

## Clase\_S14\_D1.R

isa\_r

2022-05-20

```
# AMANDA
# 04/05/2022 SEMANA 14
# CORRELACIÓN: EJERCICIO DE REGRESIÓN

#  $Y' = \alpha + \beta(x)$  "línea de la regresión lineal"
# REGRESIÓN: valores mínimos y máximos para conocer la media( $Y'$ )
# sumatoria:  $y' - y_{obs} = 0$  residuales
# varianza

semilla <- read.csv("BaseDeDatos_estadistica.csv", header = T)

# PLOT DE DIÁMETRO vs PESO (gr)
plot(semilla$Diametro_mm, semilla$Peso_gr,
     pch=20,
     col= "lightblue",
     xlab = "Diámetro de semilla (mm)",
     ylab = "Peso de semilla (gr)")

# REGRESIÓN <- lm es la función que ayuda a determinar  $\alpha$  y  $\beta$  de la
# regresión

# lm(variables dependiente, variable independientes)
sem.lm <- lm(semilla$Peso_gr ~ semilla$Diametro_mm)
#  $\alpha$  (intercept)=-4.93  $X=0$ 
#  $\beta$ = 0.5318

summary(sem.lm)

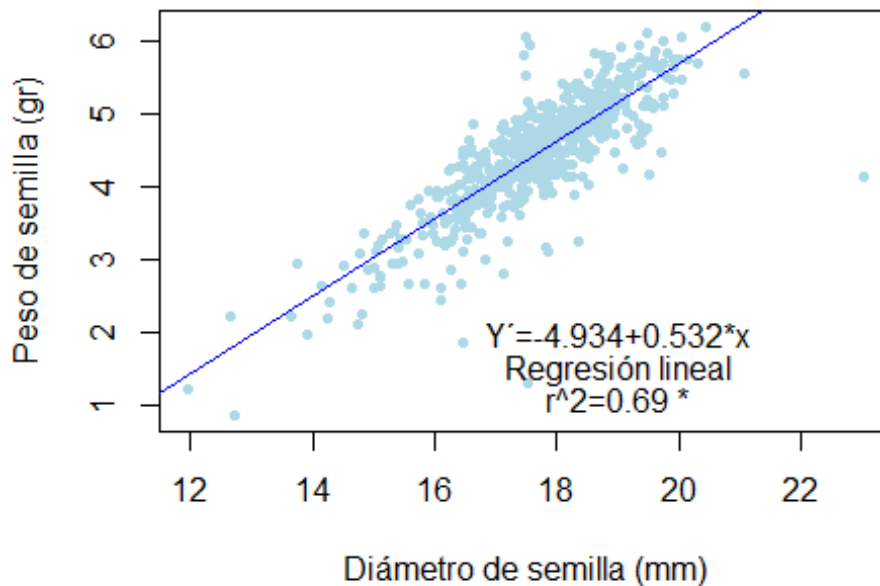
##
## Call:
## lm(formula = semilla$Peso_gr ~ semilla$Diametro_mm)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.15249 -0.19803  0.02135  0.25257  1.71292
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -4.93370    0.25602  -19.27  <2e-16 ***
## semilla$Diametro_mm  0.53178    0.01446   36.77  <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4372 on 597 degrees of freedom
## Multiple R-squared:  0.6936, Adjusted R-squared:  0.6931
## F-statistic: 1352 on 1 and 597 DF,  p-value: < 2.2e-16

# *= SINIFICATIVO

semilla$yprima <- round(-4.93370+0.53178*semilla$Diametro_mm, 2)

# AGREGAR LA LÍNEA DE TENDENCIA CENTRAL USANDO abline
plot(semilla$Diametro_mm, semilla$Peso_gr,
     pch=20,
     col= "lightblue",
     xlab = "Diámetro de semilla (mm)",
     ylab = "Peso de semilla (gr)")
abline(sem.lm,
      col= "blue")
text(19, 2, "Y'=-4.934+0.532*x")
text(19, 1.5, "Regresión lineal")
text(19, 1.1, "r^2=0.69 *")
```



```
semilla$Ajustados <- round(sem.lm$fitted.values, 2)
sem.lm$coefficients

##          (Intercept) semilla$Diametro_mm
##          -4.9337031          0.5317771
```

```
sum(sem.lm$residuals)

## [1] 6.508682e-15

# ENCONTRAR LOS SIGUIENTES VALORES APLICANDO LA FÓRMULA DE REGRESIÓN (18,
12.5, 14, 16, 15)
valores <- c(12.5, 14, 15, 16, 18)
-4.93370+0.53178*valores

## [1] 1.71355 2.51122 3.04300 3.57478 4.63834
```