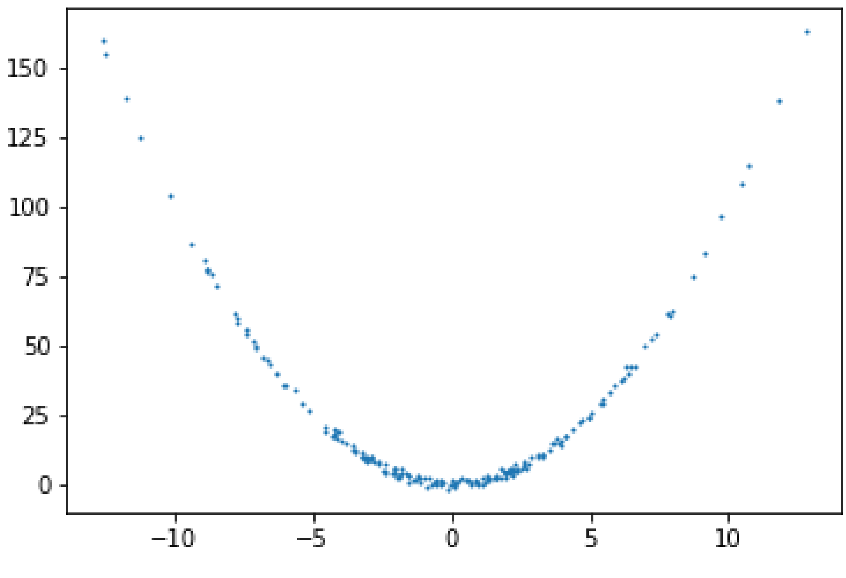
Linealidad en el modelo de regresión lineal múltiple

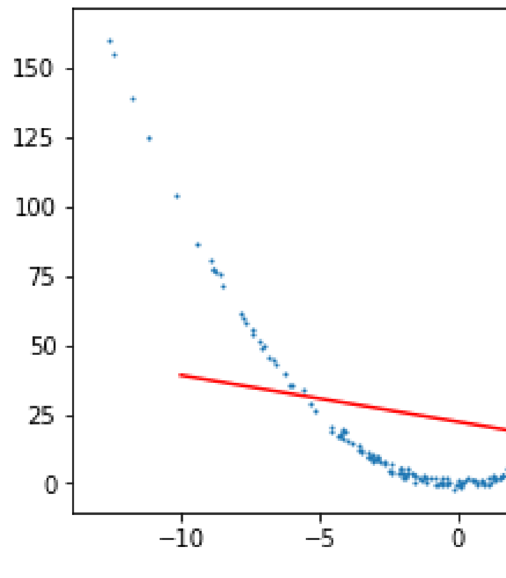
Econometría 2021-2022 Grado Ingenierías & ADE

Linealidad del Modelo

$$Y = X\beta + u$$

Los modelos de regresión lineal necesitan que la relación entre la variables dependientes e independientes sea lineal.

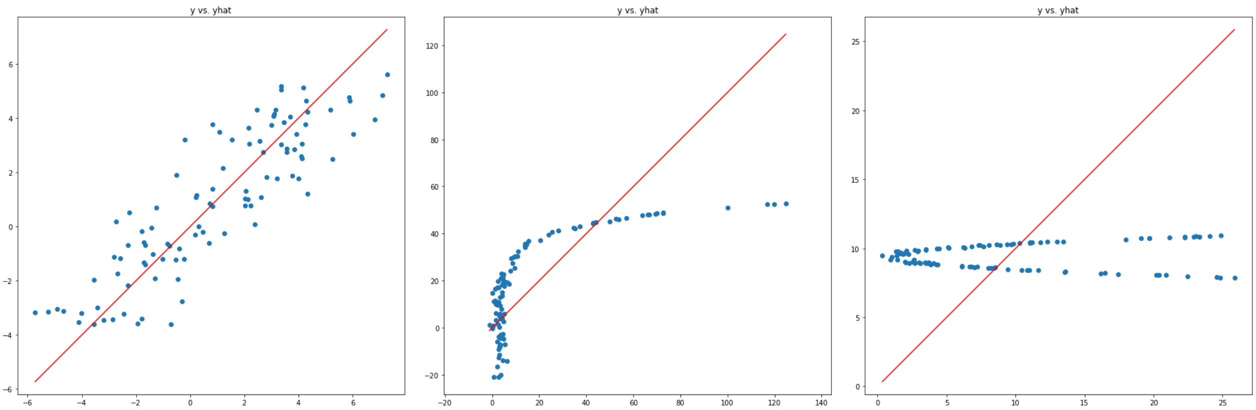




Linealidad del Modelo

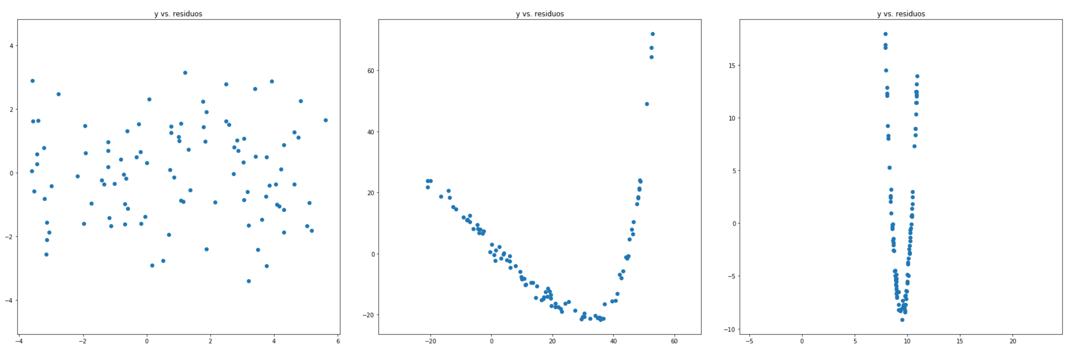
¿Cómo chequear la linealidad? Con nubes de puntos:

- Valores observados vs. Valores Predichos. Los puntos deberían estar simétricamente distribuidos alrededor de la diagonal.



Linealidad del Modelo

- Residuos vs. Valores predichos. Los puntos deberían estar situados alrededor de una línea horizontal con una varianza casi constante.



Corrección

- Si la variable dependiente es estrictamente positiva y el gráfico Residuos v.s Predicha indica que el tamaño de los errores es proporcional al tamaño de las predicciones $\Rightarrow$ transformar variables a su logaritmo.
- El modelo es claramente no lineal...

Modelos No Lineales

$$y_i = \beta_0 + \beta_1 \frac{1}{1 - e^{x_i}} + u_t$$

$$\sqrt{y_i} = \beta_0 + \beta_1 e^{x_i} + \beta_2 \sqrt[3]{\sin(z_i)} + u_t$$

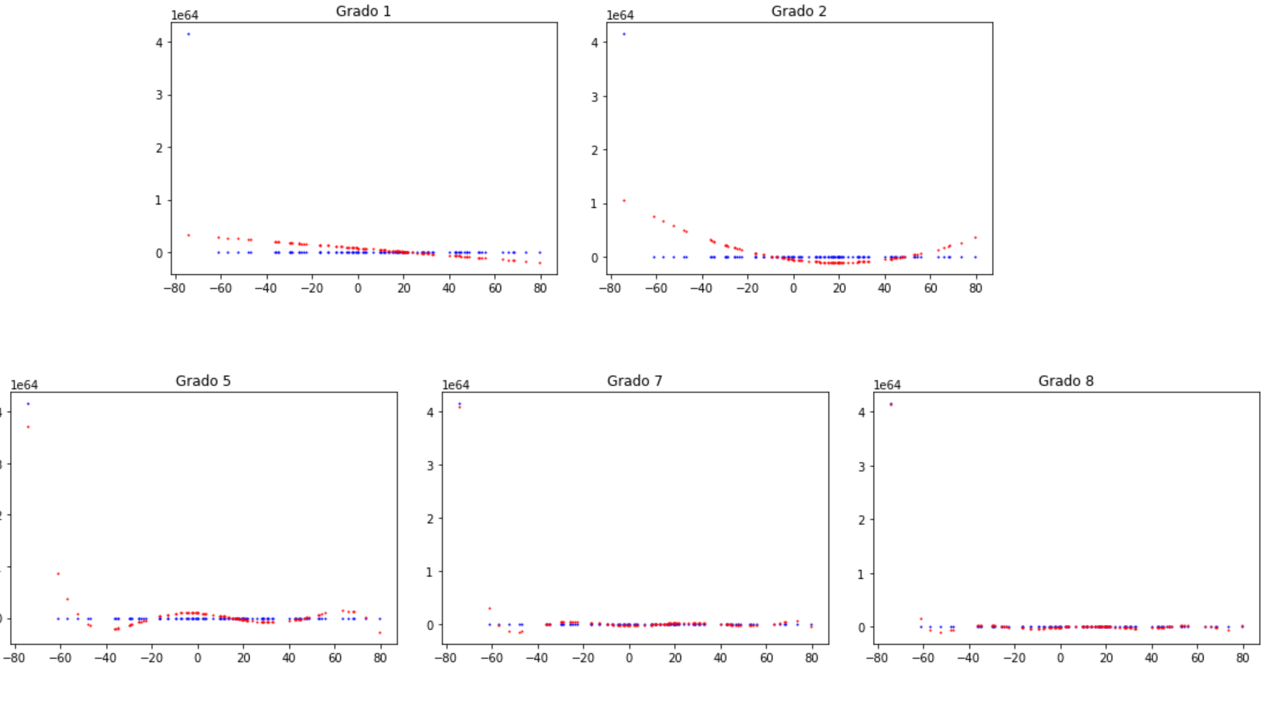
$$\sqrt{y_i} = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + u_t$$

Modelos No Lineales

$$y_i = \beta_0 + x_i^{\beta_1} + u_t$$

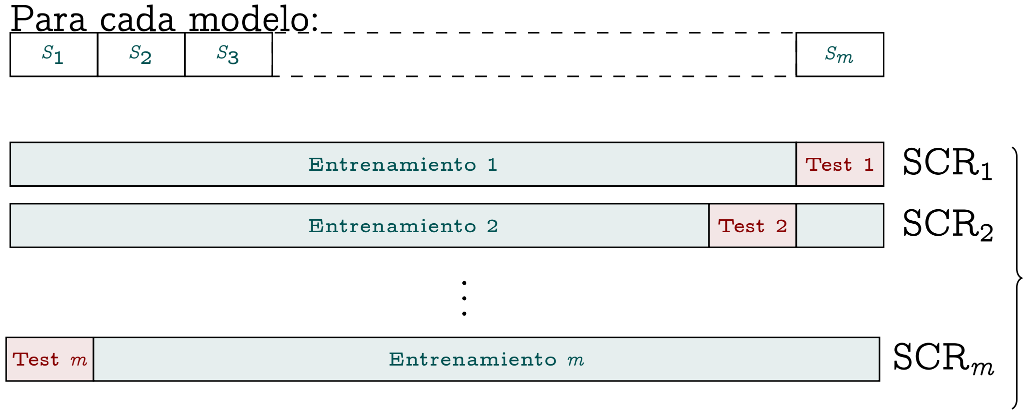
Ajustes Polinomiales

$$y = \text{polinomio}_r(X)$$



Validación Cruzada

Para cada modelo:



Nos quedaríamos con el modelo tal que $\overline{SCR}$ sea más pequeño.

Tests de Hipótesis: Harvey-Collier

$$U_j = \frac{y_j - x_j \hat{\beta}_j}{\sqrt{1 + x_j (X_j^t X_j) x_h}}, j = k+1, \ldots, n.$$

donde $x_j = [1 X_{1j} \ldots X_{kj}]$, $X_j = [x_1 \ldots x_j]$ $\hat{\beta}_j$ coeficientes estimados para las j primeras observaciones.

$$\xi = \frac{\sum_{j=1}^{k+1} U_j}{\sqrt{(n-k)(n-k-1) \sum_{j=k+1}^{n} (U_j - \overline{U}_j)^2}} \sim t_{n-k+1}$$

H_0 : El modelo de regresión es lineal (F-test)

`statsmodels.stats.diagnostic.linear_harvey_collier(mco)`

Tests de Hipótesis: RESET de Ramsey

$$y = X\beta + \gamma_2 \hat{y}^2 + \gamma_3 \hat{y}^3 + \cdots + \gamma_k \hat{y}^k + u$$

$$H_0 : \gamma_2 = \cdots = \gamma_k = 0 \quad (F - test)$$

```
import statsmodels.stats.outliers_influence as oi
statsmodels.stats.outliers_influence.reset_ramsey(mco,
degree=k)
```

Cobb-Douglas

$$P = f(K, L)$$

donde:

- P: Producción de un bien (en unidades monetarias).
- L: Trabajadores (número de personas/horas trabajadas). labor input (the total number of person-hours worked in a year)
- K: Coste de instalaciones, maquinaria, equipamientos.

Hipótesis:

1. $f(0, 0) = 0$.
2. La productividad marginal de L debe ser proporcional a la cantidad producidad por unidad de trabajo
3. La productividad marginal de K debe ser proporcional a la cantidad producidad por unidad de maquinaria.

Cobb-Douglas

$$P = f(K, L)$$

1. $f(0,0) = 0$.
2. La productividad marginal de L debe ser proporcional a la cantidad producidad por unidad de trabajo
3. La productividad marginal de K debe ser proporcional a la cantidad producidad por unidad de maquinaria.

---

1. $f(0,0) = 0$.
2. $\frac{\partial P}{\partial L} = \alpha \frac{P}{L} \Rightarrow P = C_1(K)L^{\alpha}$
3. $\frac{\partial K}{\partial L} = \beta \frac{K}{L} \Rightarrow P = C_2(L)K^{\beta}$

$$P = \kappa L^{\alpha} K^{\beta}$$

Cobb-Douglas

$$P = \kappa L^{\alpha} K^{\beta}$$

- α y β son las elasticidades de la mano de obra de la tecnología, respectivamente.
- Las elasticidades miden la respuesta a la producción a cambios en L o K, ceteris paribus.
- $\alpha + \beta = 1$: Retorno constante a escala $\rightarrow$ Si L y K incrementan en $a\%$, entonces P incrementa $a\%$.
- α y β se pueden interpretar como la proporción de productividad que comparten K y L.

¿Cómo estimar α y β?

$$\log(P) = \log(\kappa) + \alpha \log(L) + \beta \log(K)$$

$$P' = \kappa + \alpha L' + \beta K' + u$$