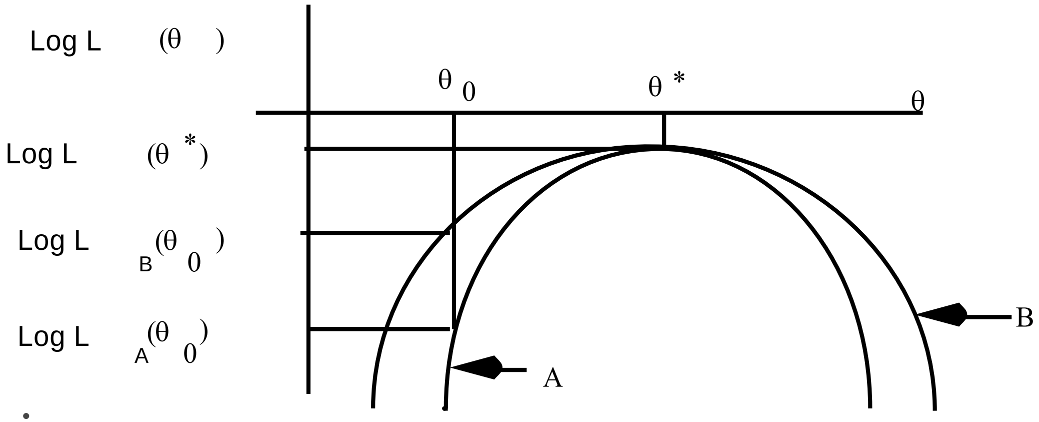
Diagnosis en el modelo de regresión lineal normal



Elementos de diagnosis, interpretación y análisis

Justificación del tema

- Hasta ahora hemos considerado que las principales suposiciones del modelo se verificaban, y en consecuencia, los estimadores MCO eran los más eficientes.
- Sin embargo, cuando trabajamos con datos reales, no siempre es esto cierto, puede ser que el modelo no sea válido y por tanto, debemos buscar alguna alternativa.
- Lo primero es detectar que ese modelo no es válido.
- Para ello necesitamos algún **instrumento** que nos permita ver que suposiciones son válidas y cuales no.

Fallos en el modelo

Universidade de Vigo

- Un modelo puede fallar por dos tipos de causas:
  - El modelo está mal especificado, es decir alguna suposición de partida no se verifica.
    - En ese caso los cálculos son válidos pero las propiedades teóricas cambian
  - Existen datos que provienen de otra población y contaminan el modelo.
    - Afecta a los cálculos y a las propiedades teóricas
- Cuando el modelo obtenido no se ve afectado por cambios en las suposiciones o los datos se dice que es **robusto**.
- Por tanto existen dos tipos de robustez y dependiendo del efecto la robustez puede ser mayor o menor a ese fallo
  - Respecto a los datos
  - Respecto a las suposiciones
- La diagnosis nos permite determinar el grado de robustez del modelo.

Necesidad de la diagnosis

Universidade de Vigo

- A continuación se van a exponer de modo gráfico cuatro modelos cuyos resultados de la regresión son casi equivalentes aparentemente.

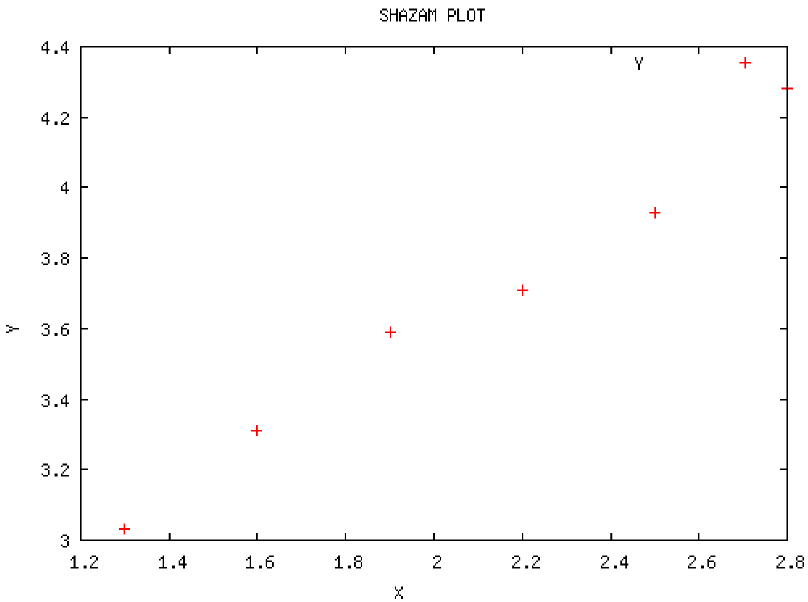
	Variable Name	Estimated Coefficient
Caso1	X	0.84
	CONSTANT	19.15
Caso2	X	0.80
	CONSTANT	19.87
Caso3	X	0.90
	CONSTANT	18.08
Caso4	X	0.80
	CONSTANT	20.00

- El coeficiente de determinación vale en todos los casos 0,985 y los coeficientes estimados son muy similares.

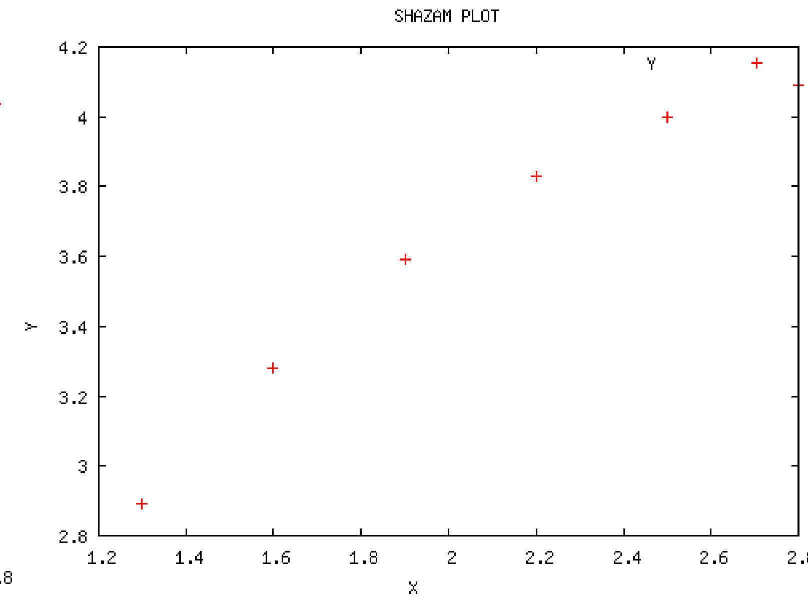
Ejemplo del efecto de la forma de la nube de puntos

Universidade de Vigo

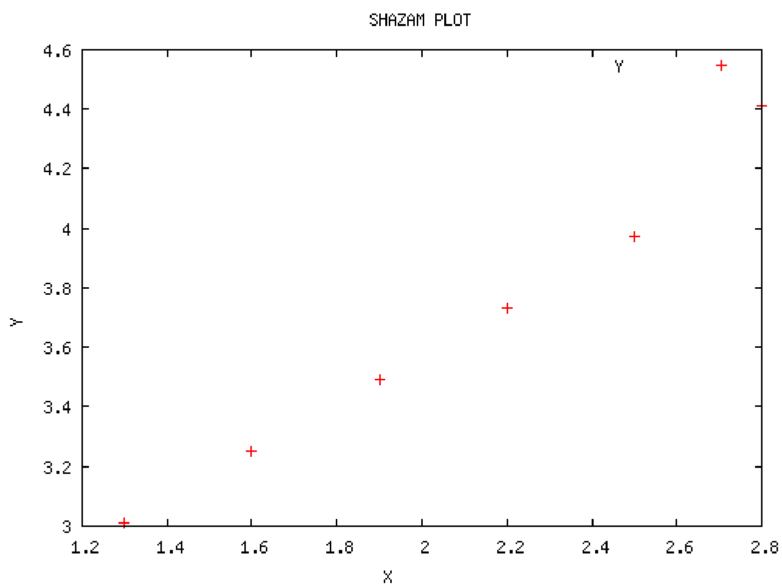
1



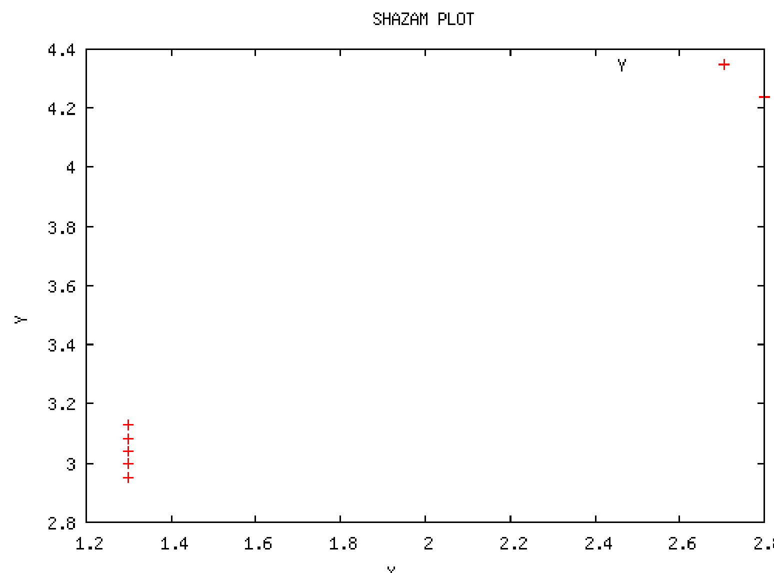
2



3

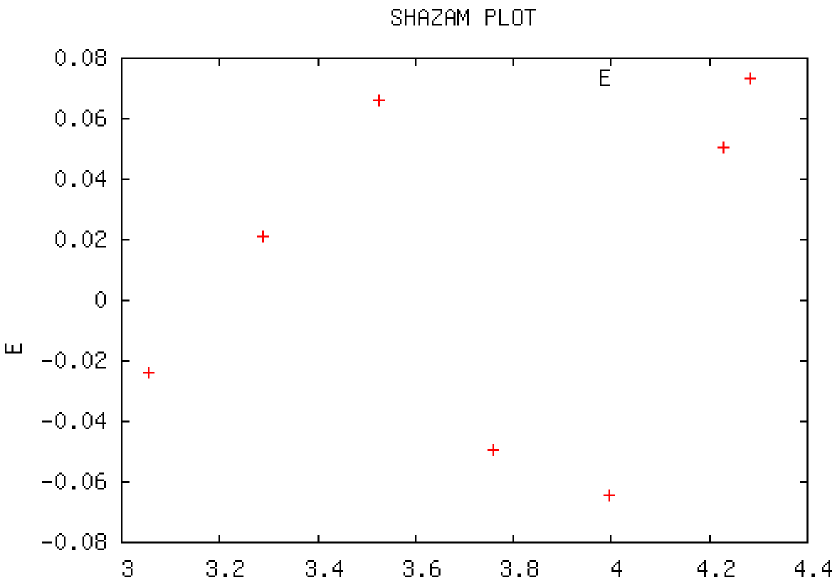


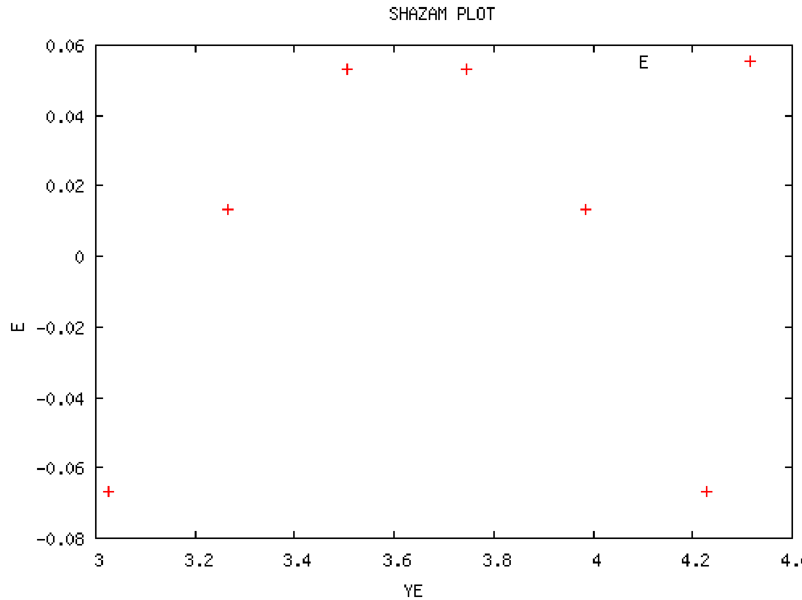
4

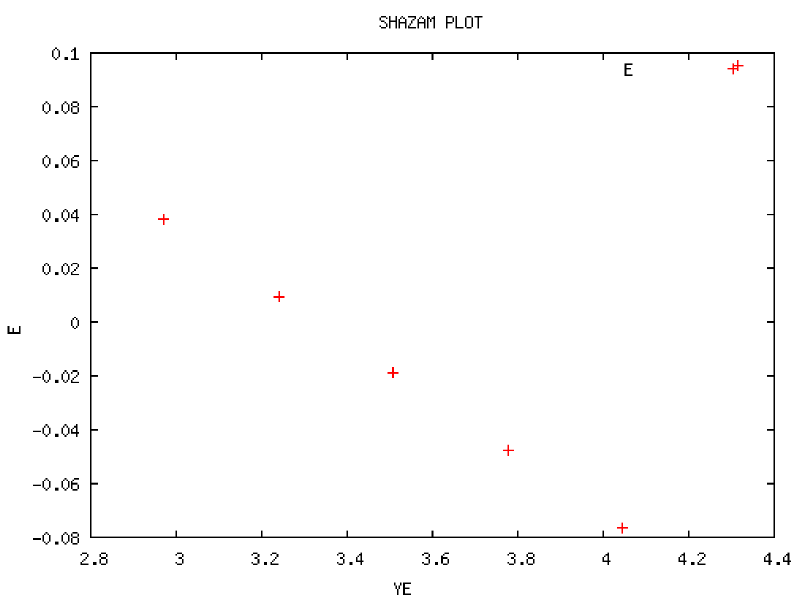


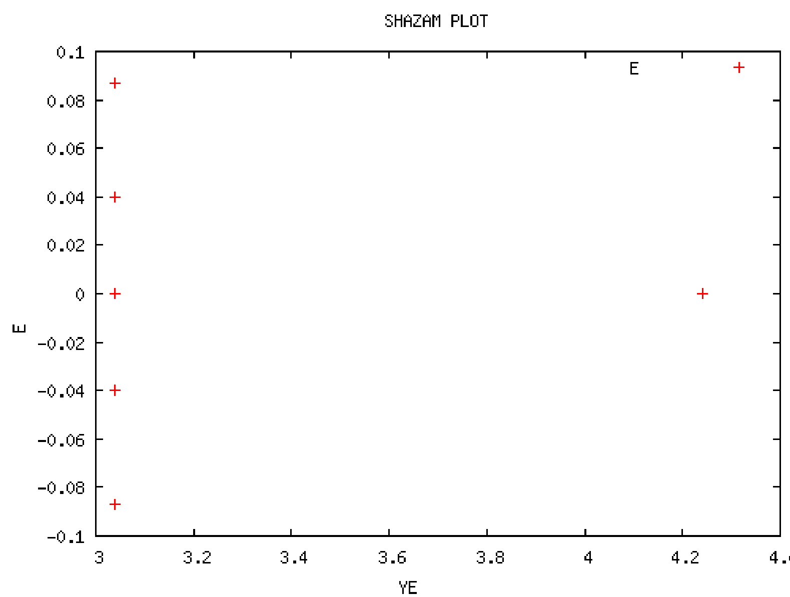
Residuos

iversidade
de Vigo









Elementos para analizar un modelo

Universidade de Vigo

- Eso nos indica que no solo se deben examinar la salida sino también diagnosticar el modelo.
- Eso significa que debemos determinar
  1. si las suposiciones son todas válidas o no
  2. Si el fallo es debido al conjunto de los datos o solo a algunos que tienen un comportamiento anómalo respecto al resto.
- Por consiguiente, los elementos que se deben tener en cuenta a la hora de la diagnosis son:

  1º El comportamiento de los datos.

  2º El efecto de ese comportamiento sobre las suposiciones.

Esquema de análisis de cada suposición

Universidade de Vigo

Interpretación

Significado teórico y las implicaciones de cada suposición

Efectos del fallo

Efecto que sobre los estimadores MCO produce su incumplimiento y grado de validez del estimador

Diagnostico

Se analizan los instrumentos para detectar si se verifica o no esa suposición y como se aplican a ese caso concreto.

Soluciones

Se sugiere algún método que permita obtener estimadores que mejoren los de MCO, en el nuevo modelo.

Proceso de diagnosis de un modelo

Universidade de Vigo

1. Se define un modelo teórico que analiza el problema enunciado.
2. Se estima.
3. Se analiza si existen datos extraños al modelo, suponiendo válidas todas las suposiciones. (si son muchos es que el modelo esta mal especificado).
4. En caso de datos extraños se eliminan esos (o se corrigen) y se repiten los pasos 1 a 3.
5. Si alguna suposición falla, se corrige y se vuelve a revisar el modelo corregido.
6. Ese proceso se repite hasta que haya una adecuación conveniente entre modelo y datos.

Elementos de Diagnosis

- **Medidas basadas en los residuos**
- **Gráficos**
- **Test de hipótesis**

Residuos de Mínimos Cuadrados Ordinarios

Universidade de Vigo

- Definición

$$e=Y-Xb$$

- Propiedades
  - $e=My$
  - Siguen leyes normales $N(0,\sigma^2 M)$
  - Son ortogonales a los regresores $Xe=0$
  - $(T-k-1)S^2/\sigma^2$ sigue una χ^2 con T-k-1 grados de libertad

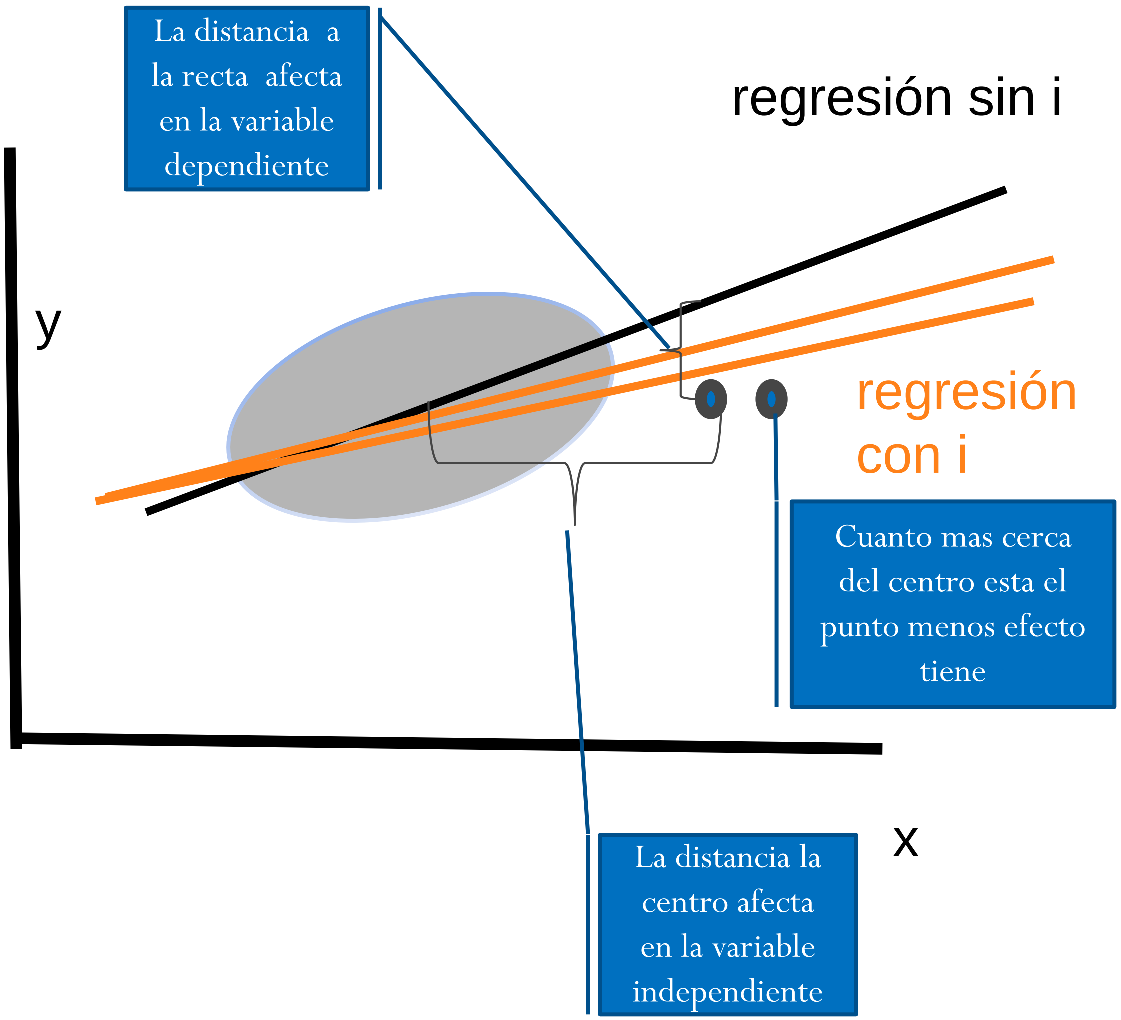
Efecto de una suposición sobre el modelo estimado

Universidade de Vigo

- Un modelo mal especificado no verifica alguna de las suposiciones de partida.
  - Por ejemplo si el modelo es cuadrático en vez de lineal, los residuos dependerán de la variable independiente al cuadrado, en vez de tener media 0.
  - Los residuos siguen leyes normales $N(\alpha x^2, \sigma^2 M)$
  - Por tanto su forma no será centrada en el 0

Efectos de una observación sobre el modelo estimado

Universidade de Vigo



- Todo valor tiene un efecto sobre la estimación del modelo:
- Unos afectan más y otros menos.
- Los que cambian mucho el modelo se denominan influyentes
- Esa influencia puede deberse a las variables independientes o a la dependiente
- Las diferentes medidas de influencia buscan también la causa

Valores especiales en la regresión

Universidade de Vigo

- Existen dos tipos de valores que pueden afectar al modelo:
  - Los **valores atípicos** que afectan al comportamiento de la variable dependiente sobre la estimación. Nos centraremos en estos. Para definirlos usaremos el concepto de residuos estandarizados
  - Los **valores influyentes** que afectan a la estimación del modelo, bien a los coeficientes, bien los valores estimados de la dependiente o bien a las varianzas.

Instrumento para medir valores atipicos

Universidade de Vigo

Residuos estandarizados

- Se definen como los residuos divididos, cada uno, por su desviación estándar estimada

$$s_t = \frac{e_t}{s_R \sqrt{1 - h_t}}$$

- siendo

$$s_R^2 = \frac{1}{T - k - 1} \sum_s e_s^2$$

- se les denomina internamente estudentizados porque incluyen el propio valor al hacer la estimación de la varianza residual.

Propiedades de los residuos estudentizados

Universidade de Vigo

- Miden el error de cada observación independientemente de las unidades de medida de las variables, por tanto permiten indicar si ese error es relativamente grande o pequeño.
- Los residuos estandarizados siguen aproximadamente una normal tipificada.
- Por consiguiente se puede decir que un residuo es muy grande si es mayor que el valor de las tablas de una normal estándar al nivel prefijado, aproximadamente 1,96 al 95% de confianza

Definición de valor atípico

Universidade de Vigo

- El concepto de valor atípico se inserta dentro del conjunto de datos con el que se trabaja:
  - Se define valor atípico individual al nivel α si verifica que

$$|s_t| > \lambda_{\frac{\alpha}{2}}$$

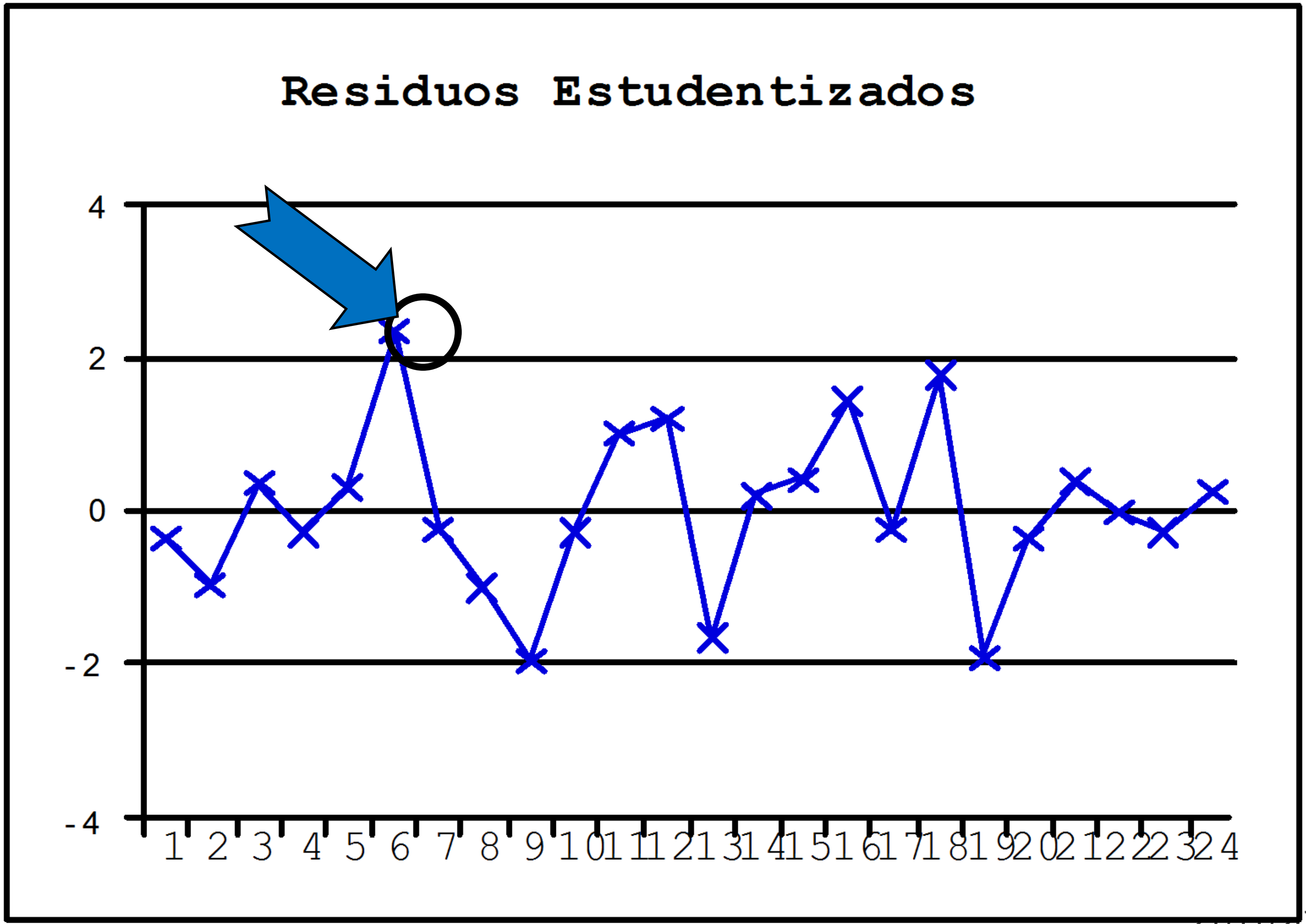
  - Se define valor atípico conjunto al nivel α si verifica que

$$|s_t| > \lambda_{\frac{\alpha}{2T}}$$

Principio de Bonferroni

Valor atípico al nivel α

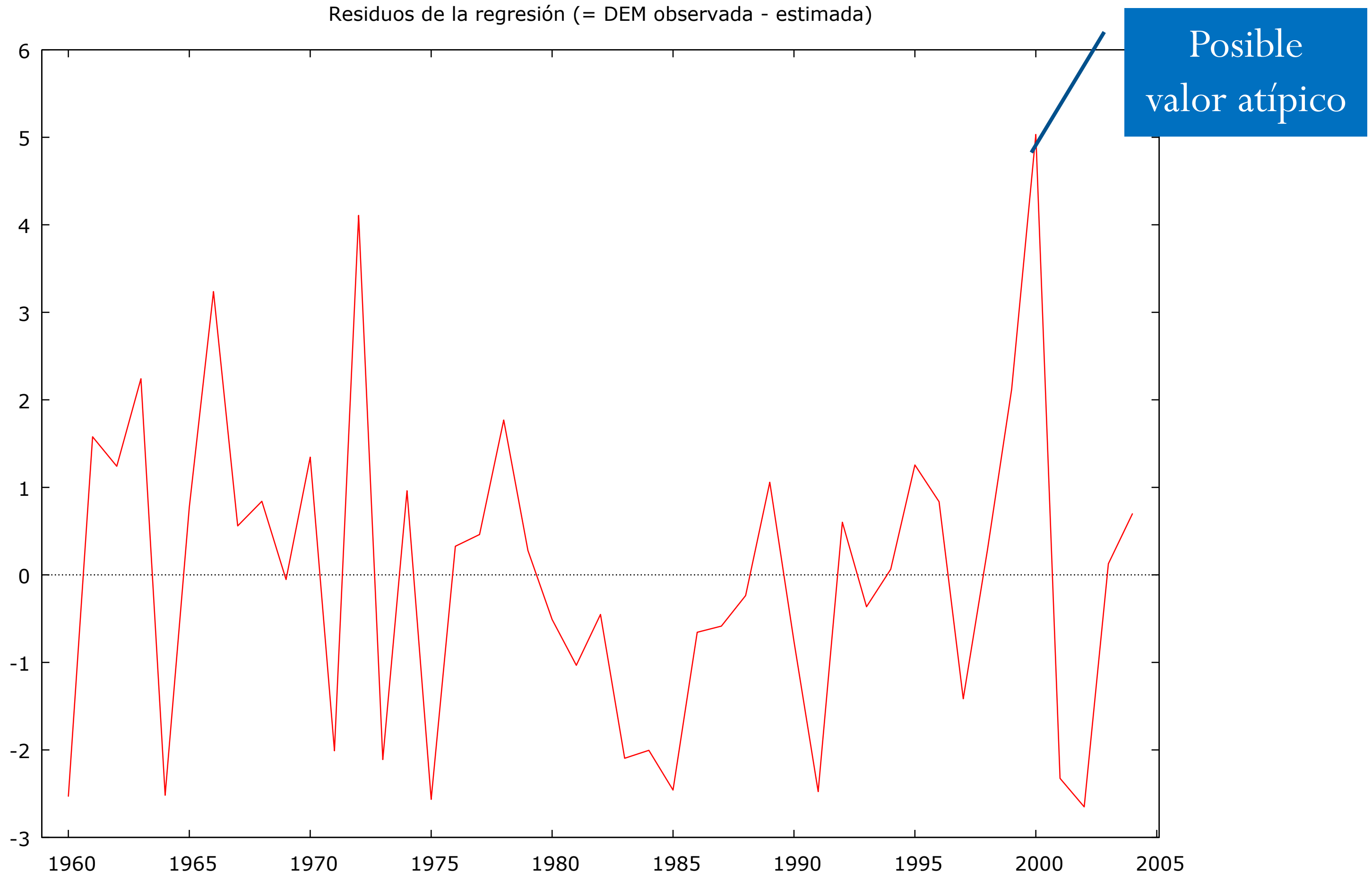
Universidade de Vigo



Atípicos en Demanda de café

Universidade de Vigo

- La sospecha se tiene al observar la lista de residuos o en el gráfico temporal que se obtienen con las opciones de graficos



Comprobación de atípicos en demanda de café

Universidade de Vigo

- Para comprobarlo hace falta calcular el residuo estandarizado. Estos se obtiene guardando los residuos (uhat1) y la desviación estándar de la regresión (sigma_1) y luego definiendo una nueva variable ustand=uhat1/sigma_1.
- En ver se obtienen los valores.

	ustand
1960	-1.366651
1961	0.851202
1962	0.669435
1963	1.208973
1964	-1.357646
1965	0.417928
1966	1.745786
1967	0.301930
1968	0.453618
1969	-0.028602
1970	0.725450
1971	-1.083848
1972	2.215312
1973	-1.138235
1974	0.518116
1975	-1.383533
…………………………	

Nos indica los residuos estandarizados. Los comparamos con el valor de las tablas de una normal estandarizada

Valor atípico individual

Efectos de los valores atípicos

Universidade de Vigo

- Afectan a la distribución de los datos.
- En muestras pequeñas sesgan la estimación.
- Pueden generar un efecto sobre los parámetros si también son influyentes.
- Aumentan el error de la estimación, generando ineficiencia en los estimadores de los coeficientes y sesgo en el estimador de la varianza.

Corrección de los atípicos

Universidade de Vigo

- Existen varias posibilidades. Las mas comunes son:
  - Eliminarlo
    - Es la mas lógica cuando el dato claramente no representa a la población estudiada
  - Incluirlo como variable ficticia
    - Cuando el dato si es de la población pero su comportamiento se aleja por una causa externa. Nos interesa evaluar cuanto cambia respecto al comportamiento normal

Variables ficticias

- Definición
- Binomiales
- Multinomiales
- Regresión con variables ficticias
- Aplicación para solucionar la normalidad

Universidade de Vigo

Definición de Variables ficticias

- Son variables que caracterizan comportamientos cualitativos de forma que indican si una determinada observación verifica o no una propiedad prefijada
- También se les denomina variables indicador de la propiedad o característica
- Generalmente se definen como variables dicotómicas, pero también pueden definirse para variables multinomiales

Variables ficticias dicotómicas

Universidade de Vigo

- Supongamos que tenemos una variable cualitativa dicotómica C, es decir, que se verifica una determinada propiedad o no, que tienen una cualidad o no, etc…, por tanto únicamente puede tomar dos valores A y B.
- Se define la variable ficticia dicotómica como

$$I_A = \begin{cases} 1 \ \text{si} \ \ C = A \\ 0 \ \text{si} \ \ C = B \end{cases}$$

De esta forma se cuantifica el efecto de la variable dicotómica, vale 1 si la cualidad se verifica y 0 si no.

Ejemplos

- En una encuesta responder si o no
- Ser valor atípico o no serlo
- Saber informática o no
- Tener un sexo u otro
- Ser conductor o no
- Ser directivo o no
- ………

Variables ficticias multinomiales

Universidade de Vigo

- Supongamos que tenemos una variable cualitativa multinomial C, es decir, que puede tomar mas de dos valores $C_1, \ldots, C_m$
- Se define una variable ficticia dicotómica para cada uno de los posibles valores.

$$I_{jt} = \begin{cases} 1 \text{ si } C_t = c_j \\ 0 \text{ si } C_t \neq c_j \end{cases} \qquad \begin{matrix} \text{j=1,...m} \\ \text{t=1,...T} \end{matrix}$$

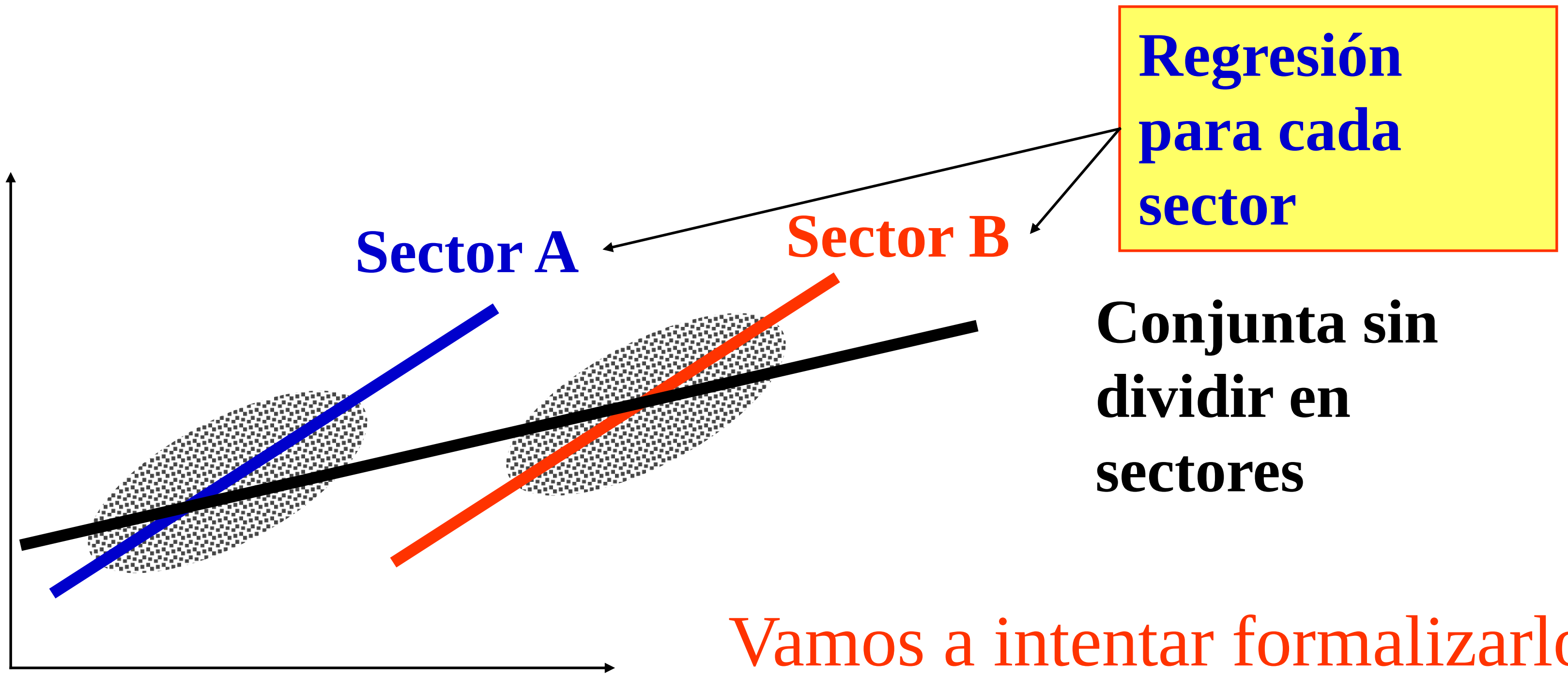
Variables ficticias en la regresión

Universidade de Vigo

- Al incluirlas en una regresión lo hacen como cualquier otra variable, con la diferencia de que el coeficiente nos mide el cambio que se produce por estar en esa categoría en vez de en otra
- **Ejemplo**
- considerar el coste de producir una pieza en dos sectores diferentes A y B

Coste de producción en dos sectores

- El hecho de incluir o no la variable ficticia cambia los parámetros de la regresión



Vamos a intentar formalizarlo

Ejemplos de regresión con variables dicotómicas

Universidade de Vigo

- El caso más habitual es cuando se responde a preguntas sobre gustos, actitudes, etc., únicamente de la forma si o no, sin respuestas intermedias.
- También se usa para medir efectos de cambios en el tiempo por legislaciones o efectos puntuales debidos a un sólo valor o a un conjunto de valores.
- Este será el caso que nos interese para resolver los problemas que se plantean con los valores atípicos, pero previamente veamos como se introducen estas variables en las ecuaciones de regresión y que efectos pueden producir.

Planteamiento de la regresión con variables dicotómicas

Universidade de Vigo

- Supongamos que tenemos una variable cualquiera C que únicamente puede tomar dos valores A y B de forma que ambos son excluyentes y exhaustivos. Entones la variable ficticia se define como

$$I_A = \begin{cases} 1 & \text{si } C = A \\ 0 & \text{si } C = B \end{cases}$$

Regresión con variables dicotómicas

Universidade de Vigo

- En el caso mas simple, se introduciría en el modelo de regresión como una variable cualquiera
- Modelo sin variable ficticia

$$y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \varepsilon$$

- Modelo con variable ficticia

$$y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \alpha I_A + \varepsilon$$

Efecto de la variable ficticia

Interpretación

Universidade de Vigo

- Las pendientes se interpretan igual, pero ahora sería el efecto independientemente del sector
- La constante β_0 sería el coste fijo en el sector B
- La suma de β_0 y α sería el coste fijo en el sector A
- Por tanto α mide la diferencia entre los costes fijos.

Efecto de las variables dicotómicas en la regresión

- Partiendo del modelo sin variable ficticia se puede medir el impacto de esta sobre cada uno de los coeficientes de la regresión cuando se sospecha que cada uno de los grupos tiene una relación diferente totalmente. En ese caso se definen una serie de variables auxiliares que miden el impacto sobre la pendiente

$$IX_j = \begin{cases} X_j \; si \; t \in A \\ 0 \; si \; t \notin A \end{cases} \quad j = 1 \ldots k$$

- Con esas variable el modelo quedaría:

$$y = \beta_0 + \beta_1 X_1 + \cdots + \beta X_k + $$
$$+ \alpha_0 I_A + \alpha_1 IX_1 + \cdots + \alpha_k IX_k + \varepsilon$$

Efecto de la variable ficticia sobre la pendiente de X_k

Interpretación

- Las pendientes se interpretan igual, pero ahora sería el efecto independientemente del sector
- La constante β_0 sería el efecto fijo en el sector B
- La suma de β_0 y α_0 sería el efecto fijo en el sector A
- Por tanto α_0 mide la diferencia entre los efectos fijos
- Cada una de las pendientes β_j sería el impacto de X_j sobre Y en el sector B
- La suma de β_j y α_j nos mediría el impacto de X_j sobre Y en el sector A
- Por tanto cada uno de los α_j nos mide la diferencia entre los impactos en los sectores A y B.

Ejemplo de sectores

Universidade de Vigo

- Consideremos las ventas en miles de euros en función de los empleados en do sectores industriales:

Ventas de cada empleado en el sector A

Ventas promedio suponiendo que no hay empleados en el sector A

Incremento de las ventas promedio suponiendo que no hay empleados del sector B respecto al sector A

```
Modelo 5: MCO, usando las observaciones 1-26
Variable dependiente: Ventas

             Coeficiente   Desv. Típica  Estadístico t   Valor p
  ---------------------------------------------------------------
  const        15.4429       1.04927        14.72       7.19e-013 ***
  Empleados     0.344958     0.175820        1.962      0.0625    *
  X2B           1.00400      0.246818        4.068      0.0005    ***
  sector_B      0.416192     1.53513         0.2711     0.7888

Media de la vble. dep.  20.50960   D.T. de la vble. dep.   4.608494
Suma de cuad. residuos  65.81072   D.T. de la regresión    1.729565
R-cuadrado              0.876052   R-cuadrado corregido    0.859150
F(3, 22)                51.83139   Valor p (de F)          3.86e-10
Log-verosimilitud      -48.96532   Criterio de Akaike      105.9306
Criterio de Schwarz     110.9630   Crit. de Hannan-Quinn   107.3798
```

Incremento de la ventas de cada empleado del sector B respecto al sector A

Gráficos para la diagnosis

Instrumentos gráficos que permiten evaluar la calidad del modelo

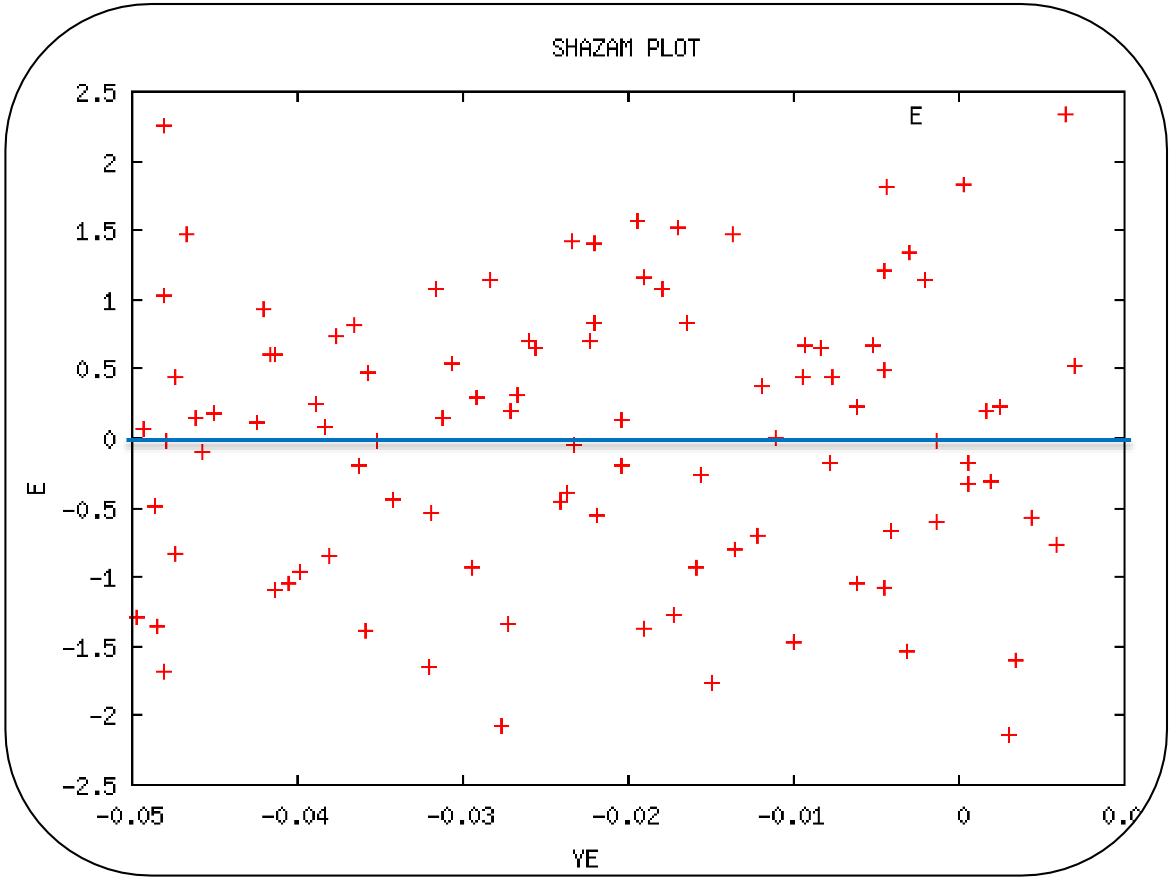
Universidade
de Vigo

Ruido Blanco

Universidade de Vigo

El modelo bien especificado tiene residuos que se comportan como un ruido blanco.

- Sucesión de variables aleatorias independientes:
  - Media cero
  - Varianza constante
  - Simétrica
  - independencia



Comportamiento de los residuos

Universidade de Vigo

- Las perturbaciones del MRLN deberían comportarse como un ruido blanco, puesto que son independientes e igualmente distribuidas con media 0 y simétricas.
- Los residuos de la regresión son aproximaciones a las perturbaciones y aunque no son independientes ni con varianza constante son bastante cercanos a ese comportamiento, por lo que suponemos que bajo todas las suposiciones del modelo deberían comportarse casi como ruido blanco.
- En todos los gráficos donde se representen los residuos respecto a alguna variable, estos deben verificar que bajo las suposiciones del MRLN debería ser cercanos a un ruido blanco.

Gráficos de residuos

Universidade de Vigo

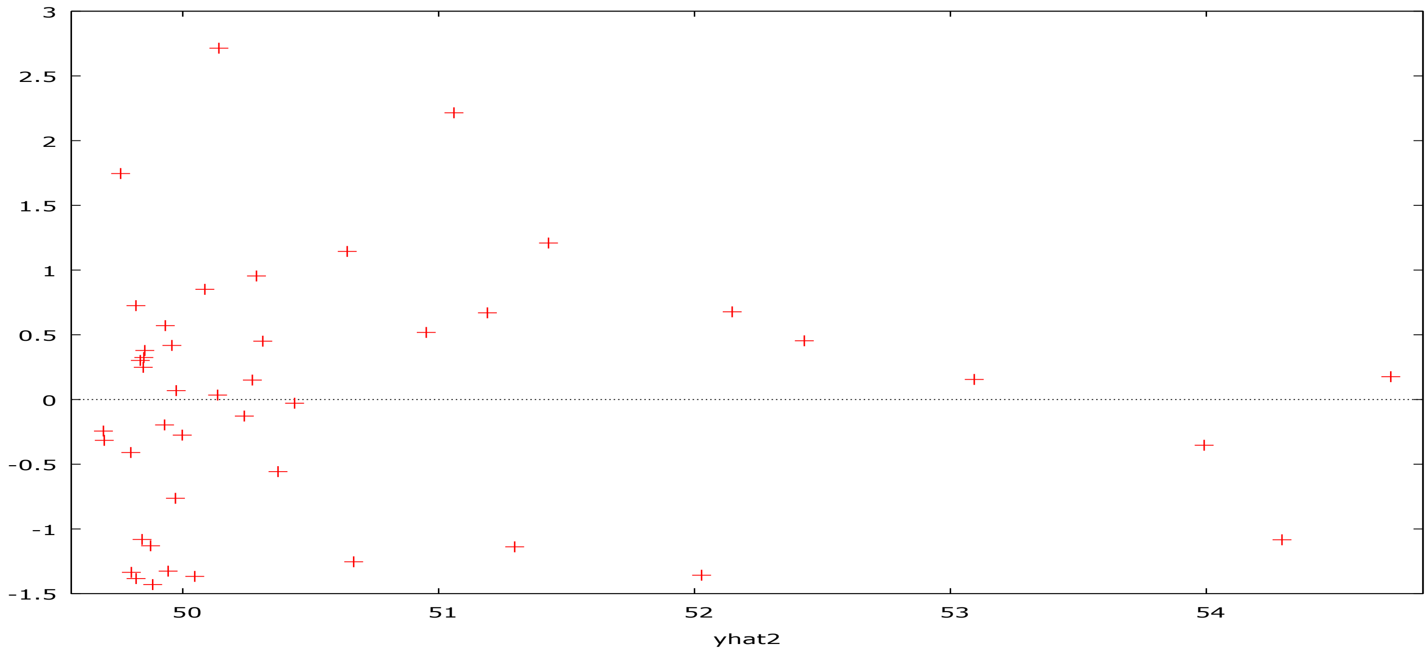
- Por tanto, los residuos debe verificar:
  - E(e)=0;
    - los residuos deben oscilar alrededor del 0, no deben presentar ninguna forma funcional
  - Var(e)= constante aproximadamente.
    - Los residuos deben dispersarse de forma homogénea, estar comprendidos entre dos bandas
  - Independencia:
    - Los residuos no deben presentar relaciones entre ellos, formas graficas encadenadas.
  - Normalidad
    - Los residuos deben estar cercanos al 0, simétricos a cada lado y las bandas no muy alejadas del 0 (alrededor de dos veces la desviación estándar).
- Eso significa que cualquier forma en que se aleje de esas características existe un fallo de alguna suposición.
- El residuo se introduce en el eje de ordenadas. La variable introducida en el eje de abscisas nos dice algo sobre ese fallo.

Gráfico resumen de calidad del modelo

Universidade de Vigo

- Se obtiene representando residuos respecto a valores estimados de la dependiente. Por tanto, en GRETL, se deben guardar los valores estimados y los residuos y luego ver el grafico de ambos

Residuos respecto a valores estimados



Conceptos asociados al gráfico

Universidade de Vigo

- Este gráfico representa los residuos frente a los valores estimados de la variable dependiente.
- Los residuos deben verificar las condiciones del ruido blanco.
- La variable independiente en este caso es el valor estimado de la dependiente, por tanto recoge los posibles efectos de esta variable sobre el comportamiento de los residuos.
  - Esa variable contiene información sobre:
    - La variable dependiente, puesto que es la mejor aproximación lineal en función de las independientes
    - Las variables independientes, puesto que todas intervienen en la estimación.
  - Eso significa que nos da información sobre como influyen tanto la variable dependiente como las independientes sobre los residuos. En ese sentido se habla de resumen de la calidad del modelo puesto que integra todos los componentes de este.

Interpretación del gráfico

Universidade de Vigo

Este grafico permite observar una serie de suposiciones del modelo, puesto que deben manifestar el ruido blanco de los residuos y la independencia con las variables

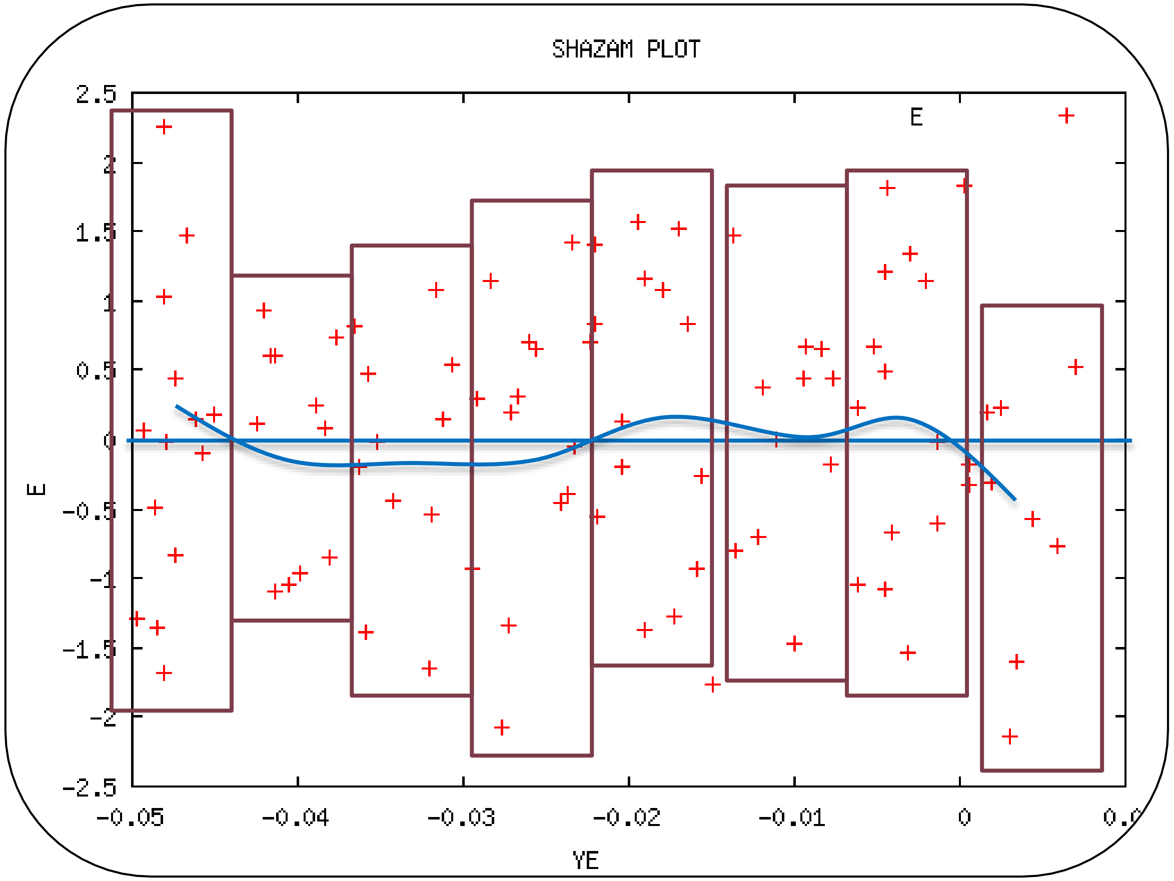
Se observa si existe:

- Linealidad,
- Normalidad
- Homocedasticidad
- Estabilidad
- Independencia

No se observan el resto de las suposiciones, puesto que las suposiciones sobre las variables independientes no se relacionan con los residuos

Linealidad,

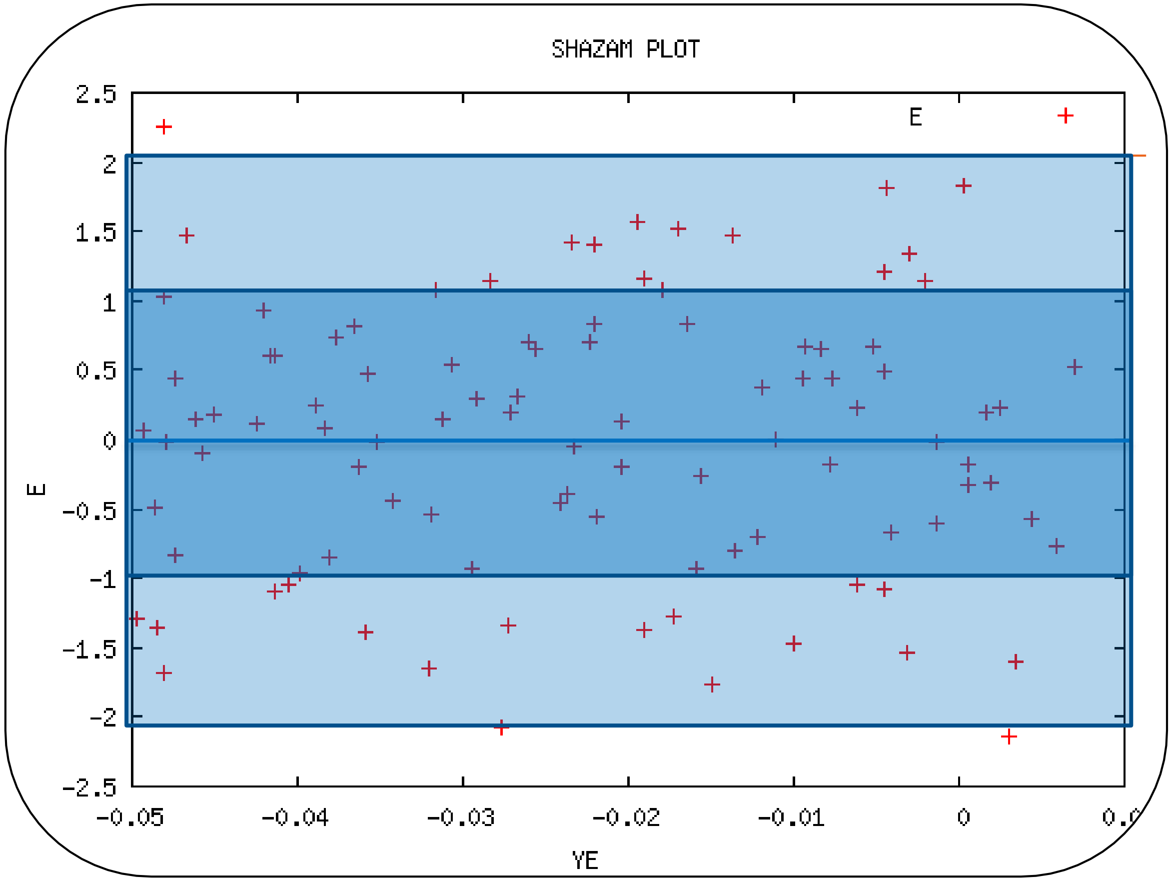
- Se trazan rodajas verticales y se unen sus centros para ver la forma funcional empírica
- no se debe observar ninguna forma funcional.
- La media es constante e igual a cero en todos los casos



Normalidad

Universidade de Vigo

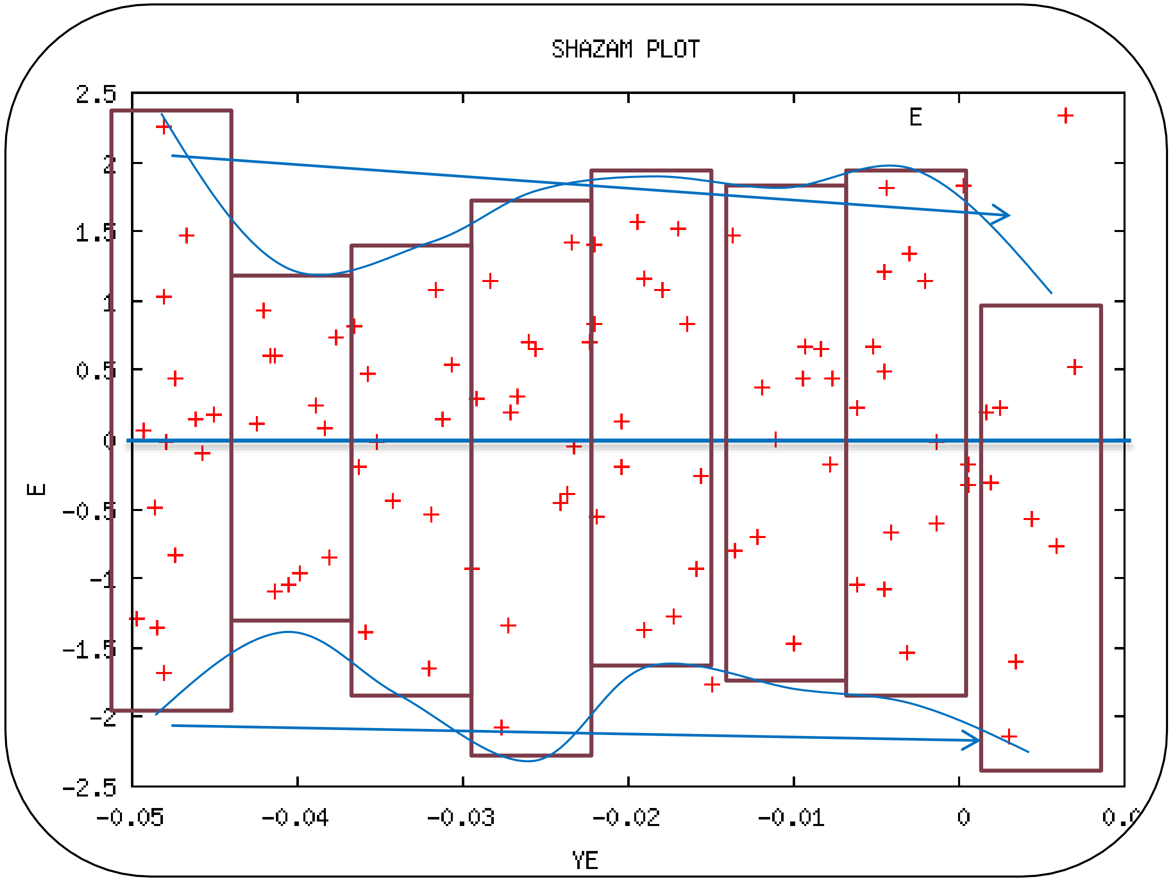
- Los datos se distribuyen de forma simétrica y más concentrados alrededor del 0.
- Se traza una banda central y una banda ampliada
- En la banda entre-s y +s deben estar la mayor parte de los valores (un 75%) y entre -2s y +2s el resto. Los que están fuera son posible s valores raros



Homocedasticidad

Universidade de Vigo

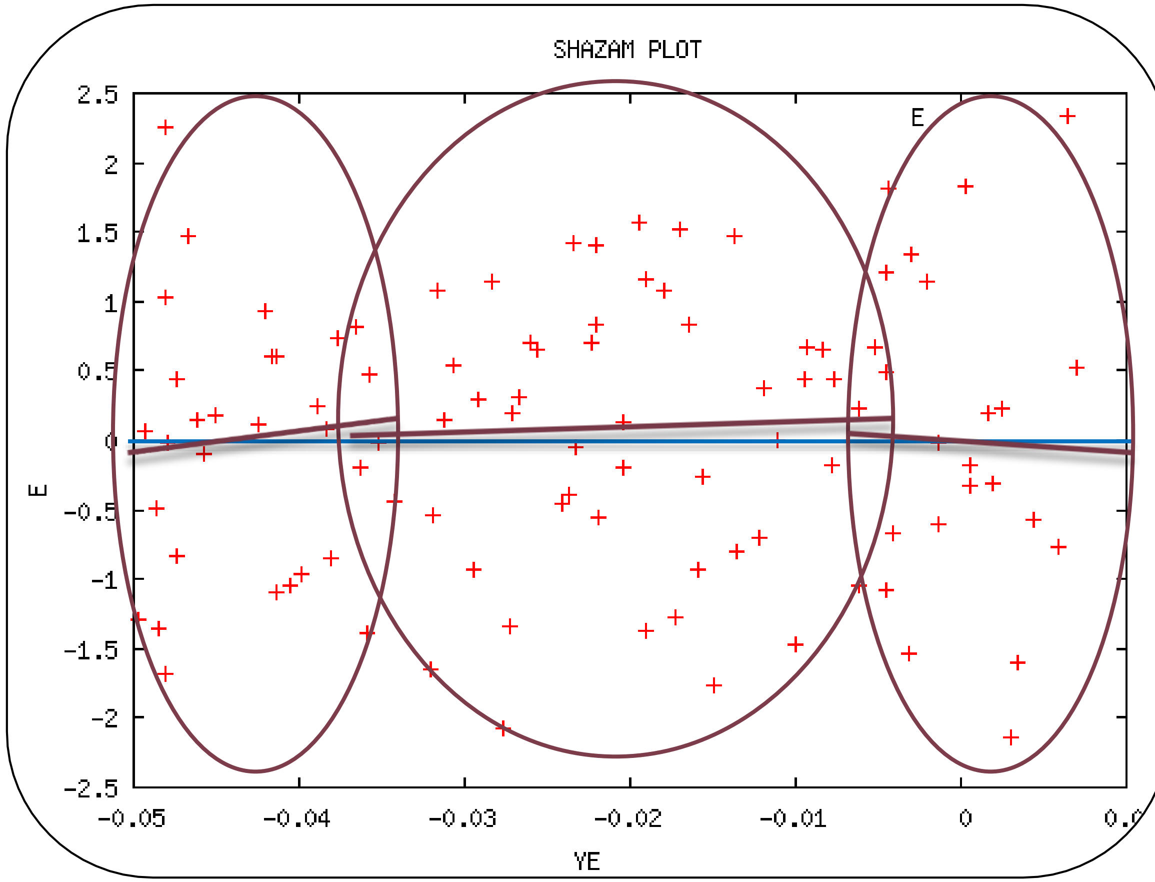
- Se trazan rodajas verticales y se unen sus extremos para ver la variación de la amplitud de esas rodajas
- La varianza es constante, dispersión igual para todos la serie.
- Las amplitudes de ambos lados deben ser mas o menos constantes



Estabilidad

Universidade de Vigo

- No se producen cambios repentinos de la serie de residuos
- Se buscan las nubes de puntos existentes, dibujando los cortes que se intuyan
- Si aparece más de una con diferente tendencia, indica fallo de estabilidad
- Si hay varias pero la tendencia de las nubes de puntos no varía, hay estabilidad



Independencia

Universidade de Vigo

- No se observan relaciones entre residuos o grupos de residuos en la evolución de los datos
- Se espera un comportamiento aleatoria, la forma en que quedarían unas monedas al tirarlas al azar al aire.
- Por consiguiente no deben verse "gusanitos" o grandes "huecos"

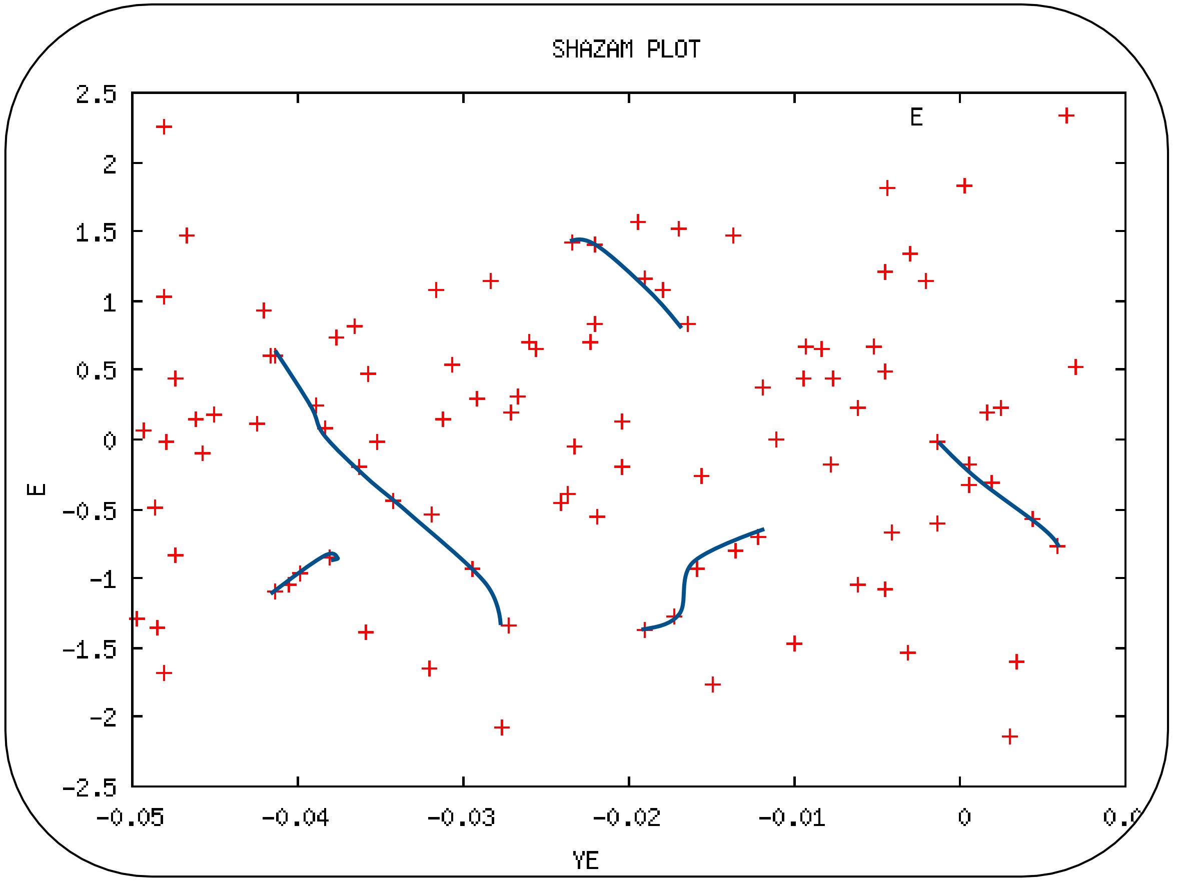
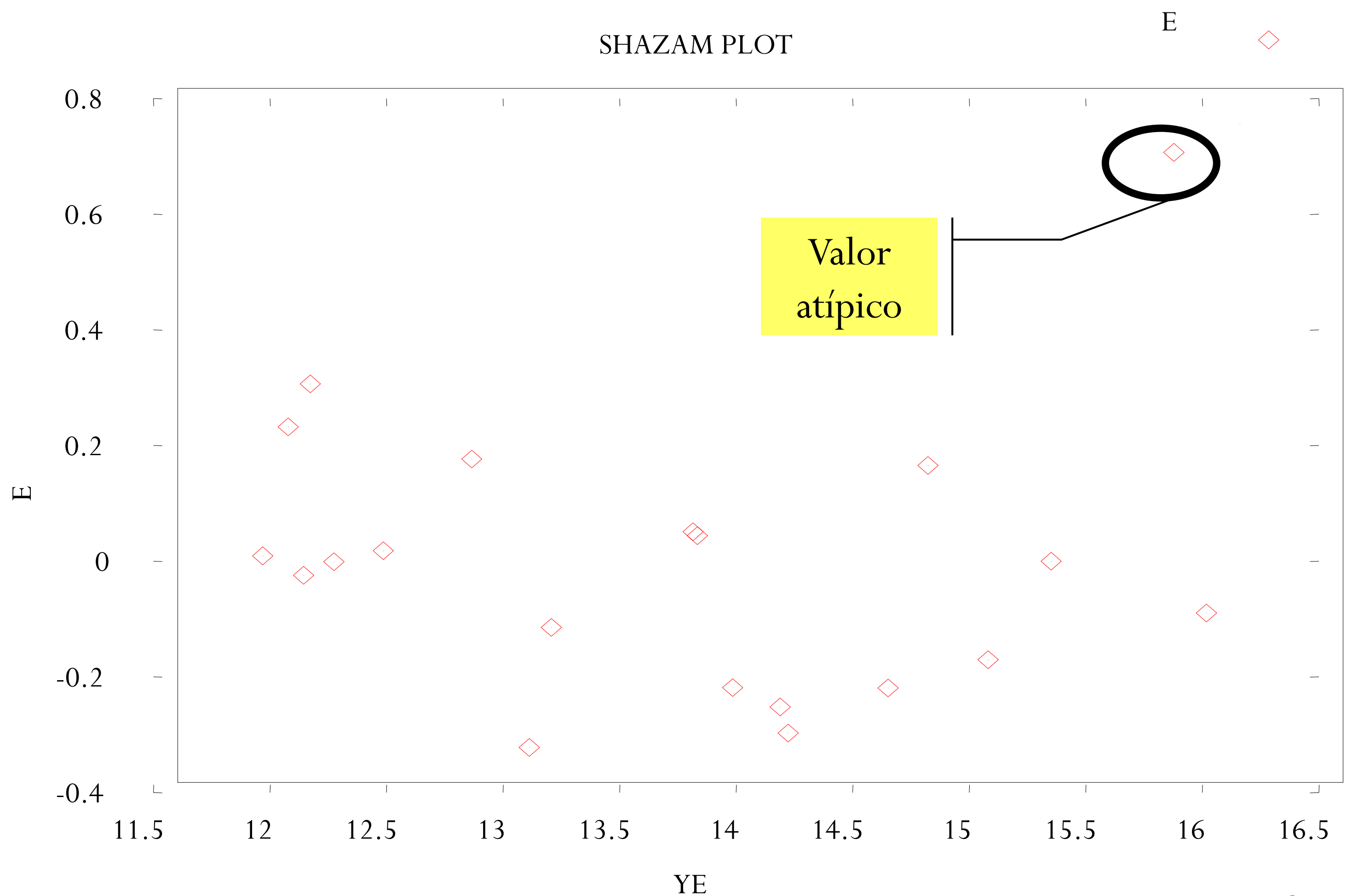


Gráfico de residuos respecto a valores predichos en la empresa XUMA

Universidade de Vigo

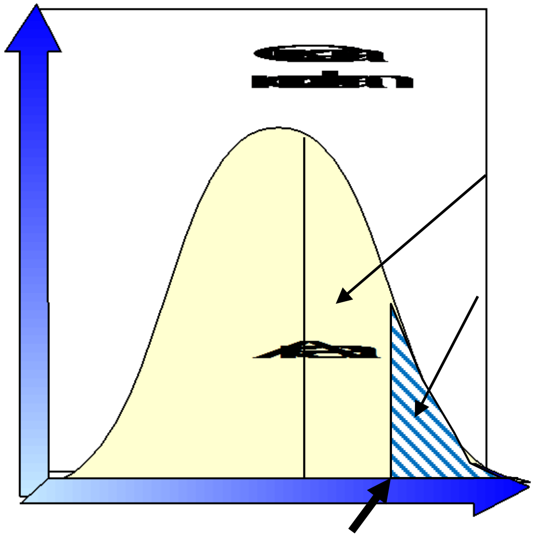


Ejercicio 3.2

La no normalidad de las perturbaciones

Normalidad

Universidade de Vigo



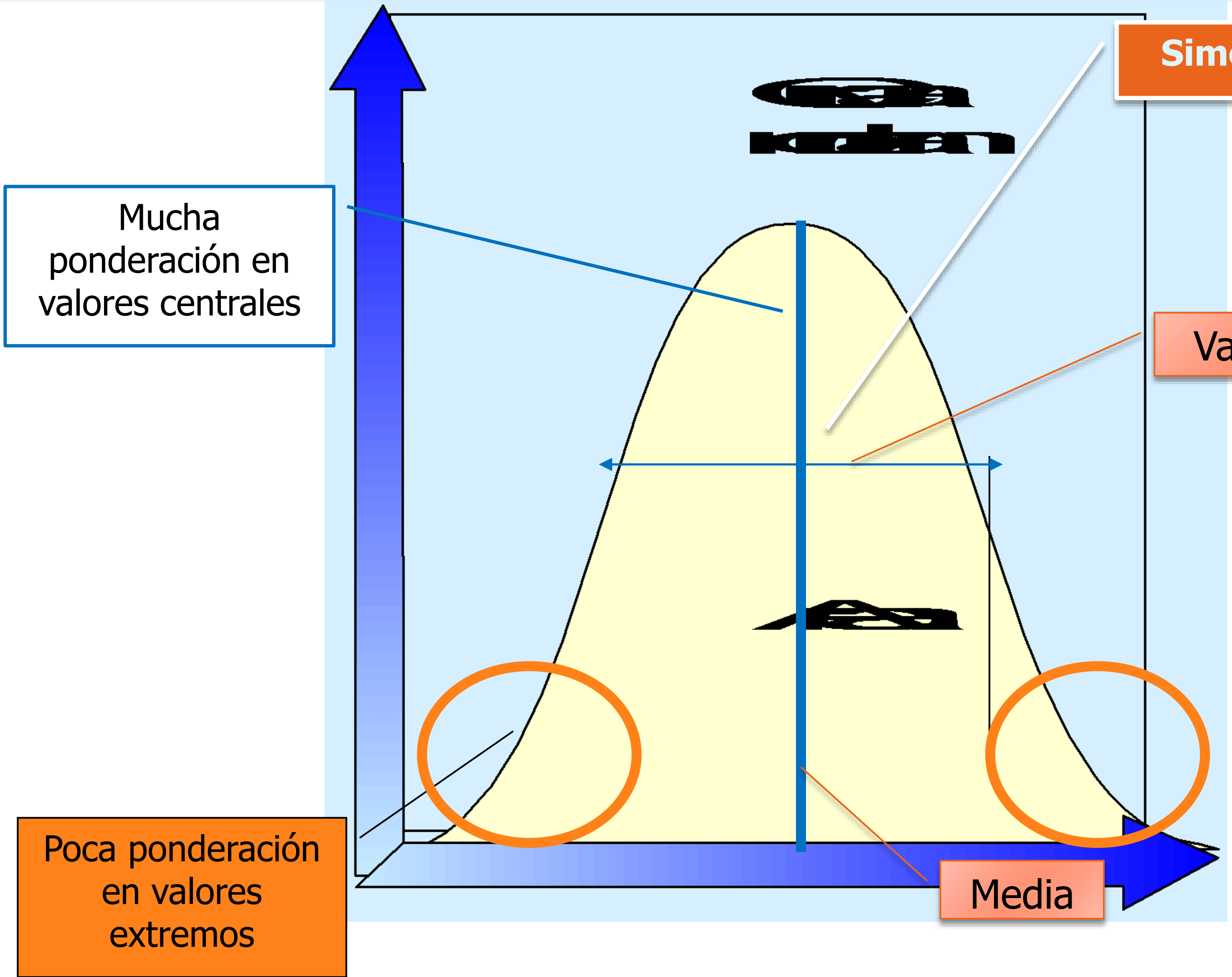
Normalidad

Universidade de Vigo

- Nos dice si los datos con los que trabajamos siguen leyes de distribución normales o no. Su comprobación es necesaria, para realizar los test de hipótesis exactos y los intervalos de confianza en el MRLC.
- El comportamiento normal se denomina así porque tiende a ponderar más los valores centrales y menos los extremos, además de ser simétrica.
- Caracterizada por media y varianza

Comportamiento normal

Universidade de Vigo



Efectos de la no normalidad

Universidade de Vigo

- Si no se verifica la normalidad del modelo, entonces los estimadores MCO dejan de ser MV y por tanto pierden la eficiencia dentro de los estimadores insesgados, sin embargo siguen siendo ELIO.
- Mantienen la consistencia y la normalidad asintótica, pero también pierden la eficiencia asintótica.
- Dejan de seguir leyes normales, pero son asintóticamente normales, es decir siguen leyes normales aproximadas cuando la muestra es grande (mayor de 30).
- El estimador de la varianza sigue siendo insesgado y consistente, pero deja de ser asintóticamente eficiente.
- Los estimadores MV en general, verificarán mejores propiedades.

Propiedades asintóticas

Universidade de Vigo

- Propiedades que se verifican solo en grandes nuestras, son aproximadas y al crecer el tamaño muestral se puede afirmar con más categoría su validez:
  - Asintóticamente normal:
    - su ley de distribución converge a una normal
  - Asintóticamente insesgado:
    - Su esperanza converge al verdadero valor del parámetro
  - Asintóticamente eficiente:
    - Su error cuadrático medio tiende a ser mínimo.
  - Consistente:
    - Converge en probabilidad al verdadero valor del parámetro, es decir al crecer el tamaño muestral, fijado un error, la probabilidad del intervalo de confianza del parámetro tiende a 1.

Ejemplo de interpretación

Universidade de Vigo

- Supongamos que en la demanda de café fallara la normalidad, entonces la interpretación de la salida cambiaría:

Su ley de distribución no es una t, es asintóticamente normal

No son los mas eficientes dentro de todos los inssesgados

Modelo 2: MCO, usando las observaciones 1960-2004 (T = 45)
Variable dependiente: DEM

	iciente	*Desv. Típica*	*Estadístico t*	*Valor p*	
const	41.7402	7.64881	5.4571	<0.0001	
P	0.0715555	0.306738	0.2333	0.8167	
RTA	0.179393	0.0396732	4.5218	<0.0001	***

Ya no es válido hay que usar una normal

Media de la vble. dep.	50.65181	D.T. de la vble. dep.	2.212548
Suma de cuad. residuos	144.3694	D.T. de la regresión	1.854014
R-cuadrado	0.329750	R-cuadrado corregido	0.297833
F(2, 42)	10.33158	Valor p (de F)	0.000224
Log-verosimilitud	-90.08078	Criterio de Akaike	186.1616
Criterio de Schwarz	191.5815	Crit. de Hannan-Quinn	188.1821
rho	-0.058935	Durbin-Watson	2.069583

Sigue siendo válido, pero ya no sirve para contrastar si el efecto conjunto es significativo

Sigue siendo válido, per su ley de distribución no es conocida

Ejercicio 3.3.

Causas de la no Normalidad

Universidade de Vigo

1. Existencia de valores atípicos
2. Distribuciones no normales
   - Formas no simétricas, no están centradas en la media:
     - Fallo de la simetría
   - Mayor masa probabilística en el centro que la normal
   - Mayor masa en los extremos que la normal
     - Fallo de la curtósis

Identificación de la Normalidad

Universidade de Vigo

- Gráficos
  - Residuos
  - Histogramas
- Test de hipótesis
  - Pretenden comprobar la distribución normal de las perturbaciones a partir de alguna regla de decisión estadística.
    - Bondad de ajuste, compara la distribución teórica con la empírica, pero se aplica a intervalos.
    - Jarque-Bera, que estudia la simetría y curtósis de la densidad empírica.

Histograma

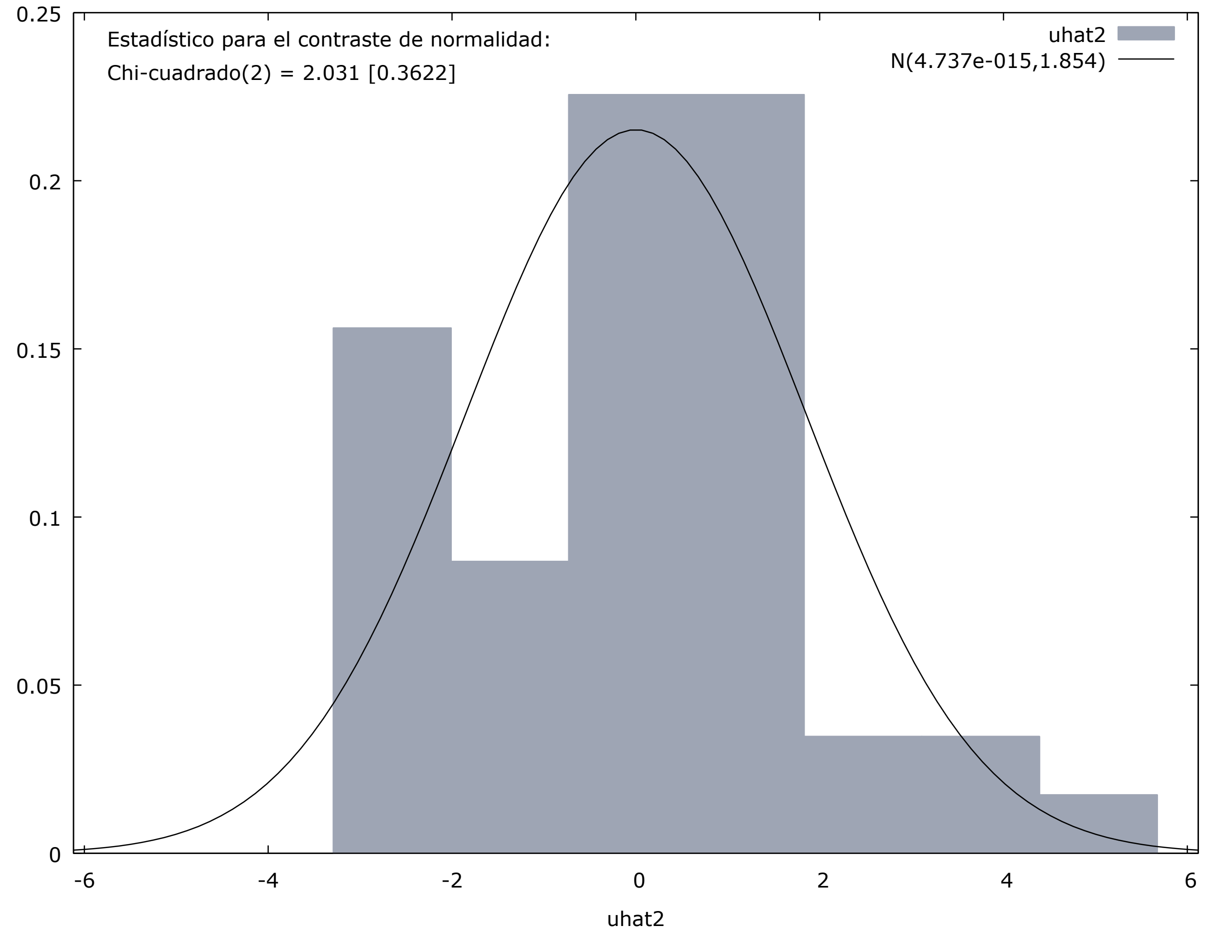
Universidade de Vigo

- Es de sobra conocido. Representa la frecuencia de cada intervalo de observaciones.
- Se define un conjunto de intervalos.
- Se busca el numero de observaciones en ese intervalo
- Se representa la altura por el cociente ente número y amplitud.
- Si los intervalos son todos iguales se puede utilizar directamente la frecuencia.

Histograma en GRETL

Universidade de Vigo

Sale con el contraste de normalidad de residuos



Test de hipótesis

Universidade de Vigo

- Los gráficos nos dan una idea de los posibles fallos, pero para contrastarlos debemos utilizar los test de hipótesis.
- Vamos a recordar algunas ideas de los test de hipótesis para contrastar suposiciones.
- Haremos uso del test de Doornik–Hansen

Test de significación para contrastar suposiciones del MRLN

Universidade de Vigo

- Cuando se quieren contrastar las suposiciones del MRLN, siempre se parte del modelo, con alguna generalización, es decir se suponen validas todas las suposiciones excepto la que se quiere contrastar.
- En el caso de la normalidad se suponen todas menos la normalidad de las perturbaciones.

Modelo de contraste de normalidad

Universidade de Vigo

$$\varepsilon_t = y_t - E\left({}^{Y_t}\!\big/_{X_{1t}\ldots X_{kt}} \right) = y_t - (\beta_0 + \beta_1 X_{1t} + \ldots + \beta_k X_{kt})$$

Donde:

- ε son independientes e igualmente distribuidas y no dependen de las X (Independencia, homocedasticidad y exogeneidad),
- β son estables y estimables (Estabilidad e identificabilidad)
- X no están relacionadas entre sí y vienen dadas sin error (no colinealidad y mensurabilidad)

Resultados del modelo

Esas suposiciones nos permiten:

- Estimar las perturbaciones a partir de los errores de MCO.
- Suponer que los residuos son aproximadamente independientes e igualmente distribuidos con leyes de media 0 y varianza constante, lo que nos permite comparar la distribución empírica con una normal teórica.
- Cuando se comparan las funciones de densidad se realiza el test de Doornik–Hansen.
- Por tanto haremos uso de este modelo para los test de hipótesis

Test de Doornik–Hansen

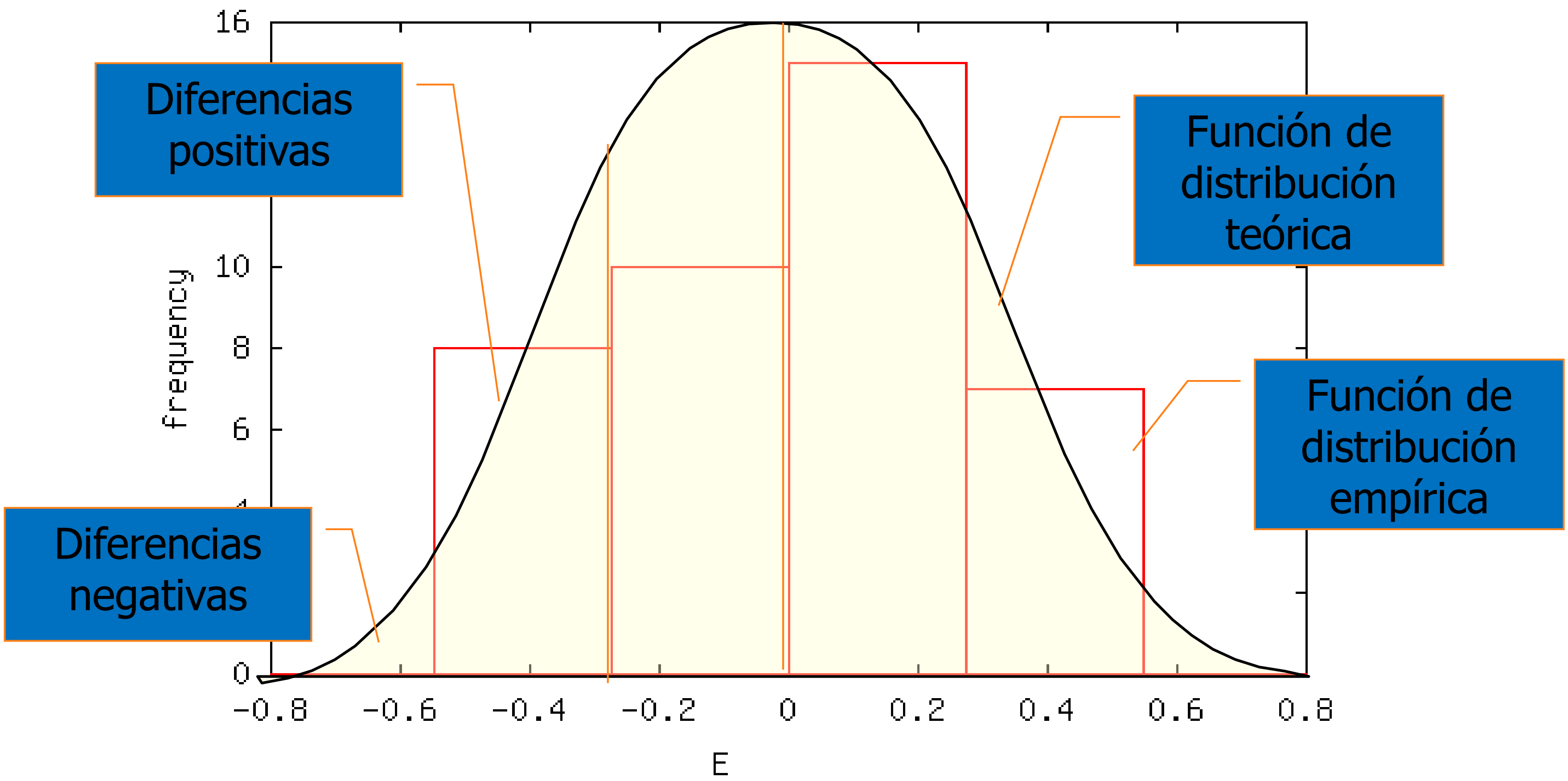
- Hipótesis
  - H_0: La simetría y el exceso de curtosis son nulos
  - H_1: existe simetría o exceso de curtosis
- Compara la distribución teórica de los residuos con la empírica pero en realidad las hipótesis se hacen sobre las perturbaciones teóricas.
- Analiza las funciones de densidad, tanto teórica como empírica, es decir hace uso de los histograma y la función de densidad gaussiana, es decir si la simetría y curtosis del histograma coinciden con el de la normal

Estadístico del Test de bondad de ajuste

Universidade de Vigo

- Como nos interesa comparar las distribuciones teórica y empírica, calculamos la simetría y la curtosis de la distribución empírica (el histograma).
- Luego comparamos esos valores con los valores teóricos de la distribución normal, simetría 0 y curtosis 3.
- Las distancias se toman al cuadrado se estandarizan y se corrigen de acuerdo a los valores estimados por Doornik–Hansen
- El estadístico sigue una ley chi cuadrado con 2 grados d elibertad.

Histograma teórico y empírico

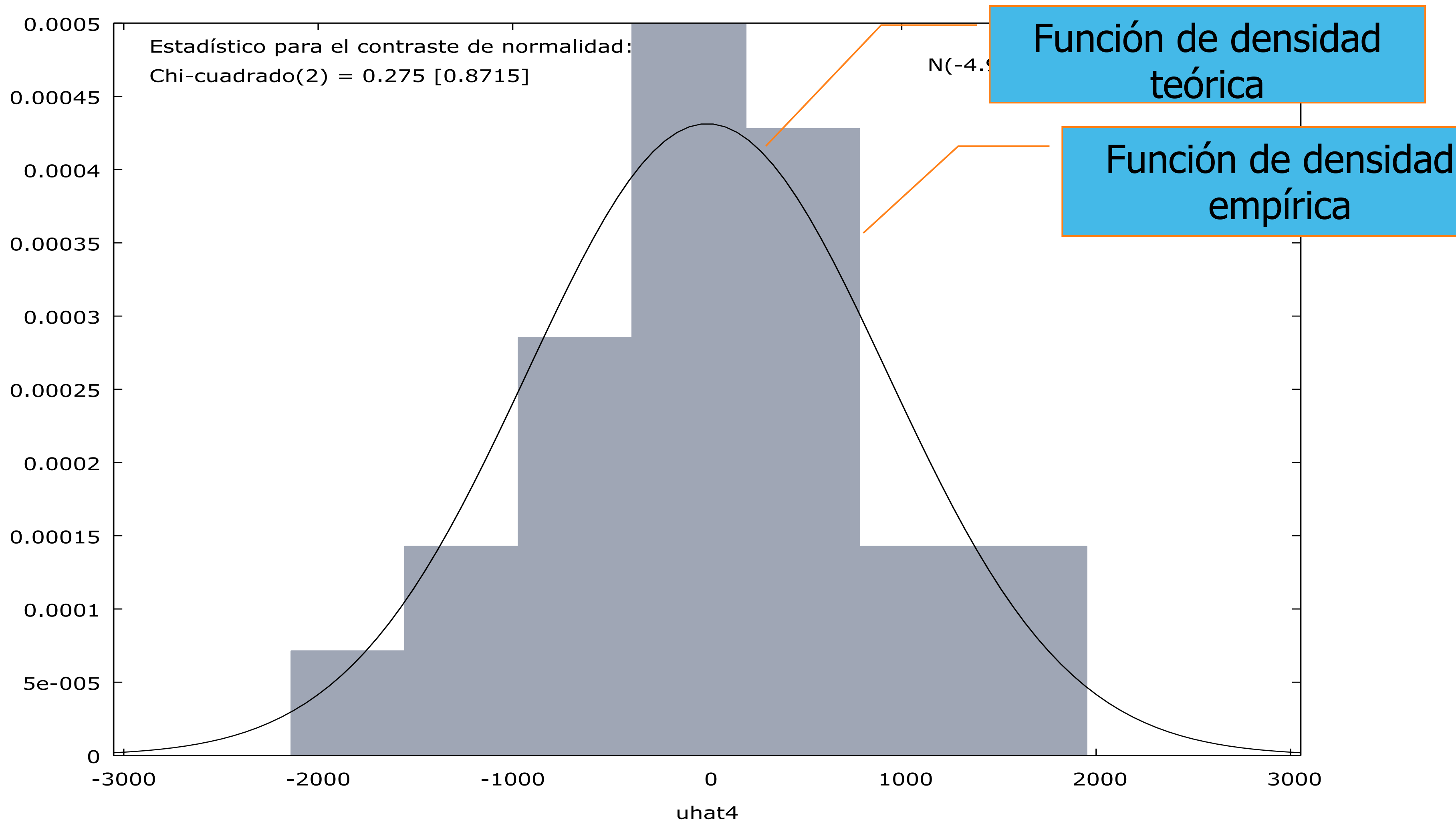


Empresas-normalidad


```
Distribución de frecuencias para uhat4, observaciones 1-24
número de cajas = 7, media = -4.92643e-013, desv.típ.=924.811

      intervalo     punto medio   frecuencia  rel     acum.

          < -1554.0   -1846.1        1       4.17%    4.17% *
   -1554.0 - -969.92   -1262.0        2       8.33%   12.50% ***
   -969.92 - -385.80   -677.86        4      16.67%   29.17% ******
   -385.80 -  198.32   -93.744        7      29.17%   58.33% **********
    198.32 -  782.44    490.38        6      25.00%   83.33% *********
    782.44 -  1366.6    1074.5        2       8.33%   91.67% ***
          >=  1366.6    1658.6        2       8.33%  100.00% ***

Contraste de la hipótesis nula de distribución normal:
Chi-cuadrado(2) = 0.275 con valor p  0.87152
```

Test de bondad de ajuste

Universidade de Vigo



Tratamiento de la normalidad

Universidade de Vigo

1. Si la distribución es conocida, aunque no sea normal, se aplica estimación MV.
2. Si la distribución es desconocida, se puede utilizar:
   a. Transformaciones buscando normalidad.
   b. Regresión robusta.
3. Si la no normalidad es debida a valores atípicos
   a. Se utilizan variables ficticias.
   b. Se eliminan si hay suficientes datos.

Ejercicio 3.4

El problema de la multicolinealidad

Universidade
de Vigo

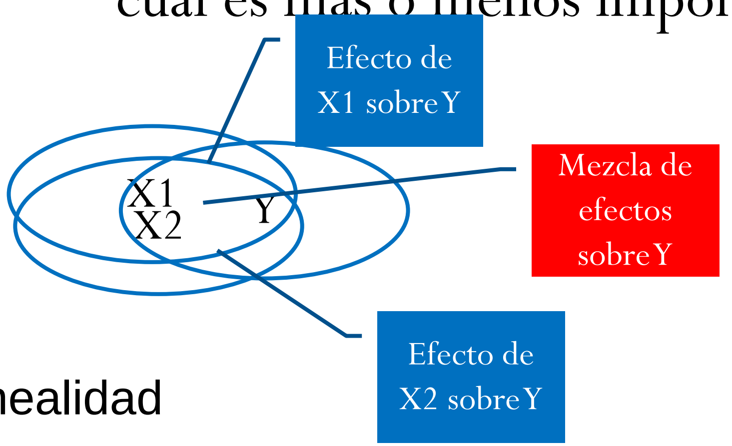
Multicolinealidad

Universidade de Vigo

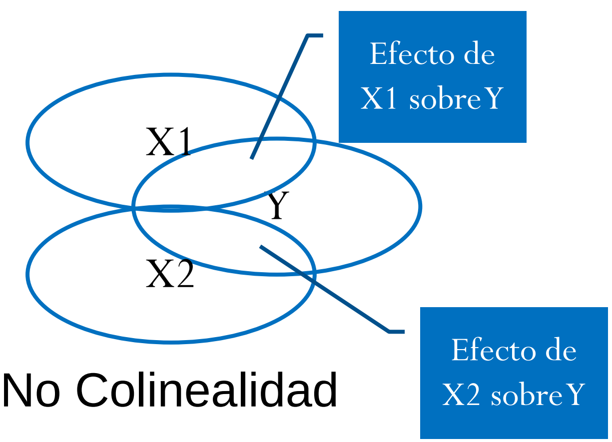
- La multicolinealidad se produce cuando las variables independientes comparten información sobre la dependiente.
- Normalmente su efecto únicamente genera una falta de eficiencia, incrementando la variabilidad de los estimadores al crecer la inversa de (X'X).
- Sin embargo, si la multicolinealidad es severa puede invalidar el modelo e incluso hacer imposible el calculo de los estimadores de MCO, puesto que la matriz (X'X) puede no tener inversa.
- Por ese motivo interesa encontrar estadísticos que detecten la multicolinealidad

Efectos de la multicolinealidad

- Incrementa el error estándar de la estimación de los coeficientes, al incrementar el determinante de la matriz inversa de (X'X).
  - Perdida de la eficiencia de los estimadores
  - Disminución de los t-estadísticos, haciendo que las variables dejen de ser significativas.
  - Mezcla de efectos entre las variables independientes, sin saber cual es mas o menos importante



Colinealidad



No Colinealidad

Detección de la multicolinealidad

Universidade de Vigo

- Se trata de buscar instrumentos que indiquen si las variables independientes están relacionadas ente sí.
- Existen dos tipos de instrumentos:
  - analizar la relación de una variable en función de las otras
    - En este caso el instrumento se basa en la dependencia de una variable en función de otras. Se suele usar el **factor de inflación de la varianza (VIF)**
  - buscar de modo conjunto las interrelaciones entre todas las variables independientes.
    - En este caso se hace uso de lo autovalores de la matriz (X'X) que miden la interrelación entre las variables. Da pie a dos indicadores:
      - **El índice de condición**
      - **La proporción de la varianza**

Factor de inflación de la varianza

- Se define como

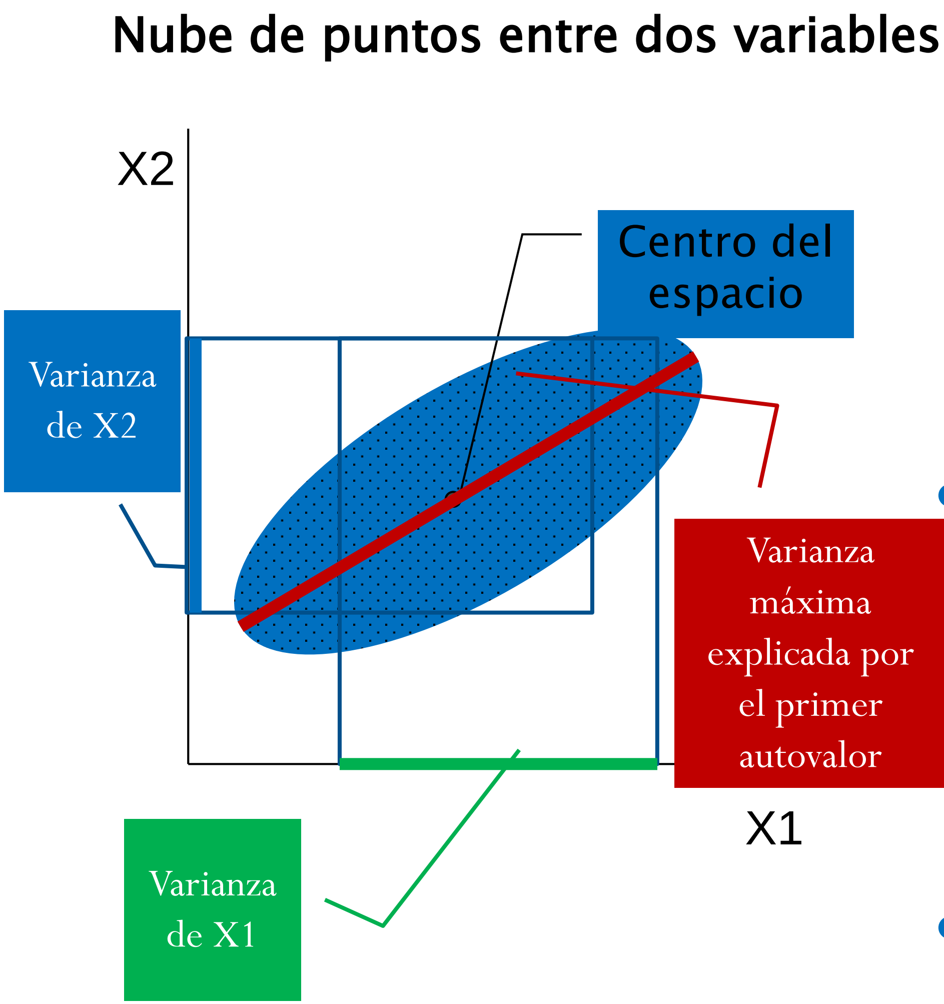
$$\mathrm{VIF(j)} = \frac{1}{1 - R_j^2}$$

donde R(j) es el coeficiente de correlación múltiple entre la variable j y las demás variables independientes

- De esa forma mide la dependencia de una variable en función de las demás.
- Lo ideal es que ese valor sea cercano a 1 pues eso indica que no hay nada de colinealidad.
- Es común recomendar que ese valor no sea mayor de 10

Autovalores y varianza explicada en una nube de puntos

Universidade de Vigo



- Los autovalores de una matriz indican la dispersión de las diferentes combinaciones de las variables independientes ordenándolas por orden de dispersión.
- El primer autovalor indica la varianza explicada por la combinación lineal de las variables que mas explica la dispersión de los datos
- Mucha dispersión explicada indica mucha relación entre las variables

Índice de condición

Universidade de Vigo

- Se calcula a partir de los autovalores de la matriz X'X,
- Se define como el cociente entre el autovalor máximo y el mínimo.
- Por tanto mide si hay mucha diferencia entre la varianza explicada por la máxima dispersión y la mínima.
- Si el valor es muy alto indica que una variable es prácticamente explicada por el resto, lo que implica multicolinealidad.
- Para Belsley índices de condición entre 5 y 10 están asociados con una colinealidad débil, mientras que índices de condición entre 30 y 100 señalan una colinealidad moderada a fuerte
- Mide la colinealidad de modo conjunto, puesto que cada autovalor no se refiere solo a una variable.

D.A. Belsley
Conditioning Diagnostics: Collinearity and Weak Data in Regression.
John Wiley &Sons. 1991

Proporciones de la varianza

- Se calculan a partir de la descomposición de la matriz X'X, puesto que esta interviene en la varianza de los estimadores.
- De esa forma, se encuentra cuanto aporta cada componente y, por tanto su colinealidad, a la varianza de cada coeficiente estimado.
- Cada columna de la proporción de la varianza suman 1, por tanto si una columna es muy alta en relación al resto indica que ese coeficiente tiene un alto incremento de la varianza debido a la multicolinealidad

Detección conjunta de la multicolinealidad

Universidade de Vigo

- Belsley propone usar conjuntamente los índices de condición y la proporción de descomposición de varianza para realizar el diagnóstico de colinealidad, usando como umbral de proporción alta 0,5 de modo que, finalmente, dicho diagnóstico se hará:
  - Los índices de condición altos (mayores que 30) indican el número de colinealidades y la magnitud de los mismos mide su importancia relativa.
  - Si un componente tiene un índice de condición mayor que 30 y dos o más variables tienen un proporción de varianza alta en el mismo, esas variables son colineales.
- En GRETL se obtiene con el comando VIF después del OLS

Ejemplo: colinealidad en Nerlov

Factores de inflación de varianza (VIF)
Mínimo valor posible = 1.0
Valores mayores que 10.0 pueden indicar un problema de colinealidad

KWH	1.055
PL	1.229
PF	1.179
PK	1.078

VIF(j) = 1/(1 - R(j)^2), donde R(j) es el coeficiente de correlación múltiple entre la variable j y las demás variables independientes

Diagnósticos de colinealidad de Belsley-Kuh-Welsch:

		--- proporciones de la varianza ---				
lambda	cond	const	KWH	PL	PF	PK
4.331	1.000	0.000	0.016	0.001	0.003	0.001
0.596	2.696	0.000	0.922	0.000	0.005	0.000
0.056	8.779	0.010	0.028	0.008	0.910	0.020
0.014	17.554	0.000	0.020	0.469	0.037	0.296
0.003	36.930	0.989	0.015	0.522	0.044	0.683

lambda = valores propios de X'X, del más grande al más pequeño
cond = índice de condición
nota: las columnas de proporciones de la varianza suman 1.0

Solución a la multicolinealidad

- La mas común es eliminar aquellas variables que están muy relacionadas con otras puesto que no aportan información importante al modelo.
- Una forma de hacer es la regresión paso a paso, es decir ir introduciendo las variables según su importancia en el modelo, a través de prueba de cada una.

Nerlov sin los precios de energia

Universidade de Vigo

Factores de inflación de varianza (VIF)

Mínimo valor posible = 1.0

Valores mayores que 10.0 pueden indicar un problema de colinealidad

KWH	1.051
PL	1.159
PF	1.131

VIF(j) = 1/(1 - R(j)^2), donde R(j) es el coeficiente de correlación múltiple entre la variable j y las demás variables independientes

Diagnósticos de colinealidad de Belsley-Kuh-Welsch:

		--- proporciones de la varianza ---			
lambda	cond	const	KWH	PL	PF
3.368	1.000	0.001	0.028	0.001	0.006
0.576	2.417	0.001	0.909	0.001	0.009
0.049	8.264	0.060	0.040	0.034	0.959
0.007	22.239	0.938	0.023	0.964	0.026

lambda = valores propios de X'X, del más grande al más pequeño

cond = índice de condición

nota: las columnas de proporciones de la varianza suman 1.0

Ejercicio 3.5

La no linealidad del modelo

Universidade
de Vigo

El concepto de linealidad

Universidade de Vigo

- Indica que el valor esperado de la variable dependiente depende linealmente de las variables independientes
- El impacto esperado por un cambio unitario de cada una de las variables independientes, manteniendo las otras constantes, es siempre el mismo.

La constante recoge el efecto básico de las X mas el efecto combinado de todas las variables no incluidas en el modelo

$$E\left(Y_t / X_{1t} \ldots X_{kt}\right) = \beta_0 + \beta_1 X_{1t} + \ldots + \beta_k X_{kt} =$$

$$= \sum_{i=0}^{k} \beta_i X_{it} \text{ con } X_{0t} = 1 \; \forall t$$

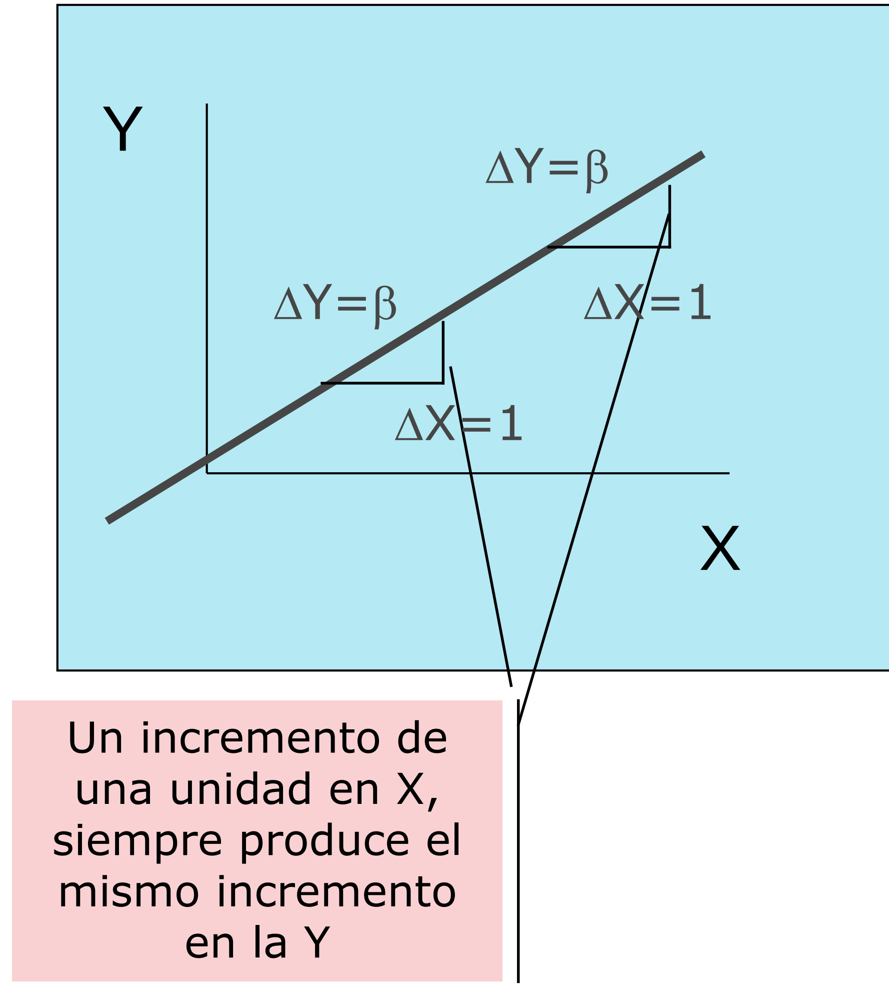
El parámetro mide la derivada parcial, es decir el efecto de cada X sobre la Y

El impacto de cada variable independiente es constante

Se define sobre el valor esperado de la variable dependiente condicionado a la información suministrada por las independientes

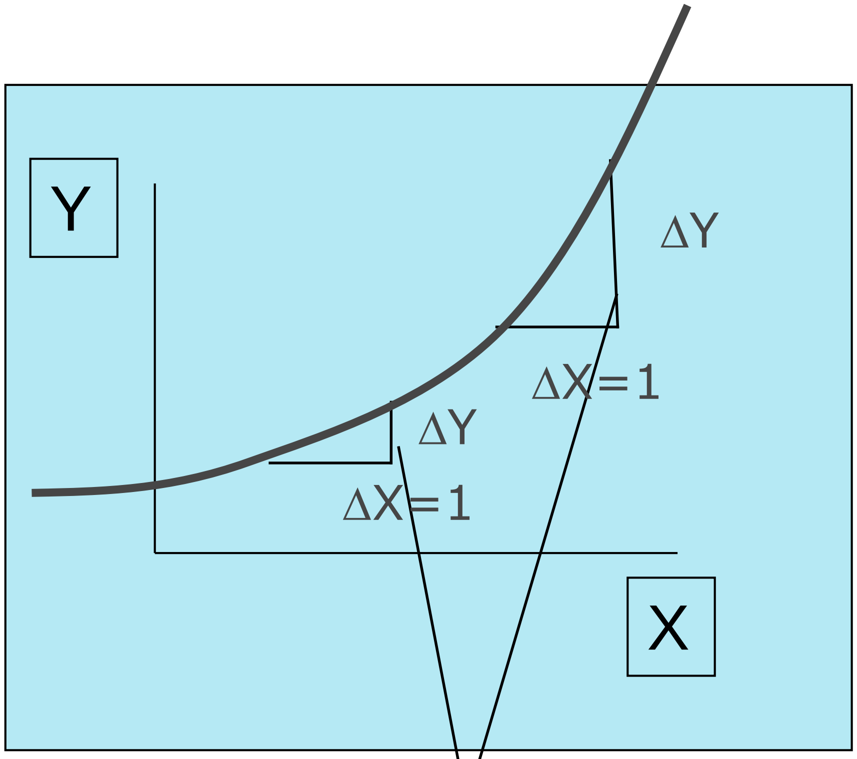
Relación lineal

- La variable respuesta depende linealmente de los regresores. El valor esperado de la variable dependiente, condicionado al conocimiento de las variables independientes, es función lineal de un conjunto de dichas variables



Relación no lineal

- La variable respuesta no depende linealmente de los regresores. El valor esperado de la variable dependiente, condicionado al conocimiento de las variables independientes, es función no lineal de un conjunto de dichas variables



Un incremento de una unidad en X, no siempre produce el mismo incremento en la Y

Ejemplo: Costes variables en CENSA

Universidade de Vigo

- Para analizar todo este tema vamos a hacer uso de un ejemplo en el que se analiza la relación entre coste y cantidad.
- El coste de fabricación de celulosa en una empresa (CENSA) depende de la cantidad de celulosa producida. Los datos se recogen en la tabla siguiente. Comprobar cual es la función que relaciona el coste con la producción y obtener una estimación de los costes fijos y los costes variables en esa empresa.
- El modelo sería el siguiente:

$$Y_t = E\left(Y_t / X_{1t}\right) + \varepsilon_t$$

Modelo general

Linealidad

$$E\left(Y_t / X_{1t}, X_{2t}\right) = \beta_0 + \beta_1 X_{1t}$$

Efectos de la no linealidad

- Las propiedades de consistencia e insesgadez se basaban en la siguiente relación:

$$b=(X'X)^{-1}X'y=(X'X)^{-1}X'(X\beta+\varepsilon)=\beta+(X'X)^{-1}X'\varepsilon$$

- Que exige la linealidad del modelo. Si el modelo no es lineal, entonces, el estimador de MCO de los coeficientes

$$b=(X'X)^{-1}X'y=(X'X)^{-1}X'(G(X,\beta)+\varepsilon)=$$
$$=(X'X)^{-1}X'G(X,\beta)+(X'X)^{-1}X'\varepsilon)\neq\beta+(X'X)^{-1}X'\varepsilon$$

- Deja de verificar la insesgadez y la consistencia. Por tanto, el estimador de MCO deja de ser válido, puesto que la relación ya no es lineal y pierde las principales propiedades que tenía.

Instrumentos de diagnostico de la linealidad

Universidade de Vigo

- Gráficos
  - Residuos respecto a valores estimados o el tiempo u otras variables que se sospechen que influyen en la forma de la regresión.
- Test de hipótesis
  - Test RESET

Modelo de contraste en el Test RESET de linealidad

Universidade de Vigo

- Definimos para cada potencia p un conjunto de variables independientes

$$W_t^{(p)} = \{\hat{Y}_t^2, \hat{Y}_t^3, \ldots, \hat{Y}_t^p\} \text{ , t=1...T}$$

- Cada uno de eso conjuntos va a servirnos como base para contrastar la no linealidad del modelo.
- El modelo de contraste para los test tiene la forma general siguiente:

$$y_t = X_t \beta + W_t^{(p)} \gamma + \varepsilon_t \quad t = 1 \ldots T$$

- Con las perturbaciones verificando todas las suposiciones del modelo de regresión lineal normal.

Hipótesis del test de linealidad

Universidade de Vigo

- H_0: $\gamma=0$, lo que significa que el error es ruido blanco
- H_1: $\gamma\neq0$, lo que significa que existe alguna variable independiente elevada a una potencia de orden superior al de la lineal influye en la variable dependiente

- El número de parámetros que intervienen en la ecuación anterior son k+1 en el β y p-1 en el γ.
- Las hipótesis que vamos a contrastar es que todos los coeficientes γ son nulos pues de esa forma el modelo que nos queda es el MRLN clásico
- Si alguno de los elementos de γ son diferentes de cero, eso indica que alguna potencia de los valores estimados de la y intervienen en el modelo y por lo tanto una combinación lineal de potencias de las X, es decir, la linealidad no se verifica.

Estadístico para el Test RESET

Universidade de Vigo

- Regresamos la variable dependiente respecto al conjunto de variables independientes y a las variables del conjunto formado por W(p)
- Contrastamos mediante test tipo LM si algún coeficiente de las nuevas variables es significativo. Para ello se compara el R^2 de la regresión inicial con el de esta nueva regresión.
- El método de comparación consiste en un cociente corregido por los grados de libertad entre ambos R^2.

$$FR = \frac{(R^2 - R_0^2) \big/ p-1}{(1-R^2) \big/ T-k-p} = \frac{(SCE_F - SCE_0) \big/ p-1}{(SCE_F) \big/ T-k-p}$$

Ley de distribución y regla de decisión en el Test RESET de linealidad

Universidade de Vigo

- Ese estadístico sigue, bajo normalidad una F con p-1 y T-k-p grados de libertad, siendo p la potencia máxima que se incluye en la ecuación.
- La regla de decisión consistirá en rechazar si FR es mayor que el valor de las tablas $F_{p-1, T-k-p, \alpha}$ o si la cola de probabilidad es menor que el nivel de significación
- El test es robusto a la suposición de normalidad, puesto que si esta falla deja de ser la distribución una F de Snedecor, pero se puede aproximar por una chi cuadrado con p-1 grados de libertad. El test pasa a ser asintótico en vez de exacto.

Test de RESET de linealidad de orden 2

Universidade de Vigo

Contrasta la no linealidad, cuando se supone válido el MRLN.

$$E(e/\hat{Y}) = \gamma_2 \hat{Y}_t^2 + \gamma_3 \hat{Y}_t^3 + \ldots \gamma_k \hat{Y}_t^k + \ldots$$

Para cada potencia realiza un test diferente.
Para el caso de potencia de orden 2, las hipótesis a contrastar serán las siguientes:

- H_0: $\gamma_2=0$, lo que significa que el error es ruido blanco
- H_1: $\gamma_2 \neq 0$, lo que significa que existe alguna variable independiente de orden superior al de la lineal que influye en la variable dependiente

Ejemplo: CENSA

- El coste de fabricación de celulosa en una empresa (CENSA) depende de la cantidad de celulosa producida. Los datos se recogen en la tabla siguiente. Comprobar cual es la función que relaciona el coste con la producción y obtener una estimación de los costes fijos y los costes variables en esa empresa.
- Interesa comprobar si la función de los costes de la fabricación es lineal.

Regresión sobre CENSA

Universidade de Vigo

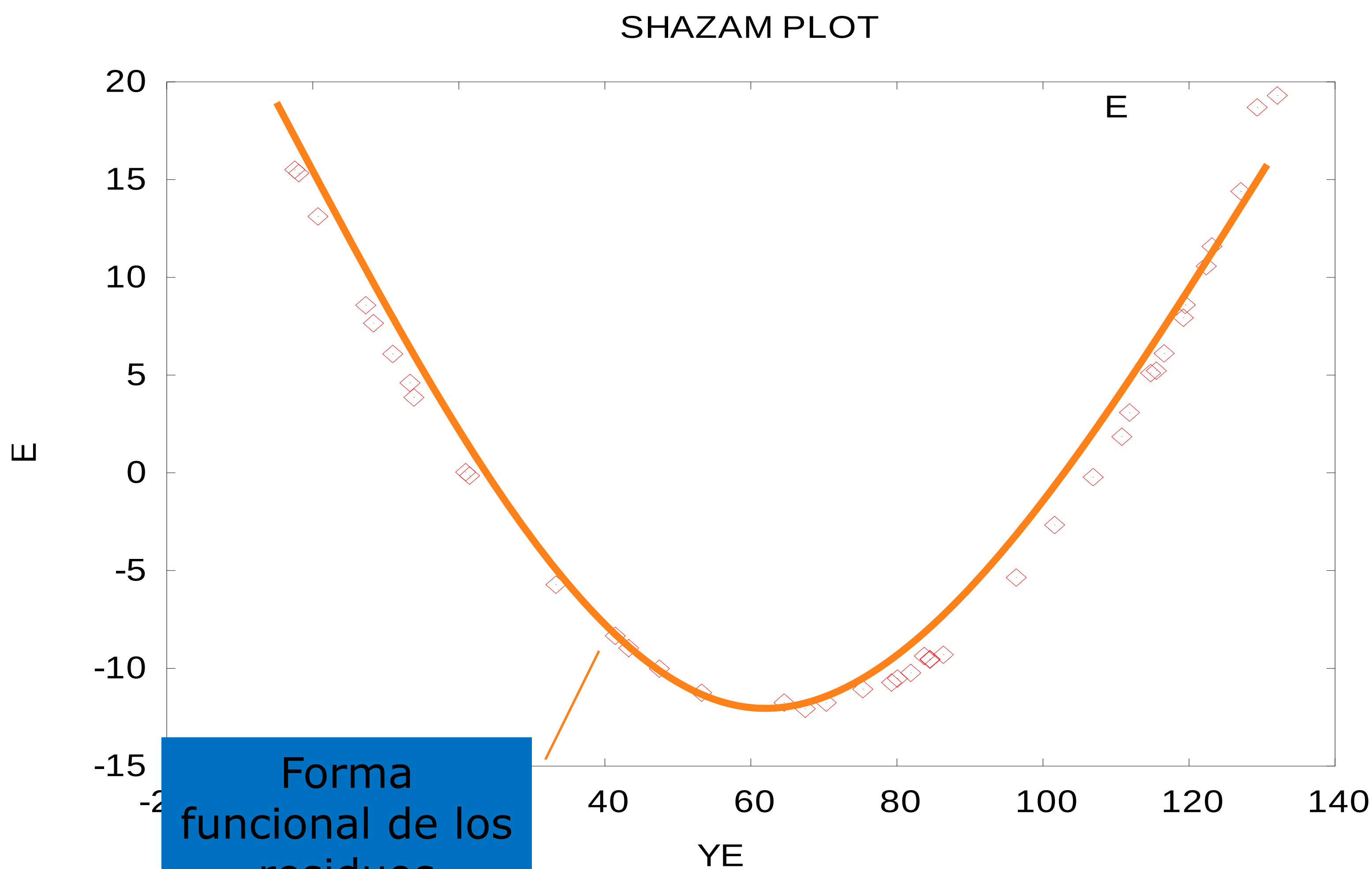
Modelo 1: MCO, usando las observaciones 1-40
Variable dependiente: Coste

	Coeficiente	*Desv. Típica*	*Estadístico t*	*Valor p*	
const	−33.8947	4.03859	−8.3927	<0.0001	***
celulosa	14.2686	0.514496	27.7331	<0.0001	***

Media de la vble. dep.	69.66747	D.T. de la vble. dep.	44.25266
Suma de cuad. residuos	3595.734	D.T. de la regresión	9.727516
R-cuadrado	0.952919	R-cuadrado corregido	0.951680
F(1, 38)	769.1226	Valor p (de F)	8.00e-27
Log-verosimilitud	−146.7300	Criterio de Akaike	297.4600
Criterio de Schwarz	300.8378	Crit. de Hannan-Quinn	298.6813

Grafico de residuos

Universidade de Vigo



Forma funcional de los residuos

Test RESET para CENSA

Regresión auxiliar para el contraste de especificación RESET
MCO, usando las observaciones 1-40
Variable dependiente: Coste

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	5.63668	0.290604	19.40	1.20e-020	***
celulosa	3.39370	0.0946079	35.87	8.98e-030	***
yhat^2	0.00521728	0.000124567	41.88	3.86e-032	***
yhat^3	6.58684e-06	6.30438e-07	10.45	1.90e-012	***

Estadístico de contraste: F = 52179.254074,
con valor p = P(F(2,36) > 52179.3) = 4.76e-063

Estadístico del test

Cola de probabilidad

Soluciones a la no linealidad

Universidade de Vigo

- La forma mas habitual de tratar la no linealidad consiste en buscar transformaciones en las variables que linealicen la relación en los parámetros. Normalmente estas transformaciones corrigen de modo conjunto la falta de linealidad y de normalidad.
- Si hay mas de una variable el proceso se hace paso a paso con las variables. Se empieza con la mas influyente en el fallo y una vez corregido se pasa a las demás.
- La forma mas habitual es hacer uso de las transformaciones más básicas basadas en los logaritmos
- En caso de que en ese caso no se linealice se puede recurrir a otras transformaciones para relacionar cada variable independiente con la dependiente.

Modelo linealizado

- Supongamos que la relación entre las variables tenga la forma siguiente:

$$h(y_t) = \beta_0 + \beta_1 h_1(X_{1t}) + \ldots + \beta_k h_1(X_{kt}) + \varepsilon_t \quad t = 1 \ldots T$$

Siendo h, $h_1, \ldots, h_k$ cualesquiera funciones de transformación.

- Hacemos los siguientes cambios de variables:

$$v_t = h(y_t)$$
$$Z_{1t} = h_1(X_{1t})$$
$$\vdots$$
$$Z_{kt} = h_1(X_{kt})$$

- El modelo nos quedaría:

$$v_t = \beta_0 + \beta_1 Z_{1t} + \ldots + \beta_k Z_{kt} + \varepsilon_t \quad t = 1 \ldots T$$

- Es decir, un MRLC, pero con otras variables

Transformaciones asociadas con los logaritmos

Universidade de Vigo

- En las transformaciones asociadas con logaritmos se consideran tres tipos de transformaciones
  - LOGLIN
    - logaritmo de la dependiente solo
  - LINLOG
    - logaritmo de las independientes solo
  - LOGLOG
    - logaritmo de la dependiente y de las independientes

Transformación LOGLIN

Universidade de Vigo

- Se toma el logaritmo de la variable dependiente y se dejan las independientes como estaban.
- El modelo es el siguiente:

$$\log y_t = \beta_0 + \beta_1 X_{1t} + \ldots + \beta_k X_{kt} + \varepsilon_t \qquad t = 1 \ldots T$$

- Interpretación de los parámetros:
  - La exponencial de la constante mide el valor promedio de la dependiente cuando las independientes son nulas.
  - Los coeficientes miden la tasa de crecimiento de la dependiente por cada unidad de las independientes

Transformación LINLOG

- Se toma el logaritmo de las variables independientes y se deja la dependiente como estaba.
- El modelo es el siguiente:

$$y_t = \beta_0 + \beta_1 \log X_{1t} + \ldots + \beta_k \log X_{kt} + \varepsilon_t \qquad t = 1 \ldots T$$

- Interpretación de los parámetros:
  - La constante mide el valor promedio de la dependiente cuando las independientes son nulas.
  - Los coeficientes miden el cambio de la dependiente por cada unidad porcentual que cambian las independientes

Transformación LOGLOG

Universidade de Vigo

- Se toma el logaritmo de la variable dependiente y de las independientes
- El modelo es el siguiente:

$$\log y_t = \beta_0 + \beta_1 \log X_{1t} + \ldots + \beta_k \log X_{kt} + \varepsilon_t \quad t = 1 \ldots T$$

- Interpretación de los parámetros:
  - La exponencial de la constante mide el valor promedio de la dependiente cuando las independientes son nulas.
  - Los coeficientes miden la elasticidad de la dependiente por cada unidad de las independientes
- Ejemplo: las funciones de Cobb-Douglas

Caso celulosa

Universidade de Vigo

```
Regresión auxiliar para el contraste de especificación RESET
Variable dependiente: Coste
              Coeficiente   Desv. Típica  Estadístico t   Valor p
  -------------------------------------------------------------
  const       -0.0228987    2.34114        -0.009781    0.9923
  l_celulosa  16.4556       2.24271         7.337       1.19e-08  ***
  yhat^2      -0.00453083   0.000675128    -6.711       7.88e-08  ***
  yhat^3       0.000109287  4.14097e-06    26.39        3.78e-025 ***
Estadístico de contraste: F = 3977.491628,
con valor p  = P(F(2,36) > 3977.49) = 5.84e-043

Regresión auxiliar para el contraste de especificación RESET
Variable dependiente: l_Coste
             Coeficiente   Desv. Típica   Estadístico t   Valor p
  -------------------------------------------------------------
  const       0.922158      0.207406         4.446        8.05e-05 ***
  celulosa    0.0750297     0.0681134        1.102        0.2780
  yhat^2      0.306705      0.0671480        4.568        5.59e-05 ***
  yhat^3     -0.0360566     0.00574684      -6.274        2.99e-07 ***

Estadístico de contraste: F = 575.514389,
con valor p  = P(F(2,36) > 575.514) = 4.71e-028

Regresión auxiliar para el contraste de especificación RESET
Variable dependiente: l_Coste
             Coeficiente  Desv. Típica  Estadístico t   Valor p
  -------------------------------------------------------------
  const        2.21604      0.0105134       210.8       2.95e-057 ***
  l_celulosa  -3.01497      0.140302       -21.49       4.01e-022 ***
  yhat^2       0.599462     0.0253478       23.65       1.59e-023 ***
  yhat^3      -0.0350832    0.00231147     -15.18       3.20e-017 ***

Estadístico de contraste: F = 8557.442761,
con valor p  = P(F(2,36) > 8557.44) = 6.26e-049
```

En este caso ninguna verifica la linealidad. Si hubiera mas de una se tomaría la transformación con mayor R2

Ejercicio 3.6

Estabilidad de parámetros

Análisis de los cambios estructurales

Universidade
de Vigo

Estabilidad de parámetros

- Queremos comprobar si los parámetros se mantienen los mismos a lo largo de toda la muestra o cambian de un valor a otro.
- En general el hecho de cambio de parámetros lleva implícito que existe otra variable independiente que condiciona el modelo y los parámetros cambian de acuerdo a esa variable.
- Normalmente el cambio se observa en el tiempo, por ello se habla de cambio estructural. Trabajaremos como si el tiempo fuera la variable que indica el cambio (es decir, el orden en que se observan los datos).
- Si fuera otra variable se reordenarían los datos de acuerdo a esa nueva variable. Por ejemplo si tenemos dos grupos de datos y queremos ver si tienen los mismos parámetros, pondríamos primero un grupo y luego otro.

Ejemplos de cambio de parámetros

- Existen muchos motivos por el que pueden cambiar los parámetros en un modelo:
  - Intervenciones externas en un momento de tiempo, por ejemplo anuncio de una subida del precio del petróleo, cambia las estructuras que relacionan variables fundamentales, renta consumo o inflación, o costes producción etc.. . El hundimiento del Prestige.
  - Existen dos grupos que tiene comportamientos diferentes , por ejemplo dos sectores industriales con relaciones de costes diferentes.
  - Modelos con parámetros cambiantes, cuando estos cambian en cada momento del tiempo.

Efecto de la no estabilidad

Universidade de Vigo

- El hecho de que no haya estabilidad en los parámetros genera el mismo efecto que la falta de linealidad, pues el modelo esta mal especificado y las estimaciones son sesgadas e inconsistentes.
- Eso significa que el método de estimación deja de ser válido, por ese motivo detectar la inestabilidad de parámetros es fundamental.

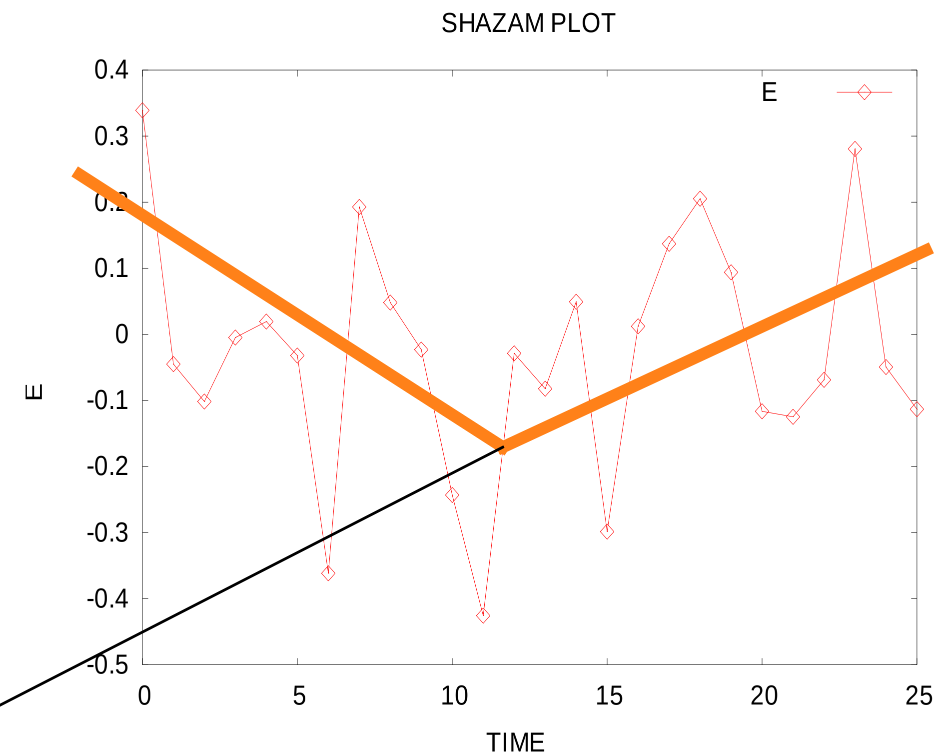
Diagnostico de la estabilidad

- Gráficos
  - Residuos respecto a valores estimados de la variable dependiente
  - Residuos respecto a variable de cambio estructural (tiempo u otra variable)
- Test de hipótesis
  - Test de Chow

Gráficos de residuos respecto a variable estimada

Universidade de Vigo

- Se representa el residuo respecto al tiempo u otra variable de la que se sospeche influye en el cambio de los parámetros.
- Si aparece un cambio de tendencia es síntoma de cambio estructural



Cambio de tendencia

Ejemplo CENSA

Universidade de Vigo

El coste de fabricación de celulosa en una empresa (CENSA) depende de la cantidad de celulosa producida. Comprobar cual es la función que relaciona el coste con la producción y obtener una estimación de los costes fijos y los costes variables en esa empresa. Los residuos se representan en los siguientes gráficos:

Grafico respecto al tiempo

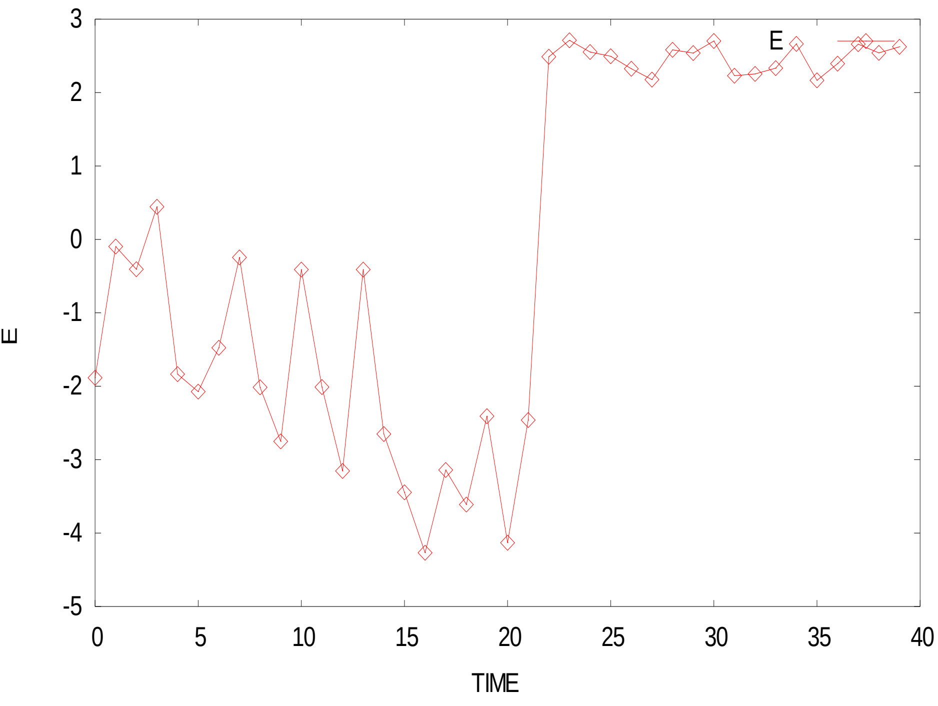
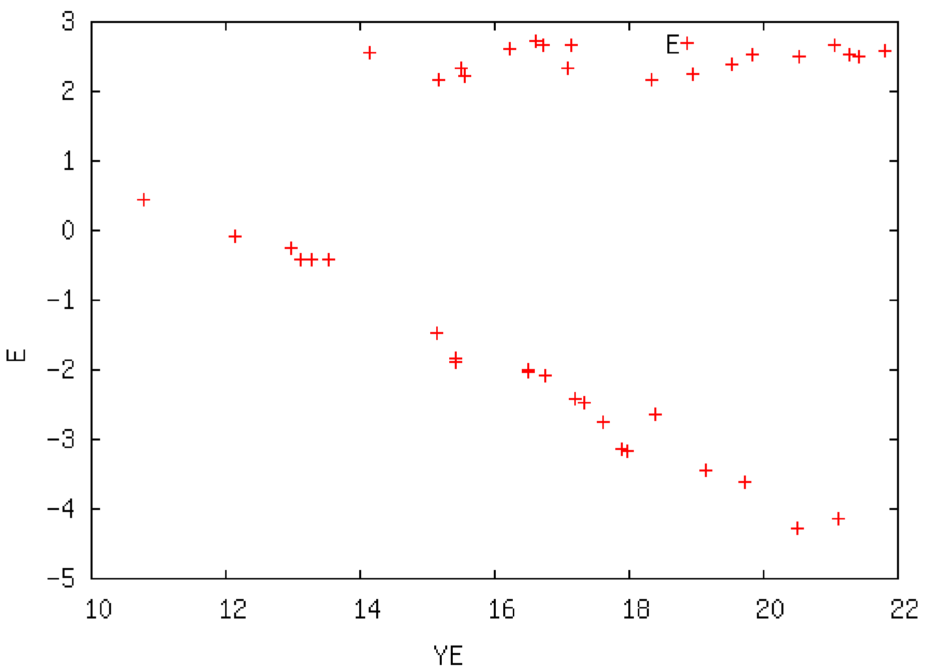


Grafico respecto al valores estimados de la dependiente



Test de estabilidad de Chow

Universidade de Vigo

- La idea intuitiva del test es muy sencilla. Se supone que existe una observación J a partir de la cual los parámetros son diferentes. Por ese motivo se debe ordenar la muestra de acuerdo a la variable que influya en el cambio.
- Inicialmente el test de Chow suponía únicamente dos casos, que delimitaban la hipótesis nula y alternativa del test:
  - Que en toda la muestra los parámetros fueran iguales.
  - A partir de la observación los parámetros son diferentes
- Si no hay cambio estructural de parámetros, ambos modelos deben explicar lo mismo y por consiguiente las sumas de cuadrados del primero caso y del segundo deberían ser similares. Si fueran muy distintos es que los parámetros son diferentes.

Suposiciones del test de estabilidad

Universidade de Vigo

- Como queremos comparar dos modelos alternativos : con parámetros iguales o diferentes, nos ponemos en el caso mas desfavorable que consiste en suponer que ambos son diferentes.
- Por consiguiente el modelo general tendría la forma siguiente:

$$y_t = X'_t \beta + \epsilon_t \quad \text{t=1…J}$$

$$y_t = X'_t \gamma + \epsilon_t \quad \text{t=J+1…T}$$

- Verificando las perturbaciones todas las suposiciones del MRLC

Test de Diferencia en todos los coeficientes (Test de Chow)

Universidade de Vigo

- Llamando β a los parámetros del modelo en las primeras J observaciones y γ al del modelo en las ultimas T-J observaciones tendremos que las hipótesis a contrastar serán:

$$H_0 : \beta = \gamma$$
$$H_1 : \beta \neq \gamma$$

- Cuando los parámetros son iguales tendremos un único vector de parámetros β, en otro caso tendremos dos diferentes. Eso significa que bajo H_0 hay estabilidad
- Fijamos el nivel de significación α

Idea del Estadístico del test de estabilidad

Universidade de Vigo

- Vimos que se consideraban dos casos, que delimitaban la hipótesis nula y alternativa del test:
  - Que en toda la muestra los parámetros fueran iguales.
    - En ese caso se estima el modelo único para todos los datos
    - se calcula la suma de cuadrados de los errores, SCE_0
  - A partir de la observación los parámetros son diferentes
    - Se estiman dos modelos diferentes; uno para las primeras J observaciones, antes del cambio y otro para los últimos datos, después del cambio sospechado.
    - Se calculan las sumas de cuadrados de cada modelo: SCE1 para el primero y SCE2 para el segundo. Luego la suma de ambos (SCE_C)que nos dará el grado de error en el modelo conjunto

Cálculo del Estadístico del test de estabilidad

Universidade de Vigo

- Para evaluar la mejora del modelo dividido respecto al modelo con parámetros iguales calculamos las diferencias entre ambas sumas de cuadrados que nos dirá cuanto se disminuye el error al dividir los parámetros en dos.
- La diferencia promedio la estandarizamos respecto al estimador de la varianza residual en el modelo conjunto.
- Si k es el número de variables independientes y T el numero de observaciones, el estadístico sería, por tanto

$$F = \frac{\dfrac{SCE_0 - SCE_C}{k+1}}{\dfrac{SCE_C}{T-2k-2}}$$

Ley de distribución y regla de decisión del test de estabilidad

Universidade de Vigo

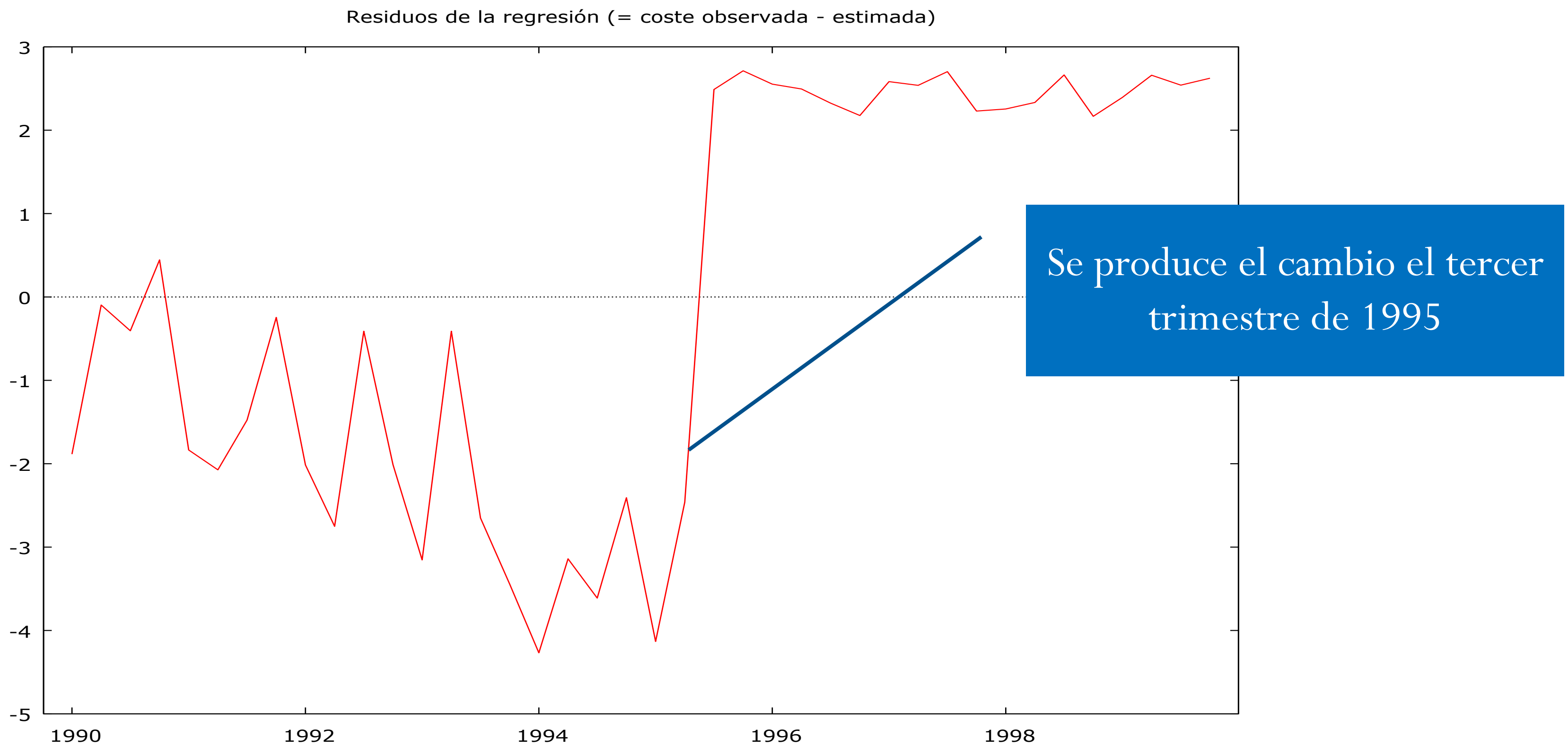
- **Ley de distribución**
  - Sigue una F de Snedecor con k+1 y T-2k-2 grados de libertad respectivamente
- **Regla de decisión:**
  - Se rechaza la hipótesis nula si el valor del estadístico es mayor que el valor de las tablas de la F con k+1 y T-2k-2 grados de libertad al nivel α.

Test de Chow de Censa

Universidade de Vigo

Modelo 1: MCO, usando las observaciones 1990:1-1999:4 (T = 40)
Variable dependiente: coste

	Coeficiente	*Desv. Típica*	*Estadístico t*	*Valor p*	
const	8.30376	1.35676	6.1203	<0.0001	***
cantidad	0.942491	0.139413	6.7604	<0.0001	***



Salida del test deChow

Universidade de Vigo

```
Regresión aumentada para el contraste de Chow
MCO, usando las observaciones 1990:1-1999:4 (T = 40)
Variable dependiente: coste

               Coeficiente  Desv. Típica  Estadístico t   Valor p
  ---------------------------------------------------------------
  const         10.1438      0.130937        77.47       1.20e-041 ***
  cantidad       0.486983    0.0146252       33.30       1.21e-028 ***
  splitdum       0.497007    0.242489         2.050      0.0477    **
  sd_cantidad    0.468259    0.0241694       19.37       1.25e-020 ***

Media de la vble. dep.  17.07562   D.T. de la vble. dep.   3.673949
Suma de cuad. residuos  1.413678   D.T. de la regresión    0.198164
R-cuadrado              0.997315   R-cuadrado corregido    0.997091
F(3, 36)                4456.498   Valor p (de F)          2.57e-46
Log-verosimilitud       10.09615   Criterio de Akaike    -12.19231
Criterio de Schwarz    -5.436792   Crit. de Hannan-Quinn  -9.749728
rho                     0.049066   Durbin-Watson           1.822978

Contraste de Chow de cambio estructural en la observación 1995:3
  F(2, 36) = 3024.93 con valor p  0.0000
```

Ejercicio 3.7

El efecto del tiempo

La dependencia en modelos de regresión lineal

Modelos autocorrelados: El efecto del pasado de forma lineal

Universidade
de Vigo

Carlos M. Jardon

La independencia

- El hecho de suponer que las observaciones son independientes nos permite analizar las perturbaciones como si fueran un conjunto de variables igualmente distribuidas y, por lo tanto, nos permite suponer que cada observación agota la información que esa observación puede dar sobre la variable dependiente.
- Este hecho nos permite suponer que todas las perturbaciones provienen de la misma población y, por consiguiente analizar los residuos como si todos ellos fueran observaciones de una misma variable
- En caso contrario existe dependencia. Eso implicaría que la matriz de varianzas covarianzas ya no va ser escalar, es decir, la identidad por un escalar, sino que los elementos de fuera de la diagonal pueden ser distintos de 0.

Ejemplos de dependencia

- Se quiere estudiar la rentabilidad en función del numero de clientes en una red de sucursales de banca, cuyos sistemas de organización están relacionados y es posible que las perturbaciones entre oficinas estén relacionadas. (correlación)
- Estudio de las ventas de una compañía en diferentes zonas que posiblemente esté afectado por al distancia a los centros de venta (correlación espacial)
- Evolución de las ventas de una empresa en función de su publicidad y sus ventas pasadas. (correlación temporal)

Modelos autocorrelados

Universidade de Vigo

- Aunque la dependencia se puede dar por múltiples causas, lo mas habitual es que tenga que ver con el tiempo, pues existe en las variables económicas un efecto temporal consecuencia de “recordar” efectos del pasado.
- Por ese motivo, vamos a estudiar un caso particular, los modelos autocorrelados, que significan que existe influencia del pasado, y esta influencia se recoge mediante un efecto lineal.
- Normalmente, el efecto directo del pasado desaparece a partir de un determinado retardo m.

Modelo lineal con autocorrelación

- Existe dependencia del pasado en las perturbaciones, por consiguiente, estas no son independientes.
- El pasado tiene un efecto lineal:
  - es siempre constante.
  - independientemente del periodo en el que se da.
  - pero depende del retardo con el que se da.
- Esto, como veremos posteriormente, indica que existe influencia en cualquier pasado, si bien de una forma determinada.

Identificación de la autocorrelación

- Gráficos
  - Algunos coinciden con los ya definidos y sólo cambia su interpretación.
  - Otros son específicos de este análisis
- Tests de hipótesis
  - Son específicos de la hipótesis alternativa de autocorrelación
  - Pueden ser de orden 1 o de orden mayor que 1.

Gráficos de residuos

Universidade de Vigo

- Se debe recordar que los residuos deben verificar:
  - E(e)=0;
    - los residuos deben oscilar alrededor del 0, no deben presentar ninguna forma funcional
  - Var(e)= constante aproximadamente.
    - Los residuos deben dispersarse de forma homogénea, estar comprendidos entre dos bandas
  - Independencia:
    - Los residuos no deben presentar relaciones entre ellos, formas graficas encadenadas.
  - Normalidad
    - Los residuos deben estar cercanos al 0, simétricos a cada lado y las bandas no muy alejadas del 0 (alrededor de dos veces la desviación estándar).
- Eso significa que cualquier forma en que se aleje de esas características existe un fallo de alguna suposición.
- El residuo se introduce en el eje de ordenadas. La variable introducida en el eje de abscisas nos dice algo sobre ese fallo.

Residuos OLS en modelos autocorrelados

Universidade de Vigo

- Para desarrollar los diferentes gráficos y test de hipótesis de autocorrelación de orden 1 es conveniente analizar el comportamiento de los residuos OLS bajo autocorrelación.

$$e_t = r e_{t-1} + u_t$$

- La esperanza del residuo de orden t condicionada al de orden t-1 depende del signo de r

$$E(e_t / e_{t-1}) = r e_{t-1}$$

- Si r es negativo el residuo se espera con el signo diferente del anterior, mientras que si es positivo el signo es igual.
- Por lo tanto

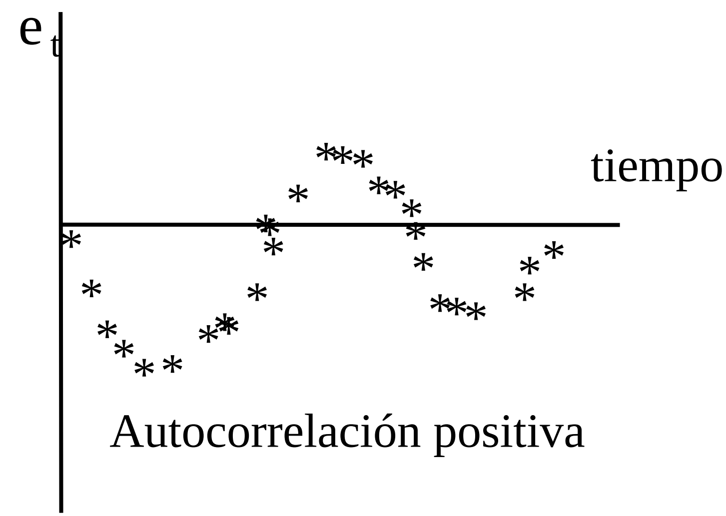
$$|E(e_t / e_{t-1}) - e_{t-1}| = |r-1|$$

- Si r es negativo la diferencia tiende a ser mayor que 1, mientras que si es positivo tiende a ser cercano a cero. Por tanto las distancias entre dos residuos consecutivos se incrementará cuando ambos tienen diferente signo
- Cambiando el retardo se generalizan los modelos a ordenes superiores

Gráficos de residuos respecto al tiempo

Universidade de Vigo

- Representamos los residuos respecto al tiempo.
- Los comportamientos de dependencia respecto al tiempo se reflejan en la evolución de los residuos:
  - Cuando existe autocorrelación positiva los residuos cortan pocas veces el eje.
  - Cuando es negativa lo cortan muchas.
  - En caso de no correlación deberían comportarse como un ruido blanco.



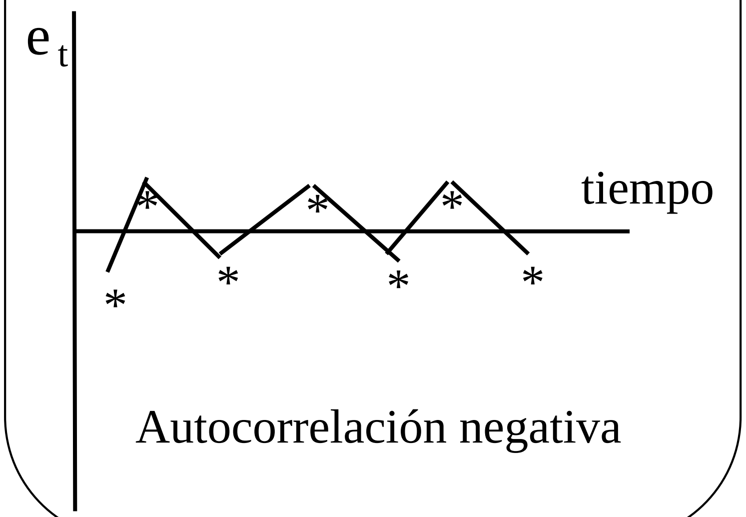
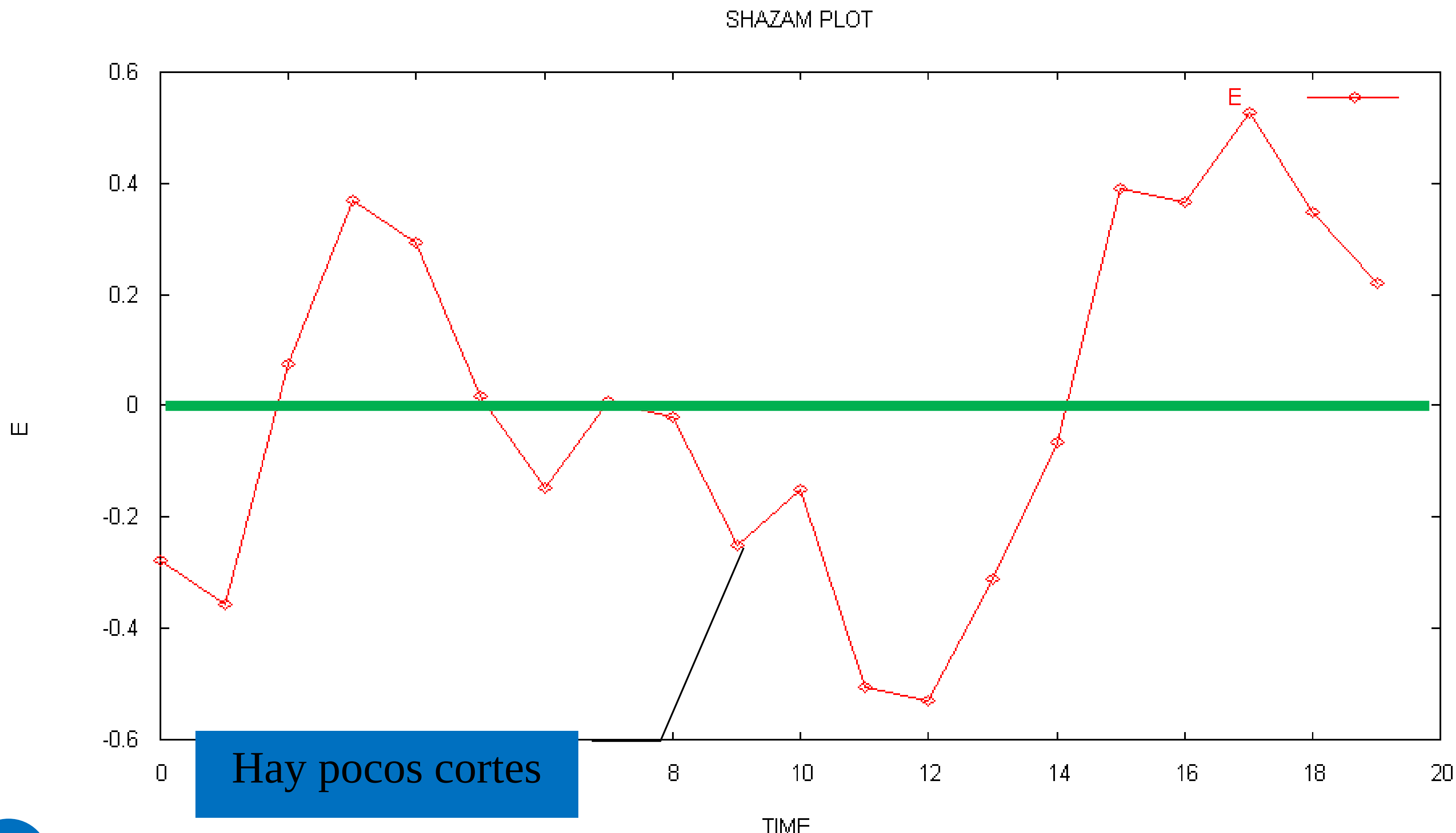


Gráfico de residuos respecto al tiempo

Universidade de Vigo



Gráficos de residuos respecto a valores retardados

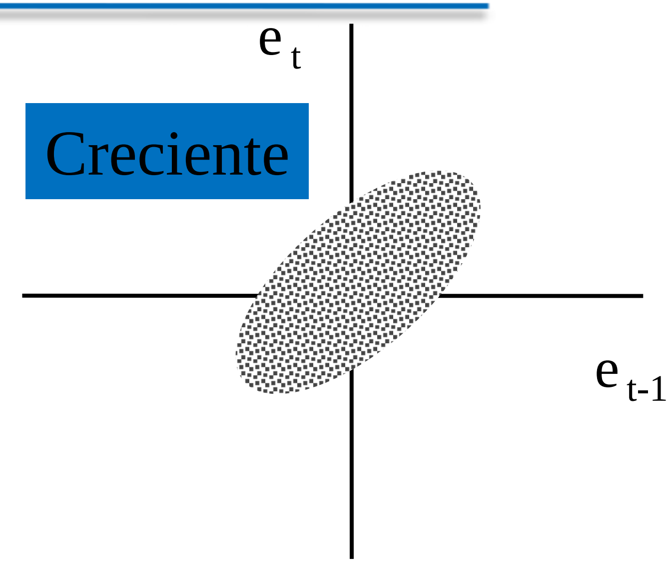
Universidade de Vigo

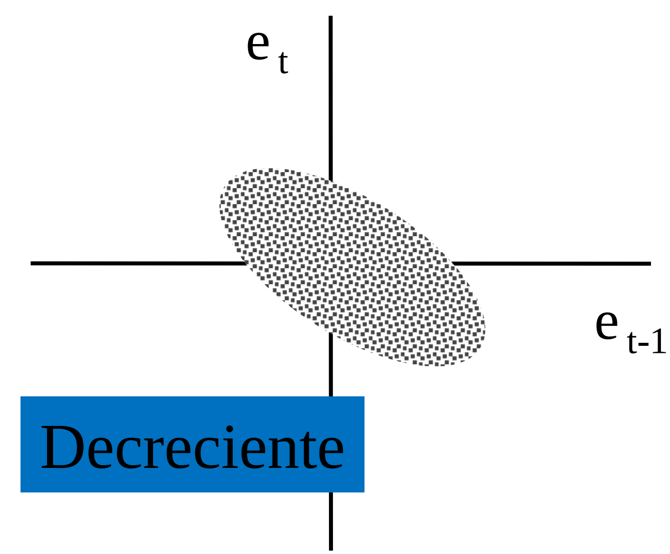
- Representamos los residuos respecto a sus valores retardados.
- Depende del orden que se quiera estudiar para hacer uso de él en la representación.
- Lo normal será estudiar autocorrelación de primer orden por lo que haremos uso de los residuos retardados una vez.
- Para órdenes superiores al primero se haría de la misma forma haciendo uso de residuos de orden superiores

Gráficos de residuos respecto a valores retardados

Universidade de Vigo

- Al estudiar autocorrelación de primer orden debería tener una forma similar a la lineal.
- Si la autocorrelación es positiva dicha forma será creciente
- Si la autocorrelación es negativa dicha forma será decreciente.
- En caso de no correlación deberían comportarse como un ruido blanco.

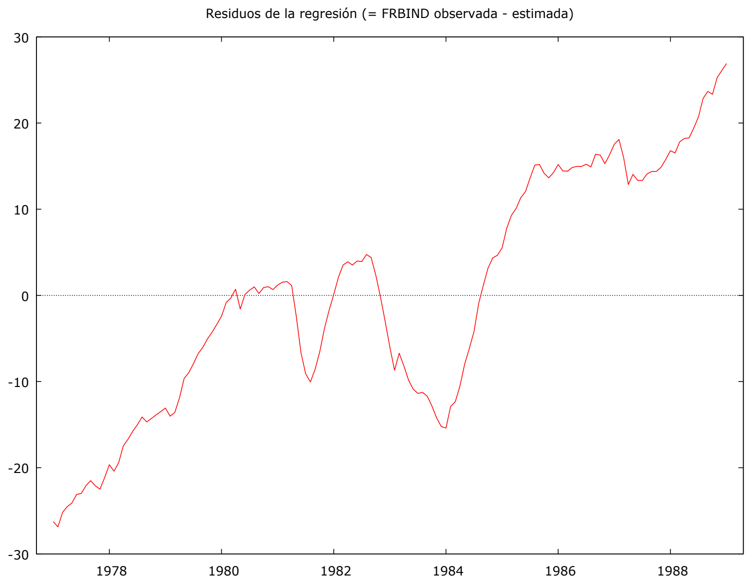


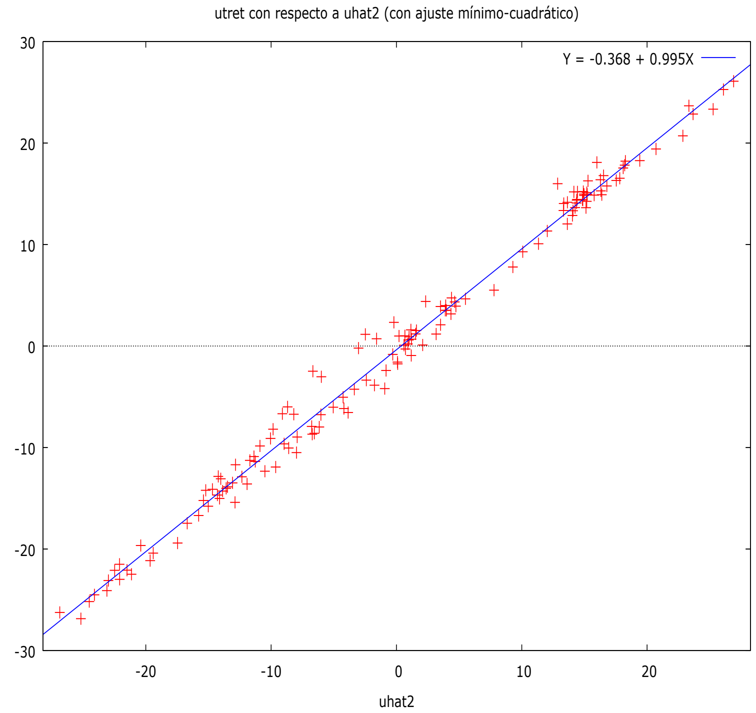


Gráficos en Gretl: Ejemplos Ross

Universidade de Vigo

```
Modelo 2: MCO, usando las observaciones 1977:01-1989:01 (T = 145)
Variable dependiente: FRBIND

            Coeficiente   Desv. Típica   Estadístico t    Valor p
  ---------------------------------------------------------------
  const       155.348        3.33596        46.57       2.66e-088 ***
  POILR       -54.5671      45.4118         -1.202      0.2315
```





Ejercicio 3.7

Tests de hipótesis para la identificación de la autocorrelación

- Para identificar la autocorrelación en un modelo, se parte de suponer que ésta no existe.
- Se realiza la estimación como si no la hubiese
- Se analizan los residuos para ver si el comportamiento que presentan es el esperado bajo no correlación, o aparecen situaciones inesperadas que, en algunos casos serán síntomas de autocorrelación.
- La hipótesis nula, en todos ellos va a ser la de no existencia de autocorrelación.
- La alternativa cambia según se sospeche autocorrelación de primer orden o de orden superior.

Test de autocorrelación

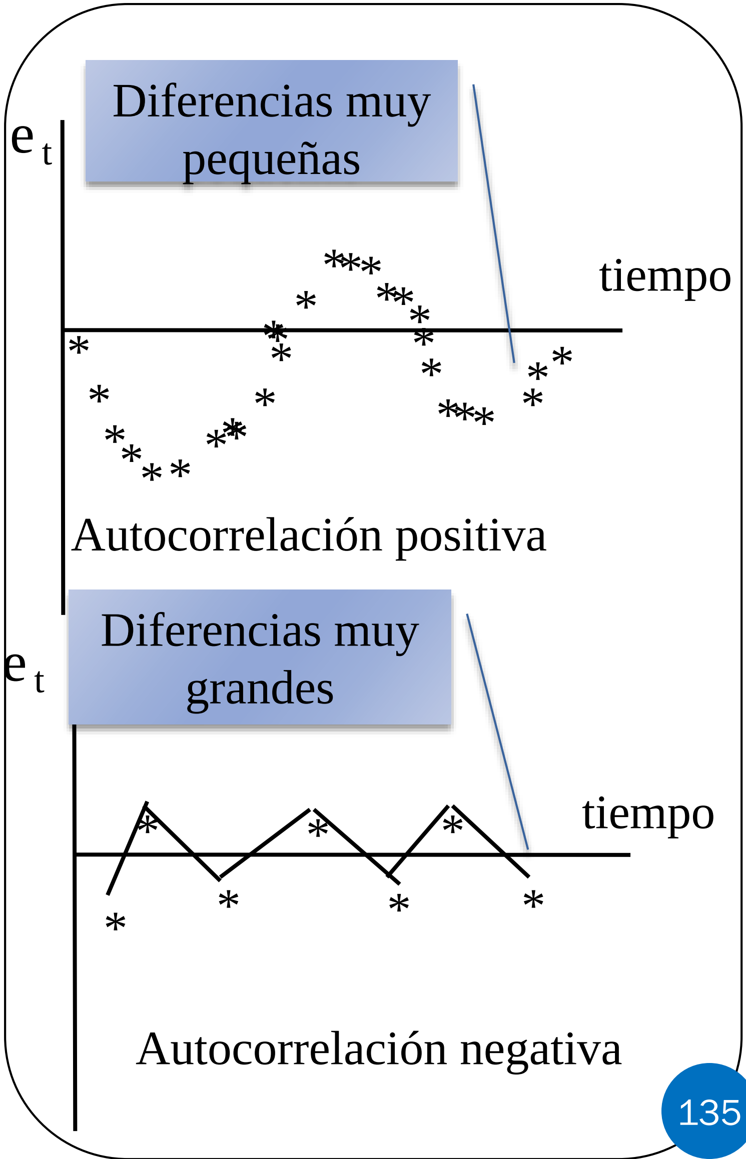
Universidade de Vigo

- Los test de autocorrelación de orden 1 se consideran todos ellos Individuales, pues únicamente se contrasta un parámetro
- Pueden ser
  - **Test de Durbin-Watson**
  - **Test de Wald**
  - **Test LM de Box Pierce**

Idea del Test de Durbin-Watson

Universidade de Vigo

- La idea de este test es comprobar si las diferencias entre los residuos observados consecutivamente es muy grande o muy pequeña.
- Si es muy grande es síntoma de que cambia de signo, y entonces la autocorrelación es negativa.
- Si es muy pequeña es que están muy seguidos y normalmente no cambiará de signo por lo que la relación entre ellos es positiva



Estadístico del Test de Durbin-Watson

Universidade de Vigo

- Bajo las suposiciones del MRLC, que es bajo la hipótesis nula, se construye el estadístico

$$DW = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

- Se puede observar como dicho estadístico se obtiene como diferencia entre dos residuos consecutivos al cuadrado, normalizado por la suma de todos ellos al cuadrado.
- Intuitivamente, ese estadístico nos dice que si un residuo está muy cercano al anterior entonces la correlación es positiva, ya que la diferencia será cercana a 0.
- Si está muy alejado, significa que irá pegando saltos de un lado al otro del 0 y por tanto la correlación en negativa.
- Cuando es un número intermedio el estimador sería aproximadamente el promedio entre 0 y el límite superior del estadístico.

Relación entre la autocorrelacion y el estadístico de Durbin-Watson

Universidade de Vigo

- A la vista del estadístico anterior, se demuestra que es aproximadamente igual a

$$DW \approx 2(1-r)$$

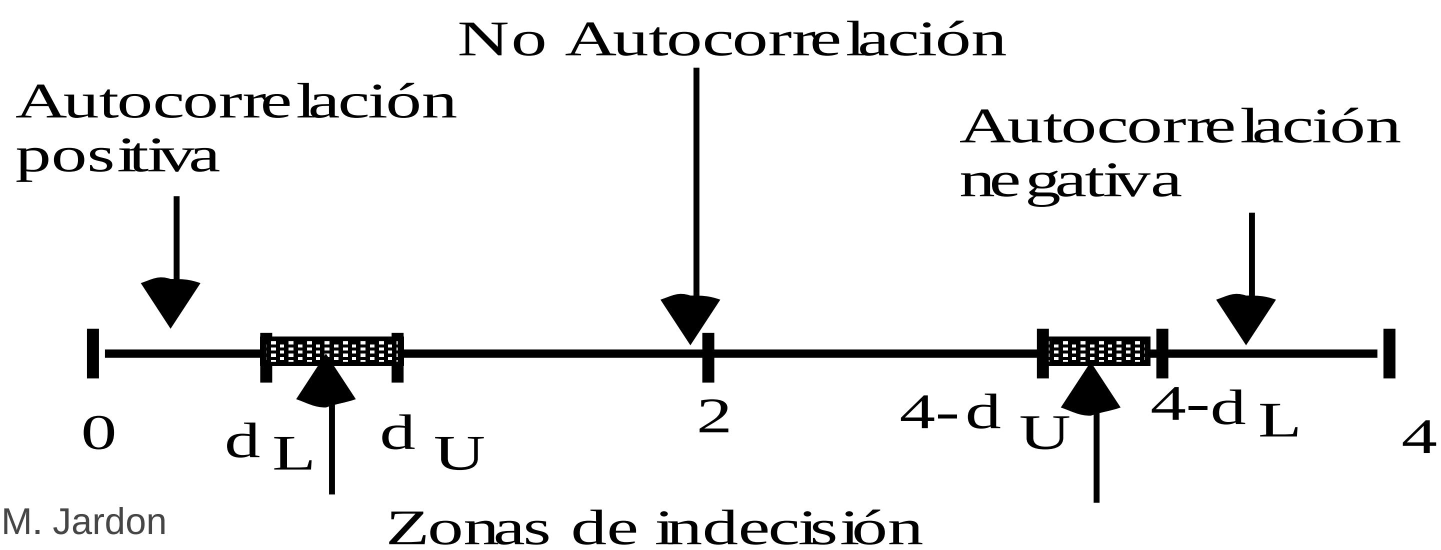
Siendo r un estimador de ρ

- Cuando DW es pequeño será cercano a 0, pues r es aproximadamente 1.
- Cuando DW es grande será cercano a 4 pues r es aproximadamente –1.
- Cuando es un número intermedio el estimador sería aproximadamente 2, que es el promedio entre 0 y 4.

Regla de decisión en el test de Durbin-Watson

Universidade de Vigo

- El estadístico DW sigue una ley tabulada por D-W.
- Dicha ley tiene una particularidad que hay una zona donde el test no permite decidir
- Las decisiones pueden ser de indecisión, según el siguiente gráfico



Valores críticos del contraste de Durbin-Watson

						k k =					
		k =	1	k =	2		3	k =	4	k =	6
n	®	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
	.01	.81	1.07	.7	1.25	.59	1.46	.49	1.7	.39	1.96
15	.025	.95	1.23	.83	1.4	.71	1.61	.59	1.84	.48	2.09
	.05	1.08	1.36	.95	1.54	.82	1.75	.69	1.97	.56	2.21
	.01	.95	1.15	.86	1.27	.77	1.41	.63	1.57	.6	1.74
20	.025	1.08	1.28	.99	1.41	.89	1.55	.79	1.7	.7	1.87
	.05	1.2	1.41	1.1	1.54	1	1.68	.9	1.83	.79	1.99
	.01	1.05	1.21	.98	1.3	.9	1.41	.83	1.52	.75	1.65
25	.025	1.13	1.34	1.1	1.43	1.02	1.54	.94	1.65	.86	1.77
	.05	1.2	1.45	1.21	1.55	1.12	1.66	1.04	1.77	.95	1.89
	.01	1.13	1.26	1.07	1.34	1.01	1.42	.94	1.51	.88	1.61
30	.025	1.25	1.38	1.18	1.46	1.12	1.54	1.05	1.63	.98	1.73
	.05	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
	.01	1.25	1.34	1.2	1.4	1.15	1.46	1.1	1.52	1.05	1.58
40	.025	1.35	1.45	1.3	1.51	1.25	1.57	1.2	1.63	1.15	1.69
	.05	1.44	1.54	1.39	1.6	1.34	1.66	1.29	1.72	1.23	1.79
	.01	1.32	1.4	1.28	1.45	1.24	1.49	1.2	1.54	1.16	1.59
50	.025	1.42	1.5	1.38	1.54	1.34	1.59	1.3	1.64	1.26	1.69
	.05	1.5	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
	.01	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.6
60	.025	1.47	1.54	1.44	1.57	1.4	1.61	1.37	1.65	1.33	1.69
	.05	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
	.01	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.6	1.36	1.62
80	.025	1.54	1.59	1.52	1.62	1.49	1.65	1.47	1.67	1.44	1.7
	.05	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
	.01	1.52	1.56	1.5	1.58	1.48	1.6	1.45	1.63	1.44	1.65
100	.025	1.59	1.63	1.57	1.65	1.55	1.67	1.53	1.7	1.51	1.72
	.05	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

k representa el número de variables regresoras

Limitaciones del test de Durbin Watson

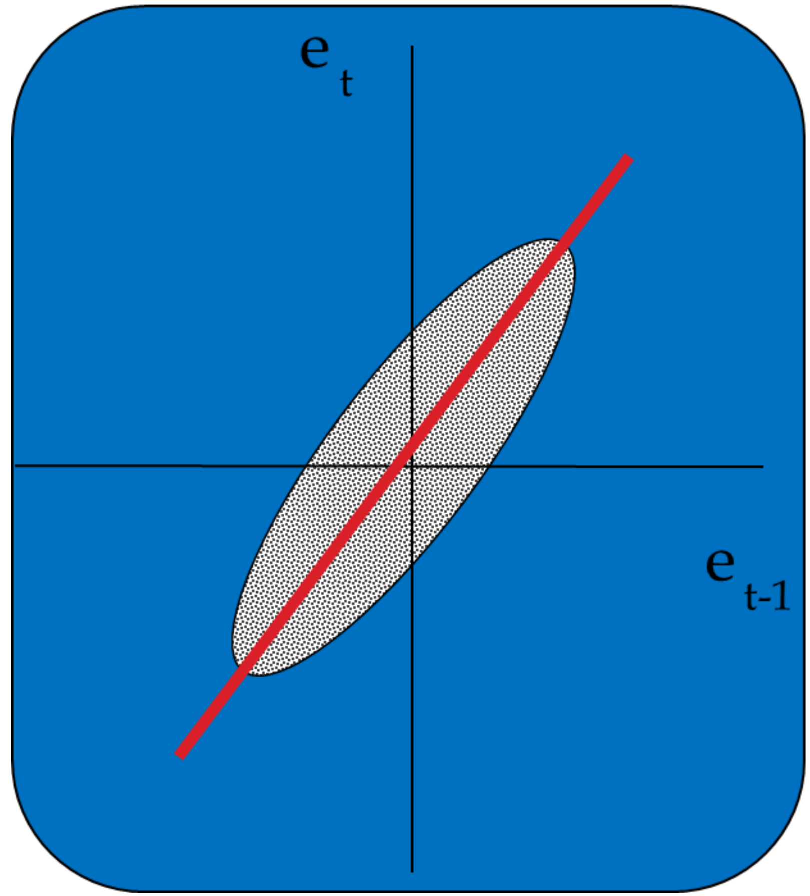
Universidade de Vigo

- Además de las suposiciones del modelo de regresión lineal normal, la ley de distribución del estadístico es válida si se verifican la siguientes condiciones:
  - los regresores son fijos
  - Si los regresores son estocásticos la ley de distribución es aproximada y la aproximación depende del grado de aleatoriedad.
  - las variables que intervienen siguen leyes normales
  - Si no existe dinámica en las variables independientes

Idea de los test generalizables

Universidade de Vigo

- La idea de estos test es comprobar que no existe relación lineal entre los residuos y los residuos retardados basándonos en el gráfico de residuos y residuos retardados.
- En dicho gráfico se observa si existe algún tipo de relación lineal entre los residuos y los residuos retardados, para comprobar si existe o no, se debe hacer una regresión de los residuos sobre los residuos retardados.
- Una vez hecho eso, se busca un estadístico que mida el grado de relación lineal existente entre las dos variables. Este es el que cambia dependiendo del test utilizado



Test de Wald

Universidade de Vigo

- El test de Wald es un test paramétrico que se realiza para contrastar la existencia de autocorrelación. Por tanto sigue el modelo explicitado previamente, y las hipótesis son:

  - **H_0: $\rho=0$, lo que significa incorrelación**
  - **H_1: $\rho\neq0$, lo que significa que existe autocorrelación lineal de orden 1**

- La idea de este test consiste en comprobar directamente si un estimador consistente de la autocorrelación se puede considerar nulo (hipótesis nula) o no. Al ser un estimador consistente, es decir que converge en probabilidad al parámetro en grandes muestras, se espera que dicho estimador esté cercano al parámetro y por consiguiente si este está cerca de cero el parámetro teórico también se puede suponer igual a cero.

Estadístico y regla de decisión del test de Wald

Universidade de Vigo

- Utilizamos como estimador del parámetro el coeficiente de correlación simple r.

$$r = \sum_{i=2}^{T} e_i e_{i-1} \Big/ \sum_{i=1}^{T} e_i^2$$

- Dicho coeficiente sigue bajo la hipótesis nula una ley de distribución asintóticamente normal de media 0 y varianza 1/T (AN(0, 1/T)), luego el estadístico como estadístico del test tomamos el valor tipificado de r que será asintóticamente N(0,1), es decir

$$t_W = \frac{r}{\sqrt{\frac{1}{T}}}$$

- La regla de decisión será rechazar la hipótesis nula si $|t_W| > \lambda_{\alpha/2}$

Extensión del test de Wald a orden superior

- Si se quiere contrastar la existencia de autocorrelación de orden superior al orden 1, el test se generaliza fácilmente, pero cambian las hipótesis a contrastar. Por ejemplo suponiendo un orden m cualquiera, el modelo ahora será el siguiente:

$$Y_t = X_t'\beta + \nu_t$$

$$\nu_t = \rho_m \nu_{t-m} + \varepsilon_t$$

$$\varepsilon \text{ sigue} \quad N(0, \sigma^2 I)$$

- Ahora las hipótesis a contrastar serían:

- H_0: $\rho_m = 0$, lo que significa incorrelación de orden m
- H_1: $\rho_m \neq 0$, lo que significa que existe autocorrelación lineal de orden m

Estadístico y regla de decisión del test de Wald de orden superior

Universidade de Vigo

- Siguiendo la misma idea intuitiva del caso de orden 1 y utilizando también de modo similar a ese caso como estimador del parámetro el coeficiente de correlación simple de orden m r_m.

$$r_m = \sum_{i=m+1}^{T} e_i e_{i-1} \Big/ \sum_{i=1}^{T} e_i^2$$

- Dicho coeficiente también sigue bajo la hipótesis nula una ley de distribución asintóticamente normal de media 0 y varianza 1/T, es decir, AN(0, 1/T), luego como estadístico del test tomamos el valor tipificado de r_m

$$t_W = \frac{r_m}{\sqrt{\frac{1}{T}}}$$

- Ese estadístico seguirá asintóticamente N(0,1)
- La regla de decisión será rechazar la hipótesis nula si $|t_w| > \lambda_{\alpha/2}$

Test de Wald para las telas

Universidade de Vigo

LAG	RHO	STD ERR	T-STAT
1	0.6521	0.1581	4.1244
2	0.3960	0.1581	2.5042
3	0.2816	0.1581	1.7810
4	0.2438	0.1581	1.5418
5	0.1434	0.1581	0.9072
6	0.0655	0.1581	0.4142
7	-0.1958	0.1581	-1.2385
8	-0.3726	0.1581	-2.3565
9	-0.3150	0.1581	-1.9923
10	-0.2053	0.1581	-1.2985
11	-0.1736	0.1581	-1.0978
12	-0.2099	0.1581	-1.3277

Se rechaza autocorrelación nula

Ejercicio 5.4

Test LM de autocorrelación

Universidade de Vigo

- El test LM de autocorrelación es un test paramétrico que se realiza para contrastar la existencia de autocorrelación de cualquier orden. Empezamos explicando el de orden 1, peor se generaliza fácilmente a ordenes superiores.
- Por tanto sigue el modelo explicitado previamente, y las hipótesis son:
  - **H_0: $\rho_1 = \ldots = \rho_\kappa = 0$, lo que significa incorrelación**
  - **H_1: existe un i =1…k/ $\rho_\iota \neq 0$, lo que significa que existe autocorrelación lineal de orden 1 o superior**
- Dado que es una regresión simple parece lógico hacer uso del coeficiente de determinación.
- Cuanto mayor sea este mas relación lineal existe independientemente del signo de la relación.

Estadístico del Test LM de autocorrelación

Universidade de Vigo

- Se calcula un estadístico estándar llamado LM que consiste en analizar el coeficiente de determinación de la regresión de residuos respecto a residuos retardados desde 1 hasta k periodos.
- Se demuestra que el estadístico LM=TR^2 sigue asintóticamente una χ^2 con 1 grado de libertad. Por consiguiente ese será el estadístico que elijamos.
- La regla de decisión será rechazar la hipótesis nula si $LM > \chi^2_\alpha$
- En la práctica se utiliza la raíz cuadrada de LM y se compara con una normal, pero el estadístico no indica el signo de la relación.

Test LM para telas

Universidade de Vigo

```
RESIDUAL CORRELOGRAM
  LM-TEST FOR HJ:RHO(J)=0, STATISTIC IS STANDARD NORMAL
  LAG     RHO      sample 2 40
ols e elag
  R-SQUARE =    0.4410     R-SQUARE ADJUSTED =    0.4259
 VARIANCE OF THE ESTIMATE-SIGMA**2 =   0.42912E-01
 STANDARD ERROR OF THE ESTIMATE-SIGMA =   0.20715
 SUM OF SQUARED ERRORS-SSE=   1.5878
 MEAN OF DEPENDENT VARIABLE =   0.72745E-02
 LOG OF THE LIKELIHOOD FUNCTION =   7.08551
 VARIABLE   ESTIMATED  STANDARD   T-RATIO        PARTIAL STANDARDIZED ELASTICITY
   NAME     COEFFICIENT   ERROR      37 DF    P-VALUE CORR. COEFFICIENT  AT MEANS
 ELAG       0.65675     0.1216       5.402     0.000 0.664     0.6640    -0.3159
 CONSTANT   0.95726E-02 0.3317E-01   0.2886     0.775 0.047     0.0000     1.3159
gen1 LM=$n*$r2
 ..NOTE..CURRENT VALUE OF $N    =    39.000
 ..NOTE..CURRENT VALUE OF $R2   =   0.44096
distrib lm / type=chi df=1
 CHI-SQUARE PARAMETERS- DF=   1.0000
 MEAN=   1.0000     VARIANCE=   2.0000     MODE=   0.0000
                    DATA         PDF          CDF        1-CDF
  LM          17.197     0.17734E-04 0.99997     0.33689E-04
```

Tratamiento de la autocorrelación

- Aunque los estimadores de MCO siguen siendo consistentes, pierden la eficiencia y por ello nos interesa buscar unos estimadores que sean mejores.
- Los estimadores de las varianzas son inconsistentes por lo que para evitar ese problema se pueden estimar las varianzas mediante estimadores robustos.
- El proceso genérico para modelos con heterocedasticidad y/o autocorrelación se denomina de mínimos cuadrados generalizados factibles (MCGF)

Ejercicio 3.8

Heterocedasticidad

Estimadores, test y posibles soluciones factibles.

Universidade
de Vigo

Concepto

- Un modelo heterocedástico es aquel en que las varianzas de las perturbaciones no son constantes, por lo tanto, la variabilidad es diferente para cada observación.
- La matriz de varianzas-covarianzas es diagonal.
- Por consiguiente, se sigue verificando independencia entre las observaciones aunque éstas no provienen de la misma población.

$Y = X\beta + \upsilon$ donde υ sigue una ley $N(0,\Sigma)$

$$\Sigma = diag(\sigma_1^2 \ldots\ldots \sigma_T^2) = \begin{pmatrix} \sigma_1^2 & \ldots & 0 \\ : & : & : \\ 0 & \ldots & \sigma_T^2 \end{pmatrix}$$

Ejemplo

- Una empresa decide invertir en bolsa. Analiza los riesgos de varios activos comprobando que estos son diferentes, de hecho, el riesgo es proporcional al valor del activo.
- En este caso se dice que existe heterocedasticidad aditiva creciente.

Diagnosis

Universidade de Vigo

- Como siempre haremos uso de dos instrumentos:
- Gráficos
  - Vamos a utilizar los gráficos de los residuos y residuos al cuadrado.
- Test de hipótesis
  - Normalmente los test de hipótesis son de tipo LM, y buscan comprobar que la varianza no es constante.

Gráficos de residuos

Se representan los residuos respecto a la variable de la que se sospecha que causa la heterocedasticidad. Dado que son ruido blanco, bajo el MRLC, deberían de verse centrados en la media y con bandas constantes. Observando la forma del gráfico nos ayuda a confirmar la sospecha. Siempre que aparezcan bandas crecientes o decrecientes en función de alguna variable es síntoma de heterocedasticidad.

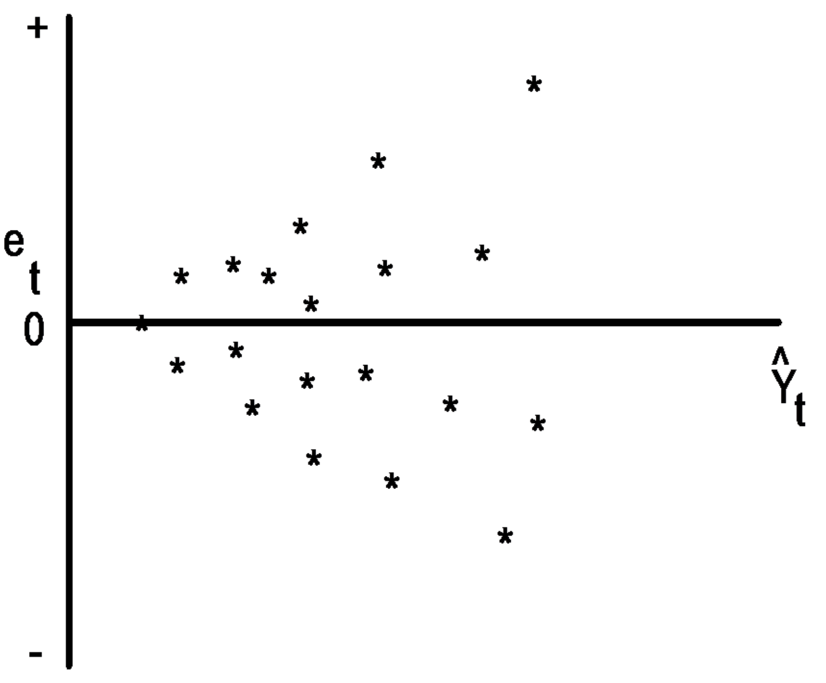
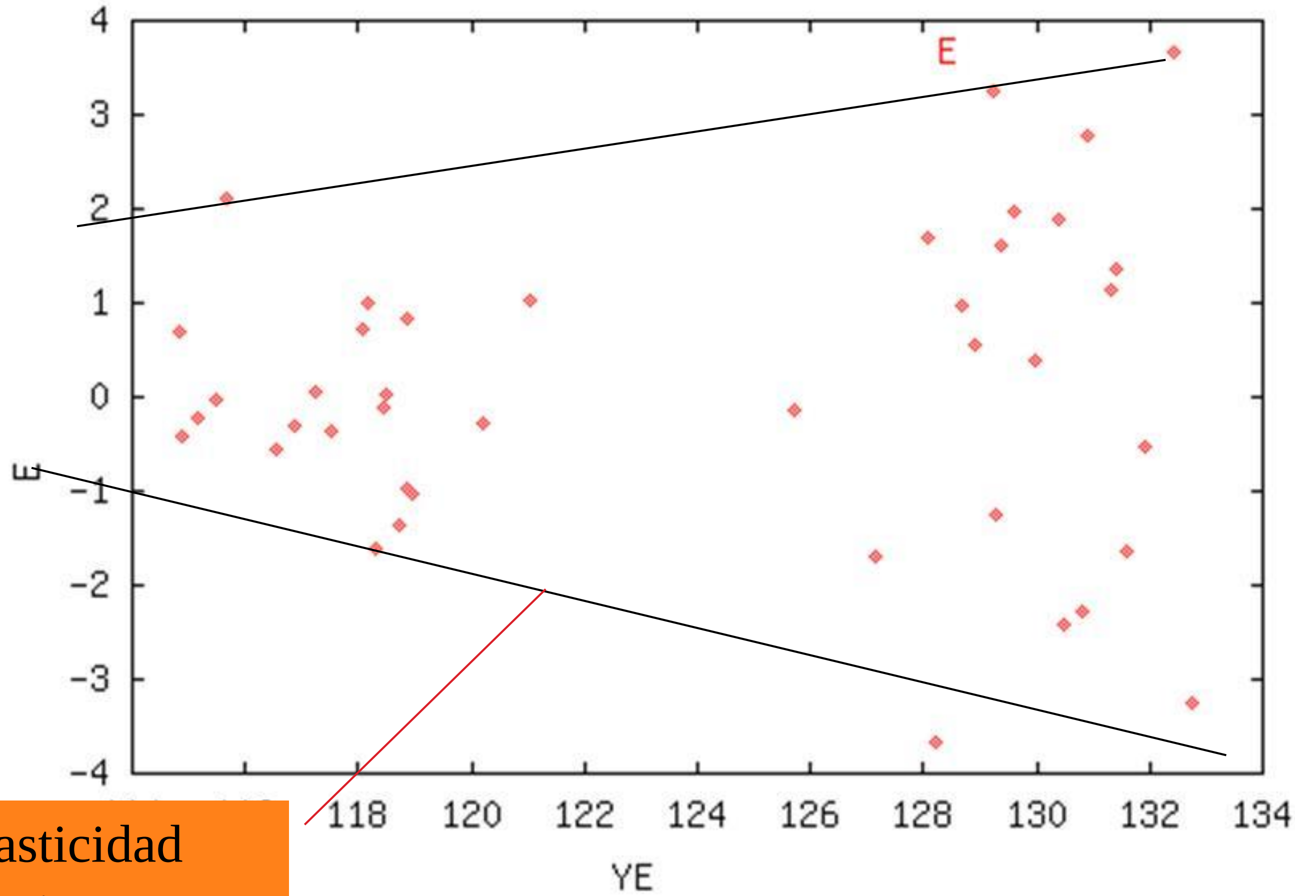


Grafico de residuos respecto a valores predichos



Heterocedasticidad
creciente

Gráficos de residuos al cuadrado

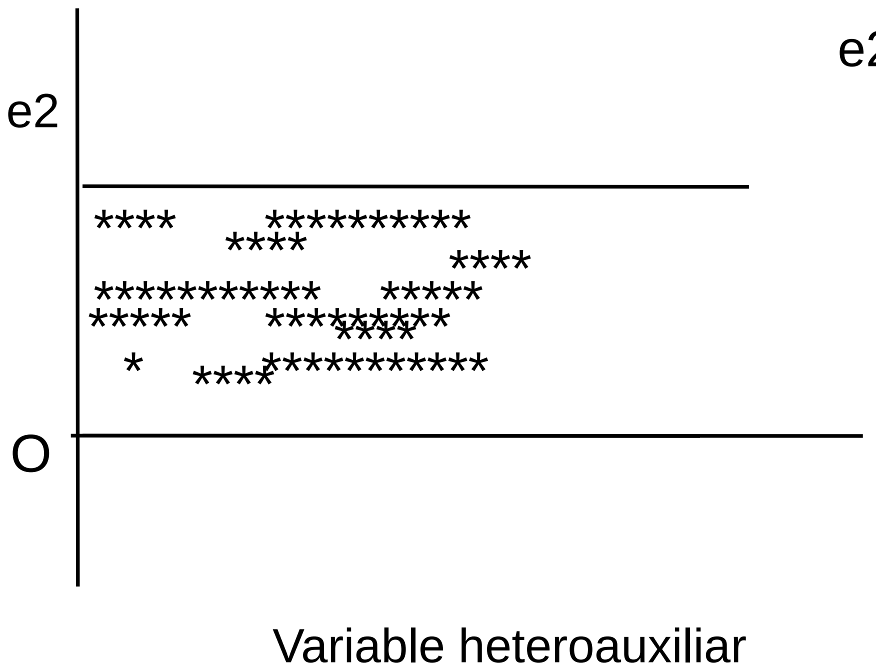
Universidade de Vigo

- Los residuos al cuadrado representan estimadores de la varianza cuando esta es diferente para cada observación. Por lo que nos dan idea de la forma funcional en que esta varía.
- Se representan los residuos al cuadrado respecto a la variable de la que se sospecha que causa la heterocedasticidad. La forma estándar, bajo el MRLC, de ese grafico debería ser de bandas constantes.
- Si e_t son $N(0,\sigma^2)$, los residuos al cuadrado estandarizados seguirán aproximadamente una chi cuadrado con 1 grado de libertad. Por tanto deberán tener de media 1 y varianza 2, por lo tanto debe ser constante.
- Si existe heterocedasticidad la forma de la nube de puntos de los residuos no es constante.
- Esa forma funcional nos sugiere el tipo de heterocedasticidad.
- La imagen de alguno de los modelos introducidos previamente serían las siguientes

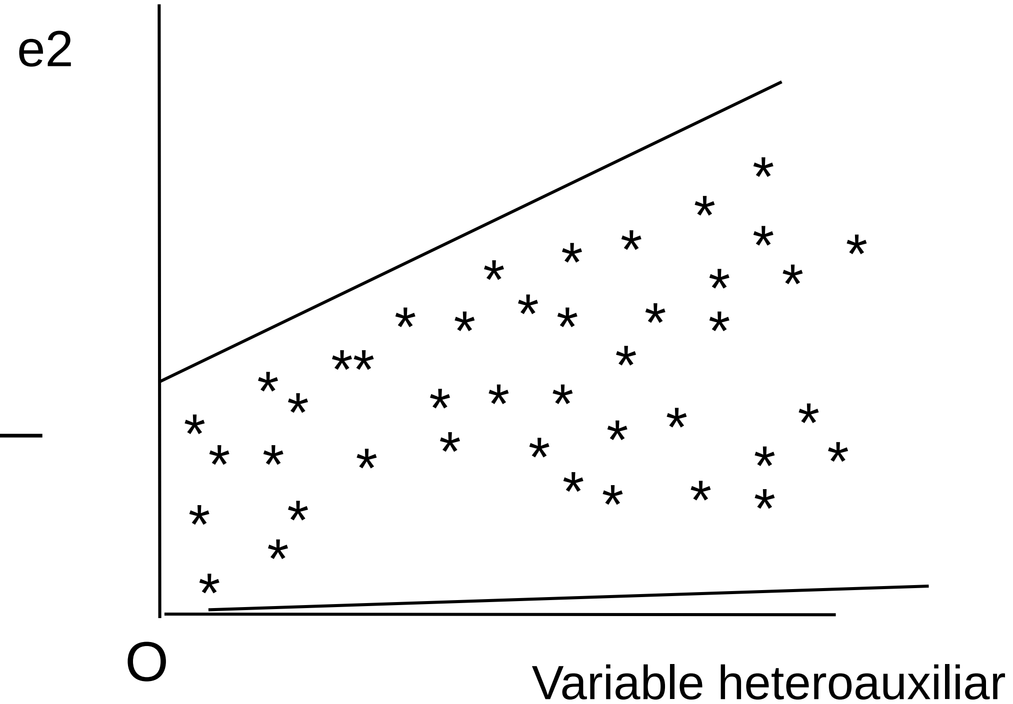
Gráficos de residuos al cuadrado

Universidade de Vigo

Homocedasticidad



Heterocedasticidad lineal



Modelo básico de los test de heterocedasticidad

Universidade de Vigo

- En general todos los test parten de un modelo general heterocedástico, como hipótesis alternativa y la nula es el homocedástico.
- Por tanto hay una serie de suposiciones comunes, salvo las que se refieren a la varianza, que en general dependerá de una función de una serie de variables. Esa función es la que da pie a los diferentes test de hipótesis.
- El modelo en general sería:

$Y=X\beta+\upsilon$ donde υ sigue una ley $N(0,\Sigma)$

$$\Sigma = diag(\sigma_1^2 .....\sigma_T^2) = \begin{pmatrix} \sigma_1^2 & \ldots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \ldots & \sigma_T^2 \end{pmatrix}$$

Las variables X son exógenas, no colineales y medibles

Test de tipo LM

- En este caso se supone que la hipótesis alternativa es una forma funcional de un conjunto de variables Z. Es conveniente especificar la función, que es la que se supone en la hipótesis alternativa .
- Las hipótesis a contrastar son:
  - H_0 : Existe homocedasticidad $\sigma_t^2=\sigma^2 \ \forall t=1,\ldots,T$
  - H_1 : Existe heterocedasticidad en función de Z, siendo Z un conjunto de variables exógenas.

$$\sigma^2_t = h^{-1}(Z'_t)$$

$$h(\sigma^2_t)=\alpha_0+\alpha_1 Z_{1t}+\ldots+\alpha_p Z_p$$

- De hecho se contrasta contra la hipótesis de que la función de la varianza sigue un modelo aditivo en función de las variables transformadas.

Idea del test de tipo LM

- La idea básica es hacer una estimación de las varianzas a partir de los residuos estandarizados.
- Bajo H_0 estos estimadores serán aproximadamente constantes.
- Mientras que bajo H_1 seguirán el modelo establecido por dicha hipótesis
- Para contrastarlo se hará uso de un estadístico que mida la relación lineal entre los estimadores de las varianzas y las variables heteroauxiliares. Eso significa que debemos indicar el modelo de contraste en cada caso, aunque la forma general de construirlo es común.

Idea gráfica del test de Breusch-Pagan

Universidade de Vigo

Gráfico de residuos respecto a las variables exógenas

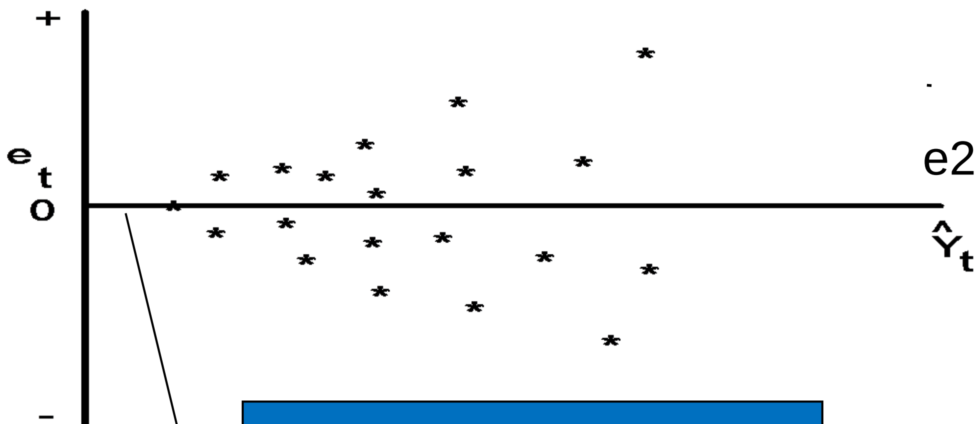
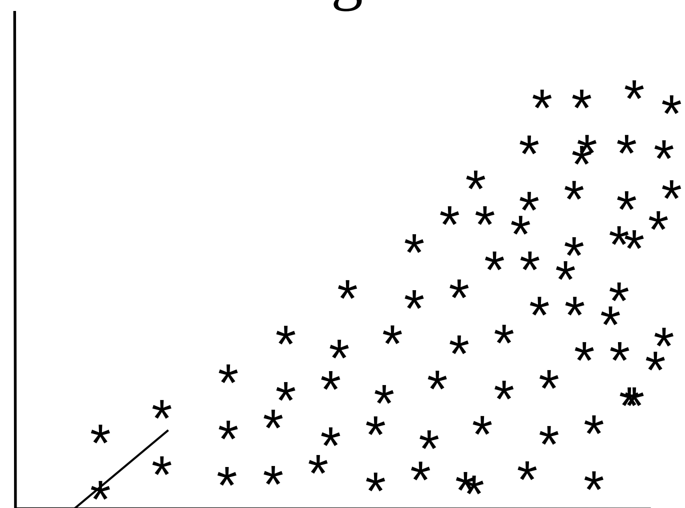


Gráfico de residuos al cuadrado respecto a las variables exógenas



La varianza crece en función de uno o mas de las variables exógenas

El gráfico de residuos al cuadrado se puede aproximar por una función lineal

El modelo sería heterocedástico aditivo de la varianza en función de las independientes

Cálculo del test de tipo LM

Universidade de Vigo

- Se estima por MCO el modelo original y se calculan los residuos del modelo.
- Se realiza una estimación de las varianzas bajo H_1 haciendo uso de los residuos al cuadrado en función de las variables indicadas en la hipótesis H1. Eso es lo que indica cada uno de los test que se realicen
- Se realiza una regresión por MCO de esos estimadores de la varianza respecto a las variables exógenas y se calcula el coeficiente de determinación de la regresión anterior R^2 .
- Se calcula el estadístico de este test, basado en el método LM, que será $LM=TR^2$. Dicho estadístico seguirá asintóticamente una ji cuadrado con p grados de libertad, bajo H_0
- La regla de decisión será rechazar H_0 si el estadístico LM es mayor que el valor de las tablas $LM > \chi^2_{p,\alpha}$

Test de heterocedasticidad

Universidade de Vigo

- Existen tras versiones genéricas
  - Test utilizando las variables independientes del modelo
    - Test de Breusch Pagan
      - Utiliza la suma de cuadrados de la regresión
    - Test de Koenker
      - Utiliza el R^2 de la regresión
  - Test utilizando cuadrados de las independientes
    - Test de White
      - Combina las variables independientes con sus cuadrados y productos cruzados obteniendo un estimado consistente de la varianza

Test en OLS en Gretl

Universidade de Vigo

```
Contraste de heterocedasticidad de Breusch-Pagan
MCO, usando las observaciones 1977:01-1989:01 (T = 145)
Variable dependiente: uhat^2 escalado

             Coeficiente   Desv. Típica   Estadístico t    Valor p
  ----------------------------------------------------------------
  const         2.33856      0.208932        11.19        2.80e-021 ***
  POILR       -19.3943       2.84416         -6.819       2.39e-010 ***

  Suma de cuadrados explicada = 34.5146

Estadístico de contraste: LM = 17.257322,
con valor p  = P(Chi-cuadrado(1) > 17.257322) = 0.000033
Contraste de heterocedasticidad de Breusch-Pagan (Koenker)
MCO, usando las observaciones 1977:01-1989:01 (T = 145)
Variable dependiente: uhat^2 escalado (variante robusta de Koenker)

             Coeficiente   Desv. Típica   Estadístico t    Valor p
  ----------------------------------------------------------------
  const         249.805       38.9913         6.407       2.01e-09  ***
  POILR       -3619.40       530.782         -6.819       2.39e-010 ***

  Suma de cuadrados explicada = 1.20207e+006

Estadístico de contraste: LM = 35.579691,
con valor p  = P(Chi-cuadrado(1) > 35.579691) = 0.000000
```

Ejercicio 3.14

Test de White

Contraste de heterocedasticidad de White

MCO, usando las observaciones 1977:01-1989:01 (T = 145)

Variable dependiente: uhat^2

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	426.573	111.450	3.827	0.0002	***
POILR	-3316.11	3255.63	-1.019	0.3101	
sq_POILR	-2052.86	21739.6	-0.09443	0.9249	

R-cuadrado = 0.245425

Estadístico de contraste: TR^2 = 35.586561,

con valor p = P(Chi-cuadrado(2) > 35.586561) = 0.000000

Tratamiento de la Heterocedasticidad

Universidade de Vigo

- Aunque los estimadores de MCO siguen siendo consistentes, pierden la eficiencia y por ello nos interesa buscar unos estimadores que sean mejores.
- Los estimadores de las varianzas son inconsistentes por lo que para evitar ese problema se pueden estimar las varianzas mediante estimadores robustos.
- El proceso genérico para modelos con heterocedasticidad se denomina de mínimos cuadrados generalizados factibles (MCGF)