

7. Regresión lineal simple con datos de “performance”.

Consideraremos los datos en la base `performance.csv` y las variables: y = academic performance of the school (`api00`) y x = percentage of students receiving free meals (`meals`). Estos datos corresponden a una muestra aleatoria de 400 escuelas primarias en California, en donde por escuela se realizaron mediciones que tienen que ver con su desempeño en el año 2000.

i) Regresión lineal simple y verificación de supuestos.

Ajustaremos un modelo de regresión lineal simple del desempeño escolar (`api00`) en función del porcentaje de estudiantes que recibieron desayunos gratuitos en las escuelas (`meals`).

Table 1:

Dependent variable:	
	api00
meals	-4.015*** s.e.(0.097) p-value: <2e-16
Constant	889.783*** (6.622) p-value: <2e-16
Observations	400
R ²	0.811
Adjusted R ²	0.811
Residual Std. Error	61.877 (df = 398)
F Statistic	1,710.691*** (df = 1; 398); p-value: < 2.2e-16
Note: *p<0.1; **p<0.05; ***p<0.01	

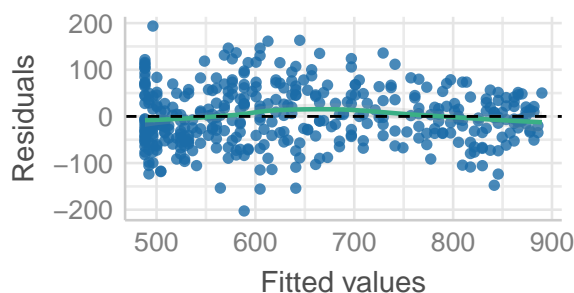
En el siguiente Cuadro se pueden observar las pruebas de Shapiro-Wilk, Breusch-Pagan y Durbin-Watson, en el primer caso de la normalidad el p-value asociado es mayor a 0.05, por lo que no hay evidencia para rechazar las hipótesis nulas de normalidad, sin embargo hay problemas de homoscedasticidad y no autocorrelación.

	1
Normality (Shapiro-Wilk)	0.618
Homoscedasticity (Breusch-Pagan)	0.002
Autocorrelation of residuals (Durbin-Watson)	0.000

También podemos observar de forma gráfica estos resultados.

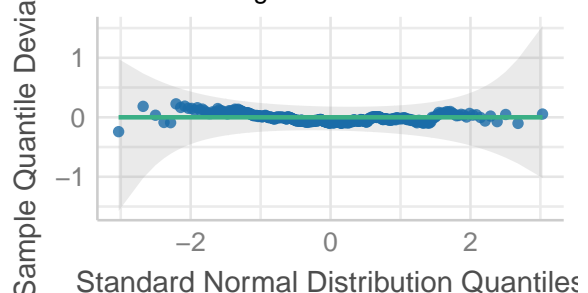
Linearity

Reference line should be flat and horizontal



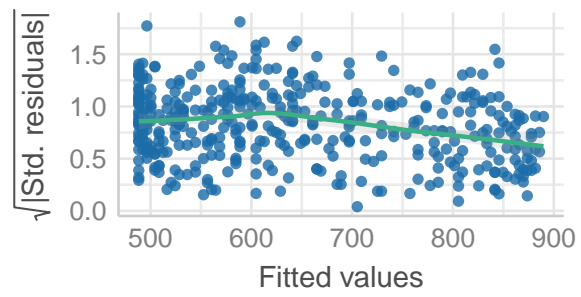
Normality of Residuals

Dots should fall along the line



Homogeneity of Variance

Reference line should be flat and horizontal



Influential Observations

Points should be inside the contour lines

