

5.1) LINEALIDAD

- Si tenemos réplicas (varias observaciones de Y para un mismo valor de los regresores, se pueden realizar contrastes de Linealidad.
- En la práctica no se suele disponer de réplicas
- Nos ayudamos de métodos gráficos:
 - En regresión lineal simple \rightarrow diagrama de dispersión de Y frente a X .
 - En regresión múltiple \rightarrow diagrama de dispersión de los residuos frente a cada regresor.

Hipótesis del Modelo de Regresión Lineal Múltiple

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \dots + \beta_k X_{i,k} + \varepsilon_i \quad i=1, \dots, n$$

$n \rightarrow$ "N° de observaciones"

1°) ε_i son NORMALES $\Rightarrow Y_i$ NORMALES

2°) $\text{Var}(\varepsilon_i) = \sigma^2 \quad \forall i$

3°) ε_i son independientes $\forall i$

⊗ Para validar el modelo de regresión múltiple, usaremos los residuos que resultan del modelo ajustado.

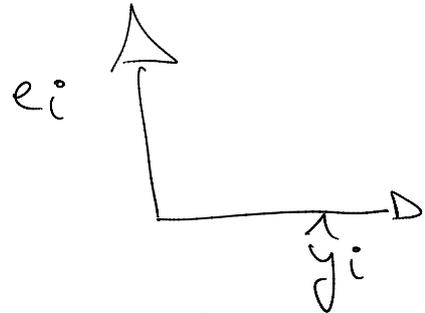
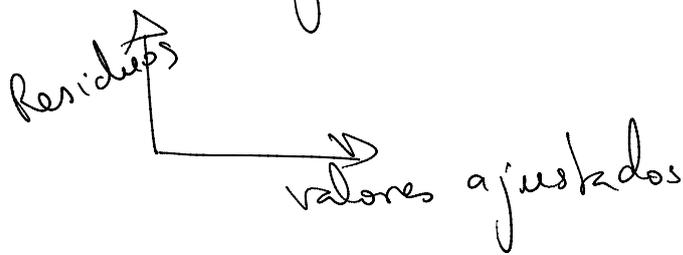
¿Podemos aceptar como válidas las hipótesis del modelo de RLM?

1) NORMALIDAD

- Las perturbaciones aleatorias ε_i deben seguir una distribución aprox. Normal.
- Usando los residuos del modelo ajustado, denotados por ε_i , haremos contrastes de Kolmogorov - Smirnov y Shapiro - Wilks.
- También podemos apoyarnos en el histograma y en los gráficos Q-Q.

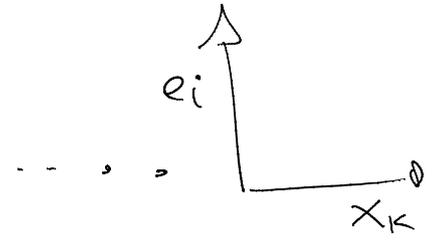
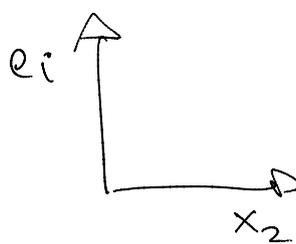
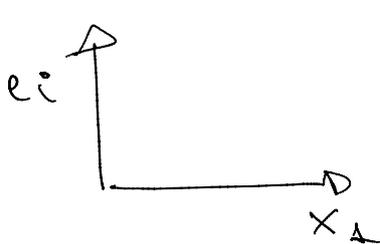
2) HOMOCEDASTICIDAD

- Las perturbaciones ϵ_i deben tener la misma varianza σ^2 .
- Se podrían realizar contrastes pero necesitaríamos réplicas.
- Usaremos métodos gráficos:
 - Diagrama de dispersión de residuos frente a valores ajustados:



Debe aparecer una dispersión aleatoria y de dispersión aprox. constante.

- Diagrama de dispersión de residuos frente a cada regresor:



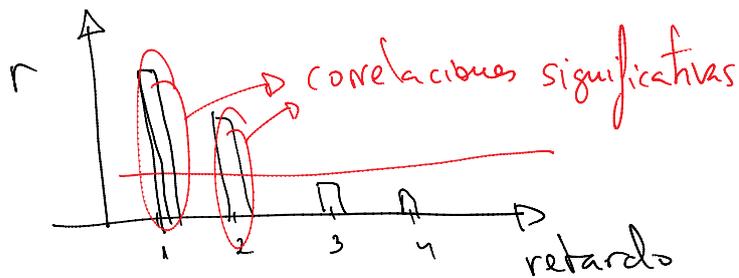
Debe aparecer una dispersión aleatoria y de dispersión aprox. constante.

3) INDEPENDENCIA

- Los ε_i deben ser v.a. independientes.
- Podemos usar 3 herramientas:
 - Gráfico secuencial de residuos \rightarrow no deben aparecer patrones de comportamiento.
 - Estadístico Durbin-Watson

$$DW \in [1.5, 2.5]$$

- Correlograma de los residuos



No deben aparecer correlaciones significativas

Multicolinealidad

viernes, 16 de octubre de 2009

15:09

Multicolinealidad: ¡ Relación lineal
entre los regresores del Modelo final!
¡ No se puede dar!

- Si $VIF(FIV) > 7 \Rightarrow$ hay multicolinealidad
- Si $VIF(FIV) < 7 \Rightarrow$ O.K.

CRITERIOS para detectar Obs. Influyentes:

1.º) Residuos tipificados $\in [-3, 3] \Rightarrow$ pts con
residuos $\notin [-3, 3]$ pueden ser influyentes

2.º) Valor de influencia centrado $< 0.5 \Rightarrow$
pts con valor Influence > 0.5 pueden ser influyentes

3.º) Distancia de Cook $< 1 \Rightarrow$ pts con
Distancia de Cook > 1 son influyentes

Ejemplos: Datos Anscombe