Notas de Clase Sobre Regresión Lineal
Regresión Lineal Simple (RLS): Parte III

Nelfi González Alvarez

Profesora Asociada Escuela de Estadística

e-mail: ngonzale@unal.edu.co

Isabel Cristina Ramírez Guevara

Profesora Asociada Escuela de Estadística

e-mail: iscramirezgu@unal.edu.co

Facultad de Ciencias, Universidad Nacional de Colombia Sede Medellín

UNIVERSIDAD NACIONAL DE COLOMBIA

Escuela de Estadística

2021

Contenido I

Contenido

Tipos de diagnósticos

Para la variable predictora

- *Observaciones x extremas que puedan influenciar el ajuste de regresión?* Un análisis descriptivo: box plots, diagramas de puntos.
- *Patrones temporales en los valores x?* Gráfico de *x* vs. tiempo o algún índice de ordenación.

Para el modelo

- *La función de regresión es lineal?* Test de carencia de ajuste (si hay réplicas), gráficos de residuos.
- *Los errores tienen varianza constante?* Tests de homocedasticidad, gráficos de residuos.
- *Los errores son independientes?* Tests de incorrelación y gráfico residuos vs. tiempo (si se conoce orden temporal).
- *Los errores son normales?* Gráfico de probabilidad normal y tests para normalidad, con residuos.
- *Hay observaciones outliers?* Análisis residuos estandarizados, residuos estudentizados.
- *No hace falta una o más variables predictoras?* Gráficos residuos vs. otras variables no consideradas.

Tipos de diagnósticos

Para la variable predictora

- *Observaciones x extremas que puedan influenciar el ajuste de regresión?* Un análisis descriptivo: box plots, diagramas de puntos.
- *Patrones temporales en los valores x?* Gráfico de x vs. tiempo o algún índice de ordenación.

Para el modelo

- *La función de regresión es lineal?* Test de carencia de ajuste (si hay réplicas), gráficos de residuos.
- *Los errores tienen varianza constante?* Tests de homocedasticidad, gráficos de residuos.
- *Los errores son independientes?* Tests de incorrelación y gráfico residuos vs. tiempo (si se conoce orden temporal).
- *Los errores son normales?* Gráficos de probabilidad normal y tests para normalidad con residuos.
- *Hay observaciones outliers?* Análisis residuos estandarizados, residuos estudentizados.
- *No hay falta una o más variables predictoras?* Gráficos residuos vs. otras variables no consideradas.

Tipos de diagnósticos

Para la variable predictora

- *Observaciones x extremas que puedan influenciar el ajuste de regresión?* Un análisis descriptivo: box plots, diagramas de puntos.
- *Patrones temporales en los valores x?* Gráfico de x vs. tiempo o algún índice de ordenación.

Para el modelo

- *La función de regresión es lineal?* Test de carencia de ajuste (si hay réplicas), gráficos de residuos.
- *Los errores tienen varianza constante?* Tests de homocedasticidad, gráficos de residuos.
- *Los errores son independientes?* Tests de incorrelación y gráfico residuos vs. tiempo (si se conoce orden temporal).
- *Los errores son normales?* Gráficos de probabilidad normal y tests para normalidad, con residuos.
- *Hay observaciones outliers?* Análisis residuos estandarizados, residuales estudentizados.
- *No hace falta una o más variables predictoras?* Gráficos residuos vs. otras variables no consideradas.

Tipos de diagnósticos

Para la variable predictora

- *Observaciones x extremas que puedan influenciar el ajuste de regresión?* Un análisis descriptivo: box plots, diagramas de puntos.
- *Patrones temporales en los valores x?* Gráfico de x vs. tiempo o algún índice de ordenación.

Tipos de diagnósticos

Para la variable predictora

- *Observaciones x extremas que puedan influenciar el ajuste de regresión?* Un análisis descriptivo: box plots, diagramas de puntos.
- *Patrones temporales en los valores x?* Gráfico de x vs. tiempo o algún índice de ordenación.

Para el modelo

- *La función de regresión es lineal?* Test de carencia de ajuste (si hay réplicas), gráficos de residuos.
- *Los errores tienen varianza constante?* Tests de homocedasticidad, gráficos de residuos.
- *Los errores son independientes?* Tests de incorrelación y gráfico residuos vs. tiempo (si se conoce orden temporal).
- *Los errores son normales?* Gráfico de probabilidad normal y tests para normalidad, con residuos.
- *Hay observaciones outliers?* Análisis residuos estandarizados, residuos estudentizados.
- *No hace falta una o más variables predictoras?* Gráficos residuos vs. otras variables no consideradas.

Tipos de diagnósticos

Para la variable predictora

- *Observaciones x extremas que puedan influenciar el ajuste de regresión?* Un análisis descriptivo: box plots, diagramas de puntos.
- *Patrones temporales en los valores x?* Gráfico de x vs. tiempo o algún índice de ordenación.

Para el modelo

- *La función de regresión es lineal?* Test de carencia de ajuste (si hay réplicas), gráficos de residuos.
- *Los errores tienen varianza constante?* Tests de homocedasticidad, gráficos de residuos.
- *Los errores son independientes?* Tests de incorrelación y gráfico residuos vs. tiempo (si se conoce orden temporal).
- *Los errores son normales?* Gráfico de probabilidad normal y tests para normalidad, con residuos.
- *Hay observaciones outliers?* Análisis residuos estandarizados, residuos estudentizados.
- *No hace falta una o más variables predictoras?* Gráficos residuos vs. otras variables no consideradas.

Tipos de diagnósticos

Para la variable predictora

- *Observaciones x extremas que puedan influenciar el ajuste de regresión?* Un análisis descriptivo: box plots, diagramas de puntos.
- *Patrones temporales en los valores x?* Gráfico de x vs. tiempo o algún índice de ordenación.

Para el modelo

- *La función de regresión es lineal?* Test de carencia de ajuste (si hay réplicas), gráficos de residuos.
- *Los errores tienen varianza constante?* Tests de homocedasticidad, gráficos de residuos.
- *Los errores son independientes?* Tests de incorrelación y gráfico residuos vs. tiempo (si se conoce orden temporal).
- *Los errores son normales?* Gráfico de probabilidad normal y tests para normalidad, con residuos.
- *Hay observaciones outliers?* Análisis residuos estandarizados, residuos estudentizados.
- *No hace falta una o más variables predictoras?* Gráficos residuos vs. otras variables no consideradas.

Tipos de diagnósticos

Para la variable predictora

- *Observaciones x extremas que puedan influenciar el ajuste de regresión?* Un análisis descriptivo: box plots, diagramas de puntos.
- *Patrones temporales en los valores x?* Gráfico de x vs. tiempo o algún índice de ordenación.

Para el modelo

- *La función de regresión es lineal?* Test de carencia de ajuste (si hay réplicas), gráficos de residuos.
- *Los errores tienen varianza constante?* Tests de homocedasticidad, gráficos de residuos.
- *Los errores son independientes?* Tests de incorrelación y gráfico residuos vs. tiempo (si se conoce orden temporal).
- *Los errores son normales?* Gráfico de probabilidad normal y tests para normalidad, con residuos.
- *Hay observaciones outliers?* Análisis residuos estandarizados, residuos estudentizados.
- *No hace falta una o más variables predictoras?* Gráficos residuos vs. otras variables no consideradas.

Tipos de diagnósticos

Para la variable predictora

- *Observaciones x extremas que puedan influenciar el ajuste de regresión?* Un análisis descriptivo: box plots, diagramas de puntos.
- *Patrones temporales en los valores x?* Gráfico de x vs. tiempo o algún índice de ordenación.

Para el modelo

- *La función de regresión es lineal?* Test de carencia de ajuste (si hay réplicas), gráficos de residuos.
- *Los errores tienen varianza constante?* Tests de homocedasticidad, gráficos de residuos.
- *Los errores son independientes?* Tests de incorrelación y gráfico residuos vs. tiempo (si se conoce orden temporal).
- *Los errores son normales?* Gráfico de probabilidad normal y tests para normalidad, con residuos.
- *Hay observaciones outliers?* Análisis residuos estandarizados, residuos estudentizados.
- *No hace falta una o más variables predictoras?* Gráficos residuos vs. otras variables no consideradas.

Tipos de diagnósticos

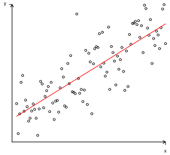
Para la variable predictora

- *Observaciones x extremas que puedan influenciar el ajuste de regresión?* Un análisis descriptivo: box plots, diagramas de puntos.
- *Patrones temporales en los valores x?* Gráfico de x vs. tiempo o algún índice de ordenación.

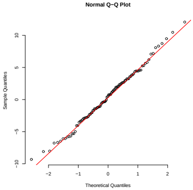
Para el modelo

- *La función de regresión es lineal?* Test de carencia de ajuste (si hay réplicas), gráficos de residuos.
- *Los errores tienen varianza constante?* Tests de homocedasticidad, gráficos de residuos.
- *Los errores son independientes?* Tests de incorrelación y gráfico residuos vs. tiempo (si se conoce orden temporal).
- *Los errores son normales?* Gráfico de probabilidad normal y tests para normalidad, con residuos.
- *Hay observaciones outliers?* Análisis residuos estandarizados, residuos estudentizados.
- *No hace falta una o más variables predictoras?* Gráficos residuos vs. otras variables no consideradas.

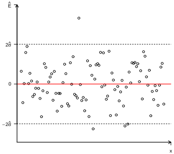
Comportamientos esperados



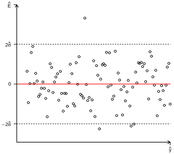
(a)



(b)



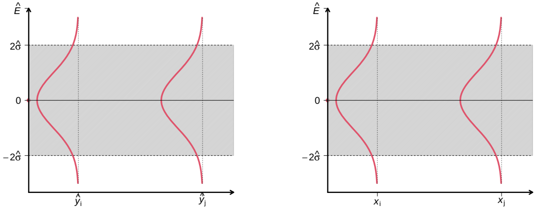
(c)



(d)

Figura 1: Ilustración de un modelo lineal adecuado con varianza constante y supuesto de normalidad válido. (a) gráfico de dispersión; (b) gráfico de probabilidad normal con residuos; (c) gráfico de residuos vs. x; (d) gráfico de residuos vs. $\widehat{y}$.

Por qué se espera un patrón rectangular en la dispersión de residuos?



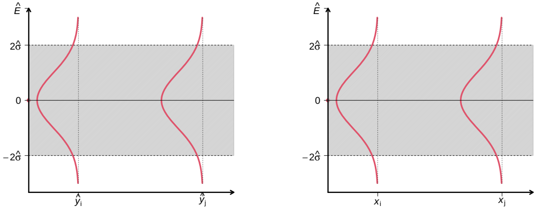
(a) (b)

Figura 2: En (a) $\widehat{E}_i$ vs. $\widehat{y}$ y (b) $\widehat{E}_i$ vs. x, con n grande, en cada nivel de $\widehat{y}$ y de x, se espera que los $\widehat{E}_i$ se comporten como v.a normales, independientes, de media cero y con misma varianza, como consecuencia de que los verdaderos errores satisfacen $E_i \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)$. También se espera que no más de un 5 % de los residuos estén por fuera de las bandas $\pm 2\widehat{\sigma}$.

Nota 1.1

No se grafican $\widehat{E}_i$ vs. y_i, desde que $Cov\left(\widehat{E}_i, Y_i\right) = \left(1 - h_{ii}\right)\sigma^2 \rightarrow \sigma^2$, con $n \rightarrow \infty$, de modo que a mayor σ^2, veríamos en esa gráfica una tendencia lineal con pendiente positiva, y no un gráfico como los ilustrados previamente.

Por qué se espera un patrón rectangular en la dispersión de residuos?



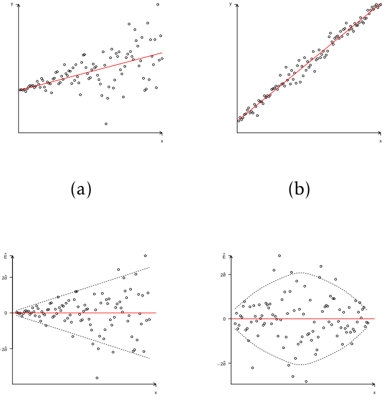
(a) (b)

Figura 2: En (a) $\widehat{E}_i$ vs. $\widehat{y}$ y (b) $\widehat{E}_i$ vs. x, con n grande, en cada nivel de $\widehat{y}$ y de x, se espera que los $\widehat{E}_i$ se comporten como v.a normales, independientes, de media cero y con misma varianza, como consecuencia de que los verdaderos errores satisfacen $E_i \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)$. También se espera que no más de un 5 % de los residuos estén por fuera de las bandas $\pm 2\widehat{\sigma}$.

Nota 1.1

No se grafican $\widehat{E}_i$ vs. y_i, desde que $Cov\left(\widehat{E}_i, Y_i\right) = \left(1 - h_{ii}\right)\sigma^2 \to \sigma^2$, con $n \to \infty$, de modo que a mayor σ^2, veríamos en esa gráfica una tendencia lineal con pendiente positiva, y no un gráfico como los ilustrados previamente.

Ejemplo patrones que indican varianza no constante

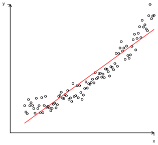


(a) (b)

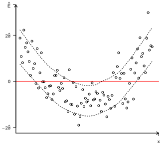
(c) (d)

Figura 3: Patrón de embudo: (a) gráfico de dispersión con recta ajustada; (c) residuos vs. x. Patrón de balón de fútbol americano: (b) gráfico de dispersión con recta ajustada; (d) residuos vs. x.

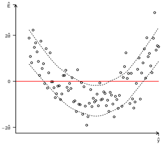
Ejemplo patrón que indica carencia de ajuste



(a)



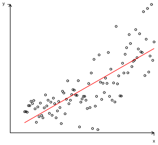
(b)



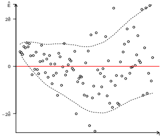
(c)

Figura 4: Ejemplo del caso donde el modelo lineal entre y y x no es adecuado, pero la varianza es constante: (a) gráfico de dispersión con recta ajustada, (b) residuos vs. x y (c) residuos vs. $\widehat{y}$.

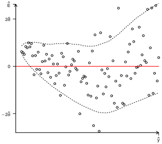
Ejemplo patrón carencia de ajuste junto con varianza no constante



(a)



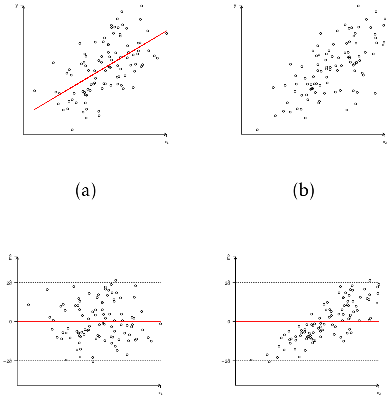
(b)



(c)

Figura 5: Ejemplo del caso donde el modelo lineal entre y y x no es adecuado, ni la varianza es constante: (a) gráfico de dispersión con recta ajustada (b) residuos vs. x y (c) residuos vs. $\widehat{y}$.

Ejemplo patrón cuando hay una variable explicatoria omitida



(a) (b)

(c) (d)

Figura 6: Modelo considerado: $Y_i = \beta_0 + \beta_1 x_1 + E_i$, $E_i \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)$. (a) Ajuste de y vs. x_1; (b) gráf. de y vs. x_2, muestra que también hay relación lineal con x_2 (predictor omitido); (c) $\widehat{E}$ vs. x_1; (d) $\widehat{E}$ vs. x_2, muestra que es necesario incluir a x_2 en el modelo.

Test de carencia de ajuste: No linealidad de la función de regresión

- *El test asume que los valores de Y dado X son:*
  - independientes
  - se distribuyen en forma normal
  - tienen varianza constante
- *Se requiere que en uno o más valores de X haya más de una observación de Y ó réplicas. Usaremos la siguiente notación:*
  - Y_{ij}, la respuesta j-ésima en el i-ésimo nivel de X;
  - x_i, i-ésimo nivel de X; supondremos $i = 1, 2, \ldots, k$;
  - n_i, número de observaciones de Y tomadas en el i-ésimo nivel de X. Por tanto, el total de observaciones n corresponde a

$$n = \sum_{i=1}^{k} n_i. \tag{1}$$

- *El test define* **el modelo lineal general** *que corresponde a*

$$Y_{ij} = \mu_i + E_{ij}, \text{ con } E_{ij} \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right), \; \forall\, i, j, \; i = 1, \ldots, k, \; j = 1, \ldots, n_i, \tag{2}$$

donde $\mu_i = \mathrm{E}\left[Y_{ij}\right]$,

Test de carencia de ajuste: No linealidad de la función de regresión

- *El test asume que los valores de Y dado X son:*
  - independientes
  - se distribuyen en forma normal
  - tienen varianza constante

- *Se requiere que en uno o más valores de X haya más de una observación de Y ó réplicas. Usaremos la siguiente notación:*
  - Y_{ij}, la respuesta j-ésima en el i-ésimo nivel de X;
  - x_i, i-ésimo nivel de X; supondremos $i = 1, 2, \ldots, k$;
  - n_i, número de observaciones de Y tomadas en el i-ésimo nivel de X. Por tanto, el total de observaciones n corresponde a

$$n = \sum_{i=1}^{k} n_i. \tag{1}$$

- *El test define* **el modelo lineal general** *que corresponde a*

$$Y_{ij} = \mu_i + E_{ij}, \text{ con } E_{ij} \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right), \ \forall\, i, j, \ i = 1, \ldots, k, \ j = 1, \ldots, n_i, \tag{2}$$

donde $\mu_i = \mathrm{E}\left[Y_{ij}\right]$,

Test de carencia de ajuste: No linealidad de la función de regresión

- *El test asume que los valores de Y dado X son:*
  - independientes
  - se distribuyen en forma normal
  - tienen varianza constante
- *Se requiere que en uno o más valores de X haya más de una observación de Y ó réplicas. Usaremos la siguiente notación:*
  - Y_{ij}: la respuesta j-ésima en el i-ésimo nivel de X;
  - x_i: i-ésimo nivel de X; supondremos $i = 1, 2, \ldots, k$;
  - n_i: número de observaciones de Y tomadas en el i-ésimo nivel de X. Por tanto, el total de observaciones n corresponde a

$$n = \sum_{i=1}^{k} n_i. \tag{1}$$

- *El test define* **el modelo lineal general** *que corresponde a*

$$Y_{ij} = \mu_i + E_{ij}, \text{ con } E_{ij} \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right), \ \forall\, i, j,\ i = 1, \ldots, k,\ j = 1, \ldots, n_i, \tag{2}$$

donde $\mu_i = \mathrm{E}\left[Y_{ij}\right]$,

Tabla 1: Comparación entre el MRLS y el modelo lineal general

Características	MRLS	Modelo lineal general
Ecuación	$Y_{ij} = \beta_0 + \beta_1 x_i + E_{ij}$ $E_{ij} \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)\ \forall\, i = 1, \ldots, k,\ j = 1, \ldots, n_i$	$Y_{ij} = \mu_i + E_{ij}$ $E_{ij} \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)\ \forall\, i = 1, \ldots, k,\ j = 1, \ldots, n_i$
Respuesta media	$\mathrm{E}\left[Y_{ij}\right] = \beta_0 + \beta_1 x_i$	$\mathrm{E}\left[Y_{ij}\right] = \mu_i$
Respuesta estimada	$\widehat{Y}_{ij} = \widehat{Y}_{i\bullet} = \widehat{\beta}_0 + \widehat{\beta}_1 x_i$ Todas las obs. en mismo nivel x_i tienen misma respuesta estimada: $\widehat{Y}_{i\bullet}$	$\widehat{Y}_{ij} = \bar{Y}_{i\bullet} = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$ Todas las obs. en mismo nivel x_i tienen misma respuesta estimada: $\bar{Y}_{i\bullet}$
Residuos de ajuste	$\widehat{E}_{ij} = Y_{ij} - \widehat{Y}_{i\bullet}$	$\widehat{E}_{ij} = Y_{ij} - \bar{Y}_{i\bullet}$
Suma cuad. residuos	$\mathrm{SSE} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \widehat{Y}_{i\bullet}\right)^2$ $\mathrm{g.l}\,(\mathrm{SSE}) = n - 2$	$\mathrm{SSPE} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \bar{Y}_{i\bullet}\right)^2$ Conocido como suma de cuadrados de error puro $\mathrm{g.l}\,(\mathrm{SSPE}) = n - k$

Tenemos la siguiente descomposición bajo supuestos $E_{ij} \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)\ \forall\, i = 1, \ldots, k\ j = 1, \ldots, n_i$:

$$\begin{aligned} \mathrm{SSE} &= \underbrace{\mathrm{SSPE}}_{\text{S.C error puro}} + \underbrace{\mathrm{SSLOF}}_{\text{S.C carencia de ajuste}} \\ \underbrace{\mathrm{g.l}\,(\mathrm{SSE})}_{n-2} &= \underbrace{\mathrm{g.l}\,(\mathrm{SSPE})}_{n-k} + \underbrace{\mathrm{g.l}\,(\mathrm{SSLOF})}_{k-2} \end{aligned} \tag{3}$$

Tabla 1: Comparación entre el MRLS y el modelo lineal general

Características	MRLS	Modelo lineal general
Ecuación	$Y_{ij} = \beta_0 + \beta_1 x_i + E_{ij}$ $E_{ij} \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)\ \forall\, i = 1, \ldots, k,\ j = 1, \ldots, n_i$	$Y_{ij} = \mu_i + E_{ij}$ $E_{ij} \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)\ \forall\, i = 1, \ldots, k,\ j = 1, \ldots, n_i$
Respuesta media	$\mathrm{E}\left[Y_{ij}\right] = \beta_0 + \beta_1 x_i$	$\mathrm{E}\left[Y_{ij}\right] = \mu_i$
Respuesta estimada	$\widehat{Y}_{ij} = \widehat{Y}_{i\bullet} = \widehat{\beta}_0 + \widehat{\beta}_1 x_i$ Todas las obs. en mismo nivel x_i tienen misma respuesta estimada: $\widehat{Y}_{i\bullet}$	$\widehat{Y}_{ij} = \bar{Y}_{i\bullet} = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$ Todas las obs. en mismo nivel x_i tienen misma respuesta estimada: $\bar{Y}_{i\bullet}$
Residuos de ajuste	$\widehat{E}_{ij} = Y_{ij} - \widehat{Y}_{i\bullet}$	$\widehat{E}_{ij} = Y_{ij} - \bar{Y}_{i\bullet}$
Suma cuad. residuos	$\mathrm{SSE} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \widehat{Y}_{i\bullet}\right)^2$ $\text{g.l}\,(\mathrm{SSE}) = n - 2$	$\mathrm{SSPE} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \bar{Y}_{i\bullet}\right)^2$ Conocido como suma de cuadrados de error puro $\text{g.l}\,(\mathrm{SSPE}) = n - k$

Tenemos la siguiente descomposición bajo supuestos $E_{ij} \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)\ \forall\, i = 1, \ldots, k\ j = 1, \ldots, n_i$:

$$\begin{aligned} \mathrm{SSE} &= \underbrace{\mathrm{SSPE}}_{\textbf{S.C error puro}} + \underbrace{\mathrm{SSLOF}}_{\textbf{S.C carencia de ajuste}} \\ \underbrace{\text{g.l}\,(\mathrm{SSE})}_{n-2} &= \underbrace{\text{g.l}\,(\mathrm{SSPE})}_{n-k} + \underbrace{\text{g.l}\,(\mathrm{SSLOF})}_{k-2} \end{aligned} \tag{3}$$

Despejando tenemos para la suma de cuadrados de carencia de ajuste:

$$\text{SSLOF} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\bar{Y}_{i\bullet} - \widehat{Y}_{i\bullet}\right)^2 = \sum_{i=1}^{k} n_i\left(\bar{Y}_{i\bullet} - \widehat{Y}_{i\bullet}\right)^2, \text{ con g.l}\,(\text{SSLOF}) = k-2. \tag{4}$$

Note que en SSLOF se comparan las estimaciones de la respuesta media de los modelos lineal general y MRLS, mediante las diferencias $\bar{Y}_{i\bullet} - \widehat{Y}_{i\bullet}$

Tabla 2: ANOVA para modelo de regresión y carencia de ajuste

Fuente	SC	gl	CM	ECM	F calculada	Valor P
Regresión	SSR	1	$\text{MSR} = \frac{\text{SSR}}{1}$	$\sigma^2 + \beta_1^2 S_{xx}$	$F_{0,\text{reg}}$	$P\left(f_{1,n-2} > F_{0,\text{reg}}\right)$
LOF	SSLOF	$k-2$	$\text{MSLOF} = \frac{\text{SSLOF}}{k-2}$	$\sigma^2 + \frac{\sum_{i=1}^{k} n_i Q_i^2}{k-2}$	$F_{0,\text{LOF}}$	$P\left(f_{k-2,n-k} > F_{0,\text{LOF}}\right)$
Error Puro	SSPE	$n-k$	$\text{MSPE} = \frac{\text{SSPE}}{n-k}$	σ^2	Test LOF: con $F_{0,\text{LOF}}$ probar que	
Error	SSE	$n-2$	$\text{MSE} = \frac{\text{SSE}}{n-2}$	σ^2	$H_0: \mu_i = \beta_0 + \beta_1 x_i$ vs.	
Total	SST	$n-1$	$\text{MST} = \frac{\text{SST}}{n-1}$		$\text{H}_1: \mu_i \neq \beta_0 + \beta_1 x_i$	

SC: Suma de cuadrados, CM: cuadrado medio y ECM cuadrado medio esperado es decir E[CM], gl: grados de libertad.
$F_{0,\text{reg}} = \frac{\text{MSR}}{\text{MSE}}$ para test de significancia del modelo de regresión lineal simple.
LOF: Carencia de ajuste, $Q_i = \mu_i - \left(\beta_0 + \beta_1 x_i\right)$, $F_{0,\text{LOF}} = \frac{\text{MSLOF}}{\text{MSPE}}$
bajo H_0 y validez supuestos sobre E_{ij}, Estadístico de prueba en test LOF cumple que $F_{0,\text{LOF}} \sim f_{k-2,n-k}$

Despejando tenemos para la suma de cuadrados de carencia de ajuste:

$$\text{SSLOF} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\bar{Y}_{i\bullet} - \widehat{Y}_{i\bullet}\right)^2 = \sum_{i=1}^{k} n_i\left(\bar{Y}_{i\bullet} - \widehat{Y}_{i\bullet}\right)^2, \text{ con g.l (SSLOF)} = k-2. \tag{4}$$

Note que en SSLOF se comparan las estimaciones de la respuesta media de los modelos lineal general y MRLS, mediante las diferencias $\bar{Y}_{i\bullet} - \widehat{Y}_{i\bullet}$

Tabla 2: ANOVA para modelo de regresión y carencia de ajuste

Fuente	SC	gl	CM	ECM	F calculada	Valor P
Regresión	SSR	1	MSR $= \frac{\text{SSR}}{1}$	$\sigma^2 + \beta_1^2 S_{xx}$	$F_{0,\text{reg}}$	$P\left(f_{1,n-2} > F_{0,\text{reg}}\right)$
LOF	SSLOF	$k-2$	MSLOF $= \frac{\text{SSLOF}}{k-2}$	$\sigma^2 + \frac{\sum_{i=1}^{k} n_i Q_i^2}{k-2}$	$F_{0,\text{LOF}}$	$P\left(f_{k-2,n-k} > F_{0,\text{LOF}}\right)$
Error Puro	SSPE	$n-k$	MSPE $= \frac{\text{SSPE}}{n-k}$	σ^2	Test LOF: con $F_{0,\text{LOF}}$ probar que	
Error	SSE	$n-2$	MSE $= \frac{\text{SSE}}{n-2}$	σ^2	$H_0: \mu_i = \beta_0 + \beta_1 x_i$ vs.	
Total	SST	$n-1$	MST $= \frac{\text{SST}}{n-1}$		$\text{H}_1: \mu_i \neq \beta_0 + \beta_1 x_i$	

SC: Suma de cuadrados, CM: cuadrado medio y ECM cuadrado medio esperado es decir E[CM], gl: grados de libertad.

$F_{0,\text{reg}} = \frac{\text{MSR}}{\text{MSE}}$ para test de significancia del modelo de regresión lineal simple.

LOF: Carencia de ajuste, $Q_i = \mu_i - \left(\beta_0 + \beta_1 x_i\right)$, $F_{0,\text{LOF}} = \frac{\text{MSLOF}}{\text{MSPE}}$

bajo H_0 y validez supuestos sobre E_{ij}, Estadístico de prueba en test LOF cumple que $F_{0,\text{LOF}} \sim f_{k-2,n-k}$

Algunas consideraciones en test LOF

- *Cálculo del SSPE sólo usa aquellos niveles x_i de X en los cuales $n_i > 1$, entonces k es el número de niveles de X donde $n_i > 1$.*
- Test LOF aplica con cualquier función de regresión lineal, g.l(SSLOF) $= k - p$, con p el número de parámetros en la función de regresión propuesta y con $k > p$.
- *Como estimador de σ^2 cuando no hay carencia de ajuste, se usa el MSE de la regresión y no el MSPE como un estimador de la varianza, debido a que el primero tiene más grados de libertad.*
- Cualquier inferencia sobre los parámetros del modelo lineal sólo debe llevarse a cabo luego de haber probado que el modelo de regresión lineal es apropiado.

Nota 1.2

Ver Código R 3.1 Capítulo 3 de notas de clase, para implementación del test de carencia de ajuste con función `anova( )` *combinada con* `aov( )` *y con función* `summary( )` *combinada con función* `rsm( )` *de la librería del mismo nombre.*

Algunas consideraciones en test LOF

- *Cálculo del SSPE sólo usa aquellos niveles x_i de X en los cuales $n_i > 1$, entonces k es el número de niveles de X donde $n_i > 1$.*
- Test LOF aplica con cualquier función de regresión lineal, $\text{g.l}(\text{SSLOF}) = k - p$, con p el número de parámetros en la función de regresión propuesta y con $k > p$.
- *Como estimador de σ^2 cuando no hay carencia de ajuste, se usa el MSE de la regresión y no el MSPE como un estimador de la varianza, debido a que el primero tiene más grados de libertad.*
- Cualquier inferencia sobre los parámetros del modelo lineal sólo debe llevarse a cabo luego de haber probado que el modelo de regresión lineal es apropiado.

Nota 1.2

Ver Código R 3.1 Capítulo 3 de notas de clase, para implementación del test de carencia de ajuste con función `anova( )` combinada con `aov( )` y con función `summary( )` combinada con función `rsm( )` de la librería del mismo nombre.

Algunas consideraciones en test LOF

- *Cálculo del SSPE sólo usa aquellos niveles x_i de X en los cuales $n_i > 1$, entonces k es el número de niveles de X donde $n_i > 1$.*
- Test LOF aplica con cualquier función de regresión lineal, $\text{g.l}(\text{SSLOF}) = k - p$, con p el número de parámetros en la función de regresión propuesta y con $k > p$.
- *Como estimador de σ^2 cuando no hay carencia de ajuste, se usa el MSE de la regresión y no el MSPE como un estimador de la varianza, debido a que el primero tiene más grados de libertad.*
- Cualquier inferencia sobre los parámetros del modelo lineal sólo debe llevarse a cabo luego de haber probado que el modelo de regresión lineal es apropiado.

Nota 1.2

Ver Código R 3.1 Capítulo 3 de notas de clase, para implementación del test de carencia de ajuste con función `anova( )` *combinada con* `aov( )` *y con función* `summary( )` *combinada con función* `rsm( )` *de la librería del mismo nombre.*

Algunas consideraciones en test LOF

- *Cálculo del SSPE sólo usa aquellos niveles x_i de X en los cuales $n_i > 1$, entonces k es el número de niveles de X donde $n_i > 1$.*
- Test LOF aplica con cualquier función de regresión lineal, $\text{g.l}(\text{SSLOF}) = k - p$, con p el número de parámetros en la función de regresión propuesta y con $k > p$.
- *Como estimador de σ^2 cuando no hay carencia de ajuste, se usa el MSE de la regresión y no el MSPE como un estimador de la varianza, debido a que el primero tiene más grados de libertad.*
- Cualquier inferencia sobre los parámetros del modelo lineal sólo debe llevarse a cabo luego de haber probado que el modelo de regresión lineal es apropiado.

Nota 1.2

Ver Código R 3.1 Capítulo 3 de notas de clase, para implementación del test de carencia de ajuste con función `anova( )` combinada con `aov( )` y con función `summary( )` combinada con función `rsm( )` de la librería del mismo nombre.

Algunas consideraciones en test LOF

- *Cálculo del SSPE sólo usa aquellos niveles x_i de X en los cuales $n_i > 1$, entonces k es el número de niveles de X donde $n_i > 1$.*
- Test LOF aplica con cualquier función de regresión lineal, g.l(SSLOF) $= k - p$, con p el número de parámetros en la función de regresión propuesta y con $k > p$.
- *Como estimador de σ^2 cuando no hay carencia de ajuste, se usa el MSE de la regresión y no el MSPE como un estimador de la varianza, debido a que el primero tiene más grados de libertad.*
- Cualquier inferencia sobre los parámetros del modelo lineal sólo debe llevarse a cabo luego de haber probado que el modelo de regresión lineal es apropiado.

Nota 1.2

Ver Código R 3.1 Capítulo 3 de notas de clase, para implementación del test de carencia de ajuste con función `anova( )` combinada con `aov( )` y con función `summary( )` combinada con función `rsm( )` de la librería del mismo nombre.

Algunas consideraciones en test LOF

- *Cálculo del SSPE sólo usa aquellos niveles x_i de X en los cuales $n_i > 1$, entonces k es el número de niveles de X donde $n_i > 1$.*
- Test LOF aplica con cualquier función de regresión lineal, $\text{g.l}(\text{SSLOF}) = k - p$, con p el número de parámetros en la función de regresión propuesta y con $k > p$.
- *Como estimador de σ^2 cuando no hay carencia de ajuste, se usa el MSE de la regresión y no el MSPE como un estimador de la varianza, debido a que el primero tiene más grados de libertad.*
- Cualquier inferencia sobre los parámetros del modelo lineal sólo debe llevarse a cabo luego de haber probado que el modelo de regresión lineal es apropiado.

Nota 1.2

Ver Código R 3.1 Capítulo 3 de notas de clase, para implementación del test de carencia de ajuste con función `anova( )` *combinada con* `aov( )` *y con función* `summary( )` *combinada con función* `rsm( )` *de la librería del mismo nombre.*

Qué hacer si hay carencia de ajuste significativa?

- Desarrollar un modelo más apropiado (tal vez un modelo de regresión no lineal).
- *Transformar las variables X y/o Y para tener un MRLS en los datos transformados.*
- Regresión no paramétrica, para explorar y/o confirmar la forma de la función de regresión.

Qué hacer si hay carencia de ajuste significativa?

- Desarrollar un modelo más apropiado (tal vez un modelo de regresión no lineal).
- *Transformar las variables X y/o Y para tener un MRLS en los datos transformados.*
- Regresión no paramétrica, para explorar y/o confirmar la forma de la función de regresión.

Qué hacer si hay carencia de ajuste significativa?

- Desarrollar un modelo más apropiado (tal vez un modelo de regresión no lineal).
- *Transformar las variables X y/o Y para tener un MRLS en los datos transformados.*
- Regresión no paramétrica, para explorar y/o confirmar la forma de la función de regresión.

Tests de homocedasticidad

Considerando el modelo: $Y_i = \beta_0 + \beta_1 x_i + E_i$, con

$$E_i \overset{\text{ind.}}{\sim} N\left(0, \sigma_i^2\right), \quad i = 1, 2, \ldots, n, \tag{5}$$

donde $\text{Var}[E_i] = \sigma_i^2$, probaremos que,

$$\begin{aligned} H_0 &: \sigma_i^2 = \sigma^2 \ \forall\, i \\ H_1 &: \sigma_i^2 \neq \sigma^2 \text{ para algún } i. \end{aligned} \tag{6}$$

Tests:

- Levene modificado o test Brown-Forsythe
- Breusch-Pagan
- Breusch-Pagan estudentizado

Tests de homocedasticidad

Considerando el modelo: $Y_i = \beta_0 + \beta_1 x_i + E_i$, con

$$E_i \overset{\text{ind.}}{\sim} N\left(0, \sigma_i^2\right), \quad i = 1, 2, \ldots, n, \tag{5}$$

donde $\text{Var}\left[E_i\right] = \sigma_i^2$, probaremos que,

$$\begin{aligned} H_0 &: \sigma_i^2 = \sigma^2 \ \forall \ i \\ H_1 &: \sigma_i^2 \neq \sigma^2 \text{ para algún } i. \end{aligned} \tag{6}$$

Tests:

- Levene modificado o test Brown-Forsythe
- Breusch-Pagan
- Breusch-Pagan estudentizado

Levene modificado o test Brown-Forsythe (BF)

Divide en tres o cuatro grupos los residuos de acuerdo a niveles observados o valores de la variable x, cuando son muchos los niveles y toma grupos de los extremos:

Grupo 1: Niveles de X bajos
con n_1 casos

Grupo 2: Niveles de X altos
con n_2 casos

$x_{(1)}$ $x_{(k)}$ $x_{(l)}$ $x_{(n)}$

$x_{(k)}, x_{(l)}$ tales que: $\quad x_{(k)} - x_{(1)} \approx x_{(n)} - x_{(1)}$

Si no son muchos los niveles o valores distintos en x, divide en sólo dos grupos,

Grupo 1: Niveles de X bajos
con n_1 casos

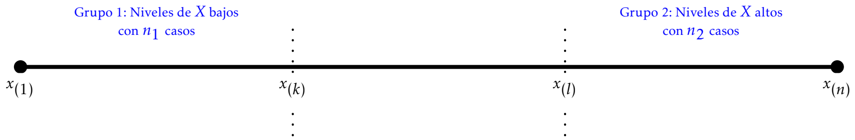
Grupo 2: Niveles de X altos
con n_2 casos

$x_{(1)}$ $x_{(k)}$ $x_{(n)}$

$x_{(k)}$ tal que: $\quad x_{(k)} - x_{(1)} \approx x_{(n)} - x_{(k)}$

Levene modificado o test Brown-Forsythe (BF)

Divide en tres o cuatro grupos los residuos de acuerdo a niveles observados o valores de la variable x, cuando son muchos los niveles y toma grupos de los extremos:



$x_{(k)}, x_{(l)}$ tales que: $\quad x_{(k)} - x_{(1)} \approx x_{(n)} - x_{(1)}$

Si no son muchos los niveles o valores distintos en x, divide en sólo dos grupos,

Grupo 1: Niveles de X bajos
con n_1 casos

Grupo 2: Niveles de X altos
con n_2 casos

$x_{(1)}$ $\quad x_{(k)}$ $\quad x_{(n)}$

$x_{(k)}$ tal que: $\quad x_{(k)} - x_{(1)} \approx x_{(n)} - x_{(k)}$

Levene modificado o test Brown-Forsythe (BF)

Divide en tres o cuatro grupos los residuos de acuerdo a niveles observados o valores de la variable x, cuando son muchos los niveles y toma grupos de los extremos:

Grupo 1: Niveles de X bajos con n_1 casos

Grupo 2: Niveles de X altos con n_2 casos

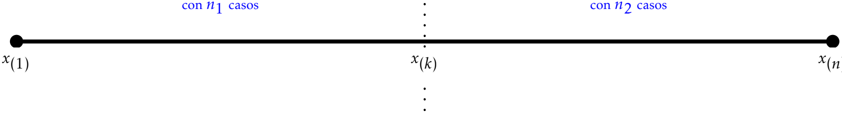
$x_{(1)}$ $x_{(k)}$ $x_{(l)}$ $x_{(n)}$

$x_{(k)}, x_{(l)}$ tales que: $x_{(k)} - x_{(1)} \approx x_{(n)} - x_{(1)}$

Si no son muchos los niveles o valores distintos en x, divide en sólo dos grupos,

Grupo 1: Niveles de X bajos con n_1 casos

Grupo 2: Niveles de X altos con n_2 casos



$x_{(1)}$ $x_{(k)}$ $x_{(n)}$

$x_{(k)}$ tal que: $x_{(k)} - x_{(1)} \approx x_{(n)} - x_{(k)}$

Consideraciones del test BF:

- *Test robusto a no normalidad*
- *Supone comportamiento monótono de σ_i^2 vs. x_i*
- *Requiere n grande para que sean despreciables correlaciones entre residuos*
- *Funciona como un test t para comparación de dos medias poblacionales.*
- *$n_1+n_2=n$ cuando se dividen residuos sólo en dos grupos; $n_1+n_2<n$ cuando se dividen residuos en 3 ó 4 grupos y se toman los grupos de los extremos de acuerdo a niveles de x.*

Construcción de la prueba:

Consideraciones del test BF:

- *Test robusto a no normalidad*
- *Supone comportamiento monótono de σ_i^2 vs. x_i*
- *Requiere n grande para que sean despreciables correlaciones entre residuos*
- *Funciona como un test t para comparación de dos medias poblacionales.*
- *$n_1 + n_2 = n$ cuando se dividen residuos sólo en dos grupos; $n_1 + n_2 < n$ cuando se dividen residuos en 3 ó 4 grupos y se toman los grupos de los extremos de acuerdo a niveles de x.*

Construcción de la prueba:

1. Separar residuos de acuerdo a grupos 1 y 2: e_{ij} i-ésimo residuo en grupo j, $j = 1, 2$, $i = 1, \ldots, n_j$.
2. Calcular medianas en los grupos: $\tilde{e}_j = \text{mediana}\left\{e_{1j}, \ldots, e_{n_j j}\right\}$, $j = 1, 2$
3. Calcular en cada grupo desv. absolutas respecto a las medianas: $d_{ij} = |e_{ij} - \tilde{e}_j|$, $j = 1, 2$, $i = 1, \ldots, n_j$
4. Calcular en cada grupo la media de las desv. absolutas respecto a la mediana: $\bar{d}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} d_{ij}$, $j = 1, 2$

Consideraciones del test BF:

- *Test robusto a no normalidad*
- *Supone comportamiento monótono de σ_i^2 vs. x_i*
- *Requiere n grande para que sean despreciables correlaciones entre residuos*
- *Funciona como un test t para comparación de dos medias poblacionales.*
- *$n_1+n_2=n$ cuando se dividen residuos sólo en dos grupos; $n_1+n_2<n$ cuando se dividen residuos en 3 ó 4 grupos y se toman los grupos de los extremos de acuerdo a niveles de x.*

Construcción de la prueba:

1. Separar residuos de acuerdo a grupos 1 y 2: e_{ij} i-ésimo residuo en grupo j, $j=1,2$, $i=1,\ldots,n_j$.
2. Calcular medianas en los grupos: $\tilde{e}_j = \text{mediana}\left\{e_{1j},\ldots,e_{n_j j}\right\}$, $j=1,2$
3. Calcular en cada grupo desv. absolutas respecto a las medianas: $d_{ij}=|e_{ij}-\tilde{e}_j|$, $j=1,2$, $i=1,\ldots,n_j$
4. Calcular en cada grupo la media de las desv. absolutas respecto a la mediana: $\bar{d}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} d_{ij}$, $j=1,2$

Consideraciones del test BF:

- *Test robusto a no normalidad*
- *Supone comportamiento monótono de σ_i^2 vs. x_i*
- *Requiere n grande para que sean despreciables correlaciones entre residuos*
- *Funciona como un test t para comparación de dos medias poblacionales.*
- *$n_1 + n_2 = n$ cuando se dividen residuos sólo en dos grupos; $n_1 + n_2 < n$ cuando se dividen residuos en 3 ó 4 grupos y se toman los grupos de los extremos de acuerdo a niveles de x.*

Construcción de la prueba:

Consideraciones del test BF:

- *Test robusto a no normalidad*
- *Supone comportamiento monótono de σ_i^2 vs. x_i*
- *Requiere n grande para que sean despreciables correlaciones entre residuos*
- *Funciona como un test t para comparación de dos medias poblacionales.*
- *$n_1+n_2=n$ cuando se dividen residuos sólo en dos grupos; $n_1+n_2<n$ cuando se dividen residuos en 3 ó 4 grupos y se toman los grupos de los extremos de acuerdo a niveles de x.*

Construcción de la prueba:

1. Separar residuos de acuerdo a grupos 1 y 2: e_{ij} i-ésimo residuo en grupo j, $j=1,2$, $i=1,\ldots,n_j$.
2. Calcular medianas en los grupos $\tilde{e}_j = \text{mediana}\left\{e_{1j},\ldots,e_{n_j j}\right\}$, $j=1,2$
3. Calcular en cada grupo desv. absolutas respecto a las medianas: $d_{ij}=\left|e_{ij}-\tilde{e}_j\right|$, $j=1,2$, $i=1,\ldots,n_j$
4. Calcular en cada grupo la media de las desv. absolutas respecto a la mediana: $\bar{d}_j=\frac{1}{n_j}\sum_{i=1}^{n_j} d_{ij}$, $j=1,2$

Consideraciones del test BF:

- *Test robusto a no normalidad*
- *Supone comportamiento monótono de σ_i^2 vs. x_i*
- *Requiere n grande para que sean despreciables correlaciones entre residuos*
- *Funciona como un test t para comparación de dos medias poblacionales.*
- *$n_1+n_2=n$ cuando se dividen residuos sólo en dos grupos; $n_1+n_2<n$ cuando se dividen residuos en 3 ó 4 grupos y se toman los grupos de los extremos de acuerdo a niveles de x.*

Construcción de la prueba:

1. Separar residuos de acuerdo a grupos 1 y 2: e_{ij} i-ésimo residuo en grupo j, $j=1,2$, $i=1,\ldots,n_j$.
2. Calcular medianas en los grupos: $\tilde{e}_j = \text{mediana}\left\{e_{1j},\ldots,e_{n_jj}\right\}$, $j=1,2$
3. Calcular en cada grupo desv. absolutas respecto a las medianas: $d_{ij}=|e_{ij}-\tilde{e}_j|$, $j=1,2$, $i=1,\ldots,n_j$
4. Calcular en cada grupo la media de las desv. absolutas respecto a la mediana: $\bar{d}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} d_{ij}$, $j=1,2$

Consideraciones del test BF:

- *Test robusto a no normalidad*
- *Supone comportamiento monótono de σ_i^2 vs. x_i*
- *Requiere n grande para que sean despreciables correlaciones entre residuos*
- *Funciona como un test t para comparación de dos medias poblacionales.*
- *$n_1+n_2=n$ cuando se dividen residuos sólo en dos grupos; $n_1+n_2<n$ cuando se dividen residuos en 3 ó 4 grupos y se toman los grupos de los extremos de acuerdo a niveles de x.*

Construcción de la prueba:

1. Separar residuos de acuerdo a grupos 1 y 2: e_{ij} i-ésimo residuo en grupo j, $j=1,2$, $i=1,\ldots,n_j$.
2. Calcular medianas en los grupos: $\widetilde{e}_j = \text{mediana}\left\{e_{1j},\ldots,e_{n_j j}\right\}$, $j=1,2$
3. Calcular en cada grupo desv. absolutas respecto a las medianas: $d_{ij}=\left|e_{ij}-\widetilde{e}_j\right|$, $j=1,2$, $i=1,\ldots,n_j$
4. Calcular en cada grupo la media de las desv. absolutas respecto a la mediana: $\bar{d}_j=\frac{1}{n_j}\sum_{i=1}^{n_j} d_{ij}$, $j=1,2$

Consideraciones del test BF:

- *Test robusto a no normalidad*
- *Supone comportamiento monótono de σ_i^2 vs. x_i*
- *Requiere n grande para que sean despreciables correlaciones entre residuos*
- *Funciona como un test t para comparación de dos medias poblacionales.*
- *$n_1+n_2=n$ cuando se dividen residuos sólo en dos grupos; $n_1+n_2<n$ cuando se dividen residuos en 3 ó 4 grupos y se toman los grupos de los extremos de acuerdo a niveles de x.*

Construcción de la prueba:

1. Separar residuos de acuerdo a grupos 1 y 2: e_{ij} i-ésimo residuo en grupo j, $j=1,2$, $i=1,\ldots,n_j$.
2. Calcular medianas en los grupos: $\widetilde{e}_j=\text{mediana}\left\{e_{1j},\ldots,e_{n_jj}\right\}$, $j=1,2$
3. Calcular en cada grupo desv. absolutas respecto a las medianas: $d_{ij}=\left|e_{ij}-\widetilde{e}_j\right|$, $j=1,2$, $i=1,\ldots,n_j$
4. Calcular en cada grupo la media de las desv. absolutas respecto a la mediana: $\bar{d}_j=\frac{1}{n_j}\sum_{i=1}^{n_j}d_{ij}$, $j=1,2$

Consideraciones del test BF:

- *Test robusto a no normalidad*
- *Supone comportamiento monótono de σ_i^2 vs. x_i*
- *Requiere n grande para que sean despreciables correlaciones entre residuos*
- *Funciona como un test t para comparación de dos medias poblacionales.*
- *$n_1 + n_2 = n$ cuando se dividen residuos sólo en dos grupos; $n_1 + n_2 < n$ cuando se dividen residuos en 3 ó 4 grupos y se toman los grupos de los extremos de acuerdo a niveles de x.*

Construcción de la prueba:

1. Separar residuos de acuerdo a grupos 1 y 2: e_{ij} i-ésimo residuo en grupo j, $j = 1, 2$, $i = 1, \ldots, n_j$.
2. Calcular medianas en los grupos: $\widetilde{e}_j = \text{mediana}\left\{e_{1j}, \ldots, e_{n_j j}\right\}$, $j = 1, 2$
3. Calcular en cada grupo desv. absolutas respecto a las medianas: $d_{ij} = \left|e_{ij} - \widetilde{e}_j\right|$, $j = 1, 2$, $i = 1, \ldots, n_j$
4. Calcular en cada grupo la media de las desv. absolutas respecto a la mediana: $\bar{d}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} d_{ij}$, $j = 1, 2$

Consideraciones del test BF:

- *Test robusto a no normalidad*
- *Supone comportamiento monótono de σ_i^2 vs. x_i*
- *Requiere n grande para que sean despreciables correlaciones entre residuos*
- *Funciona como un test t para comparación de dos medias poblacionales.*
- *$n_1+n_2=n$ cuando se dividen residuos sólo en dos grupos; $n_1+n_2<n$ cuando se dividen residuos en 3 ó 4 grupos y se toman los grupos de los extremos de acuerdo a niveles de x.*

Construcción de la prueba:

1. Separar residuos de acuerdo a grupos 1 y 2: e_{ij} i-ésimo residuo en grupo j, $j=1,2$, $i=1,\ldots,n_j$.
2. Calcular medianas en los grupos: $\widetilde{e}_j = \text{mediana}\left\{e_{1j},\ldots,e_{n_jj}\right\}$, $j=1,2$
3. Calcular en cada grupo desv. absolutas respecto a las medianas: $d_{ij} = \left|e_{ij}-\widetilde{e}_j\right|$, $j=1,2$, $i=1,\ldots,n_j$
4. Calcular en cada grupo la media de las desv. absolutas respecto a la mediana: $\bar{d}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} d_{ij}$, $j=1,2$

Tabla 3: Test BF: Prueba indirectamente (6), realizando el siguiente test:

Test	Estadístico de prueba	Criterio de rechazo
$H_0: \mu_{d_1} = \mu_{d_2}$ $H_1: \mu_{d_1} \neq \mu_{d_2}$	$t_L^* = \dfrac{\bar{d}_1 - \bar{d}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \overset{a}{\sim} t_{n_1+n_2-2}$, con, $S_p^2 = \dfrac{\sum_{i=1}^{n_1}\left(d_{i1}-\bar{d}_1\right)^2 + \sum_{i=1}^{n_2}\left(d_{i2}-\bar{d}_2\right)^2}{n-2}$	si $P\left(\left\lvert t_{n_1+n_2-2}\right\rvert > \left\lvert t_L^*\right\rvert\right)$ es pequeño, o a un nivel α si : $\left\lvert t_L^*\right\rvert > t_{\alpha/2, n_1+n_2-2}$
donde: $\mu_{d_1} = \mathrm{E}\left[d_{i1}\right], \quad \mu_{d_2} = \mathrm{E}\left[d_{i2}\right]$		

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
  - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)^2$

  Note que si $\gamma = 0$, se tendría que $\sigma_i^2 = \sigma^2$.

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

1. Se ajustan los $\widehat{E}_i^2$ vs. x_i.
2. Se prueba la significancia del parámetro γ. El estadístico de la prueba se construye con base en SSR*: la suma de cuadrados de la regresión hecha en el paso 1, y el SSE de la regresión de y_i vs. x_i de la cual se obtuvieron inicialmente los residuos $\widehat{E}_i$.

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
  - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)^2$

  Note que si $\gamma = 0$, se tendría que $\sigma_i^2 = \sigma^2$.

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

1. Se ajustan los $\widehat{E}_i^2$ vs. x_i

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
  - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

1. Se ajustan los E_i^2 vs. x_i

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
    - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$
    - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)$
    - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)^2$

  Note que si $\gamma = 0$, se tendría que $\sigma_i^2 = \sigma^2$.

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
  - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)^2$

Note que si $\gamma = 0$, se tendría que $\sigma_i^2 = \sigma^2$.

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
  - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)^2$

  Note que si $\gamma = 0$, se tendría que $\sigma_i^2 = \sigma^2$.

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

1. Se ajustan los $\hat{E}_i^2$ vs. x_i.

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
  - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)^2$

Note que si $\gamma = 0$, se tendría que $\sigma_i^2 = \sigma^2$.

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
    - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$
    - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)$
    - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)^2$

Note que si $\gamma = 0$, se tendría que $\sigma_i^2 = \sigma^2$.

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

1. Se ajustan los $\widehat{E}_i^2$ vs. x_i.

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
  - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)^2$

  Note que si $\gamma = 0$, se tendría que $\sigma_i^2 = \sigma^2$.

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

1. Se ajustan los $\widehat{E}_i^2$ vs. x_i.
2. Se prueba la significancia del parámetro γ. El estadístico de la prueba se construye con base en SSR*: la suma de cuadrados de la regresión hecha en el paso 1, y el SSE de la regresón de y_i vs. x_i de la cual se obtuvieron inicialmente los residuos $\widehat{E}_i$.

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
  - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)^2$

  Note que si $\gamma = 0$, se tendría que $\sigma_i^2 = \sigma^2$.

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

1. Se ajustan los $\widehat{E}_i^2$ vs. x_i.
2. Se prueba la significancia del parámetro γ. El estadístico de la prueba se construye con base en SSR*: la suma de cuadrados de la regresión hecha en el paso 1, y el SSE de la regresón de y_i vs. x_i de la cual se obtuvieron inicialmente los residuos $\widehat{E}_i$.

Test Breusch-Pagan (BP)

Consideraciones del test BP:

- *Requiere validez de incorrelación y normalidad en los errores*
- *n debe ser grande*
- *Supone que $\sigma_i^2 = \sigma^2 h(x_i)$ con $h(\cdot)$ t.q $\sigma_i^2 > 0$.*
- *algunos modelos para la varianza son:*
  - $\sigma_i^2 = \sigma^2 \exp(\gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)$
  - $\sigma_i^2 = \sigma^2 (1 + \gamma x_i)^2$

  Note que si $\gamma = 0$, se tendría que $\sigma_i^2 = \sigma^2$.

Construcción de la prueba: Teniendo en cuenta que $\sigma_i^2 = \mathrm{E}\left[E_i^2\right]$,

1. Se ajustan los $\widehat{E}_i^2$ vs. x_i.
2. Se prueba la significancia del parámetro γ. El estadístico de la prueba se construye con base en SSR*: la suma de cuadrados de la regresión hecha en el paso 1, y el SSE de la regresón de y_i vs. x_i de la cual se obtuvieron inicialmente los residuos $\widehat{E}_i$.

Tabla 4: Test BP en el MRLS: Prueba indirectamente (6), realizando el siguiente test:

Test	Estadístico de prueba	Criterio de rechazo
$H_0 : \gamma = 0$ $H_1 : \gamma \neq 0$	$\chi^2_{BP} = \frac{1}{2}\frac{\text{SSR}^*}{(\text{SSE}/n)^2} \overset{\text{aprox}}{\sim} \chi^2_1$	si $P\left(\chi^2_1 > \chi^2_{BP}\right)$ es pequeño, o a un nivel α si : $\chi^2_{BP} > \chi^2_{\alpha,1}$
	Test BP estudentizado: $\chi^{*2}_{BP} = nR^{*2} \overset{\text{aprox}}{\sim} \chi^2_1$	si $P\left(\chi^2_1 > \chi^{*2}_{BP}\right)$ es pequeño, o a un nivel α si : $\chi^{*2}_{BP} > \chi^2_{\alpha,1}$
R^{*2}: Coeficiente de determinación muestral de la regresión lineal de $\widehat{E}^2_i$ vs. x_i SSR*: Suma de cuadrados de la regresión hecha en el paso 1		

Nota 1.3

Ver Código R 3.2 en el Capítulo 3 de notas de clase, para la implementación del Test Breusch-Pagan mediante función `ncvTest( )` de la librería `car` y con la función `bptest( )` de la librería `lmtest`.

Tabla 4: Test BP en el MRLS: Prueba indirectamente (6), realizando el siguiente test:

Test	Estadístico de prueba	Criterio de rechazo
$H_0: \gamma = 0$ $H_1: \gamma \neq 0$	$\chi^2_{BP} = \frac{1}{2}\frac{\text{SSR}^*}{(\text{SSE}/n)^2} \overset{\text{aprox}}{\sim} \chi^2_1$	si $P\left(\chi^2_1 > \chi^2_{BP}\right)$ es pequeño, o a un nivel α si : $\chi^2_{BP} > \chi^2_{\alpha,1}$
	Test BP estudentizado: $\chi^{*2}_{BP} = nR^{*2} \overset{\text{aprox}}{\sim} \chi^2_1$	si $P\left(\chi^2_1 > \chi^{*2}_{BP}\right)$ es pequeño, o a un nivel α si : $\chi^{*2}_{BP} > \chi^2_{\alpha,1}$
R^{*2}: Coeficiente de determinación muestral de la regresión lineal de $\widehat{E}^2_i$ vs. x_i SSR*: Suma de cuadrados de la regresión hecha en el paso 1		

Nota 1.3

Ver Código R 3.2 en el Capítulo 3 de notas de clase, para la implementación del Test Breusch-Pagan mediante función `ncvTest( )` *de la librería* `car` *y con la función* `bptest( )` *de la librería* `lmtest`.

Qué hacer si hay heterocedasticidad?

- Ajustar por mínimos cuadrados ponderados: Se buscan β_0, β_1 que minimicen a

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} \omega_i E_i^2 = \sum_{i=1}^{n} \omega_i (Y_i - \beta_0 - \beta_1 x_i)^2. \tag{7}$$

  con w_i inversamente proporcional a $\mathrm{Var}[Y_i]$.
- *Usar transformaciones sobre Y que estabilicen la varianza: Con respuesta continua existen las transformaciones de potencia o Box-Cox.*
- Modelos heterocedásticos (pero ya no son modelos lineales).

Qué hacer si hay heterocedasticidad?

- Ajustar por mínimos cuadrados ponderados: Se buscan β_0, β_1 que minimicen a

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} \omega_i E_i^2 = \sum_{i=1}^{n} \omega_i (Y_i - \beta_0 - \beta_1 x_i)^2 . \qquad (7)$$

con w_i inversamente proporcional a $\text{Var}[Y_i]$.

- *Usar transformaciones sobre Y que estabilicen la varianza: Con respuesta continua existen las transformaciones de potencia o Box-Cox.*
- Modelos heterocedásticos (pero ya no son modelos lineales).

Qué hacer si hay heterocedasticidad?

- Ajustar por mínimos cuadrados ponderados: Se buscan β_0, β_1 que minimicen a

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} \omega_i E_i^2 = \sum_{i=1}^{n} \omega_i (Y_i - \beta_0 - \beta_1 x_i)^2. \quad (7)$$

  con w_i inversamente proporcional a $\mathrm{Var}[Y_i]$.
- *Usar transformaciones sobre Y que estabilicen la varianza: Con respuesta continua existen las transformaciones de potencia o Box-Cox.*
- Modelos heterocedásticos (pero ya no son modelos lineales).

Evaluación de la independencia

Nota 1.4

Recuerde que

- *La correlación sólo es una medida de dependencia lineal entre dos variables;*
- *Por lo anterior, aunque dos variables aleatorias sean incorrelacionadas, podrían no ser independientes;*
- *Sin embargo, si dos variables aleatorias son independientes, entonces son incorrelacionadas.*
- *Se procederá a evaluar si hay evidencia de correlación entre los errores del MRLS, y de hallar tal evidencia, entonces se rechazaría la independencia, en caso contrario, se dice que no se ha hallado evidencia en contra de la independencia.*

Tests: Requieren conocer orden de observación de los datos y n grande,

- Tests de la función de autocorrelación (ACF)
- Tests Ljung-Box, Box-Pierce
- Tests de la función de autocorrelación parcial (PACF)
- *Test Durbin-Watson para autocorrelación de orden 1*

Evaluación de la independencia

Nota 1.4

Recuerde que

- *La correlación sólo es una medida de dependencia lineal entre dos variables;*
- *Por lo anterior, aunque dos variables aleatorias sean incorrelacionadas, podrían no ser independientes;*
- *Sin embargo, si dos variables aleatorias son independientes, entonces son incorrelacionadas.*
- *Se procederá a evaluar si hay evidencia de correlación entre los errores del MRLS, y de hallar tal evidencia, entonces se rechazaría la independencia, en caso contrario, se dice que no se ha hallado evidencia en contra de la independencia.*

Tests: Requieren conocer orden de observación de los datos y n grande,

- Tests de la función de autocorrelación (ACF)
- Tests Ljung-Box, Box-Pierce
- Tests de la función de autocorrelación parcial (PACF)
- *Test Durbin-Watson para autocorrelación de orden 1*

Evaluación de la independencia

Nota 1.4

Recuerde que

- *La correlación sólo es una medida de dependencia lineal entre dos variables;*
- *Por lo anterior, aunque dos variables aleatorias sean incorrelacionadas, podrían no ser independientes;*
- *Sin embargo, si dos variables aleatorias son independientes, entonces son incorrelacionadas.*
- *Se procederá a evaluar si hay evidencia de correlación entre los errores del MRLS, y de hallar tal evidencia, entonces se rechazaría la independencia, en caso contrario, se dice que no se ha hallado evidencia en contra de la independencia.*

Tests: Requieren conocer orden de observación de los datos y n grande,

- Tests de la función de autocorrelación (ACF)
- Tests Ljung-Box, Box-Pierce
- Tests de la función de autocorrelación parcial (PACF)
- *Test Durbin-Watson para autocorrelación de orden 1*

Evaluación de la independencia

Nota 1.4

Recuerde que

- *La correlación sólo es una medida de dependencia lineal entre dos variables;*
- *Por lo anterior, aunque dos variables aleatorias sean incorrelacionadas, podrían no ser independientes;*
- *Sin embargo, si dos variables aleatorias son independientes, entonces son incorrelacionadas.*
- *Se procederá a evaluar si hay evidencia de correlación entre los errores del MRLS, y de hallar tal evidencia, entonces se rechazaría la independencia, en caso contrario, se dice que no se ha hallado evidencia en contra de la independencia.*

Tests: Requieren conocer orden de observación de los datos y n grande,

- Tests de la función de autocorrelación (ACF)
- Tests Ljung-Box, Box-Pierce
- Tests de la función de autocorrelación parcial (PACF)
- *Test Durbin-Watson para autocorrelación de orden 1*

Evaluación de la independencia

Nota 1.4

Recuerde que

- *La correlación sólo es una medida de dependencia lineal entre dos variables;*
- *Por lo anterior, aunque dos variables aleatorias sean incorrelacionadas, podrían no ser independientes;*
- *Sin embargo, si dos variables aleatorias son independientes, entonces son incorrelacionadas.*
- *Se procederá a evaluar si hay evidencia de correlación entre los errores del MRLS, y de hallar tal evidencia, entonces se rechazaría la independencia, en caso contrario, se dice que no se ha hallado evidencia en contra de la independencia.*

Tests: Requieren conocer orden de observación de los datos y n grande,

- Tests de la función de autocorrelación (ACF)
- Tests Ljung-Box, Box-Pierce
- Tests de la función de autocorrelación parcial (PACF)
- *Test Durbin-Watson para autocorrelación de orden 1*

Evaluación de la independencia

Nota 1.4

Recuerde que

- *La correlación sólo es una medida de dependencia lineal entre dos variables;*
- *Por lo anterior, aunque dos variables aleatorias sean incorrelacionadas, podrían no ser independientes;*
- *Sin embargo, si dos variables aleatorias son independientes, entonces son incorrelacionadas.*
- *Se procederá a evaluar si hay evidencia de correlación entre los errores del MRLS, y de hallar tal evidencia, entonces se rechazaría la independencia, en caso contrario, se dice que no se ha hallado evidencia en contra de la independencia.*

Tests: Requieren conocer orden de observación de los datos y n grande,

- Tests de la función de autocorrelación (ACF)
- Tests Ljung-Box, Box-Pierce
- Tests de la función de autocorrelación parcial (PACF)
- *Test Durbin-Watson para autocorrelación de orden 1*

Evaluación de la independencia

Nota 1.4

Recuerde que

- *La correlación sólo es una medida de dependencia lineal entre dos variables;*
- *Por lo anterior, aunque dos variables aleatorias sean incorrelacionadas, podrían no ser independientes;*
- *Sin embargo, si dos variables aleatorias son independientes, entonces son incorrelacionadas.*
- *Se procederá a evaluar si hay evidencia de correlación entre los errores del MRLS, y de hallar tal evidencia, entonces se rechazaría la independencia, en caso contrario, se dice que no se ha hallado evidencia en contra de la independencia.*

Tests: Requieren conocer orden de observación de los datos y n grande,

- Tests de la función de autocorrelación (ACF)
- Tests Ljung-Box, Box-Pierce
- Tests de la función de autocorrelación parcial (PACF)
- *Test Durbin-Watson para autocorrelación de orden 1*

Evaluación de la independencia

Nota 1.4

Recuerde que

- *La correlación sólo es una medida de dependencia lineal entre dos variables;*
- *Por lo anterior, aunque dos variables aleatorias sean incorrelacionadas, podrían no ser independientes;*
- *Sin embargo, si dos variables aleatorias son independientes, entonces son incorrelacionadas.*
- *Se procederá a evaluar si hay evidencia de correlación entre los errores del MRLS, y de hallar tal evidencia, entonces se rechazaría la independencia, en caso contrario, se dice que no se ha hallado evidencia en contra de la independencia.*

Tests: Requieren conocer orden de observación de los datos y n grande,

- Tests de la función de autocorrelación (ACF)
- Tests Ljung-Box, Box-Pierce
- Tests de la función de autocorrelación parcial (PACF)
- *Test Durbin-Watson para autocorrelación de orden 1*

Evaluación de la independencia

Nota 1.4

Recuerde que

- *La correlación sólo es una medida de dependencia lineal entre dos variables;*
- *Por lo anterior, aunque dos variables aleatorias sean incorrelacionadas, podrían no ser independientes;*
- *Sin embargo, si dos variables aleatorias son independientes, entonces son incorrelacionadas.*
- *Se procederá a evaluar si hay evidencia de correlación entre los errores del MRLS, y de hallar tal evidencia, entonces se rechazaría la independencia, en caso contrario, se dice que no se ha hallado evidencia en contra de la independencia.*

Tests: Requieren conocer orden de observación de los datos y n grande,

- Tests de la función de autocorrelación (ACF)
- Tests Ljung-Box, Box-Pierce
- Tests de la función de autocorrelación parcial (PACF)
- *Test Durbin-Watson para autocorrelación de orden 1*

Evaluación de la independencia

Nota 1.4

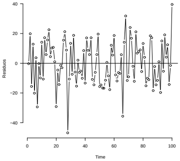
Recuerde que

- *La correlación sólo es una medida de dependencia lineal entre dos variables;*
- *Por lo anterior, aunque dos variables aleatorias sean incorrelacionadas, podrían no ser independientes;*
- *Sin embargo, si dos variables aleatorias son independientes, entonces son incorrelacionadas.*
- *Se procederá a evaluar si hay evidencia de correlación entre los errores del MRLS, y de hallar tal evidencia, entonces se rechazaría la independencia, en caso contrario, se dice que no se ha hallado evidencia en contra de la independencia.*

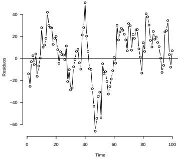
Tests: Requieren conocer orden de observación de los datos y n grande,

- Tests de la función de autocorrelación (ACF)
- Tests Ljung-Box, Box-Pierce
- Tests de la función de autocorrelación parcial (PACF)
- *Test Durbin-Watson para autocorrelación de orden 1*

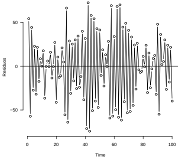
Algunos ejemplos de patrones en residuos vs. orden de observación



(a)



(b)



(c)

Figura 7: (a) Sin patrón claro; (b) con patrón cíclico; (c) con patrón de rachas de cambio sistemático en signos ±.

Test Durbin-Watson (DW) para autocorrelación de orden 1

Consideraciones del test DW:

- *Aplica sólo en modelos de regresión lineal*
- *Para los errores del modelo en el orden del tiempo, E_t, considera modelo autorregresivo de orden 1 (conocido en series de tiempo como AR(1)):*

$$E_t = \phi_1 E_{t-1} + a_t, \text{ con } a_t \overset{iid}{\sim} N\left(0, \sigma_a^2\right) \text{ y } \left|\phi_1\right| < 1. \tag{8}$$

- *Bajo este modelo se cumple que la autocorrelación de orden 1, es: $\rho(1) = Corr\left(E_t, E_{t-1}\right) = \phi_1$.*

Construcción de la prueba:

Test Durbin-Watson (DW) para autocorrelación de orden 1

Consideraciones del test DW:

- *Aplica sólo en modelos de regresión lineal*
- *Para los errores del modelo en el orden del tiempo, E_t, considera modelo autorregresivo de orden 1 (conocido en series de tiempo como AR(1)):*

$$E_t = \phi_1 E_{t-1} + a_t, \text{ con } a_t \stackrel{iid}{\sim} N\left(0, \sigma_a^2\right) \text{ y } \left|\phi_1\right| < 1. \tag{8}$$

- *Bajo este modelo se cumple que la autocorrelación de orden 1, es:* $\rho(1) = Corr\left(E_t, E_{t-1}\right) = \phi_1$.

Construcción de la prueba:

Test Durbin-Watson (DW) para autocorrelación de orden 1

Consideraciones del test DW:

- *Aplica sólo en modelos de regresión lineal*
- *Para los errores del modelo en el orden del tiempo, E_t, considera modelo autorregresivo de orden 1 (conocido en series de tiempo como AR(1)):*

$$E_t = \phi_1 E_{t-1} + a_t, \text{ con } a_t \stackrel{iid}{\sim} N\left(0, \sigma_a^2\right) \text{ y } \left|\phi_1\right| < 1. \tag{8}$$

- *Bajo este modelo se cumple que la autocorrelación de orden 1, es:* $\rho(1) = Corr\left(E_t, E_{t-1}\right) = \phi_1$.

Construcción de la prueba:

1. Se ajusta el modelo de regresión ordinario,
2. Con residuos de ajuste $\hat{E}_t$ en orden de t, se calcula estadístico de prueba d_1:

$$d_1 = \frac{\sum_{t=2}^{n}\left(\hat{E}_t - \hat{E}_{t-1}\right)^2}{\sum_{t=1}^{n}\hat{E}_t^2} \approx 2(1-\hat{\rho}(1)), \quad \text{donde} \quad \hat{\rho}(1) = \frac{\sum_{t=2}^{n}\hat{E}_t\hat{E}_{t-1}}{\sum_{t=1}^{n}\hat{E}_t^2}. \tag{9}$$

3. El juego de hipótesis a considerar se escoge de acuerdo a valores de d_1: si $0 < d_1 < 2$, se prueba autocorrelación positiva de orden 1; si $2 < d_1 < 4$, se prueba autocorrelación negativa de orden 1.

Test Durbin-Watson (DW) para autocorrelación de orden 1

Consideraciones del test DW:

- *Aplica sólo en modelos de regresión lineal*
- *Para los errores del modelo en el orden del tiempo, E_t, considera modelo autorregresivo de orden 1 (conocido en series de tiempo como AR(1)):*

$$E_t = \phi_1 E_{t-1} + a_t, \text{ con } a_t \overset{iid}{\sim} N\left(0, \sigma_a^2\right) \text{ y } \left|\phi_1\right| < 1. \tag{8}$$

- *Bajo este modelo se cumple que la autocorrelación de orden 1, es:* $\rho(1) = Corr\left(E_t, E_{t-1}\right) = \phi_1$.

Construcción de la prueba:

1. Se ajusta el modelo de regresión ordinario,
2. Con residuos de ajuste $\widehat{E}_t$ en orden de t, se calcula estadístico de prueba d_1:

$$d_1 = \frac{\sum_{t=2}^{n}\left(\widehat{E}_t - \widehat{E}_{t-1}\right)^2}{\sum_{t=1}^{n}\widehat{E}_t^2} \approx 2\left(1-\widehat{\rho}(1)\right), \quad \text{donde} \quad \widehat{\rho}(1) = \frac{\sum_{t=2}^{n}\widehat{E}_t\widehat{E}_{t-1}}{\sum_{t=1}^{n}\widehat{E}_t^2}, \tag{9}$$

Test Durbin-Watson (DW) para autocorrelación de orden 1

Consideraciones del test DW:

- *Aplica sólo en modelos de regresión lineal*
- *Para los errores del modelo en el orden del tiempo, E_t, considera modelo autorregresivo de orden 1 (conocido en series de tiempo como AR(1)):*

$$E_t = \phi_1 E_{t-1} + a_t, \text{ con } a_t \stackrel{iid}{\sim} N\left(0, \sigma_a^2\right) \text{ y } \left|\phi_1\right| < 1. \tag{8}$$

- *Bajo este modelo se cumple que la autocorrelación de orden 1, es:* $\rho(1) = Corr\left(E_t, E_{t-1}\right) = \phi_1$.

Construcción de la prueba:

1. Se ajusta el modelo de regresión ordinario,
2. Con residuos de ajuste $\widehat{E}_t$ en orden de t, se calcula estadístico de prueba d_1:

$$d_1 = \frac{\sum_{t=2}^{n}\left(\widehat{E}_t - \widehat{E}_{t-1}\right)^2}{\sum_{t=1}^{n}\widehat{E}_t^2} \approx 2\left(1 - \widehat{\rho}(1)\right), \quad \text{donde} \quad \widehat{\rho}(1) = \frac{\sum_{t=2}^{n}\widehat{E}_t\widehat{E}_{t-1}}{\sum_{t=1}^{n}\widehat{E}_t^2}, \tag{9}$$

3. El juego de hipótesis a considerar se escoge de acuerdo a valores de d_1: Si $0 < d_1 < 2$, se prueba autocorrelación positiva de orden 1; si $2 < d_1 < 4$, se prueba autocorrelación negativa de orden 1.

Test Durbin-Watson (DW) para autocorrelación de orden 1

Consideraciones del test DW:

- *Aplica sólo en modelos de regresión lineal*
- *Para los errores del modelo en el orden del tiempo, E_t, considera modelo autorregresivo de orden 1 (conocido en series de tiempo como AR(1)):*

$$E_t = \phi_1 E_{t-1} + a_t, \text{ con } a_t \overset{iid}{\sim} N\left(0, \sigma_a^2\right) \text{ y } \left|\phi_1\right| < 1. \tag{8}$$

- *Bajo este modelo se cumple que la autocorrelación de orden 1, es:* $\rho(1) = Corr\left(E_t, E_{t-1}\right) = \phi_1$.

Construcción de la prueba:

1. Se ajusta el modelo de regresión ordinario,
2. Con residuos de ajuste $\widehat{E}_t$ en orden de t, se calcula estadístico de prueba d_1:

$$d_1 = \frac{\sum_{t=2}^{n}\left(\widehat{E}_t - \widehat{E}_{t-1}\right)^2}{\sum_{t=1}^{n}\widehat{E}_t^2} \approx 2\left(1 - \widehat{\rho}(1)\right), \quad \text{donde} \quad \widehat{\rho}(1) = \frac{\sum_{t=2}^{n}\widehat{E}_t\widehat{E}_{t-1}}{\sum_{t=1}^{n}\widehat{E}_t^2}, \tag{9}$$

3. El juego de hipótesis a considerar se escoge de acuerdo a valores de d_1: Si $0 < d_1 < 2$, se prueba autocorrelación positiva de orden 1; si $2 < d_1 < 4$, se prueba autocorrelación negativa de orden 1.

Test Durbin-Watson (DW) para autocorrelación de orden 1

Consideraciones del test DW:

- *Aplica sólo en modelos de regresión lineal*
- *Para los errores del modelo en el orden del tiempo, E_t, considera modelo autorregresivo de orden 1 (conocido en series de tiempo como AR(1)):*

$$E_t = \phi_1 E_{t-1} + a_t, \text{ con } a_t \overset{iid}{\sim} N\left(0, \sigma_a^2\right) \text{ y } \left|\phi_1\right| < 1. \tag{8}$$

- *Bajo este modelo se cumple que la autocorrelación de orden 1, es:* $\rho(1) = Corr\left(E_t, E_{t-1}\right) = \phi_1$.

Construcción de la prueba:

1. Se ajusta el modelo de regresión ordinario,
2. Con residuos de ajuste $\widehat{E}_t$ en orden de t, se calcula estadístico de prueba d_1:

$$d_1 = \frac{\sum_{t=2}^{n}\left(\widehat{E}_t - \widehat{E}_{t-1}\right)^2}{\sum_{t=1}^{n}\widehat{E}_t^2} \approx 2\left(1 - \widehat{\rho}(1)\right), \quad \text{donde} \quad \widehat{\rho}(1) = \frac{\sum_{t=2}^{n}\widehat{E}_t\widehat{E}_{t-1}}{\sum_{t=1}^{n}\widehat{E}_t^2}, \tag{9}$$

3. El juego de hipótesis a considerar se escoge de acuerdo a valores de d_1: Si $0 < d_1 < 2$, se prueba autocorrelación positiva de orden 1; si $2 < d_1 < 4$, se prueba autocorrelación negativa de orden 1.

Tabla 5: Tests de hipótesis Durbin Watson para autocorrelación de orden 1. De acuerdo a d_1 se debe elegir entre uno de los dos juegos de hipótesis que se describen en esta tabla.

Test de auto correlación de orden 1 positiva: Si $0 < d_1 < 2$	
Juego de hipótesis	Criterio de rechazo
$H_0 : \rho(1) = 0$, o equivalentemente $H_0 : \phi_1 = 0$ vs. $H_1 : \rho(1) > 0$, o equivalentemente $H_1 : \phi_1 > 0$	Si $P(DW_1 < d_1)$ es pequeño
Conclusiones posibles: Si no se rechaza H_0, se concluye que no se ha encontrado evidencia de auto correlación de orden 1; si se rechaza H_0 se concluye que los errores consecutivos (o sea que distan una unidad de tiempo) están positivamente correlacionados.	

Test de auto correlación de orden 1 negativa: Si $2 < d_1 < 4$	
Juego de hipótesis	Criterio de rechazo
$H_0 : \rho(1) = 0$, o equivalentemente $H_0 : \phi_1 = 0$ vs. $H_1 : \rho(1) < 0$, o equivalentemente $H_1 : \phi_1 < 0$	Si $P(DW_1 > d_1)$ es pequeño
Conclusiones posibles: Si no se rechaza H_0, se concluye que no se ha encontrado evidencia de auto correlación de orden 1; si se rechaza H_0 se concluye que los errores consecutivos (o sea que distan una unidad de tiempo) están negativamente correlacionados.	

Nota 1.5

Ver Código R 3.3 en el Capítulo 3 de notas de clase, para la ejecución del test Durbin-Watson con la función `durbinWatsonTest( )` *de la librería* `car`.

Tabla 5: Tests de hipótesis Durbin Watson para autocorrelación de orden 1. De acuerdo a d_1 se debe elegir entre uno de los dos juegos de hipótesis que se describen en esta tabla.

Test de auto correlación de orden 1 positiva: Si $0 < d_1 < 2$	
Juego de hipótesis	Criterio de rechazo
$H_0 : \rho(1) = 0$, o equivalentemente $H_0 : \phi_1 = 0$ vs. $H_1 : \rho(1) > 0$, o equivalentemente $H_1 : \phi_1 > 0$	Si $P(DW_1 < d_1)$ es pequeño
Conclusiones posibles: Si no se rechaza H_0, se concluye que no se ha encontrado evidencia de auto correlación de orden 1; si se rechaza H_0 se concluye que los errores consecutivos (o sea que distan una unidad de tiempo) están positivamente correlacionados.	

Test de auto correlación de orden 1 negativa: Si $2 < d_1 < 4$	
Juego de hipótesis	Criterio de rechazo
$H_0 : \rho(1) = 0$, o equivalentemente $H_0 : \phi_1 = 0$ vs. $H_1 : \rho(1) < 0$, o equivalentemente $H_1 : \phi_1 < 0$	Si $P(DW_1 > d_1)$ es pequeño
Conclusiones posibles: Si no se rechaza H_0, se concluye que no se ha encontrado evidencia de auto correlación de orden 1; si se rechaza H_0 se concluye que los errores consecutivos (o sea que distan una unidad de tiempo) están negativamente correlacionados.	

Nota 1.5

Ver Código R 3.3 en el Capítulo 3 de notas de clase, para la ejecución del test Durbin-Watson con la función `durbinWatsonTest( )` *de la librería* `car`.

Tabla 5: Tests de hipótesis Durbin Watson para autocorrelación de orden 1. De acuerdo a d_1 se debe elegir entre uno de los dos juegos de hipótesis que se describen en esta tabla.

Test de auto correlación de orden 1 positiva: Si $0 < d_1 < 2$	
Juego de hipótesis	Criterio de rechazo
$H_0 : \rho(1) = 0$, o equivalentemente $H_0 : \phi_1 = 0$ vs. $H_1 : \rho(1) > 0$, o equivalentemente $H_1 : \phi_1 > 0$	Si $P(DW_1 < d_1)$ es pequeño
Conclusiones posibles: Si no se rechaza H_0, se concluye que no se ha encontrado evidencia de auto correlación de orden 1; si se rechaza H_0 se concluye que los errores consecutivos (o sea que distan una unidad de tiempo) están positivamente correlacionados.	

Test de auto correlación de orden 1 negativa: Si $2 < d_1 < 4$	
Juego de hipótesis	Criterio de rechazo
$H_0 : \rho(1) = 0$, o equivalentemente $H_0 : \phi_1 = 0$ vs. $H_1 : \rho(1) < 0$, o equivalentemente $H_1 : \phi_1 < 0$	Si $P(DW_1 > d_1)$ es pequeño
Conclusiones posibles: Si no se rechaza H_0, se concluye que no se ha encontrado evidencia de auto correlación de orden 1; si se rechaza H_0 se concluye que los errores consecutivos (o sea que distan una unidad de tiempo) están negativamente correlacionados.	

Nota 1.5

Ver Código R 3.3 en el Capítulo 3 de notas de clase, para la ejecución del test Durbin-Watson con la función `durbinWatsonTest( )` *de la librería* `car`.

Qué hacer si no se cumple independencia?

- *Modelar la dependencia entre los errores: Modelos de regresión con errores correlacionados.*
- Incluir funciones del índice de tiempo t para modelar tendencias, patrones periodicos o estacionales.
- *Trabajar con primeras diferencias:*
  1. Calcular las primeras diferencias: $Y_t^{*} = Y_t - Y_{t-1}$, $X_t^{*} = X_t - X_{t-1}$
  2. Ajustar Y_t^{*} vs. X_t^{*} bajo el modelo (sin intercepto):

$$Y_t^{*} = \beta_1^{*} X_t^{*} + u_t, \quad u_t \overset{iid}{\sim} N(0, \sigma_u^2) \tag{10}$$

  3. Construir ec. ajustada para Y_t, $\widehat{Y}_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_t$, tomando

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1^{*} \bar{X} \tag{11}$$

$$\widehat{\beta}_1 = \widehat{\beta}_1^{*} \tag{12}$$

Nota 1.6

Detalles en Capítulo 12 de: Kutner, M. H., Nachtsheim, C. J., Neter, J. and Li. W. (2005). Applied Linear Statistical Models, 5th ed. McGraw-Hill Irwing, New York.

Qué hacer si no se cumple independencia?

- *Modelar la dependencia entre los errores: Modelos de regresión con errores correlacionados.*
- Incluir funciones del índice de tiempo t para modelar tendencias, patrones periodicos o estacionales.
- *Trabajar con primeras diferencias:*
  1. Calcular las primeras diferencias: $Y_t' = Y_t - Y_{t-1}$, $X_t' = X_t - X_{t-1}$

Nota 1.6

Detalles en Capítulo 12 de: Kutner, M. H., Nachtsheim, C. J., Neter, J. and Li. W. (2005). Applied Linear Statistical Models, 5th ed. McGraw-Hill Irwing, New York.

Qué hacer si no se cumple independencia?

- *Modelar la dependencia entre los errores: Modelos de regresión con errores correlacionados.*
- Incluir funciones del índice de tiempo t para modelar tendencias, patrones periodicos o estacionales.
- *Trabajar con primeras diferencias:*
  1. Calcular las primeras diferencias: $Y_t^* = Y_t - Y_{t-1}$, $X_t^* = X_t - X_{t-1}$
  2. Ajustar Y_t^* vs. X_t^* bajo el modelo (sin intercepto)

$$Y_t^* = \beta_1^* X_t^* + u_t, \quad u_t \stackrel{\text{iid}}{\sim} N\left(0, \sigma_u^2\right) \tag{10}$$

  3. Construir ec. ajustada para Y_t, $\widehat{Y_t} = \widehat{\beta_0} + \widehat{\beta_1} X_t$, tomando

$$\widehat{\beta_0} = \bar{Y} - \widehat{\beta_1^*} \bar{X} \tag{11}$$

$$\widehat{\beta_1} = \widehat{\beta_1^*} \tag{12}$$

Nota 1.6

Detalles en Capítulo 12 de: Kutner, M. H., Nachtsheim, C. J., Neter, J. and Li. W. (2005). Applied Linear Statistical Models, 5th ed. McGraw-Hill Irwing, New York.

Qué hacer si no se cumple independencia?

- *Modelar la dependencia entre los errores: Modelos de regresión con errores correlacionados.*
- Incluir funciones del índice de tiempo t para modelar tendencias, patrones periodicos o estacionales.
- *Trabajar con primeras diferencias:*
  1. Calcular las primeras diferencias: $Y_t^* = Y_t - Y_{t-1}$, $X_t^* = X_t - X_{t-1}$
  2. Ajustar Y_t^* vs. X_t^* bajo el modelo (sin intercepto)

$$Y_t^* = \beta_1^* X_t^* + u_t, \quad u_t \overset{\text{iid}}{\sim} N\left(0, \sigma_u^2\right) \tag{10}$$

  3. Construir ec. ajustada para Y_t, $\widehat{Y}_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_t$, tomando

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1^* \bar{X} \tag{11}$$

$$\widehat{\beta}_1 = \widehat{\beta}_1^* \tag{12}$$

Nota 1.6

Detalles en Capítulo 12 de: Kutner, M. H., Nachtsheim, C. J., Neter, J. and Li. W. (2005). Applied Linear Statistical Models, 5th ed. McGraw-Hill Irwing, New York.

Qué hacer si no se cumple independencia?

- *Modelar la dependencia entre los errores: Modelos de regresión con errores correlacionados.*
- Incluir funciones del índice de tiempo t para modelar tendencias, patrones periodicos o estacionales.
- *Trabajar con primeras diferencias:*
  1. Calcular las primeras diferencias: $Y_t^* = Y_t - Y_{t-1}$, $X_t^* = X_t - X_{t-1}$
  2. Ajustar Y_t^* vs. X_t^* bajo el modelo (sin intercepto)

$$Y_t^* = \beta_1^* X_t^* + u_t, \quad u_t \overset{\text{iid}}{\sim} N\left(0, \sigma_u^2\right) \tag{10}$$

  3. Construir ec. ajustada para Y_t, $\widehat{Y}_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_t$, tomando

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1^* \bar{X} \tag{11}$$

$$\widehat{\beta}_1 = \widehat{\beta}_1^* \tag{12}$$

Nota 1.6

Detalles en Capítulo 12 de: Kutner, M. H., Nachtsheim, C. J., Neter, J. and Li, W. (2005). Applied Linear Statistical Models, 5th ed. McGraw-Hill Irwing, New York.

Qué hacer si no se cumple independencia?

- *Modelar la dependencia entre los errores: Modelos de regresión con errores correlacionados.*
- Incluir funciones del índice de tiempo t para modelar tendencias, patrones periodicos o estacionales.
- *Trabajar con primeras diferencias:*
  1. Calcular las primeras diferencias: $Y_t^* = Y_t - Y_{t-1}$, $X_t^* = X_t - X_{t-1}$
  2. Ajustar Y_t^* vs. X_t^* bajo el modelo (sin intercepto)

$$Y_t^* = \beta_1^* X_t^* + u_t, \quad u_t \overset{\text{iid}}{\sim} N\left(0, \sigma_u^2\right) \tag{10}$$

  3. Construir ec. ajustada para Y_t, $\widehat{Y}_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_t$, tomando

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1^* \bar{X} \tag{11}$$

$$\widehat{\beta}_1 = \widehat{\beta}_1^* \tag{12}$$

Nota 1.6

Detalles en Capítulo 12 de: Kutner, M. H., Nachtsheim, C. J., Neter, J. and Li. W. (2005). Applied Linear Statistical Models, 5th ed. McGraw-Hill Irwing, New York.

Qué hacer si no se cumple independencia?

- *Modelar la dependencia entre los errores: Modelos de regresión con errores correlacionados.*
- Incluir funciones del índice de tiempo t para modelar tendencias, patrones periodicos o estacionales.
- *Trabajar con primeras diferencias:*
  1. Calcular las primeras diferencias: $Y_t^* = Y_t - Y_{t-1}$, $X_t^* = X_t - X_{t-1}$
  2. Ajustar Y_t^* vs. X_t^* bajo el modelo (sin intercepto)

$$Y_t^* = \beta_1^* X_t^* + u_t, \quad u_t \overset{\text{iid}}{\sim} N\left(0, \sigma_u^2\right) \tag{10}$$

  3. Construir ec. ajustada para Y_t, $\widehat{Y}_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_t$, tomando

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1^* \bar{X} \tag{11}$$

$$\widehat{\beta}_1 = \widehat{\beta}_1^* \tag{12}$$

Nota 1.6

Detalles en Capítulo 12 de: Kutner, M. H., Nachtsheim, C. J., Neter, J. and Li. W. (2005). Applied Linear Statistical Models, 5th ed. McGraw-Hill Irwing, New York.

Qué hacer si no se cumple independencia?

- *Modelar la dependencia entre los errores: Modelos de regresión con errores correlacionados.*
- Incluir funciones del índice de tiempo t para modelar tendencias, patrones periodicos o estacionales.
- *Trabajar con primeras diferencias:*
  1. Calcular las primeras diferencias: $Y_t^* = Y_t - Y_{t-1}$, $X_t^* = X_t - X_{t-1}$
  2. Ajustar Y_t^* vs. X_t^* bajo el modelo (sin intercepto)

$$Y_t^* = \beta_1^* X_t^* + u_t, \quad u_t \overset{\text{iid}}{\sim} N\left(0, \sigma_u^2\right) \tag{10}$$

  3. Construir ec. ajustada para Y_t, $\widehat{Y}_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_t$, tomando

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1^* \bar{X} \tag{11}$$

$$\widehat{\beta}_1 = \widehat{\beta}_1^* \tag{12}$$

Nota 1.6

Detalles en Capítulo 12 de: Kutner, M. H., Nachtsheim, C. J., Neter, J. and Li. W. (2005). Applied Linear Statistical Models, 5th ed. McGraw-Hill Irwing, New York.

Evaluación de normalidad

Consideraciones

- *Como cualquier test de bondad de ajuste, los tests de normalidad exigen que el conjunto de valores sobre los que se aplican provengan de una muestra aleatoria.*

Definición 1.1

Un conjunto de variables aleatorias, $W_1, W_2, \ldots, W_n$, constituyen una muestra aleatoria de tamaño n si y solo si

1. *Son mutuamente independientes*
2. *Son idénticamente distribuidas*

- La correlación entre variables en una muestra puede afectar significativamente el desempeño de los tests de bondad de ajuste.
- *El supuesto de independencia debería verificarse antes de la evaluación de normalidad.*
- La no normalidad frecuentemente va de la mano con la no homogeneidad de la varianza.
- *Aunque se aplican estos tests usando los valores de los residuos de ajuste, no debe perder de vista que el supuesto y las conclusiones deben formularse para los errores del modelo:*

$$H_0 : E_i \sim N\left(0, \sigma^2\right)$$
$$H_1 : E_i \nsim N\left(0, \sigma^2\right). \qquad (13)$$

Evaluación de normalidad

Consideraciones

- *Como cualquier test de bondad de ajuste, los tests de normalidad exigen que el conjunto de valores sobre los que se aplican provengan de una muestra aleatoria.*

Definición 1.1

Un conjunto de variables aleatorias, $W_1, W_2, \ldots, W_n$, constituyen una muestra aleatoria de tamaño n si y solo si

1. *Son mutuamente independientes*
2. *Son idénticamente distribuidas*

- La correlación entre variables en una muestra puede afectar significativamente el desempeño de los tests de bondad de ajuste.
- *El supuesto de independencia debería verificarse antes de la evaluación de normalidad.*
- La no normalidad frecuentemente va de la mano con la no homogeneidad de la varianza.
- *Aunque se aplican estos tests usando los valores de los residuos de ajuste, no debe perder de vista que el supuesto y las conclusiones deben formularse para los errores del modelo:*

$$H_0 : E_i \sim N\left(0, \sigma^2\right)$$

$$H_1 : E_i \nsim N\left(0, \sigma^2\right). \tag{13}$$

Evaluación de normalidad

Consideraciones

- *Como cualquier test de bondad de ajuste, los tests de normalidad exigen que el conjunto de valores sobre los que se aplican provengan de una muestra aleatoria.*

Definición 1.1

Un conjunto de variables aleatorias, $W_1, W_2, \ldots, W_n$, constituyen una muestra aleatoria de tamaño n si y solo si

1. *Son mutuamente independientes*
2. *Son idénticamente distribuidas*

- La correlación entre variables en una muestra puede afectar significativamente el desempeño de los tests de bondad de ajuste.
- *El supuesto de independencia debería verificarse antes de la evaluación de normalidad.*
- La no normalidad frecuentemente va de la mano con la no homogeneidad de la varianza.
- *Aunque se aplican estos tests usando los valores de los residuos de ajuste, no debe perder de vista que el supuesto y las conclusiones deben formularse para los errores del modelo:*

$$H_0 : E_i \sim N\left(0, \sigma^2\right)$$
$$H_1 : E_i \nsim N\left(0, \sigma^2\right). \tag{13}$$

Evaluación de normalidad

Consideraciones

- *Como cualquier test de bondad de ajuste, los tests de normalidad exigen que el conjunto de valores sobre los que se aplican provengan de una muestra aleatoria.*

Definición 1.1

Un conjunto de variables aleatorias, $W_1, W_2, \ldots, W_n$, constituyen una muestra aleatoria de tamaño n si y solo si

1. *Son mutuamente independientes*
2. *Son idénticamente distribuidas*

- La correlación entre variables en una muestra puede afectar significativamente el desempeño de los tests de bondad de ajuste.
- *El supuesto de independencia debería verificarse antes de la evaluación de normalidad.*
- La no normalidad frecuentemente va de la mano con la no homogeneidad de la varianza.
- *Aunque se aplican estos tests usando los valores de los residuos de ajuste, no debe perder de vista que el supuesto y las conclusiones deben formularse para los errores del modelo:*

$$H_0 : E_i \sim N\left(0, \sigma^2\right)$$
$$H_1 : E_i \nsim N\left(0, \sigma^2\right). \qquad (13)$$

Evaluación de normalidad

Consideraciones

- *Como cualquier test de bondad de ajuste, los tests de normalidad exigen que el conjunto de valores sobre los que se aplican provengan de una muestra aleatoria.*

Definición 1.1

Un conjunto de variables aleatorias, $W_1, W_2, \ldots, W_n$, constituyen una muestra aleatoria de tamaño n si y solo si

1. *Son mutuamente independientes*
2. *Son idénticamente distribuidas*

- La correlación entre variables en una muestra puede afectar significativamente el desempeño de los tests de bondad de ajuste.
- *El supuesto de independencia debería verificarse antes de la evaluación de normalidad.*
- La no normalidad frecuentemente va de la mano con la no homogeneidad de la varianza.
- *Aunque se aplican estos tests usando los valores de los residuos de ajuste, no debe perder de vista que el supuesto y las conclusiones deben formularse para los errores del modelo:*

$$H_0 : E_i \sim N\left(0, \sigma^2\right)$$

$$H_1 : E_i \nsim N\left(0, \sigma^2\right). \quad (13)$$

Evaluación de normalidad

Consideraciones

- *Como cualquier test de bondad de ajuste, los tests de normalidad exigen que el conjunto de valores sobre los que se aplican provengan de una muestra aleatoria.*

Definición 1.1

Un conjunto de variables aleatorias, $W_1, W_2, \ldots, W_n$, constituyen una muestra aleatoria de tamaño n si y solo si

1. *Son mutuamente independientes*
2. *Son idénticamente distribuidas*

- La correlación entre variables en una muestra puede afectar significativamente el desempeño de los tests de bondad de ajuste.
- *El supuesto de independencia debería verificarse antes de la evaluación de normalidad.*
- La no normalidad frecuentemente va de la mano con la no homogeneidad de la varianza.
- *Aunque se aplican estos tests usando los valores de los residuos de ajuste, no debe perder de vista que el supuesto y las conclusiones deben formularse para los errores del modelo:*

$$H_0 : E_i \sim N\left(0, \sigma^2\right)$$
$$H_1 : E_i \nsim N\left(0, \sigma^2\right). \quad (13)$$

Evaluación de normalidad

Consideraciones

- *Como cualquier test de bondad de ajuste, los tests de normalidad exigen que el conjunto de valores sobre los que se aplican provengan de una muestra aleatoria.*

Definición 1.1

Un conjunto de variables aleatorias, $W_1, W_2, \ldots, W_n$, constituyen una muestra aleatoria de tamaño n si y solo si

1. *Son mutuamente independientes*
2. *Son idénticamente distribuidas*

- La correlación entre variables en una muestra puede afectar significativamente el desempeño de los tests de bondad de ajuste.
- *El supuesto de independencia debería verificarse antes de la evaluación de normalidad.*
- La no normalidad frecuentemente va de la mano con la no homogeneidad de la varianza.
- *Aunque se aplican estos tests usando los valores de los residuos de ajuste, no debe perder de vista que el supuesto y las conclusiones deben formularse para los errores del modelo:*

$$H_0 : E_i \sim N\left(0, \sigma^2\right)$$
$$H_1 : E_i \nsim N\left(0, \sigma^2\right). \tag{13}$$

Evaluación de normalidad

Consideraciones

- *Como cualquier test de bondad de ajuste, los tests de normalidad exigen que el conjunto de valores sobre los que se aplican provengan de una muestra aleatoria.*

Definición 1.1

Un conjunto de variables aleatorias, $W_1, W_2, \ldots, W_n$, constituyen una muestra aleatoria de tamaño n si y solo si

1. *Son mutuamente independientes*
2. *Son idénticamente distribuidas*

- La correlación entre variables en una muestra puede afectar significativamente el desempeño de los tests de bondad de ajuste.
- *El supuesto de independencia debería verificarse antes de la evaluación de normalidad.*
- La no normalidad frecuentemente va de la mano con la no homogeneidad de la varianza.
- *Aunque se aplican estos tests usando los valores de los residuos de ajuste, no debe perder de vista que el supuesto y las conclusiones deben formularse para los errores del modelo:*

$$\begin{aligned} H_0 &: E_i \sim N\left(0, \sigma^2\right) \\ H_1 &: E_i \nsim N\left(0, \sigma^2\right). \end{aligned} \tag{13}$$

Tabla 6: Pruebas para normalidad

Pruebas especiales para normalidad			
Test	**Característica**	**Estadístico**	**Valor P**
Jarque-Bera:	Basada en asimetría y kurtosis	$\text{JB} = \frac{n}{6}\left[S^2 + \frac{(k-3)^2}{4}\right] \sim \chi^2_2$ S es el coef. de asimetría muestral, k es la kurtosis	$P\left(\chi^2_2 > \text{JB}\right)$
Shapiro-Wilk:	Compara estimador varianza basado en estadísticos de orden muestrales vs. estimador basado en Suma de cuadrados corregidos.	$W_0 = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$, con $0 < W_0 < 1$ y los a_i funciones de las medias, varianzas y covarianzas de los estadísticos de orden de v.a $N(0,1)$.	$P(W \leq W_0)$

Pruebas generales			
Test	**Característica**	**Estadístico**	**Valor P**
Kolmogorov-Smirnov:	Para cualquier distribución continua	$D_0 = \max_{1 \leq i \leq n} \left\vert F_n(x_i) - F_0(x_i)\right\vert$	$P(D \geq D_0)$
Anderson-Darling:	Mejora sensibilidad de K-S en las colas de la distribución, pero sólo está disponible para las distribuciones normal, lognormal, weibull, exponencial, de valor extremo y logística.	$A^2_n = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F_0(x))^2}{F_0(x)(1-F_0(x))} dF_0(x)$	$P\left(A^2 \geq A^2_0\right)$
Cramer-Von Mises:	Similar a K-S, pero más complejo computacionalmente.	$W^2_0 = n \int_{-\infty}^{\infty} [F_n(x) - F_0(x)]^2 \, dF_0(x)$	$P\left(W^2 \geq W^2_0\right)$

Nota 1.7 (Otros tests)

- *Test Lilliefors (Kolmogorov-Smirnov para normalidad):*

$$D = \max\left\{D^+, D^-\right\}$$

con

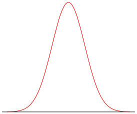
$$D^+ = \max_{i=1,\ldots,n}\left\{\frac{i}{n} - p_{(i)}\right\}, \quad D^- = \max_{i=1,\ldots,n}\left\{p_{(i)} - \frac{(i-1)}{n}\right\}$$

$$p_{(i)} = \Phi\left(\frac{x_{(i)} - \overline{x}}{s}\right)$$

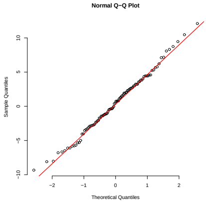
$\overline{x}$ *y* s *son la media y desviación estándar de los valores ingresados a la prueba y* $x_{(i)}$ *el iésimo valor de menor a mayor. Disponible en librería* `nortest` *bajo la función:*

`lillie.test( )`.

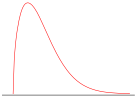
- *También en esta librería se hayan implementadas las pruebas Anderson-Darling y Shapiro-Francia, para normalidad, en las funciones* `ad.test( )` *y* `sf.test`*, respectivamente.*



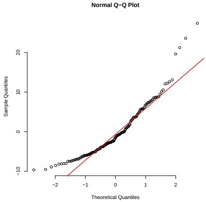
(a)



(b)

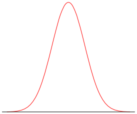


(c)

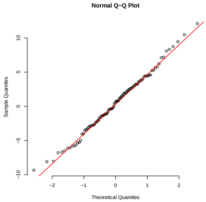


(d)

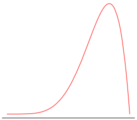
Figura 8: Densidades poblacionales y gráficos de probabilidad normal con muestras provenientes de: (a) (b) una distribución normal de media cero; (c) y (d) con una distribución no normal y asimétrica a derecha



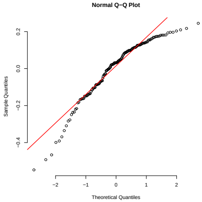
(a)



(b)

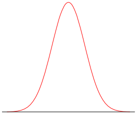


(c)

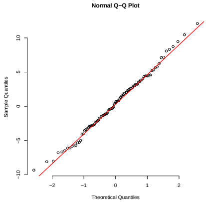


(d)

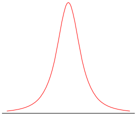
Figura 9: Densidades poblacionales y gráficos de probabilidad normal con muestras provenientes de: (a) (b) una distribución normal de media cero; (c) y (d) con una distribución no normal asimétrica a izquierda



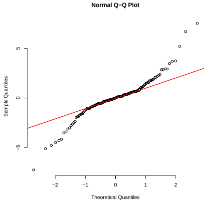
(a)



(b)



(c)



(d)

Figura 10: Densidades poblacionales y gráficos de probabilidad normal con muestras provenientes de: (a) (b) una distribución normal de media cero; (c) y (d) con una distribución no normal, simétrica pero de colas pesadas.

Algunas soluciones a no normalidad

La no normalidad puede estar vinculada a no homogeneidad de varianza. En algunos casos las transformaciones sobre Y que estabilizan varianza también logran una variables respuesta normal en escala transformada.

- *Transformaciones sobre la variable respuesta: las transformaciones de potencia Box-Cox:* Y^{λ} *(ver Sección 3.5, Capítulo 3 Notas de Clase).*
- *Regresión no paramétrica: Por ejemplo la regresión local polinomial o LOESS.*
- *Modelos lineales (o no lineales) generalizados: Donde considera para* $Y|x$ *distribuciones diferentes a la normal.*

Algunas soluciones a no normalidad

La no normalidad puede estar vinculada a no homogeneidad de varianza. En algunos casos las transformaciones sobre Y que estabilizan varianza también logran una variables respuesta normal en escala transformada.

- *Transformaciones sobre la variable respuesta: las transformaciones de potencia Box-Cox: Y^{λ} (ver Sección 3.5, Capítulo 3 Notas de Clase).*
- *Regresión no paramétrica: Por ejemplo la regresión local polinomial o LOESS.*
- *Modelos lineales (o no lineales) generalizados: Donde considera para $Y|x$ distribuciones diferentes a la normal.*

Algunas soluciones a no normalidad

La no normalidad puede estar vinculada a no homogeneidad de varianza. En algunos casos las transformaciones sobre Y que estabilizan varianza también logran una variables respuesta normal en escala transformada.

- *Transformaciones sobre la variable respuesta: las transformaciones de potencia Box-Cox:* Y^{λ} *(ver Sección 3.5, Capítulo 3 Notas de Clase).*
- *Regresión no paramétrica: Por ejemplo la regresión local polinomial o LOESS.*
- *Modelos lineales (o no lineales) generalizados: Donde considera para* $Y|x$ *distribuciones diferentes a la normal.*

Outliers u observaciones atípicas en la respuesta

Definición 1.2

Un outlier es aquella observación en donde la respuesta toma un valor bastante alejado o atípico con respecto al resto de valores.

Procedimientos gráficos para identificación

- Residuos ordinarios $\widehat{E}_i$
- Residuos estandarizados e_i

$$e_i = \frac{\widehat{E}_i}{\sqrt{\text{MSE}}} \tag{14}$$

Outliers u observaciones atípicas en la respuesta

Definición 1.2

Un outlier es aquella observación en donde la respuesta toma un valor bastante alejado o atípico con respecto al resto de valores.

Procedimientos gráficos para identificación

- Residuos ordinarios $\widehat{E}_i$
- Residuos estandarizados e_i

$$e_i = \frac{\widehat{E}_i}{\sqrt{\text{MSE}}} \tag{14}$$

- Residuos estudentizados (internamente estudentizados) r_i, en R con función `rstandard( )`.

$$r_i = \frac{\widehat{E}_i}{\sqrt{\widehat{\text{Var}}\left(\widehat{E}_i\right)}} = \frac{\widehat{E}_i}{\sqrt{(1-h_{ii})\text{MSE}}}, \text{ con } h_{ii} = \frac{1}{n} + \frac{(x_i-\bar{x})^2}{S_{xx}}. \tag{15}$$

- Residuos externamente estudentizados t_i, en R con función `rstudent( )`.

$$t_i = r_i\left[\frac{n-p-1}{n-p-r_i^2}\right]^{1/2} \sim t_{n-p-1}, \tag{16}$$

p número de parámetros en el modelo de regresión, siendo 2 para la RLS.

Outliers u observaciones atípicas en la respuesta

Definición 1.2

Un outlier es aquella observación en donde la respuesta toma un valor bastante alejado o atípico con respecto al resto de valores.

Procedimientos gráficos para identificación

- Residuos ordinarios $\widehat{E}_i$
- Residuos estandarizados e_i

$$e_i = \frac{\widehat{E}_i}{\sqrt{\text{MSE}}} \tag{14}$$

- Residuos estudentizados (internamente estudentizados) r_i, en R con función `rstandard( )`.

$$r_i = \frac{\widehat{E}_i}{\sqrt{\widehat{\text{Var}}\left(\widehat{E}_i\right)}} = \frac{\widehat{E}_i}{\sqrt{(1-h_{ii})\text{MSE}}}, \text{ con } h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}. \tag{15}$$

- Residuos externamente estudentizados t_i, en R con función `rstudent( )`.

$$t_i = r_i \left[\frac{n-p-1}{n-p-r_i^2}\right]^{1/2} \sim t_{n-p-1}, \tag{16}$$

p número de parámetros en el modelo de regresión, siendo 2 para la RLS.

Outliers u observaciones atípicas en la respuesta

Definición 1.2

Un outlier es aquella observación en donde la respuesta toma un valor bastante alejado o atípico con respecto al resto de valores.

Procedimientos gráficos para identificación

- Residuos ordinarios $\widehat{E}_i$
- Residuos estandarizados e_i

$$e_i = \frac{\widehat{E}_i}{\sqrt{\text{MSE}}} \tag{14}$$

- Residuos estudentizados (internamente estudentizados) r_i, en R con función `rstandard( )`.

$$r_i = \frac{\widehat{E}_i}{\sqrt{\widehat{\text{Var}}\left(\widehat{E}_i\right)}} = \frac{\widehat{E}_i}{\sqrt{(1-h_{ii})\text{MSE}}}, \text{ con } h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}. \tag{15}$$

- Residuos externamente estudentizados t_i, en R con función `rstudent( )`.

$$t_i = r_i\left[\frac{n-p-1}{n-p-r_i^2}\right]^{1/2} \sim t_{n-p-1}, \tag{16}$$

p número de parámetros en el modelo de regresión, siendo 2 para la RLS.

Outliers u observaciones atípicas en la respuesta

Definición 1.2

Un outlier es aquella observación en donde la respuesta toma un valor bastante alejado o atípico con respecto al resto de valores.

Procedimientos gráficos para identificación

- Residuos ordinarios $\widehat{E}_i$
- Residuos estandarizados e_i

$$e_i = \frac{\widehat{E}_i}{\sqrt{\text{MSE}}} \tag{14}$$

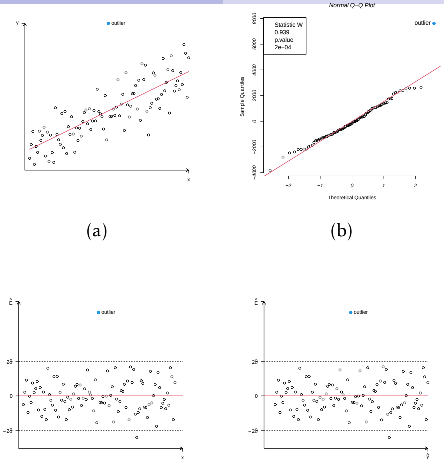
- Residuos estudentizados (internamente estudentizados) r_i, en R con función `rstandard( )`.

$$r_i = \frac{\widehat{E}_i}{\sqrt{\widehat{\text{Var}}\left(\widehat{E}_i\right)}} = \frac{\widehat{E}_i}{\sqrt{(1-h_{ii})\text{MSE}}}, \text{ con } h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}. \tag{15}$$

- Residuos externamente estudentizados t_i, en R con función `rstudent( )`.

$$t_i = r_i \left[\frac{n-p-1}{n-p-r_i^2}\right]^{1/2} \sim t_{n-p-1}, \tag{16}$$

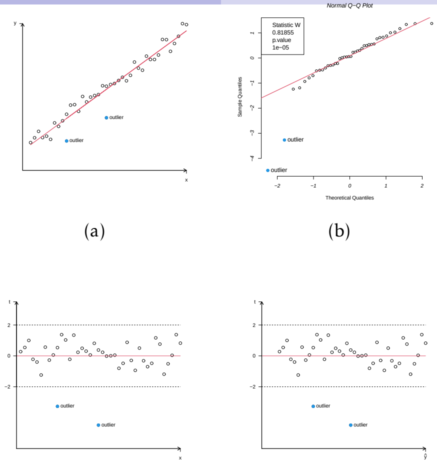
p número de parámetros en el modelo de regresión, siendo 2 para la RLS.



(a) (b)

(c) (d)

Figura 11: Observación outlier identificada como un valor atípico en la variable respuesta (a) Gráfico de dispersión y recta ajustada (b) gráfico de probabilidad normal con residuos de ajuste; (c) y (d) gráficos de residuos vs x y vs. $\widehat{y}$, respectivamente.



(a) (b) (c) (d)

Figura 12: Dos bservaciones outliers identificadas como valores atípicos en la variable respuesta (a) Gráfico de dispersión y recta ajustada (b) gráfico de probabilidad normal con residuos estudentizados externamente; (c) y (d) gráficos de residuos estudentizados externamente vs x y vs. $\widehat{y}$, respectivamente.