

Comparación de concentraciones de estroncio en cuerpos de agua

Andrea Michelle Luna Vasconcelos – 1950889

22/09/2025

Descripción de los datos

```
estroncio <- read.csv("C:/Repositorio GitHub/Posgrado_Estadistica_2025/Tarea 22_09/Estroncio mg_ml.csv")
library(knitr)
kable(head(estroncio), caption = "Concentración de estroncio (mg/ml) en cinco cuerpos de agua (n = 6)")
```

Concentración de estroncio (mg/ml) en cinco cuerpos de agua (n = 6)

Muestra	Grayson.s.Pond	Beaver.Lake	Angler.s.Cove	Appletree.Lake	Rock.River
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

```
library(tidyverse)
```

```
estroncio_long <- estroncio %>%
  pivot_longer(cols = -Muestra,
               names_to = "Cuerpo_agua",
               values_to = "Concentracion")
```

```
estroncio_long$Cuerpo_agua <- as.factor(estroncio_long$Cuerpo_agua)
```

```
kable(estroncio_long, caption = "Datos reorganizados de concentraciones de estroncio (mg/ml) en cinco cuerpos de agua como factor")
```

Datos reorganizados de concentraciones de estroncio (mg/ml) en cinco cuerpos de agua como factor

Muestra	Cuerpo_agua	Concentracion
1	Grayson.s.Pond	28.2
1	Beaver.Lake	39.6
1	Angler.s.Cove	46.3
1	Appletree.Lake	41.0

Muestra	Cuerpo_agua	Concentracion
1	Rock.River	56.3
2	Grayson.s.Pond	33.2
2	Beaver.Lake	40.8
2	Angler.s.Cove	42.1
2	Appletree.Lake	44.1
2	Rock.River	54.1
3	Grayson.s.Pond	36.4
3	Beaver.Lake	37.9
3	Angler.s.Cove	43.5
3	Appletree.Lake	46.4
3	Rock.River	59.4
4	Grayson.s.Pond	34.6
4	Beaver.Lake	37.1
4	Angler.s.Cove	48.8
4	Appletree.Lake	40.2
4	Rock.River	62.7
5	Grayson.s.Pond	29.1
5	Beaver.Lake	43.6
5	Angler.s.Cove	43.7
5	Appletree.Lake	38.6
5	Rock.River	60.0
6	Grayson.s.Pond	31.0
6	Beaver.Lake	42.4
6	Angler.s.Cove	40.1
6	Appletree.Lake	36.3
6	Rock.River	57.3

```
medias <- tapply(estroncio_long$Concentracion, estroncio_long$Cuerpo_agua, mean)
kable(medias, caption = "Medias de concentración de estroncio (mg/ml) por cuerpo de agua",
      col.names = c("Cuerpo de agua", "Media (mg/ml)"))
```

Medias de concentración de estroncio (mg/ml) por cuerpo de agua

Cuerpo de agua	Media (mg/ml)
Angler.s.Cove	44.08333
Appletree.Lake	41.10000
Beaver.Lake	40.23333
Grayson.s.Pond	32.08333
Rock.River	58.30000

Hipótesis del ANOVA

- **H0:** La media de concentración de estroncio en todos los cuerpos de agua es igual.
- **H1:** Al menos una media de concentración de estroncio es diferente.

Cálculo del ANOVA

```
bartlett.test(estroncio_long$Concentracion ~ estroncio_long$Cuerpo_agua)

##
## Bartlett test of homogeneity of variances
##
## data:  estroncio_long$Concentracion by estroncio_long$Cuerpo_agua
## Bartlett's K-squared = 0.63895, df = 4, p-value = 0.9586

estroncio_long.aov <- aov(estroncio_long$Concentracion ~ estroncio_long$C
uerpo_agua)
summary(estroncio_long.aov)

##
##              Df Sum Sq Mean Sq F value    Pr(>F)
## estroncio_long$Cuerpo_agua  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

library(broom)
anova <- tidy(estroncio_long.aov)
kable(anova, caption = "Resultados de análisis de varianza (ANOVA de una
vía)")
```

Resultados de análisis de varianza (ANOVA de una vía)

term	df	sumsq	meansq	statistic	p.value
estroncio_long\$Cuerpo_agua	4	2193.442	548.3605	56.15456	0
Residuals	25	244.130	9.7652	NA	NA

- H_0 = (medias iguales) = se rechaza
- H_1 = (al menos una media distinta) = se acepta

Gráfica de distribución

```
library(ggplot2)
```

```
ggplot(estruncio_long, aes(x = Cuerpo_agua, y = Concentracion, fill = Cuerpo_agua)) +
  geom_violin(trim = FALSE, alpha = 0.5) +
  geom_boxplot(width = 0.1, fill = "white", outlier.shape = NA) +
  geom_jitter(width = 0.1, size = 2, alpha = 0.7) +
  labs(
    title = "Concentraciones de estruncio en cuerpos de agua",
    x = "Cuerpo de agua",
    y = "Concentración (mg/ml)",
    fill = "Cuerpos de agua",
    caption = "Figura 1. Concentraciones de estruncio en cinco cuerpos de agua"
  ) +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5),
        plot.caption = element_text(hjust = 0.5))
```

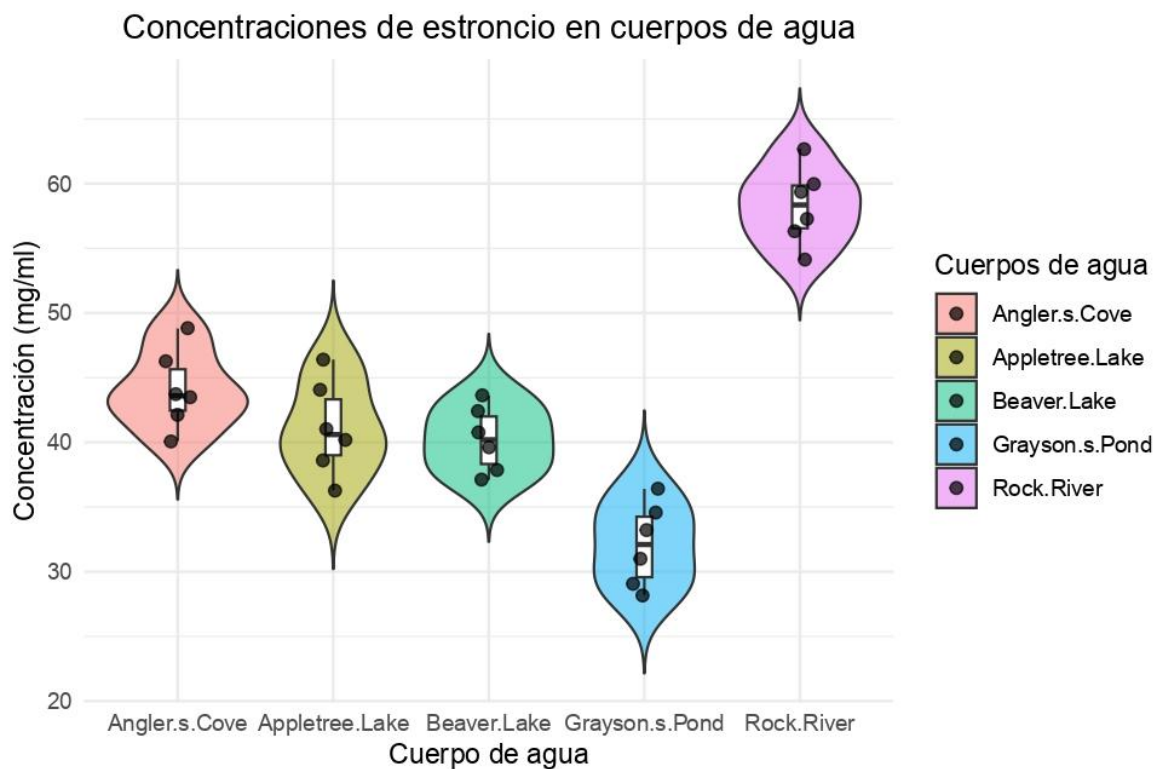


Figura 1. Concentraciones de estruncio en cinco cuerpos de agua

Prueba LSD

```
LSD <- sqrt((2*9.8)/6) * qt(0.975, 25)
LSD
```

```
## [1] 3.722394
```

Comparación por pares con LSD:

```
resultados <- data.frame(
  Comparacion = c(
    "Grayson's Pond vs Beaver Lake",
    "Grayson's Pond vs Angler's Cove",
    "Grayson's Pond vs Appletree Lake",
    "Grayson's Pond vs Rock River",
    "Beaver Lake vs Angler's Cove",
    "Beaver Lake vs Appletree Lake",
    "Beaver Lake vs Rock River",
    "Angler's Cove vs Appletree Lake",
    "Angler's Cove vs Rock River",
    "Appletree Lake vs Rock River"
  ),
  Diferencia = c(-8.15, -12.00, -9.017, -26.217, -3.85, -0.867, -18.067,
2.983, -14.217, -17.2),
  Significancia = c("Significativa", "Significativa", "Significativa", "S
ignificativa",
                    "Significativa", "No significativa", "Significativa",
                    "No significativa", "Significativa", "Significativa")
)
kable(resultados, caption = "Resultados LSD de comparación entre pares de
medias",
      col.names = c("Pares de medias", "Diferencia (mg/ml)", "Resultado L
SD"))
```

Resultados LSD de comparación entre pares de medias

Pares de medias	Diferencia (mg/ml)	Resultado LSD
Grayson's Pond vs Beaver Lake	-8.150	Significativa
Grayson's Pond vs Angler's Cove	-12.000	Significativa
Grayson's Pond vs Appletree Lake	-9.017	Significativa
Grayson's Pond vs Rock River	-26.217	Significativa
Beaver Lake vs Angler's Cove	-3.850	Significativa
Beaver Lake vs Appletree Lake	-0.867	No significativa
Beaver Lake vs Rock River	-18.067	Significativa
Angler's Cove vs Appletree Lake	2.983	No significativa
Angler's Cove vs Rock River	-14.217	Significativa
Appletree Lake vs Rock River	-17.200	Significativa

Prueba Tukey HSD

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

## [1] 5.308078

TukeyHSD(estroncio_long.aov)

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = estroncio_long$Concentracion ~ estroncio_long$Cuerpo
o_agua)
##
## $`estroncio_long$Cuerpo_agua`
##              diff              lwr              upr              p a
dj
## Appletree.Lake-Angler.s.Cove -2.9833333 -8.281979  2.315312 0.47911
00
## Beaver.Lake-Angler.s.Cove    -3.8500000 -9.148645  1.448645 0.23762
17
## Grayson.s.Pond-Angler.s.Cove -12.0000000 -17.298645 -6.701355 0.00000
53
## Rock.River-Angler.s.Cove     14.2166667  8.918021 19.515312 0.00000
03
## Beaver.Lake-Appletree.Lake   -0.8666667 -6.165312  4.431979 0.98848
03
## Grayson.s.Pond-Appletree.Lake -9.0166667 -14.315312 -3.718021 0.00033
39
## Rock.River-Appletree.Lake     17.2000000 11.901355 22.498645 0.00000
00
## Grayson.s.Pond-Beaver.Lake    -8.1500000 -13.448645 -2.851355 0.00112
93
## Rock.River-Beaver.Lake       18.0666667 12.768021 23.365312 0.00000
00
## Rock.River-Grayson.s.Pond    26.2166667 20.918021 31.515312 0.00000
00
```

Tabla de comparaciones con Tukey:

```
resultadosHSD <- data.frame(
  Comparacion = c(
    "Grayson's Pond vs Beaver Lake",
    "Grayson's Pond vs Angler's Cove",
    "Grayson's Pond vs Appletree Lake",
    "Grayson's Pond vs Rock River",
    "Beaver Lake vs Angler's Cove",
    "Beaver Lake vs Appletree Lake",
    "Beaver Lake vs Rock River",
    "Angler's Cove vs Appletree Lake",
```

```

      "Angler's Cove vs Rock River",
      "Appletree Lake vs Rock River"
    ),
    Diferencia = c(-8.15, -12.00, -9.017, -26.217, -3.85, -0.867, -18.067,
2.983, -14.217, -17.2),
    Significancia = c("Significativa", "Significativa", "Significativa", "S
ignificativa",
                      "No significativa", "No significativa", "Significativ
a",
                      "No significativa", "Significativa", "Significativa")
  )
kable(resultadosHSD, caption = "Resultados Tukey HSD de comparación entre
pares de medias",
      col.names = c("Pares de medias", "Diferencia (mg/ml)", "Resultado T
ukey HSD"))

```

Resultados Tukey HSD de comparación entre pares de medias

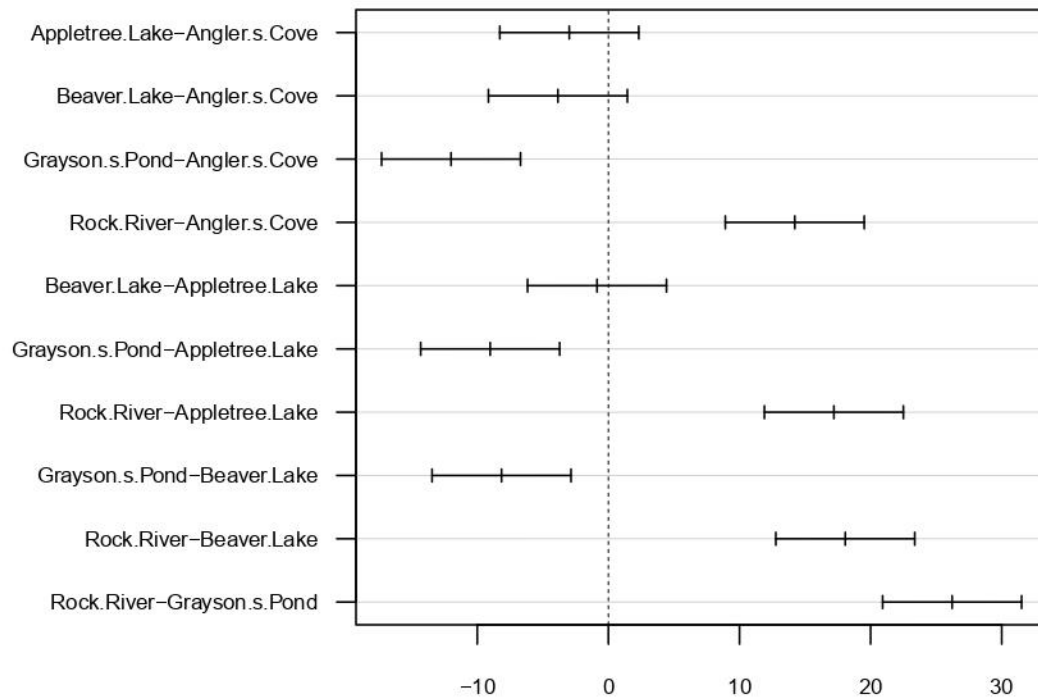
Pares de medias	Diferencia (mg/ml)	Resultado Tukey HSD
Grayson's Pond vs Beaver Lake	-8.150	Significativa
Grayson's Pond vs Angler's Cove	-12.000	Significativa
Grayson's Pond vs Appletree Lake	-9.017	Significativa
Grayson's Pond vs Rock River	-26.217	Significativa
Beaver Lake vs Angler's Cove	-3.850	No significativa
Beaver Lake vs Appletree Lake	-0.867	No significativa
Beaver Lake vs Rock River	-18.067	Significativa
Angler's Cove vs Appletree Lake	2.983	No significativa
Angler's Cove vs Rock River	-14.217	Significativa
Appletree Lake vs Rock River	-17.200	Significativa

```

plot(TukeyHSD(estroncio_long.aov), las = 1, cex.axis = 0.8)

```


95% family-wise confidence level



Differences in mean levels of estroncio_long\$Cuerpo_agua

Comparación LSD vs Tukey HSD

```
comparaciones <- c(
  "Grayson's Pond vs Beaver Lake",
  "Grayson's Pond vs Angler's Cove",
  "Grayson's Pond vs Appletree Lake",
  "Grayson's Pond vs Rock River",
  "Beaver Lake vs Angler's Cove",
  "Beaver Lake vs Appletree Lake",
  "Beaver Lake vs Rock River",
  "Angler's Cove vs Appletree Lake",
  "Angler's Cove vs Rock River",
  "Appletree Lake vs Rock River"
)

diferencias <- c(-8.15, -12.00, -9.017, -26.217, -3.85, -0.867, -18.067,
2.983, -14.217, -17.2)

resultado_LSD <- c("Significativa", "Significativa", "Significativa", "Si
nificativa",
                  "Significativa", "No significativa", "Significativa",
                  "No significativa", "Significativa", "Significativa")

resultado_Tukey <- c("Significativa", "Significativa", "Significativa", "
Significativa",
                   "No significativa", "No significativa", "Significati
va",
                   "No significativa", "Significativa", "Significativa"
)

resultadosComparados <- data.frame(
  `Pares de medias` = comparaciones,
  `Diferencia (mg/ml)` = diferencias,
  `Resultado LSD` = resultado_LSD,
  `Resultado Tukey HSD` = resultado_Tukey
)

kable(resultadosComparados,
      caption = "Comparación de resultados LSD vs Tukey HSD entre pares d
e medias",
      col.names = c("Pares de medias", "Diferencia (mg/ml)", "Resultado L
SD", "Resultado Tukey HSD"))
```

Comparación de resultados LSD vs Tukey HSD entre pares de medias

Pares de medias	Diferencia (mg/ml)	Resultado LSD	Resultado Tukey HSD
Grayson's Pond vs Beaver Lake	-8.150	Significativa	Significativa
Grayson's Pond vs Angler's Cove	-12.000	Significativa	Significativa
Grayson's Pond vs Appletree Lake	-9.017	Significativa	Significativa
Grayson's Pond vs Rock River	-26.217	Significativa	Significativa
Beaver Lake vs Angler's Cove	-3.850	Significativa	No significativa
Beaver Lake vs Appletree Lake	-0.867	No significativa	No significativa
Beaver Lake vs Rock River	-18.067	Significativa	Significativa
Angler's Cove vs Appletree Lake	2.983	No significativa	No significativa
Angler's Cove vs Rock River	-14.217	Significativa	Significativa
Appletree Lake vs Rock River	-17.200	Significativa	Significativa

Interpretación ambiental

¿Qué cuerpo de agua presenta las concentraciones más altas?

- Grayson's Pond y Beaver Lake.

¿Qué sitios no difieren entre sí?

- Según LSD: Beaver Lake vs Appletree Lake y Angler's Cove vs Appletree Lake.
- Según Tukey HSD: los mismos de LSD, más Beaver Lake vs Angler's Cove.

¿Qué implicaciones podrían tener estas diferencias en la calidad del agua?

- Las diferencias en las concentraciones de estroncio entre cuerpos de agua indican una variabilidad en la calidad. Los sitios con valores más altos podrían estar expuestos a contaminación local o condiciones específicas, lo que representa un riesgo potencial para la fauna y flora local. Asimismo, estas diferencias pueden tener implicaciones en el uso humano del agua (consumo, riego o recreación), por lo que los sitios con mayores concentraciones requieren mayor monitoreo y gestión ambiental.
-