

# Ejemplo Ilustrativo de medición de variables químicas en suelos

En un estudio realizado por Franco y otros en septiembre de 2021 a 35 fincas de la Sierra ecuatoriana, parroquia Tupigachi, Pichincha, a 3000 msnm bajo clima Mesotérmico Semi-húmedo. Se analizó el impacto del uso pecuario intensivo entre 2008-2017 sobre algunas propiedades químicas del suelo, repitiendo los análisis de suelos en ambos años bajo procedimientos standard de laboratorio este diagnóstico sirvió para formular recomendaciones que contribuyan al mejoramiento el sistema agrícola de las fincas. La matriz de datos tiene la estructura siguiente:

```
data1<- data.frame(read_excel("Datos dos Poblaciones.xlsx",
                             sheet = "Biofran"))
```

```
## New names:
## * `` -> `...1`
```

```
str(data1)
```

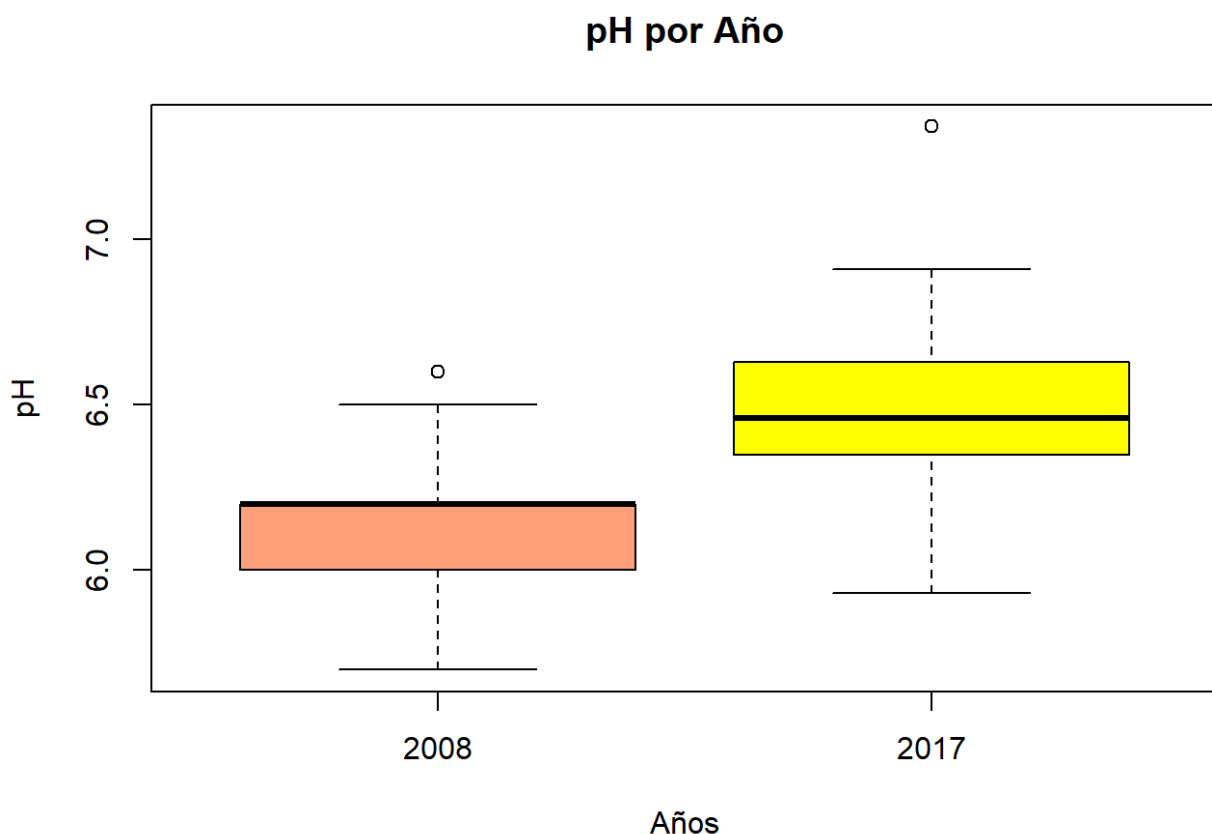
```
## 'data.frame': 70 obs. of 24 variables:
## $ ...1 : chr "SA17-161" "SA17-121" "SA17-120" "SA17-196" ...
## $ PH : num 6.3 6.5 6.2 5.7 5.8 6 6 6.3 6.2 6.1 ...
## $ C.E : num 0.12 0.23 0.19 0.17 0.23 0.11 0.15 0.19 0.14 0.2 ...
## $ M.O.. : num 1.42 2.79 1.68 1.84 2.58 2.62 2.15 2.89 2.24 2.21 ...
## $ N : num 12.59 29.99 3.26 21.83 25.33 ...
## $ P : num 3 13.2 5.9 15 20.6 1.9 1.2 8 13 10.7 ...
## $ K : num 0.18 0.58 0.53 0.38 0.28 0.31 0.19 0.44 0.15 0.24 ...
## $ Ca : num 8.09 10.07 9.8 6.9 6.96 ...
## $ Mg : num 2.88 3.18 3.24 1.72 1.83 2.24 2.72 3.01 2.69 2.28 ...
## $ Na : num 0.07 0.08 0.04 0.03 0.06 0.04 0.05 0.07 0.07 0.06 ...
## $ CICE : num 11.22 13.91 13.61 9.03 9.13 ...
## $ Cu : num 8.5 6.8 9.4 6.3 8.4 8.4 8.4 8.3 7.7 6 ...
## $ Fe : num 207 217 98.8 357 641 484 429 340 512 537 ...
## $ Mn : num 3.2 34 4.5 4.6 4 8.3 6.1 5 9 4.7 ...
## $ Zn : num 4.5 7.7 7.3 5.6 5.7 3.4 3.4 5.2 3.9 3.2 ...
## $ B : num 0.26 0.2 0.38 0.26 0.18 0.43 0.15 0.25 0.25 0.31 ...
## $ S : num 0.957 1.386 1.683 1.188 1.452 ...
## $ Fe.Mn : num 64.69 6.38 21.96 77.61 160.25 ...
## $ Ca.Mg : num 2.81 3.17 3.02 4.01 3.8 ...
## $ Mg.K : num 16 5.48 6.11 4.53 6.54 ...
## $ Ca.Mg.K: num 60.9 22.8 24.6 22.7 31.4 ...
## $ ARENA : num 60 70 58 60 78 76 56 58 68 58 ...
## $ LIMO : num 30 24 30 32 20 26 30 32 24 36 ...
## $ ARCILLA: num 10 6 12 8 2 2 14 10 8 6 ...
```

se trata de 70 mediciones realizadas a las 35 fincas en dos oportunidades diferentes (año 2008 y año 2017), que conforman la `(data.frame)`, esta contiene 23 variables numéricas. Entre las variables de estudio se encuentra el `(pH)`; una característica química muy importante para medir el impacto del cambio pecuario en el suelo estudiado. Presentaremos unos elementos básicos para el contraste de media de los datos y una interpretación “introdutoria” de los resultados.

# Representación del \(\text{pH}\).

El primer paso es crear un vector que hemos denominado "año" para destacar las diferencias del cambio pecuario; posteriormente, como análisis exploratorio hemos creado un boxplot que nos permite visualizar las diferencias de las dos ditribuciones.

```
año<-factor(c(rep("2008",35),(rep("2017",35))))
pH<-data1$PH
#EXPLORACIÓN
boxplot(pH ~ año,
        main="pH por Año",xlab="Años",
        ylab="pH",col=c("lightSalmon","yellow"))
```



```
summary(pH[1:35])
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	5.700	6.000	6.200	6.123	6.200	6.600

```
summary(pH[35:70])
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	5.930	6.345	6.460	6.492	6.625	7.340

El gráfico muestra que aparentemente hay diferencias entre los valores de las soluciones de suelos en sus distribuciones de \(\text{pH}\) en los años estudiados, al parecer en el año 2017 hubo un cambio en el promedio de \(\text{pH}\) en la solución del suelo. El año 2008 muestra una asimetría negativa, lo que nos

indica una marcada tendencia de valores centrales bajos de pH que oscilan entre 6 a 6.2. Mientras que, en el año 2017 el  $\text{pH}$  aumenta a valores medios centrales que van entre 6.35 y 6.63.

## Contraste de normalidad

En esta parte se contrasta la hipótesis nula de normalidad de los datos en contraposición de que los datos no son normales:

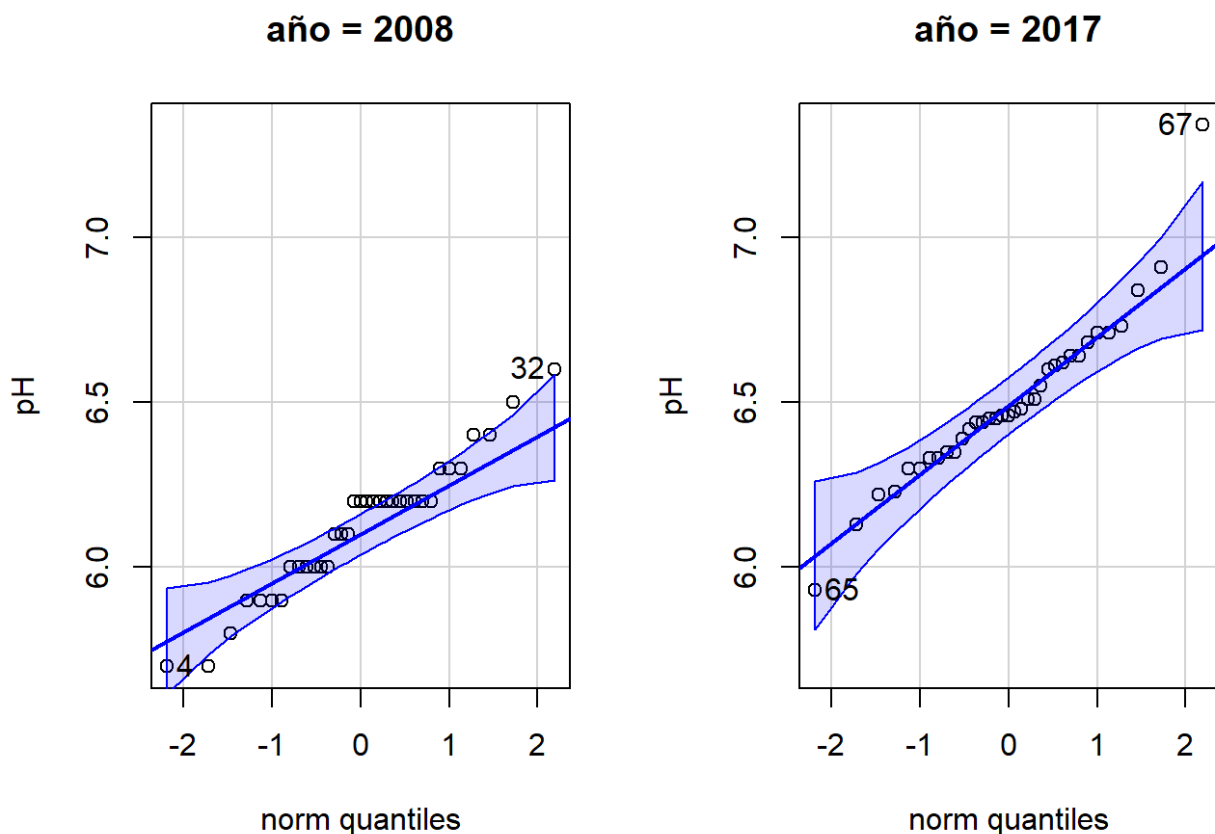
$H_0$ : los datos se distribuyen como una normal vs

$H_1$ : los datos no se distribuyen como una normal

## Exploración de la normalidad

Los gráficos qqPlot muestran que “aparentemente” los datos se distribuyen como una normal; sin embargo, hay datos que se escapan de los intervalos de confianza, por ello se procede a realizar un contraste confirmatorio con el test de Shapiro-Wilk.

```
#NORMALIDAD
library(car)
qqPlot(pH ~ año,)
```



Los valores  $p$  para los datos son mayores de  $\alpha=0.05$ . Por lo tanto, no hay evidencia estadística para rechazar hipótesis nula  $H_0$ , luego podemos asegurar que tanto en el año 2008 como en el año 2017 el  $\text{pH}$  en suelo presenta distribuciones normales.

```
shapiro.test(pH[1:35])
```

```
##
##  Shapiro-Wilk normality test
##
## data:  pH[1:35]
## W = 0.95825, p-value = 0.2025
```

```
shapiro.test(pH[36:70])
```

```
##
##  Shapiro-Wilk normality test
##
## data:  pH[36:70]
## W = 0.93935, p-value = 0.05345
```

Por otro lado, para poder realizar un contraste de medias adecuado debemos determinar si las poblaciones de estudio tienen o no varianza iguales; es decir, debemos realizar el siguiente contraste:

$H_0: \sigma_1^2 = \sigma_2^2$  vs

$H_0: \sigma_1^2 \neq \sigma_2^2$  la instrucción de R es como sigue:

```
#Contraste para dos varianzas
var.test(data1$PH~año)
```

```
##
##  F test to compare two variances
##
## data:  data1$PH by año
## F = 0.69572, num df = 34, denom df = 34, p-value = 0.295
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.3511732 1.3782963
## sample estimates:
## ratio of variances
##           0.695716
```

El valor  $p$  (0.295) del contraste de homogeneidad, nos indica que no hay evidencia estadística para rechazar la hipótesis de que las varianzas de son iguales.

## Contraste de medias para el pH

Una vez contrastadas las hipótesis de normalidad de las poblaciones y la homogeneidad de la varianza pasamos a realizar el contraste de medias del pH. Cabe preguntarse si las muestras son independientes o relacionadas. Por un lado, podría pensarse que las muestras son independiente ya que han pasado 9 años entre las mediciones, otra postura es que los datos provienen del mismo suelo, y por lo tanto, las muestras están correlacionadas (pareadas). Nosotros haremos los dos contrastes, uno para muestras independientes, y otra para muestras dependientes, queda como investigación para los estudiantes sustentar cuál de las dos pruebas es la más adecuada para nuestros datos.

Hipótesis a contrastar:

$H_0: \mu_1 = \mu_2$  vs

$H_0: \mu_1 \neq \mu_2$

## Contraste al tratar las muestras como independientes

```
tt<-t.test(pH~año,alternative = "two.sided",
           var.equal = T);tt
```

```
##
## Two Sample t-test
##
## data: pH by año
## t = -6.9561, df = 68, p-value = 1.702e-09
## alternative hypothesis: true difference in means between group 2008 and group 2017 i
s not equal to 0
## 95 percent confidence interval:
## -0.4864355 -0.2695645
## sample estimates:
## mean in group 2008 mean in group 2017
##          6.122857          6.500857
```

## Contraste al tratar las muestras como dependientes

```
tt<-t.test(pH~año,alternative = "two.sided",
           paired = TRUE,
           var.equal = T);tt
```

```
##
## Paired t-test
##
## data: pH by año
## t = -7.8498, df = 34, p-value = 3.871e-09
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.4758603 -0.2801397
## sample estimates:
## mean of the differences
##          -0.378
```

En ambos contrastes el valor  $\backslash(p\backslash)$  es más pequeño que el nivel de significación  $\backslash(\alpha= 0,05\backslash)$ , por lo tanto, hay evidencia para rechazar la hipótesis nula de que la media de los  $\backslash(pH\backslash)$  en las soluciones de suelo en los años 2008 y 2017 son iguales.

## Elementos referenciales para interpretar los resultados del Ejemplo de estudio.

Dado que el  $\backslash(pH\backslash)$  es una propiedad química que mide el grado de acidez o alcalinidad de las soluciones acuosas, N.W. Osorio (2012) asegura: "... la solución del suelo usualmente fluctúa entre 4.0 a 8.0. En general, se considera que los suelos con  $\backslash(pH<7\backslash)$  son ácidos y si el  $\backslash(pH> 7\backslash)$  son alcalinos" p.2. También añade que la acidez del suelo es una condición muy común en los suelos de las regiones húmedas tropicales; mientras que, la condición alcalina predomina en suelos de regiones secas-

tropicales. Esta característica del suelo es sumamente importante puesto que es un buen indicador de la disponibilidad de nutrientes. A continuación se suministra una tabla que nos permitirá la interpretación de nuestros datos.

Tabla 1: Interpretación del pH del suelo (agua, 1:1, V:V). Fuente: ICA en N.W. Osorio (2012).

pH	Categoría	Interpretación
< 5.0	Extremadamente ácido	Severa toxicidad por Al y quizá por Mn; Alta probabilidad de deficiencia de P, S, Mo y bases intercambiables; se esperan altos niveles de algunos micronutrientes. Muchos cultivos requieren encalamiento
5.0-5.5	Fuertemente ácido	Toxicidad moderada por Al y Mn; deficiencia de P, S, Mo y bases; altos niveles de algunos micronutrientes. Muchos cultivos requieren encalamiento.
5.5-6.0	Moderadamente ácido	No se espera la toxicidad por Al; mayor disponibilidad de P, S, Mo y bases. Algunos cultivos susceptibles a la acidez del suelo requieren encalamiento.
6.0-6.5	Ligeramente ácido	Adecuada condición para la disponibilidad de nutrientes para las plantas.
6.5-7.3	Neutro	Altos niveles de Ca, Mg. Algunos cultivos pueden mostrar deficiencias de micronutrientes. La disponibilidad de P puede ser baja.
7.4-8.0	Alcalino	Baja disponibilidad de P y micronutrientes. Altos niveles de Ca, Mg. El Na puede ser un problema.
> 8.0	Muy alcalino	Severas limitaciones en la disponibilidad de algunos nutrientes. El nivel de Na puede ser tóxico

Por medio de la tabla podemos concluir que hubo un cambio significativo de un suelo que en año 2008 se consideraba moderadamente ácido a ligeramente ácido, a un suelo ligeramente ácido a neutro, y por lo tanto, con este cambio el suelo tiende a tener una adecuada condición para la disponibilidad de nutrientes de las plantas en la mayoría de las fincas, aunque algunas pueden presentar altos niveles de  $\text{Ca}$  y  $\text{Mg}$ ; así como también, deficiencias de micronutrientes y baja disponibilidad de fósforo.