

## Problema 1

Se busca estudiar la relación entre dos variables X e Y.

- X e Y variables cuantitativas. ¿Qué herramientas estadísticas y gráficas utilizaría? ¿Qué tipo de relaciones podría detectar? Describa.
- Al realizar una regresión de la variable Y en función de una variable X se obtiene el gráfico (1) cómo el grafo entre  $y$  e  $\hat{y}$ . El gráfico (2), corresponde a la gráfica de  $y$  v/s  $\hat{\varepsilon}$ . Según su intuición, ¿Es posible mejorar el modelo? Explique claramente que es lo que haría para mejorarlo.

Gráfico 1:  $y$  v/s  $\hat{y}$

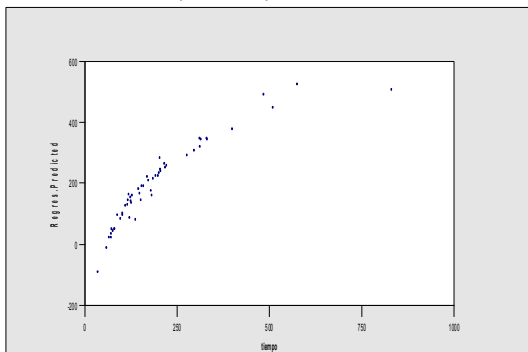
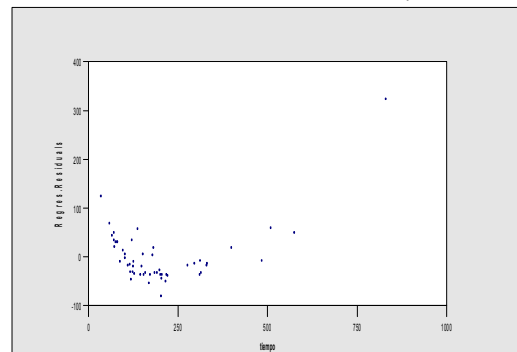


Gráfico 2:  $y$  v/s  $\hat{\varepsilon}$



PAUTA:

- Para medir la relación entre dos variables cuantitativas conviene hacer un gráfico de dispersión. Este permite, en primer lugar, comparar las observaciones, ver en qué difieren y por otra parte, estudiar la relación entre las dos variables. Como el coeficiente de correlación lineal permite solamente de medir el grado de relación de tipo **lineal**, el gráfico de dispersión podrá orientar en buscar eventualmente otra relación si se detecta que no hay linealidad, pero otro tipo de relación. Si vale +1 se tiene una relación estrictamente lineal con pendiente positiva. Si vale -1 se tiene una relación estrictamente lineal con pendiente negativa. Si está cercano a +1 hay una relación de tendencia lineal con pendiente positiva (salvo excepción con mezcla de poblaciones o observaciones atípicas). Si vale 0 no hay relación lineal, pero puede haber una relación funcional no lineal como, por ejemplo, una cuadrática.
- En el gráfico 1, se esperaría los puntos vecinos a la primera bisectriz y en el gráfico 2 que los puntos se distribuyen uniformemente, es decir que los residuos no dependen de la variable Y. Aquí observa una clara tendencia de los residuos con Y. Esto indica probablemente que la relación entre X e Y no es lineal. Es posible que sea logarítmica. Se podría probar de hacer la regresión de  $\log(Y)$  con X.