Estadística (Q) - Clase 22
Mínimos cuadrados pesados

---

1. Varias inmobiliarias han recolectado datos viendo el valor real (Y en miles de pesos) de las ventas de departamentos que han realizado en el último año en el barrio de Palermo y los metros cuadrados de los departamentos (X). Desean plantear un modelo lineal que relacione dichos valores, con el objeto de estimar mediante mínimos cuadrados la media del valor del metro cuadrado en dicho barrio. A continuación un gráfico

   Si planteamos el modelo

   $Y_i = \alpha + \beta X_i + \varepsilon_i,$

   a) ¿Hay alguno de los supuestos del modelo lineal del que se pueda sospechar con el gráfico de XY observado?

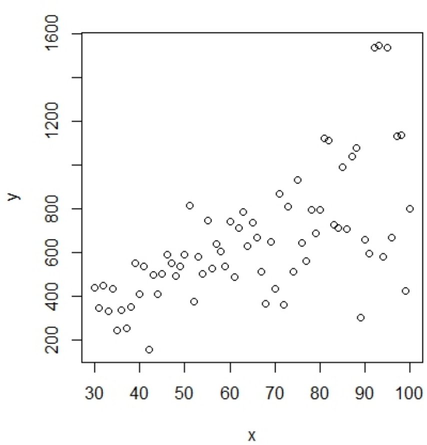
   b) Se sospecha que $s.d.(\varepsilon_i)$ depende linealmente de X_i. ¿Cómo se podría intentar estimar dicha dependencia?

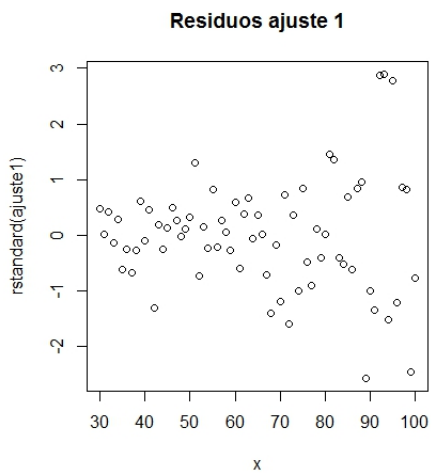
   c) Supongamos que se sabe que $S.D.(\varepsilon_i) = 3X_i$. Transforme los datos para armar un modelo lineal que cumpla las hipótesis y estime los parámetros del nuevo modelo y úselos para construir la recta de mínimos cuadrados del viejo modelo. ¿Cuál es finalmente la recta de mínimos cuadrados que elegiría para estimar?

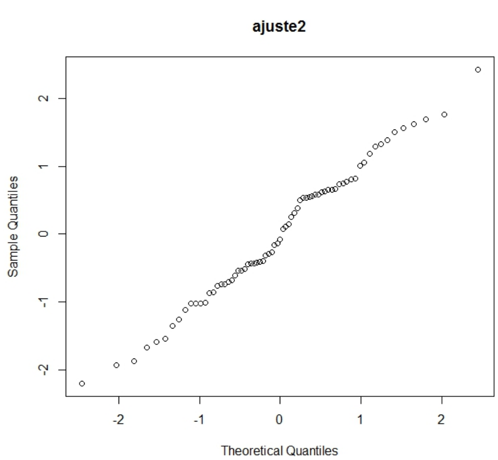
```
w<-1/(3*x)^2  #es el vector de pesos
ajuste1<-lm(y~x)
ajuste2<-lm(y~x,weights=w)
ajuste3<-lm(y~x-1,weights=w)

############################################
> shapiro.test(rstandard(ajuste1))

Shapiro-Wilk normality test

data:  rstandard(ajuste1)
W = 0.985, p-value = 0.5612
############################################
> shapiro.test(rstandard(ajuste2))

Shapiro-Wilk normality test

data:  rstandard(ajuste2)
W = 0.9865, p-value = 0.6477

############################################
> shapiro.test(rstandard(ajuste3))

Shapiro-Wilk normality test
```

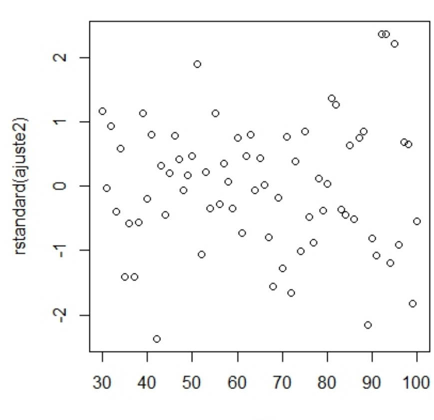
```
data:  rstandard(ajuste3)
W = 0.9867, p-value = 0.6592

###########################################
> summary(ajuste1)

Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max 
-670.30 -148.85   14.46  163.17  856.47 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  -25.977    111.369  -0.233    0.816    
x             10.999      1.634   6.731 4.07e-09 ***
---
Signif. codes:  0 ?***? 0.001 ?**? 0.01 ?*? 0.05 ?.? 0.1 ? ? 1 

Residual standard error: 282.2 on 69 degrees of freedom
Multiple R-squared:  0.3964,Adjusted R-squared:  0.3876 
F-statistic: 45.31 on 1 and 69 DF,  p-value: 4.074e-09 

##########################################
> summary(ajuste2)

Call:
lm(formula = y ~ x, weights = w)

Weighted Residuals:
    Min      1Q  Median      3Q     Max 
-2.9199 -0.9556 -0.1035  0.8760  3.1923 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   24.545     76.010   0.323    0.748    
x             10.182      1.396   7.294 3.88e-10 ***
---
Signif. codes:  0 ?***? 0.001 ?**? 0.01 ?*? 0.05 ?.? 0.1 ? ? 1 

Residual standard error: 1.337 on 69 degrees of freedom
Multiple R-squared:  0.4354,Adjusted R-squared:  0.4272 
F-statistic: 53.21 on 1 and 69 DF,  p-value: 3.882e-10 

#########################################
> summary(ajuste3)
```

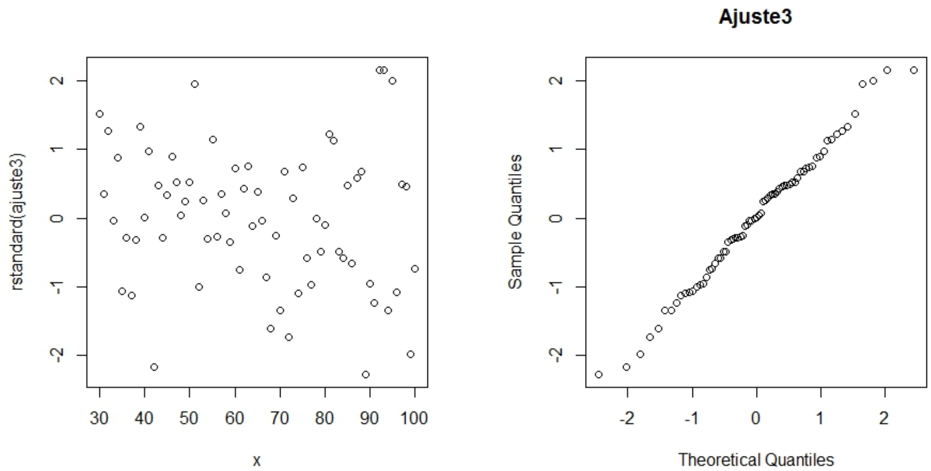
```
Call:
lm(formula = y ~ x - 1, weights = w)

Weighted Residuals:
     Min       1Q   Median       3Q      Max
-2.93132 -0.97881  0.01097  0.87134  3.14000

Coefficients:
  Estimate Std. Error t value Pr(>|t|)
x  10.6060     0.4729   22.43   <2e-16 ***
---
Signif. codes:  0 ?***? 0.001 ?**? 0.01 ?*? 0.05 ?.? 0.1 ? ? 1

Residual standard error: 1.328 on 70 degrees of freedom
Multiple R-squared:  0.8778,Adjusted R-squared:  0.8761
F-statistic: 502.9 on 1 and 70 DF,  p-value: < 2.2e-16
```










Ajuste3
rstandard(ajuste3)
x
Sample Quantiles
Theoretical Quantiles