Universidad Autónoma de Nuevo León

Facultad de Ciencias Forestales

Análisis Estadístico

Docente: Dr. Marco A. González Tagle

Tarea 5: Correlación

Jonathan de Jesus Pere de la Rosa

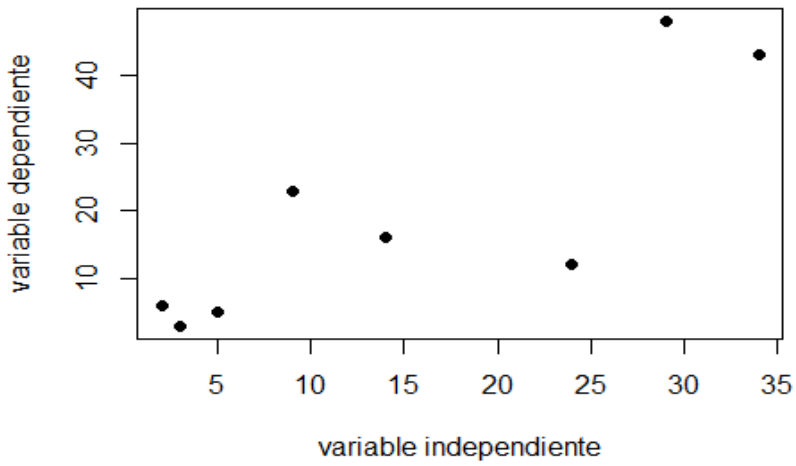
Primer Semestre

Tarea_5_JonathandeJesusPerezdelaRosa.R

k_ren

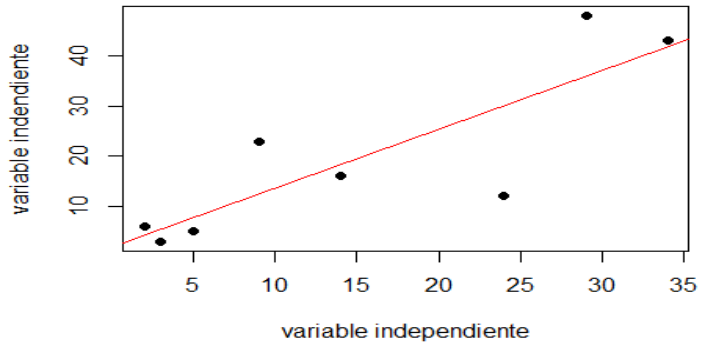
2023-10-05

```
#JONATHAN DE JESUS PEREZ DE LA ROSA
# 06/10/2022
# Tarea 1
# Ejercicio 1 ------------------------------------------------------------
Speed <- c(2,3,5,9,14,24,29,34)
Abundance <- c(6,3,5,23,16,12,48,43)
efimeras <- matrix(0,2,8)
efimeras[1,] <- Speed
efimeras[2,] <- Abundance

plot(Speed,Abundance,
     pch=19,
       xlab = "variable independiente",
       ylab = "variable dependiente")
```




```
##¿Es estadisticamente significativa la correlacion?
###Respuesta: si es significativo la correlacion
cor.ef <- cor.test(efimeras[1,],efimeras[2,])
cor.ef

##
## Pearson's product-moment correlation
##
## data:  efimeras[1, ] and efimeras[2, ]
## t = 3.8568, df = 6, p-value = 0.008393
## alternative hypothesis: true correlation is not equal to 0
```

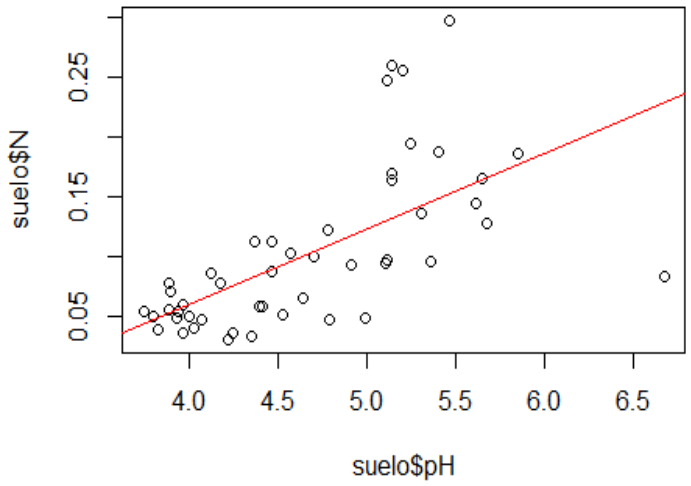
```
## 95 percent confidence interval:
##  0.3442317 0.9711386
## sample estimates:
##       cor
## 0.8441408

##
## Pearson's product-moment correlation
##
## data: efimeras[1, ] and efimeras[2, ]
## t = 3.8568, df = 6, p-value = 0.008393
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.3442317 0.9711386
  ## sample estimates:
## cor
## 0.8441408
# funcion de lm lineal modem
ef.lm <- lm(efimeras[2,]~efimeras[1,])
ef.lm

##
## Call:
## lm(formula = efimeras[2, ] ~ efimeras[1, ])
##
## Coefficients:
##   (Intercept)  efimeras[1, ]
##         1.867          1.176

##
## Call:
## lm(formula = efimeras[2, ] ~ efimeras[1, ])
##
## Coefficients:
## (Intercept) efimeras[1, ]
## 1.867 1.176
#grafica de dispercion
plot(Speed,Abundance,pch=19,
     xlab = "variable independiente",
     ylab = "variable indendiente")
abline (ef.lm,col="red")
```



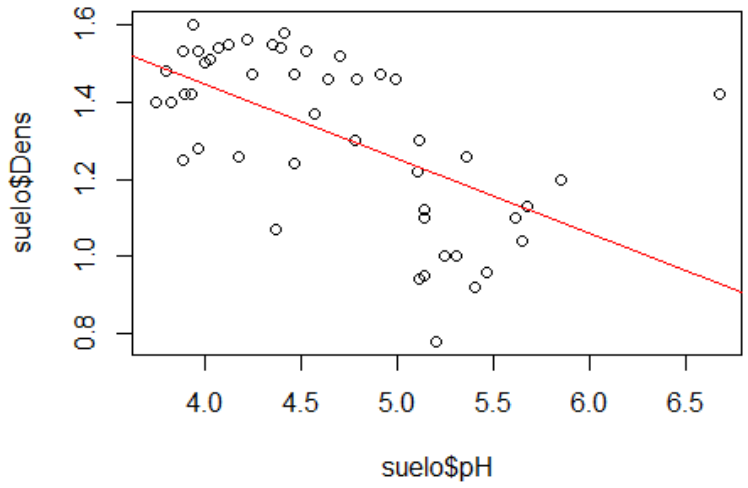

```
## Examinar la relacion que existe entre dos muestras mediante una
correlacion.
### Respuesta: es una relacion lineal.

## Explore los datos graficamente y explique.
### Respuesta: es una relacion lineal, con una correlacion positiva

##Establezca la Hipotesis nula y la Hipotesis alternativa,
### Respuesta:
### hipotesis alternativa:Existe una correlación positiva entre la
velocidad de los arroyos y la abundancia de efimeras (Ecdyonurus dispar).
### hipotesis nula:No existe una correlación entre la velocidad del

## Aplique la prueba correspondiente.
###Prueba de zafiro
shapiro.test(efimeras[1,]) #Los datos son normales.

##
##  Shapiro-Wilk normality test
##
## data:  efimeras[1, ]
## W = 0.89444, p-value = 0.2572

##
## Shapiro-Wilk normality test
##
## data: efimeras[1, ]
## W = 0.89444, p-value = 0.2572
###Coeficientes de correlacion (r).
cor.ef <- cor.test(efimeras[1,],efimeras[2,])
cor.ef

##
##  Pearson's product-moment correlation
##
## data:  efimeras[1, ] and efimeras[2, ]
## t = 3.8568, df = 6, p-value = 0.008393
```

```
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.3442317 0.9711386
## sample estimates:
##       cor
## 0.8441408

##
## Pearson's product-moment correlation
##
## data: efimeras[1, ] and efimeras[2, ]
## t = 3.8568, df = 6, p-value = 0.008393
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.3442317 0.9711386
## sample estimates:
## cor
## 0.8441408
summary(ef.lm)

##
## Call:
## lm(formula = efimeras[2, ] ~ efimeras[1, ])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -18.080  -2.481  -0.580   3.975  12.042
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     1.8667     5.7912   0.322  0.75813
## efimeras[1, ]   1.1756     0.3048   3.857  0.00839 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.05 on 6 degrees of freedom
## Multiple R-squared:  0.7126, Adjusted R-squared:  0.6647
## F-statistic: 14.87 on 1 and 6 DF,  p-value: 0.008393

##
## Call:
## lm(formula = efimeras[2, ] ~ efimeras[1, ])
##
## Residuals:
## Min 1Q Median 3Q Max
## -18.080 -2.481 -0.580 3.975 12.042
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.8667 5.7912 0.322 0.75813
## efimeras[1, ] 1.1756 0.3048 3.857 0.00839 **
```

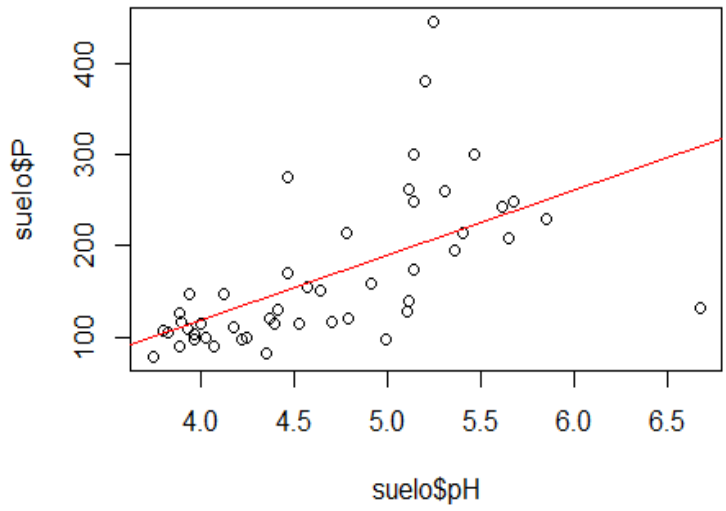
```
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.05 on 6 degrees of freedom
## Multiple R-squared: 0.7126, Adjusted R-squared: 0.6647
## F-statistic: 14.87 on 1 and 6 DF, p-value: 0.008393
##Reporte los datos (indicar valor de r, grados de libertad y
probabilidad, asi como significancia
##de la correlacion).
### Respuesta: r=0.8441408 ; df=6; p-value=0.008393.
### Hipotesis aceptada es la Hipotesis alterna=Existe una correlación
positiva entre la velocidad de los arroyos y la abundancia de efimeras.


# Ejercicio 2 ------------------------------------------------------------
--
#Conjunto de datos: Composiciones del suelo, caracteristicas fisicas y
quimicas
suelo <- read.csv("E:\\suelo.csv", header = T)
N <- cor.test(suelo$pH, suelo$N)
plot(suelo$pH, suelo$N)
abline(lm(suelo$N ~ suelo$pH),
       col = "red")
```



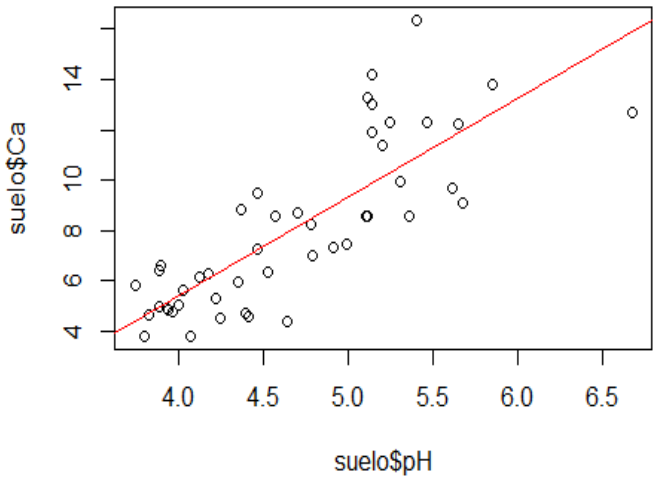

```
De <- cor.test(suelo$pH,suelo$Dens)
plot(suelo$pH,suelo$Dens)
abline(lm(suelo$Dens ~ suelo$pH),
       col="red")
```



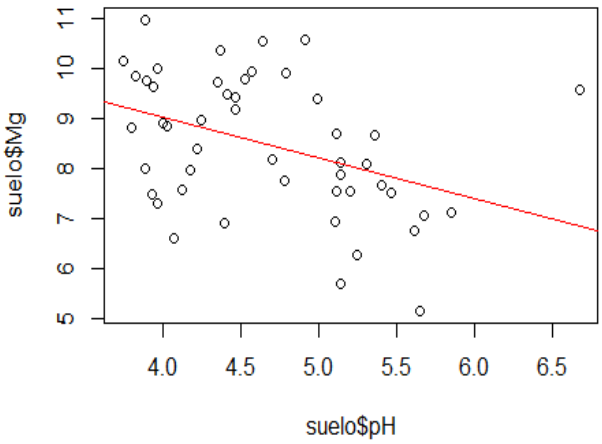

```r
P <- cor.test(suelo$pH,suelo$P)
plot(suelo$pH,suelo$P)
abline(lm(suelo$P ~ suelo$pH),
       col="red")
```



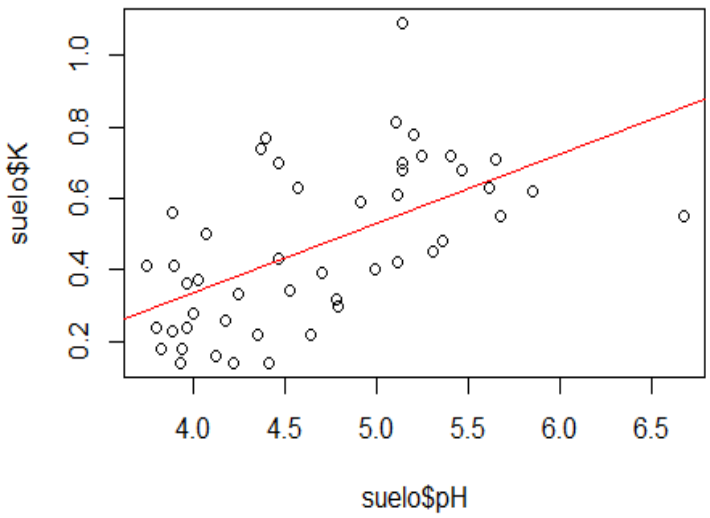

```r
Ca <- cor.test(suelo$pH,suelo$Ca)
plot(suelo$pH,suelo$Ca)
abline(lm(suelo$Ca ~ suelo$pH),
       col="red")
```



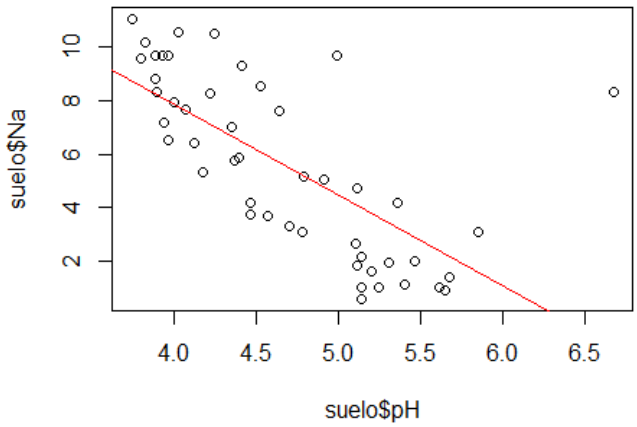

```r
Mg <- cor.test(suelo$pH,suelo$Mg)
plot(suelo$pH,suelo$Mg)
abline(lm(suelo$Mg ~ suelo$pH),
       col="red")
```



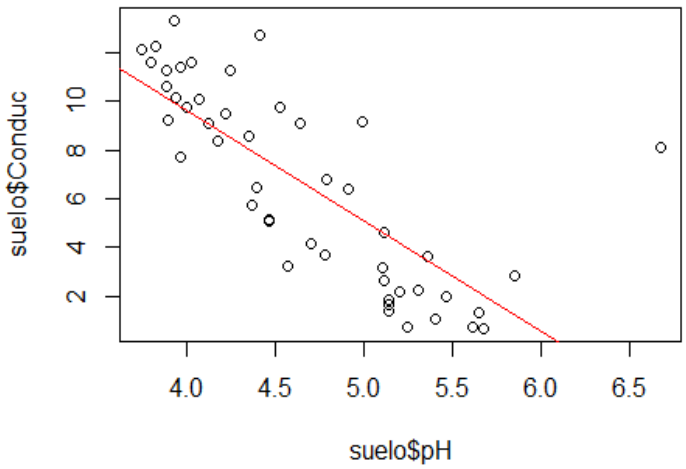

```r
K <- cor.test(suelo$pH,suelo$K)
plot(suelo$pH,suelo$K)
abline(lm(suelo$K ~ suelo$pH),
       col="red")
```




```
Na <- cor.test(suelo$pH,suelo$Na)
plot(suelo$pH,suelo$Na)
abline(lm(suelo$Na ~ suelo$pH),
       col="red")
```




```
Con <- cor.test(suelo$pH,suelo$Conduc)
plot(suelo$pH,suelo$Conduc)
abline(lm(suelo$Conduc ~ suelo$pH),
       col="red")
```




```
r <- c(N$estimate,De$estimate,P$estimate,Ca$estimate,
       Mg$estimate,K$estimate,Na$estimate,Con$estimate)
v_p <- c(N$p.value,De$p.value,P$p.value,Ca$p.value,
         Mg$p.value,K$p.value,Na$p.value,Con$p.value)
conjuntos <- c("pH - N","pH - Dens","pH - P","pH - Ca","pH -Mg","pH -
K","pH -Na","pH - Conductividad")

##Cuadro de datos con los estadísticos de interés.
M <- matrix(0,8,3)
M[,1] <- conjuntos
M[,2] <- r
M[,3] <- v_p
colnames(M) <- c("Conjunto", "r","valor de P")
M

## Conjunto r valor de P
## [1,] "pH - N" "0.636654010590012" "1.14882323161906e-06"
## [2,] "pH - Dens" "-0.589026444818348" "1.0620204782642e-05"
## [3,] "pH - P" "0.591030281447051" "9.74033722057688e-06"
## [4,] "pH - Ca" "0.808629294673671" "3.61374548769419e-12"
## [5,] "pH -Mg" "-0.395782110416998" "0.00536142977803124"
## [6,] "pH -K" "0.57957268749503" "1.58489239417252e-05"
## [7,] "pH -Na" "-0.693261417881367" "4.72420741991162e-08"
## [8,] "pH - Conductividad" "-0.764810364711197" "2.48377504440975e-10"
```