
## 15   anglercove      43.5
## 16   anglercove      48.8
## 17   anglercove      43.7
## 18   anglercove      40.1
## 19   appletreelake    41.0
## 20   appletreelake    44.1
## 21   appletreelake    46.4
## 22   appletreelake    40.2
## 23   appletreelake    38.6
## 24   appletreelake    36.3
## 25    rockriver      56.3
## 26    rockriver      54.1
## 27    rockriver      59.4
## 28    rockriver      62.7
## 29    rockriver      60.0

```

```
## 30      rockriver      57.3
# PRUEBA DE NORMALIDAD

# p-value = 0.1052 no hay diferencias significativas datos normales
shapiro.test(estroncio2$concentracion)

##
## Shapiro-Wilk normality test
##
## data:  estroncio2$concentracion
## W = 0.94235, p-value = 0.1052

# p values mayores a 0.05, no hay diferencias significativas en cada sitio
shapiro.test(subset(estroncio2$concentracion, estroncio2$sitio == "graysonpond"))

##
## Shapiro-Wilk normality test
##
## data:  subset(estroncio2$concentracion, estroncio2$sitio == "graysonpond")
## W = 0.95674, p-value = 0.7943

shapiro.test(subset(estroncio2$concentracion, estroncio2$sitio == "beaverlake"))

##
## Shapiro-Wilk normality test
##
## data:  subset(estroncio2$concentracion, estroncio2$sitio == "beaverlake")
## W = 0.96163, p-value = 0.8322

shapiro.test(subset(estroncio2$concentracion, estroncio2$sitio == "anglercove"))

##
## Shapiro-Wilk normality test
##
## data:  subset(estroncio2$concentracion, estroncio2$sitio == "anglercove")
## W = 0.97181, p-value = 0.9044

shapiro.test(subset(estroncio2$concentracion, estroncio2$sitio == "appletreelake"))

##
## Shapiro-Wilk normality test
##
## data:  subset(estroncio2$concentracion, estroncio2$sitio == "appletreelake")
## W = 0.9784, p-value = 0.9433

shapiro.test(subset(estroncio2$concentracion, estroncio2$sitio == "rockriver"))

##
## Shapiro-Wilk normality test
##
## data:  subset(estroncio2$concentracion, estroncio2$sitio == "rockriver")
## W = 0.98937, p-value = 0.9876

# PRUEBA DE ANOVA

estroncio.aov <- aov(estroncio2$concentracion ~ estroncio2$sitio)
summary(estroncio.aov)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## estroncio2$sitio  4 2193.4    548.4    56.16 3.95e-12 ***
## Residuals       25  244.1      9.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Con un valor de p value de 3.95e-12 ***, se demuestra que existen diferencias
# significativas en todas las concentraciones de estroncio entre todos los sitios.
# Se rechaza la HIPOTESIS NULA.

# Calcule el valor de LSD con alpha = 0.05.

qt(0.975, df= 25)

## [1] 2.059539

sqrt((2*9.8)/6) * qt(0.975, df= 25)

## [1] 3.722394

# Compare las diferencias entre medias de los sitios

tapply(estroncio2$concentracion, estroncio2$sitio, mean)

##      anglercove appletreelake    beaverlake    graysonpond    rockriver
##      44.08333      41.10000      40.23333      32.08333      58.30000

# Determine cuales pares son significativamente diferentes

# LSD = 3.722394

# El supuesto es que si el resultado entre la diferencia de medias, supera el LSD
# existen diferencias significativas

# AnglerCove vs AppletreeLake
44.08333 - 41.10000

## [1] 2.98333

# 2.98333 no supera el valor de LSD, NO EXISTE DIF SIGNIFICATIVAS

# AnglerCove vs BeaverLake
44.08333 - 40.23333

## [1] 3.85

# 3.85 Supera el valor de LSD, DIFRENECIAS SIGNIFICATIVAS

# AnglerCove vs GraysonPond
44.08333 - 32.08333

## [1] 12

# 12 Supera el valor de LSD, DIFRENECIAS SIGNIFICATIVAS

# AnglerCove vs RockRiver
44.08333 - 58.30000

## [1] -14.21667
```

```

# 14.21 Supera el valor de LSD, DIFERENCIAS SIGNIFICATIVAS

# AppletreeLake vs BeaverLake
41.10000 - 40.23333

## [1] 0.86667
# 0.86667 no supera el valor de LSD, NO EXISTE DIF SIGNIFICATIVAS

# AppletreeLake vs GraysonPond
41.10000 - 32.08333

## [1] 9.01667
# 9.01667 Supera el valor de LSD, DIFERENCIAS SIGNIFICATIVAS

# AppletreeLake vs RockRiver
41.10000 - 58.30000

## [1] -17.2
# 17.2 Supera el valor de LSD, DIFERENCIAS SIGNIFICATIVAS

# BeaverLake vs GraysonPond
40.23333 - 32.08333

## [1] 8.15
# 8.15 Supera el valor de LSD, DIFERENCIAS SIGNIFICATIVAS

# BeaverLake vs RockRiver
40.23333 - 58.30000

## [1] -18.06667
# 18.06667 Supera el valor de LSD, DIFERENCIAS SIGNIFICATIVAS

# GraysonPond vs RockRiver
32.08333 - 58.30000

## [1] -26.21667
# 26.21667 Supera el valor de LSD, DIFERENCIAS SIGNIFICATIVAS

# Solo AnglerCove vs AppletreeLake y AppletreeLake vs BeaverLake
# no muestran diferencias significativas.

# Todos los demás pares superan el LSD y son significativamente diferentes.

# PRUEBA DE TUKEY HSD

# Obtenga el valor critico

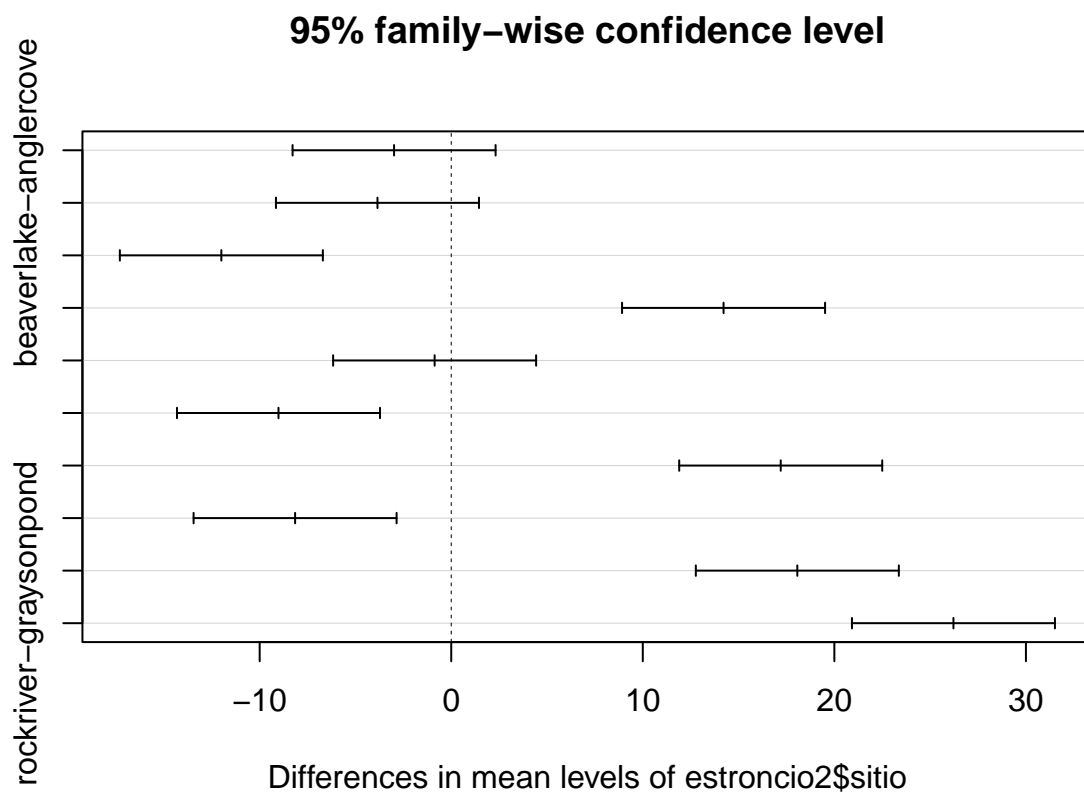
TukeyHSD(estroncio.aov)

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = estroncio2$concentracion ~ estroncio2$sitio)
##

```

```
## $`estroncio2$sitio`
##               diff      lwr      upr    p adj
## appletreelake-anglercove -2.9833333 -8.281979  2.315312 0.4791100
## beaverlake-anglercove    -3.8500000 -9.148645  1.448645 0.2376217
## graysonpond-anglercove  -12.0000000 -17.298645 -6.701355 0.0000053
## rockriver-anglercove     14.2166667  8.918021 19.515312 0.0000003
## beaverlake-appletreelake -0.8666667 -6.165312  4.431979 0.9884803
## graysonpond-appletreelake -9.0166667 -14.315312 -3.718021 0.0003339
## rockriver-appletreelake   17.2000000 11.901355 22.498645 0.0000000
## graysonpond-beaverlake    -8.1500000 -13.448645 -2.851355 0.0011293
## rockriver-beaverlake     18.0666667 12.768021 23.365312 0.0000000
## rockriver-graysonpond     26.2166667 20.918021 31.515312 0.0000000
```

```
plot(TukeyHSD(estroncio.aov))
```



```
# Calcule la diferencia minima significativa con Tukey
qtukey(0.95, nmeans = 5, df = 25)
```

```
## [1] 4.153363
```

```
sqrt((9.8/6)) * qtukey(0.95, nmeans = 5, df = 25)
```

```
## [1] 5.308078
```

```
# Compare los resultados con la prueba LSD
```

```
# DMS = 5.308078
```

```
# El supuesto es que si el resultado entre la diferencia de medias, supera el DMS
# existen diferencias significativas
```

```
# AnglerCove vs AppletreeLake
44.08333 - 41.10000
```

```
## [1] 2.98333
```

```
# 2.98333 no supera el valor de DMS, NO EXISTE DIF SIGNIFICATIVAS
```

```
# AnglerCove vs BeaverLake
44.08333 - 40.23333
```

```
## [1] 3.85
```

```
# 3.85 no supera el valor de DMS, NO EXISTE DIF SIGNIFICATIVAS
```

```
# AnglerCove vs GraysonPond
44.08333 - 32.08333
```

```
## [1] 12
```

```
# 12 Supera el valor de DMS, DIFRENECIAS SIGNIFICATIVAS
```

```
# AnglerCove vs RockRiver
44.08333 - 58.30000
```

```
## [1] -14.21667
```

```
# 14.21 Supera el valor de DMS, DIFRENECIAS SIGNIFICATIVAS
```

```
# AppletreeLake vs BeaverLake
41.10000 - 40.23333
```

```
## [1] 0.86667
```

```
# 0.86667 no supera el valor de DMS, NO EXISTE DIF SIGNIFICATIVAS
```

```
# AppletreeLake vs GraysonPond
41.10000 - 32.08333
```

```
## [1] 9.01667
```

```
# 9.01667 Supera el valor de DMS, DIFRENECIAS SIGNIFICATIVAS
```

```
# AppletreeLake vs RockRiver
41.10000 - 58.30000
```

```
## [1] -17.2
```

```
# 17.2 Supera el valor de DMS, DIFRENECIAS SIGNIFICATIVAS
```

```
# BeaverLake vs GraysonPond
40.23333 - 32.08333
```

```
## [1] 8.15
```

```
# 8.15 Supera el valor de DMS, DIFRENECIAS SIGNIFICATIVAS
```

```
# BeaverLake vs RockRiver
40.23333 - 58.30000
```



```
## [1] -18.06667
```

```
# 18.06667 Supera el valor de DMS, DIFERENCIAS SIGNIFICATIVAS
```

```
# GraysonPond vs RockRiver
```

```
32.08333 - 58.30000
```

```
## [1] -26.21667
```

```
# 26.21667 Supera el valor de DMS, DIFERENCIAS SIGNIFICATIVAS
```

```
# ¿Los mismos pares resultan significativos?
```

```
# En este caso, a excepcion de AnglerCove vs BeaverLake que ahora no presenta  
# diferencias significativas, todos los demas siguen igual
```

```
# INTERPRETACION -----
```

```
# ¿ QUE CUERPO DE AGUA PRESENTA LAS CONCENTRACIONES MAS ALTAS ?
```

```
# Rock River, con un promedio de 58.3 mg/ml, es el sitio con las concentraciones  
# más altas de estroncio.
```

```
# En contraste, GraysonPond tiene las concentraciones más bajas (32.08 mg/ml).
```

```
tapply(estroncio2$concentracion, estroncio2$sitio, mean)
```

```
##      anglercove appletreelake    beaverlake  graysonpond    rockriver  
##      44.08333      41.10000      40.23333      32.08333      58.30000
```

```
# ¿ QUE SITIOS NO DIFIEREN ENTRE SI ?
```

```
# Según la prueba LSD ( $\alpha = 0.05$ ,  $LSD = 3.722$ ):
```

```
# AnglerCove vs AppletreeLake → no difieren ( $2.983 < LSD$ )
```

```
# AppletreeLake vs BeaverLake → no difieren ( $0.867 < LSD$ )
```

```
# Según la prueba Tukey HSD ( $DMS = 5.308$ ), más conservadora:
```

```
# AnglerCove vs AppletreeLake → no difieren ( $2.983 < DMS$ )
```

```
# AppletreeLake vs BeaverLake → no difieren ( $0.867 < DMS$ )
```

```
# AnglerCove vs BeaverLake → ahora tampoco difieren ( $3.85 < DMS$ )
```

```
# Los sitios intermedios (BeaverLake, AppletreeLake, AnglerCove)
```

```
# presentan algunas diferencias no significativas entre ellos.
```

```
# RockRiver y GraysonPond siempre difieren significativamente de los demás.
```

```
# DESDE EL PUNTO DE VISTA AMBIENTAL ¿ QUE IMPLICACIONES PODRIAN TENER ESTAS  
# DIFERENCIAS EN LA CALIDAD DEL AGUA?
```

```
# En Rock River, al tener concentraciones mas altas de estroncio, puede indicar
```

# una mayori disolucion de este mineral en suelo y rocas por la esorrentia o  
# presencia de agua, o por actividades mineras o procesamiento de este.

# En concentraciones mayores o sobre todo del estroncio isotopo 90 puede  
# representar un mayor riesgo a la salud humana y ambiental de la fauna, al  
# llegar a ser radioactivo

# Los sitios con menores concentraciones o intermedias, pudieran estar dentro  
# de los rangos normales o naturales.