Modelo de regresión lineal clásico

Modelo
Suposiciones
Estimadores MCO
Propiedades
Interpretación

Ejemplo Salarios-Productividad

Universidade de Vigo

Ejemplo SP1*: La Tabla presenta información sobre Y, el índice de compensación real por hora y X, el índice de productividad por hora, para el sector comercial de la economía de los EE.UU. entre los años 1960 y 1983.*

Puesto que se espera que la compensación real y la productividad estén relacionadas, se puede utilizar el siguiente modelo sencillo para averiguar hasta qué punto las dos variables están relacionadas

Año	Y	X	Año	Y	X
1960	69.5	65.2	1972	95.7	92.4
1961	71.4	67.4	1973	97.3	94.8
1962	73.8	69.9	1974	95.9	92.5
1963	75.6	72.5	1975	96.4	94.6
1964	78.4	75.6	1976	98.9	97.6
1965	80.1	78.3	1977	100	100
1966	83.3	80.8	1978	100.8	100.5
1967	85.3	82.6	1979	99.1	99.3
1968	88.3	85.3	1980	96.4	98.8

sigue

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Este es un ejemplo de modelo lineal de regresión clásico

B_0 nos mediría el índice de compensación salarial cuando el de productividad fuera 0

B_1 nos mediría la variación del índice de compensación salarial cuando el de productividad varía una unidad

El modelo de regresión

Universidade de Vigo

- Trata de evaluar la relación de causa efecto
- Considera diferentes factores que pueden producir una determinada respuesta.
  - A los factores se les considera variables independientes (las representaremos por la letra X)
  - a la variable respuesta como variable dependiente (la representaremos por la letra Y)
- Por tanto siempre presenta una relación asimétrica, unas variables son factores y otra variable es respuesta
- El nombre viene del primer trabajo realizado con este método.

Formalizando

Universidade de Vigo

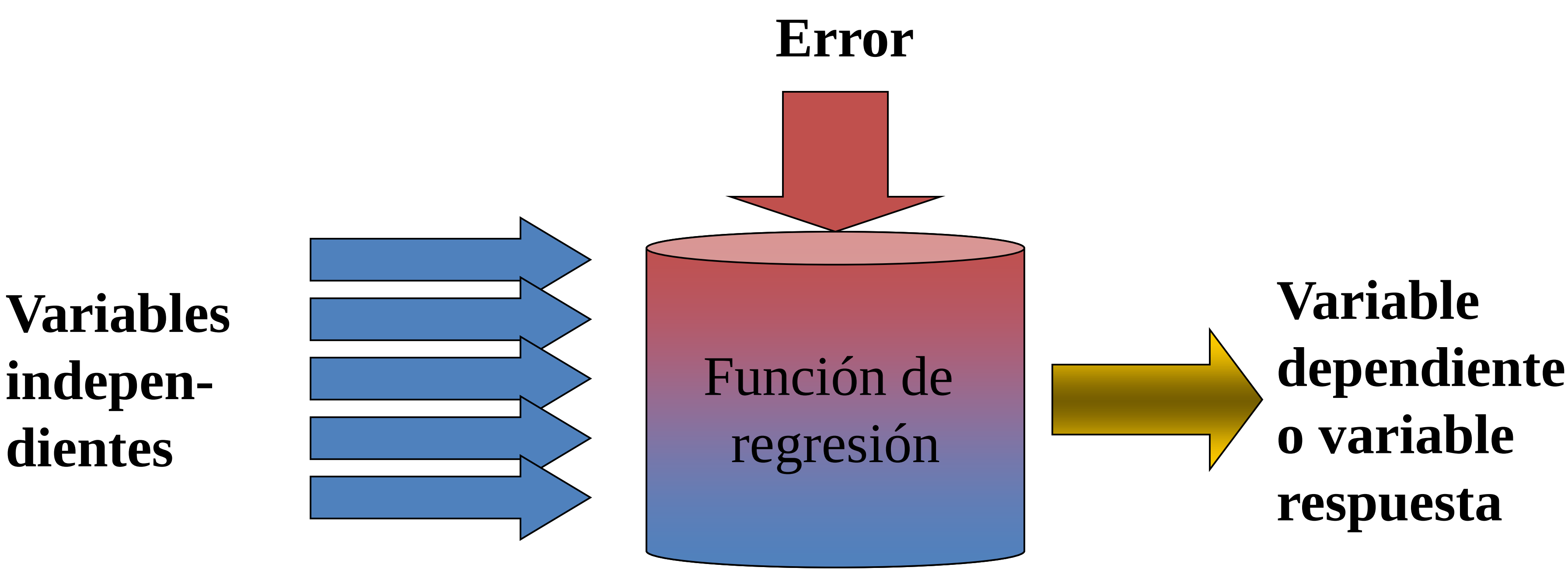
$$y = f(X_1, \ldots X_k) + g(X_{k+1}, \ldots)$$

f *regresion*   *error*

- La función **f** recoge la información de los factores relevantes, que se denominan *independientes o explicativas*, para explicar la dependiente o *variable respuesta*
- La función **g** recoge el efecto de las variables no relevantes que se denomina *perturbación*.

Flujos del modelo lineal general

- En formato de flujos el modelo quedaría de la siguiente forma:



Modelo y suposiciones

Parte A

Universidade
de Vigo

Suposiciones

- **Sobre la forma funcional**
  1. Linealidad:

- **Sobre las perturbaciones:**
  2. Esperanza nula
  3. Homocedasticidad
  4. Independencia
  5. Normalidad

- **Sobre los regresores:**
  6. Exogeneidad
  7. No colinealidad
  8. Mensurabilidad

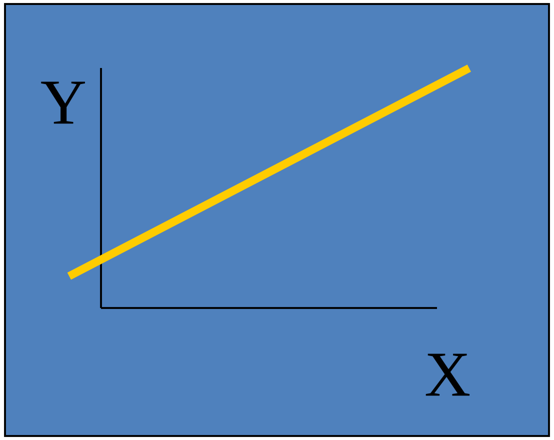
- **Sobre los parámetros:**
  9. Estabilidad
  10. Identificabilidad

1. Linealidad

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

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

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



Ejemplo de compensación salarial

- En el ejemplo del índice de compensación real por hora, esta suposición se expresaría diciendo que :
- *Un incremento del índice de productividad por hora en una unidad genera siempre el mismo incremento esperado del índice de compensación real por hora, independientemente del resto de efectos considerados, es decir, el efecto del índice de productividad sobre el índice de compensación real es siempre constante*

$$\frac{\partial Y}{\partial X} = \beta_1$$

- Por eso el modelo puede escribirse como:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

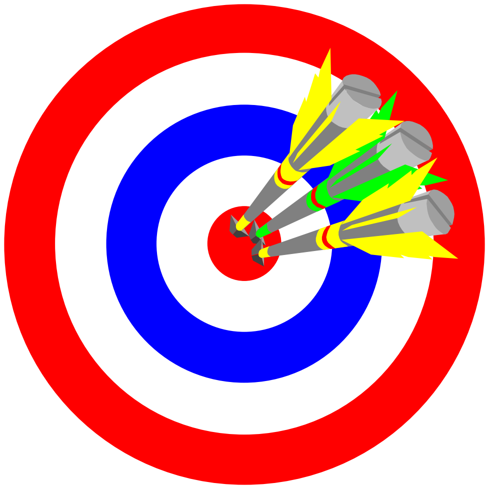
2. Esperanza nula

$$E(\varepsilon)=0$$

El valor esperado de las perturbaciones es 0.

Por tanto el efecto esperado de las variables no introducidas en el modelo es despreciable.

Cuando eso no ocurre, la constante asume dicho efecto, por lo que dificulta su interpretación.



Ejemplo de compensación salarial

Universidade de Vigo

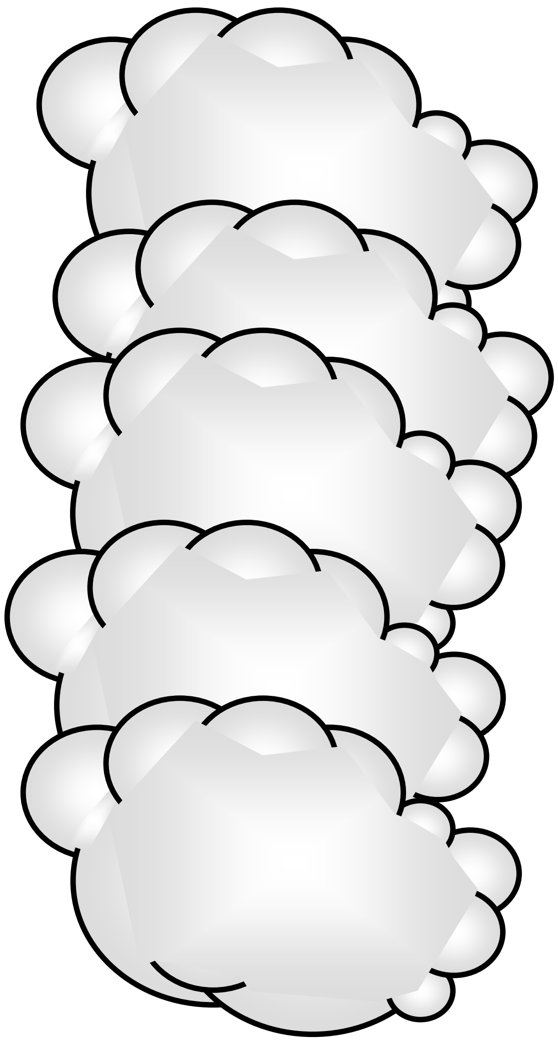
- En el ejemplo del índice de compensación real por hora, esta suposición se expresaría diciendo que :
- *La única variable que tiene efecto sobre el valor esperado del índice de productividad por hora es el índice de compensación real por hora. El resto de efectos son despreciables*

$$E(Y/X) = \beta_0 + \beta_1 X$$

$$E(\varepsilon) = 0$$

3. Homocedasticidad

Universidade de Vigo



$$Var(\varepsilon_t)=\sigma^2 \quad \forall t$$

La varianza de las perturbaciones es constante en todas las observaciones, por tanto la variabilidad alrededor del valor esperado de la variable respuesta, para cada posible combinación de valores de las variables independientes es la misma.

El riesgo esperado de predecir el valor de la dependiente conociendo todas las independientes es el mismo y se mantiene constante para todas las observaciones.

Ejemplo de compensación salarial

Universidade de Vigo

- En el ejemplo del índice de compensación real por hora, esta suposición se expresaría diciendo que :
- *El riesgo asociado a la predicción del índice de productividad por hora condicionado a la información suministrada por el índice de compensación real por hora es el mismo para todos los años*

$$Var(Y/X) = \mathrm{var}(\varepsilon) = \sigma^2$$

4. Independencia

$$E(\varepsilon_t \varepsilon_s)=0 \ \forall t \neq s$$

Las observaciones son todas independientes. Ninguna da información sobre el resto.

En consecuencia, toda la información que se tiene sobre la variable respuesta está en los regresores, pues las perturbaciones son totalmente imprevisibles, en cada observación.

Ejemplo de compensación salarial

Universidade de Vigo

- En el ejemplo del índice de compensación real por hora, esta suposición se expresaría diciendo que :
- *El índice de productividad por hora condicionado a la información suministrada por el índice de compensación real por hora no depende de su pasado. Cada año, la información sobre el índice de productividad por hora se agota con la información suministrada por el índice de compensación real por hora.*

$$Cov(Y_t, Y_{t-s} / X) = 0$$

3’. Varianza escalar

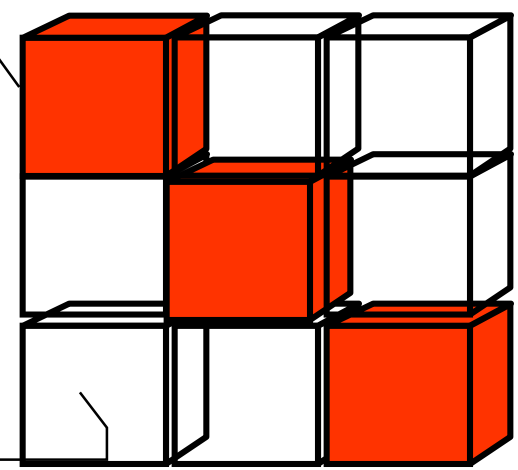
Universidade de Vigo

Las dos anteriores se pueden resumir en una, denominada matriz de varianzas de las perturbaciones escalar, y que escribiríamos matemáticamente como

$$Var(\varepsilon)=E(\varepsilon\varepsilon')=\sigma^2 I$$

De esta forma aseguramos que todas las observaciones se comportan de modo semejante

Las varianzas son todas iguales



Las covarianzas son todas nulas

5. Normalidad

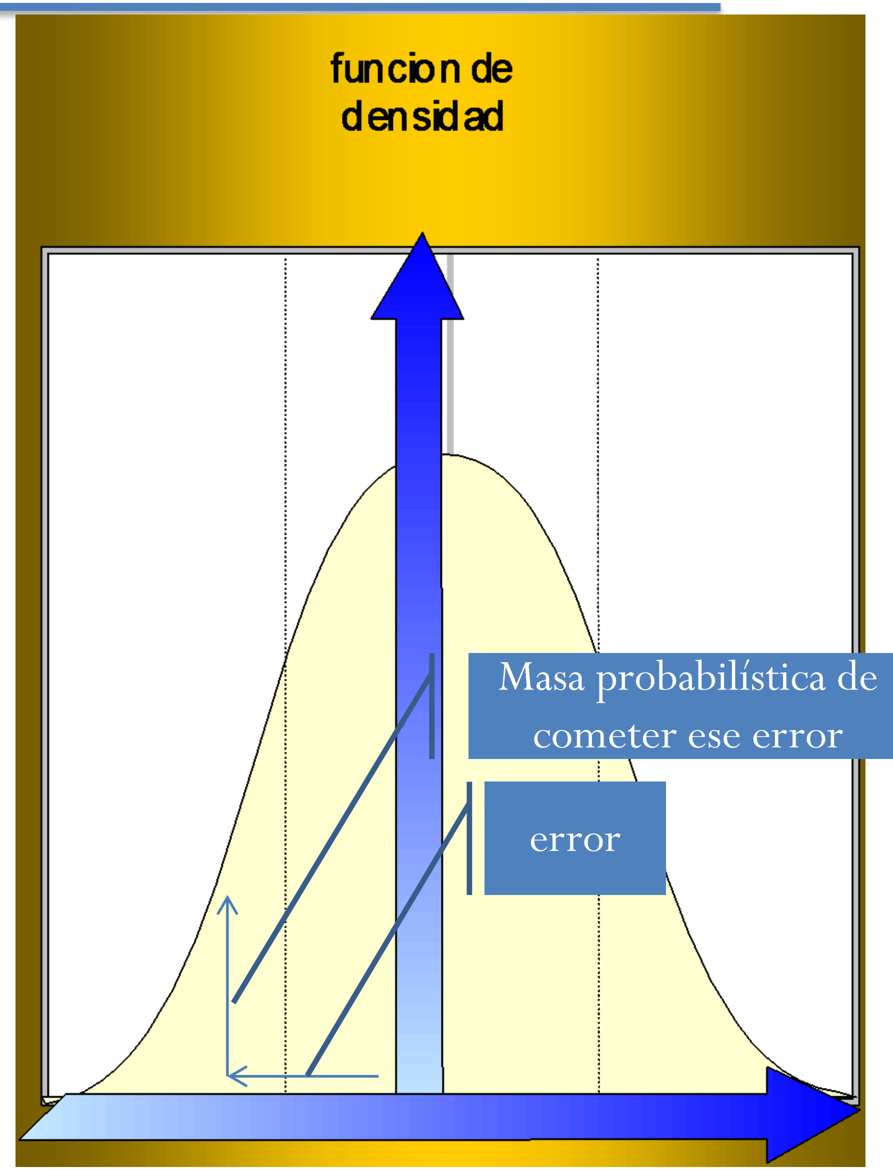
Universidade de Vigo

Las perturbaciones siguen todas leyes normales.

Nos indica que el valor más probable es el valor esperado y la probabilidad irá disminuyendo a medida que nos alejamos del valor promedio, de forma simétrica.

Cuanto mayor sea el error menor probabilidad de ocurrir tiene.

Esto nos permite hacer inferencias sobre la estimación del modelo.



Ejemplo de compensación salarial

Universidade de Vigo

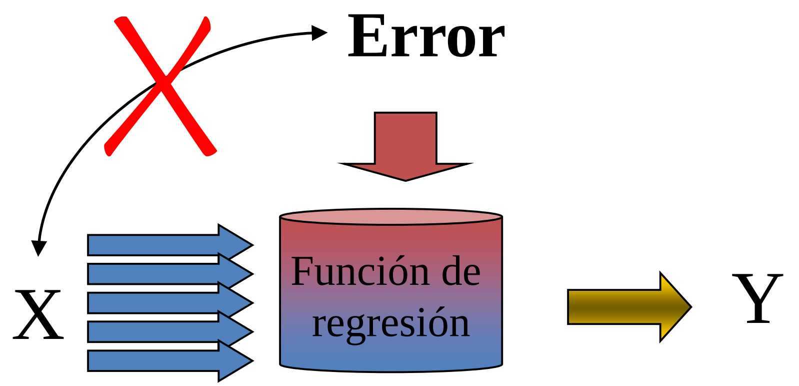
- En el ejemplo del índice de compensación real por hora, esta suposición se expresaría diciendo que :
- *El valor esperado del índice de productividad por hora condicionado a la información suministrada por el índice de compensación real es el valor mas probable.*
- *Cuanto mayor divergencia haya respecto a ese valor mas improbable es que ocurra, pero esa probabilidad no depende del signo sino del tamaño de la divergencia.*

6. Exogeneidad

Universidade de Vigo

Los regresores son fijos o son exógenos al modelo. Inicialmente supondremos que son determinísticos. Como consecuencia las variables independientes no están correladas con las perturbaciones y podemos considerarlas como si fueran fijas en muestras repetidas.

Las variables no introducidas en el modelo no afectan a los factores independientes.



Ejemplo de compensación salarial

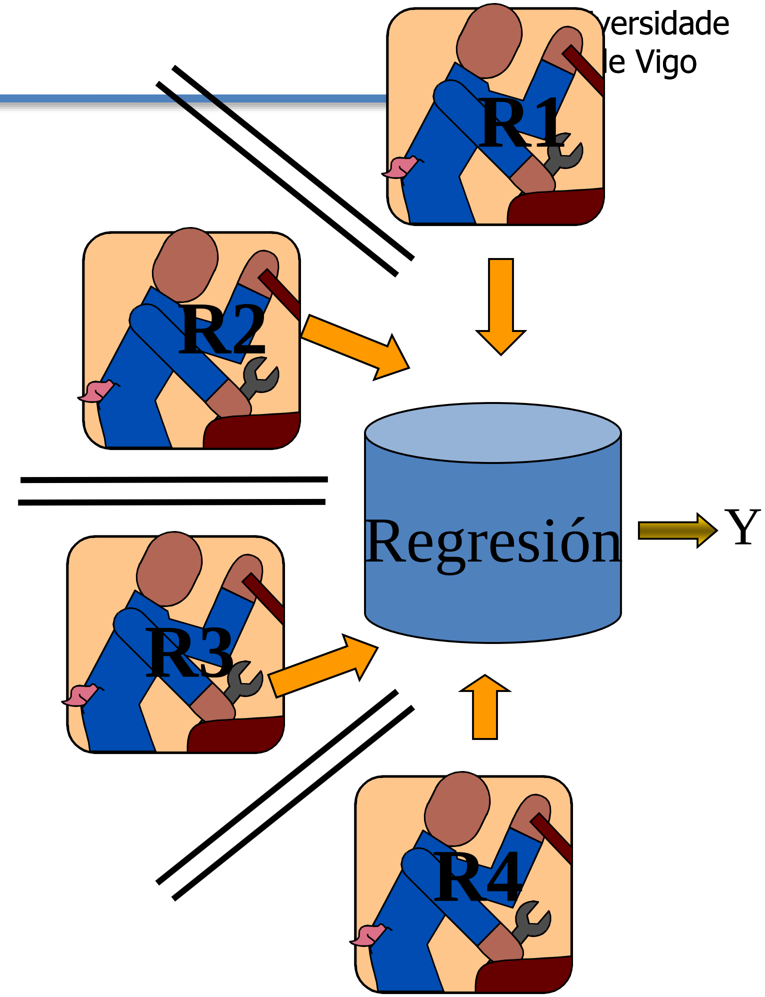
Universidade de Vigo

- En el ejemplo del índice de compensación real por hora, esta suposición se expresaría diciendo que :
- *El índice de compensación real por hora no depende del índice de productividad por hora ni de ningún otro factor que condicione a éste último.*

7. No colinealidad

No existe dependencia lineal entre los regresores. En realidad la suposición incluye aún más, pues exige, no sólo que no exista dependencia lineal exacta entre los regresores, sino que ésta no sea ni siquiera aproximada.

La información suministrada por cada variable independiente no está contenida en el resto de las variables independientes.



Ejemplo de compensación salarial

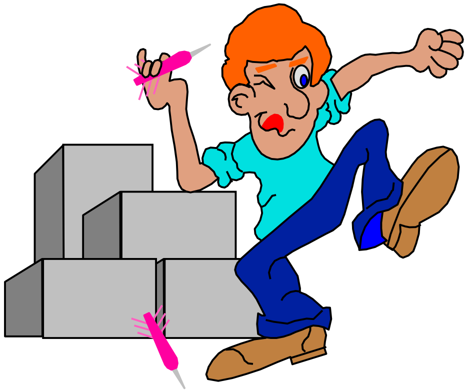
Universidade de Vigo

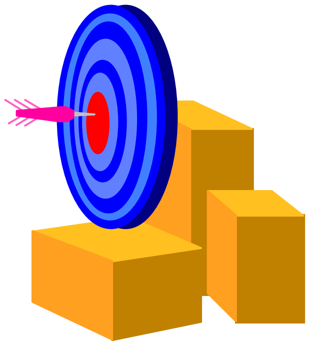
- En el ejemplo del índice de compensación real por hora, esta suposición no es necesaria puesto que sólo se incluye una variable independiente. Si hubiera mas se diría que:
- *No existe relación entre ellas, ni comparten información relativa al índice de compensación salarial por hora.*

8. Mensurabilidad

Universidade de Vigo

Las variables independientes se miden sin error. De esa forma se asegura que los errores de medición no afectan al modelo, y todo el error está concentrado en las variables no utilizadas para predecir la dependiente.





Ejemplo de compensación salarial

Universidade de Vigo

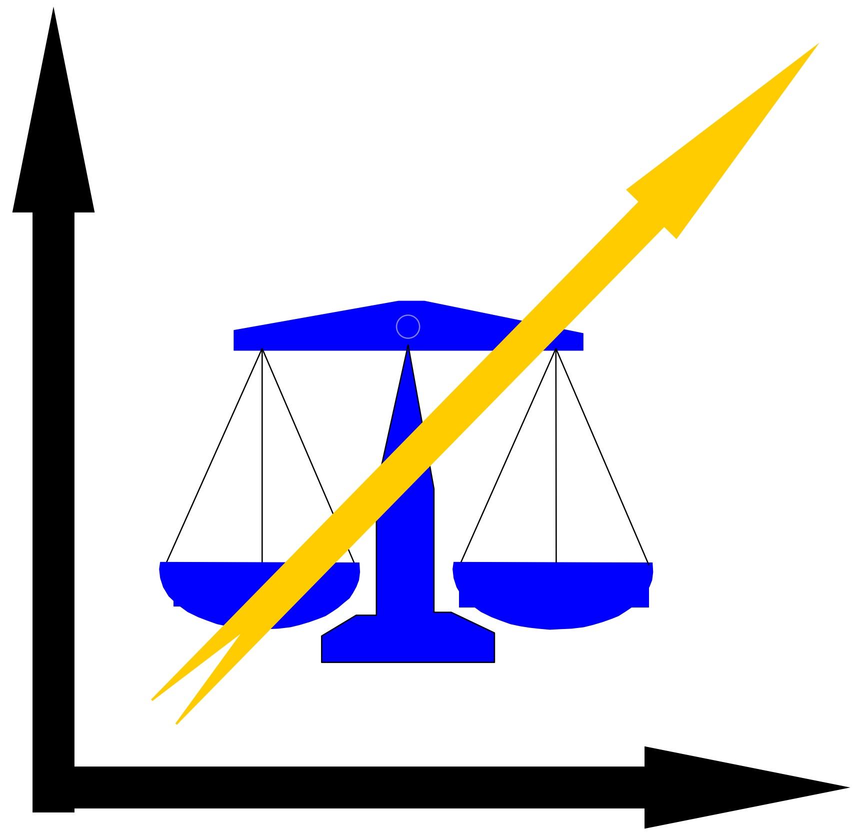
- En el ejemplo del índice de compensación real por hora, esta suposición se expresaría diciendo que :
- *Los datos obtenidos sobre el índice de compensación real por hora están medidos sin error, es decir se conoce con exactitud sus valores, no son aproximaciones.*

9. Estabilidad

Universidade de Vigo

Los parámetros no varían al considerar las distintas observaciones. Esta suposición exige que el modelo permanezca inalterable para todo el periodo muestral.

De esta forma aseguramos que el modelo sea válido y que su comportamiento sea similar para todas las observaciones.



Ejemplo de compensación salarial

Universidade de Vigo

- En el ejemplo del índice de compensación real por hora, esta suposición se expresaría diciendo que :
- *La relación existente entre el índice de productividad por hora y el índice de compensación real por hora se mantiene estable durante todos los años.*

10. Identificabilidad

Universidade de Vigo

De acuerdo con el método de estimación fijado, existen estimadores únicos de los parámetros, para la muestra observada. Nos asegura que los estimadores sean únicos y así cuando los calculemos no tenemos que decidir entre varias alternativas sin justificación científica.

En el caso del modelo lineal se concreta en que el número de parámetros del modelo sea menor que el de observaciones.

Ejemplo de compensación salarial

Universidade de Vigo

- En el ejemplo del índice de compensación real por hora, como se estima por mínimos cuadrados ordinarios solo se necesita que haya mas de tres años, cosa que ocurre.

Ejercicio 2.1.

Forma matricial del modelo (1)

Universidade de Vigo

Suponemos T observaciones y k variables independientes

$$y_1 = \beta_0 + \beta_1 x_{11} + \ldots + \beta_k x_{k1} + \varepsilon_1$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$y_T = \beta_0 + \beta_1 x_{1T} + \ldots + \beta_k x_{kT} + \varepsilon_T$$

Forma matricial del modelo (2)

Universidade de Vigo

- Reescribiendo en forma de matrices y vectores los ecuaciones anteriores nos quedaría:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_T \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots x_{k1} \\ \vdots & \vdots & \ddots \quad \vdots \\ 1 & x_{1T} & \dots x_{kT} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \end{pmatrix}$$

- Que es la forma matricial del modelo.
- A la matriz X se le denomina **matriz de diseño**.
- Escrito de forma compacta señalando los vectores y las matrices por letras nos quedará el modelo siguiente.

Forma matricial del modelo (3)

Universidade de Vigo

El modelo se escribiría matricialmente como

$$\mathbf{y} = X\,\beta + \boldsymbol{\varepsilon}$$

donde

- $\mathbf{y} = (y_1 \ldots y_T)'$ es un vector Tx1 de observaciones de la variable dependiente
- X es una matriz Tx(k+1) de observaciones de las variables independientes, con la primera columna de 1.

□ $\boldsymbol{\beta} = (\beta_0,\beta_1,\ldots,\beta_k)'$ es un vector (k+1)x1 de parámetros a estimar en el modelo.

- $\boldsymbol{\varepsilon} = (\varepsilon_1 \ldots \varepsilon_T)'$ es un vector Tx1 de perturbaciones aleatorias

Modelo matricial

En el caso del problema de la relación entre salarios y productividad, tendríamos una ecuación para cada observación, es decir:

$$y_1 = \beta_0 + \beta_1 x_{11} + \varepsilon_1 \qquad 69,5 = \beta_0 + 65,2\beta_1 + \varepsilon_1$$

$$\cdots\cdots\cdots\cdots \qquad \cdots\cdots\cdots\cdots$$

$$y_T = \beta_0 + \beta_1 x_{124} + \varepsilon_{24} \qquad 98,4 = \beta_0 + 103,7\beta_1 + \varepsilon_{24}$$

En forma matricial se escribiría

Ejemplo de modelo matricial

Universidade de Vigo

$$\mathbf{y} = X\beta + \varepsilon$$

Escrito de forma desarrollada seria

Donde:

$$Y = \begin{pmatrix} 69.5 \\ 71.4 \\ 73.8 \\ 75.6 \\ 78.4 \\ 80.1 \\ 83.3 \\ 85.3 \\ 88.3 \\ 89.7 \\ 90.8 \\ 92.8 \\ 95.7 \\ 97.3 \\ 95.9 \\ 96.4 \\ 98.9 \\ 100 \\ 100.8 \\ 99.1 \\ 96.4 \\ 95.5 \\ 97.3 \\ 98.4 \end{pmatrix} \quad X = \begin{pmatrix} 1 & 65.2 \\ 1 & 67.4 \\ 1 & 69.9 \\ 1 & 72.5 \\ 1 & 75.6 \\ 1 & 78.3 \\ 1 & 80.8 \\ 1 & 82.6 \\ 1 & 85.3 \\ 1 & 85.5 \\ 1 & 86.2 \\ 1 & 89.3 \\ 1 & 92.4 \\ 1 & 94.8 \\ 1 & 92.5 \\ 1 & 94.6 \\ 1 & 97.6 \\ 1 & 100 \\ 1 & 100.5 \\ 1 & 99.3 \\ 1 & 98.8 \\ 1 & 100.7 \\ 1 & 100.9 \\ 1 & 103.7 \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \vdots \\ \varepsilon_{24} \end{pmatrix}$$

$$\begin{pmatrix} 69.5 \\ 71.4 \\ 73.8 \\ 75.6 \\ 78.4 \\ 80.1 \\ 83.3 \\ 85.3 \\ 88.3 \\ 89.7 \\ 90.8 \\ 92.8 \\ 95.7 \\ 97.3 \\ 95.9 \\ 96.4 \\ 98.9 \\ 100 \\ 100.8 \\ 99.1 \\ 96.4 \\ 95.5 \\ 97.3 \\ 98.4 \end{pmatrix} = \begin{pmatrix} 1 & 65.2 \\ 1 & 67.4 \\ 1 & 69.9 \\ 1 & 72.5 \\ 1 & 75.6 \\ 1 & 78.3 \\ 1 & 80.8 \\ 1 & 82.6 \\ 1 & 85.3 \\ 1 & 85.5 \\ 1 & 86.2 \\ 1 & 89.3 \\ 1 & 92.4 \\ 1 & 94.8 \\ 1 & 92.5 \\ 1 & 94.6 \\ 1 & 97.6 \\ 1 & 100 \\ 1 & 100.5 \\ 1 & 99.3 \\ 1 & 98.8 \\ 1 & 100.7 \\ 1 & 100.9 \\ 1 & 103.7 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \vdots \\ \varepsilon_{24} \end{pmatrix}$$

Resumen del modelo

Universidade de Vigo

- De modo resumido el modelo de regresión lineal normal clásico se puede escribir como
- Las observaciones de la variable dependiente siguen una ley normal con media una combinación lineal de las variables independientes y varianza constante .
- Las **suposiciones sobre las variables independientes** se presuponen válidas en todos los casos, es decir, son exógenas, no colineales y medibles.

$$Y \text{ sigue ley } N(X\beta, \sigma^2 I)$$

suposiciones sobre distribución

suposiciones sobre el valor esperado

suposiciones sobre la varianza

Ejercicio 2.2

Estimadores de Mínimos cuadrados Ordinarios en el modelo de regresión lineal clásico

Universidade de Vigo

Parte B: Estimación y Propiedades

Universidade de Vigo

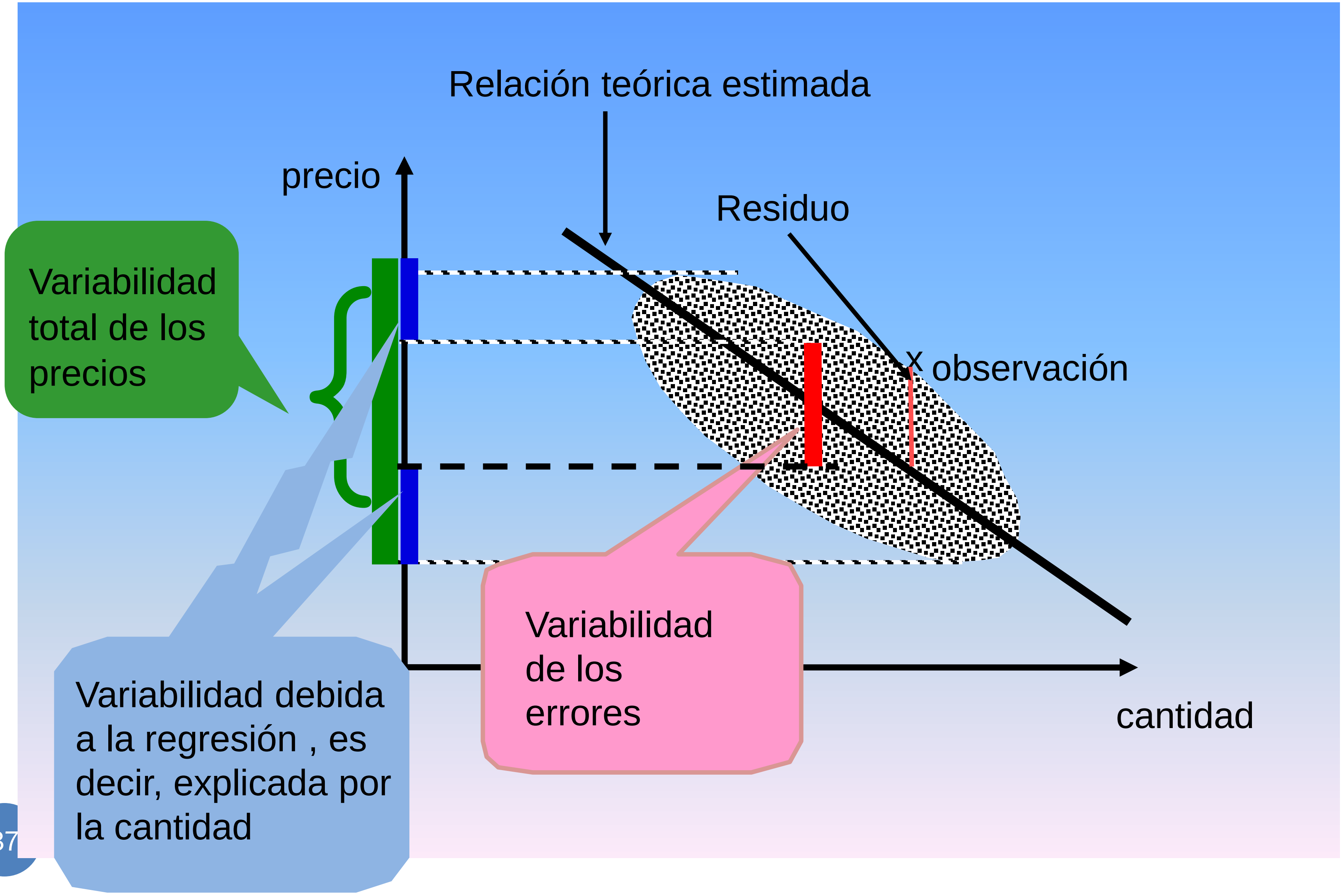
Método de mínimos cuadrados ordinarios

Universidade de Vigo

- Este método es un método geométrico que consiste buscar la recta que minimice la distancia respecto al eje de ordenadas a todos los puntos del espacio.
- Si y es el valor observado de la variable dependiente y Xb es el valor esperado de la y si fuera igual a la recta, la diferencia es el residuo e=y-Xb.
- Por consiguiente el método busca el valor de b que minimiza la suma de los cuadrados de e, es decir de las diferencias entre el valor observado y el esperado, por tanto la distancia entre el vector de observaciones (y) y el de estimaciones (Xb).

Idea gráfica de mínimos cuadrados

Universidade de Vigo



Estimadores de mínimos cuadrados ordinarios (MCO)

Universidade de Vigo

- Buscamos minimizar la suma de cuadrados de los errores

$$SCE = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon'\varepsilon = (y - X\beta)'(y - X\beta)$$

- Para minimizar en β debemos derivar respecto a β e igualar a cero. Derivando de modo similar a como se hace en el caso de una ecuación se obtiene:

$$\frac{\partial}{\partial \beta} SCE = \frac{\partial}{\partial \beta}(y - X\beta)'(y - X\beta) = -2X'(y - X\beta)$$

- Por tanto

$$-2X'(y - X\hat{\beta}) = 0 \rightarrow X'y = X'X\hat{\beta} \rightarrow \hat{\beta} = (X'X)^{-1}X'y$$

- Que es el estimador MCO de los parámetros β

Estimación de mínimos cuadrados ordinarios

Universidade de Vigo

Mide el índice de compensación salarial independientemente de la productividad de los trabajadores

Modelo 1: MCO, usando las observaciones 1960-1983 (T = 24)

Variable dependiente: Y

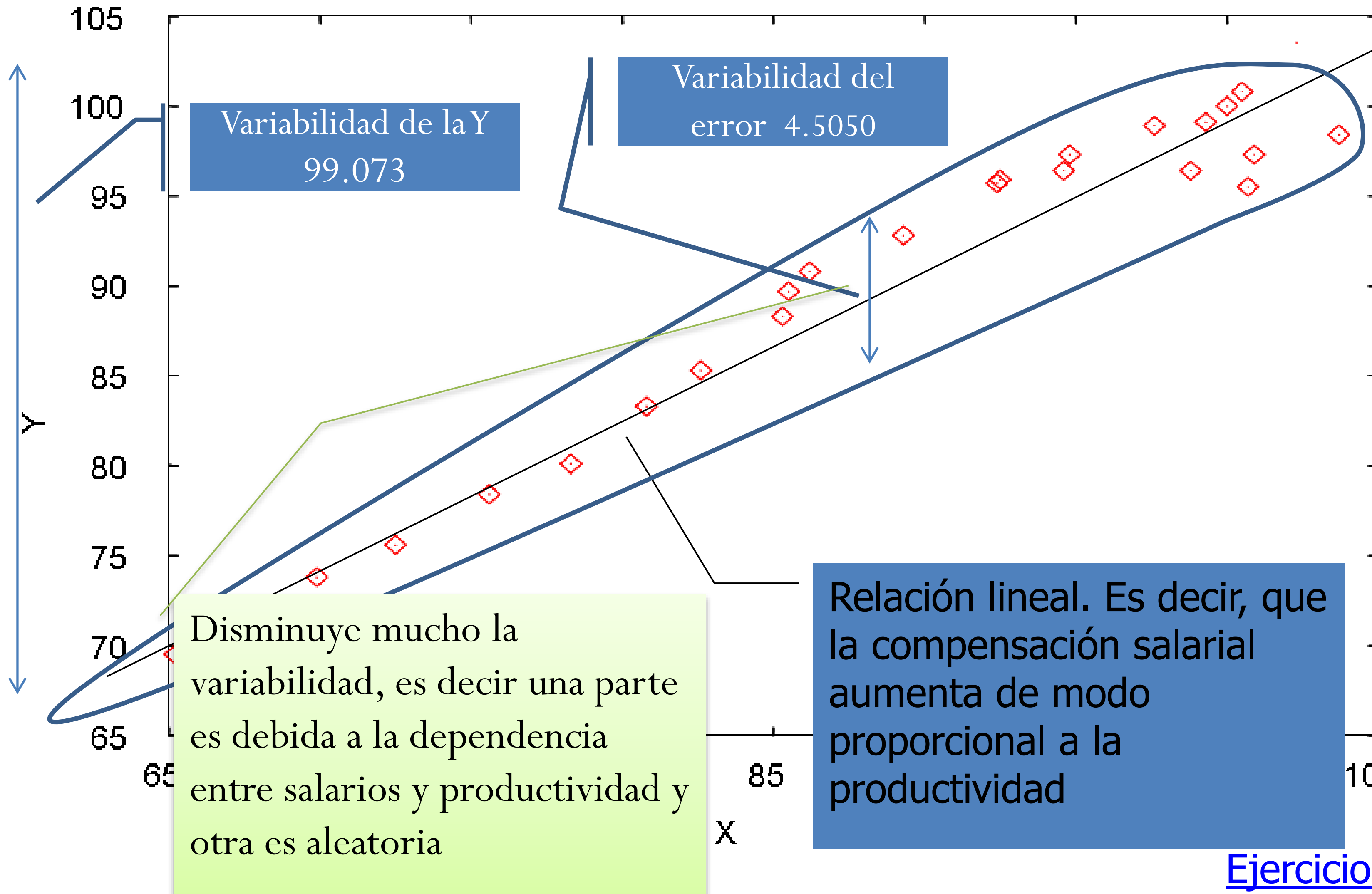
Mide el incremento porcentual de la compensación salarial por cada punto porcentual que se incrementa la productividad de los trabajadores

	Coeficiente	*Desv. Típica*	*Estadístico t*	*Valor p*	
const	16.8978	3.33414	5.0681	<0.0001	***
X	0.825365	0.037524	21.9956	<0.0001	***

Media de la vble. dep.	89.61250	D.T. de la vble. dep.	9.953558
Suma de cuad. residuos	99.11090	D.T. de la regresión	2.122508
R-cuadrado	0.956505	R-cuadrado corregido	0.954528
F(1, 22)	483.8081	Valor p (de F)	1.81e-16
Log-verosimilitud	−51.07275	Criterio de Akaike	106.1455
Criterio de Schwarz	108.5016	Crit. de Hannan-Quinn	106.7706
rho	0.948311	Durbin-Watson	0.239765

Representación gráfica de compensación salarial respecto a productividad

Universidade de Vigo



Ejercicio 2.3

Linealidad de los estimadores de MCO

Universidade de Vigo

- **1) El estimador de mínimos cuadrados ordinarios de los coeficientes de regresión es un estimador lineal, esto es, es combinación lineal de las observaciones de la variable dependiente. Esto es Si cambiamos una observación el estimador cambia proporcionalmente**

Demostración

$$b = (X'X)^{-1}X'y = DY$$

$$siendo\ D = (X'X)^{-1}X'$$

$$luego \Rightarrow b_i = \sum_{t=1}^{T} d_{it}Y_t$$

Importancia:

•Permite relacionar las propiedades de los estimadores con las de la variable dependiente. Por ejemplo, la normalidad, el valor esperado, la varianza, etc…

Insesgadez de los estimadores de MCO

Universidade de Vigo

2) El estimador de mínimos cuadrados ordinarios de los coeficientes de regresión es un estimador insesgado esto es, el valor esperado del estimador coincide con el verdadero valor del parámetro.

Demostración

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

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

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

$$= \beta + (X'X)^{-1}X'0 = \beta$$

Importancia:

- Asegura que nuestro estimador se orienta hacia el verdadero valor del parámetro y por lo tanto no debería estar lejos de dicho valor, pero es necesario saber si el error es también bueno o no.

Precisión de los estimadores de MCO

Universidade de Vigo

3) La varianza de los estimadores depende de la inversa del producto de la matriz de diseño premultiplicada por su traspuesta y de la varianza de las perturbaciones.

$$\mathbf{Var(b) = \sigma^2 (X'X)^{-1}}\ .$$

Importancia:

- Nos indica cual es el error esperado del estimador y por tanto si la aproximación es buena o no, el grado de precisión del estimador. Debe combinarse con la insesgadez.

Demostración

$$Var(b) = E\left\{(b-\beta)(b-\beta)'\right\} = E\left\{(X'X)^{-1}X'\varepsilon\varepsilon' X(X'X)^{-1}\right\} =$$

$$= (X'X)^{-1}X'E(\varepsilon\varepsilon')X(X'X)^{-1} = (X'X)^{-1}X'\sigma^2 IX(X'X)^{-1} =$$

$$= \sigma^2 (X'X)^{-1}X'X(X'X)^{-1} = \sigma^2 (X'X)^{-1}$$

Eficiencia de los estimadores de MCO

Universidade de Vigo

4) El estimador de mínimos cuadrados ordinarios es el mejor estimador lineal insesgado (ELIO), esto es, dicho estimador tiene varianza mínima dentro del conjunto de los estimadores lineales e insesgados.

Demostración

Por ser lineal

Por ser C una matriz no estocástica

Sea c un estimador lineal e insesgado cualquiera, debe verificar $E(c)=\beta$

Por otro lado $E(c)=E(Cy) = CE(y) = CX\beta$

Por la suposición de linealidad en el modelo

Por tanto $CX\beta=\beta \Rightarrow (CX-I)\beta=0 \Rightarrow CX-I=0 \Rightarrow \boxed{CX=I}$

$c=Cy=(D+(C-D))y= Dy+(C-D)y=b+Ly$

Pero $LD'=(C-D)D'= CX(X'X)^{-1} - (X'X)^{-1} X'X(X'X)^{-1}= (X'X)^{-1} - (X'X)^{-1} =0$

Luego $Var(c)= Var(Cy)=Var((D+L)y)= var(b)+var(Ly)$

Como la varianza de Ly siempre es positiva o nula se sigue que la varianza de c es mayor o igual que la de b

Importancia:

Estas cuatro propiedades se resumen en la última que se conoce como teorema de Gaus-Markov. Nos asegura que los estimadores MCO son los mejores dentro de todos los estimadores lineales e insesgados. Son los mejores en el sentido de que son los más precisos y están enfocados hacia el verdadero valor del parámetro

Valor estimado de la dependiente en los estimadores de MCO

Universidade de Vigo

5) El vector de valores estimados viene dado por

Y estimado= HY

siendo H=X(X'X)$^{-1}$X'. Por tanto los valores estimados de la y son una combinación lineal de las propias observaciones de la variable dependiente.

Importancia:

- Nos indica que el valor estimado de la dependiente es una combinación lineal de la dependiente

Demostración

$$\hat{y} = Xb = X(X'X)^{-1}X'y = Hy$$

Propiedades de los estimadores de MCO (3)

Universidade de Vigo

6) Los residuos vienen dados en forma matricial por

e = MY con M=I-H

Demostración

$$e = y - \hat{y} = y - Hy = (I - H)y = My$$

Importancia:

- Nos indica que el residuo también es una combinación lineal de la dependiente.

Comentarios

Universidade de Vigo

- Volviendo al modelo inicial, descomponemos la información de la variable dependiente en dos fuentes:
  - la que aportan las variables independientes
  - la que no aportan esas variables.

Información de las variables independientes

$$y = f(X_1, \ldots X_k) + g(X_{k+1}, \ldots)$$

$$\beta_1 X_1 + \ldots * \beta_k X_k + \beta_0 + \varepsilon = cte + error$$

Información residual

$$y = \hat{y} + e = Hy + My \Longrightarrow I = H + M$$

- La primera se recoge en el espacio de variables independientes mediante la matriz H. El resto se recoge en el espacio residual mediante la matriz M.

Conclusiones

Universidade de Vigo

- **Relaciones directas**
  - Y=Y estimado+e
- **Leyes de distribución**
  - Y es $N(X\beta,\sigma^2 I)$
  - Y estimado es $N(X\beta,\sigma^2 H)$
  - e es $N(0,\sigma^2 M)$
  - De los coeficientes: b es $N(\beta,\sigma^2(X'X)^{-1})$
  - De la varianza:

    $W = \frac{(T-k-1)S2_R}{\sigma^2}$ sigue una χ^2_{T-k-1}
  - De los t-estadísticos: $t_i=(b_i-\beta_i)/s_{bi}$ es una t_{T-k-1}
- **Relaciones entre valores esperados**
  - E(y)=E(y estimado)+E(e)

  = E(y estimado)
- **Relaciones entre varianzas**
  - Var(y)=Var(y estimado)+Var(e)
- **Relaciones entre sumas de cuadrados**
  - y'y=y estimado' y estimado+e'e
  - SCY=SCReg+SCE

Ejercicio 2.4

Notas y comentarios

Universidade de Vigo

- La suma de cuadrados de la variable dependiente centrada se descompone en dos sumandos:
  - la suma de cuadrados de los residuos
  - otra cantidad que denominaremos suma de cuadrados debida a la regresión
- **Representa la relación teórica que se verifica al descomponer la varianza de la y en suma de la varianza de los residuos y la varianza de la y estimada.**
- Esto nos va a permitir descomponer la variabilidad (el riesgo) de la variable dependiente en dos efectos:
  - el efecto debido a la regresión (información aportada por las variables independientes)
  - el efecto debido al error (información desconocida)
- **Esta descomposición se conoce como tabla ANOVA (ANalysis Of VAriance)**

Tabla ANOVA de la Regresión

respecto a la media

La variabilidad total proviene de dos fuentes: las variables independientes (regresión) y el error

FUENTE	SUMA DE CUADRADOS	DF	MEDIA DE CUADRADOS	E(CMe)	F VALUE	H_0
REGRESIÓN	$\hat{y}'\hat{y} = \hat{Y}'\hat{Y} - T\bar{Y}^2$	k	SCRegr/k	$\sigma_\varepsilon^2 + \frac{\beta_1'(X'X)\beta_1}{k}$	$\frac{\text{CMeRegr}}{\text{CME}}$	$\beta=0$
ERROR	e'e	T-k-1	SCE/(T-k-1)	σ_ε^2		
TOTAL	$y'y = Y'Y - T\bar{Y}^2$	T -1				

Comparación de estimadores: vale 1 si $\beta = 0$

Cuando $\beta = 0$ es insesgado

Estimadores de la varianza

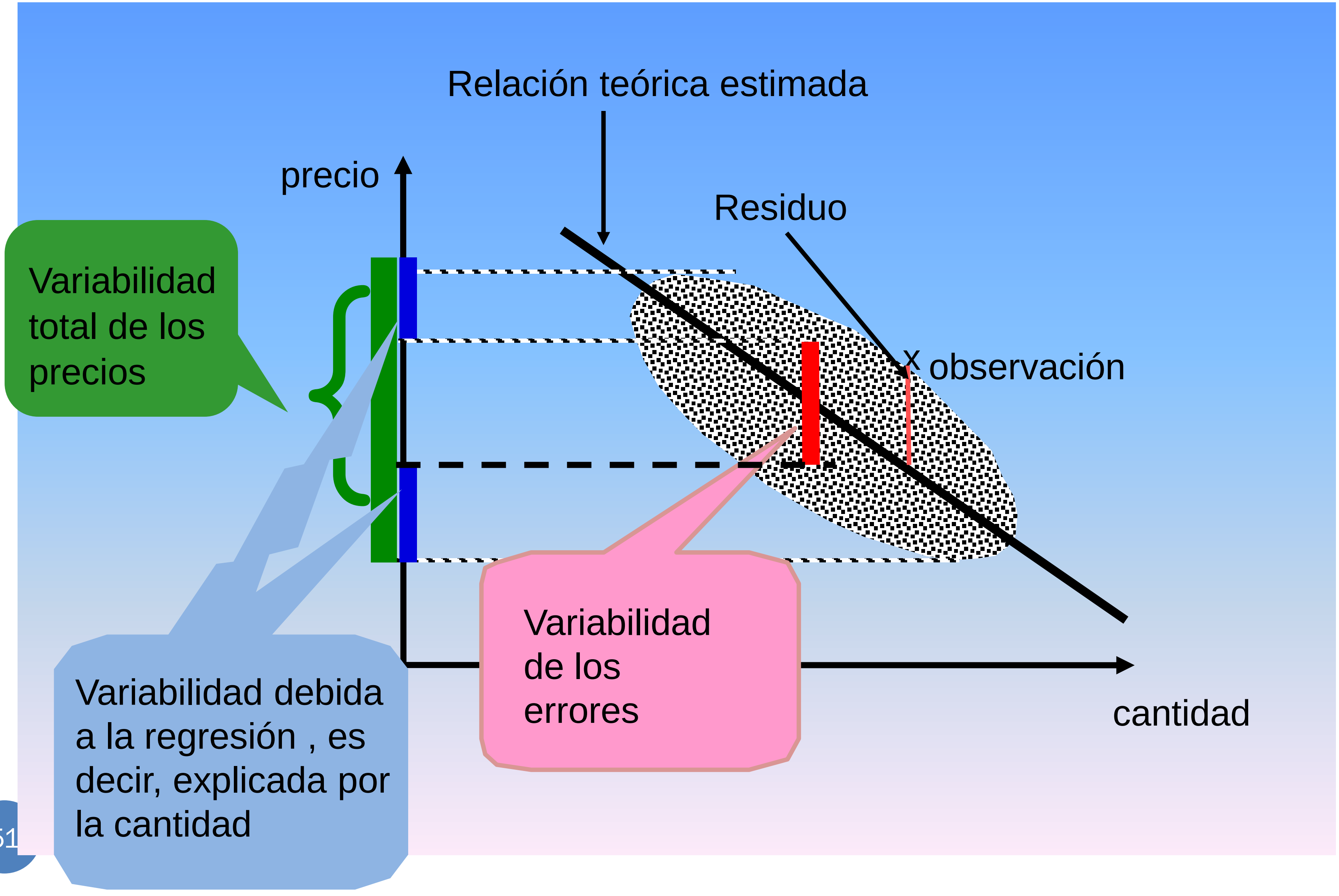
Estimador insesgado

El cociente entre ambos nos da el R2

$\frac{\text{Regresión} + \text{error}}{\text{Total}}$

Idea gráfica de mínimos cuadrados

Universidade de Vigo



Notas y comentarios (2)

Universidade de Vigo

- En la tabla ANOVA se obtiene un estadístico, la F, que es la ratio entre la varianza estimada por la regresión y la varianza estimada por los residuos. Cuanto mejor sea la regresión mayor valor tendrá esa ratio.
- No obstante para evaluar mejor la importancia de las variables independientes en la dependiente, se define, a partir de esta descomposición, una medida de la bondad de ajuste de la regresión, desde un punto de vista puramente geométrico, que denominaremos *Coeficiente de determinación* o R^2.
- Este coeficiente -multiplicado por 100- nos da el porcentaje de variación total de la variable Y - en el sentido de la suma de las desviaciones respecto de la media total - explicado por la regresión de la variable dependiente respecto de las independientes.

Propiedades del coeficiente de determinación

Universidade de Vigo

- Se define como la parte de variación de la variable dependiente no explicada por los residuos, coincidiendo con la parte explicada por la regresión, es decir:

$$R^2 = 1 - \frac{SCE}{SCY} = \frac{SC\,\mathrm{Re}\,g}{SCY}$$

- Coincide con el coeficiente de correlación entre la dependiente y su valor estimado al cuadrado. Por tanto nos indica el grado de relación lineal entre ambas variables.

Demostración: Como se verifica que

$$\sum_{t=1}^{T}(\hat{y}_t - \overline{y})(y_t - \overline{y}) = \sum_{t=1}^{T}(\hat{y}_t - \overline{y})(\hat{y}_t - \overline{y} + e_t) = \sum_{t=1}^{T}(\hat{y}_t - \overline{y})(\hat{y}_t - \overline{y}) + \sum_{t=1}^{T}(\hat{y}_t - \overline{y})e_t = SC\,\mathrm{Re}\,g + 0$$

$$Cor(\hat{y}, y)^2 = \frac{\left(\sum_{t=1}^{T}(\hat{y}_t - \overline{y})(y_t - \overline{y})\right)^2}{SCY * SC\,\mathrm{Re}\,g} = \frac{SC\,\mathrm{Re}\,g^2}{SCY * SC\,\mathrm{Re}\,g} = \frac{SC\,\mathrm{Re}\,g}{SCY} = R^2$$

Coeficiente de determinación en el origen

Universidade de Vigo

- Este coeficiente se usa cuando no hay constante en el modelo, pues se supone que la recta de regresión pasa por el origen.
- Al ser respecto al origen, ya no coincide con la variación explicada de la Y pues la media de la Y no tiene porque ser nula,, y los productos cruzados no coinciden con los productos cruzados centrados. Eso mismo ocurre con los residuos. La suma de cuadrados de los errores ya no aproxima la variación de los errores, pues estos ya no tienen porque tener media cero.
- La consecuencia es que en este caso el coeficiente de determinación no mide la parte de variación explicada por la regresión y se pueden definir diferentes medidas de ajuste.

Comparación entre los coeficientes de determinación

Universidade de Vigo

Definición general del coeficiente de determinación

$$R^2 = 1 - \frac{SCE}{SCY} = 1 - \frac{\sum_{t=1}^{T} e_t^2}{\sum_{t=1}^{T} (y_t - \bar{y})^2}$$

Cálculos relacionados con el coeficiente de determinación cuando no hay constante en la regresión

$$R^2_{Raw} = 1 - \frac{\sum_{t=1}^{T} e_t^2}{\sum_{t=1}^{T} y_t^2}$$

Se observa que tanto el $R2_{raw}$ como la correlación al cuadrado son distintos entre sí y del R^2 tradicional

$$Cor(\hat{y}, y)^2 = \frac{SC\operatorname{Re} g}{SCY} = \frac{\sum_{t=1}^{T} (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^{T} (y_t - \bar{y})^2} = 1 - \frac{\sum_{t=1}^{T} (e_t - \bar{e})^2}{\sum_{t=1}^{T} (y_t - \bar{y})^2} \neq R^2$$

Ejercicio 2.5

Salida del OLS en Gretl

Universidade de Vigo

- En Gretl el comando OLS es el que permite calcular los estimadores de MCO.
- Veamos como se interpreta cada uno de los elementos de sus salida estándar.
- Algunos no los vamos a usar en este curso
- Otros los hemos visto anteriormente
- Otros lo vemos a continuación
- El resto los veremos al final

```
Modelo 1: MCO, usando las observaciones 1960-1983 (T = 24)
Variable dependiente: Y

              Coeficiente    Desv. Típica   Estadístico t    Valor p
  ---------------------------------------------------------------------
  const        16.8978        3.33414           5.068        4.47e-05  ***
  X             0.825365      0.0375240        22.00         1.81e-016 ***

Media de la vble. dep.   89.61250   D.T. de la vble. dep.    9.953558
Suma de cuad. residuos   99.11090   D.T. de la regresión     2.122508
R-cuadrado               0.956505   R-cuadrado corregido     0.954528
F(1, 22)                 483.8081   Valor p (de F)           1.81e-16
Log-verosimilitud       -51.07275   Criterio de Akaike       106.1455
Criterio de Schwarz      108.5016   Crit. de Hannan-Quinn    106.7706
rho                      0.948311   Durbin-Watson            0.239765
```

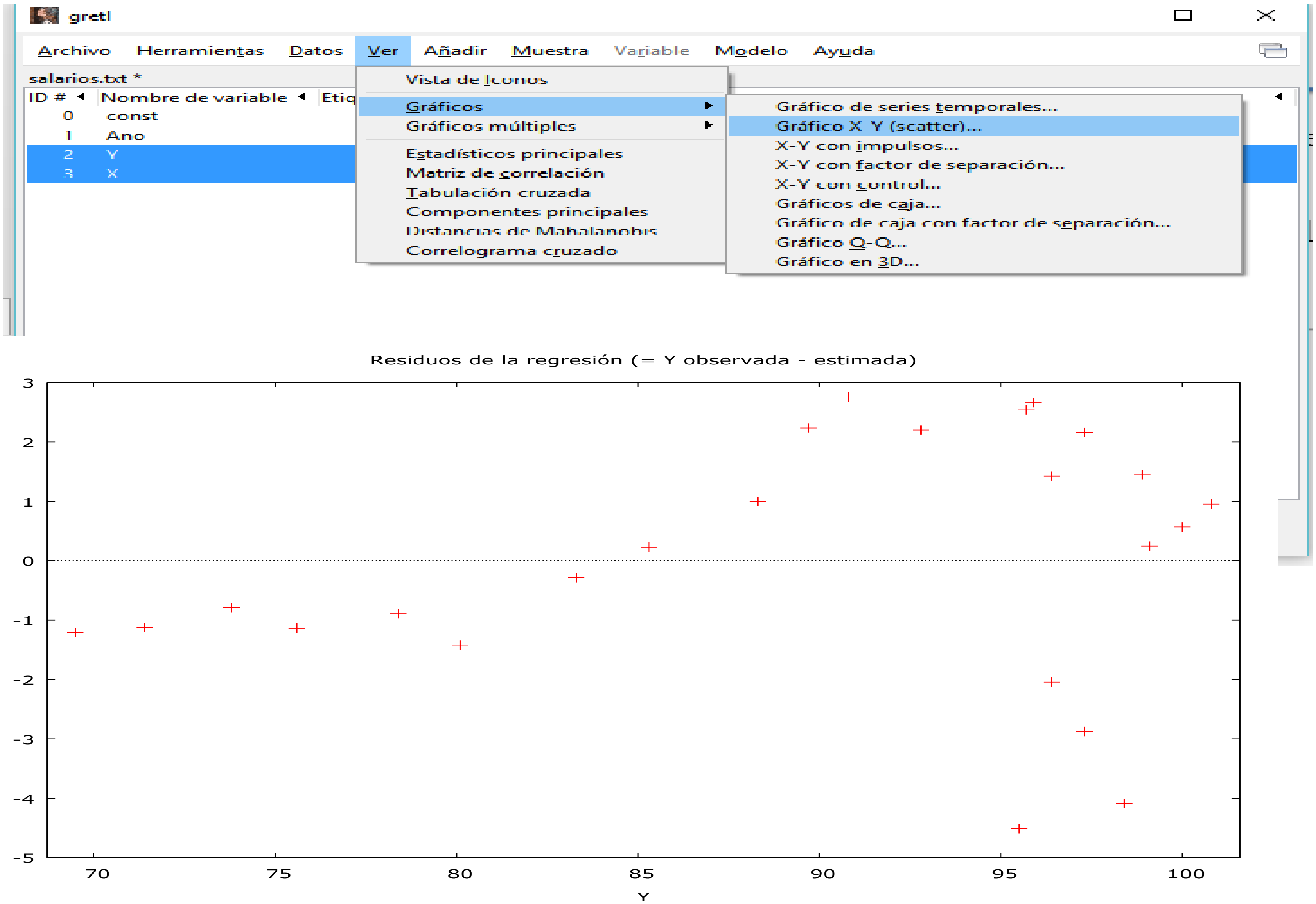
ANOVA

Universidade de Vigo

```
Análisis de Varianza:

                  Suma de cuadrados       gl  Media de cuadrados

  Regresión                 2179.58        1             2179.58
  Residuo                   99.1109       22             4.50504
  Total                     2278.69       23             99.0733

  R^2 = 2179.58 / 2278.69 = 0.956505
  F(1, 22) = 2179.58 / 4.50504 = 483.808 [Valor p 1.81e-016]
```

Gráficos

Universidade de Vigo

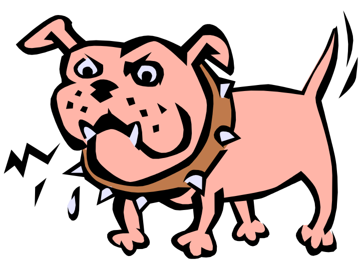


Universidade
de Vigo

Interpretación de los coeficientes que intervienen en la regresión

Parte C: Coeficientes de regresión, elasticidades y estadísticos, que nos ayudarán a interpretar desde el punto de vista económico los modelos de regresión

Universidade
de Vigo



Coeficientes de regresión parcial

Universidade de Vigo

- Son los parámetros β_i. Miden la variación unitaria de la variable dependiente respecto a cada variable independiente (aproximación de la derivada). Lo que denominaremos Efectividad.
- Se estiman directamente por b_i. Dependerán del número de variables que se incluyan en el modelo.
- Tienen unidades de medida que dependen de la variable dependiente y de cada independiente

Ejemplo 1

Universidade de Vigo

- Queremos medir el incremento que una campaña de publicidad durante los tres meses previos produce en las ventas actuales (medida en miles de euros). El modelo sería

Ventas promedio sin publicidad

$$Ventas = \beta_0 + \beta_1 GastosPub + \varepsilon$$

Incremento de las ventas por cada euro gastado en publicidad

Supongamos que el modelo estimado fuera el siguiente

$$Ventas = 20 + 0{,}4 GastosPub + e$$

El promedio de ventas sin realizar la campaña publicitaria es de 20.000 euros

Por cada euro invertido en publicidad se incrementan 40 céntimos las ventas

Ejemplo 2

Universidade de Vigo

- Queremos saber si realmente la experiencia laboral en un mismo puesto incrementa la productividad (medida en unidades fabricadas por hora de trabajo). El modelo sería

$$Productividad = \beta_0 + \beta_1 Añosdeexperiencia + \varepsilon$$

Productividad independiente de la experiencia

Incremento de las productividad por cada año de experiencia

Supongamos que el modelo estimado fuera el siguiente

$$Productividad = 1{,}2 + 2\,Añosdeexperiencia + e$$

El promedio de productividad que tendría una persona sin experiencia laboral sería de 1,2 unidades por hora

Por cada año de experiencia se incrementa la productividad en 2 unidades por hora

Ejemplo 3

Universidade de Vigo

- La relación entre la demanda de Kg de café de 45 familias, en el año 1996, junto con los precios por Kg en euros que había en los distintos lugares y la renta de que disponían en euros muestran el siguiente modelo:

$$Demanda = \beta_0 + \beta_1 precios + \beta_2 renta + \varepsilon$$

- Las estimaciones de MCO dan los siguientes resultados

Variable Name	Estimated Coefficient	Standard Error	T-Ratio	p-Value (42 DF)
Y	0.48484	0.3669E-01	13.21	0.000
P	-0.41519	0.5565E-01	-7.461	0.000
CONSTANT	2.0209	2.214	0.9128	0.367

Efecto renta en la demanda. Por cada € de renta adicional la demanda de café se incrementa en 0,48Kg

Efecto precio en la demanda. Por cada € que aumente el precio, la demanda d café disminuye en 0,41Kg

Demanda básica independientemente de la renta y el precio. Independientemente del precio y de la renta, la demanda de café es de 2 Kg

La interpretación de los coeficientes siempre depende de las unidades de medida

Errores estándar de los estimadores de los coeficientes de regresión

Universidade de Vigo

- Se calcula como la raíz cuadrada de la diagonal de la matriz de varianzas covarianzas de los estimadores, o sea, $\sigma^2(X'X)^{-1}$. Si llamamos q_i a los elementos de la diagonal de $(X'X)^{-1}$ el error estándar sería $\sigma(q_i)^{1/2}$
- Miden la fiabilidad de los estimadores de los coeficientes de regresión, en el sentido de decirnos si están cercanos a la media o alejados de esta, hablando probabilísticamente.
- Sus estimadores vienen dados por $s(q_i)^{1/2}$ siendo s la raíz cuadrada de la varianza residual. Por tanto verificarán la mismas propiedades que s.

Ejemplo 3 (continuación)

Universidade de Vigo

En la salida del ejemplo 3 teníamos que:

Variable Name	Estimated Coefficient	Standard Error	T-Ratio	p-Value (42 DF)
Y	0.48484	0.3669E-01	13.21	0.000
P	-0.41519	0.5565E-01	-7.461	0.000
CONSTANT	2.0209	2.214	0.9128	0.367

Indica el error de la estimación del efecto renta sobre la demanda

Indica el error de la estimación del efecto precio sobre la demanda

Indica el error de la estimación de la demanda promedio básica

Los dos primeros valores son muy pequeños lo que indica que esos estimadores son muy precisos

T-estadísticos

Universidade de Vigo

- Miden la distancia de los estimadores de los coeficientes de regresión al origen, en el sentido de decirnos si están cercanos o alejados de 0, y por consiguiente, si la variables tiene importancia en el modelo o no.
- Es una distancia estandarizada, por tanto se obtiene como la diferencia entre el estimador y 0 estandarizando por la desviación estándar del estimador

$$t_i=(b_i-0)/\ s(q_i)^{1/2} = b_i\ /\ s(q_i)^{1/2}$$

- Nos mide la precisión relativa, al ser un cociente entre el valor y su error. Nos dice si el error es muy grande en relación al valor.
- También es una forma de medir la importancia estadística de una variable en el modelo. Hasta que punto se puede afirmar que el coeficiente es diferente de 0.

Ejemplo 3 (continuación)

Universidade de Vigo

En la salida del ejemplo 3 teníamos que:

Variable Name	Estimated Coefficient	Standard Error	T-Ratio	p-Value (42 DF)
Y	0.48484	0.3669E-01	13.21	0.000
P	-0.41519	0.5565E-01	-7.461	0.000
CONSTANT	2.0209	2.214	0.9128	0.367

Indica la precisión relativa de la estimación del efecto renta sobre la demanda

Indica la precisión relativa de la estimación del efecto precio sobre la demanda

Indica la precisión relativa de la estimación de la demanda promedio básica

Los dos primeros valores son muy grandes lo que indica que esos estimadores son muy precisos y relativamente importantes

Cola de probabilidad

Universidade de Vigo

- Para entender la cola de probabilidad o p-valor se parte de que son ciertas todas las suposiciones del modelo. Además se incluye una hipótesis relativa a la importancia de la variable. Se parte de que el coeficiente teórico es nulo, es decir no tiene importancia en absoluto.
- En ese caso, el p-valor mide la probabilidad teórica de que ocurran los datos observados o algo más extraño bajo esas hipótesis. Por tanto:
  - Si el p-valor es muy pequeño quiere decir que es muy raro ese comportamiento de los datos bajo la hipótesis de importancia nula; eso indica que la variable es importante.
  - Si el p-valor entra dentro de lo normal, quiere decir que los datos son coherentes con la hipótesis de que la variable no es importante.

Ejemplo 3 (continuación)

Universidade de Vigo

En la salida del ejemplo 3 teníamos que:

Variable Name	Estimated Coefficient	Standard Error	T-Ratio	p-Value (42 DF)
Y	0.48484	0.3669E-01	13.21	0.000
P	-0.41519	0.5565E-01	-7.461	0.000
CONSTANT	2.0209	2.214	0.9128	0.367

Es muy poco probable que se den estos datos si el efecto de la renta sobre la demanda es nulo

Es muy poco probable que se den estos datos si el efecto del precio sobre la demanda es nulo

Es algo probable que se den estos datos si el efecto la demanda promedio básica es nula

Los dos primeros valores son muy pequeños lo que indica que esas variables son muy importantes en el modelo

Ejercicio 2.7

Estimación de máxima verosimilitud

Parte D: Intervalos y test de hipótesis

Universidade de Vigo

Universidade de Vigo

La información muestral en la estimación de una función económica

Universidade de Vigo

- La información muestral se recoge en la función de verosimilitud.
- Dicha función nos dice el grado de verosimilitud que tiene un parámetro para una muestra dada.
- Un parámetro es mas verosímil que otro -dada una muestra- si le asigna una probabilidad mayor de ocurrir a la muestra que el otro.
- La función de verosimilitud nos mide la probabilidad (o densidad) de observar unos datos, para un conjunto de parámetros dados. Es función del parámetro para unos datos concretos, normalmente los que hemos observado.

Función de log-verosimilitud en el MRLC

Universidade de Vigo

- Aplicando el logaritmo obtendremos la función de log verosimilitud

$$\ln L(y_{1,} \ldots y_T ; \beta ; \sigma^2) =$$

Es el error o perturbación

$$= -\frac{T}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta) =$$

$$= -\frac{T}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}SCE$$

- Depende de los parámetros y de la suma de cuadrados de las perturbaciones.
- Maximizar la función para β es equivalente a minimizar el termino último que está restando y coincide con la SCE.
- Por tanto los estimadores MCO y MV coinciden ene el MRLC

Estimadores de máxima verosimilitud de σ^2

Universidade de Vigo

- El estimador de máxima verosimilitud se obtendría igualando la derivada de la función de log verosimilitud a 0, por consiguiente

$$-\frac{T}{2}\frac{1}{\sigma^2}+\frac{1}{2\sigma^4}(y-X\beta)'(y-X\beta)=0$$

$$-\frac{\sigma^2}{2}+\frac{1}{2T}(y-X\beta)'(y-X\beta)=0$$

$$\hat{\sigma}^2=\frac{1}{T}(y-X\hat{\beta})'(y-X\hat{\beta})=\frac{1}{T}e'e$$

Luego

$$\hat{\sigma}^2=\frac{\sum_{t=1}^{T} e_t^2}{T}$$

- Que es similar al de MCO pero algo diferente pues se divide por T en vez de por T-k-1

Propiedades de los estimadores

1. Los estimadores de máxima verosimilitud de los coeficientes, bajo normalidad, son insesgados, lineales y óptimos (ELIO)
2. Alcanzan la cota de Cramer-Rao, por lo que serán insesgados y de mínima varianza y consistentes.
3. El estimador de máxima verosimilitud de la varianza de la perturbación "σ^2 " es sesgado, y por consiguiente no alcanza la cota de Cramer-Rao, por lo que no será eficiente. No obstante sigue siendo consistente.
4. Cuando se conoce el valor de los coeficientes el estimador de la varianza también alcanza la cota de Cramer-Rao.

Ley de distribución de los estimadores

Universidade de Vigo

- b es $N(\beta,\sigma^2(X'X)^{-1})$
- El estadístico

$$t_i=(b_i-\beta_i)/s_{bi} ,$$

- sigue una t de Student con T-k-1 grados de libertad
- El estadístico

$$W=\frac{(T-k-1)S_R^2}{\sigma^2}$$

Sigue una ji cuadrado con T-k-1 grados de libertad

Matriz de varianzas covarianzas

$$I^*(\hat{\beta},\hat{\sigma}^2) = \begin{pmatrix} \dfrac{-X'X}{\hat{\sigma}^2} & 0 \\ 0 & \dfrac{T}{2\hat{\sigma}^4} \end{pmatrix}$$

- Por tanto, los estimadores MV de β y σ^2 son asintóticamente incorrelados, lo que nos va a permitir estimar cada uno de ellos por su lado.

Intervalos de confianza de los coeficientes

Universidade de Vigo

- Se llama **intervalo de confianza** a un par de números entre los cuales se estima que estará cierto valor desconocido con una determinada probabilidad de acierto. Formalmente, estos números determinan un intervalo, que se calcula a partir de datos de una muestra, y el valor desconocido es un parámetro poblacional. La probabilidad de éxito en la estimación se representa con $1 - \alpha$ y se denomina *nivel de confianza*.
- Intervalo de confianza de coeficientes individuales

$$(b_i - t_{T-k-1,\alpha/2} s_{bi}, b_i + t_{T-k-1,\alpha/2} s_{bi})$$

- **Propiedades**
  - R1. Cuanto mayor sea el error de la estimación, menos precisión tendremos. Si S_R aumenta entonces aumenta la amplitud del intervalo
  - R2. Al aumentar el tamaño muestral, aumentamos la precisión. Al aumentar T la t de Student se aproxima a la normal tipificada, que tiene valores críticos menores.

Ejemplo de la demanda de café

Universidade de Vigo

Intervalo de confianza para el efecto de la renta sobre la demanda de café

Partiendo de la información del OLS

Variable Name	Estimated Coefficient	Standard Error
Y	0.17939	0.3967E-01
P	0.71556E-01	0.3067
CONSTANT	41.740	7.649

Calculamos los valores $(b_i - t_{T-k-1,\alpha/2} s_{bi}, b_i + t_{T-k-1,\alpha/2} s_{bi})$, obteniendo:

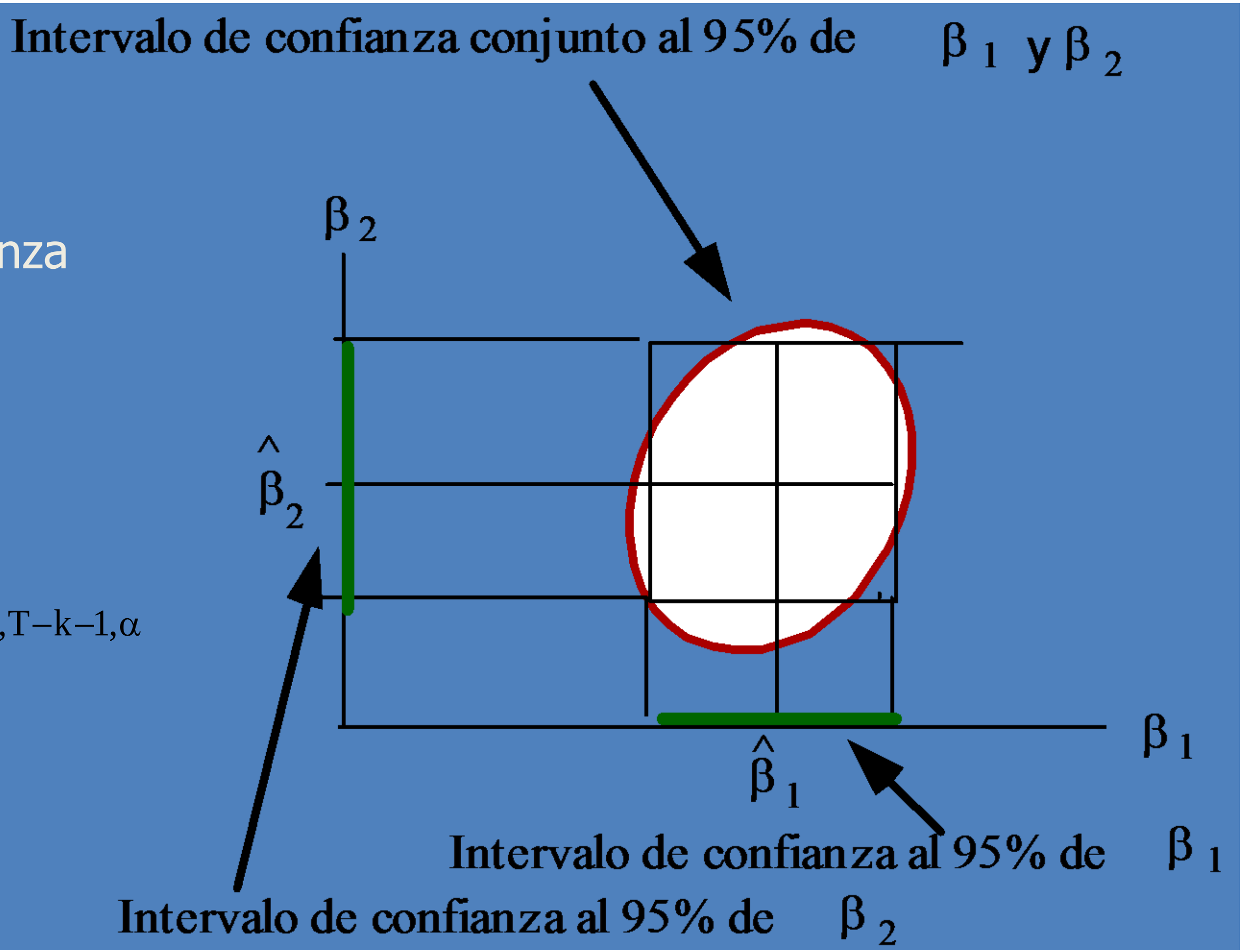
```
USING 95% CONFIDENCE INTERVALS

CONFIDENCE INTERVALS BASED ON T-DISTRIBUTION WITH  42 D.F.
     - T CRITICAL VALUES =   2.014 AND   1.680
NAME          LOWER 2.5%           COEFFICIENT     UPPER 2.5%   STD. ERROR
Y            0.9949E-01           0.17939        0.2593            0.040
```

Ejemplo gráfico de test conjunto

Universidade de Vigo

Estando la región de confianza definida por

$$\frac{\frac{1}{2}\begin{pmatrix} b_i & b_j \end{pmatrix}' Q_{ij}^{-1}\begin{pmatrix} b_i \\ b_j \end{pmatrix}}{S_R^2} < F_{2,T-k-1,\alpha}$$

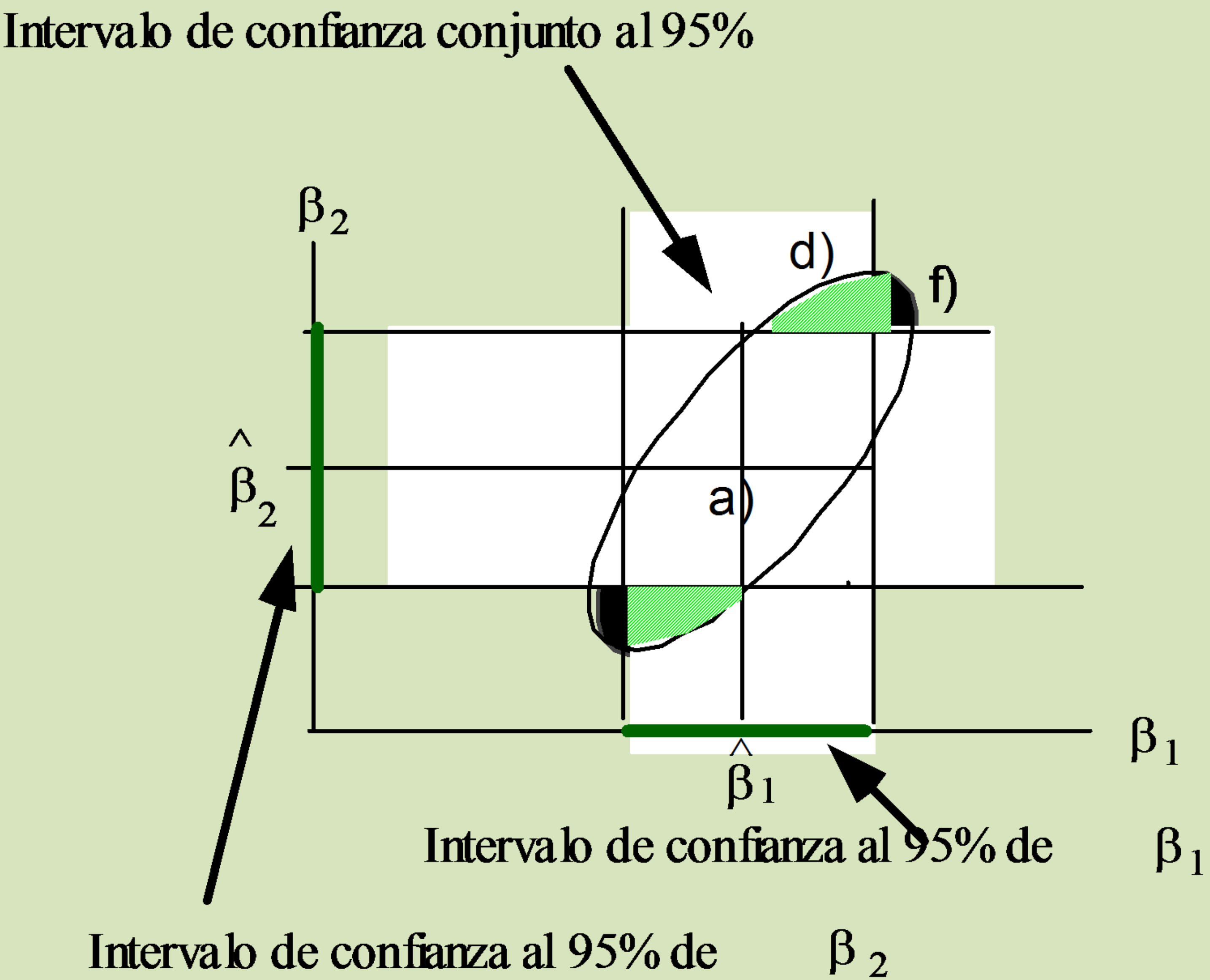


Casos posibles

- **a)** (β_1,β_2) pertenecen al área blanca del elipsoide, que se incluye en los intervalos unidimensionales y en el conjunto.
- **b)** (β_1,β_2) pertenecen al área rayada horizontalmente por lo que pertenece al intervalo conjunto y a uno de los unidimensionales pero no a los dos.
- **c)** (β_1,β_2) pertenecen al área oscura por lo que pertenece a ambos intervalos individuales pero no al conjunto. Corresponde a las esquinas del cuadrado que no están en el círculo.
- **d)** (β_1,β_2) pertenece al área rayada inclinada, por lo que no estará ni en los individuales ni en el conjunto.
- **e)** (β_1,β_2) pertenece al área blanca fuera del elipsoide, por lo que no estará en el intervalo conjunto, pero sí en alguno de los individuales.
- **f)** Es imposible en el caso de ortogonalidad que (β_1,β_2) pertenezca al intervalo conjunto y no se pertenezca a ninguno de los dos individuales. Sin embargo en el caso de no ortogonalidad si se puede dar ese caso en la zona negra.

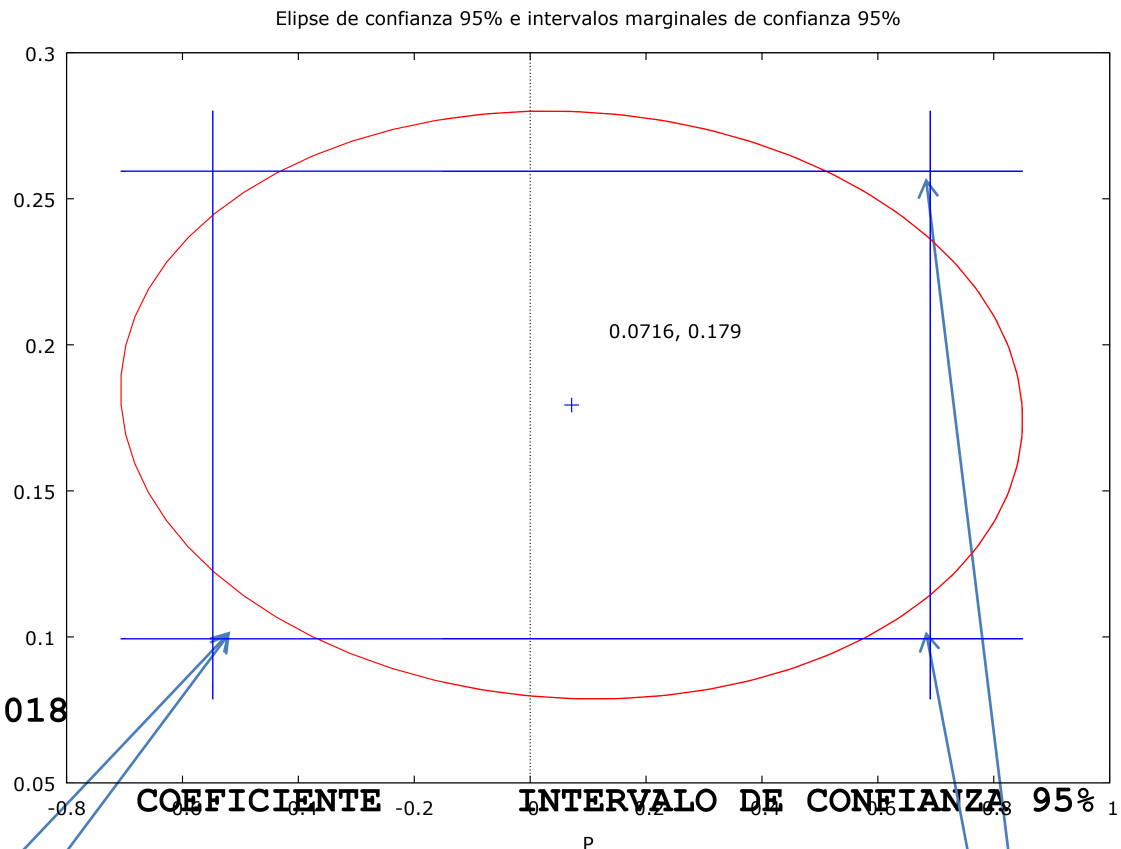
NO ORTOGONALIDAD

Universidade de Vigo



Ejemplo de demanda de café

Universidade de Vigo



t(42, 0.025) = 2.018

VARIABLE	COEFICIENTE	INTERVALO DE CONFIANZA 95%	
const	41.7402	26.3043	57.1761
P	0.0715555	-0.547466	0.690577
RTA	0.179393	0.0993291	0.259457

Ejercicio 2.9

Test de hipótesis

Universidade de Vigo

- Tratan de comprobar si una determinada hipótesis es aceptable o no por unos determinados datos bajo una serie de suposiciones previas.
- Son la clave de la demostración empírica científica y por ello son necesarios en las ciencias sociales.
- En todos los test de significación se tienen en cuenta los siguientes aspectos:
  1. Definir modelo de análisis e indicar suposiciones del test
  2. Definir hipótesis nula y alternativa
  3. Fijar el nivel de significación
  4. Estadístico de la prueba
  5. Ley de distribución del estadístico
  6. Regla de decisión

Modelo y suposiciones

- Para poder contrastar una determinada hipótesis es necesario presuponer un cierto comportamiento de los datos, puesto que el proceso de generación de estos siempre está determinado por los factores que los condicionan.
  - Cuando ese modelo está determinado por un número finito de parámetros, se dice paramétrico. En otro caso es no paramétrico.
- Estas suposiciones son previas, eso quiere decir que no se contrastan en el test y por consiguiente los resultados de ese test están condicionados a la validez de las suposiciones.
  - Cuando el fallo de esas suposiciones inhabilitan totalmente el test se dice que es poco robusto, en otro caso se dice que es robusto.
- Los test que vamos a ver son dentro del MRLN, por lo que siempre se partirá de que se verifican las suposiciones del MRLN, por consiguiente no las especificaremos de modo concreto, pero deben ser conocidas.

Hipótesis

Universidade de Vigo

- En todos los test de significación se contrastan dos hipótesis: la nula y la alternativa.
  - La hipótesis nula es aquella que se presupone inicialmente y los datos deben comprobar si se rechaza o no hay argumentos suficientes para ello. Eso indica que es una hipótesis conservadora en el sentido que se mantiene mientras no se demuestre lo contrario.
  - La hipótesis alternativa es la que tienen valor probatorio en el sentido de que dice que los datos rechazan claramente la hipótesis nula. Indica cual es la línea de fallo de la hipótesis nula y por consiguiente la que le da capacidad de discriminación al test. Por ese motivo es importante elegir bien la hipótesis alternativa para tener más capacidad de rechazar cuando haya que hacerlo. Esto es lo que se denomina potencia del test.
- Normalmente la hipótesis nula es un caso particular de la alternativa. En ese caso el test se denomina anidado. En otro caso es no anidado.

El nivel de significación

Universidade de Vigo

- Indica un limite a la probabilidad de cometer un error de tipo I. Para ello debemos definir previamente los diferentes tipos de errores, tal como se hace en la tabla siguiente, según sea cierta o no determinada hipótesis.

	H0 es cierta	*H1 es cierta*
Decido aceptar *H0*	***BIEN***	***ERROR II***
Decido aceptar *H1*	***ERROR I***	***BIEN***

- La idea es buscar la mínima probabilidad de cometer un error de tipo II, una vez fijado un límite a la probabilidad de cometer un error de tipo de I.
- La elección del nivel de significación es subjetiva, por lo que se suele elegir entre el 5% o el 10% en ciencias sociales, mientras que en las naturales se suele optar por el 1% o el 5%.

Estadístico de prueba

Universidade de Vigo

- Consiste en formalizar la idea intuitiva del test, plasmando en una función de los datos muestrales (estadístico) que nos dé información sobre como discriminar entre cada una de las hipótesis.
- Normalmente este estadístico tiene un comportamiento bajo la hipótesis nula y otro bajo la alternativa y su resultado nos va a permitir decidirnos por una u otra de las dos hipótesis según con cual sea mas coherente.
- Para definirlo es conveniente conocer cual es la idea intuitiva del test, es decir, la forma en la que comparamos las hipótesis con los datos en el test.

Ley de distribución del estadístico

Universidade de Vigo

- Para elaborar las reglas de decisión del test es conveniente conocer cual es la ley que sigue el estadístico de prueba bajo la hipótesis nula.
- Dicha ley se determina a partir de las suposiciones de partida y debe ser independiente de las hipótesis o por lo menos conocida si la hipótesis nula se verifica.
- A partir de esa ley de distribución se pueden definir dos regiones en el espacio muestral:
  - la región de aceptación que es aquella en la que se espera que caiga el estadístico cuando la hipótesis nula es cierta
  - el resto de los valores muestrales que será la región de rechazo

Regla de decisión

Universidade de Vigo

- Una vez determinada la ley simplemente se construye la regla de decisión que siempre suele ser del mismo tipo:
  - Si el estadístico muestral cae en la región de aceptación se acepta la hipótesis nula, en caso contrario se rechaza.
- La región de rechazo nos indica aquellos valores muestrales que hacen improbable (con una probabilidad menor que el nivel de significación) que sea cierta la hipótesis nula. Es decir, si la hipótesis nula fuera cierta la probabilidad que ocurra lo que ocurrió o algo mas distante es prácticamente cero, por consiguiente supongo que la hipótesis nula no es consistente con los datos, esto significa que la rechazo.

Test de hipótesis de los coeficientes individuales

Universidade de Vigo

- Suposiciones: (Las del MRLN)
- Hipótesis
    - H_0: $\beta_i = \beta$
    - H_1: $\beta_i \neq \beta$
- Nivel de significación (prefijado α)
- Estadístico de prueba
    - $t_i = (b_i - \beta)/s_{bi}$

El impacto de la variable X_i sobre la dependiente es β

El impacto de la variable X_i sobre la dependiente no es β

Mide la distancia estandarizada entre el estimador del impacto de la variable X_i y el valor de la hipótesis nula, dado por β

Por tanto la idea intuitiva del test es que si esa distancia es muy grande en valor absoluto, se rechaza puesto que el estimador que teóricamente es insesgado debería estar cerca del parámetro. En caso contrario se acepta.
La ley de distribución nos permite decir cuando es muy grande o no.

Test de hipótesis de los coeficientes individuales (2)

Universidade de Vigo

- Ley de distribución
  - Ya vimos que b_i sigue una ley $N(\beta_i,\sigma^2 X^{ii})$ donde X^{ii} es el elemento i-ésimo de la diagonal de $(X'X)^{-1}$. Por lo tanto, el estadístico t seguirá una t de Student con T-k-1 grados de libertad
- Regla de decisión

Este valor, que se deduce de la ley de distribución, es el que nos permite afirmar si la distancia es muy grande (significativa) o no

$$\text{rechazamos } H_0 \text{ si } |t_i| > t_{T-k-1,\alpha/2},$$

- O lo que es equivalente:

rechazamos H_0 si la cola de probabilidad o p-Valor $<\alpha$,

Cuando la probabilidad de observar los datos o algo más raro bajo H_0 (suponiendo que es cierta) es muy pequeña (menor que el riesgo que admito α). Rechazo si bajo esa hipótesis los datos son muy raros

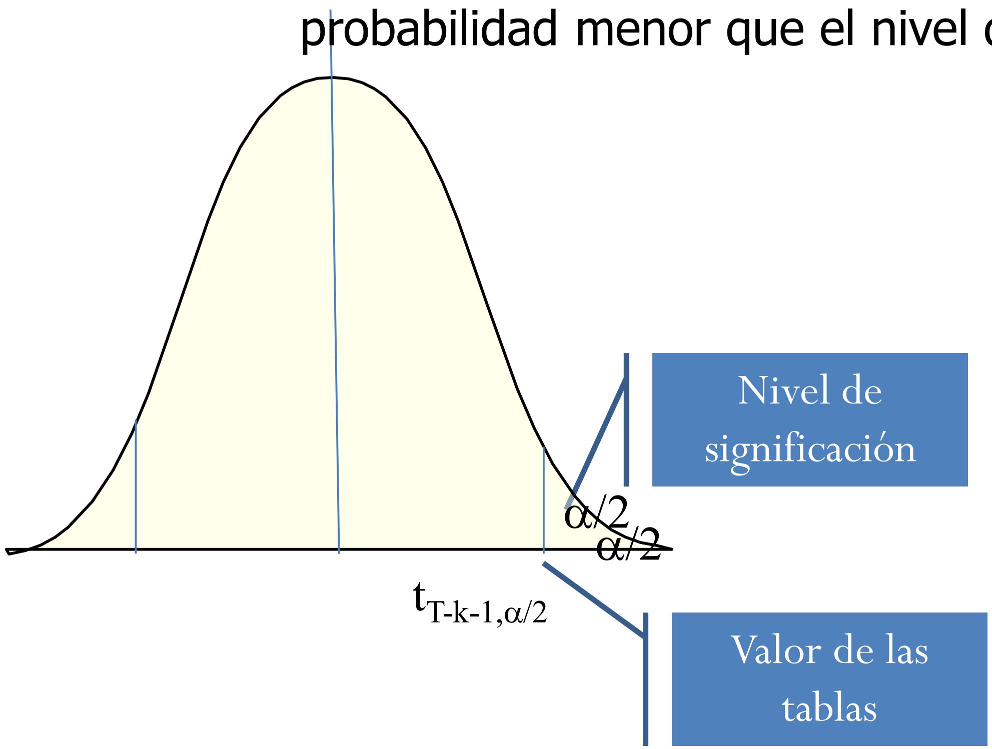
Comparación entre las reglas de decisión

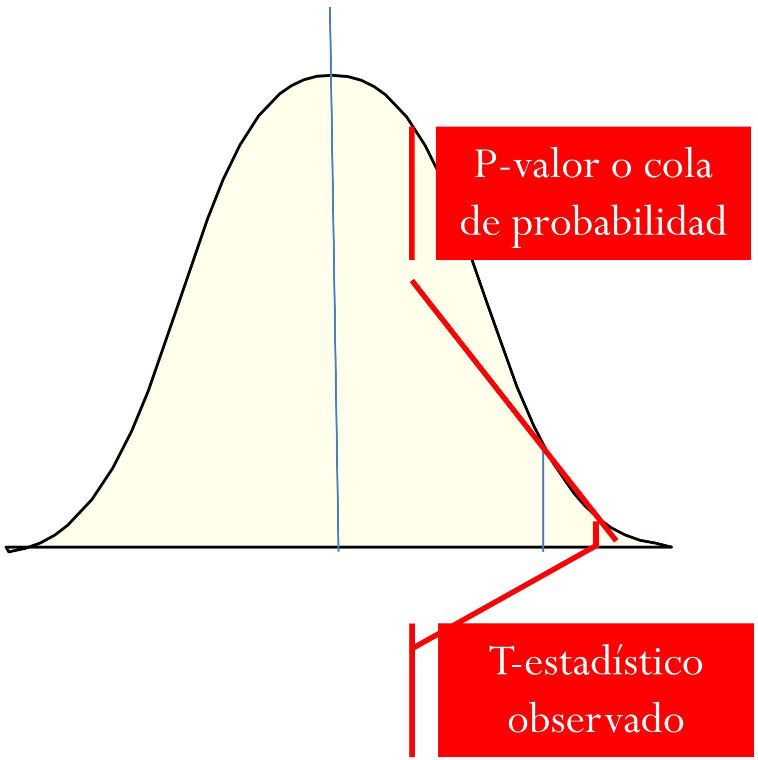
Universidade de Vigo

Regla de decisión

Observaciones

T-estadístico mayor que las tablas o cola de probabilidad menor que el nivel de significación





Resultados

Universidade de Vigo

- R1.Cuanto mayor sea el error de la estimación, menor es la potencia del test. Si S_R aumenta, aumenta la varianza y disminuye la potencia.
- R2. Al aumentar el tamaño muestral, aumentamos la potencia. Al aumentar T la t de Student se aproxima a la normal tipificada, que tiene valores críticos menores, y por lo tanto la potencia va ser mayor.
- R3. Es fácil observar que existe una relación directa entre potencia del t-test y precisión del intervalo de confianza.

Ejemplo de la demanda de café

Universidade de Vigo

- Para contrastar si el efecto de la renta es nulo, debemos hacer uso del t-estadístico y la cola de probabilidad que aparece en la salida:

Variable Name	Estimated Coefficient	Standard Error	T-Ratio	p-Value (42 DF)
P	0.71556E-01	0.3067	0.2333	0.817
RTA	0.17939	0.3967E-01	4.522	0.000
CONSTANT	41.740	7.649	5.457	0.000

- Valor de la $t_{42,0,025} = 2.0181$
- En este caso el efecto de la renta y el de la constante es significativo, mientras que el del precio no.

Se puede comparar este valor con el de las tablas

O este otro con el nivel de significación

Test de hipótesis de los coeficientes conjuntos

Universidade de Vigo

- Hipótesis

  H_0: $\beta_i = 0$ para todos los $i=i_1,..,i_j$

  H_1: $\beta_i \neq 0$ para alguno de los $i=i_1,..,i_j$
- Nivel de significación (prefijado α)
- Suposiciones: (Las del MRLN)
- Estadístico de prueba:
  - Se hace una regresión con todas las variables y se calcula la suma de los errores (SSE_k)
  - Se hace una regresión con todas las variables menos aquellas que suponemos nulas en H_0 y se calcula la suma de los errores al cuadrado (SSE_{1j}).
  - El estadístico es

$$F_{i_{1j}} = \frac{\dfrac{SSE_{1j} - SSE_k}{j}}{\dfrac{SSE_k}{T-k-1}}$$

Test de hipótesis de los coeficientes conjuntos (2)

Universidade de Vigo

- Ley de distribución
  - El estadístico anterior seguirá una F de Snedecor con j y T-k-1 grados de libertad respectivamente
- Regla de decisión

$$\text{rechazamos } H_0 \text{ si } F > F_{j,T\text{-}k\text{-}1,\alpha/2}$$

- El test mas común es el de nulidad conjunta de todos los coeficientes cuyos estadísticos salen en la tabla ANOVA

Test conjuntos de todos los coeficientes en SHAZAM

Universidade de Vigo

- Cuando el test es de todos los coeficientes, se puede calcular directamente en la tabla ANOVA.

Análisis de Varianza:

	Suma de cuadrados	gl	Media de cuadrados
Regresión	71.0269	2	35.5134
Residuo	144.369	42	3.43737
Total	215.396	44	4.89537

```
R^2 = 71.0269 / 215.396 = 0.329750
F(2, 42) = 35.5134 / 3.43737 = 10.3316 [Valor p 0.0002]
```

Nos permite contrastar si todos los coeficientes de las variables son nulos o alguno no

Test de hipótesis de restricciones unidimensionales

Universidade de Vigo

- Suposiciones: (Las del MRLN)
- Hipótesis

$$H_0 : a'\beta = r$$

$$H_1 : a'\beta \neq r$$

- Nivel de significación (prefijado α)
- Estadístico de prueba: t

$$t_a = \frac{\mathbf{a}'\mathbf{b} - r}{S_R \sqrt{\mathbf{a}'(X'X)^{-1}\mathbf{a}}}$$

- Ley de distribución
  - El estadístico anterior seguirá una t de Student con T-k-1 grados de libertad.
- Regla de decisión
  - Rechazamos H_0 si $|t| > t_{,T-k-1,\alpha/2}$

Ejemplo de la demanda de café:

Universidade de Vigo

- Queremos contrastar que la suma de los efectos de precio y renta es la unidad.
- La hipótesis nula es

$$H_0 : \beta_1+\beta_2=1$$

- contra la alternativa

$$H_1 : \beta_1+\beta_2\neq 1$$

- que escrito en forma matricial sería

$$H_0 : \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = 1$$

Ejemplo en GRETL

Universidade de Vigo

- En GRETL se escriben los coeficientes con el numero de orden

```
Restricción:
 b[RTA] + b[P] = 1Estimaciones restringidas:
Estadístico de contraste: F(1, 42) = 5.9427, con valor p = 0.0190915
```

- Por lo que se rechaza la hipótesis nula de que el efecto conjunto es unitario.
- También muestra las estimaciones restringidas

```
             Coeficiente   Desv. Típica   Estadístico t    Valor p
  ---------------------------------------------------------------
  const        23.1641       0.698639         33.16        2.94e-032 ***
  P             0.813067     0.0417638        19.47        6.03e-023 ***
  RTA           0.186933     0.0417638         4.476       5.51e-05  ***
  Desviación típica de la regresión = 1.95767
```

Ejercicio 2.10