

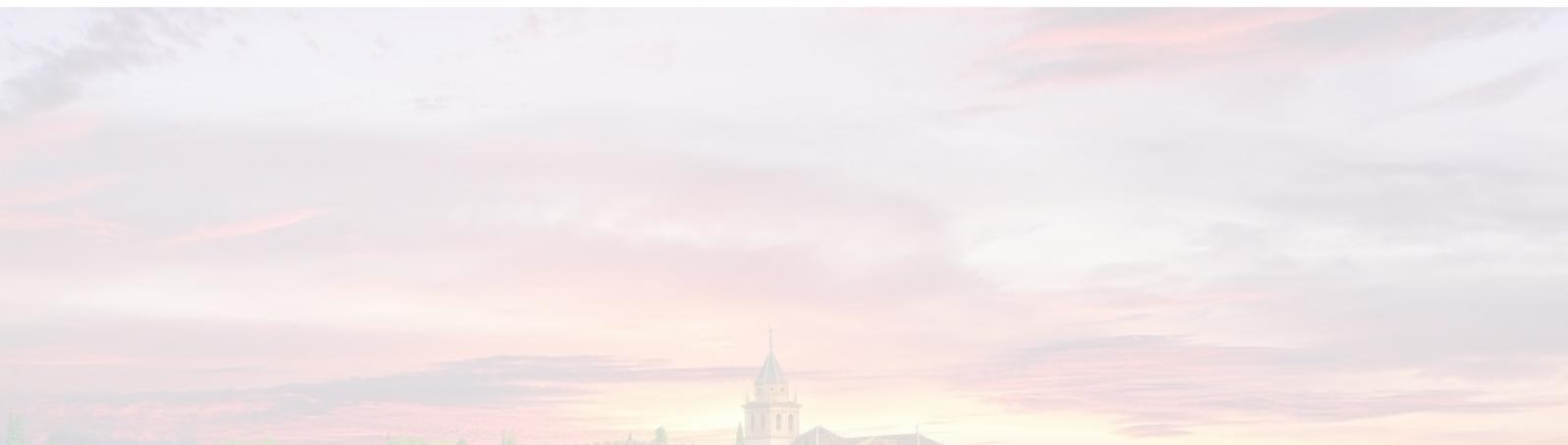


# Ingeniería Informática + ADE

Universidad de Granada (UGR)

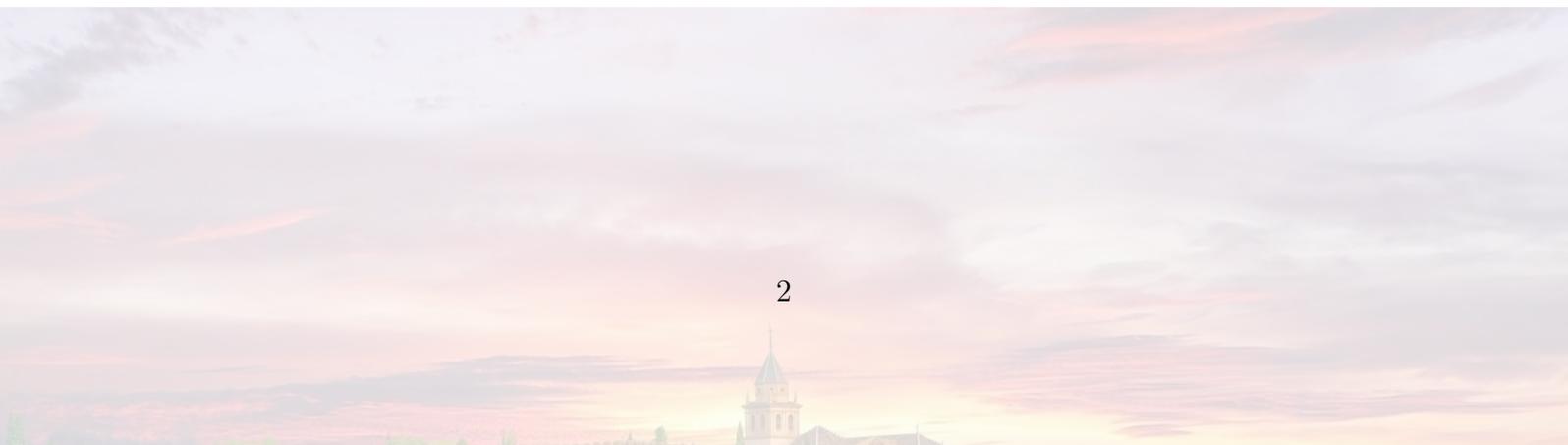
**Autor:** Ismael Sallami Moreno

**Asignatura:** Econometría

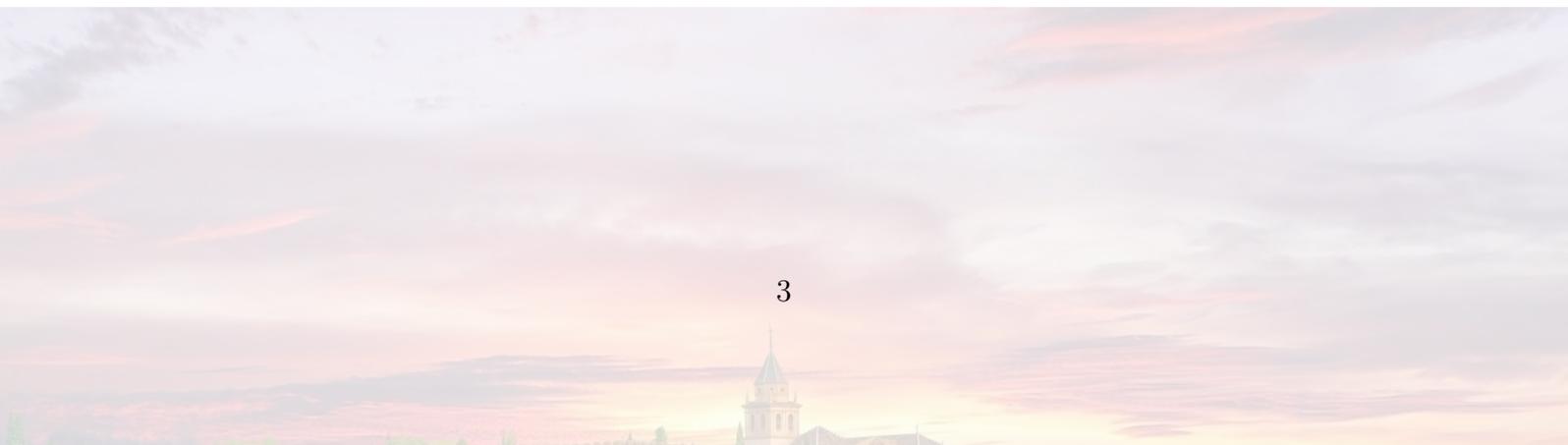


# Índice

<b>1. Tema 1</b>	<b>3</b>
<b>2. Tema 2</b>	<b>43</b>
<b>3. Tema 3</b>	<b>103</b>
3.1. Apendice Variables Ficticias . . . . .	152
<b>4. Tema 4</b>	<b>157</b>
<b>5. Tema 5</b>	<b>175</b>
<b>6. Tema 6</b>	<b>227</b>
<b>7. Referencias</b>	<b>258</b>



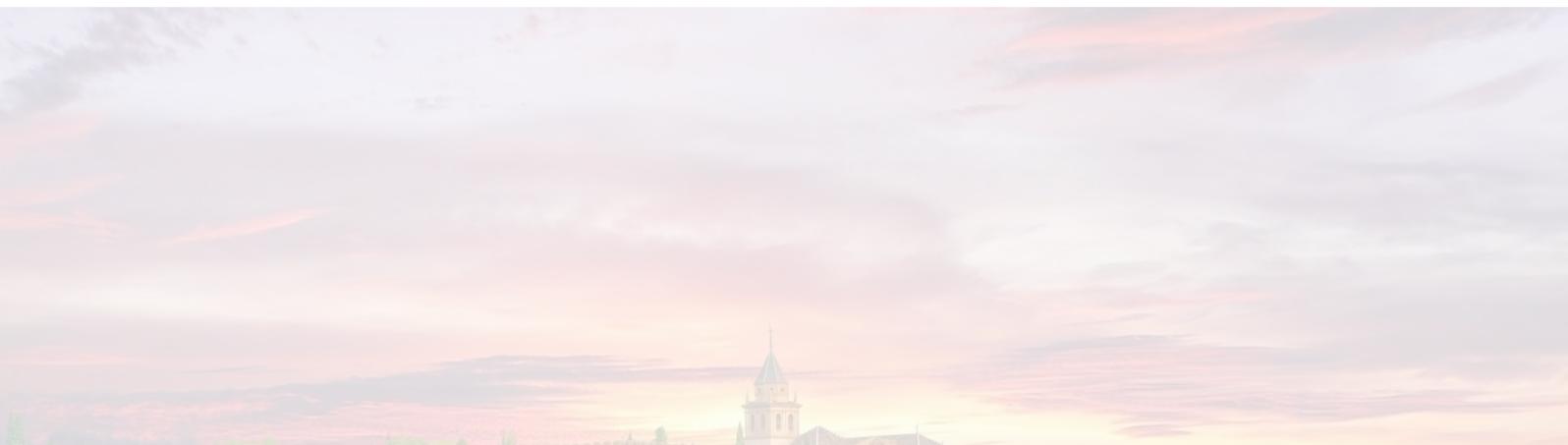
# 1 Tema 1



# ECONOMETRÍA

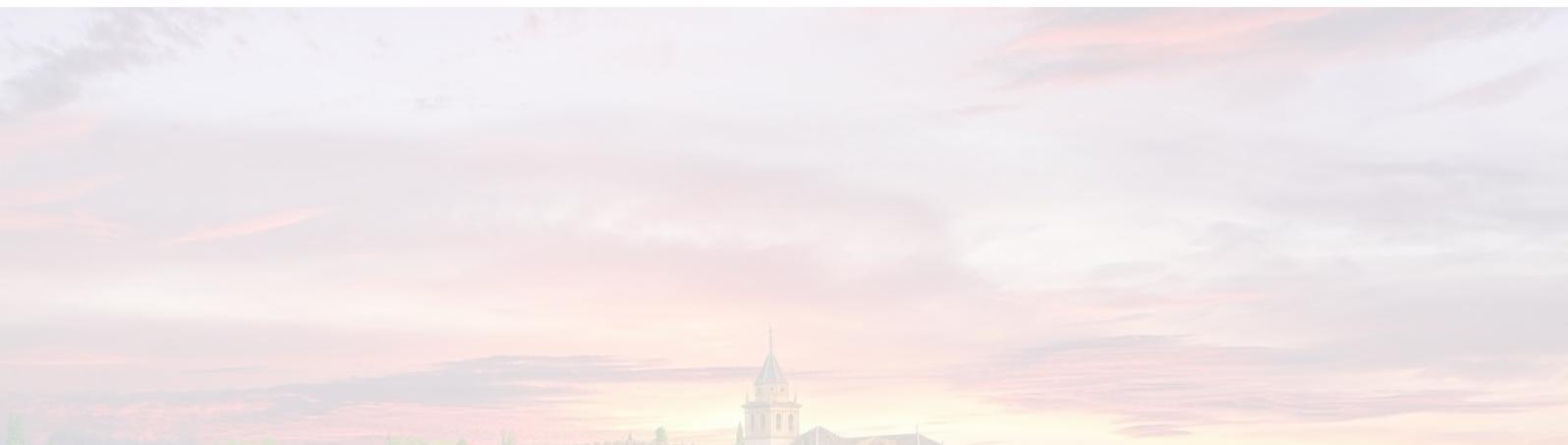
## TEMA 1: INTRODUCCIÓN A LA ECONOMETRÍA

2024-2025



# INTRODUCCIÓN A LA ECONOMETRÍA

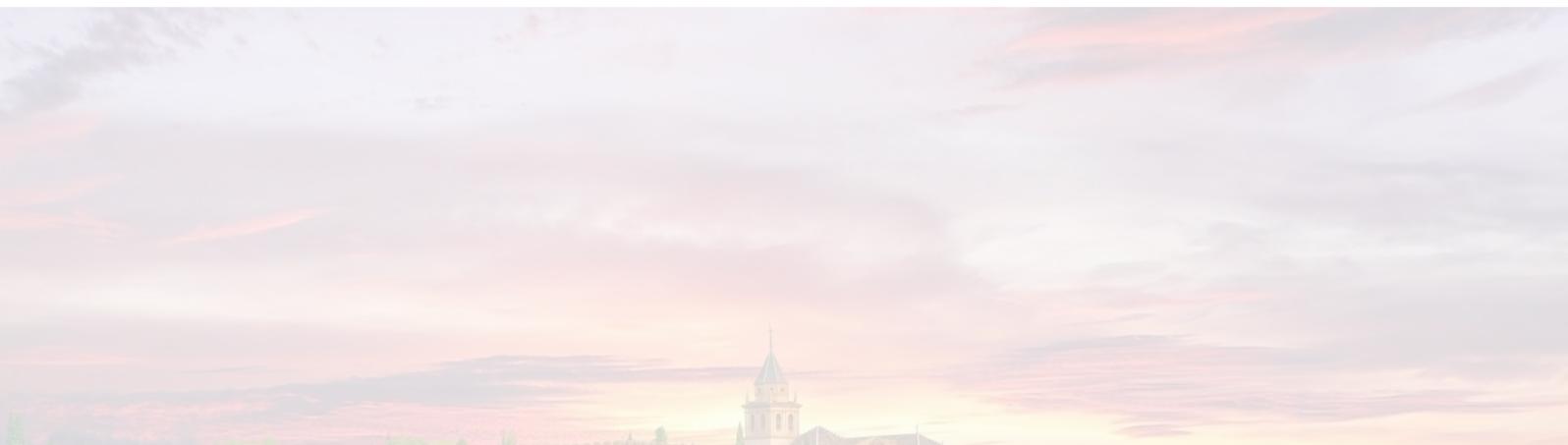
<b>1</b>	<b>El método econométrico. . . . .</b>	<b>2</b>
<b>2</b>	<b>Modelos económicos y modelos econométricos. . . . .</b>	<b>6</b>
<b>3</b>	<b>Fases del Método Econométrico . . . . .</b>	<b>8</b>
<b>4</b>	<b>Componentes de un modelo econométrico. . . . .</b>	<b>10</b>
<b>5</b>	<b>Diversas formas de expresar un Modelo Econométrico . . . . .</b>	<b>13</b>
<b>6</b>	<b>Naturaleza de la información utilizada en Econometría . . . . .</b>	<b>16</b>



## EL MÉTODO ECONOMETRICO

El método científico econométrico está construido sobre un proceso de razonamiento lógico utilizando una información inicial y un conocimiento empírico. El método está estructurado en:

**REALIDAD** El estudio de los hechos económicos que surgen de la misma sociedad.

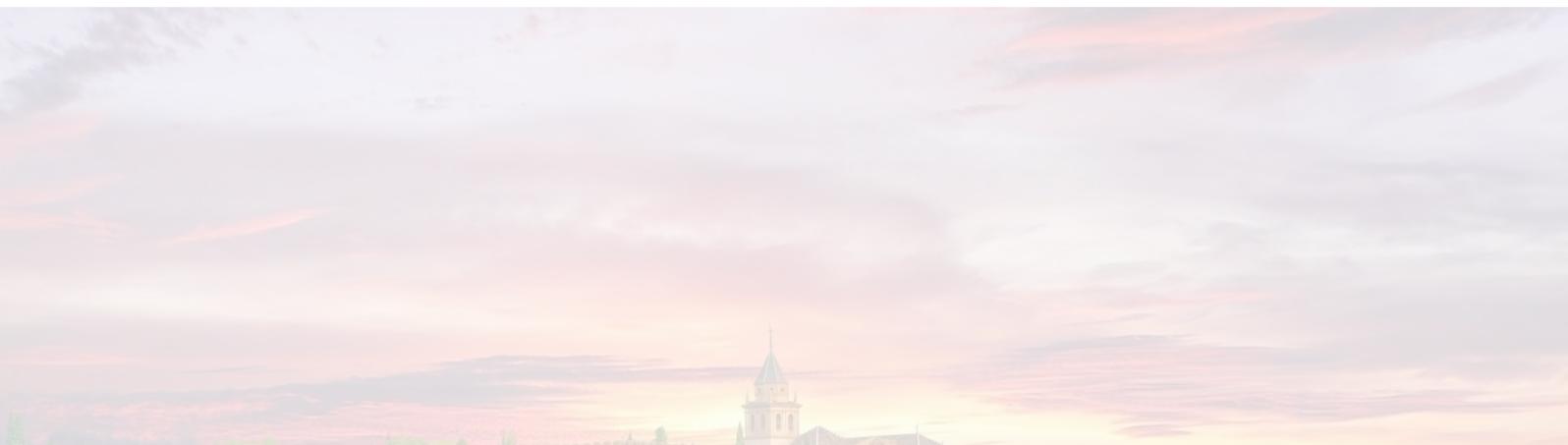


## EL MÉTODO ECONOMETRICO

El método científico econométrico está construido sobre un proceso de razonamiento lógico utilizando una información inicial y un conocimiento empírico. El método está estructurado en:

**REALIDAD** El estudio de los hechos económicos que surgen de la misma sociedad.

**DATOS** Recopilación de datos obtenidos del conocimiento que el investigador tiene sobre la propia realidad.



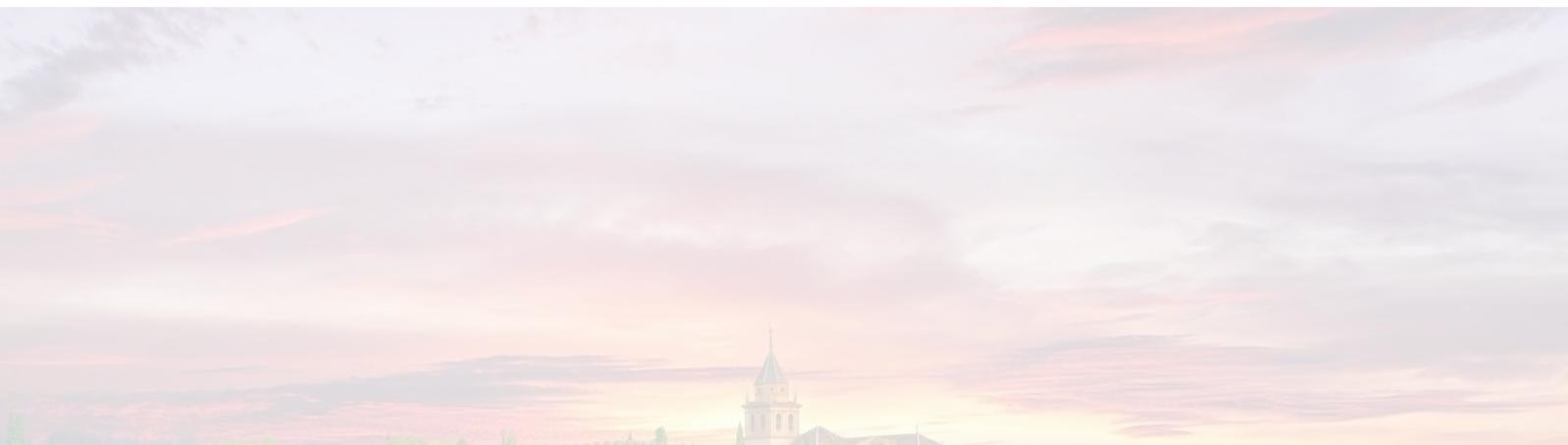
## EL MÉTODO ECONOMETRICO

El método científico econométrico está construido sobre un proceso de razonamiento lógico utilizando una información inicial y un conocimiento empírico. El método está estructurado en:

**REALIDAD** El estudio de los hechos económicos que surgen de la misma sociedad.

**DATOS** Recopilación de datos obtenidos del conocimiento que el investigador tiene sobre la propia realidad.

**MODELO** A partir de la base de datos, el investigador formula ciertas hipótesis dando lugar a la construcción de modelos, teorías y leyes.



## EL MÉTODO ECONOMETRICO

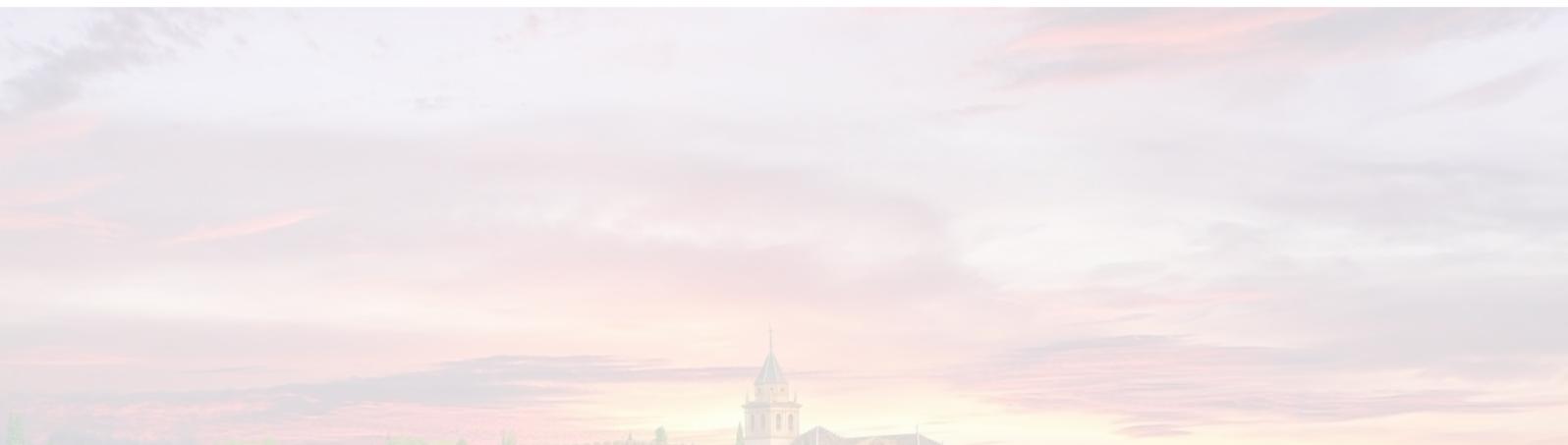
El método científico econométrico está construido sobre un proceso de razonamiento lógico utilizando una información inicial y un conocimiento empírico. El método está estructurado en:

**REALIDAD** El estudio de los hechos económicos que surgen de la misma sociedad.

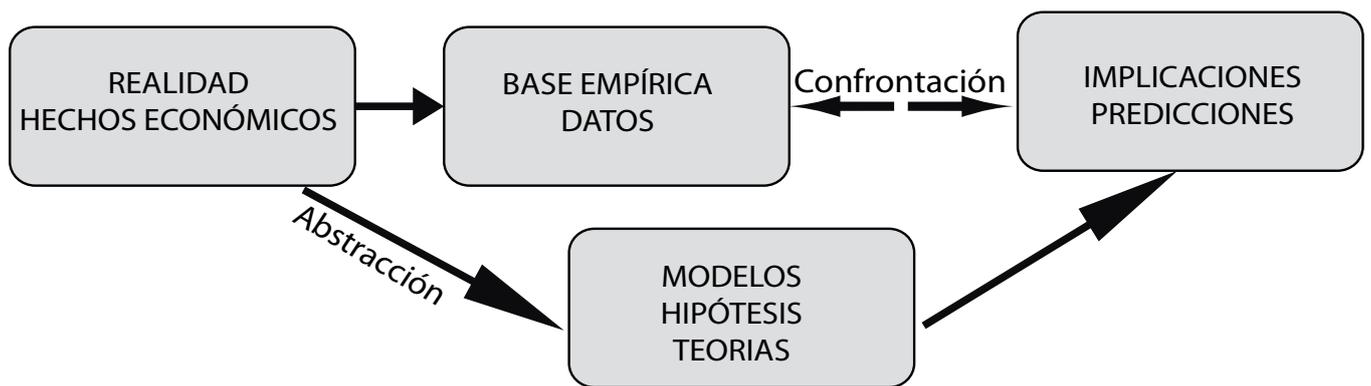
**DATOS** Recopilación de datos obtenidos del conocimiento que el investigador tiene sobre la propia realidad.

**MODELO** A partir de la base de datos, el investigador formula ciertas hipótesis dando lugar a la construcción de modelos, teorías y leyes.

**PREDICCIONES** Los modelos descritos permite al investigador predecir determinados comportamientos del sistema.

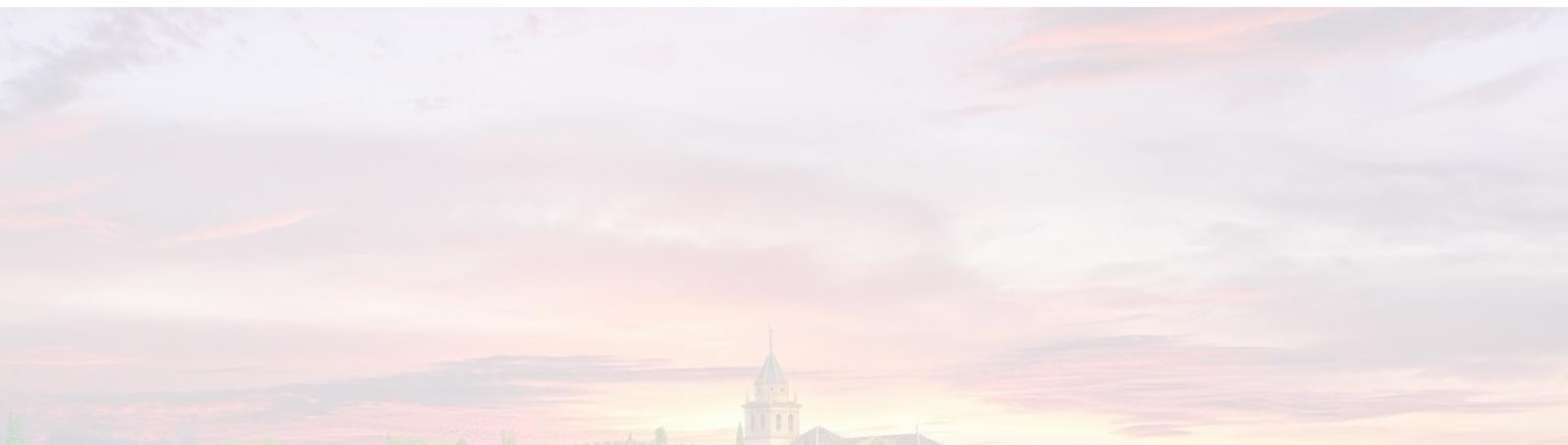


# EL MÉTODO ECONOMETRICO



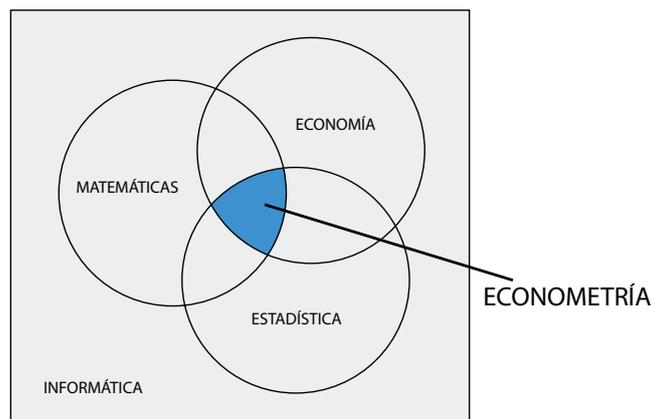
# CONCEPTOS DE ECONOMETRÍA

- Frisch (1933)** La Econometría no es lo mismo que Estadística Económica, tampoco es lo mismo que la Teoría Económica General, aunque ésta tenga un carácter cuantitativo fundamental. Tampoco debe ser considerada la Econometría como un sinónimo de la aplicación de las matemáticas a la economía. La experiencia ha demostrado que cada uno de estos tres puntos de vista son necesarios pero ninguno por sí mismo suficiente para entender las relaciones cuantitativas de la vida económica moderna. Es la unificación de los tres lo que es poderoso. Y es esa unificación lo que constituye la Econometría.
- Tintner (1952)** Indica que la Econometría consiste en la aplicación de la teoría económica, matemática y de los métodos estadísticos a los datos económicos, para establecer resultados numéricos en el campo de la Economía y verificar los teoremas económicos.
- Gollnick (1973)** El análisis econométrico combina la teoría económica, las matemáticas y la estadística para cuantificar y verificar las relaciones entre las variables económicas.



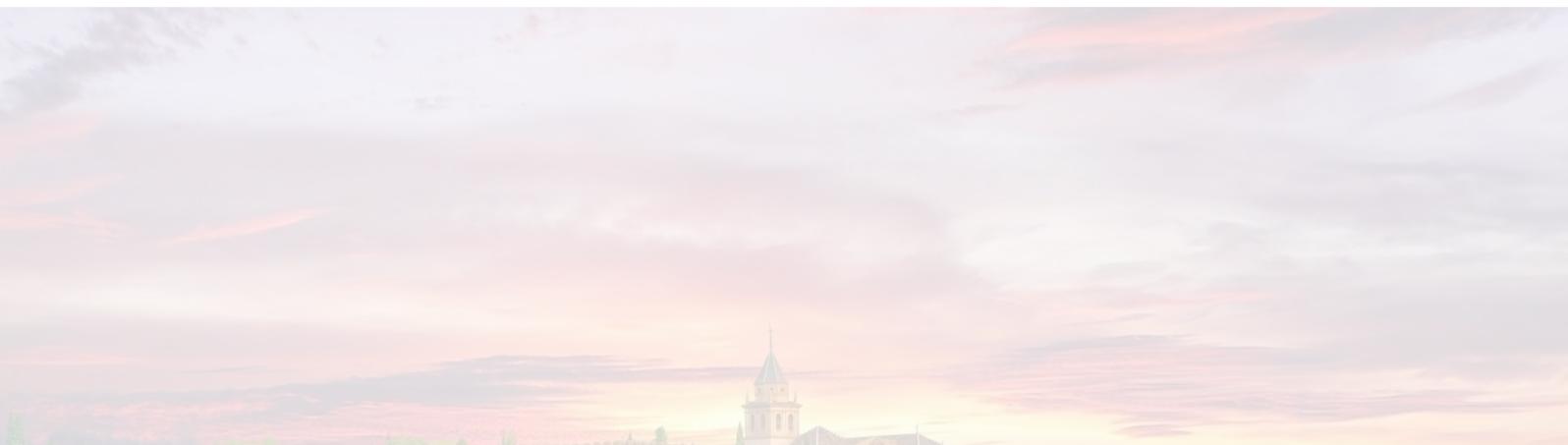
## CONCEPTO DE ECONOMETRÍA

La **Econometría** es la ciencia que utiliza la Teoría Económica, las Matemáticas y la Inferencia Estadística de forma conjunta utilizando como herramienta la Informática.



# MODELOS ECONÓMICOS Y ECONOMETRÍCOS

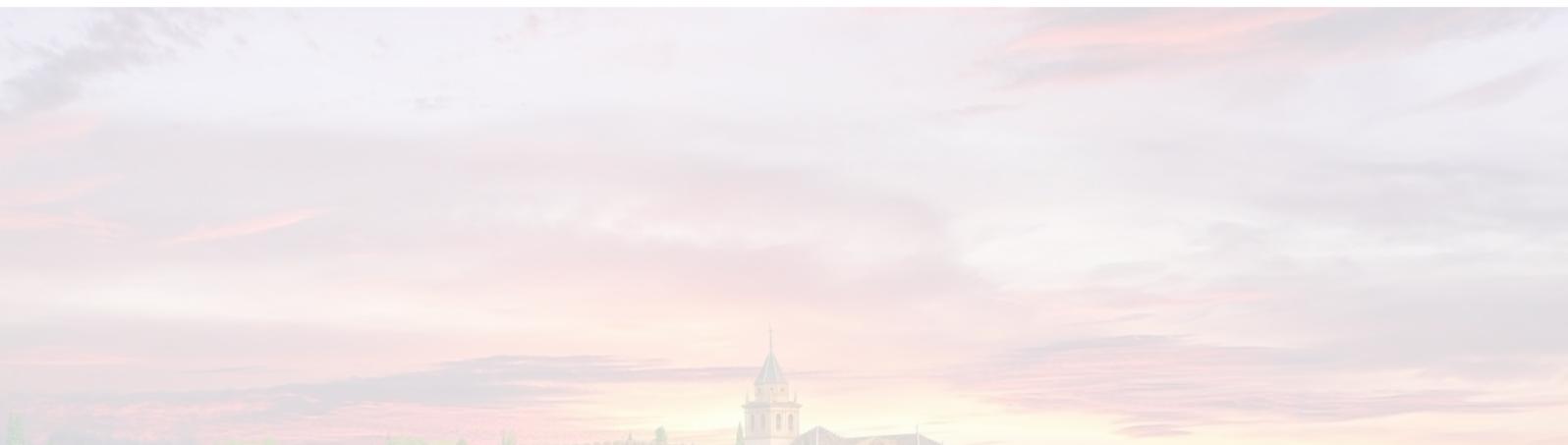
**Modelos Económicos** Los modelos económicos se basan en la representación algebraica simplificada de una realidad económica. Generalmente vienen descritos por ecuaciones matemáticas donde se relacionan distintos tipos de variables.



# MODELOS ECONÓMICOS Y ECONOMETRICOS

**Modelos Económicos** Los modelos económicos se basan en la representación algebraica simplificada de una realidad económica. Generalmente vienen descritos por ecuaciones matemáticas donde se relacionan distintos tipos de variables.

**Modelos Econométricos** Los modelos econométricos son aquellos modelos económicos que contienen todos los elementos necesarios para realizar su aplicación práctica. Es decir, un modelo econométrico es un un modelo económico en el que se ha descrito matemáticamente la relación entre una o varias variables explicadas en función a una serie de variables explicativas y un término aleatorio (perturbación).

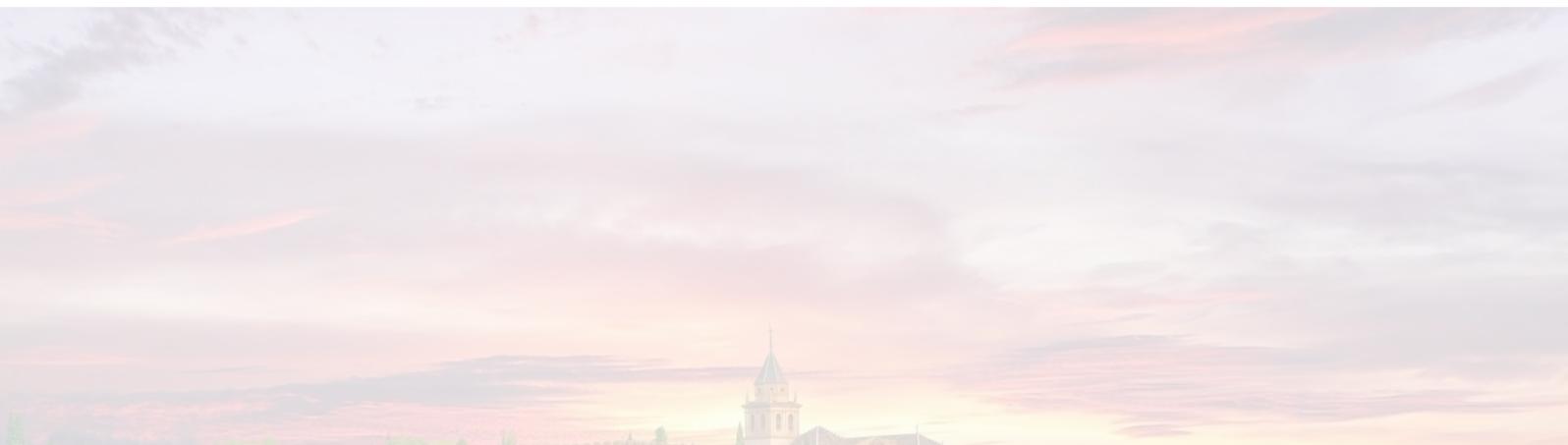


# MODELOS ECONÓMICOS Y ECONOMETRÍCOS

## Ejemplo 1

*Por ejemplo, un modelo económico que explique el consumo de energía eléctrica en España (CEE) en función del PIB vendría dada por la expresión*

$$CEE = f(PIB)$$



# MODELOS ECONÓMICOS Y ECONOMETRICOS

## Ejemplo 1

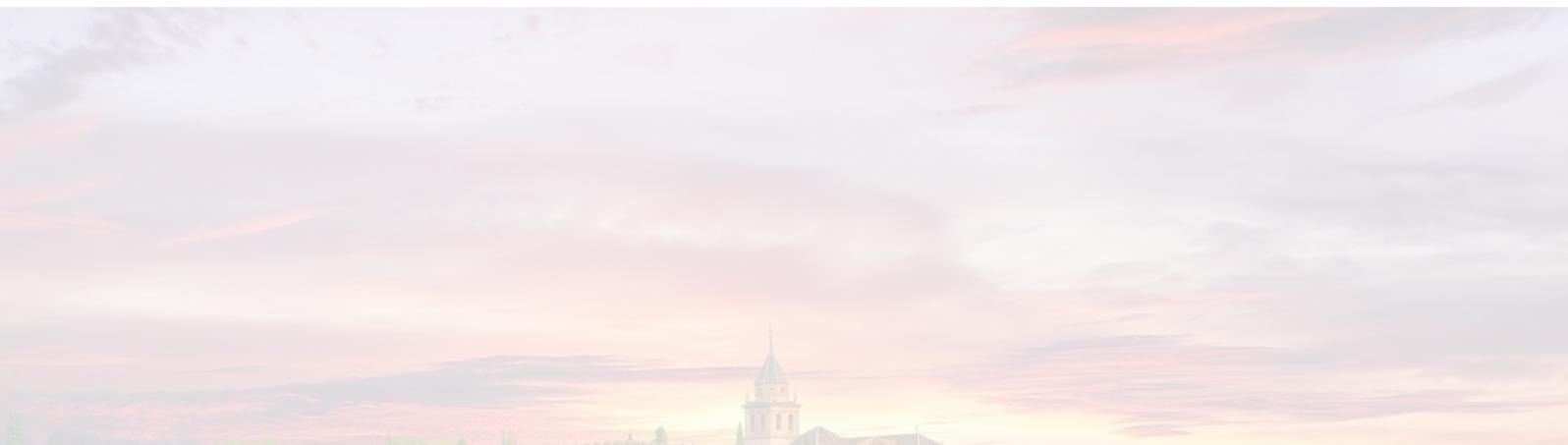
*Por ejemplo, un modelo económico que explique el consumo de energía eléctrica en España (CEE) en función del PIB vendría dada por la expresión*

$$CEE = f(PIB)$$

*Sin embargo, el modelo econométrico quedaría determinado por la expresión lineal*

$$CEE = \beta_0 + \beta_1 \cdot PIB + u_t$$

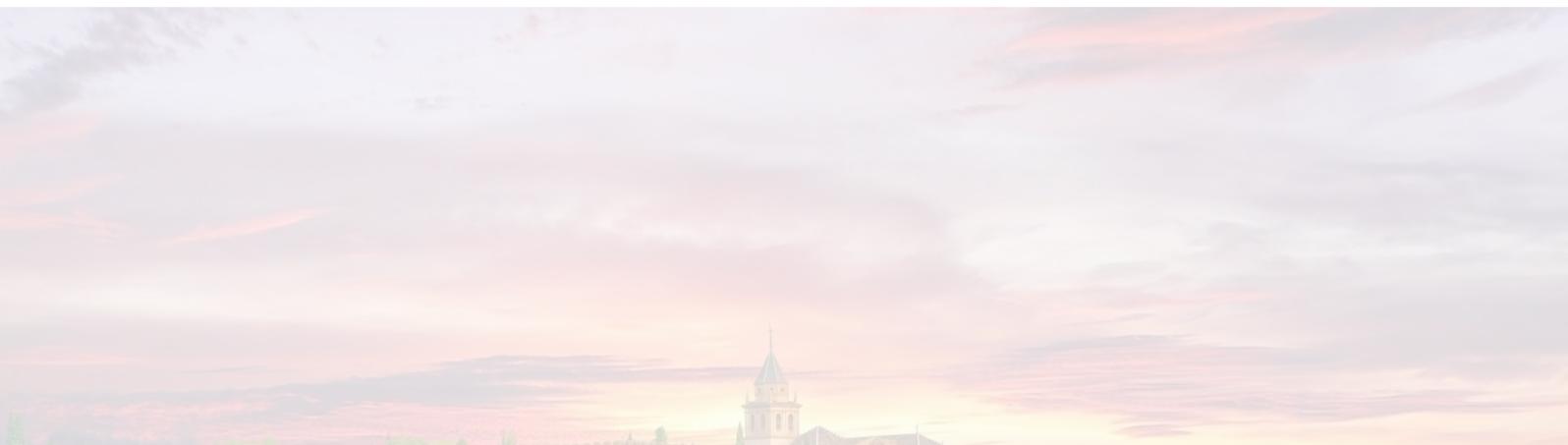
*donde  $\beta_0$ ,  $\beta_1$  son los parámetros a estimar y  $u_t$  es el término aleatorio.*



## FASES DE MÉTODO ECONOMETRICO

A la hora de elaborar un modelo econométrico deben ser considerados los siguientes pasos:

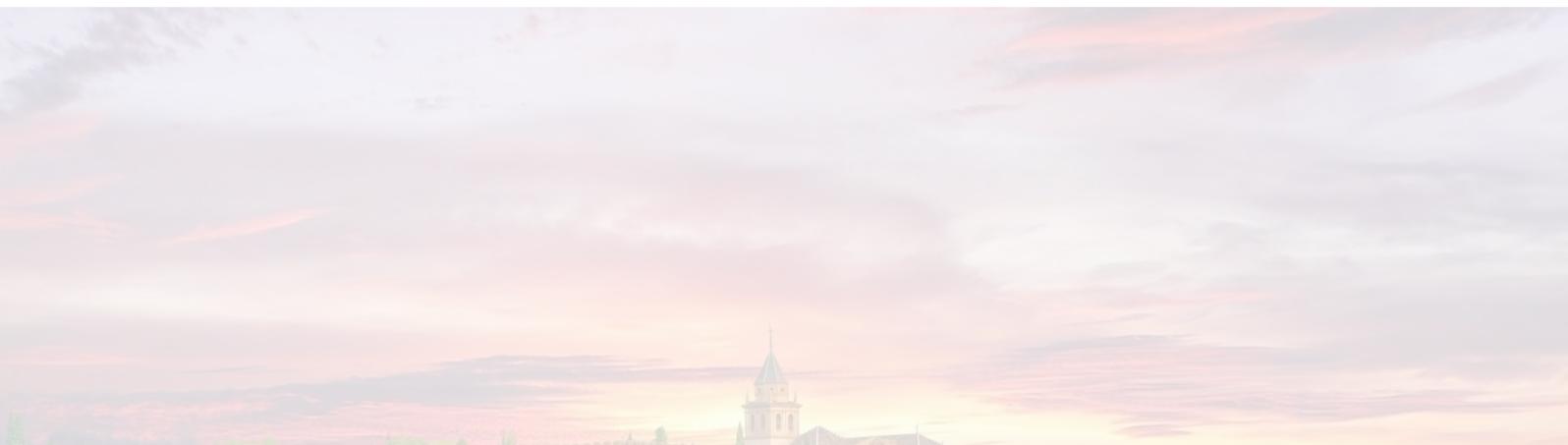
1. **ESPECIFICACIÓN.** Se establece en forma matemática, mediante una o varias ecuaciones, la relación entre las variables en estudio.



## FASES DE MÉTODO ECONOMETRICO

A la hora de elaborar un modelo econométrico deben ser considerados los siguientes pasos:

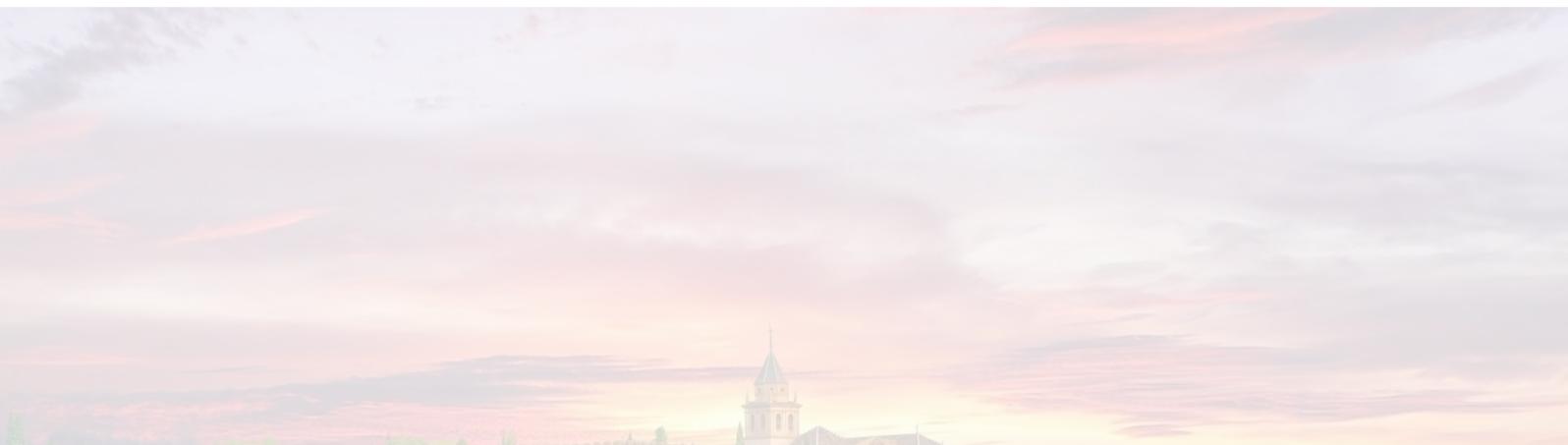
1. **ESPECIFICACIÓN.** Se establece en forma matemática, mediante una o varias ecuaciones, la relación entre las variables en estudio.
2. **ESTIMACIÓN.** Una vez especificado matemáticamente el modelo, se procede a la estimación de cada uno de los parámetros que figuran en él.



## FASES DE MÉTODO ECONOMETRICO

A la hora de elaborar un modelo econométrico deben ser considerados los siguientes pasos:

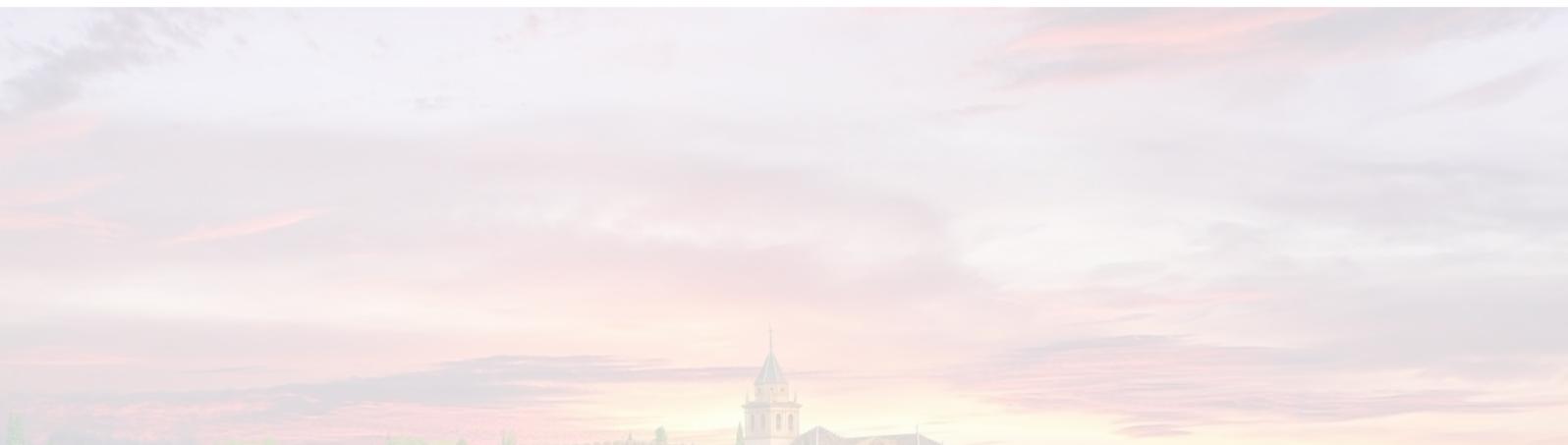
1. **ESPECIFICACIÓN.** Se establece en forma matemática, mediante una o varias ecuaciones, la relación entre las variables en estudio.
2. **ESTIMACIÓN.** Una vez especificado matemáticamente el modelo, se procede a la estimación de cada uno de los parámetros que figuran en él.
3. **VERIFICACIÓN.** Es la diagnosis del modelo con el fin de detectar errores de especificación del modelo. En el caso de que el modelo no supere satisfactoriamente este paso, será necesario reformular el modelo.



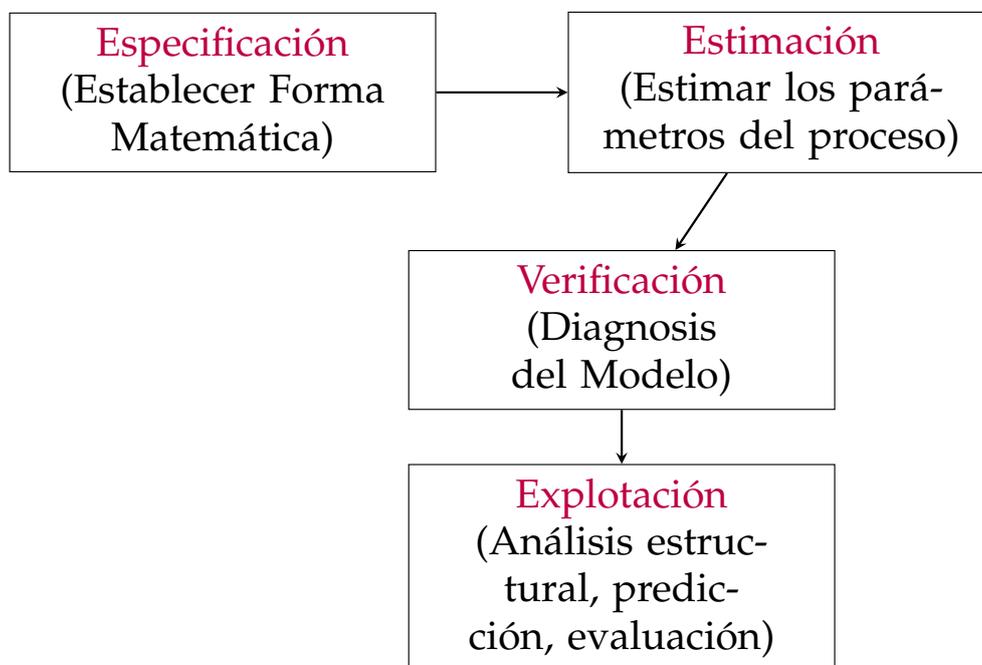
## FASES DE MÉTODO ECONOMETRICO

A la hora de elaborar un modelo econométrico deben ser considerados los siguientes pasos:

1. **ESPECIFICACIÓN.** Se establece en forma matemática, mediante una o varias ecuaciones, la relación entre las variables en estudio.
2. **ESTIMACIÓN.** Una vez especificado matemáticamente el modelo, se procede a la estimación de cada uno de los parámetros que figuran en él.
3. **VERIFICACIÓN.** Es la diagnosis del modelo con el fin de detectar errores de especificación del modelo. En el caso de que el modelo no supere satisfactoriamente este paso, será necesario reformular el modelo.
4. **EXPLOTACIÓN.** Una vez que el modelo ha superado las pruebas de diagnosis, se convierte en un modelo útil para extraer la riqueza que contiene: análisis estructural, predicción condicionada y evaluación de medidas de política económica.



## FASES DE MÉTODO ECONOMETRICO

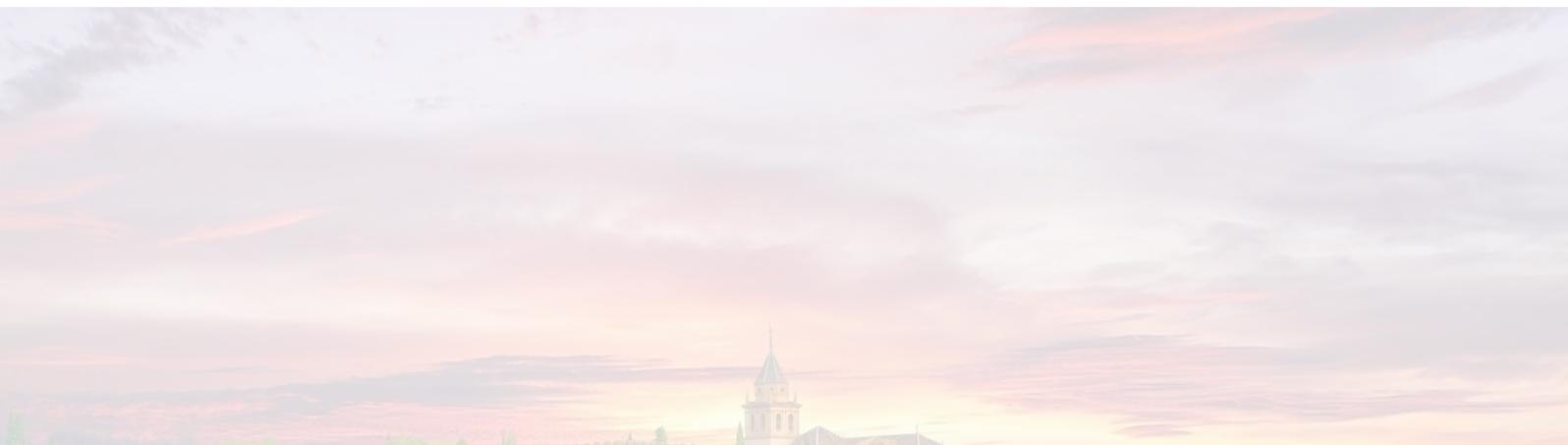


# VARIABLES

Las **VARIABLES** que intervienen en un modelo econométrico las podemos clasificar en dos grandes grupos:

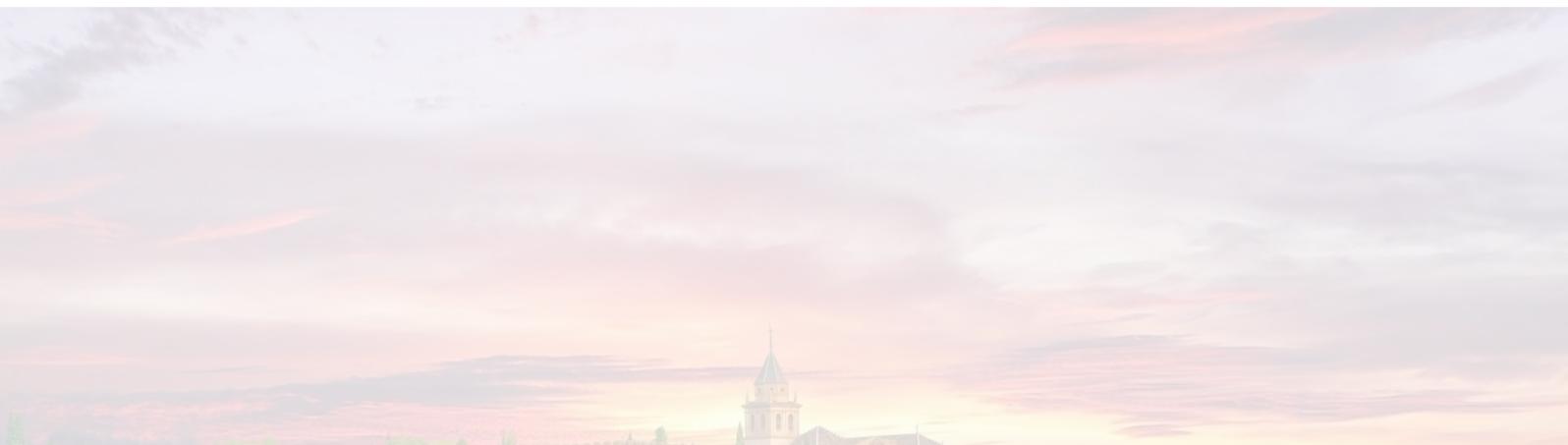
**Observables** Son aquellas variables de las que se disponen de datos muestrales.

**No Observables** Caracterizadas porque se carece de datos → **perturbación**.



## TIPOS DE VARIABLES OBSERVABLES

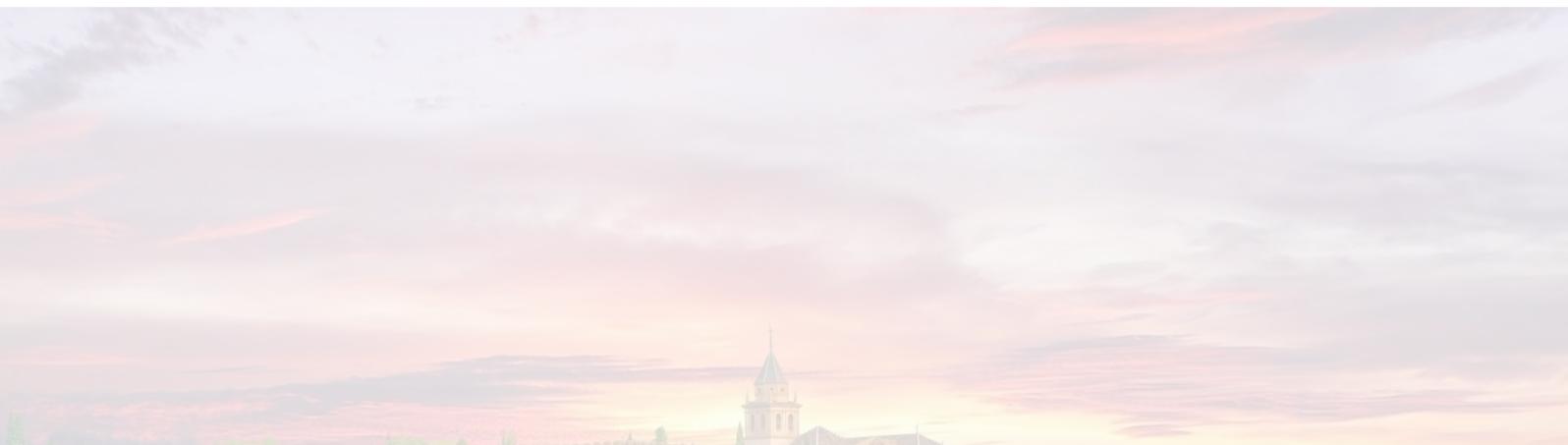
**Endógenas corrientes** También conocidas como variables explicadas o regresandos. Son aquellas que vienen explicadas por el modelo econométrico.



## TIPOS DE VARIABLES OBSERVABLES

**Endógenas corrientes** También conocidas como variables explicadas o regresandos. Son aquellas que vienen explicadas por el modelo econométrico.

**Predeterminadas** El papel fundamental de este tipo de variables es describir el comportamiento de la variable explicada. Por ello, además, reciben el nombre de variables explicativas y regresores.

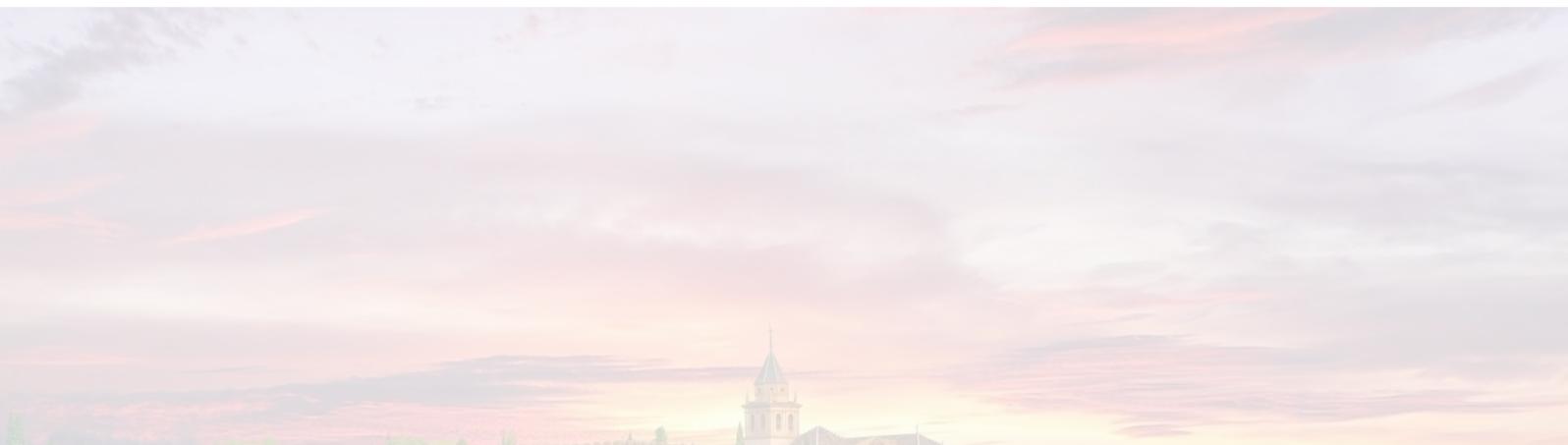


## TIPOS DE VARIABLES OBSERVABLES

**Endógenas corrientes** También conocidas como variables explicadas o regresandos. Son aquellas que vienen explicadas por el modelo econométrico.

**Predeterminadas** El papel fundamental de este tipo de variables es describir el comportamiento de la variable explicada. Por ello, además, reciben el nombre de variables explicativas y regresores. Dependiendo de la naturaleza de las variables predeterminadas se puede realizar una subdivisión:

- ▶ Endógenas retardadas.
- ▶ Exógenas corrientes.
- ▶ Exógenas retardadas.

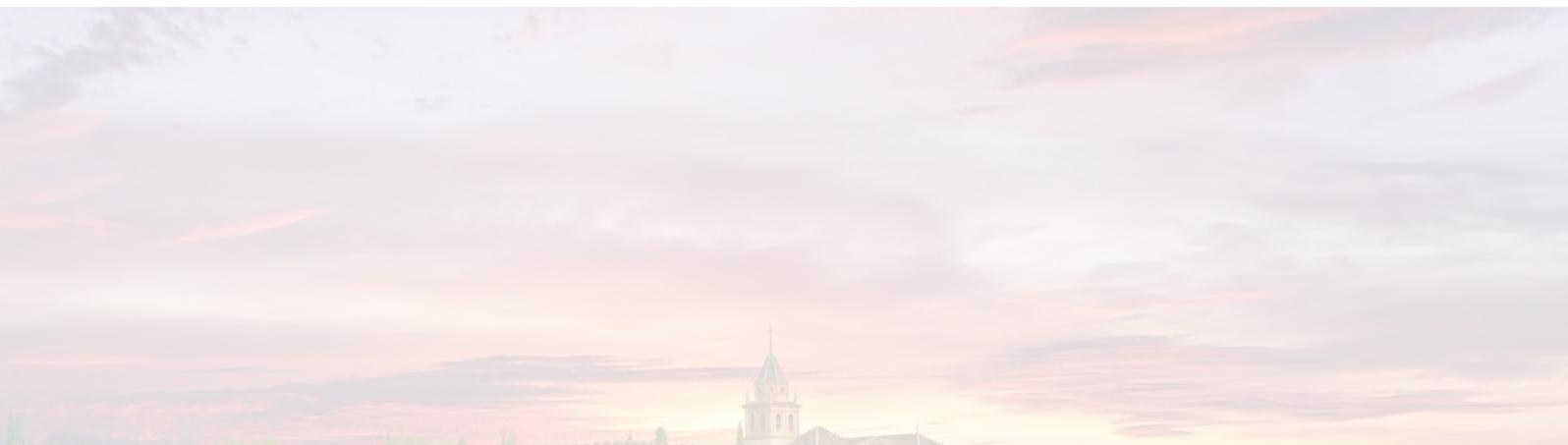


# PARÁMETROS Y RELACIONES

## Ejemplo 2

$$\begin{aligned}c_t &= \alpha_0 + \alpha_1(1 - \tau)y_t + \alpha_2r_t + u_t, \\i_t &= \beta_0 + \beta_1(y_{t-1} - y_{t-2}) + \beta_2r_{t-1} + v_t, \\y_t &\equiv c_t + i_t + g_t,\end{aligned}$$

*donde  $y$  es el producto interior bruto,  $r$  el tipo de interés,  $c$  son los gastos de consumo,  $i$  los gastos de inversión,  $g$  son los gastos de gobierno y  $\tau$  el tipo impositivo.*



# PARÁMETROS Y RELACIONES

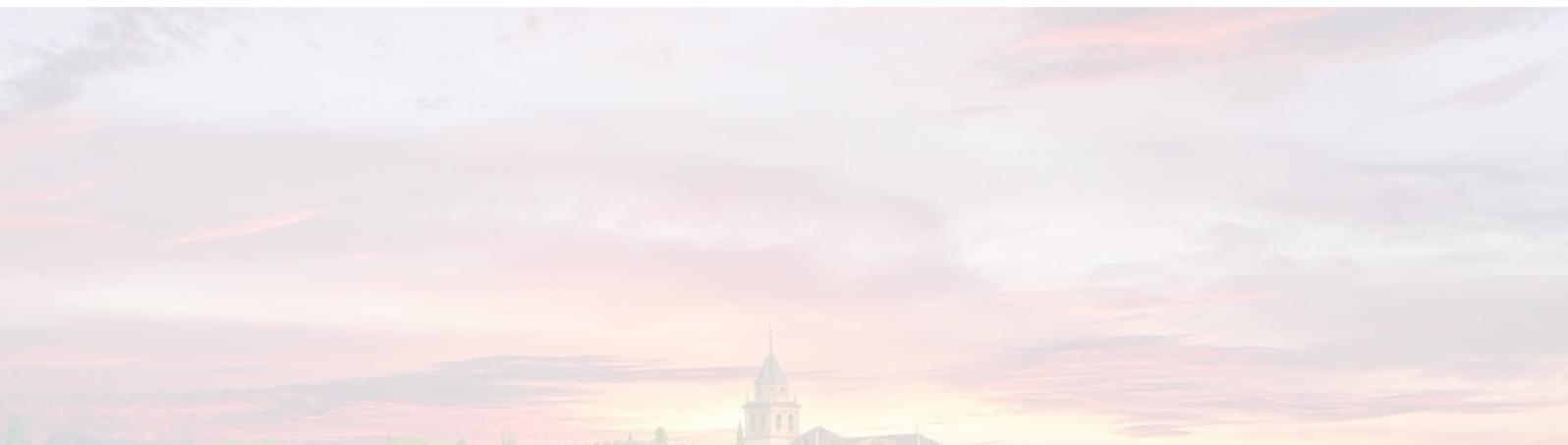
## Ejemplo 2

$$\begin{aligned}c_t &= \alpha_0 + \alpha_1(1 - \tau)y_t + \alpha_2r_t + u_t, \\i_t &= \beta_0 + \beta_1(y_{t-1} - y_{t-2}) + \beta_2r_{t-1} + v_t, \\y_t &\equiv c_t + i_t + g_t,\end{aligned}$$

donde  $y$  es el producto interior bruto,  $r$  el tipo de interés,  $c$  son los gastos de consumo,  $i$  los gastos de inversión,  $g$  son los gastos de gobierno y  $\tau$  el tipo impositivo.

¿quiénes son  $\alpha_0$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\beta_0$ ,  $\beta_1$  y  $\beta_2$ ?. Los **PARÁMETROS**: fijos y desconocidos.

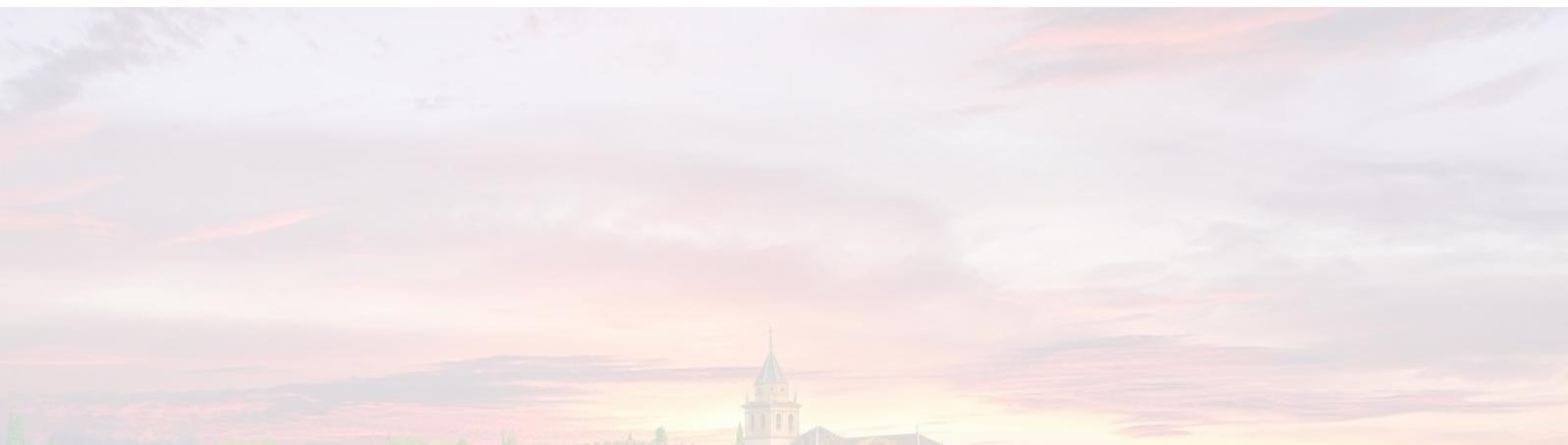
Las **RELACIONES** que existen entre los distintos tipos de variables vienen especificadas mediante un sistema de ecuaciones.



## FORMAS DE EXPRESAR UN MODELO ECONOMÉTRICO

Clasificación de los modelos econométricos según el **NÚMERO DE ECUA-  
CIONES:**

**Uniecuacionales** En este caso el sistema queda descrito por una ecuación.

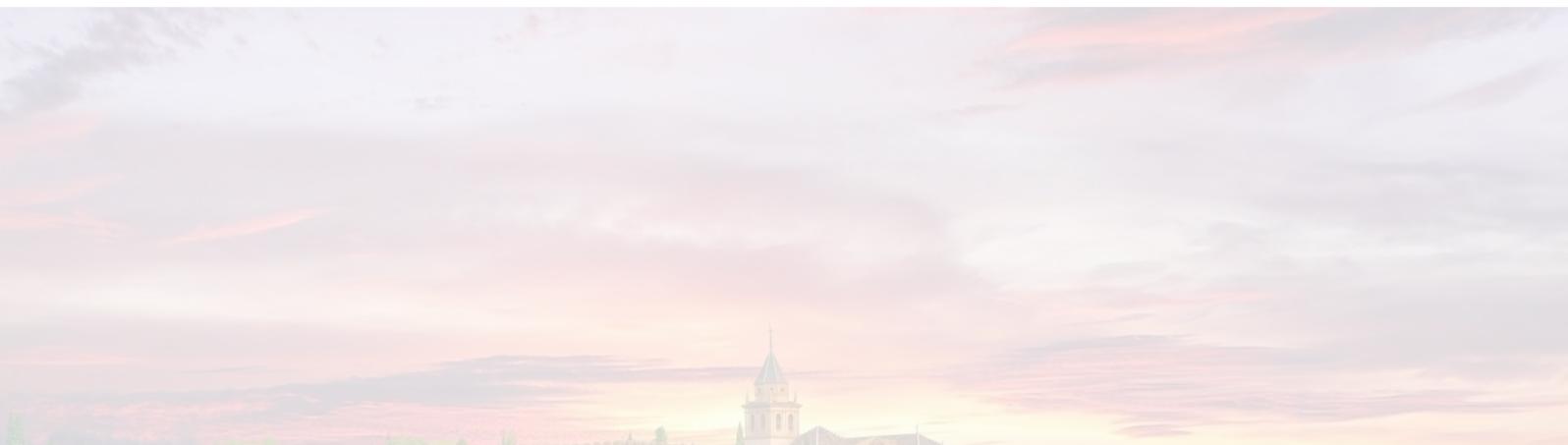


## FORMAS DE EXPRESAR UN MODELO ECONOMÉTRICO

Clasificación de los modelos econométricos según el **NÚMERO DE ECUACIONES**:

**Uniecuacionales** En este caso el sistema queda descrito por una ecuación. Dependiendo del número de variables se puede dividir en:

1. Uniecuacional simple  $Y_t = \beta_0 + \beta_1 X_t + u_t$ .
2. Uniecuacional múltiple  
 $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + u_t$ .



## FORMAS DE EXPRESAR UN MODELO ECONOMÉTRICO

Clasificación de los modelos econométricos según el **NÚMERO DE ECUACIONES**:

**Uniecuacionales** En este caso el sistema queda descrito por una ecuación. Dependiendo del número de variables se puede dividir en:

1. Uniecuacional simple  $Y_t = \beta_0 + \beta_1 X_t + u_t$ .
2. Uniecuacional múltiple  
 $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + u_t$ .

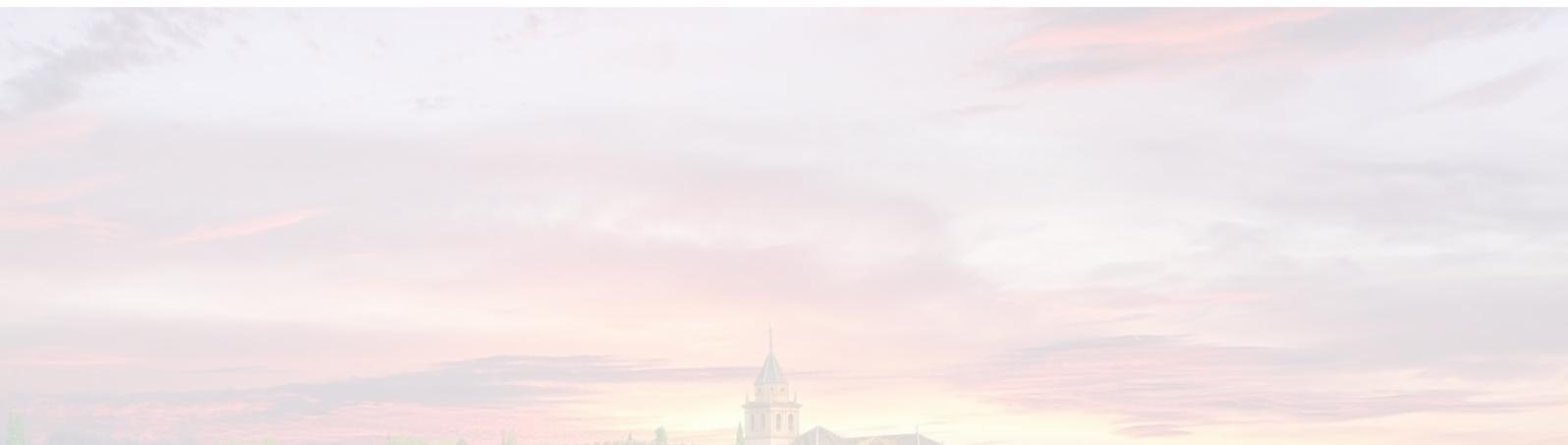
**Multiecuacionales** El modelo está formado por un sistema de ecuaciones donde puede existir relación o no entre las variables.

$$Y_{1t} = \gamma_{12} Y_{2t} + \beta_{11} X_{1t} + u_{1t}$$

$$Y_{2t} = \gamma_{21} Y_{1t} + \beta_{21} X_{1t} + \beta_{22} X_{2t} + u_{2t}$$

$$Y_{1t} = \beta_0 + \beta_1 X_{1t} + u_t$$

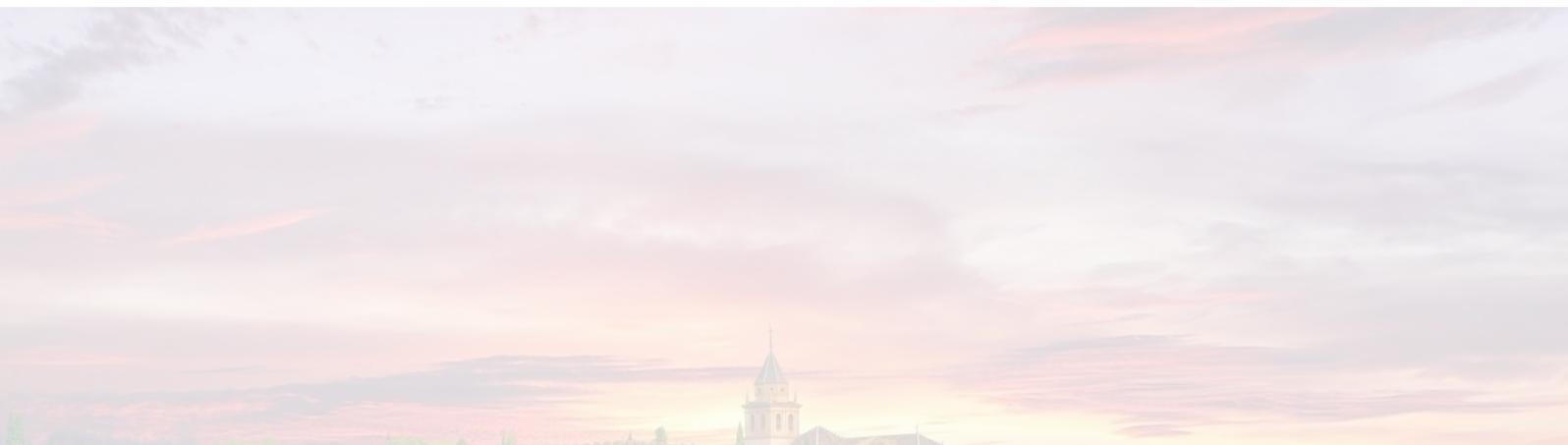
$$Y_{2t} = \alpha_0 + \alpha_1 X_{2t} + v_t$$



# FORMAS DE EXPRESAR UN MODELO ECONOMÉTRICO

Según la **FORMA FUNCIONAL**, dependiendo de la expresión del modelo encontramos:

**Modelos lineales**  $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$ .



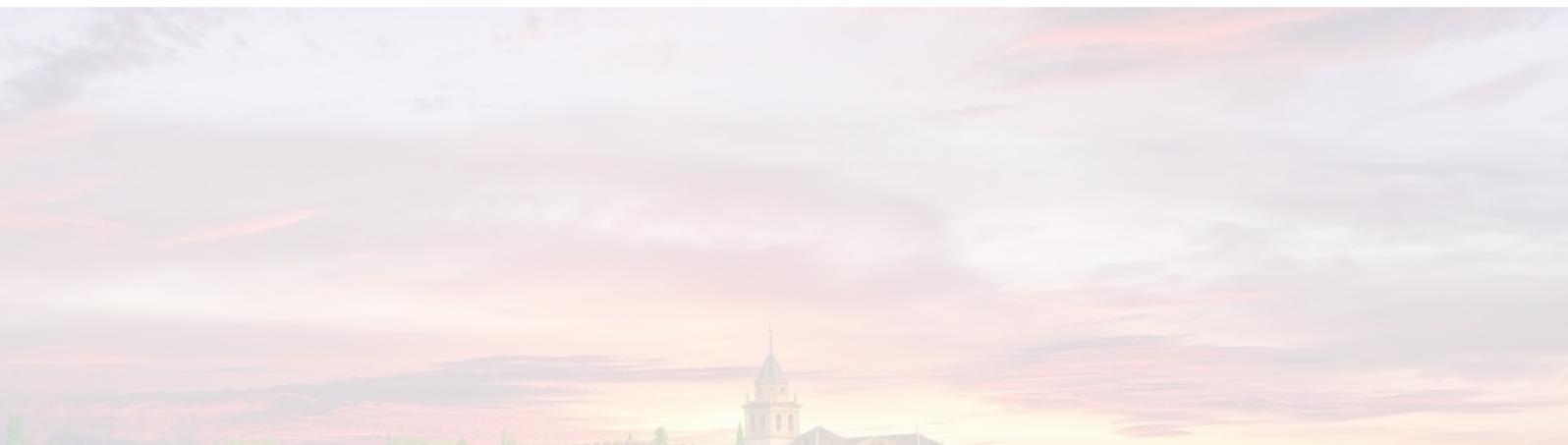
# FORMAS DE EXPRESAR UN MODELO ECONOMÉTRICO

Según la **FORMA FUNCIONAL**, dependiendo de la expresión del modelo encontramos:

**Modelos lineales**  $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$ .

**Modelos no lineales intrínsecamente linealizables**

$$Y_t = \beta_0 + \beta_1 e^{X_{1t}} + \beta_2 X_{2t} X_{3t} + u_t.$$



## FORMAS DE EXPRESAR UN MODELO ECONOMÉTRICO

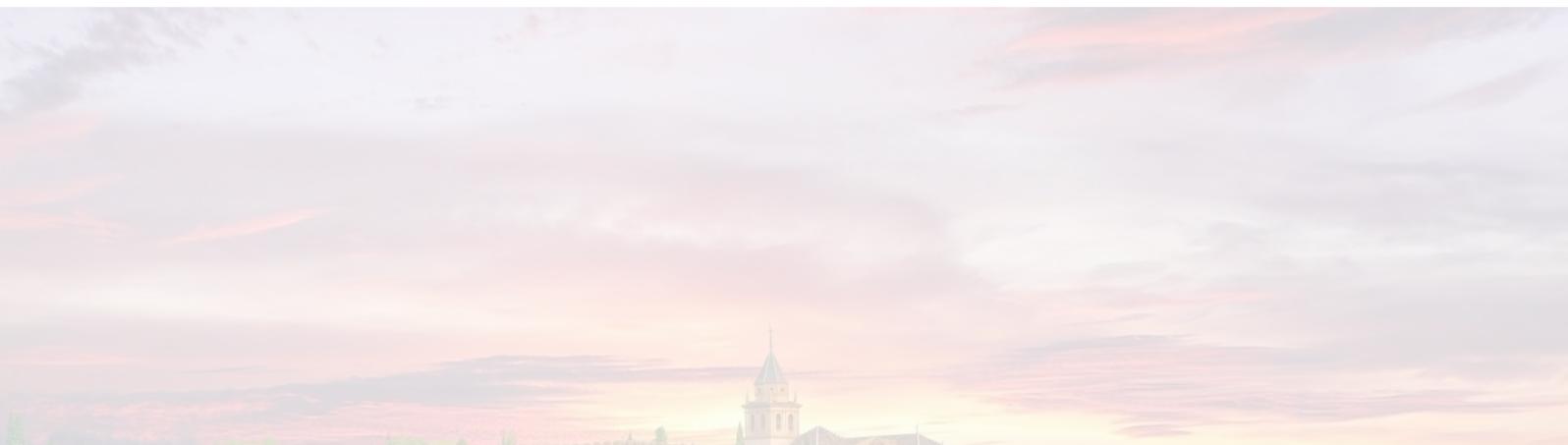
Según la **FORMA FUNCIONAL**, dependiendo de la expresión del modelo encontramos:

**Modelos lineales**  $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t$ .

**Modelos no lineales intrínsecamente linealizables**

$$Y_t = \beta_0 + \beta_1 e^{X_{1t}} + \beta_2 X_{2t} X_{3t} + u_t.$$

**Modelos no lineales intrínsecamente no linealizables**  $Y_t = \beta_0 + X_t^{\beta_1} + u_t$ .

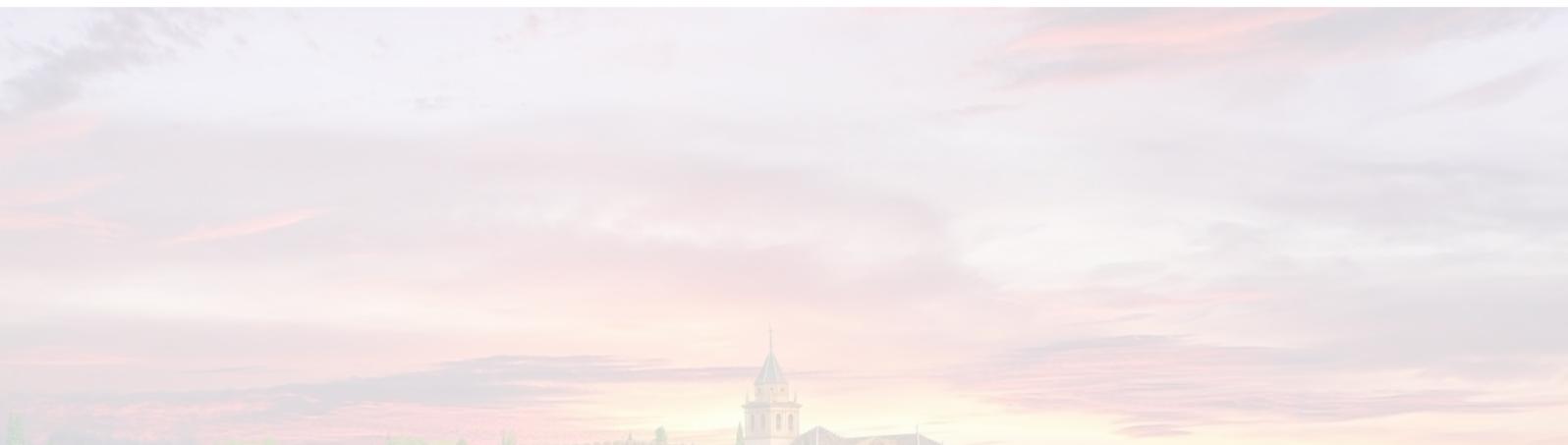


# FORMAS DE EXPRESAR UN MODELO ECONOMÉTRICO

Según las **CARACTERÍSTICAS DE LA PERTURBACIÓN**:

**Modelos clásicos** Se considera que la perturbación es ruido blanco, es decir,

1.  $E[u_t] = 0 \forall t$
2.  $\text{var}[u_t] = \sigma^2 \forall t$
3.  $\text{cov}[u_t u_s] = 0 \forall t \neq s$ .



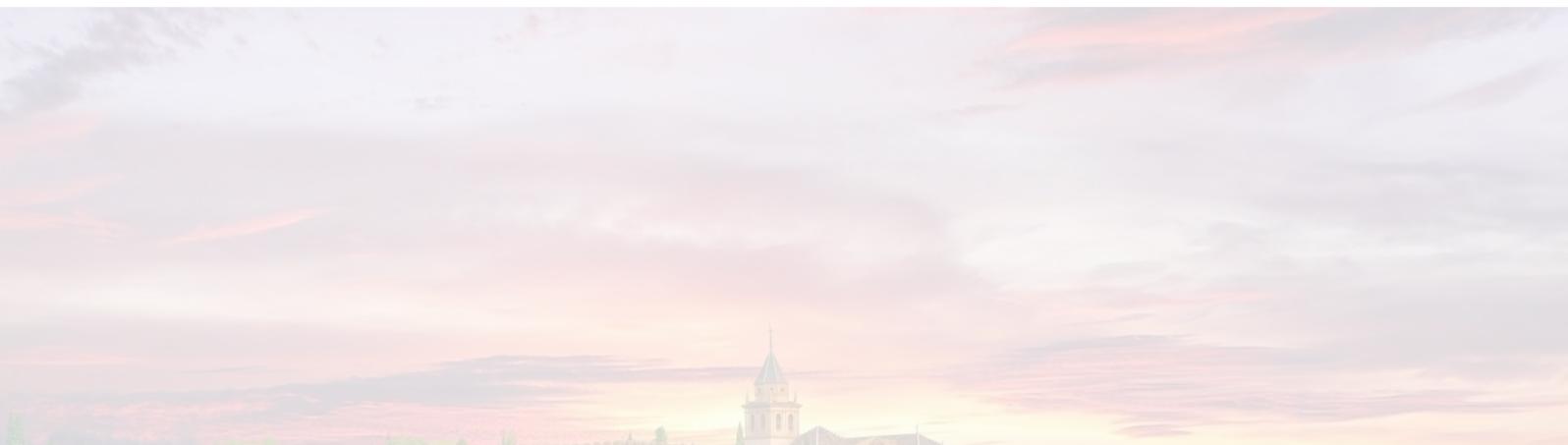
## FORMAS DE EXPRESAR UN MODELO ECONOMÉTRICO

Según las **CARACTERÍSTICAS DE LA PERTURBACIÓN**:

**Modelos clásicos** Se considera que la perturbación es ruido blanco, es decir,

1.  $E[u_t] = 0 \forall t$
2.  $\text{var}[u_t] = \sigma^2 \forall t$
3.  $\text{cov}[u_t u_s] = 0 \forall t \neq s$ .

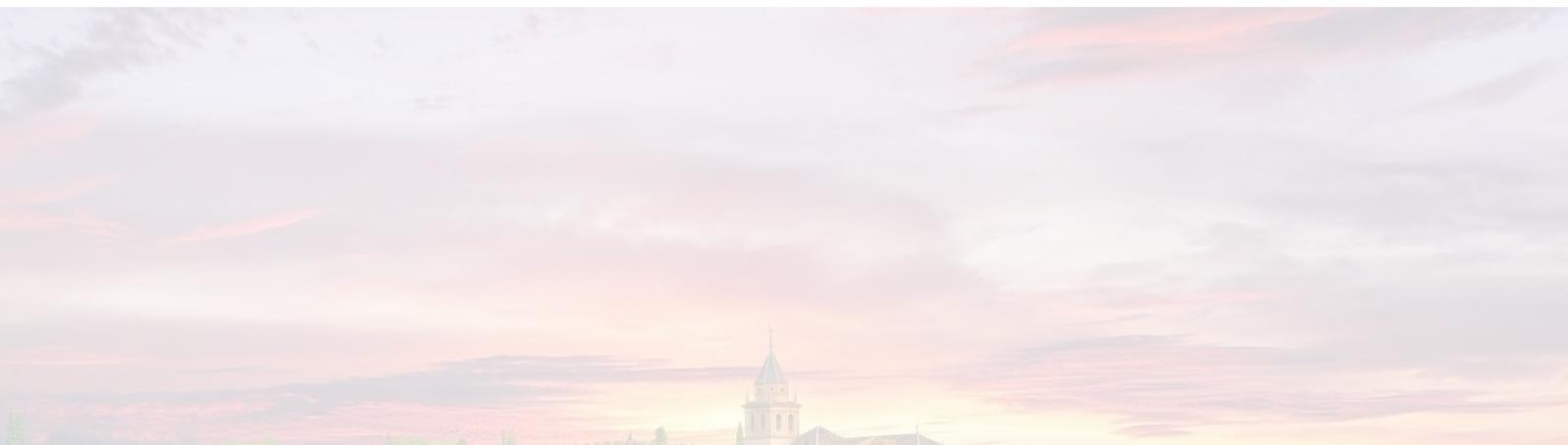
**Modelos generalizados** En estos casos las perturbaciones incumplen alguna de las dos últimas hipótesis descritas en los modelos clásicos, es decir, puede existir un problema de heterocedasticidad (la varianza no es constante) o autocorrelación (la covarianza es distinta de cero para  $t \neq s$ ).



## NATURALEZA DE LA INFORMACIÓN

Las bases de datos sobre las que se basa la econometría pueden ser clasificadas en tres tipos:

**Sección Cruzada** Los datos de sección cruzada o de corte transversal consiste en datos de múltiples individuos observados en un mismo instante de tiempo. Ej: Calificaciones de los alumnos de Econometría en la convocatoria de Febrero durante el curso 2023/2024.

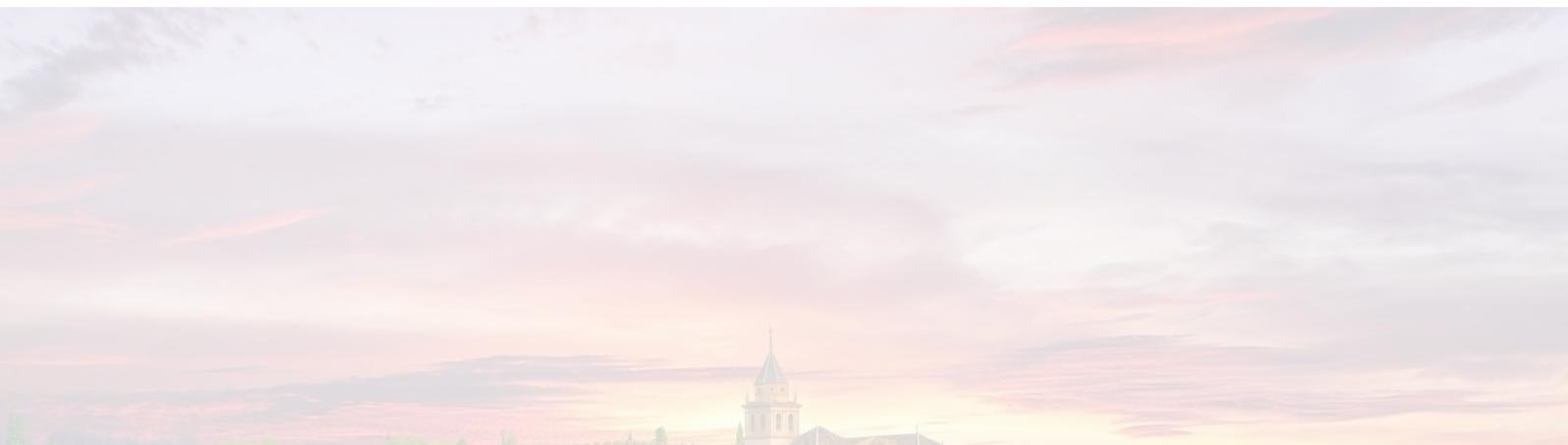


## NATURALEZA DE LA INFORMACIÓN

Las bases de datos sobre las que se basa la econometría pueden ser clasificadas en tres tipos:

**Sección Cruzada** Los datos de sección cruzada o de corte transversal consiste en datos de múltiples individuos observados en un mismo instante de tiempo. Ej: Calificaciones de los alumnos de Econometría en la convocatoria de Febrero durante el curso 2023/2024.

**Series Temporales** Los datos de series temporales es un conjunto de datos de un único individuo observado en diferentes instantes de tiempo. Ej: Cotizaciones diarias de las acciones del Banco Santander.



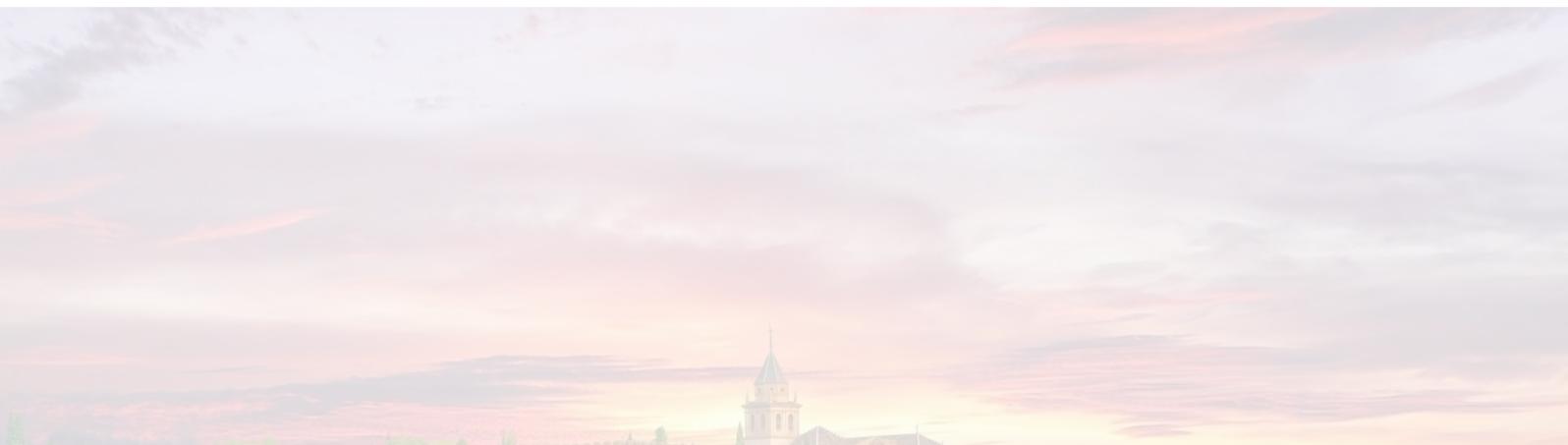
## NATURALEZA DE LA INFORMACIÓN

Las bases de datos sobre las que se basa la econometría pueden ser clasificadas en tres tipos:

**Sección Cruzada** Los datos de sección cruzada o de corte transversal consiste en datos de múltiples individuos observados en un mismo instante de tiempo. Ej: Calificaciones de los alumnos de Econometría en la convocatoria de Febrero durante el curso 2023/2024.

**Series Temporales** Los datos de series temporales es un conjunto de datos de un único individuo observado en diferentes instantes de tiempo. Ej: Cotizaciones diarias de las acciones del Banco Santander.

**Datos de Panel** Los datos de panel, o también conocidos como datos longitudinales, consiste en múltiples individuos observados cada uno de ellos en distintos instantes de tiempo. Ej: Cotizaciones diarias de todas las empresas que componen el Ibex 35.



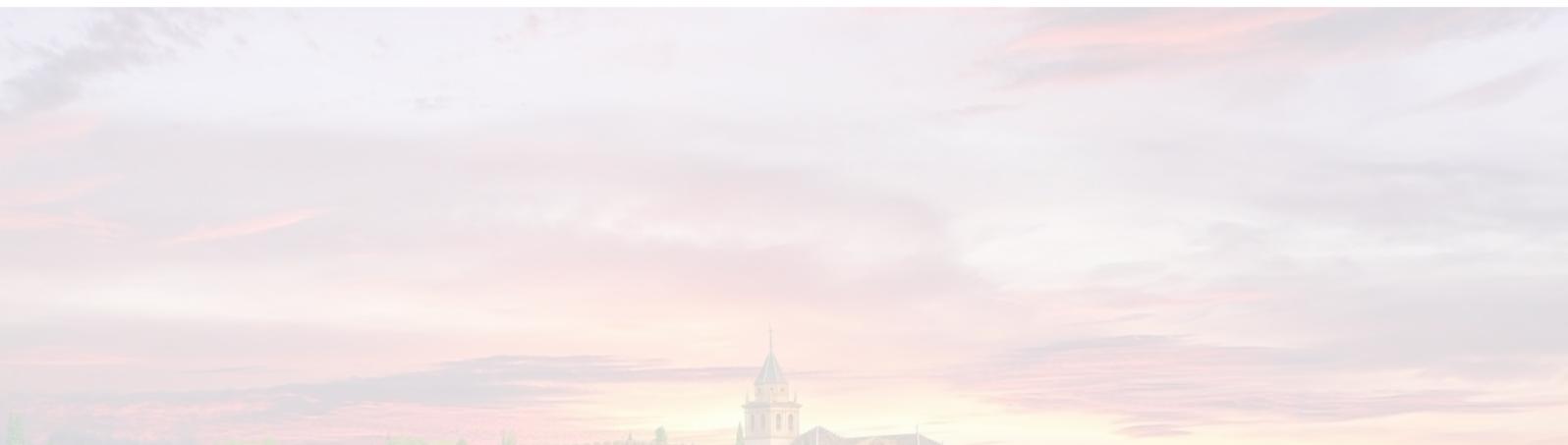
# EJEMPLOS CORTE TRANSVERSAL Y TEMPORAL

Ejemplo de corte transversal:

N. de obs.	Calif. media/dist(Y)	Ratio estud-maest(X <sub>1</sub> )	G/estud.(X <sub>2</sub> )	% estud. ingls(X <sub>3</sub> )
1	690.8	17.89	6.385	0.8
2	661.2	21.52	5.099	4.6
3	643.6	18.70	5.502	30.8
...	...	...	...	...
419	672.2	20.20	4.776	3.0
420	765.8	19.04	5.993	5.8

Ejemplo de series temporales:

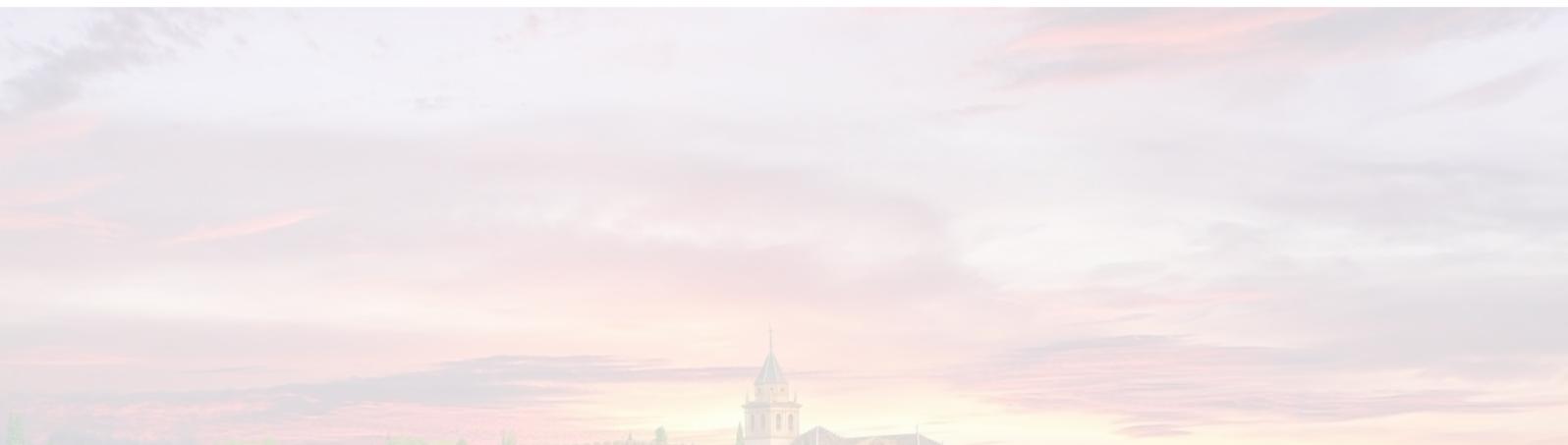
Ao	Consumo (Y)	Renta (X)
1982	3081.5	4620.3
1983	3240.6	4803.7
1984	3407.6	5140.1
1985	3566.5	5323.5
1986	3708.7	5487.7
1987	3822.3	5649.5
1988	3972.7	5865.2
1989	4064.6	6062.0
1990	4132.2	6136.3
1991	4105.8	6079.4
1992	4219.8	6244.4
1993	4343.6	6389.6
1994	4486.0	6610.7
1995	4595.3	6742.1
1996	4714.1	6928.4



## EJEMPLO DATOS DE PANEL

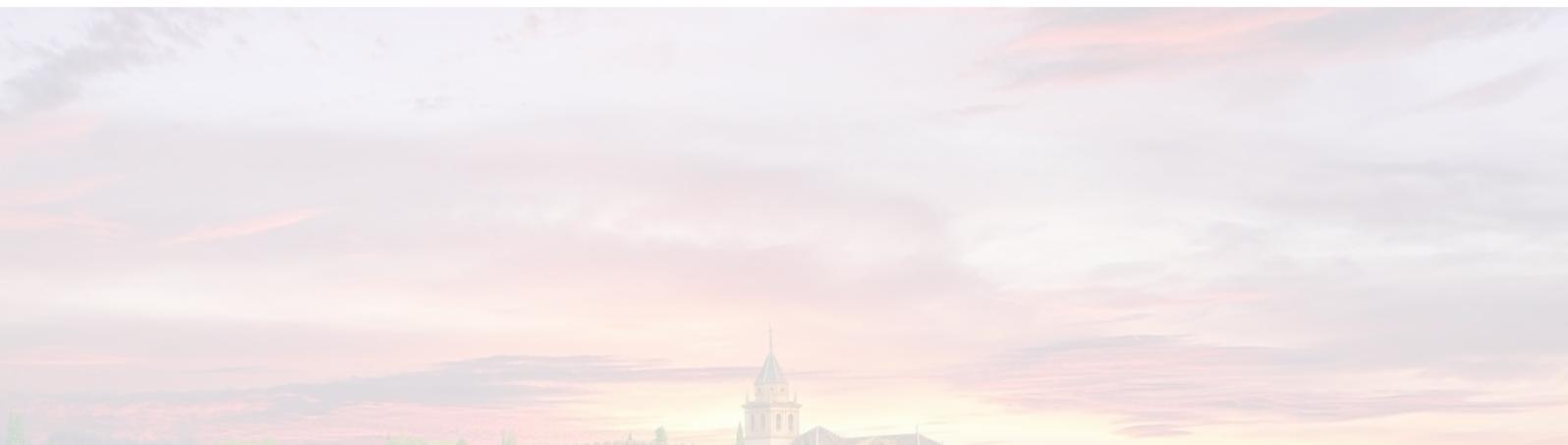
Ejemplo de datos de panel:

N. de obs.	Estado	Ao	Vta de cig (Y)	P. medio/paquete( $X_1$ )	Imptos tot.( $X_2$ )
1	Alabama	1985	116.5	1.022	0.333
2	Arkansas	1985	128.5	1.015	0.370
...	...	...	...	...	...
48	Wyoming	1985	129.4	0.935	0.240
49	Alabama	1986	117.2	1.080	0.334
...	...	...	...	...	...
96	Wyoming	1986	127.8	1.007	0.240
97	Alabama	1987	115.8	1.135	0.335
...	...	...	...	...	...
528	Wyoming	1995	127.8	1.007	0.240



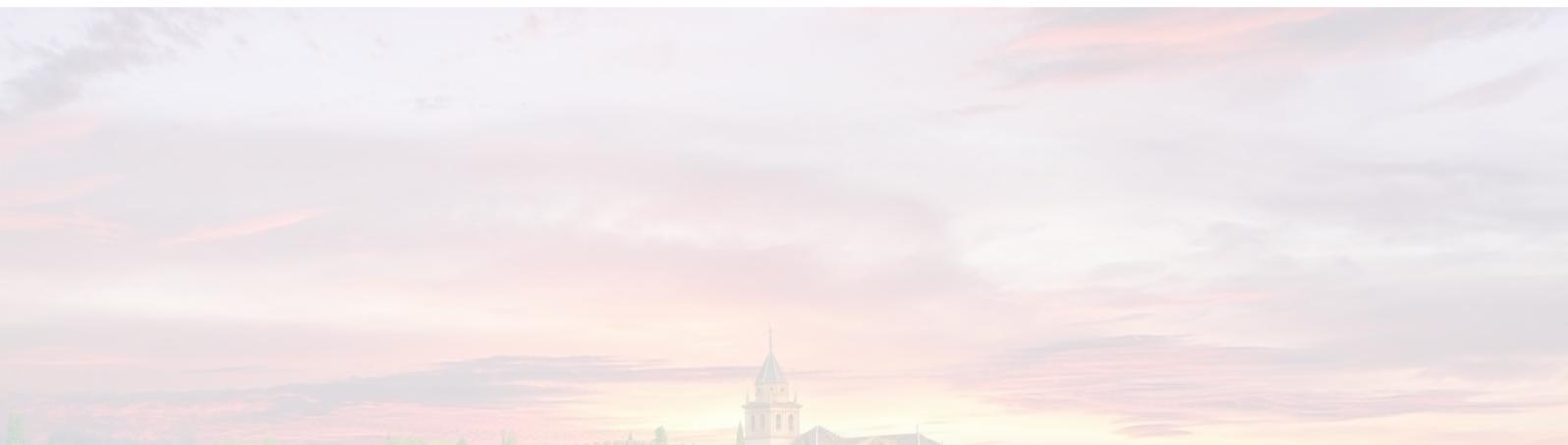
## PRACTICAMOS

Sugiera un modelo para explicar la calificación final que se obtendrá en Econometría. Defina las variables (incluya alguna cualitativa), sus unidades de medida y las dificultades que tendría para disponer de esa información.

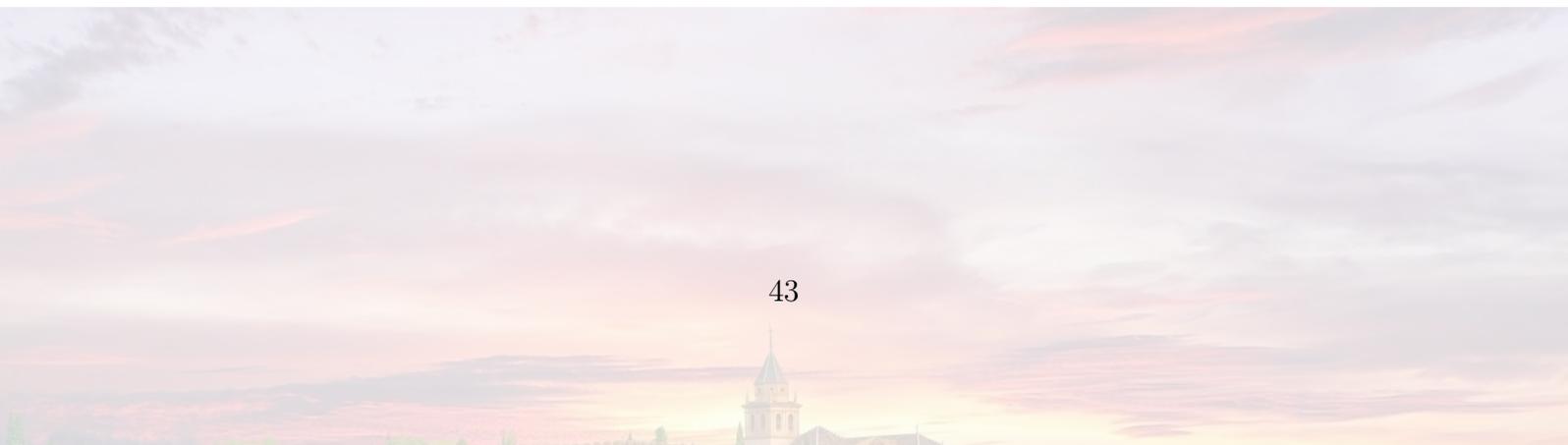


## BIBLIOGRAFÍA

- ▶ García, C.B., Pérez-sánchez, J.M. y Salmerón-Gómez, R. (2017). *Econometría Básica para la economía y la empresa*. Ed. Fleming. Tema 1.
- ▶ Sánchez, C., López, M.d.M y García, T. *Econometría*. Ed. Fleming. Capítulo 1
- ▶ Gujarati, D. (1997). *Econometría*. Ed. McGraw Hill. Capítulo 1.
- ▶ Muñoz, A. y Parra, F. (2007). *Econometría aplicada*. Ed. Académicas S.A. Capítulo 1.
- ▶ Stock, J. y Watson, M. (2012). *Introducción a la Econometría*. Ed. Pearson. Capítulo 1.
- ▶ Uriel, E., Contreras, D., Moltó, M.L. y Peiró, A. (1990). *Econometría. El Modelo Lineal*. Editorial AC. Capítulo 1.
- ▶ Wooldridge, J.M. (2005). *Introducción a la Econometría: Un enfoque moderno*. Thomson. Capítulo 1.
- ▶ Matilla García, m., Pérez Pascual, Pedro y Sanz Carnero, B. (2013) *Econometría y predicción*. Mc Graw Hill.



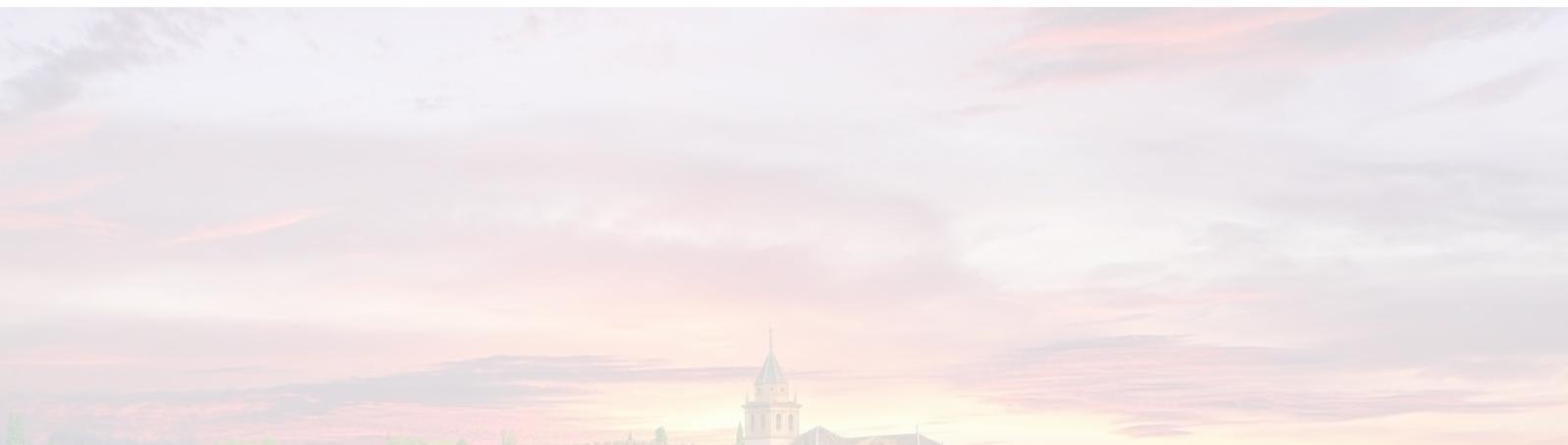
## 2 Tema 2



# ECONOMETRÍA

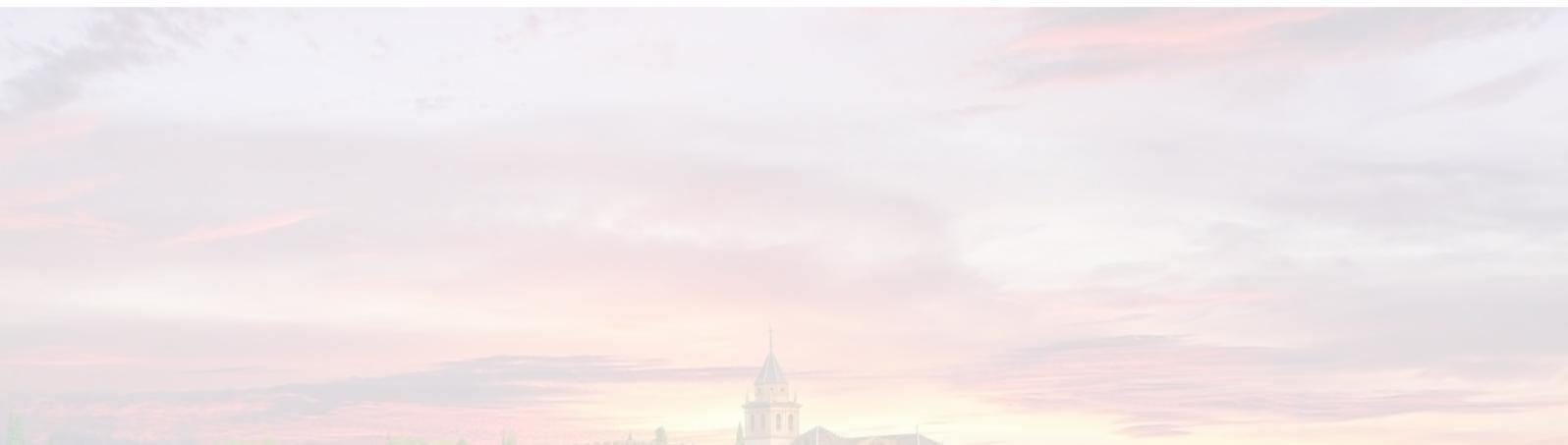
## TEMA 2: EL MODELO LINEAL (I)

2024-2025



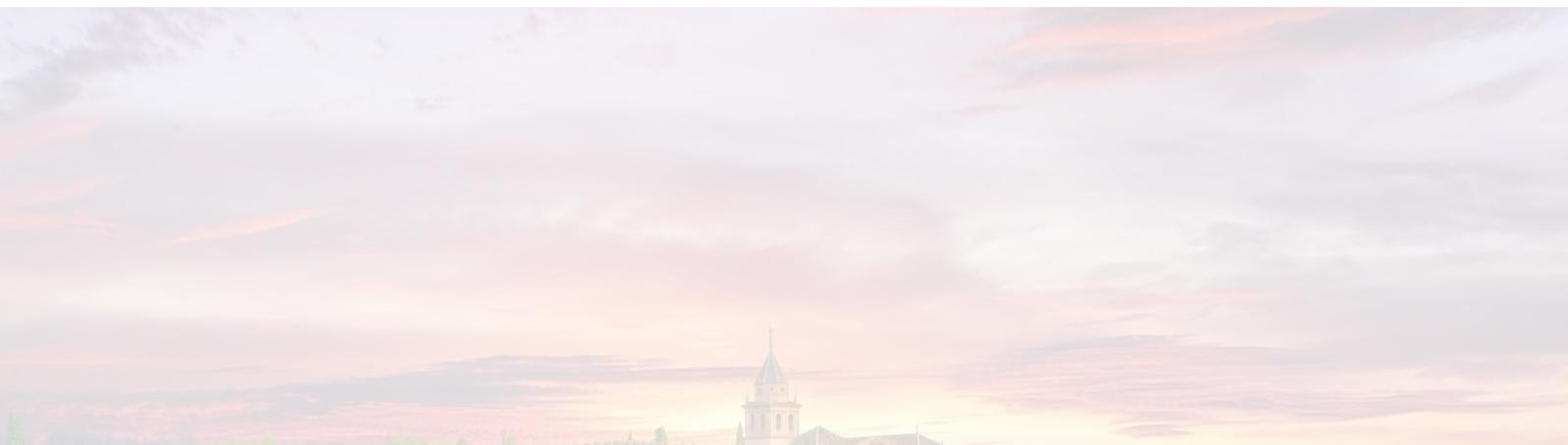
# EL MODELO LINEAL (I)

<b>1</b>	<b>INTRODUCCIÓN</b>	<b>2</b>
<b>2</b>	<b>HIPÓTESIS DEL MODELO</b>	<b>7</b>
<b>3</b>	<b>ESTIMACIÓN: MÍNIMOS CUADRADOS ORDINARIOS</b>	<b>14</b>
3.1	PROPIEDADES ALGEBRAICAS DE LOS MCO	19
3.2	LAS ECUACIONES NORMALES	21
<b>4</b>	<b>BONDAD DEL AJUSTE</b>	<b>23</b>
4.1	BONDAD DEL AJUSTE: $R^2$ Y $\bar{R}^2$	23
4.2	CRITERIOS DE INFORMACIÓN	26
<b>5</b>	<b>BIBLIOGRAFÍA</b>	<b>32</b>



# INTRODUCCIÓN

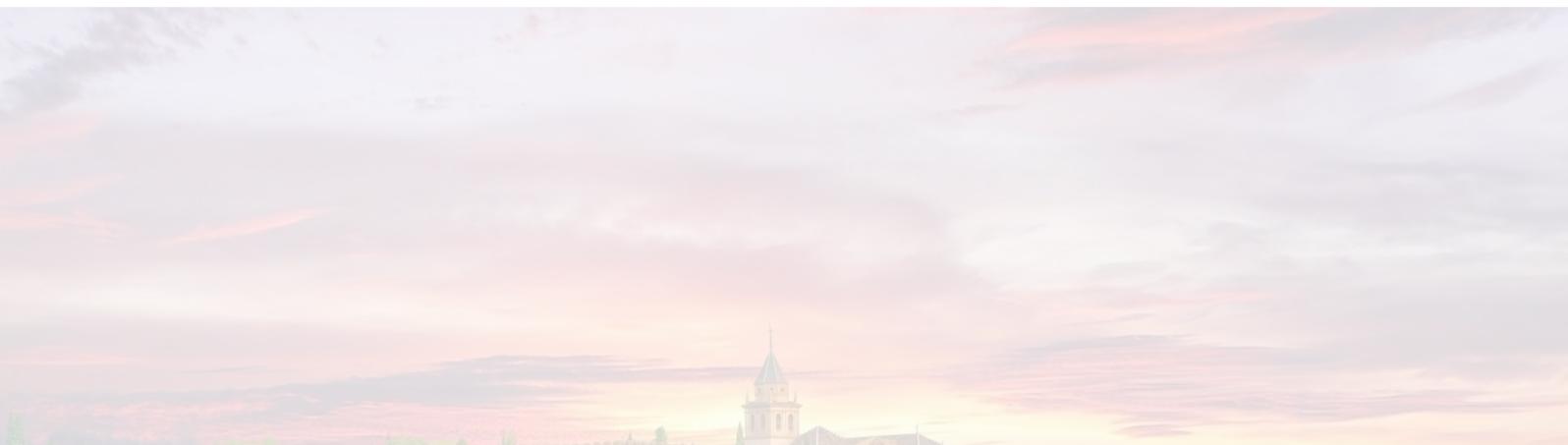
Modelo Lineal Simple:



# INTRODUCCIÓN

Modelo Lineal Simple:

$$y_i \sim a + bx_i$$

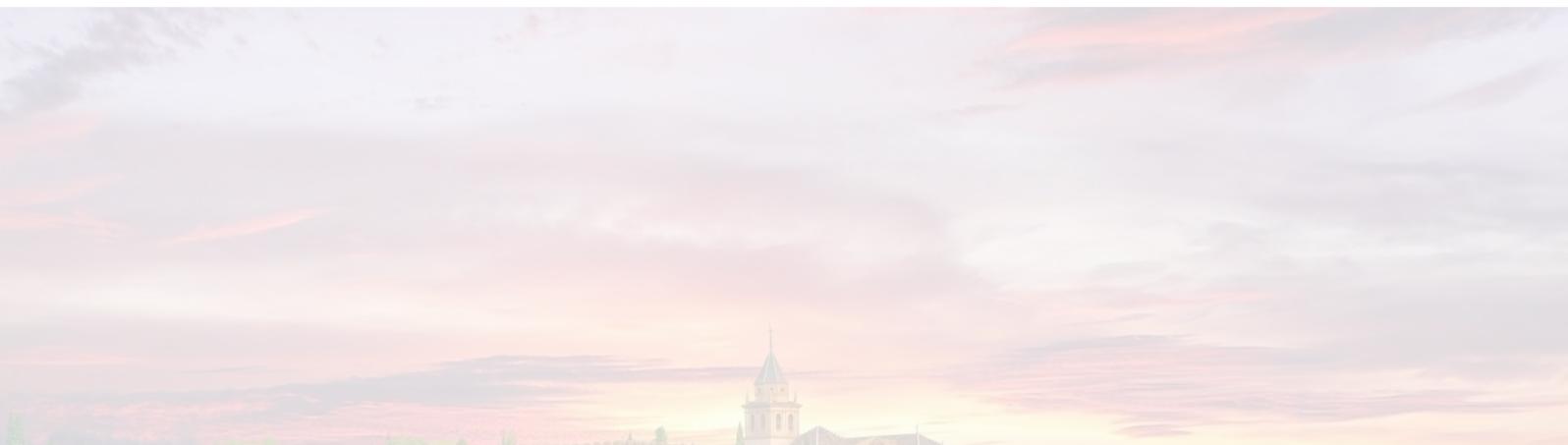


# INTRODUCCIÓN

Modelo Lineal Simple:

$$y_i \sim a + bx_i$$

Modelo Lineal Múltiple:



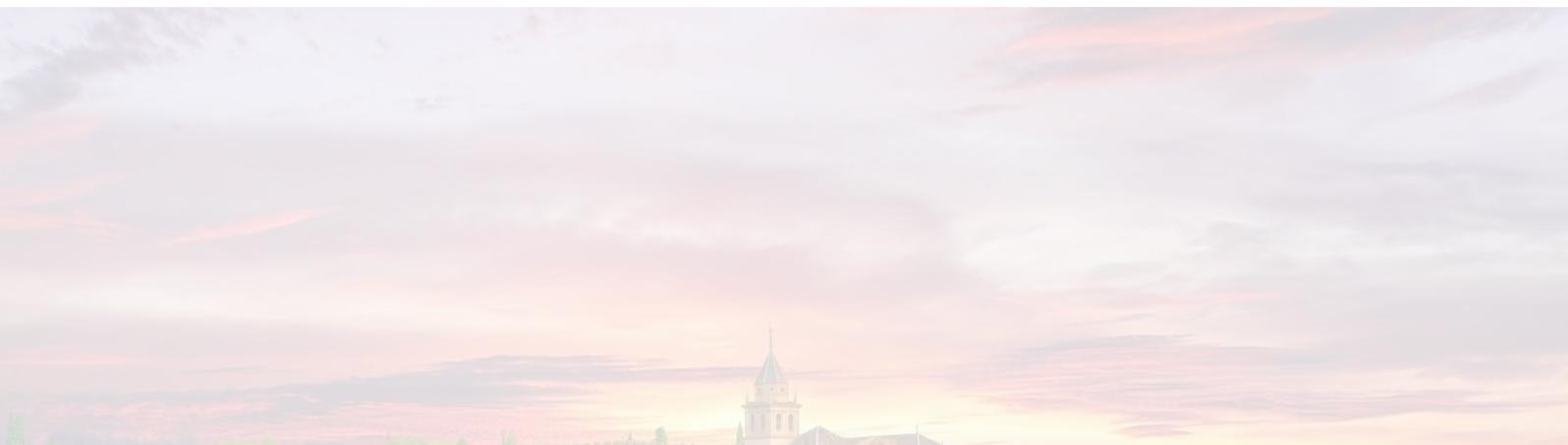
# INTRODUCCIÓN

**Modelo Lineal Simple:**

$$y_i \sim a + bx_i$$

**Modelo Lineal Múltiple:**

Permite estudiar el comportamiento de **una** variable endógena  $Y$ , a partir de  $k$  variables independientes  $(X_2, X_3, \dots, X_k)$



# INTRODUCCIÓN

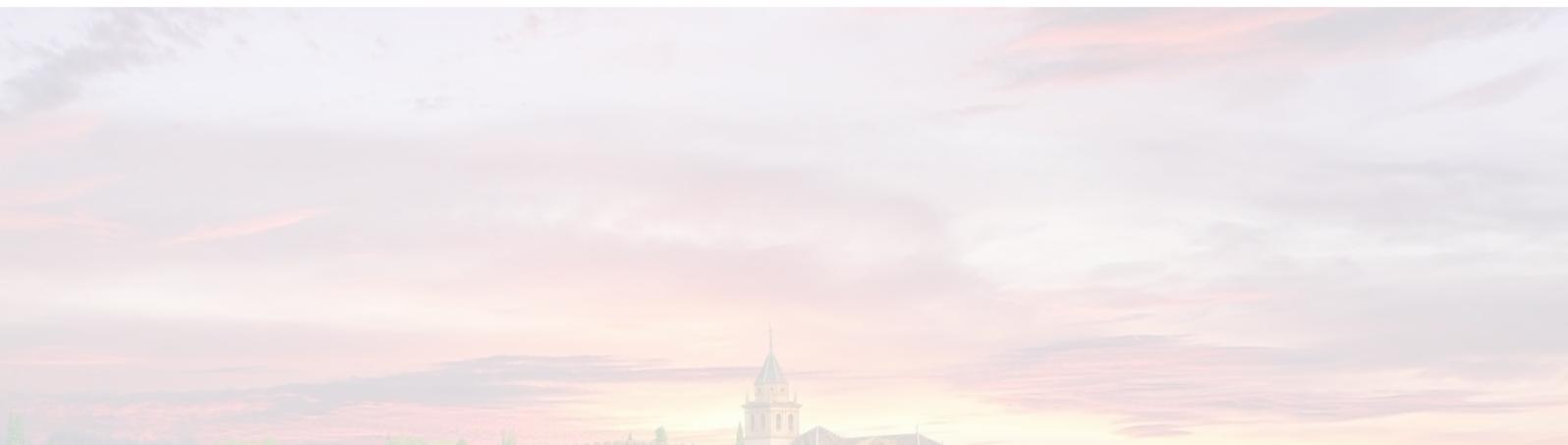
## Modelo Lineal Simple:

$$y_i \sim a + bx_i$$

## Modelo Lineal Múltiple:

Permite estudiar el comportamiento de **una** variable endógena  $Y$ , a partir de  $k$  variables independientes ( $X_2, X_3, \dots, X_k$ )

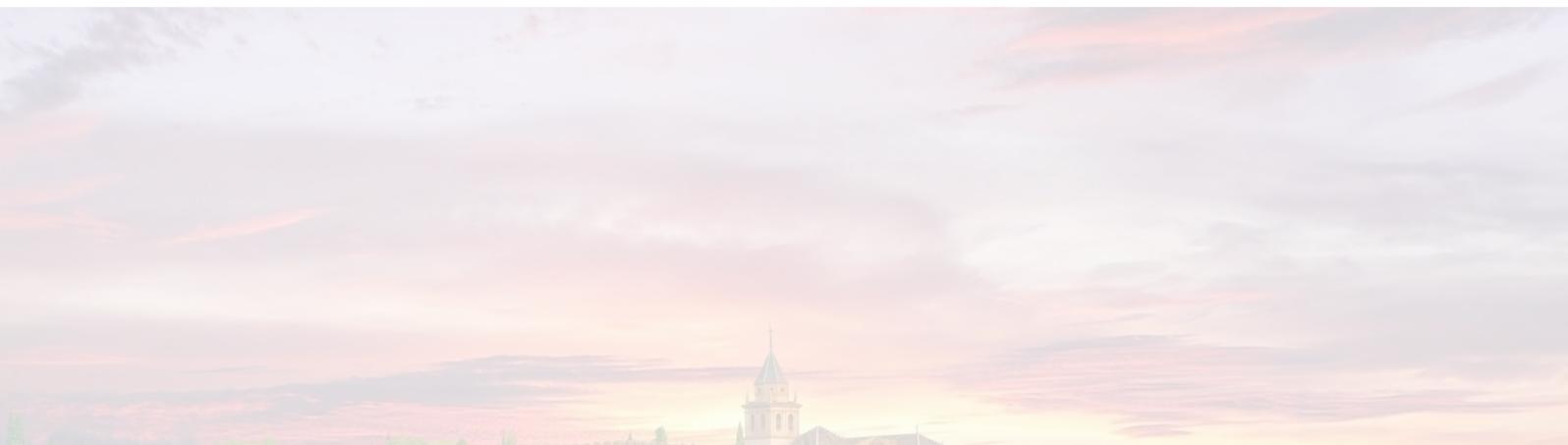
$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i$$



## INTRODUCCIÓN: NOTACIÓN MLP

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i$$

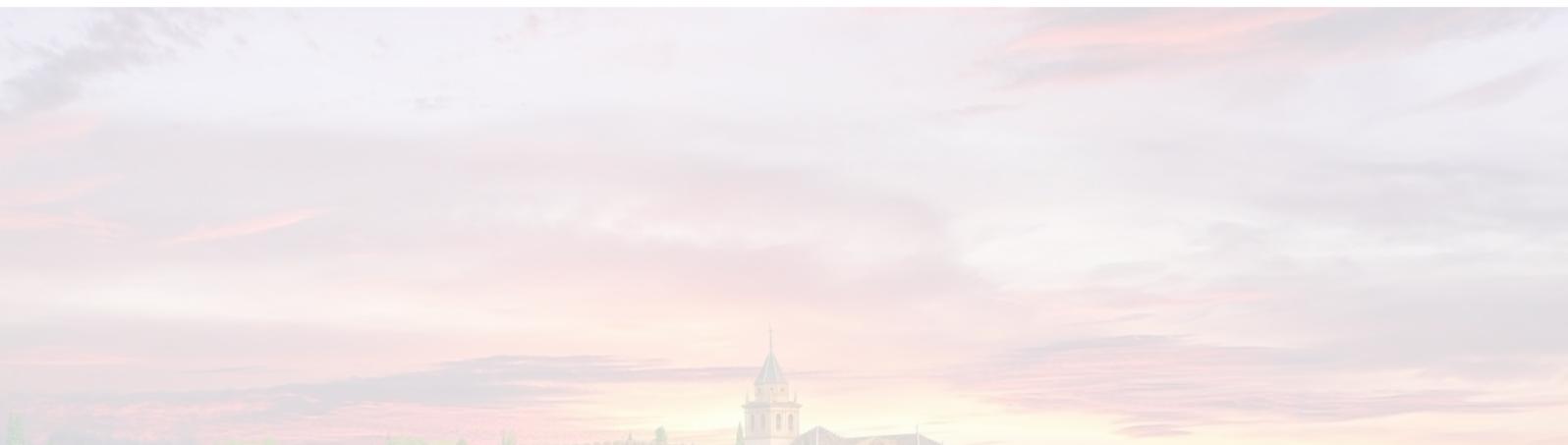
- ▶ el subíndice  $i = 1, \dots, n$ , indica la  $i$ -ésima de las  $n$  observaciones de la muestra.



## INTRODUCCIÓN: NOTACIÓN MLP

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i$$

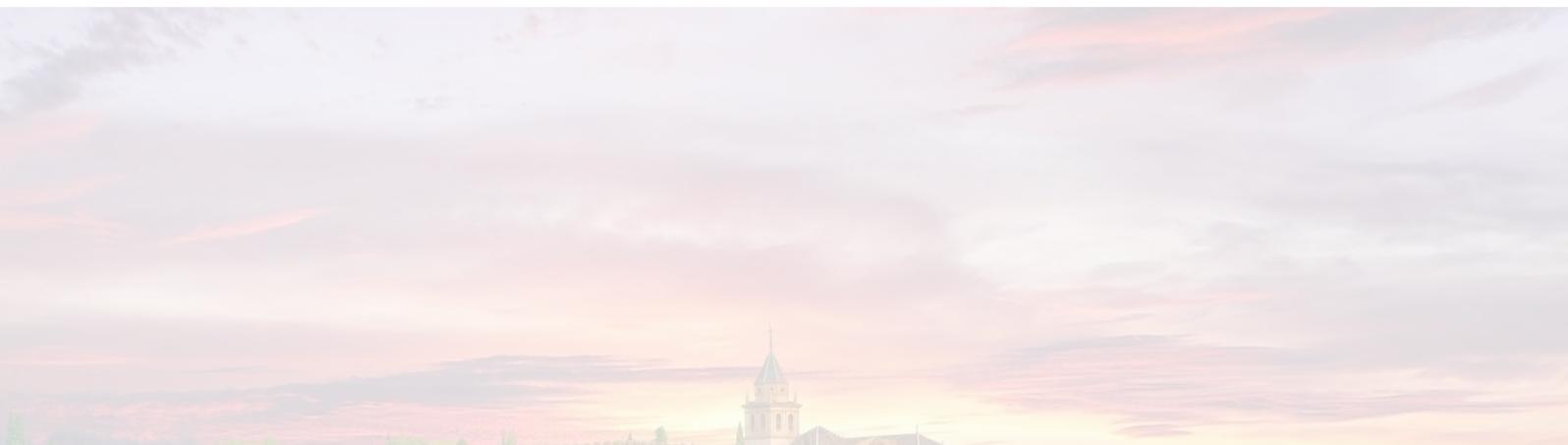
- ▶ el subíndice  $i = 1, \dots, n$ , indica la  $i$ -ésima de las  $n$  observaciones de la muestra.
- ▶ El subíndice  $i$  será empleado cuando trabajemos con observaciones de sección cruzada o de corte transversal, en el caso de trabajar con información de carácter temporal se empleará el subíndice  $t$ .



## INTRODUCCIÓN: NOTACIÓN MLP

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i$$

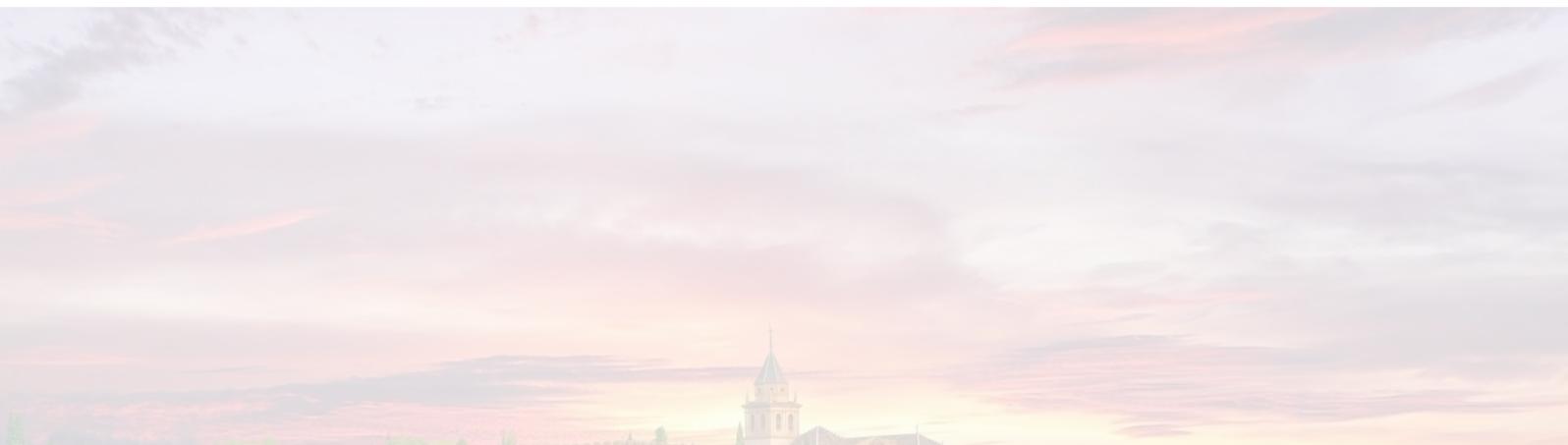
- ▶ el subíndice  $i = 1, \dots, n$ , indica la  $i$ -ésima de las  $n$  observaciones de la muestra.
- ▶ El subíndice  $i$  será empleado cuando trabajemos con observaciones de sección cruzada o de corte transversal, en el caso de trabajar con información de carácter temporal se empleará el subíndice  $t$ .
- ▶  $Y_i$  es la  $i$ -ésima observación de la variable dependiente.



## INTRODUCCIÓN: NOTACIÓN MLP

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i$$

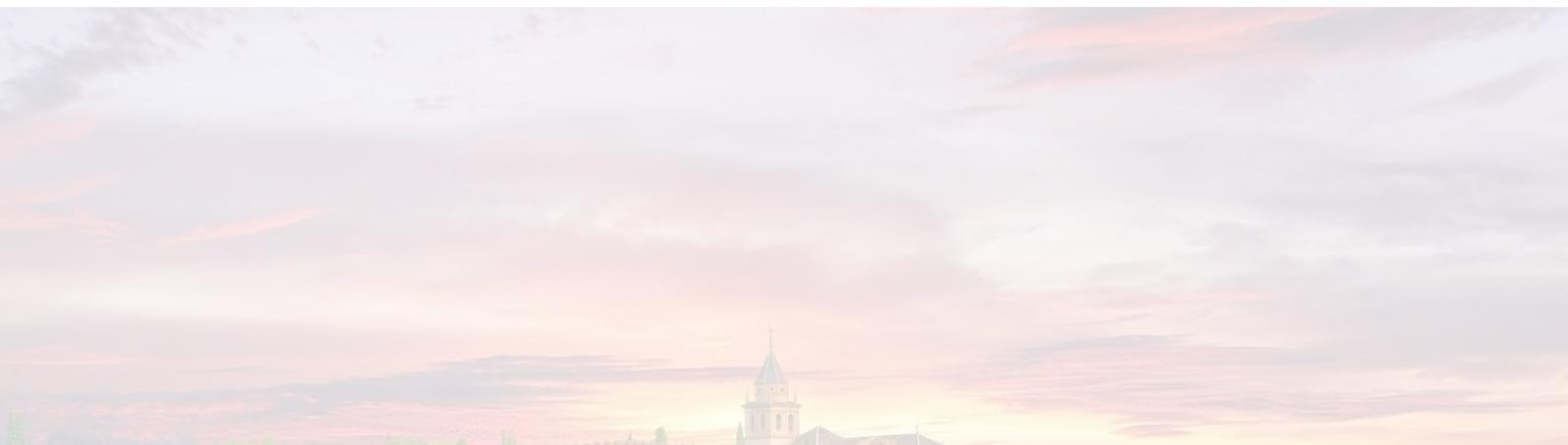
- ▶ el subíndice  $i = 1, \dots, n$ , indica la  $i$ -ésima de las  $n$  observaciones de la muestra.
- ▶ El subíndice  $i$  será empleado cuando trabajemos con observaciones de sección cruzada o de corte transversal, en el caso de trabajar con información de caracter temporal se empleará el subíndice  $t$ .
- ▶  $Y_i$  es la  $i$ -ésima observación de la variable dependiente.
- ▶  $X_{2i}, \dots, X_{ki}$  son las  $i$ -ésimas observaciones de las  $k$  variables independientes.



## INTRODUCCIÓN: NOTACIÓN MLP

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i$$

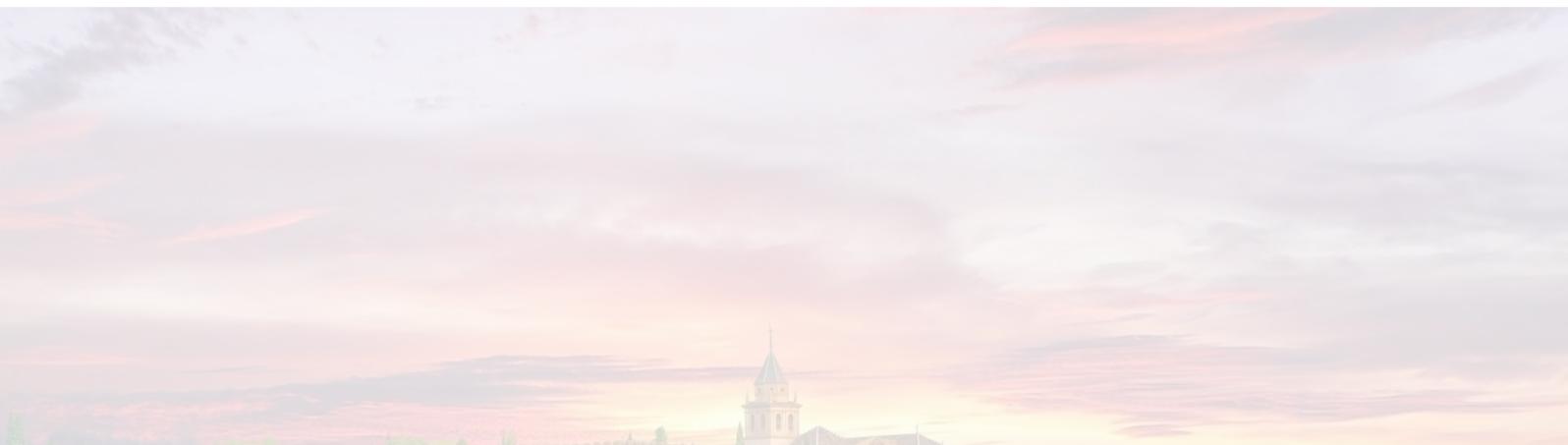
- ▶ el subíndice  $i = 1, \dots, n$ , indica la  $i$ -ésima de las  $n$  observaciones de la muestra.
- ▶ El subíndice  $i$  será empleado cuando trabajemos con observaciones de sección cruzada o de corte transversal, en el caso de trabajar con información de carácter temporal se empleará el subíndice  $t$ .
- ▶  $Y_i$  es la  $i$ -ésima observación de la variable dependiente.
- ▶  $X_{2i}, \dots, X_{ki}$  son las  $i$ -ésimas observaciones de las  $k$  variables independientes.
- ▶  $u_i$  es la perturbación de la  $i$ -ésima observación, encargada de recoger aquella parte de la variable  $Y_i$  que no es explicada por  $X_{1i}, X_{2i}, \dots, X_{ki}$ .



## INTRODUCCIÓN: NOTACIÓN MLP

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i$$

- ▶ el subíndice  $i = 1, \dots, n$ , indica la  $i$ -ésima de las  $n$  observaciones de la muestra.
- ▶ El subíndice  $i$  será empleado cuando trabajemos con observaciones de sección cruzada o de corte transversal, en el caso de trabajar con información de carácter temporal se empleará el subíndice  $t$ .
- ▶  $Y_i$  es la  $i$ -ésima observación de la variable dependiente.
- ▶  $X_{2i}, \dots, X_{ki}$  son las  $i$ -ésimas observaciones de las  $k$  variables independientes.
- ▶  $u_i$  es la perturbación de la  $i$ -ésima observación, encargada de recoger aquella parte de la variable  $Y_i$  que no es explicada por  $X_{1i}, X_{2i}, \dots, X_{ki}$ .
- ▶  $\beta_1, \dots, \beta_k$  son los parámetros del modelo. Además,  $\beta_1$  es el denominado término constante.



## INTRODUCCIÓN: NOTACIÓN MLP

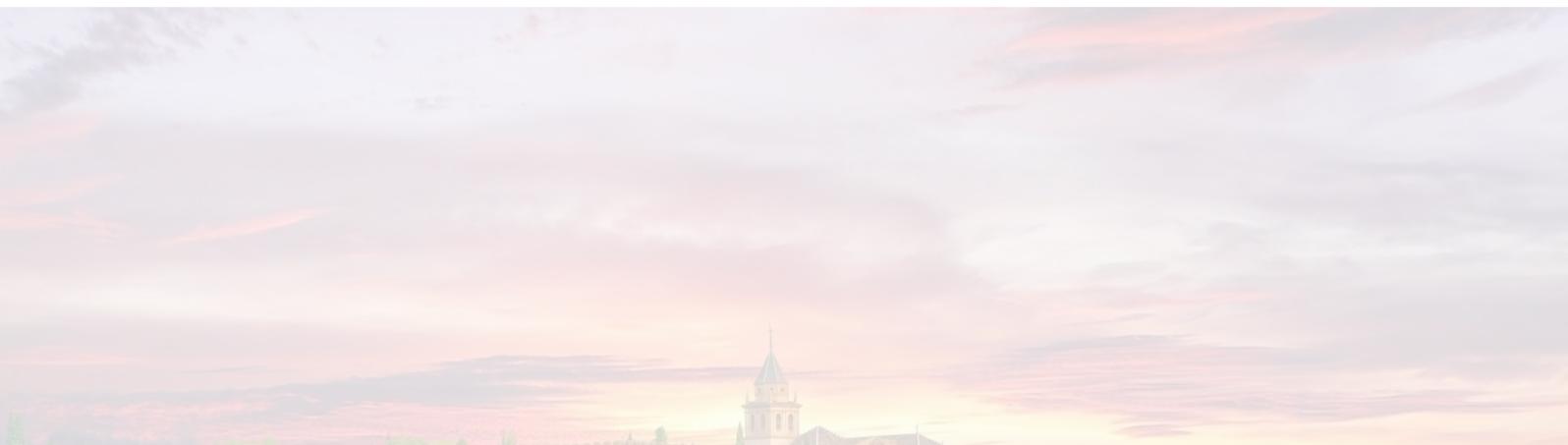
La descripción completa es:

$$Y_1 = \beta_1 + \beta_2 X_{21} + \dots + \beta_k X_{k1} + u_1$$

$$Y_2 = \beta_1 + \beta_2 X_{22} + \dots + \beta_k X_{k2} + u_2$$

⋮

$$Y_n = \beta_1 + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + u_n$$



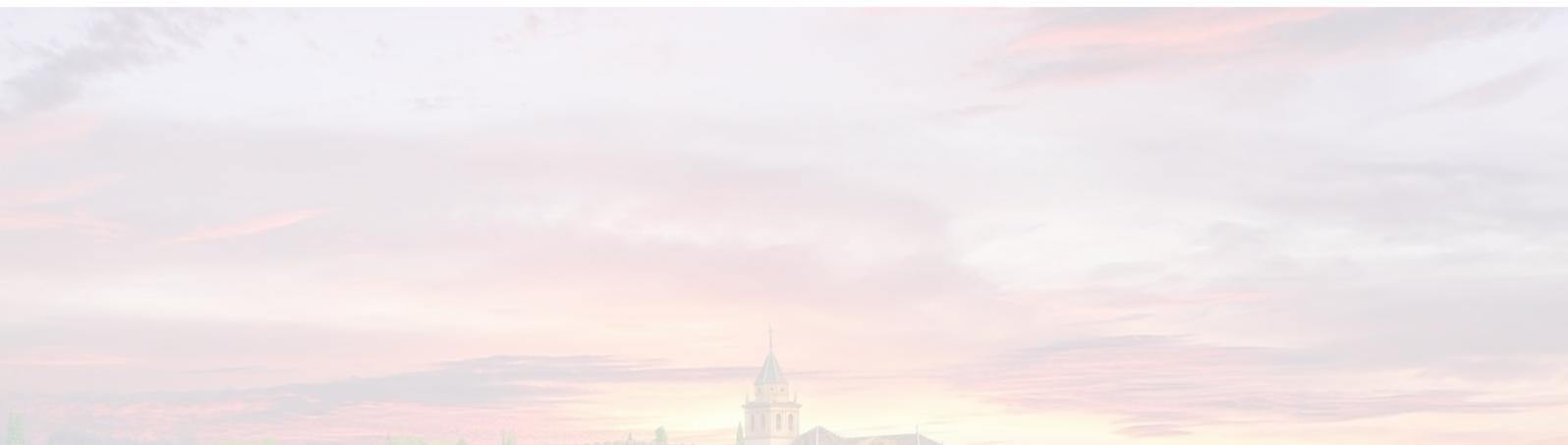
## INTRODUCCIÓN: NOTACIÓN MLP

La descripción completa es:

$$\begin{aligned} Y_1 &= \beta_1 + \beta_2 X_{21} + \dots + \beta_k X_{k1} + u_1 \\ Y_2 &= \beta_1 + \beta_2 X_{22} + \dots + \beta_k X_{k2} + u_2 \\ &\vdots \\ Y_n &= \beta_1 + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + u_n \end{aligned}$$

Equivalentemente:

$$\begin{aligned} Y_1 &= \beta_1 X_{11} + \beta_2 X_{21} + \dots + \beta_k X_{k1} + u_1 \\ Y_2 &= \beta_1 X_{12} + \beta_2 X_{22} + \dots + \beta_k X_{k2} + u_2 \\ &\vdots \\ Y_n &= \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + u_n \end{aligned}$$



## INTRODUCCIÓN: NOTACIÓN MLP

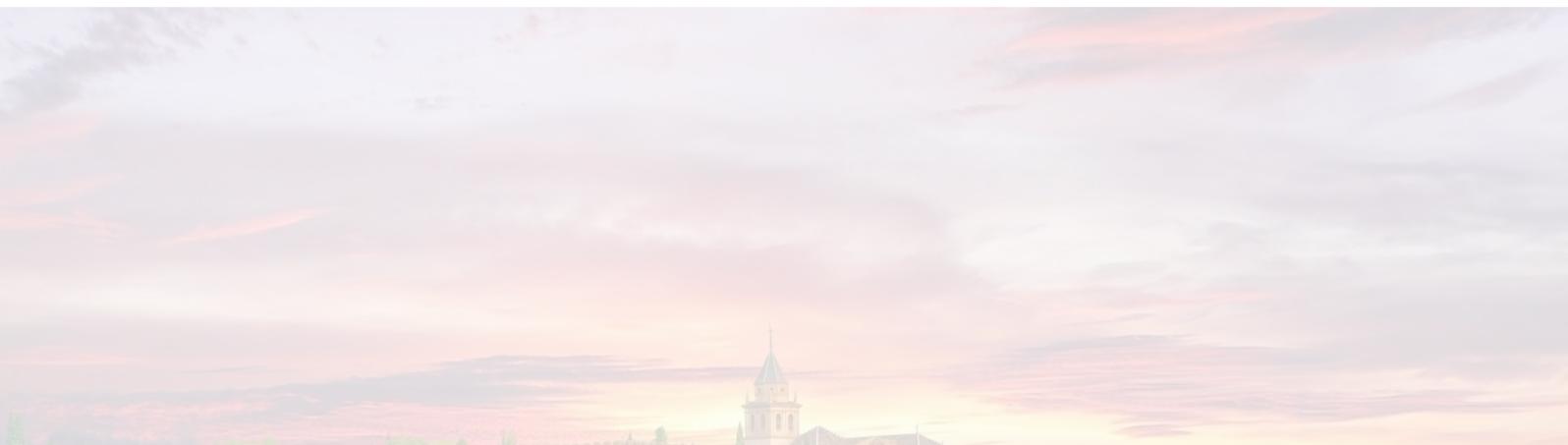
La descripción completa es:

$$\begin{aligned} Y_1 &= \beta_1 + \beta_2 X_{21} + \dots + \beta_k X_{k1} + u_1 \\ Y_2 &= \beta_1 + \beta_2 X_{22} + \dots + \beta_k X_{k2} + u_2 \\ &\vdots \\ Y_n &= \beta_1 + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + u_n \end{aligned}$$

Equivalentemente:

$$\begin{aligned} Y_1 &= \beta_1 X_{11} + \beta_2 X_{21} + \dots + \beta_k X_{k1} + u_1 \\ Y_2 &= \beta_1 X_{12} + \beta_2 X_{22} + \dots + \beta_k X_{k2} + u_2 \\ &\vdots \\ Y_n &= \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + u_n \end{aligned}$$

donde  $X_1 = (1, \dots, 1)$ .



## INTRODUCCIÓN: NOTACIÓN MLP

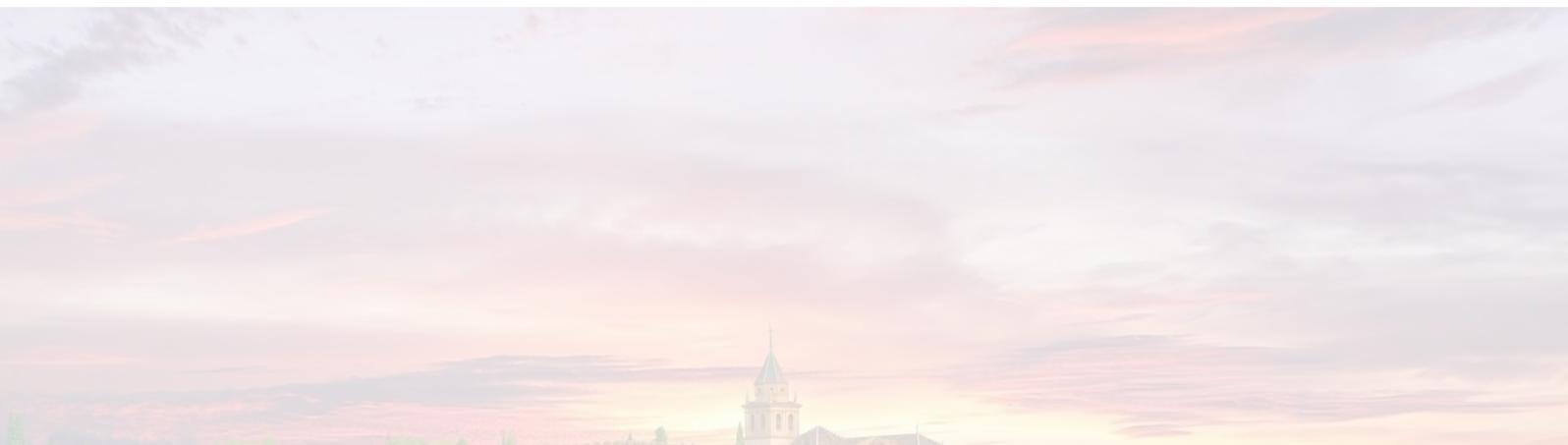
La descripción completa es:

$$\begin{aligned} Y_1 &= \beta_1 + \beta_2 X_{21} + \dots + \beta_k X_{k1} + u_1 \\ Y_2 &= \beta_1 + \beta_2 X_{22} + \dots + \beta_k X_{k2} + u_2 \\ &\vdots \\ Y_n &= \beta_1 + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + u_n \end{aligned}$$

Equivalentemente:

$$\begin{aligned} Y_1 &= \beta_1 X_{11} + \beta_2 X_{21} + \dots + \beta_k X_{k1} + u_1 \\ Y_2 &= \beta_1 X_{12} + \beta_2 X_{22} + \dots + \beta_k X_{k2} + u_2 \\ &\vdots \\ Y_n &= \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + u_n \end{aligned}$$

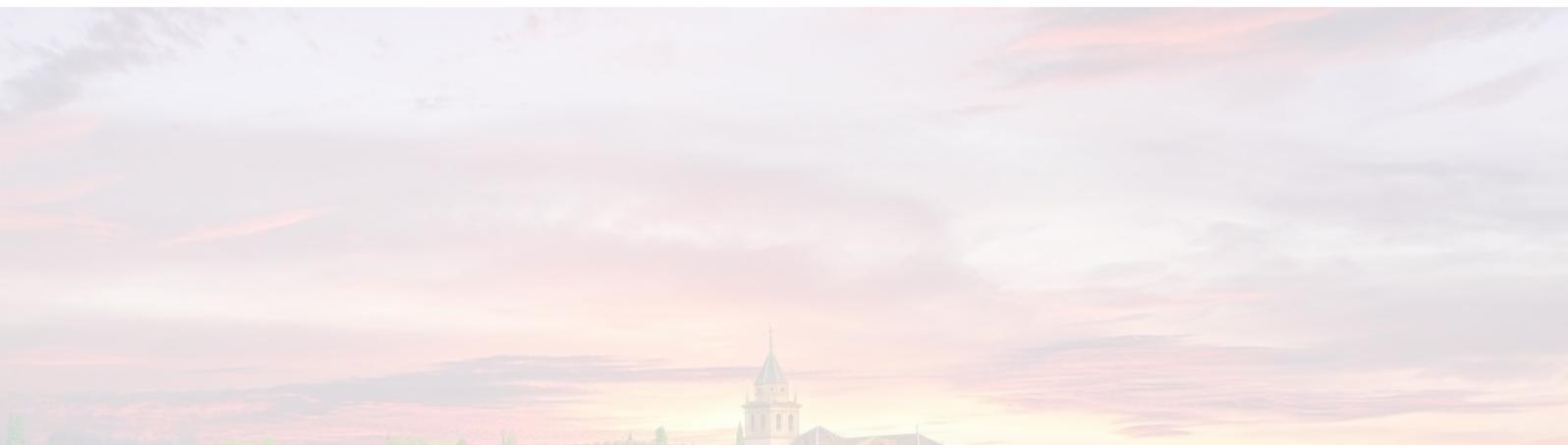
donde  $X_1 = (1, \dots, 1)$ .



## INTRODUCCIÓN: NOTACIÓN MLP

$$\vec{y} = X\vec{\beta} + \vec{u}$$

$$\vec{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad X = \begin{pmatrix} 1 & X_{21} & \dots & X_{k1} \\ 1 & X_{22} & \dots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & \dots & X_{kn} \end{pmatrix} \quad \vec{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \quad \vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$



## EJEMPLO: DATOS INDITEX

Modelo 1:

$$Y_t = \beta_1 + \beta_2 X_{2t} + u_t$$

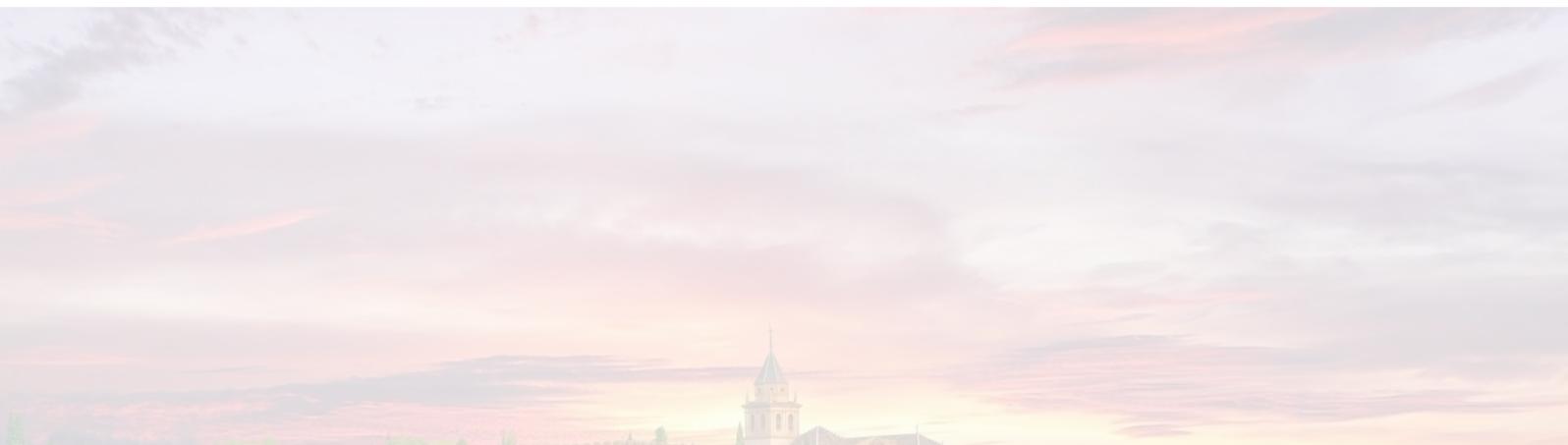
Modelo 2:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$$

Año	Y	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
2006	8196	69240	3.131	NO
2007	9435	79517	3691	NO
2008	10407	89112	4.264	NO
2009	11084	92301	4607	NO
2010	12597	100138	5044	SI
2011	13793	109512	5527	SI
2012	15946	120314	6009	SI
2013	16724	128313	6340	SI
2014	18117	137054	6683	SI
2015	20900	152854	7013	SI
2016	23311	162450	7292	SI
2017	25336	171839	7448	SI
2018	26145	174386	7490	SI
2019	28286	176611	7469	SI
2020	20402	144116	6829	SI

Y: Ingresos (millones de euros)

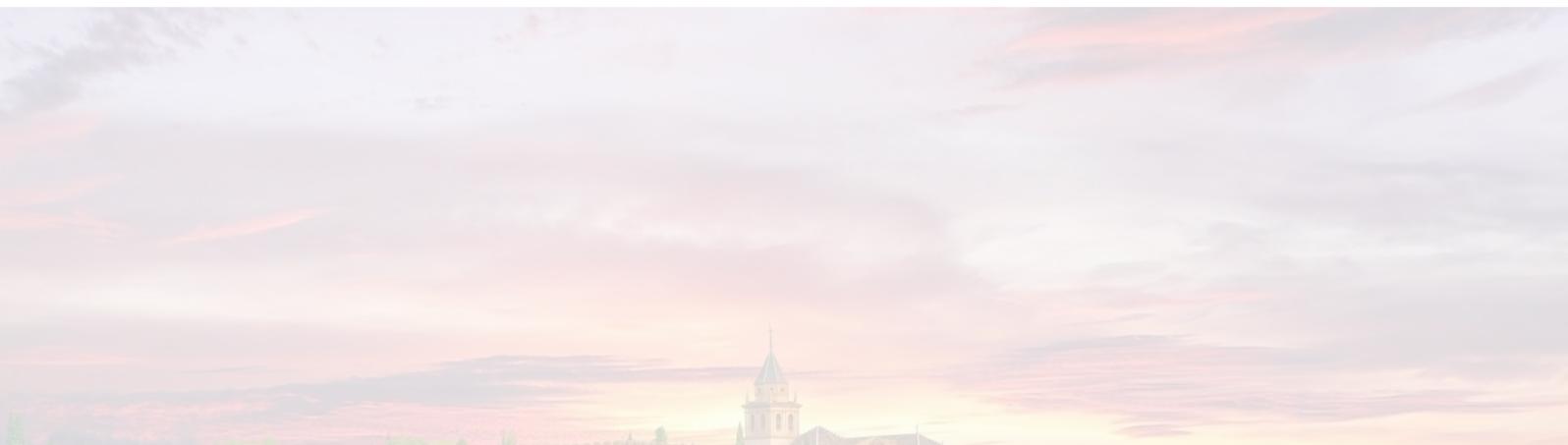
X<sub>2</sub>: Número de empleados; X<sub>3</sub>: Número de tiendas; X<sub>4</sub>: Comercio electrónico (de momento no la usamos)



## SUPUESTOS

Para el análisis estadístico del modelo, comenzaremos considerando las siguiente **hipótesis básicas** sobre el modelo:

- ▶ El supuesto de linealidad.
- ▶ El supuesto rango completo por columnas.
- ▶ El supuesto de exogeneidad.
- ▶ El supuesto de causalidad. El mecanismo de generación de las observaciones.
- ▶ Supuestos sobre el término perturbación.
- ▶ El supuesto de normalidad del término de perturbación.

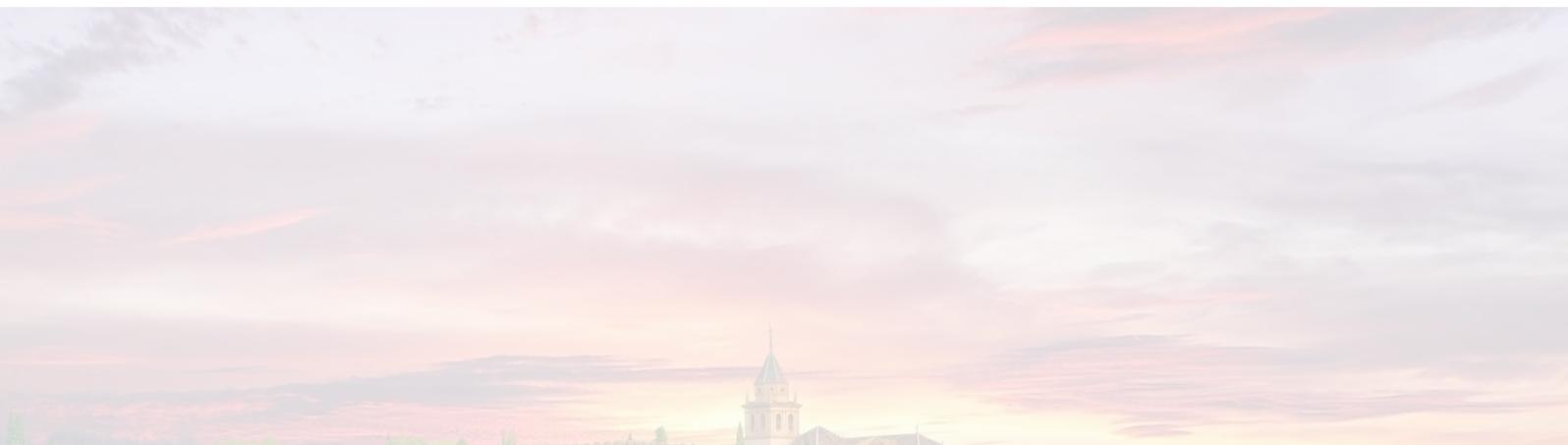


## SUPUESTO DE LINEALIDAD

Sobre el vector  $\vec{y}$  existe una relación entre las variables a nivel poblacional  $X$  y el resto de factores omitidos que son relevantes en la explicación de la variable dependiente,  $\vec{u}$ . Supondremos que esta relación de causalidad es lineal, permitiéndonos así utilizar procedimientos de álgebra de matrices y vectores, lo que simplifica enormemente la operatoria. La relación muestral podemos escribirla como:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i$$

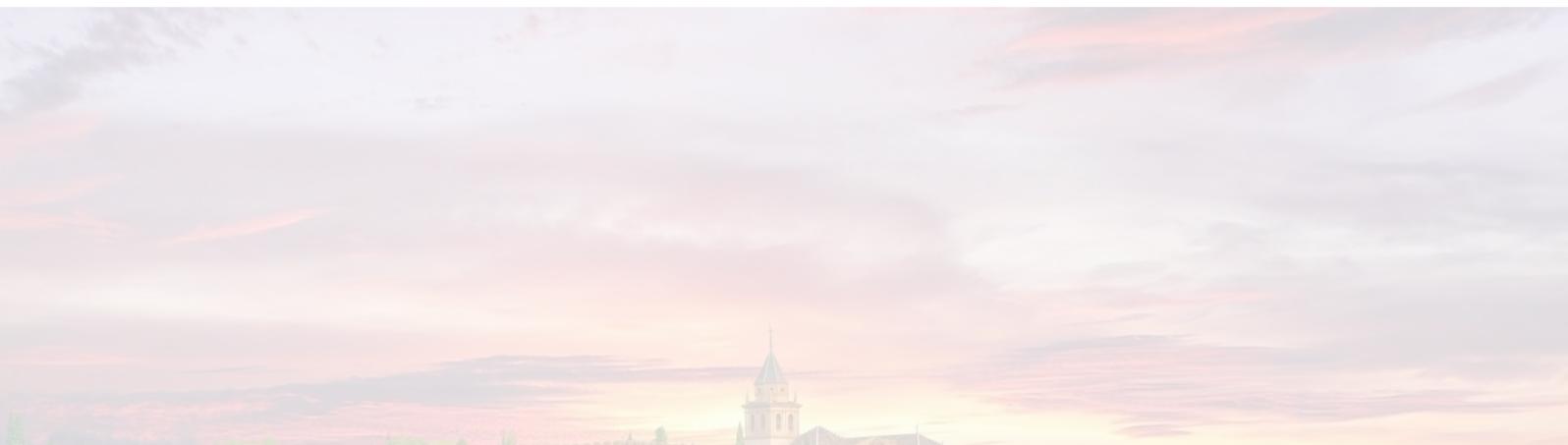
con  $i = 1, \dots, n$  o equivalentemente  $Y_i = \vec{X}_i^t \vec{\beta} + u_i$ .



## SUPUESTO DE RANGO COMPLETO POR COLUMNAS

La matriz de observaciones muestrales  $X_{n \times k}$  tiene rango completo por columnas. Al verificarse que  $n > k$  se tiene que  $\rho(X) = k$ . Por tanto, a nivel muestral, las  $k$  columnas de la matriz  $X$  son linealmente independientes, reflejando la no existencia de una relación lineal entre las variables explicativas.

Este supuesto a efectos del procedimiento de estimación del modelo, nos permite efectuar el cálculo de la inversa de la matriz de productos cruzados de los regresores del modelo, lo que denotaremos en el próximo tema como la matriz  $(X^t X)^{-1}$ . En el caso de que no se verifique el supuesto mencionado, implicaría que alguna de las variables explicativas sería combinación lineal de otras de las variables, produciéndose un problema de **multicolinealidad**.



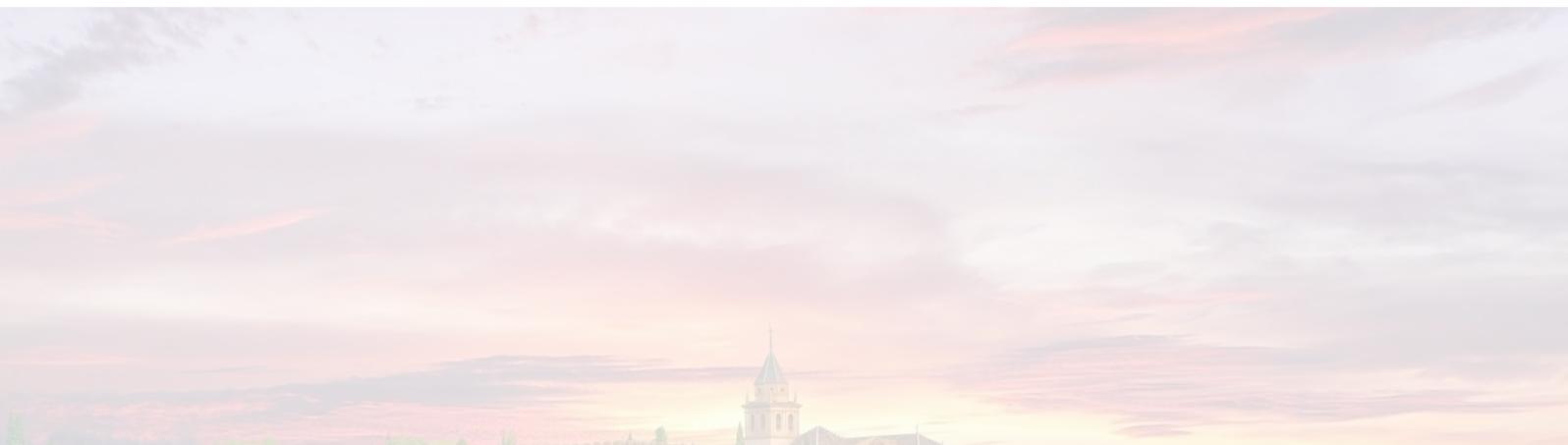
## SUPUESTO DE EXOGENEIDAD

Las variables explicativas del modelo  $X_{1i}, X_{2i}, \dots, X_{ki}$ , con  $i = 1, \dots, n$ , no incorporan información alguna que nos permita predecir el término perturbación,  $u_i$ . Equivalentemente, se tiene que para cualquier valor que tomen las variables explicativas, el valor esperado de la perturbación es cero. Es decir,

$$E [u_i | (X_{1i}, X_{2i}, \dots, X_{ki})] = 0$$

para  $i = 1, \dots, n$ . Por tanto, teniendo en cuenta que  $Y_i = \vec{X}_i^t \vec{\beta} + u_i$ , se obtiene de forma inmediata que

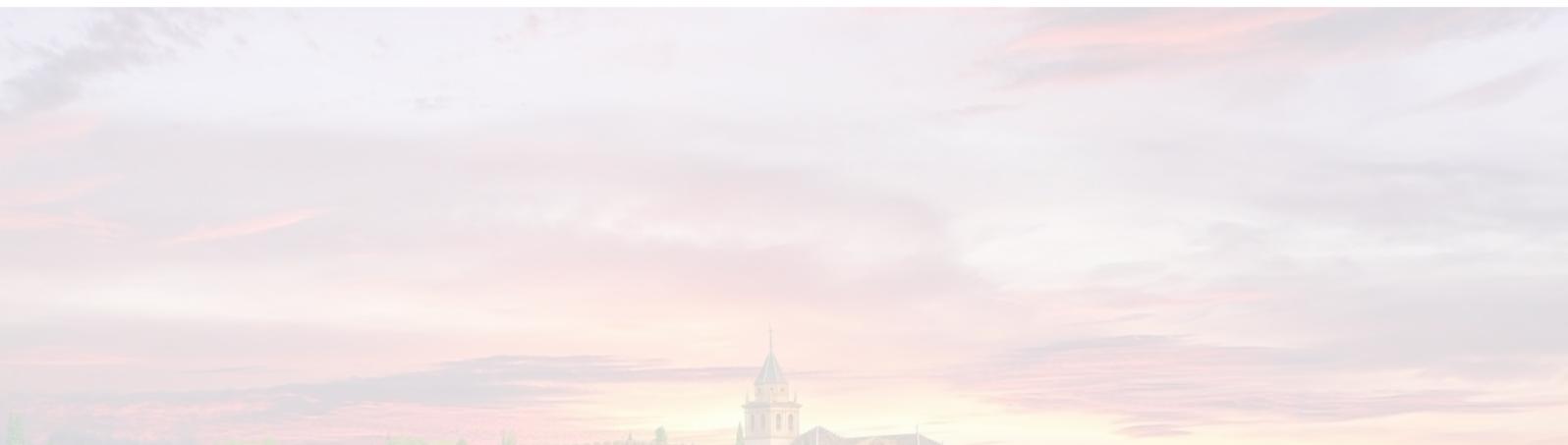
$$E [Y_i | (X_{1i}, X_{2i}, \dots, X_{ki})] = \vec{X}_i^t \vec{\beta}.$$



## SUPUESTO DE CAUSALIDAD

La relación de causalidad será siempre unidireccional, en el sentido de que serán las variables que aparecen en el lado derecho de la ecuación las que expliquen el comportamiento de la variable dependiente pero no a la inversa. Teniendo en cuenta que la variable explicada,  $Y_i$ , depende de la perturbación, se verifica que la variable dependiente adquirirá también un carácter aleatorio.

La matriz de las variables explicativas,  $X$ , se supone no estocástica. Por tanto, el proceso generador de estas observaciones será ajeno al proceso generador del modelo y de la perturbación como consecuencia del supuesto de exogeneidad.



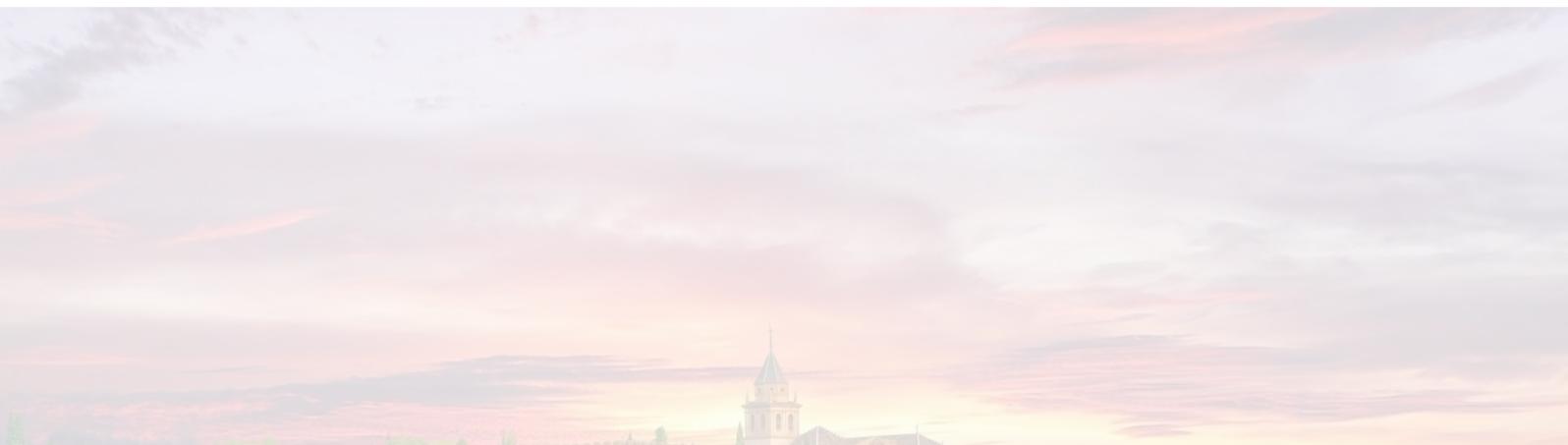
## SUPUESTO SOBRE LA PERTURBACIÓN

La perturbación aleatoria,  $u_i$ , está centrada, es homocedástica e incorrelada. Es decir:

1.  $u_i$  está centrada:  $E[u_i] = 0 \forall i = 1, \dots, n$ .
2. La perturbación es homocedástica: la varianza de la misma se mantiene constante para todas las observaciones, es decir,  $\text{var}[u_i] = E[u_i^2] = \sigma^2 \forall i = 1, \dots, n$ . En el caso de que no se verifique lo indicado, estaríamos ante un problema de **heterocedasticidad**.
3. Si la perturbación es incorrelada provoca que  $\text{cov}[u_i u_j] = E[u_i u_j] = 0 \forall i \neq j$  con  $i = 1, \dots, n$  y  $i = j, \dots, n$ . Así que la perturbación en el momento  $i$  no está correlacionada con el su valor en otro momento determinado  $j$ . Cuando  $\text{cov}[u_i u_j] \neq 0$  entonces estamos ante un problema de **autocorrelación**.

Como consecuencia se tiene que la matriz de varianzas-covarianzas del término perturbación es:

$$\text{var}[\vec{u}] = E[(\vec{u} - E[\vec{u}])(\vec{u} - E[\vec{u}])^t] = E[\vec{u} \vec{u}^t]$$



## EL SUPUESTO DE NORMALIDAD DE LA PERTURBACIÓN

Teniendo en cuenta las hipótesis relacionadas con el término perturbación, y con el fin de realizar inferencia sobre los parámetros que intervienen en el modelo, se considera que la perturbación aleatoria  $u_i$  se distribuye según una normal con media cero y varianza  $\sigma^2$ , es decir

$$u_i \sim N(0, \sigma^2) \quad \forall i = 1, \dots, n$$

Por el momento supondremos que el vector  $\vec{u}$  se distribuye según una distribución normal multivariante con vector de medias  $\vec{0}$  y matriz de varianzas-covarianzas  $\sigma^2 \mathbf{I}_n$ , es decir  $\vec{u} \sim N(\vec{0}, \sigma^2 \mathbf{I}_n)$

La función de densidad asociada a dicha distribución viene dada por la siguiente expresión:

$$f(\vec{u} | \sigma^2) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp \left[ -\frac{1}{2\sigma^2} \vec{u}^t \vec{u} \right]$$

## ESTIMACIÓN: MÍNIMOS CUADRADOS ORDINARIOS

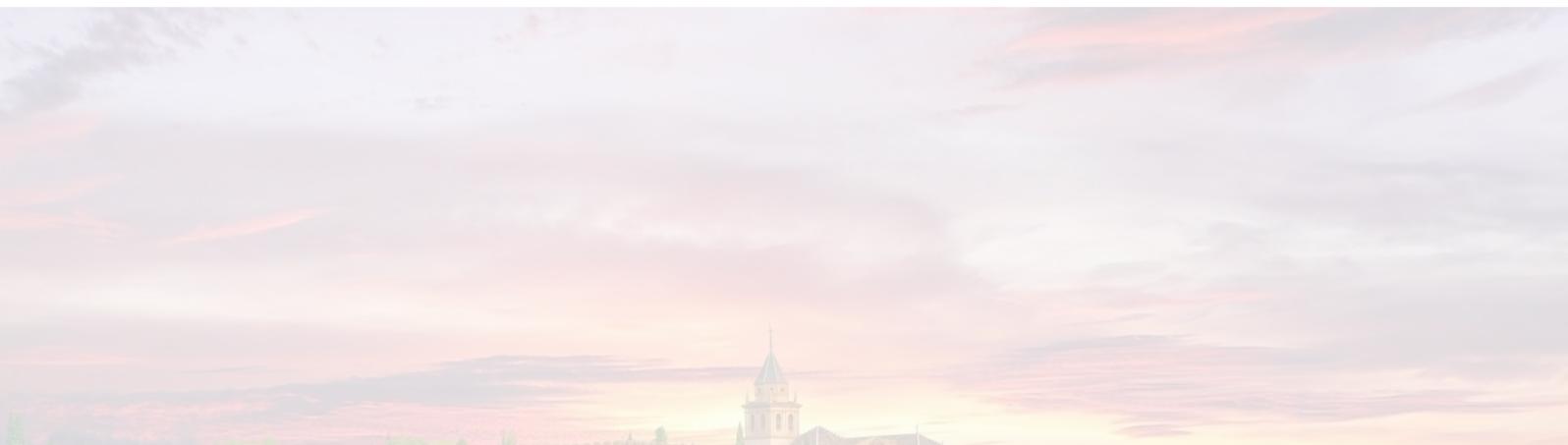
Consideremos el modelo lineal múltiple clásico en su notación matricial

$$\vec{y} = X\vec{\beta} + \vec{u}$$

Para estimar los parámetros  $\vec{\beta}$  utilizaremos el **Método de los Mínimos Cuadrados Ordinarios (MCO)**.

Denotando por  $\vec{\hat{y}}$  el valor **ajustado** de  $\vec{y}$  por el modelo:

$$\vec{\hat{y}} = X\vec{\hat{\beta}}$$



## ESTIMACIÓN: MÍNIMOS CUADRADOS ORDINARIOS

Consideremos el modelo lineal múltiple clásico en su notación matricial

$$\vec{y} = X\vec{\beta} + \vec{u}$$

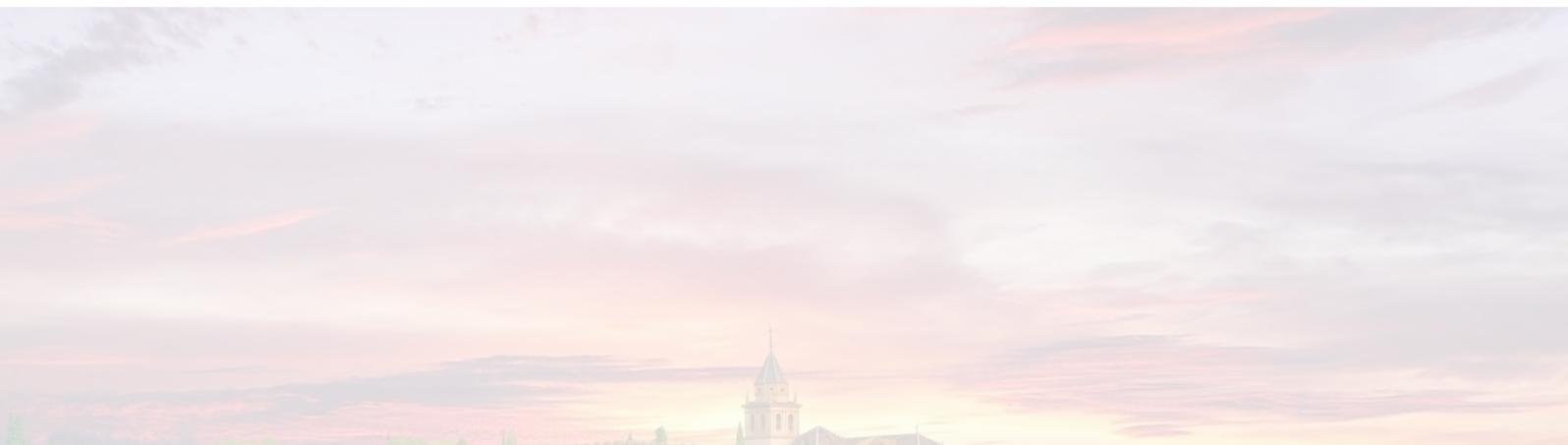
Para estimar los parámetros  $\vec{\beta}$  utilizaremos el **Método de los Mínimos Cuadrados Ordinarios (MCO)**.

Denotando por  $\vec{\hat{y}}$  el valor **ajustado** de  $\vec{y}$  por el modelo:

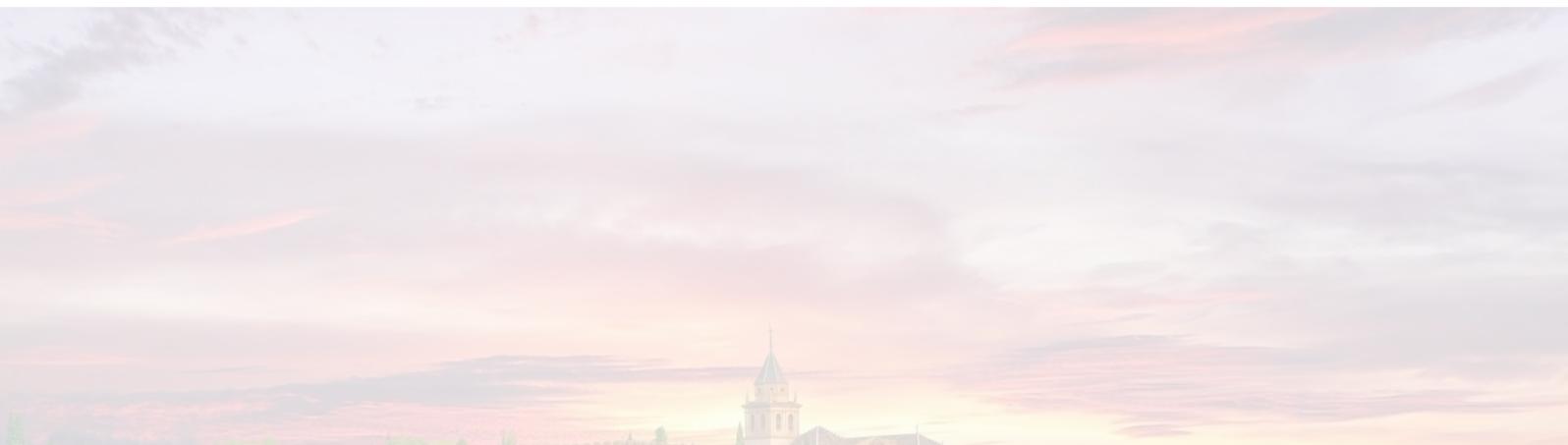
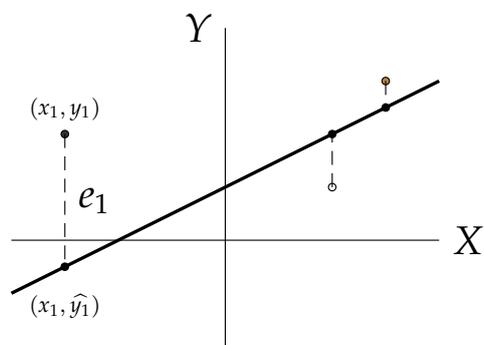
$$\vec{\hat{y}} = X\vec{\hat{\beta}}$$

[Residuo] Diferencia existente entre el valor observado de la variable explicada y su estimación:

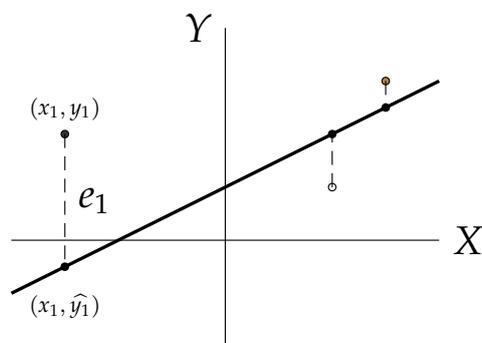
$$\vec{e} = \vec{y} - \vec{\hat{y}} = \vec{y} - X\vec{\hat{\beta}}$$



# ESTIMACIÓN: MÍNIMOS CUADRADOS ORDINARIOS



## ESTIMACIÓN: MÍNIMOS CUADRADOS ORDINARIOS



Minimizaremos la suma de los cuadrados de los residuos del ajuste:

$$\min_{\vec{\beta}} \sum_{i=1}^n e_i^2$$

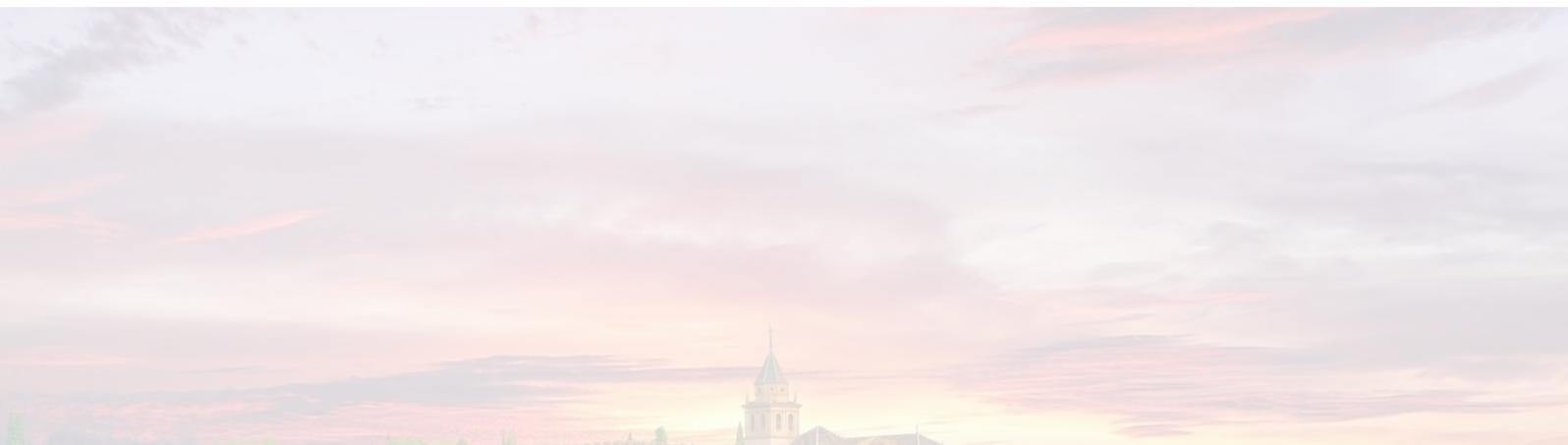
donde  $e_i = y_i - \hat{y}_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_{2i} - \dots - \hat{\beta}_k x_{ki}$  (componentes de  $\vec{e}$ )

Denotaremos a la función a minimizar por:  $f(\vec{\beta}) = \vec{e}^t \vec{e}$ .

## ESTIMACIÓN: MÍNIMOS CUADRADOS ORDINARIOS

**Paso 1** Definimos la expresión a minimizar. Se define la suma de los cuadrados de los residuos como:

$$\begin{aligned} f(\vec{\beta}) &= (\vec{y} - X\vec{\beta})^t (\vec{y} - X\vec{\beta}) \\ &= \vec{y}^t \vec{y} - \vec{y}^t X \vec{\beta} - \vec{\beta}^t X^t \vec{y} + \vec{\beta}^t X^t X \vec{\beta} \\ &= \vec{y}^t \vec{y} - 2 \vec{\beta}^t X^t \vec{y} + \vec{\beta}^t X^t X \vec{\beta} \end{aligned}$$



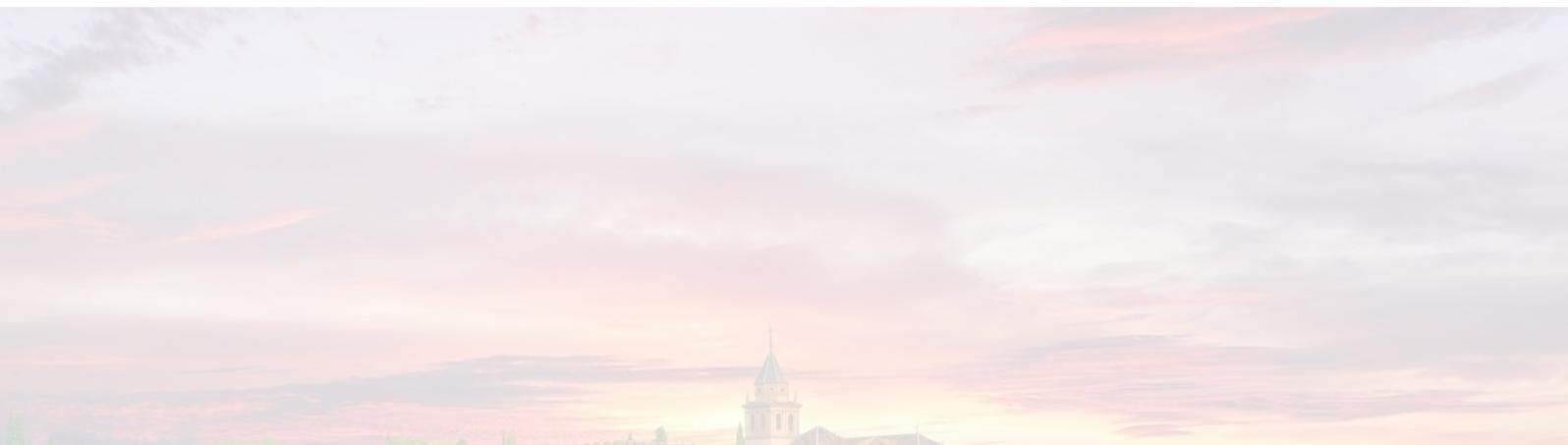
## ESTIMACIÓN: MÍNIMOS CUADRADOS ORDINARIOS

**Paso 1** Definimos la expresión a minimizar. Se define la suma de los cuadrados de los residuos como:

$$\begin{aligned} f(\vec{\beta}) &= (\vec{y} - X\vec{\beta})^t (\vec{y} - X\vec{\beta}) \\ &= \vec{y}^t \vec{y} - \vec{y}^t X \vec{\beta} - \vec{\beta}^t X^t \vec{y} + \vec{\beta}^t X^t X \vec{\beta} \\ &= \vec{y}^t \vec{y} - 2 \vec{\beta}^t X^t \vec{y} + \vec{\beta}^t X^t X \vec{\beta} \end{aligned}$$

**Paso 2** Derivamos con respecto a  $\vec{\beta}$  e igualamos a cero.

$$\nabla f(\vec{\beta}) = -2X^t \vec{y} + 2X^t X \vec{\beta} = \vec{0}$$



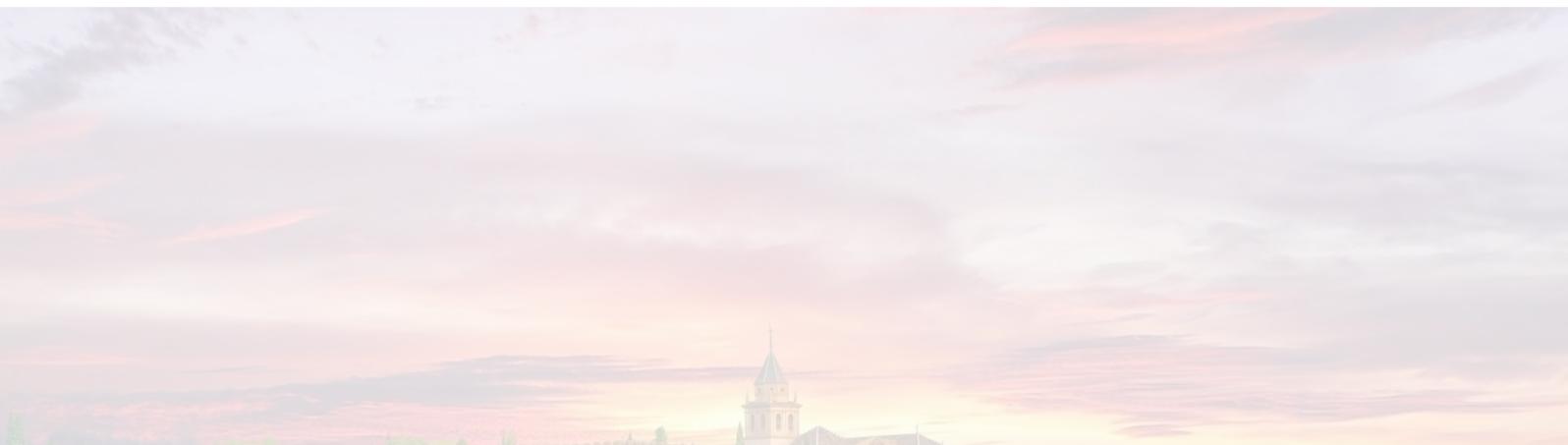
## ESTIMACIÓN: MÍNIMOS CUADRADOS ORDINARIOS

Paso 3 Despejamos el vector de parámetros.

$$-2X^t\vec{y} + 2X^tX\vec{\beta} = \vec{0} \Leftrightarrow X^tX\vec{\beta} = X^t\vec{y}$$

Considerando el supuesto de rango completo por columnas, sabemos que se verifica que  $\rho(X^tX) = k$ , condición necesaria para la existencia de la matriz  $(X^tX)^{-1}$ , luego:

$$\vec{\beta} = (X^tX)^{-1} (X^t\vec{y})$$



## ESTIMACIÓN: MÍNIMOS CUADRADOS ORDINARIOS

Paso 3 Despejamos el vector de parámetros.

$$-2X^t\vec{y} + 2X^tX\vec{\beta} = \vec{0} \Leftrightarrow X^tX\vec{\beta} = X^t\vec{y}$$

Considerando el supuesto de rango completo por columnas, sabemos que se verifica que  $\rho(X^tX) = k$ , condición necesaria para la existencia de la matriz  $(X^tX)^{-1}$ , luego:

$$\vec{\beta} = (X^tX)^{-1} (X^t\vec{y})$$

Paso 4 Para que la solución sea mínimo se ha de cumplir:

$$Hess(f) = 2X^tX \succ 0 \quad (A \succ 0 \text{ sii } z^tAz > 0)$$

$$\vec{\beta} = (X^tX)^{-1} (X^t\vec{y})$$

es el **Estimador Mínimo Cuadrático Ordinario (EMCO)** del vector de parámetros  $\vec{\beta}$  del modelo de regresión lineal múltiple  $\vec{y} = X\vec{\beta} + \vec{u}$ .

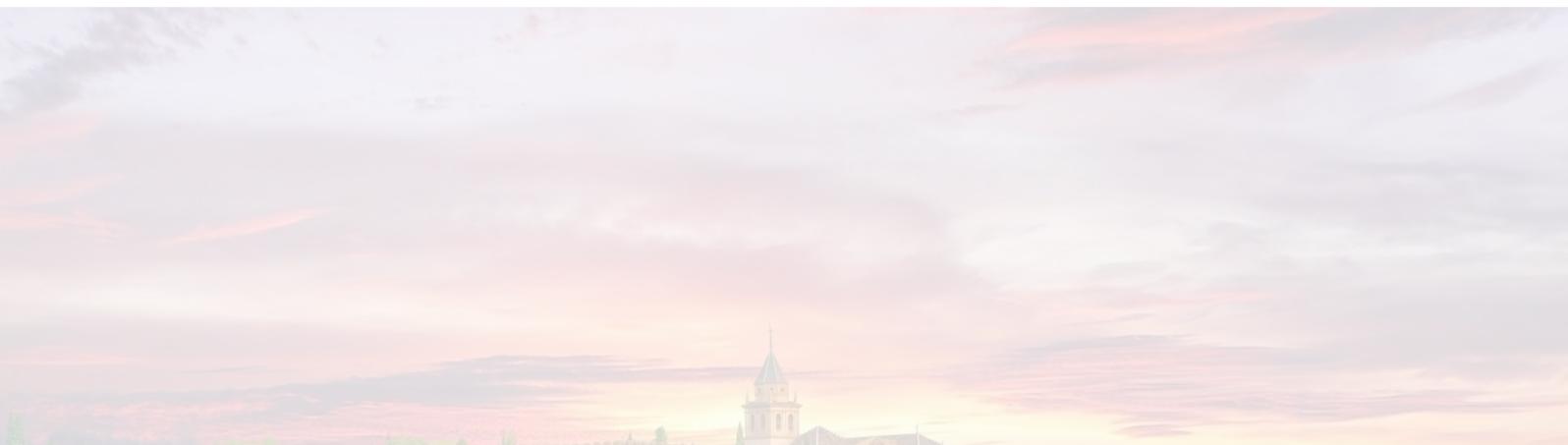
## ESTIMACIÓN: MÍNIMOS CUADRADOS ORDINARIOS

Si el modelo tiene término independiente entonces se tiene

$$\vec{\beta} = \begin{pmatrix} n & \sum_{i=1}^n X_{2i} & \dots & \sum_{i=1}^n X_{ki} \\ \sum_{i=1}^n X_{2i} & \sum_{i=1}^n X_{2i}^2 & \dots & \sum_{i=1}^n X_{2i}X_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n X_{ki} & \sum_{i=1}^n X_{ki}X_{2i} & \dots & \sum_{i=1}^n X_{ki}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_{2i}Y_i \\ \vdots \\ \sum_{i=1}^n X_{ki}Y_i \end{pmatrix}$$

En el caso de que el modelo carezca del término independiente entonces:

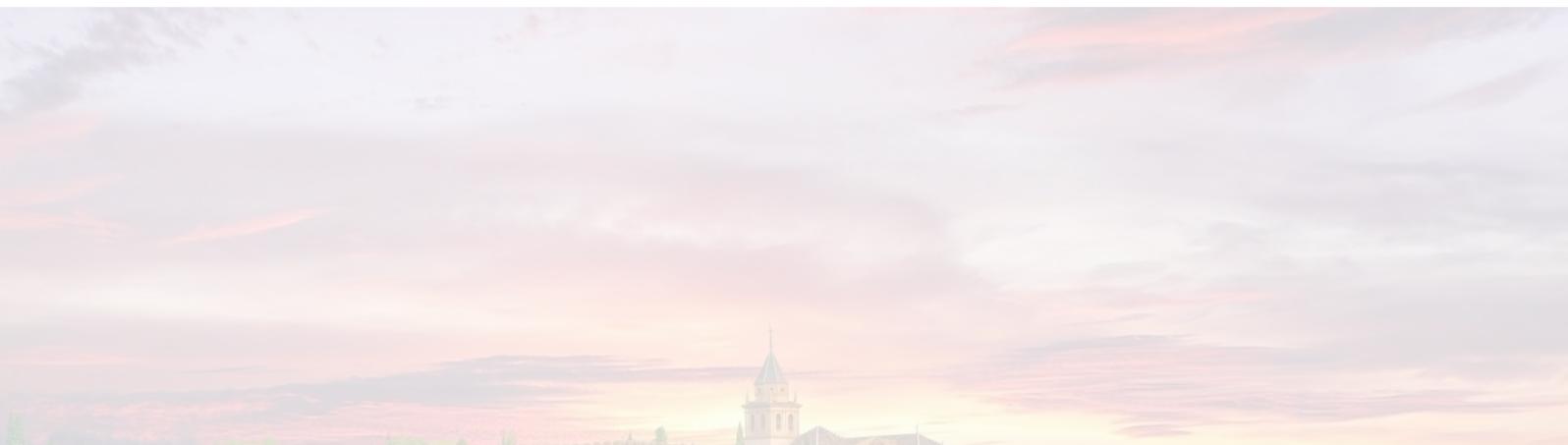
$$\vec{\beta} = \begin{pmatrix} \sum_{i=1}^n X_{1i}^2 & \sum_{i=1}^n X_{1i}X_{2i} & \dots & \sum_{i=1}^n X_{1i}X_{ki} \\ \sum_{i=1}^n X_{2i}X_{1i} & \sum_{i=1}^n X_{2i}^2 & \dots & \sum_{i=1}^n X_{2i}X_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n X_{ki}X_{1i} & \sum_{i=1}^n X_{ki}X_{2i} & \dots & \sum_{i=1}^n X_{ki}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^n X_{1i}Y_i \\ \sum_{i=1}^n X_{2i}Y_i \\ \vdots \\ \sum_{i=1}^n X_{ki}Y_i \end{pmatrix}$$



## PROPIEDADES ALGEBRAICAS DE LOS MCO

Considerando que **SÍ** hay término independiente en el modelo ( $X_{1i} = 1$ ), se tiene:

- ▶ Las variables exógenas son ortogonales al vector de los residuos:  
 $X^t \vec{e} = \vec{0}$ .



## PROPIEDADES ALGEBRAICAS DE LOS MCO

Considerando que **SÍ** hay término independiente en el modelo ( $X_{1i} = 1$ ), se tiene:

- ▶ Las variables exógenas son ortogonales al vector de los residuos:  
 $X^t \vec{e} = \vec{0}$ .

La condición necesaria de EMCO era:

$$-2X^t \vec{y} + 2X^t X \vec{\beta} = \vec{0}$$

Equivalentemente:

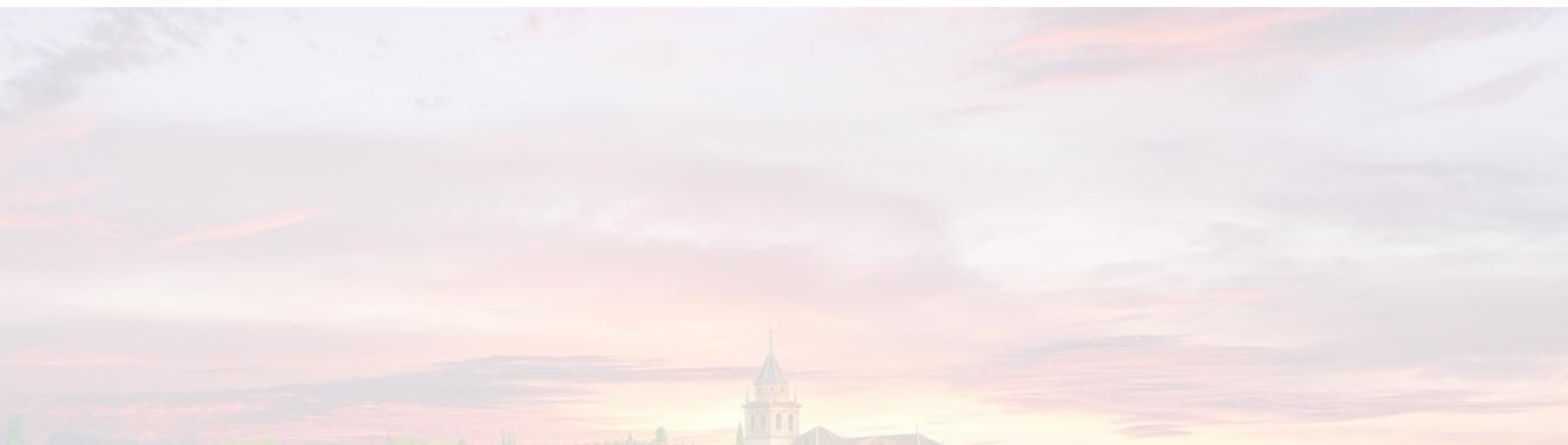
$$X^t \left( \vec{y} - X \vec{\beta} \right) = \vec{0}$$

Como  $\vec{\hat{y}} = X \vec{\hat{\beta}}$ , tenemos que:  $X^t \left( \vec{y} - \vec{\hat{y}} \right) = \vec{0}$ , ó equivalentemente

$$X^t \vec{e} = \vec{0}.$$

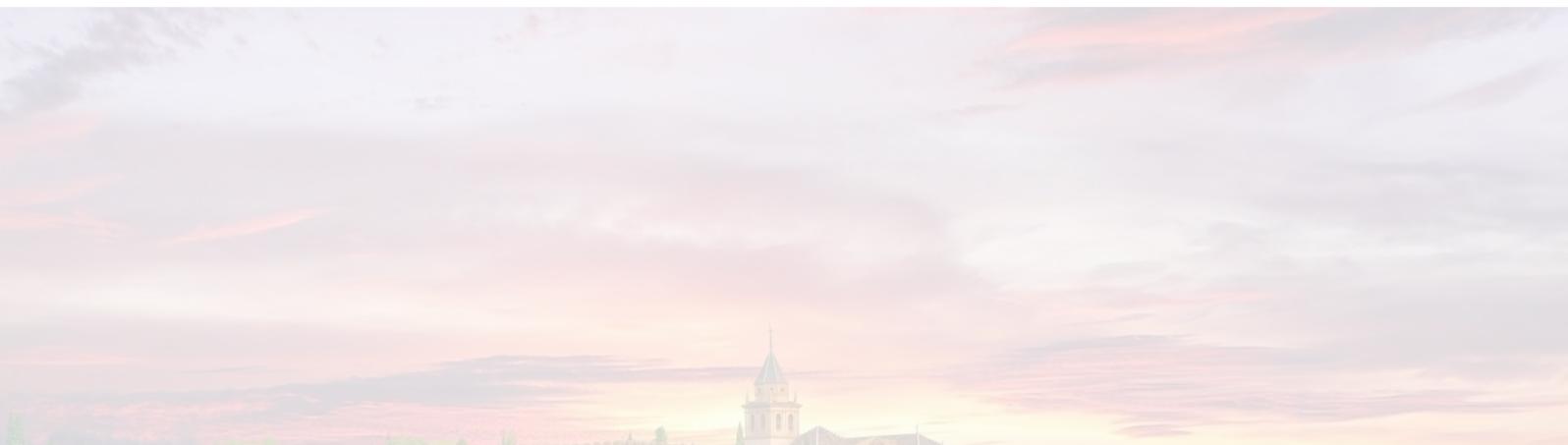
## PROPIEDADES ALGEBRAICAS DE LOS MCO

- ▶ La suma de los residuos mínimo cuadráticos es cero:  $\sum_{i=1}^n e_i = 0$ .



## PROPIEDADES ALGEBRAICAS DE LOS MCO

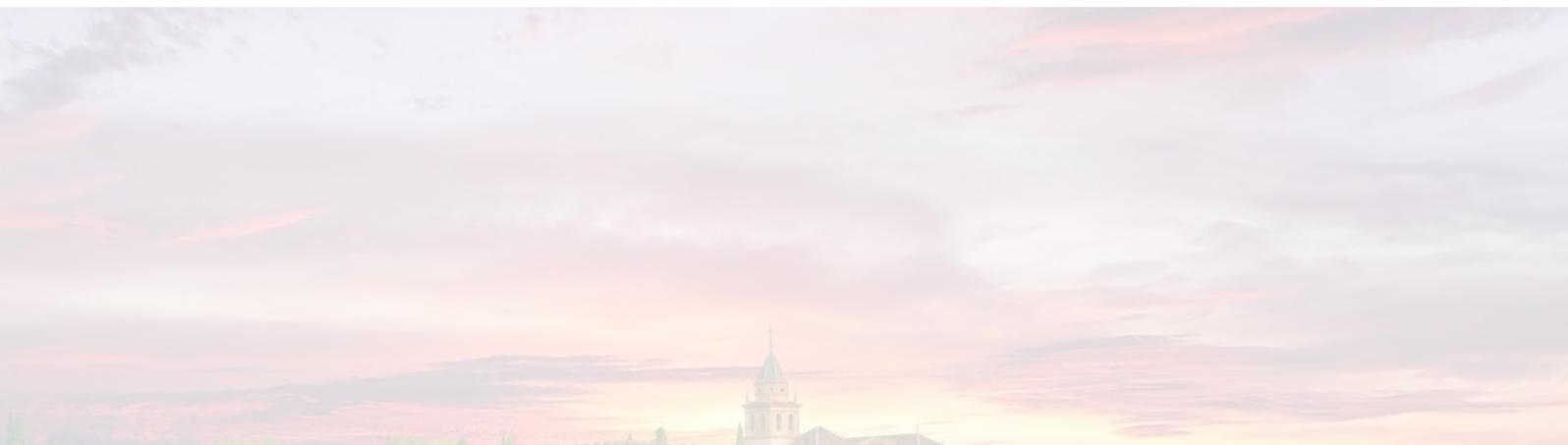
- ▶ La suma de los residuos mínimo cuadráticos es cero:  $\sum_{i=1}^n e_i = 0$ .
- ▶ La suma de los valores observados coincide con la suma de los valores estimados:  $\sum_{i=1}^n Y_i = \sum_{i=1}^n \hat{Y}_i$ .



## PROPIEDADES ALGEBRAICAS DE LOS MCO

- ▶ La suma de los residuos mínimo cuadráticos es cero:  $\sum_{i=1}^n e_i = 0$ .
- ▶ La suma de los valores observados coincide con la suma de los valores estimados:  $\sum_{i=1}^n Y_i = \sum_{i=1}^n \hat{Y}_i$ .
- ▶ Definiendo:

$$\begin{aligned}SCT &= \sum_{i=1}^n (y_i - \bar{y})^2 = \vec{y}^t \vec{y} - n \cdot \bar{Y}^2 \\SCR &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \vec{y}^t \vec{y} - \vec{\beta}^t \mathbf{X}^t \vec{y} \\SCE &= \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \vec{\beta}^t \mathbf{X}^t \vec{y} - n \cdot \bar{Y}^2\end{aligned}$$



## PROPIEDADES ALGEBRAICAS DE LOS MCO

- ▶ La suma de los residuos mínimo cuadráticos es cero:  $\sum_{i=1}^n e_i = 0$ .
- ▶ La suma de los valores observados coincide con la suma de los valores estimados:  $\sum_{i=1}^n Y_i = \sum_{i=1}^n \hat{Y}_i$ .
- ▶ Definiendo:

$$\begin{aligned}SCT &= \sum_{i=1}^n (y_i - \bar{y})^2 = \vec{y}^t \vec{y} - n \cdot \bar{Y}^2 \\SCR &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \vec{y}^t \vec{y} - \vec{\beta}^t \mathbf{X}^t \vec{y} \\SCE &= \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \vec{\beta}^t \mathbf{X}^t \vec{y} - n \cdot \bar{Y}^2\end{aligned}$$

Luego:  $SCT = SCR + SCE$ , donde  $SCT$  es la suma de los cuadrados totales,  $SCR$  es la suma de los cuadrados de los residuos y  $SCE$  es la suma de los cuadrados explicada.

## PROPIEDADES ALGEBRAICAS DE LOS MCO

- ▶ La suma de los residuos mínimo cuadráticos es cero:  $\sum_{i=1}^n e_i = 0$ .
- ▶ La suma de los valores observados coincide con la suma de los valores estimados:  $\sum_{i=1}^n Y_i = \sum_{i=1}^n \hat{Y}_i$ .
- ▶ Definiendo:

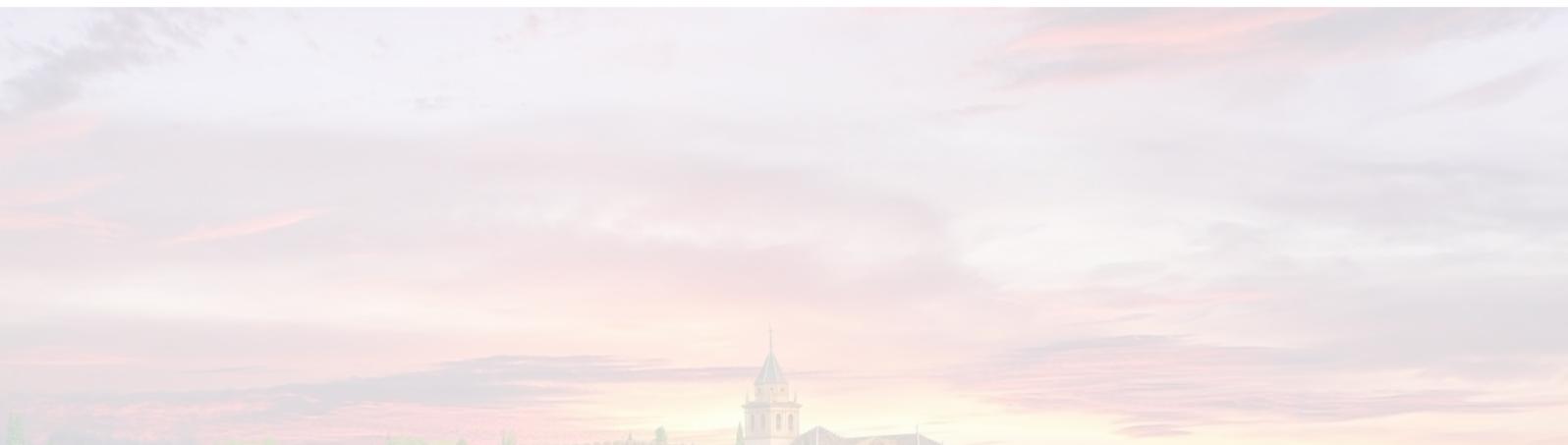
$$\begin{aligned}SCT &= \sum_{i=1}^n (y_i - \bar{y})^2 = \vec{y}^t \vec{y} - n \cdot \bar{Y}^2 \\SCR &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \vec{y}^t \vec{y} - \vec{\beta}^t \mathbf{X}^t \vec{y} \\SCE &= \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \vec{\beta}^t \mathbf{X}^t \vec{y} - n \cdot \bar{Y}^2\end{aligned}$$

Luego:  $SCT = SCR + SCE$ , donde  $SCT$  es la suma de los cuadrados totales,  $SCR$  es la suma de los cuadrados de los residuos y  $SCE$  es la suma de los cuadrados explicada.

- ▶ Se cumple que  $\vec{\hat{y}}^t \vec{e} = 0$ .

## LAS ECUACIONES NORMALES

$$-2X^t\vec{y} + 2X^tX\vec{\beta} = \vec{0} \Rightarrow X^t(\mathbf{y} - \hat{\mathbf{y}}) = 0$$

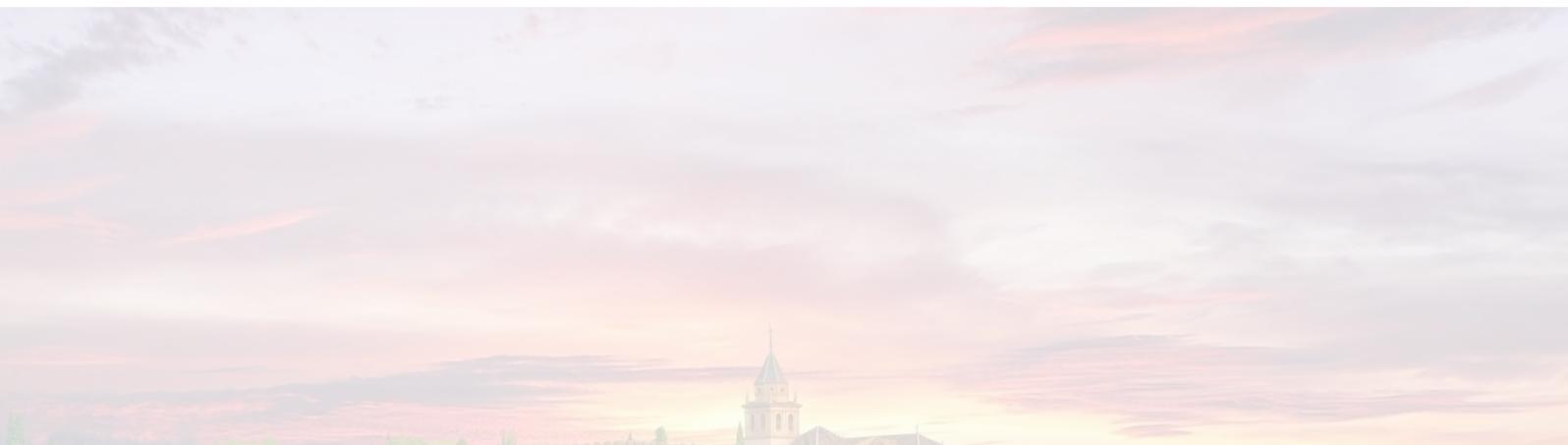


## LAS ECUACIONES NORMALES

$$-2X^t\vec{y} + 2X^tX\vec{\beta} = \vec{0} \Rightarrow X^t(\mathbf{y} - \hat{\mathbf{y}}) = 0$$

Las siguientes expresiones son equivalentes:

►  $X^t\vec{e} = 0$ .



## LAS ECUACIONES NORMALES

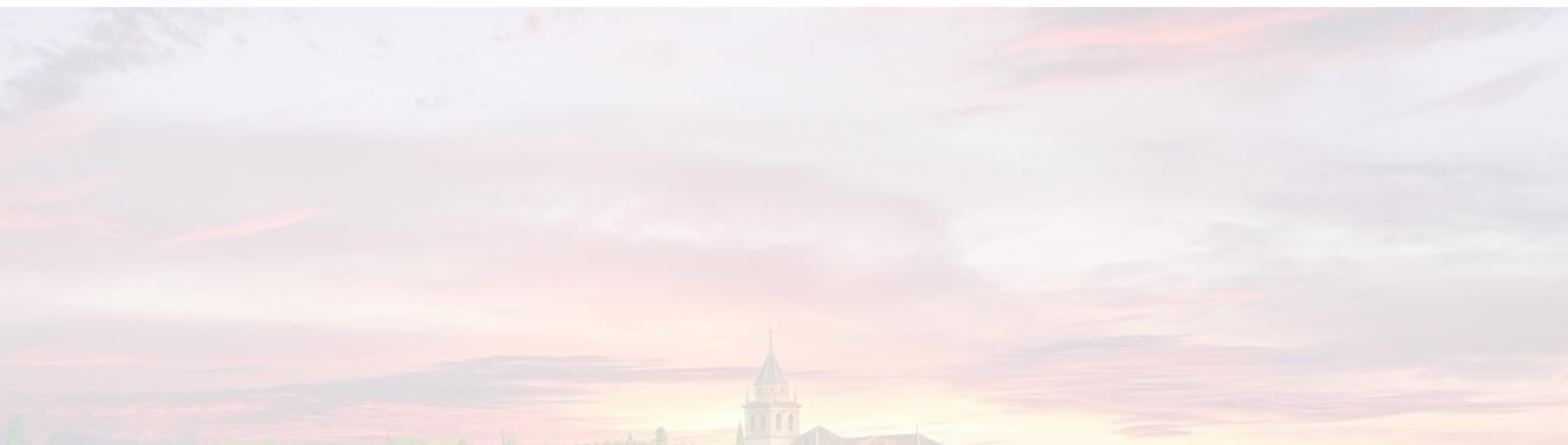
$$-2X^t\vec{y} + 2X^tX\vec{\beta} = \vec{0} \Rightarrow X^t(\mathbf{y} - \hat{\mathbf{y}}) = 0$$

Las siguientes expresiones son equivalentes:

▶  $X^t\vec{e} = 0$ .



$$\begin{pmatrix} 1 & X_{21} & \dots & X_{k1} \\ 1 & X_{22} & \dots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & \dots & X_{kn} \end{pmatrix}^t \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$



## LAS ECUACIONES NORMALES

$$-2X^t\vec{y} + 2X^tX\vec{\beta} = \vec{0} \Rightarrow X^t(\mathbf{y} - \hat{\mathbf{y}}) = 0$$

Las siguientes expresiones son equivalentes:

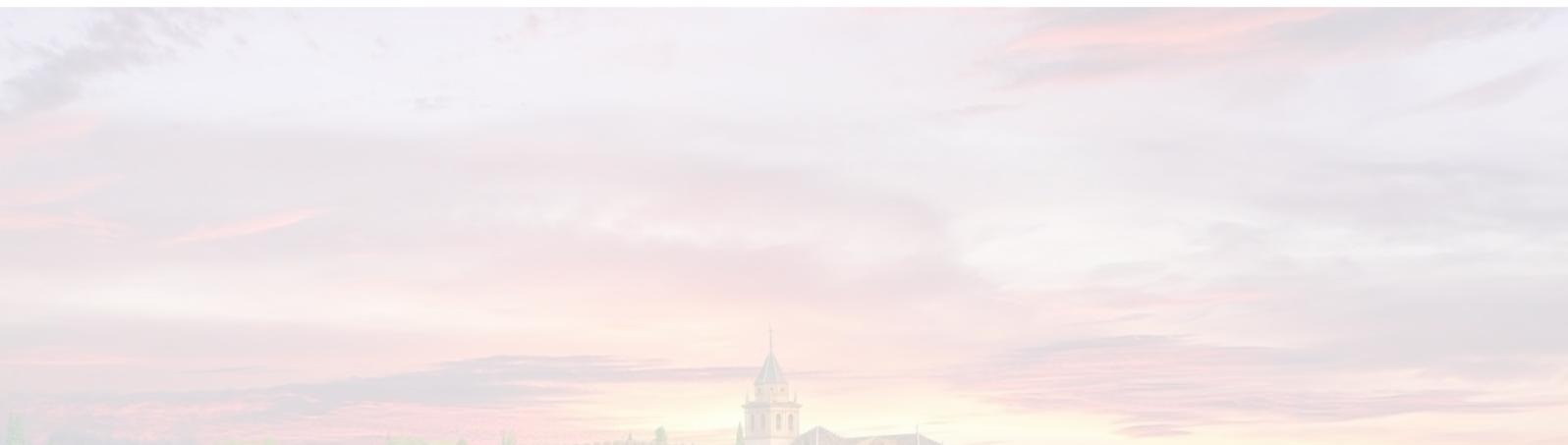
▶  $X^t\vec{e} = 0$ .



$$\begin{pmatrix} 1 & X_{21} & \dots & X_{k1} \\ 1 & X_{22} & \dots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & \dots & X_{kn} \end{pmatrix}^t \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

▶ Ecuaciones normales:

$$\begin{cases} e_1 + \dots + e_n = 0, \\ X_{21}e_1 + \dots + X_{2n}e_n = 0, \\ \vdots = \vdots \\ X_{k1}e_1 + \dots + X_{kn}e_n = 0, \end{cases}$$



## EJEMPLO: DATOS INDITEX

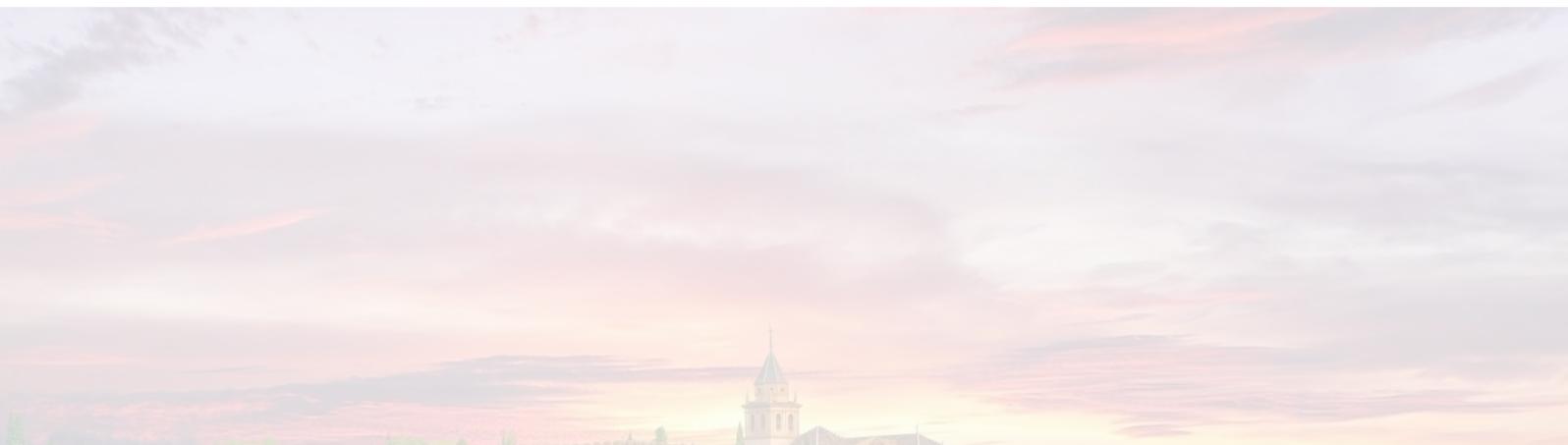
A partir de los datos anteriores, se pueden obtener los estimadores para el modelo 2:

$$\begin{aligned}\hat{\beta} &= \begin{pmatrix} 15 & 1907757 & 88.837 \\ 1907757 & 260.975.868.213 & 12.024.098.223 \\ 88.837 & 12.024.098.223 & 556.138.061 \end{pmatrix}^{-1} \begin{pmatrix} 260679 \\ 36.417.993.981 \\ 1.670.301.943 \end{pmatrix} \\ &= \begin{pmatrix} -3435,533337 \\ 0,258708553 \\ -2,041278432 \end{pmatrix}\end{aligned}$$

De manera que el modelo estimado viene dado por:

$$\hat{y}_t = -3435,533337 + 0.258708553X_{2t} - 2.041278432X_{3t}$$

Se pide obtener la estimación para la variable ingresos, obtener los residuos y comprobar que se verifican las propiedades algebraicas descritas anteriormente.

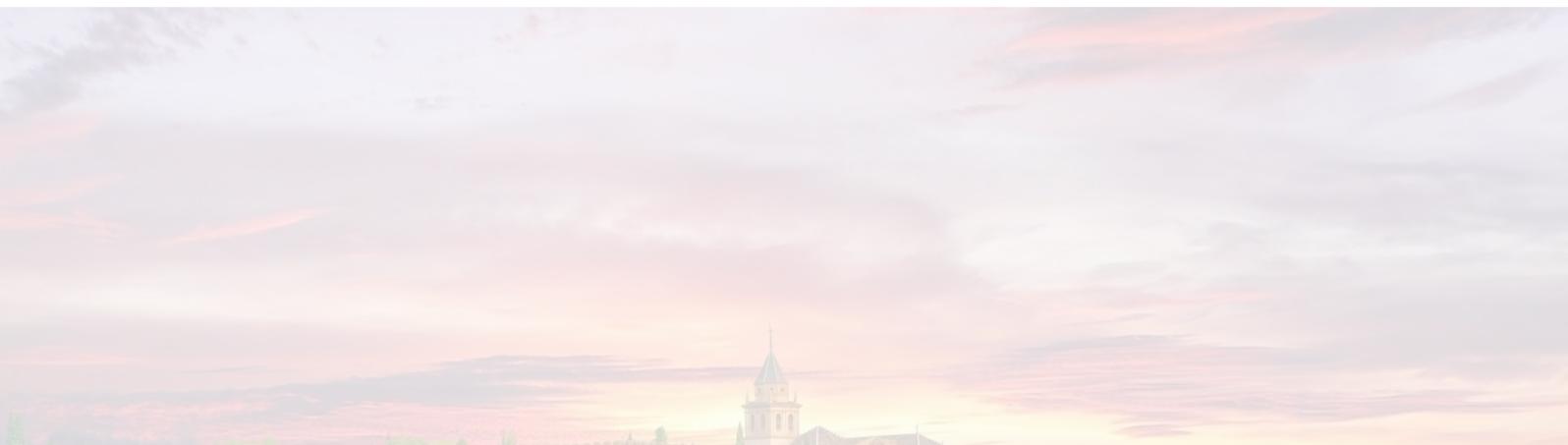


## BONDAD DEL AJUSTE: $R^2$ Y $\bar{R}^2$

[Coeficiente de determinación  $R^2$ ] Porcentaje de variabilidad explicada por el modelo. Por tanto, éste se obtendrá como el cociente entre la varianza explicada por la estimación y la total:

$$R^2 = \frac{\frac{1}{n} \cdot \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\frac{1}{n} \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2} = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} = \frac{SCE}{SCT}$$

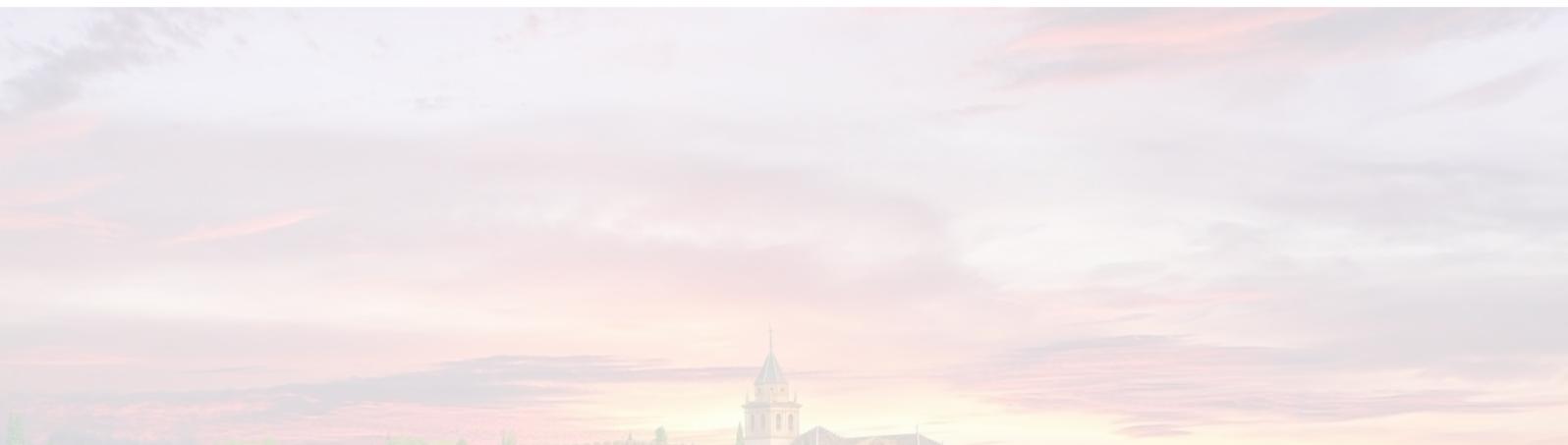
Si el modelo tiene término independiente:  $R^2 = 1 - \frac{SCR}{SCT}$  y los valores del coeficiente estarán comprendidos entre 0 y 1.



## BONDAD DEL AJUSTE: $R^2$ Y $\bar{R}^2$

[Coeficiente de determinación corregido  $\bar{R}^2$ ] Porcentaje de variación de la variable explicada considerando el número de variables incluidas en el modelo, es decir, considerando el valor de  $k$ . Se define como:

$$\bar{R}^2 = 1 - \frac{SCR/(n - k)}{SCT/(n - 1)} = 1 - (1 - R^2) \cdot \frac{n - 1}{n - k}.$$

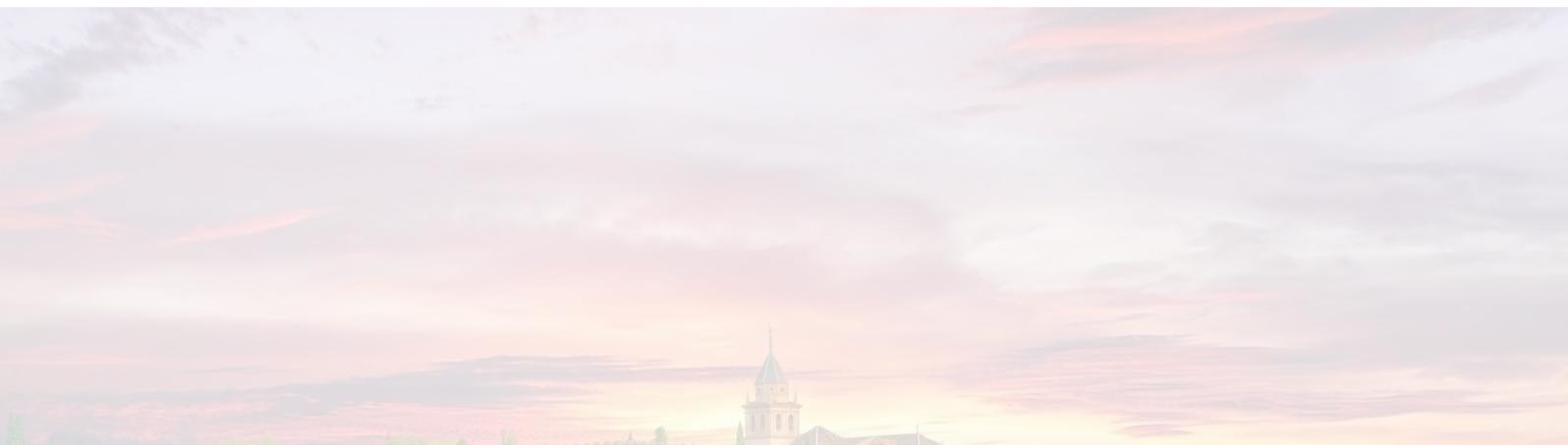


## BONDAD DEL AJUSTE: $R^2$ Y $\bar{R}^2$

Adviértase que obtener un  $R^2$  o  $\bar{R}^2$  cercano a 1 no indica que los resultados sean fiables, ya que, por ejemplo, puede ser que no se cumpla alguna de las hipótesis básicas y los resultados no ser válidos.

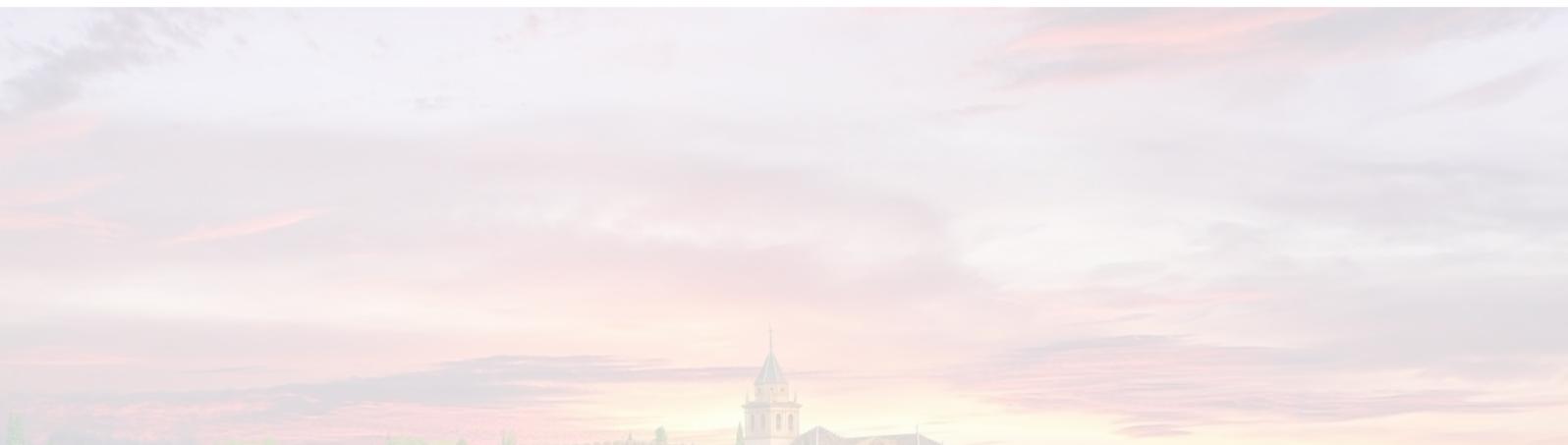
Es necesario señalar que ambos coeficientes no son capaces de detectar si alguna de las variables incluidas en el modelo son o no estadísticamente significativas.

Por tanto, estos indicadores han de ser considerados como una herramienta más a tener en cuenta dentro del análisis.



## CRITERIOS DE AKAIKE Y SCHWARZ

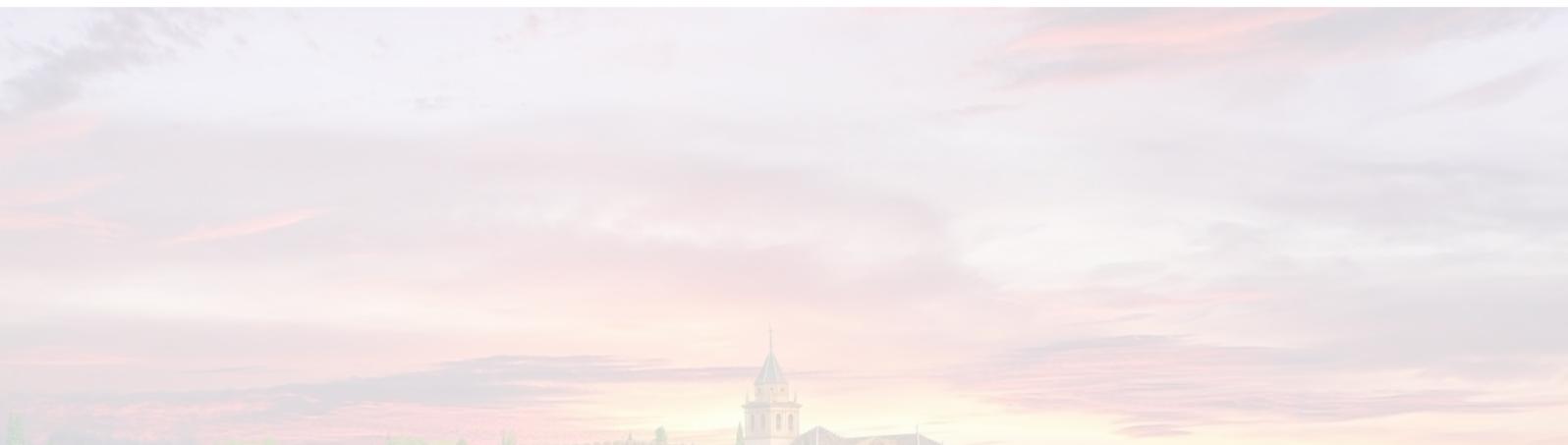
Para comparar distintos modelos podríamos utilizar el coeficiente de determinación. Sin embargo, para poder llevar cabo dicha comparación será necesario que tengan la misma variable explicada.



## CRITERIOS DE AKAIKE Y SCHWARZ

Para comparar distintos modelos podríamos utilizar el coeficiente de determinación. Sin embargo, para poder llevar cabo dicha comparación será necesario que tengan la misma variable explicada.

Con el fin de buscar una solución al problema utilizaremos los criterios de selección de modelos de Akaike (AIC) y el bayesiano de Schwarz (BIC). Estos criterios se obtienen a partir de la suma de cuadrados de los residuos y de un factor que penaliza la inclusión de parámetros.



## CRITERIOS DE AKAIKE Y SCHWARZ

Para comparar distintos modelos podríamos utilizar el coeficiente de determinación. Sin embargo, para poder llevar cabo dicha comparación será necesario que tengan la misma variable explicada.

Con el fin de buscar una solución al problema utilizaremos los criterios de selección de modelos de Akaike (AIC) y el bayesiano de Schwarz (BIC). Estos criterios se obtienen a partir de la suma de cuadrados de los residuos y de un factor que penaliza la inclusión de parámetros.

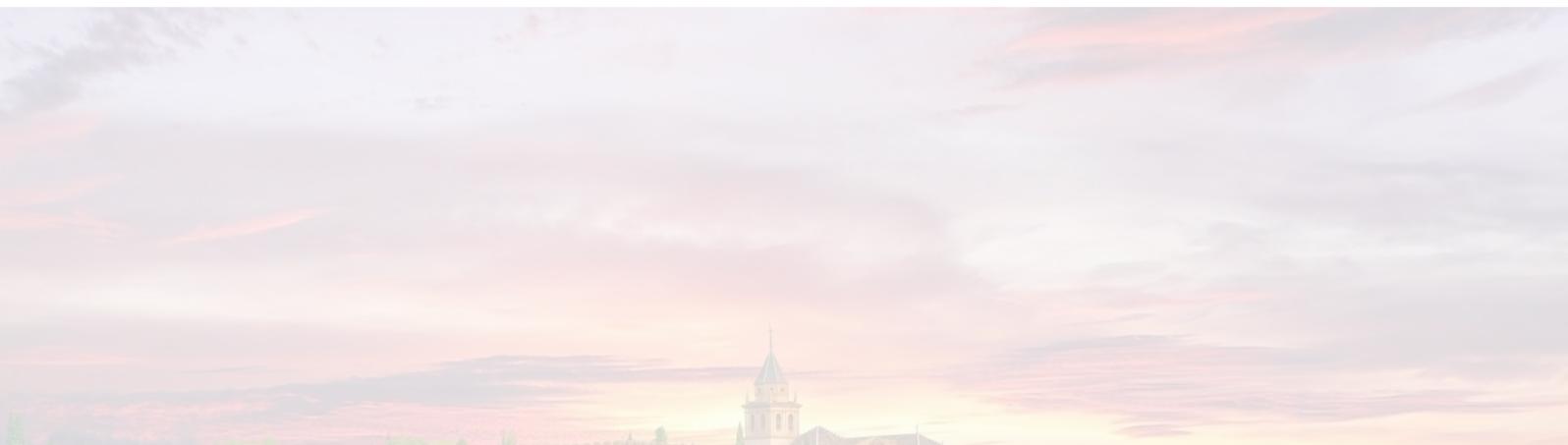
El criterio de información de Akaike:

$$AIC = \ln \left( \frac{SCR}{n} \right) + \frac{2k}{n},$$

El criterio de información de Schwarz:

$$BIC = \ln \left( \frac{SCR}{n} \right) + \frac{k}{n} \cdot \ln(n).$$

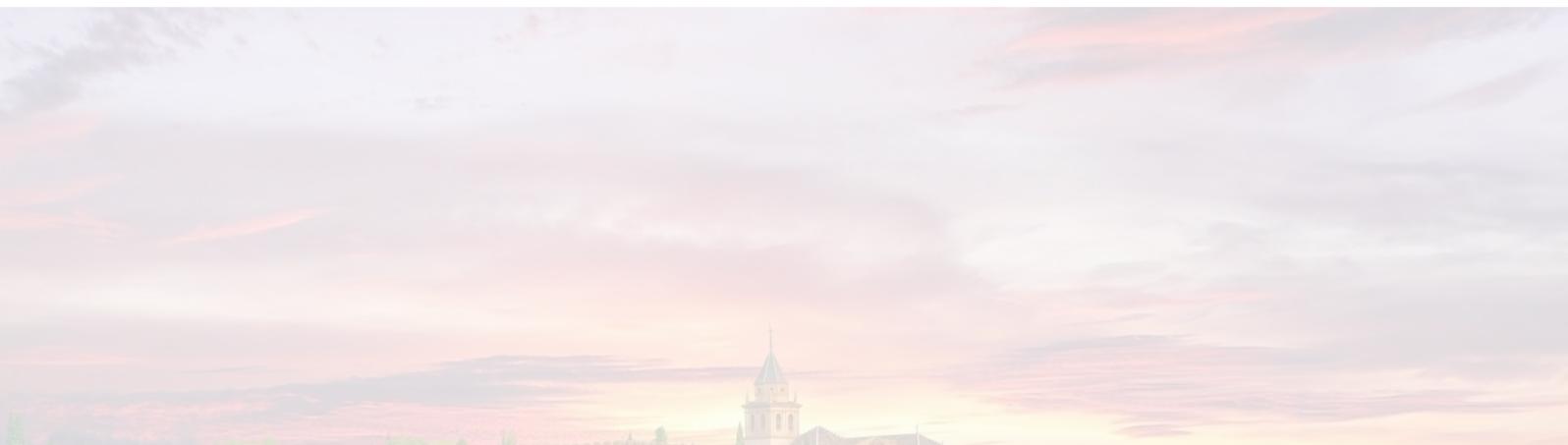
Utilizando estos criterios se escogería aquel modelo con un menor valor de AIC o BIC.



## EJEMPLO: DATOS INDITEX

Se pide:

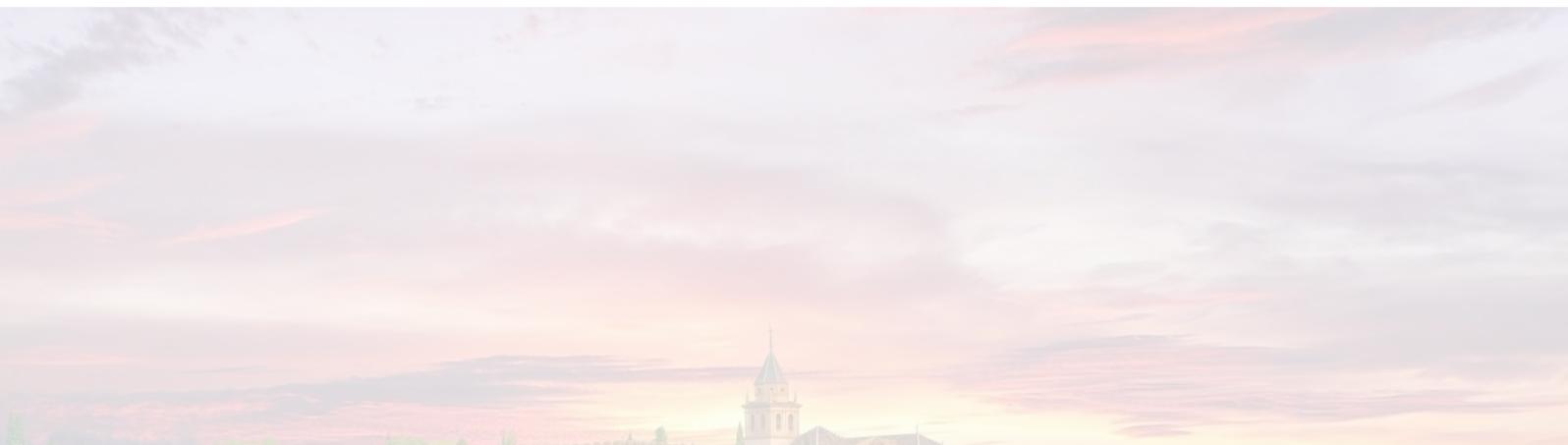
1. Obtener el coeficiente de determinación para el modelo 1 y para el modelo 2.
2. Obtener el coeficiente de determinación corregido para el modelo 1 y para el modelo 2.
3. Obtener los criterios de información para el modelo 1 y para el modelo 2.
4. Concluir sobre que modelo es mejor en cuanto a la bondad del ajuste.



## EJERCICIO PROPUESTO

Usando los siguientes datos, consumo nacional ( $C_t$ ) y renta nacional ( $R_t$ ) en España para el periodo 1995-2005 a precios corrientes ( $10^9$  euros), obtenga las estimaciones por MCO, así como las sumas de cuadrados total, explicada y residual, y el coeficiente de determinación, para el modelo de regresión  $C_t = \beta_1 + \beta_2 R_t + u_t$ .

Año	$C_t$	$R_t$
1995	349	388
1996	368	408
1997	388	433
1998	414	465
1999	444	498
2000	484	538
2001	518	574
2002	550	614
2003	586	656
2004	635	699
2005	686	748



## EJERCICIO PROPUESTO

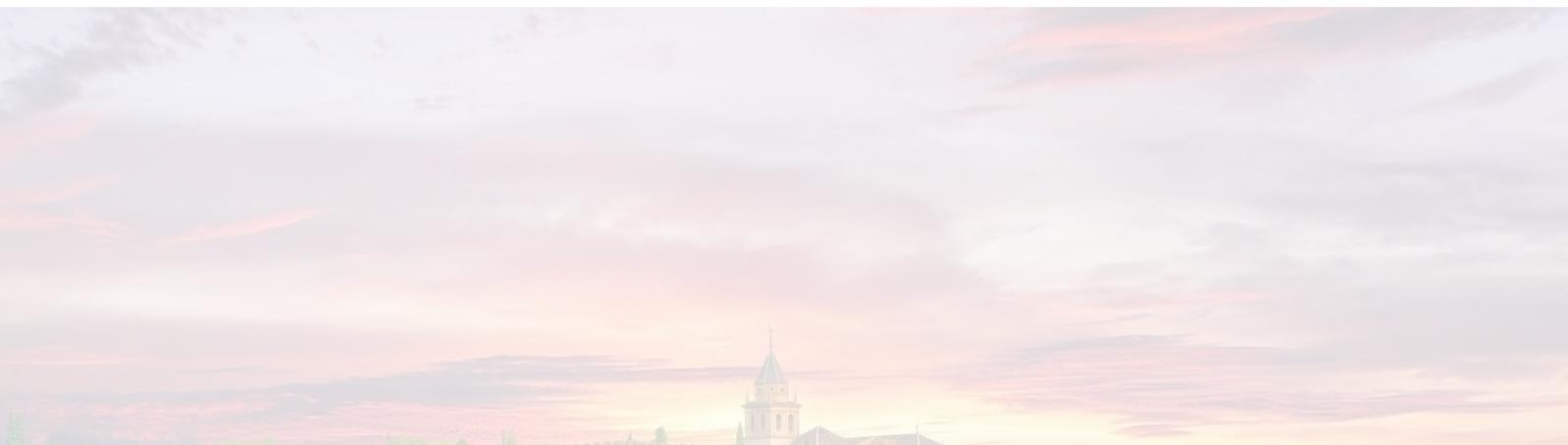
Para el modelo  $Y_t = \beta_1 + \beta_2 v_t + \beta_3 w_t + u_t$  se tienen los siguientes datos:

$$n = 12, \quad SCT = 104'9167,$$

$$(X^t X)^{-1} = \begin{pmatrix} 0'6477 & -0'041 & -0'0639 \\ -0'041 & 0'0071 & -0'0011 \\ -0'0639 & -0'0011 & 0'0152 \end{pmatrix}, \quad X^t Y = \begin{pmatrix} 91 \\ 699 \\ 448 \end{pmatrix}.$$

Se pide:

- Ajustar el modelo por el método de MCO y calcular el coeficiente de determinación.



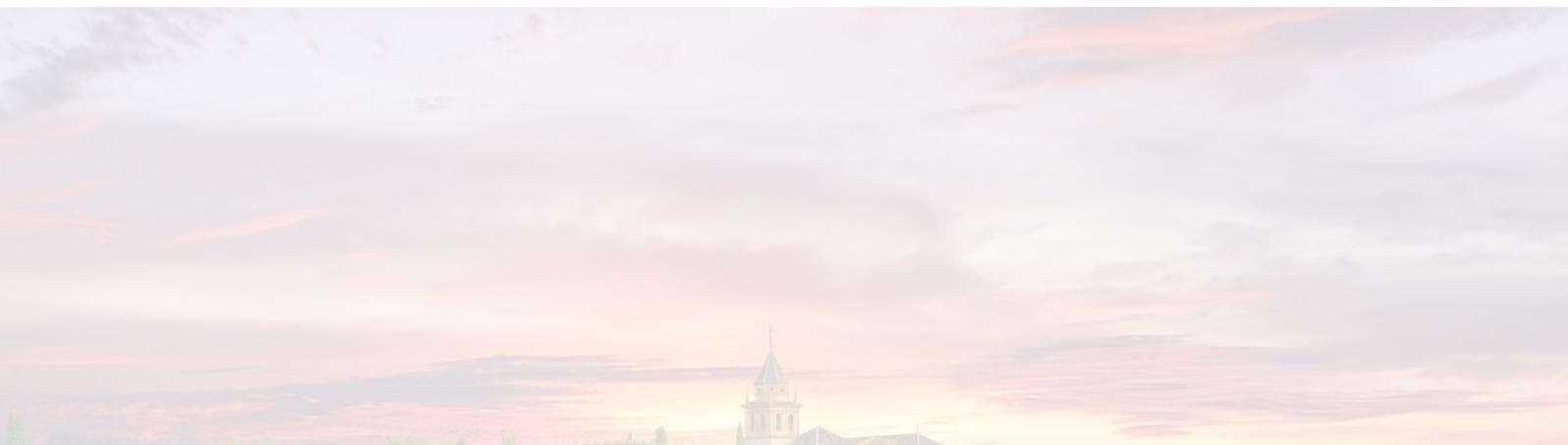
## EJERCICIO PROPUESTO

En un estudio de los determinantes de la inversión se usaron 20 datos anuales, correspondientes a las siguientes variables: inversión anual en billones de pesetas ( $Y$ ), tipo de interés en porcentaje ( $X_1$ ) y variación anual de PIB en billones de pesetas ( $X_2$ ). Se dispone de la siguiente información:

$$\begin{array}{lll} \sum X_{1t} = 100 & \sum X_{2t} = 24 & \sum Y_t = 5 \\ \sum X_{1t}Y_t = -255 & \sum X_{2t}Y_t = 146 & \sum X_{1t}X_{2t} = 100 \\ \sum X_{1t}^2 = 680 & \sum X_{2t}^2 = 48'8 & \sum (Y_t - \bar{Y})^2 = 1200 \end{array}$$

Se pide:

- a) Obtenga las estimaciones por MCO del modelo  $Y_t = \alpha + \beta X_{1t} + \delta X_{2t} + u_2$ .



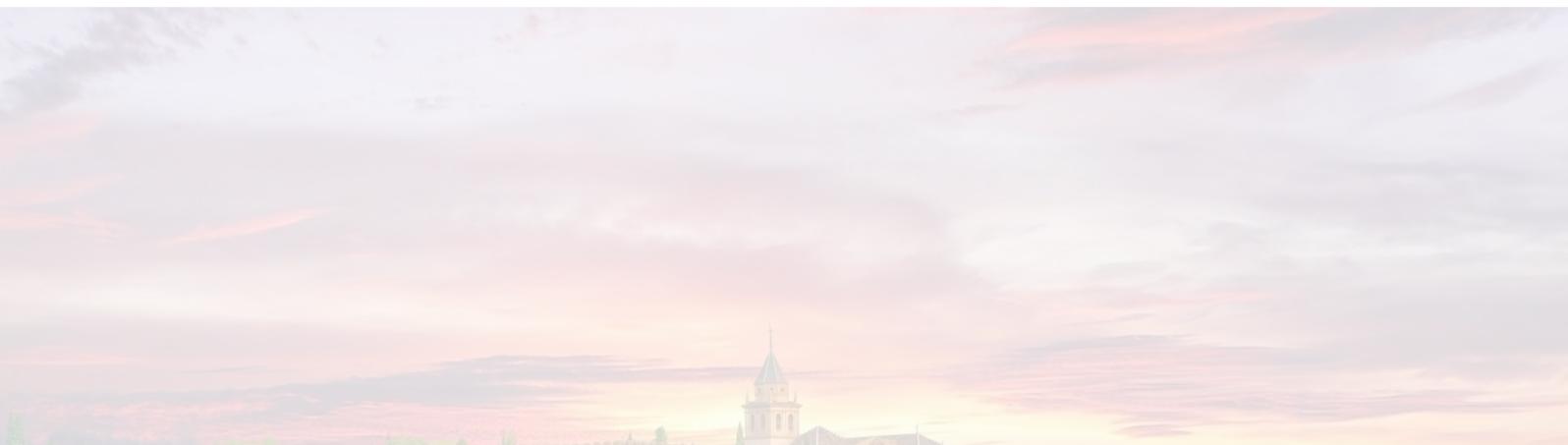
## EJERCICIO PROPUESTO

Para estimar el modelo  $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$  se ha obtenido una muestra de la cual ha resultado:

$$X^t X = \begin{pmatrix} 14 & 7 & 14 \\ 7 & 4'5 & 7 \\ 14 & 7 & 15 \end{pmatrix}, \quad X^t Y = \begin{pmatrix} 10 \\ 6 \\ 12 \end{pmatrix}, \quad Y^t Y = 14.$$

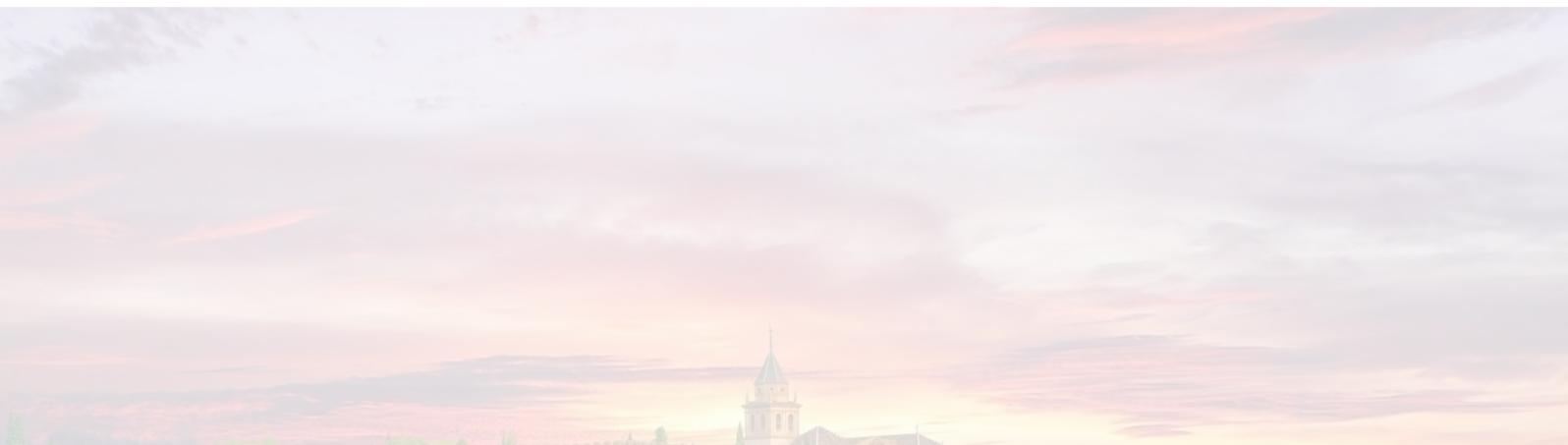
Se pide:

a) Estimar los coeficientes del modelo por MCO.

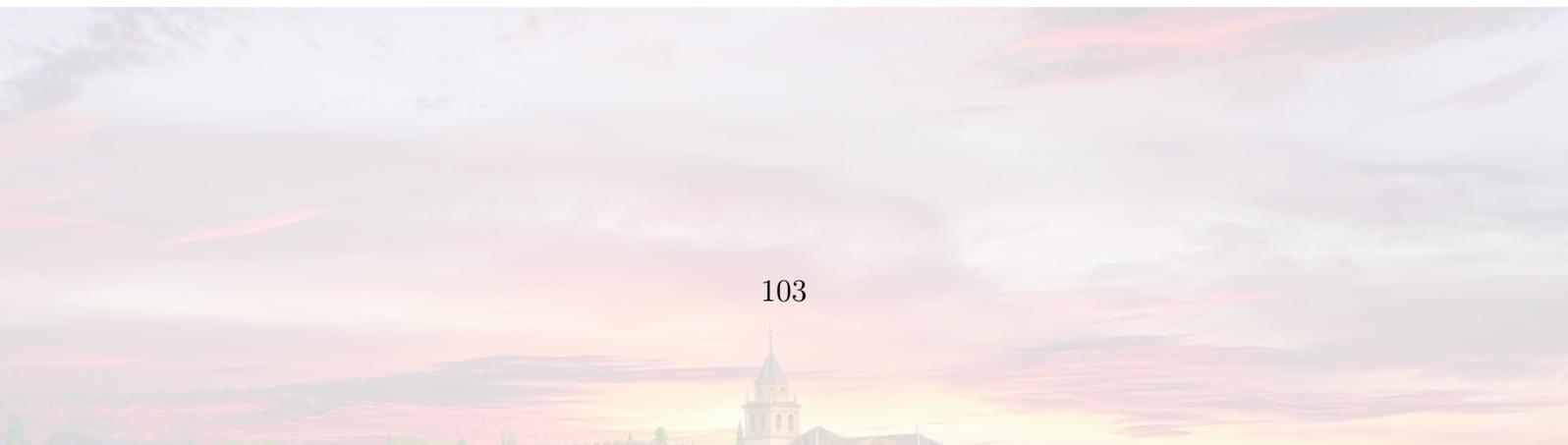


## BIBLIOGRAFÍA

- ▶ García, C.B., Pérez-sánchez, J.M. y Salmerón-Gómez, R. (2017). *Econometría Básica para la economía y la empresa*. Ed. Fleming. Tema 2.
- ▶ Sánchez, C., López, M.d.M y García, T. *Econometría*. Ed. Fleming. Capítulos 2 y 3
- ▶ Gujarati, D. (1997). *Econometría*. Ed. McGraw Hill. Capítulo 2.
- ▶ Matilla García, m., Pérez Pascual, Pedro y Sanz Carnero, B. (2013) *Econometría y predicción*. Mc Graw Hill. Capítulo 2.



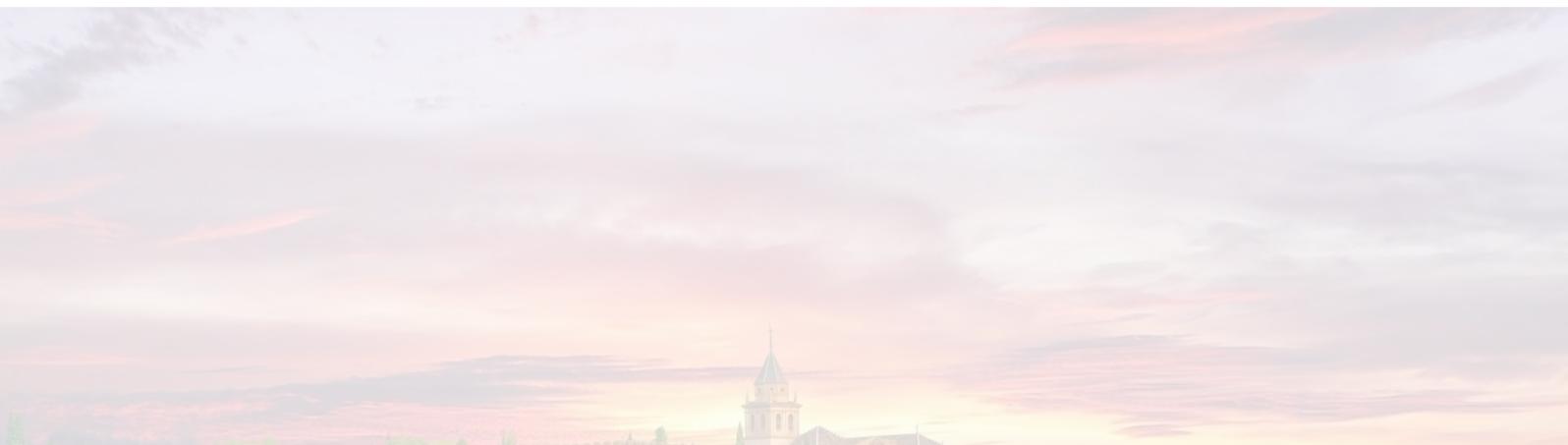
### 3 Tema 3



# ECONOMETRÍA

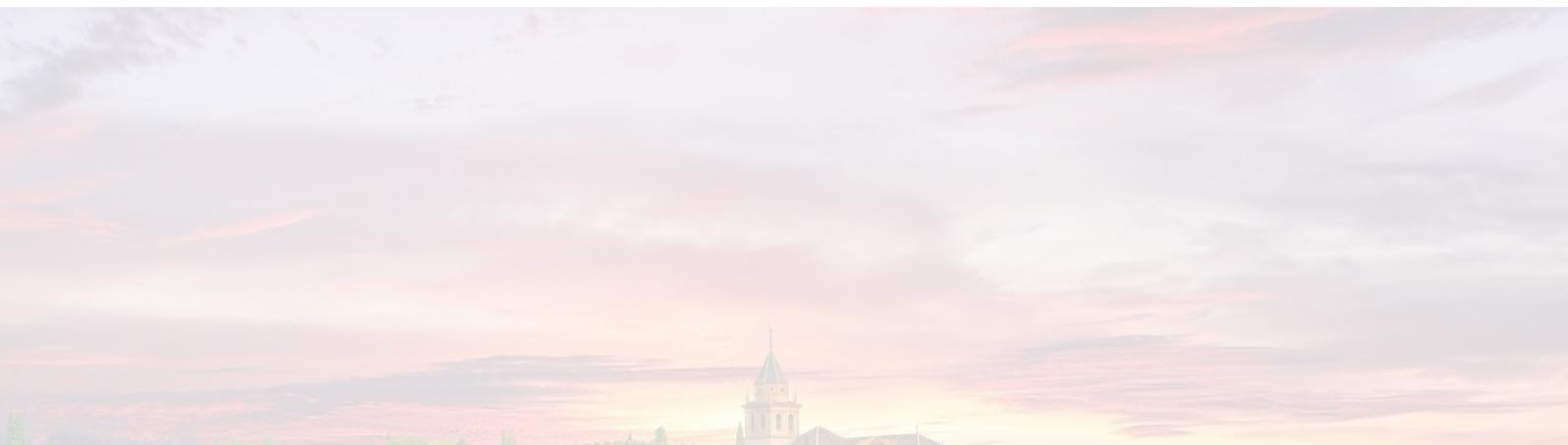
## TEMA 3: EL MODELO LINEAL (II)

2024-2025



# EL MODELO LINEAL (II)

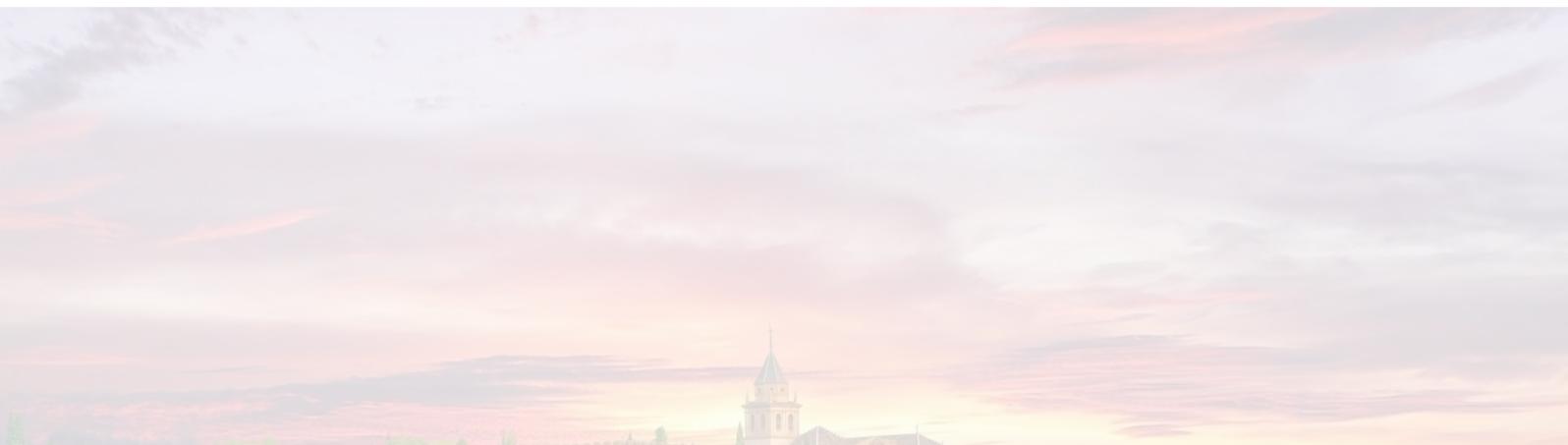
<b>1</b>	<b>Propiedades estadísticas del estimador</b>	<b>2</b>
1.1	Valor Esperado y Varianza de los EMCO	2
1.2	Teorema de Gauss-Markov	4
1.3	Estimación de $\sigma_u^2$	9
1.4	Estimación de la varianza de los EMCO	11
<b>2</b>	<b>El supuesto de normalidad y la inferencia sobre los parámetros del modelo</b>	<b>12</b>
2.1	Intervalos de confianza para los parámetros del modelo	13
2.2	Contraste de hipótesis acerca de los parámetros del modelo. Test General	14
2.2.1	Casos particulares	17
<b>3</b>	<b>Predicción</b>	<b>26</b>



## 1.1. VALOR ESPERADO Y VARIANZA DE LOS EMCO

El estimador EMCO de  $\vec{\beta}$  del modelo  $\vec{y} = X\vec{\beta} + \vec{u}$  es:

$$\vec{\hat{\beta}} = (X^t X)^{-1} (X^t \vec{y})$$



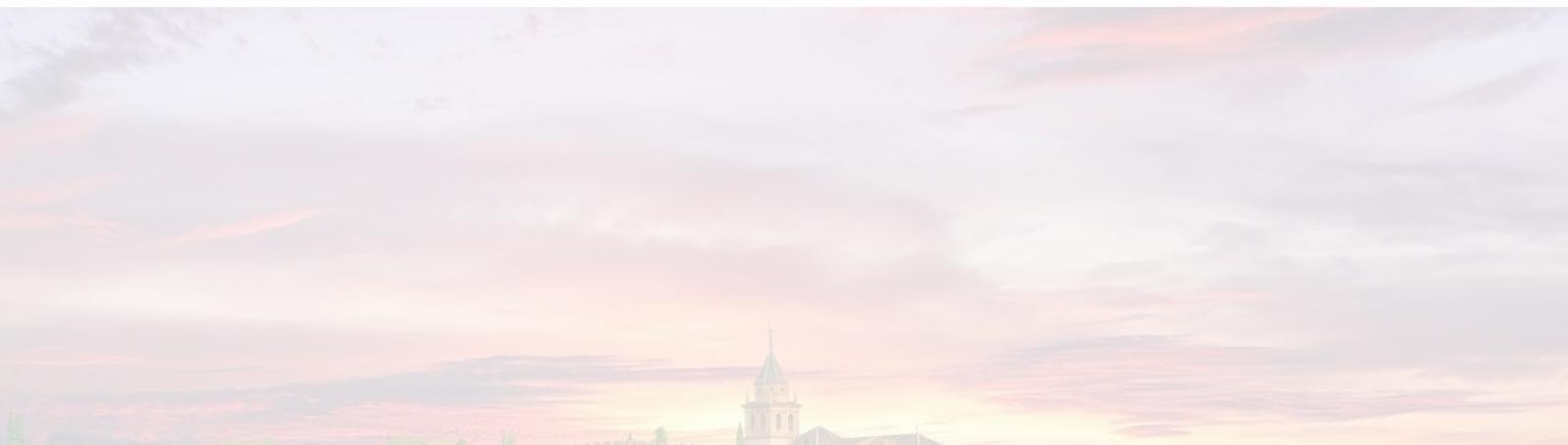
## 1.1. VALOR ESPERADO Y VARIANZA DE LOS EMCO

El estimador EMCO de  $\vec{\beta}$  del modelo  $\vec{y} = X\vec{\beta} + \vec{u}$  es:

$$\vec{\hat{\beta}} = (X^t X)^{-1} (X^t \vec{y})$$

Sustituyendo  $\vec{y}$  por  $\vec{y} = X\vec{\beta} + \vec{u}$ :

$$\begin{aligned}\vec{\hat{\beta}} &= (X^t X)^{-1} (X^t \vec{y}) = (X^t X)^{-1} X^t (X\vec{\beta} + \vec{u}) \\ &= \vec{\beta} + (X^t X)^{-1} X^t \vec{u}.\end{aligned}$$



## 1.1. VALOR ESPERADO Y VARIANZA DE LOS EMCO

El estimador EMCO de  $\vec{\beta}$  del modelo  $\vec{y} = X\vec{\beta} + \vec{u}$  es:

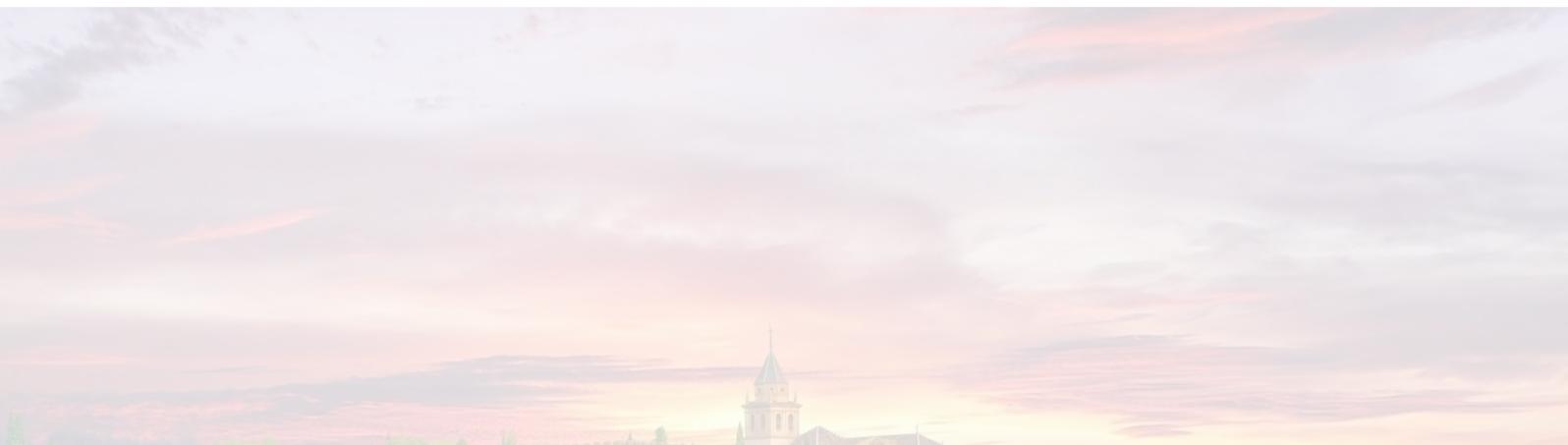
$$\vec{\hat{\beta}} = (X^t X)^{-1} (X^t \vec{y})$$

Sustituyendo  $\vec{y}$  por  $\vec{y} = X\vec{\beta} + \vec{u}$ :

$$\begin{aligned}\vec{\hat{\beta}} &= (X^t X)^{-1} (X^t \vec{y}) = (X^t X)^{-1} X^t (X\vec{\beta} + \vec{u}) \\ &= \vec{\beta} + (X^t X)^{-1} X^t \vec{u}.\end{aligned}$$

Por tanto,

$$\vec{\hat{\beta}} = \vec{\beta} + (X^t X)^{-1} X^t \vec{u}.$$



## 1.1. VALOR ESPERADO Y VARIANZA DE LOS EMCO

El estimador EMCO de  $\vec{\beta}$  del modelo  $\vec{y} = X\vec{\beta} + \vec{u}$  es:

$$\vec{\hat{\beta}} = (X^t X)^{-1} (X^t \vec{y})$$

Sustituyendo  $\vec{y}$  por  $\vec{y} = X\vec{\beta} + \vec{u}$ :

$$\begin{aligned}\vec{\hat{\beta}} &= (X^t X)^{-1} (X^t \vec{y}) = (X^t X)^{-1} X^t (X\vec{\beta} + \vec{u}) \\ &= \vec{\beta} + (X^t X)^{-1} X^t \vec{u}.\end{aligned}$$

Por tanto,

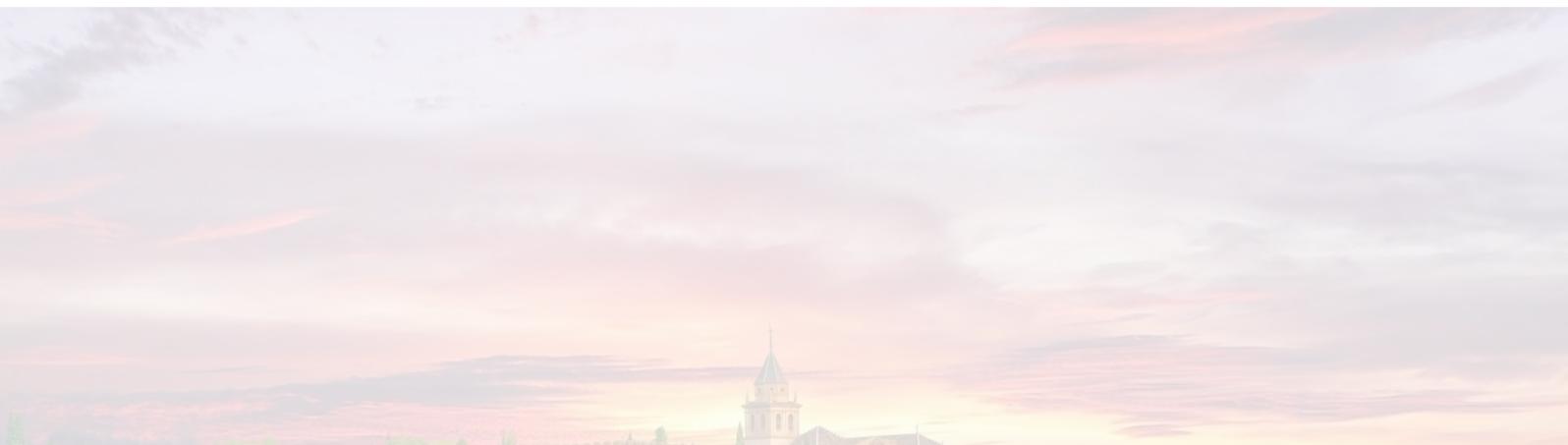
$$\vec{\hat{\beta}} = \vec{\beta} + (X^t X)^{-1} X^t \vec{u}.$$

$$E \left[ \vec{\hat{\beta}} \right] = E \left[ \vec{\beta} + (X^t X)^{-1} (X^t \vec{u}) \right] = \vec{\beta} + (X^t X)^{-1} X^t E[\vec{u}] = \vec{\beta}$$

ya que  $E[\vec{u}] = \vec{0} \Rightarrow \vec{\hat{\beta}}$  es un estimador **insesgado** de  $\vec{\beta}$ .

## 1.1. VALOR ESPERADO Y VARIANZA DE LOS EMCO

$$\begin{aligned}\text{var} \left( \vec{\hat{\beta}} \right) &= E \left[ \left( \vec{\hat{\beta}} - E \left[ \vec{\hat{\beta}} \right] \right) \cdot \left( \vec{\hat{\beta}} - E \left[ \vec{\hat{\beta}} \right] \right)^t \right] = \\ &= E \left[ \left( \vec{\hat{\beta}} - \vec{\beta} \right) \cdot \left( \vec{\hat{\beta}} - \vec{\beta} \right)^t \right] = \\ &= E \left[ \left( X^t X \right)^{-1} \left( X^t \vec{u} \right) \cdot \left( \vec{u}^t X \right) \left( X^t X \right)^{-1} \right] = \\ &= \left( X^t X \right)^{-1} X^t \cdot E \left[ \vec{u} \cdot \vec{u}^t \right] \cdot X \left( X^t X \right)^{-1} = \\ &= \sigma^2 \cdot \left( X^t X \right)^{-1} X^t X \left( X^t X \right)^{-1} = \sigma^2 \cdot \left( X^t X \right)^{-1},\end{aligned}$$

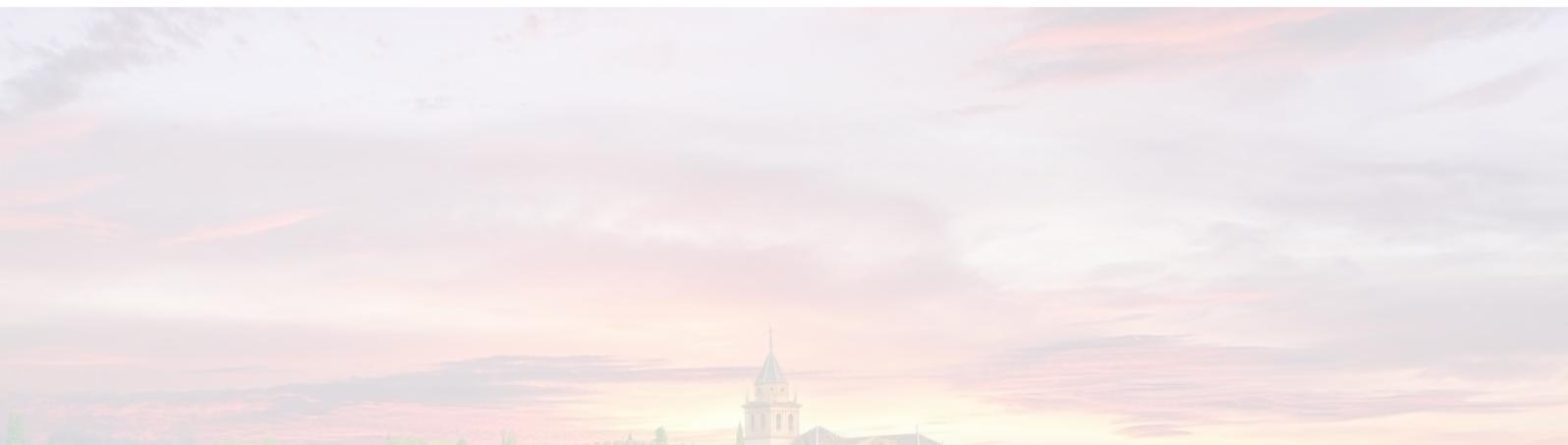


## 1.2. TEOREMA DE GAUSS-MARKOV

[Gauss–Markov] Los estimadores mínimos cuadrados ordinarios de  $\vec{\beta}$  son BLUE (ELIO en español):

- ▶ óptimos, (Best)
- ▶ lineales, (Linear) y
- ▶ insesgados (Unbiased)

Según el Teorema de Gauss-Markov podemos decir que entre todos los estimadores lineales e insesgados de  $\vec{\beta}$ , el EMCO es el que presenta menor matriz de varianzas-covarianzas.

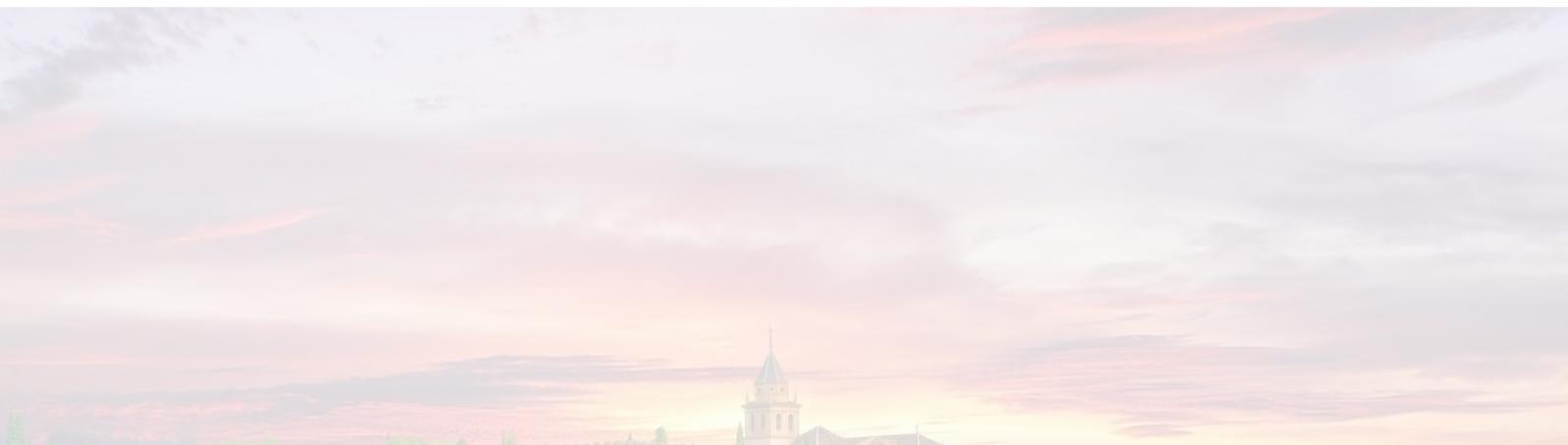


## 1.2. TEOREMA DE GAUSS-MARKOV: DEMO (I)

- ▶  $\vec{\beta}$  es lineal con respecto al vector  $\vec{y}$ .

Definiendo la matriz de dimensión  $k \times n$  como  $W = (X^t X)^{-1} X^t$  es fácil comprobar que  $\vec{\beta}$  es un estimador lineal con respecto a las observaciones  $\vec{y}$ , ya que

$$\vec{\beta} = (X^t X)^{-1} (X^t \vec{y}) = W \vec{y}$$



## 1.2. TEOREMA DE GAUSS-MARKOV: DEMO (I)

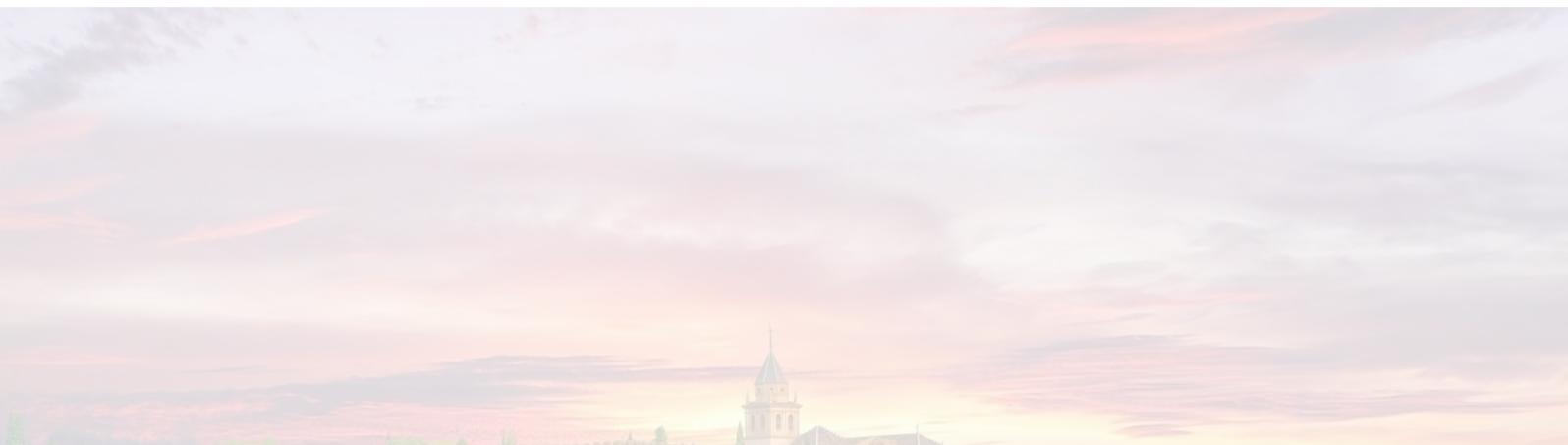
- ▶  $\vec{\beta}$  es lineal con respecto al vector  $\vec{y}$ .

Definiendo la matriz de dimensión  $k \times n$  como  $W = (X^t X)^{-1} X^t$  es fácil comprobar que  $\vec{\beta}$  es un estimador lineal con respecto a las observaciones  $\vec{y}$ , ya que

$$\vec{\beta} = (X^t X)^{-1} (X^t \vec{y}) = W \vec{y}$$

- ▶  $\vec{\beta}$  es un estimador insesgado de  $\vec{\beta}$ .

Si  $\vec{\beta}$  es un estimador insesgado de  $\vec{\beta}$  entonces es porque se verifica que  $E \left[ \vec{\beta} \right] = \vec{\beta}$  (ya lo sabemos).



## 1.2. TEOREMA DE GAUSS–MARKOV: DEMO (II)

- $\vec{\beta}$  es un estimador óptimo de  $\vec{\beta}$ .

Consideremos otro estimador lineal e insesgado, alternativo a  $\vec{\beta}$ :  $\vec{\beta}^*$ .

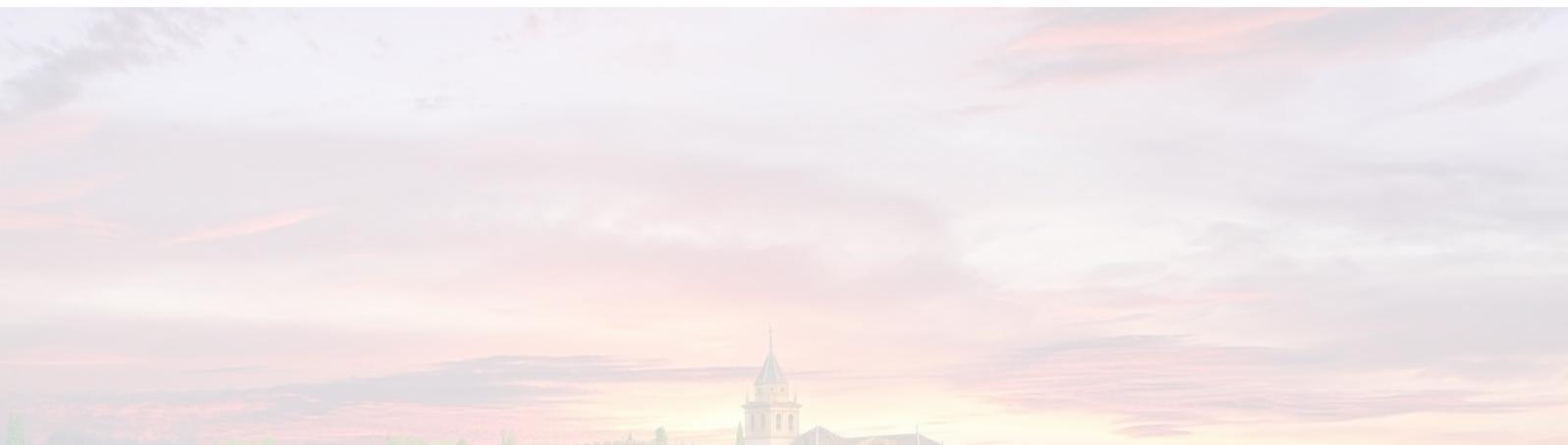
- Como  $\vec{\beta}^*$  es lineal en  $\vec{y}$  entonces se verifica que existe  $C_{k \times n}$  tal que  $\vec{\beta}^* = C\vec{y}$ .
- Al verificarse que  $\vec{\beta}^*$  es un estimador insesgado de  $\vec{\beta}$  entonces  $E[\vec{\beta}^*] = \vec{\beta}$ . Luego:

$$\begin{aligned} E[\vec{\beta}^*] &= E[C\vec{y}] = E[C(X\vec{\beta} + \vec{u})] = E[CX\vec{\beta}] + \\ &+ E[C\vec{u}] = CE[X\vec{\beta}] + CE[\vec{u}] = CX\vec{\beta} \end{aligned}$$

Así que para que  $\vec{\beta}^*$  sea insesgado se debe verificar:  $CX = I_k$ .

## 1.2. TEOREMA DE GAUSS–MARKOV: DEMO (III)

$$\begin{aligned}\text{var} \left( \vec{\beta}^* \right) &= E \left[ \left( \vec{\beta}^* - E \left[ \vec{\beta}^* \right] \right) \cdot \left( \vec{\beta}^* - E \left[ \vec{\beta}^* \right] \right)^t \right] = \\ &= E \left[ \left( \vec{\beta}^* - \vec{\beta} \right) \cdot \left( \vec{\beta}^* - \vec{\beta} \right)^t \right] = E \left[ C \vec{u} \cdot \vec{u}^t C^t \right] = \\ &= CE \left[ \vec{u} \cdot \vec{u}^t \right] C^t = C \sigma^2 I_n C^t = \sigma^2 C C^t\end{aligned}$$

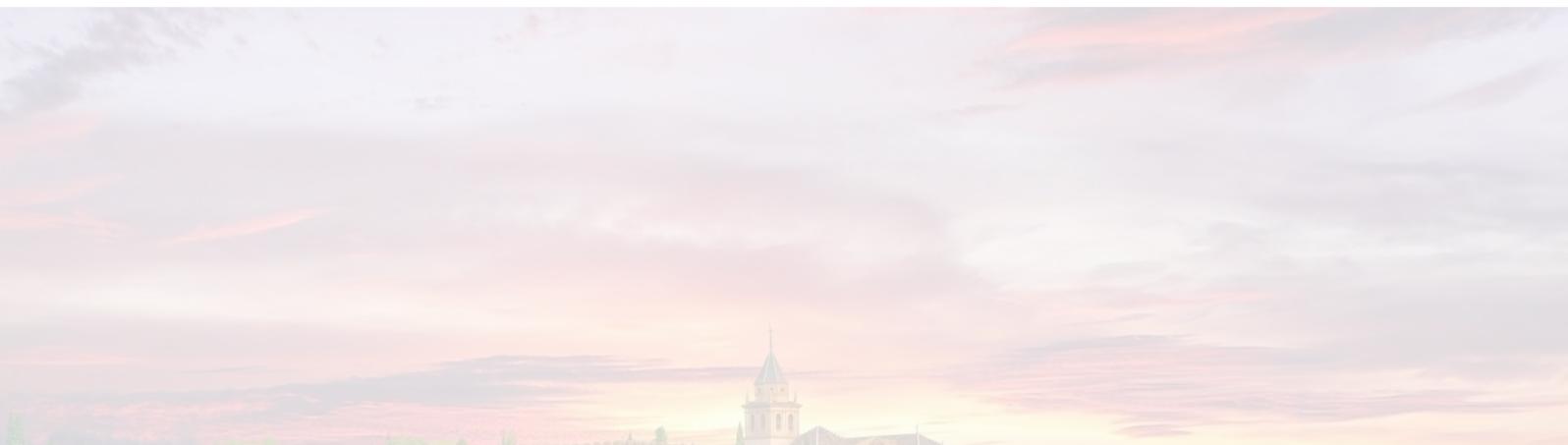


## 1.2. TEOREMA DE GAUSS–MARKOV: DEMO (III)

$$\begin{aligned}\text{var} \left( \vec{\beta}^* \right) &= E \left[ \left( \vec{\beta}^* - E \left[ \vec{\beta}^* \right] \right) \cdot \left( \vec{\beta}^* - E \left[ \vec{\beta}^* \right] \right)^t \right] = \\ &= E \left[ \left( \vec{\beta}^* - \vec{\beta} \right) \cdot \left( \vec{\beta}^* - \vec{\beta} \right)^t \right] = E \left[ C \vec{u} \cdot \vec{u}^t C^t \right] = \\ &= CE \left[ \vec{u} \cdot \vec{u}^t \right] C^t = C \sigma^2 I_n C^t = \sigma^2 C C^t\end{aligned}$$

Tenemos que  $\vec{\beta} = W \vec{y}$  y que  $\vec{\beta}^* = C \vec{y}$ .  
Si denotamos por  $D = C - W$ :

$$DX = (C - W) X = CX - WX = CX - (X^t X)^{-1} X^t X = I_k - I_k = 0_k$$



## 1.2. TEOREMA DE GAUSS–MARKOV: DEMO (III)

$$\begin{aligned}\text{var} \left( \vec{\beta}^* \right) &= E \left[ \left( \vec{\beta}^* - E \left[ \vec{\beta}^* \right] \right) \cdot \left( \vec{\beta}^* - E \left[ \vec{\beta}^* \right] \right)^t \right] = \\ &= E \left[ \left( \vec{\beta}^* - \vec{\beta} \right) \cdot \left( \vec{\beta}^* - \vec{\beta} \right)^t \right] = E \left[ C \vec{u} \cdot \vec{u}^t C^t \right] = \\ &= CE \left[ \vec{u} \cdot \vec{u}^t \right] C^t = C \sigma^2 I_n C^t = \sigma^2 C C^t\end{aligned}$$

Tenemos que  $\vec{\beta} = W \vec{y}$  y que  $\vec{\beta}^* = C \vec{y}$ .  
Si denotamos por  $D = C - W$ :

$$DX = (C - W) X = CX - WX = CX - (X^t X)^{-1} X^t X = I_k - I_k = 0_k$$

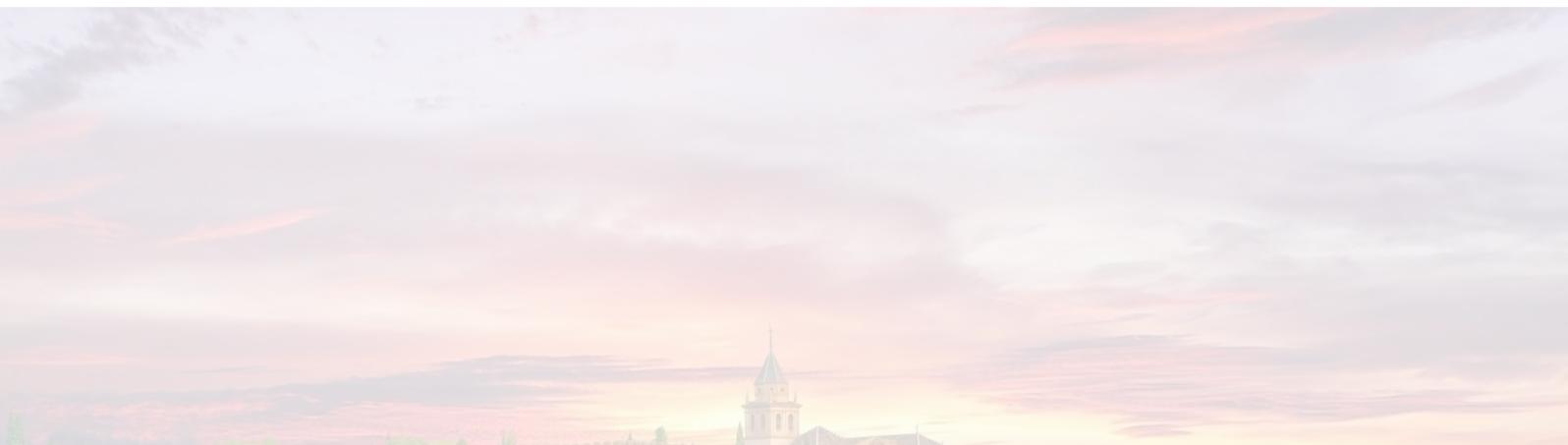
$$\Rightarrow \text{var} \left( \vec{\beta}^* \right) = \sigma^2 C C^t = \sigma^2 (W + D) (W + D)^t = \sigma^2 (W W^t + W D^t + D W^t + D D^t)$$

## 1.2. TEOREMA DE GAUSS–MARKOV: DEMO (IV)

$$WW^t = \left( (X^t X)^{-1} X^t \right) \left( (X^t X)^{-1} X^t \right)^t = (X^t X)^{-1},$$

$$WD^t = \left( (X^t X)^{-1} X^t \right) D^t = 0_{k \times k},$$

$$DW^t = D \left( (X^t X)^{-1} X^t \right)^t = 0_{k \times k},$$



## 1.2. TEOREMA DE GAUSS–MARKOV: DEMO (IV)

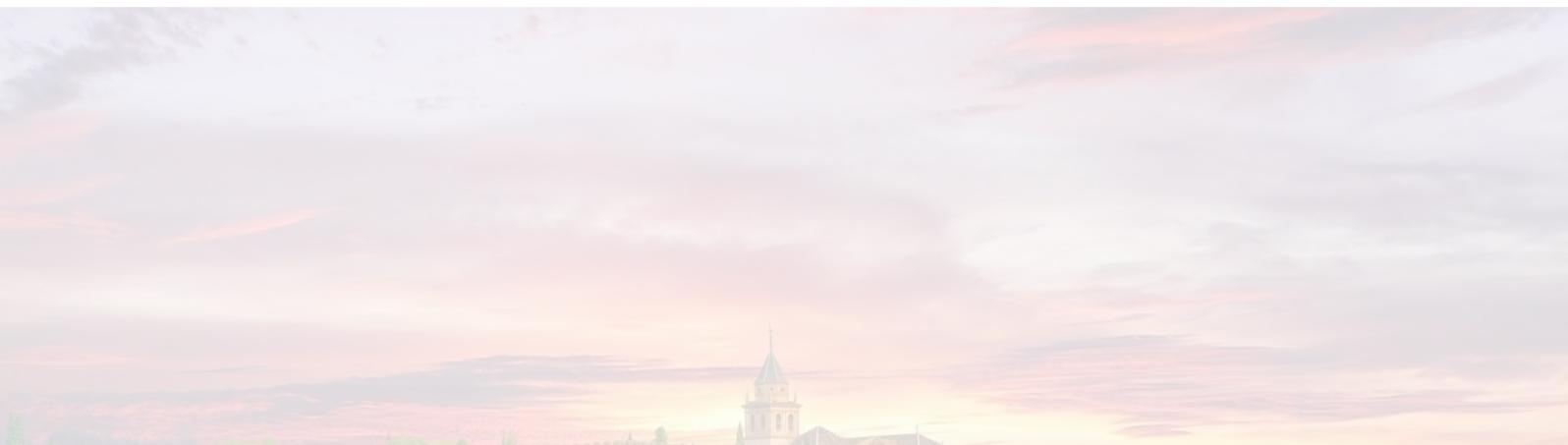
$$WW^t = \left( (X^t X)^{-1} X^t \right) \left( (X^t X)^{-1} X^t \right)^t = (X^t X)^{-1},$$

$$WD^t = \left( (X^t X)^{-1} X^t \right) D^t = 0_{k \times k},$$

$$DW^t = D \left( (X^t X)^{-1} X^t \right)^t = 0_{k \times k},$$

$$\begin{aligned} \text{var} \left( \vec{\hat{\beta}}^* \right) &= \sigma^2 (WW^t + WD^t + DW^t + DD^t) = \\ &= \sigma^2 WW^t + \sigma^2 DD^t = \sigma^2 (X^t X)^{-1} + \sigma^2 DD^t = \\ &= \text{var} \left( \vec{\hat{\beta}} \right) + \sigma^2 DD^t \end{aligned}$$

$$\text{var} \left( \vec{\hat{\beta}}^* \right) - \text{var} \left( \vec{\hat{\beta}} \right) = \sigma^2 DD^t \succ 0 \Rightarrow \text{var} \left( \vec{\hat{\beta}}^* \right) \succ \text{var} \left( \vec{\hat{\beta}} \right).$$



### 1.3. ESTIMACIÓN DE $\sigma_u^2$

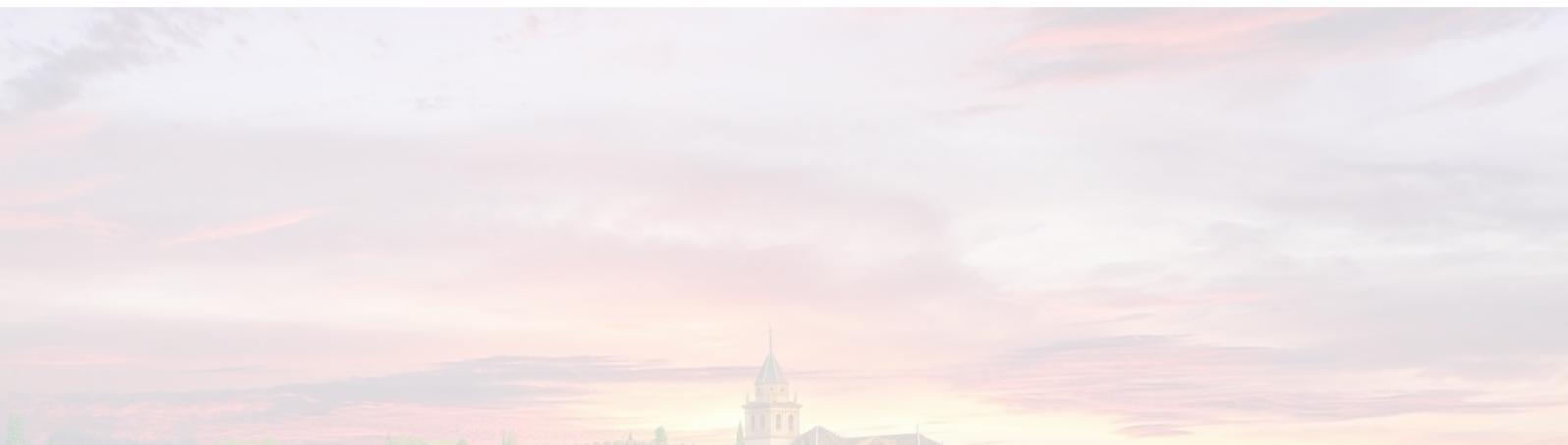
El vector de residuos es:

$$\begin{aligned}\vec{e} &= \vec{y} - \vec{\hat{y}} = \vec{y} - X \vec{\hat{\beta}} = \vec{y} - X \left( (X^t X)^{-1} (X^t \vec{y}) \right) = \\ &= \left( I_n - X (X^t X)^{-1} X^t \right) \vec{y} = M_X \vec{y}\end{aligned}$$

donde  $M_X = I_n - X (X^t X)^{-1} X^t$  es la matriz complemento del proyector ortogonal.

La matriz  $M_X$  verifica:

- ▶  $M_X^t = M_X$  (simétrica)
- ▶  $M_X^2 = M_X$  (idempotente)
- ▶  $\text{traza}(M_X) = n - k$ .
- ▶  $M_X X = \left( I_n - X (X^t X)^{-1} X^t \right) X = X - X = 0_{n \times k}$ .



### 1.3. ESTIMACIÓN DE $\sigma_u^2$

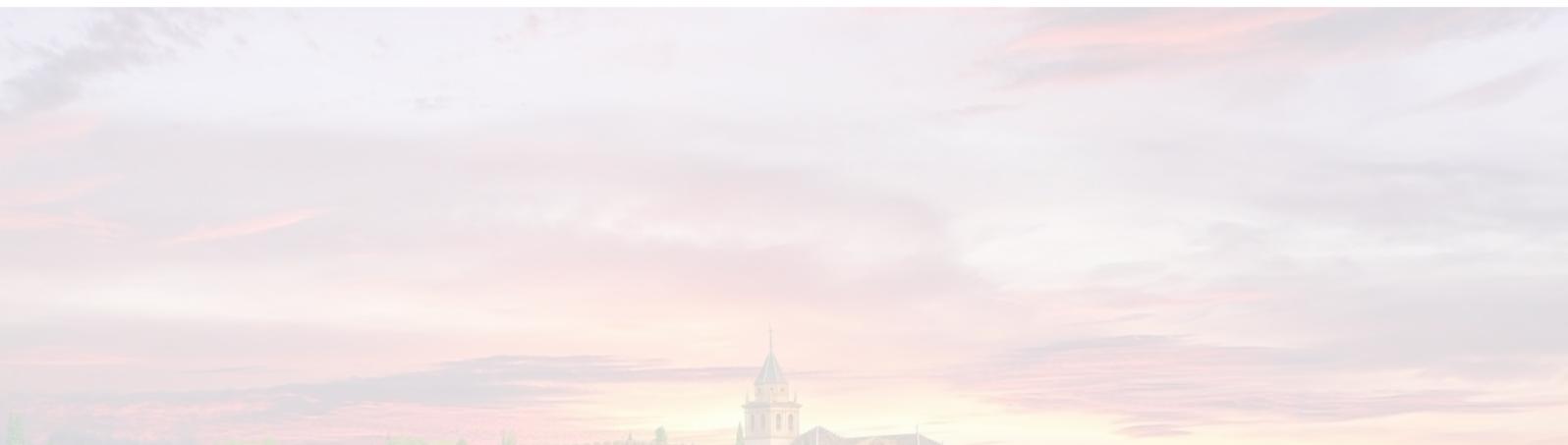
Como  $\vec{y} = X\vec{\beta} + \vec{u}$ :

$$\vec{e} = M_X \vec{y} = M_X (X\vec{\beta} + \vec{u}) = M_X X\vec{\beta} + M_X \vec{u} = M_X \vec{u}$$

Tomando valores esperados se tiene:

$$\begin{aligned} E[\vec{e}^t \vec{e}] &= E[(M_X \vec{u})^t M_X \vec{u}] &&= E[\vec{u}^t M_X \vec{u}] \\ &= E[\text{traza}(\vec{u}^t M_X \vec{u})] &&= \text{traza}(M_X E[\vec{u} \vec{u}^t]) \\ &= \text{traza}(M_X \sigma^2 I_n) &&= \sigma^2 \text{traza}(M_X) = \sigma^2 (n - k) \end{aligned}$$

(Nota:  $\vec{u}^t M_X \vec{u}$  es un escalar.)



### 1.3. ESTIMACIÓN DE $\sigma_u^2$

Como  $\vec{y} = X\vec{\beta} + \vec{u}$ :

$$\vec{e} = M_X \vec{y} = M_X (X\vec{\beta} + \vec{u}) = M_X X\vec{\beta} + M_X \vec{u} = M_X \vec{u}$$

Tomando valores esperados se tiene:

$$\begin{aligned} E[\vec{e}^t \vec{e}] &= E[(M_X \vec{u})^t M_X \vec{u}] &&= E[\vec{u}^t M_X \vec{u}] \\ &= E[\text{traza}(\vec{u}^t M_X \vec{u})] &&= \text{traza}(M_X E[\vec{u} \vec{u}^t]) \\ &= \text{traza}(M_X \sigma^2 I_n) &&= \sigma^2 \text{traza}(M_X) = \sigma^2 (n - k) \end{aligned}$$

(Nota:  $\vec{u}^t M_X \vec{u}$  es un escalar.)

Como  $E\left[\frac{\vec{e}^t \vec{e}}{n - k}\right] = \sigma^2$ , se concluye que un estimador insesgado de la varianza de las perturbaciones,  $\sigma_u^2$ , será:

$$\hat{\sigma}_u^2 = \frac{\vec{e}^t \vec{e}}{n - k} = \frac{SCR}{n - k}$$

## 1.4. ESTIMACIÓN DE LA VARIANZA DE LOS EMCO

La varianza del EMCO  $\vec{\beta}$  era:

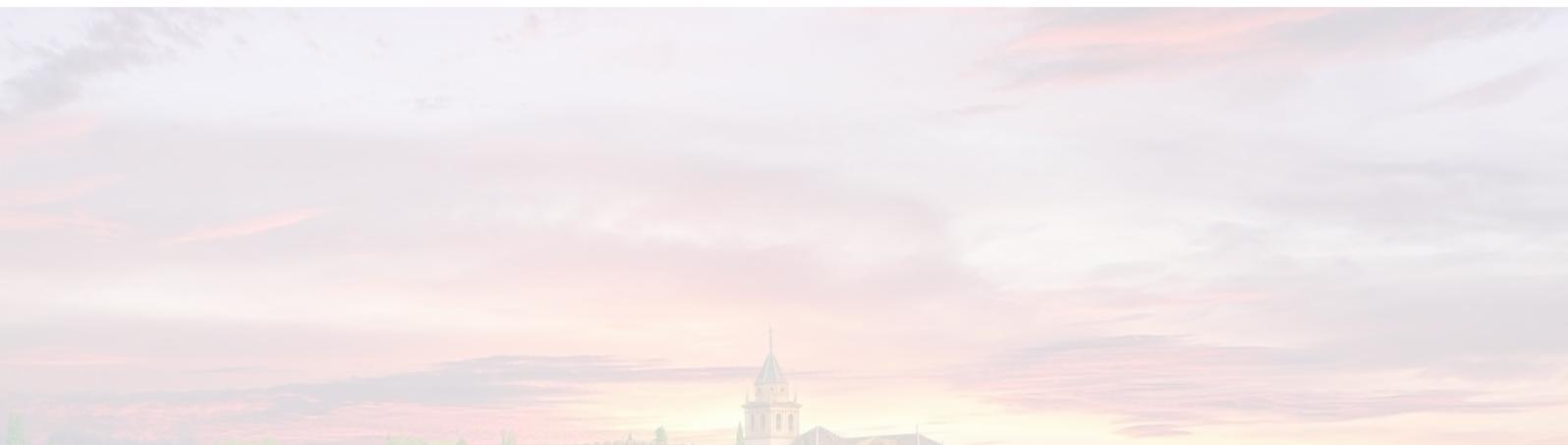
$$\text{var} \left( \vec{\beta} \right) = \sigma^2 \cdot (X^t X)^{-1},$$

Como dicha expresión depende de la varianza de la perturbación, **no se puede determinar**. Teniendo en cuenta que el estimador insesgado de la varianza del término de perturbación es

$$\hat{\sigma}^2 = \frac{\vec{e}^t \vec{e}}{n - k} = \frac{SCR}{n - k},$$

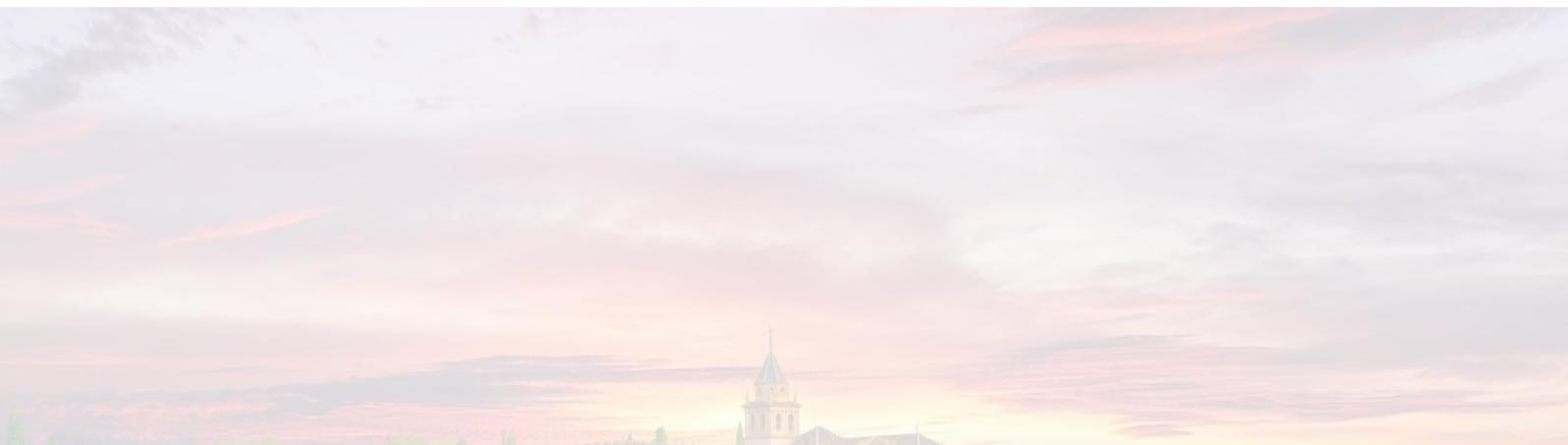
se obtiene que la estimación de la matriz de varianzas-covarianzas de  $\vec{\beta}$  es:

$$\widehat{\text{var}} \left( \vec{\beta} \right) = \frac{SCR}{n - k} \cdot (X^t X)^{-1}.$$



## 2.SUPUESTO DE NORMALIDAD E INFERENCIA

- Suponíamos que  $\vec{u} \sim \mathcal{N}(\vec{0}, \sigma^2 I_n)$ .



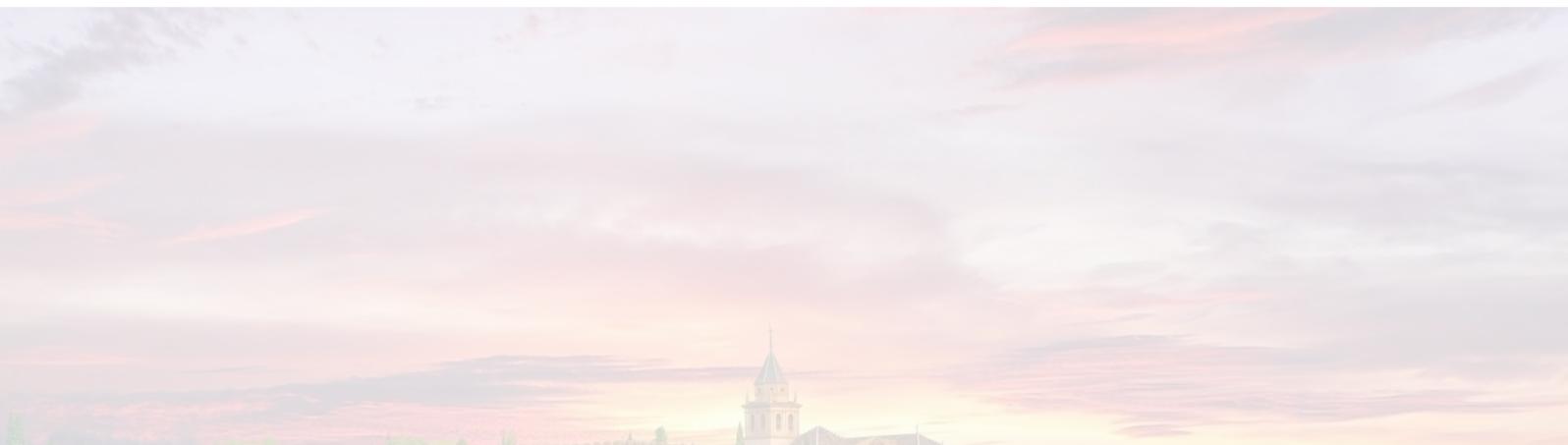
## 2.SUPUESTO DE NORMALIDAD E INFERENCIA

- ▶ Suponíamos que  $\vec{u} \sim \mathcal{N}(\vec{0}, \sigma^2 I_n)$ .
- ▶ Como  $\vec{\hat{\beta}} = \vec{\beta} + (X^t X)^{-1} X^t \vec{u}$ , y por T. Gauss-Markov:

$$\vec{\hat{\beta}} \sim \mathcal{N}\left(E\left[\vec{\hat{\beta}}\right], \text{var}\left(\vec{\hat{\beta}}\right)\right) = \mathcal{N}\left(\vec{\beta}, \sigma^2 \cdot (X^t X)^{-1}\right)$$

o equivalentemente:

$$\left(\vec{\hat{\beta}} - \vec{\beta}\right) \sim \mathcal{N}\left(\vec{0}, \sigma^2 \cdot (X^t X)^{-1}\right)$$



## 2. SUPUESTO DE NORMALIDAD E INFERENCIA

► Suponíamos que  $\vec{u} \sim \mathcal{N}(\vec{0}, \sigma^2 I_n)$ .

► Como  $\vec{\hat{\beta}} = \vec{\beta} + (X^t X)^{-1} X^t \vec{u}$ , y por T. Gauss-Markov:

$$\vec{\hat{\beta}} \sim \mathcal{N}\left(E\left[\vec{\hat{\beta}}\right], \text{var}\left(\vec{\hat{\beta}}\right)\right) = \mathcal{N}\left(\vec{\beta}, \sigma^2 \cdot (X^t X)^{-1}\right)$$

o equivalentemente:

$$\left(\vec{\hat{\beta}} - \vec{\beta}\right) \sim \mathcal{N}\left(\vec{0}, \sigma^2 \cdot (X^t X)^{-1}\right)$$

► Como se verifica que:

$$\vec{e}^t \vec{e} = \vec{u}^t M_X \vec{u}$$

$$\vec{u} \sim \mathcal{N}\left(\vec{0}, \sigma^2 I_n\right)$$

$M_X$  es simétrica, idempotente y  $\rho(M_X) = n - k$

Concluimos que  $\frac{1}{\sigma^2} \vec{u}^t M_X \vec{u} \sim \chi_{n-k}^2$ .

## 2. SUPUESTO DE NORMALIDAD E INFERENCIA

- ▶ Suponíamos que  $\vec{u} \sim \mathcal{N}(\vec{0}, \sigma^2 I_n)$ .
- ▶ Como  $\vec{\hat{\beta}} = \vec{\beta} + (X^t X)^{-1} X^t \vec{u}$ , y por T. Gauss-Markov:

$$\vec{\hat{\beta}} \sim \mathcal{N}\left(E\left[\vec{\hat{\beta}}\right], \text{var}\left(\vec{\hat{\beta}}\right)\right) = \mathcal{N}\left(\vec{\beta}, \sigma^2 \cdot (X^t X)^{-1}\right)$$

o equivalentemente:

$$\left(\vec{\hat{\beta}} - \vec{\beta}\right) \sim \mathcal{N}\left(\vec{0}, \sigma^2 \cdot (X^t X)^{-1}\right)$$

- ▶ Como se verifica que:

$$\vec{e}^t \vec{e} = \vec{u}^t M_X \vec{u}$$

$$\vec{u} \sim \mathcal{N}\left(\vec{0}, \sigma^2 I_n\right)$$

$M_X$  es simétrica, idempotente y  $\rho(M_X) = n - k$

Concluimos que  $\frac{1}{\sigma^2} \vec{u}^t M_X \vec{u} \sim \chi_{n-k}^2 \Rightarrow \frac{(n-k)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-k}^2$

## 2.1. INTERVALOS DE CONFIANZA

► IC para  $\beta_i$ :

$$\hat{\beta}_i \pm t_{n-k, 1-\frac{\alpha}{2}} \cdot \sqrt{\widehat{\text{var}}[\hat{\beta}_i]}, \quad i = 1, \dots, k.$$

siendo  $\widehat{\text{var}}[\hat{\beta}_i]$  el elemento  $(i, i)$  de la matriz de varianzas-covarianzas estimada del estimador  $\vec{\hat{\beta}}$ , es decir, el elemento  $(i, i)$  de  $\widehat{\text{var}}\left(\vec{\hat{\beta}}\right)$ .

► IC para  $\sigma^2$

$$\left[ \frac{(n-k) \cdot \hat{\sigma}^2}{\chi_{n-k, 1-\frac{\alpha}{2}}^2}, \frac{(n-k) \cdot \hat{\sigma}^2}{\chi_{n-k, \frac{\alpha}{2}}^2} \right],$$

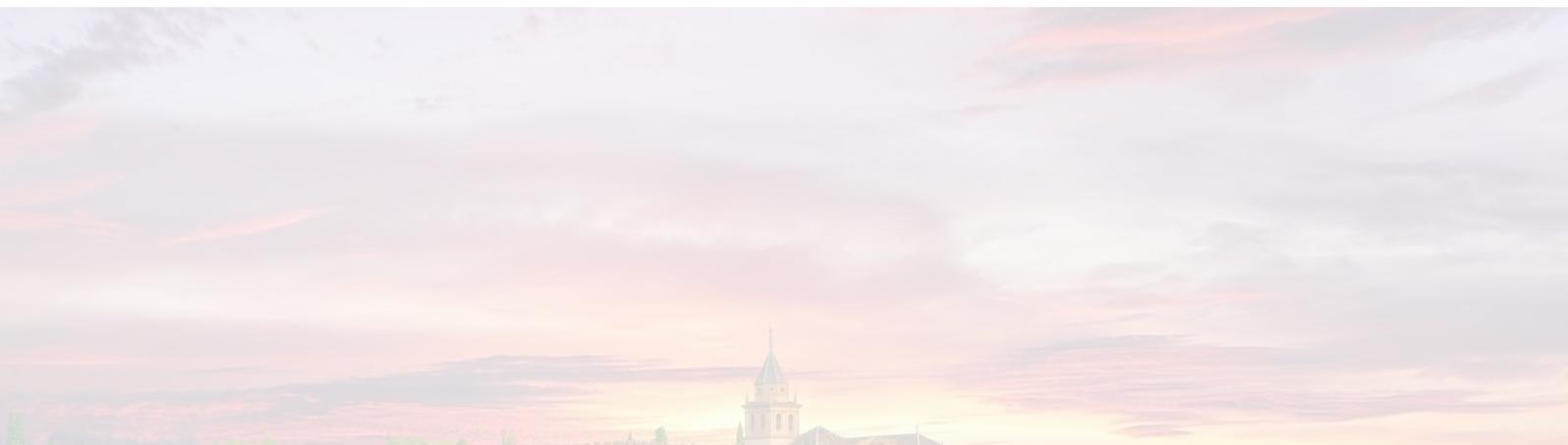
donde  $\chi_{n-k, 1-\frac{\alpha}{2}}^2$  y  $\chi_{n-k, \frac{\alpha}{2}}^2$  son los cuantiles de una distribución chi-cuadrado con  $n-k$  grados de libertad tal que  $P[\chi \leq \chi_{n-k, 1-\frac{\alpha}{2}}^2] = 1 - \frac{\alpha}{2}$  y  $P[\chi \leq \chi_{n-k, \frac{\alpha}{2}}^2] = \frac{\alpha}{2}$ .



## 2.2.CONTRASTE DE HIPÓTESIS

El estadístico de contraste asociado a la hipótesis  $H_0 : R \vec{\beta} = \vec{r}$  es

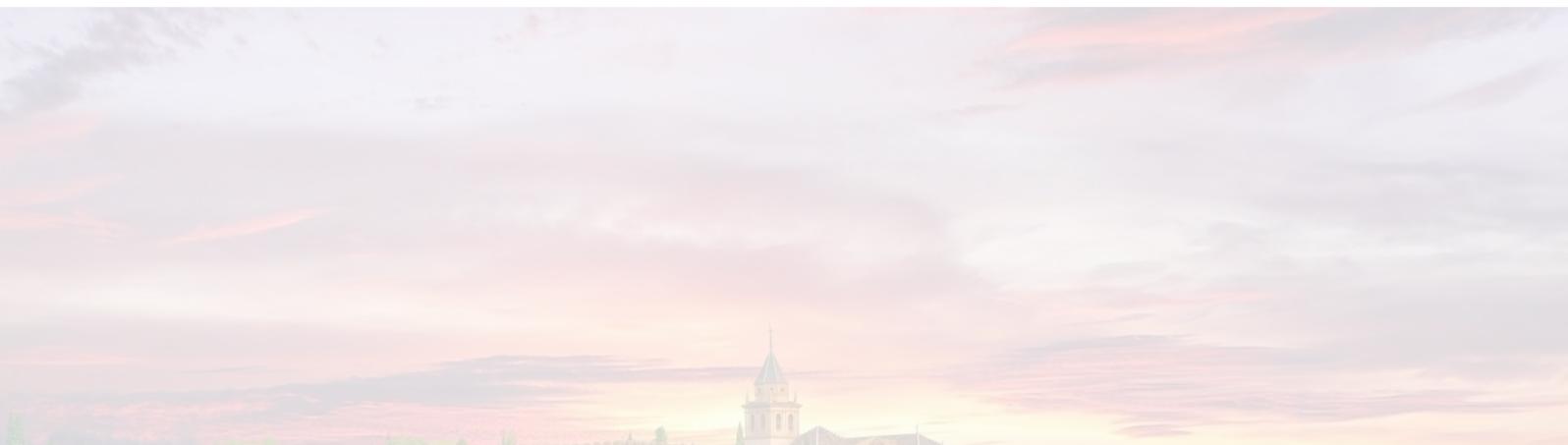
$$\begin{aligned} F_{exp} &= \frac{\left( R \vec{\hat{\beta}} - \vec{r} \right)^t \cdot \left[ R (X^t X)^{-1} R^t \right]^{-1} \cdot \left( R \vec{\hat{\beta}} - \vec{r} \right)}{\frac{m}{\frac{\vec{e}^t \vec{e}}{n-k}}} \\ &= \left( R \vec{\hat{\beta}} - \vec{r} \right)^t \cdot \frac{\left[ R (X^t X)^{-1} R^t \right]^{-1}}{m \hat{\sigma}^2} \cdot \left( R \vec{\hat{\beta}} - \vec{r} \right) \sim F_{m,n-k}. \end{aligned}$$



## 2.2.CONTRASTE DE HIPÓTESIS. TEST GENERAL

La hipótesis nula se rechazará si el valor del estadístico cae en la región de rechazo.

Es decir, si  $F_{exp} > F_{m,n-k,1-\alpha}$  entonces se rechaza la hipótesis  $H_0 : R\vec{\beta} = \vec{r}$ .



## 2.2.1. CASOS PARTICULARES (I): SIGNIFICACIÓN INDIVIDUAL

Si  $m = 1$  y  $r_i = 0 \forall i$ .

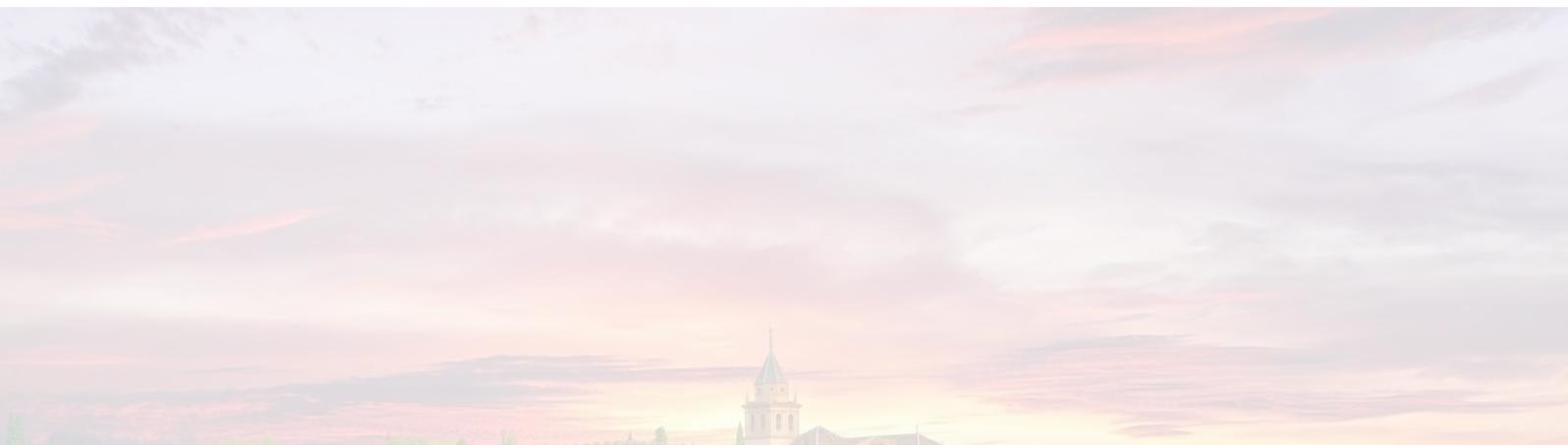
► **Hipótesis:**

$$\begin{cases} H_0 : \beta_i = 0 \\ H_1 : \beta_i \neq 0 \end{cases}$$

► **Estadístico de contraste:**

$$t_{exp} = \left| \frac{\hat{\beta}_i}{\sqrt{\widehat{\text{var}}[\hat{\beta}_i]}} \right|$$

► **Conclusión:** Se rechaza  $H_0$  si  $t_{exp} > t_{n-k, 1-\frac{\alpha}{2}}$ .



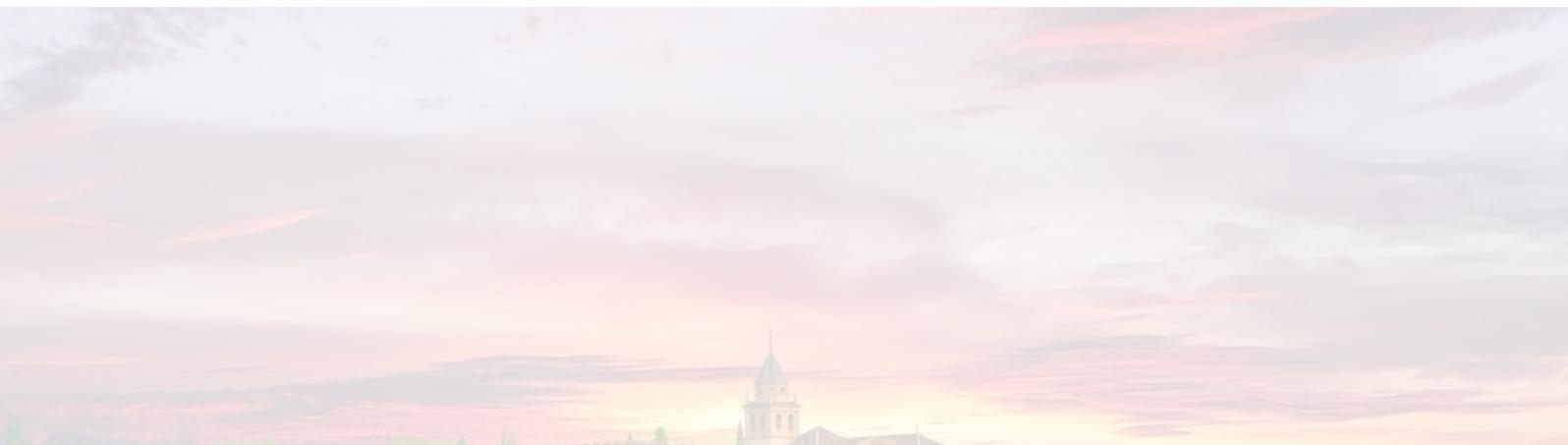
## 2.2.1. CASOS PARTICULARES (II): ÚNICA RESTRICCIÓN LINEAL

$m = 1$  y  $r_i \neq 0 \forall i$ .

- ▶ **Hipótesis:**  $\begin{cases} H_0 : \beta_i = r_i \\ H_1 : \beta_i \neq r_i \end{cases}$
- ▶ **Estadístico de contraste:**

$$t_{exp} = \left| \frac{\hat{\beta}_i - r_i}{\sqrt{\widehat{\text{var}}[\hat{\beta}_i]}} \right|$$

- ▶ **Conclusión:** Se rechaza  $H_0$  si  $t_{exp} > t_{n-k, 1-\frac{\alpha}{2}}$ .



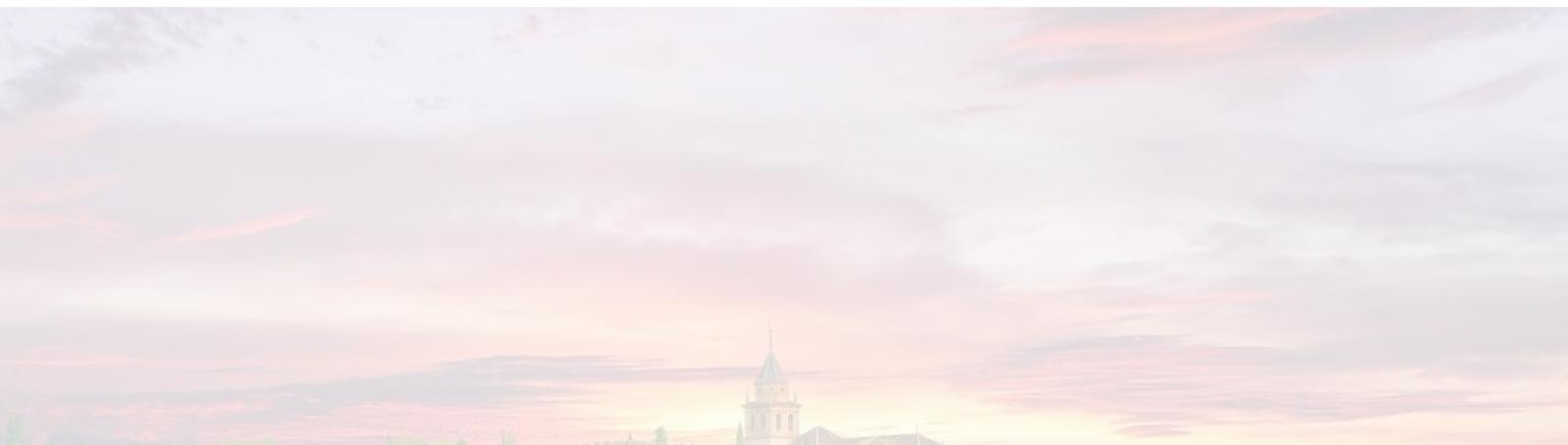
## 2.2.1. CASOS PARTICULARES (III): SIGNIFICACIÓN GLOBAL

Si  $m = k - 1$  y  $r_i = 0$   $i = 2, 3, \dots, k$ .

- ▶ **Hipótesis:** 
$$\begin{cases} H_0 : \beta_2 = \beta_3 = \dots = \beta_k = 0 \\ H_1 : \exists \beta_i \neq 0 \text{ con } i = 2, 3, \dots, k \end{cases}$$
- ▶ **Estadístico de contraste:**

$$F_{exp} = \left( R \begin{matrix} \vec{\beta} \\ \hat{\beta} \end{matrix} \right)^t \cdot \frac{\left[ R (X^t X)^{-1} R^t \right]^{-1}}{(k-1) \cdot \hat{\sigma}^2} \cdot \left( R \begin{matrix} \vec{\beta} \\ \hat{\beta} \end{matrix} \right)$$

- ▶ **Conclusión:** Se rechaza  $H_0$  si  $F_{exp} > F_{k-1, n-k, 1-\alpha}$ .



## 2.2.1. CONTRASTE DE SIGNIFICACIÓN GLOBAL

Teníamos:

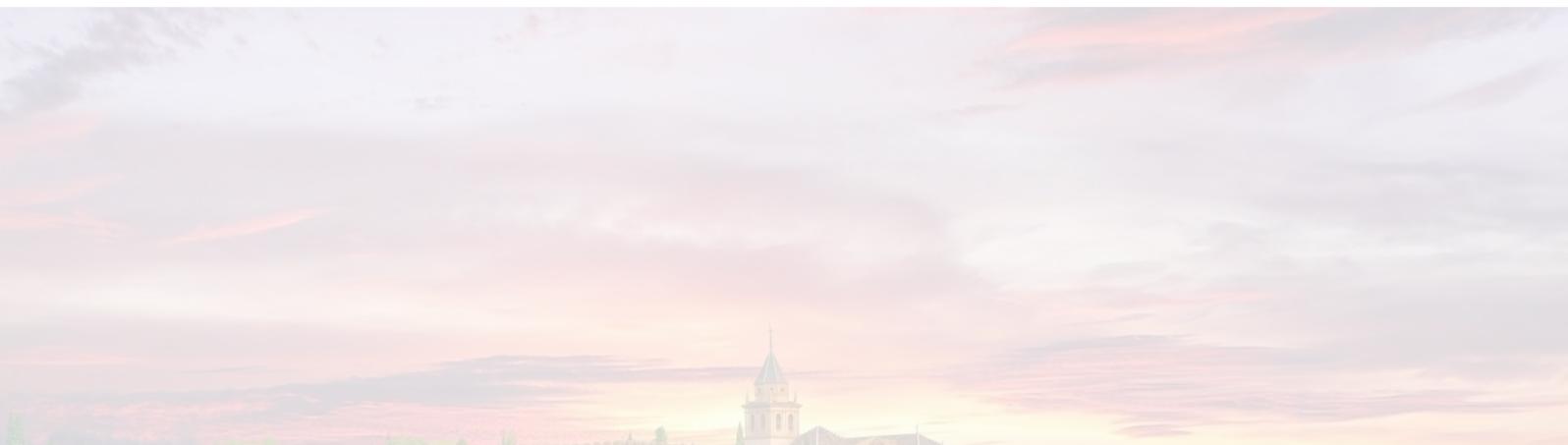
$$\begin{cases} H_0 : \beta_2 = \beta_3 = \dots = \beta_k = 0 \\ H_1 : \exists \beta_i \neq 0 \text{ con } i = 2, 3, \dots, k \end{cases}$$

Equivalentemente:

$$\begin{cases} H_0 : R\vec{\beta} = \vec{r} \\ H_1 : \text{No se verifica } H_0 \end{cases}$$

donde

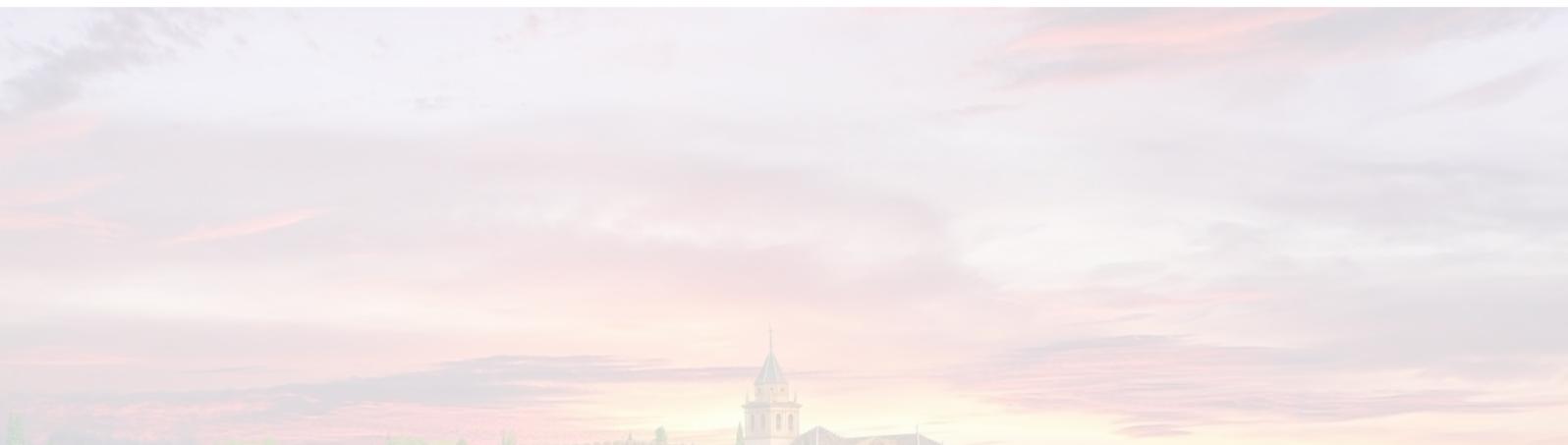
$$R = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}_{(k-1) \times k} \quad \vec{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}_{k \times 1} \quad \vec{r} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{(k-1) \times 1}$$



## 2.2.1. CONTRASTE DE SIGNIFICACIÓN GLOBAL

Realizando la partición de las matrices:

$$X^t X = \begin{bmatrix} \mathbf{1}^t \\ X_2^t \end{bmatrix} \begin{bmatrix} \mathbf{1}^t & X_2^t \end{bmatrix} = \begin{bmatrix} n & \mathbf{1}^t X_2 \\ X_2^t \mathbf{1} & X_2^t X_2 \end{bmatrix} \quad \vec{\hat{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \vec{\hat{\beta}}_2 \end{bmatrix}$$

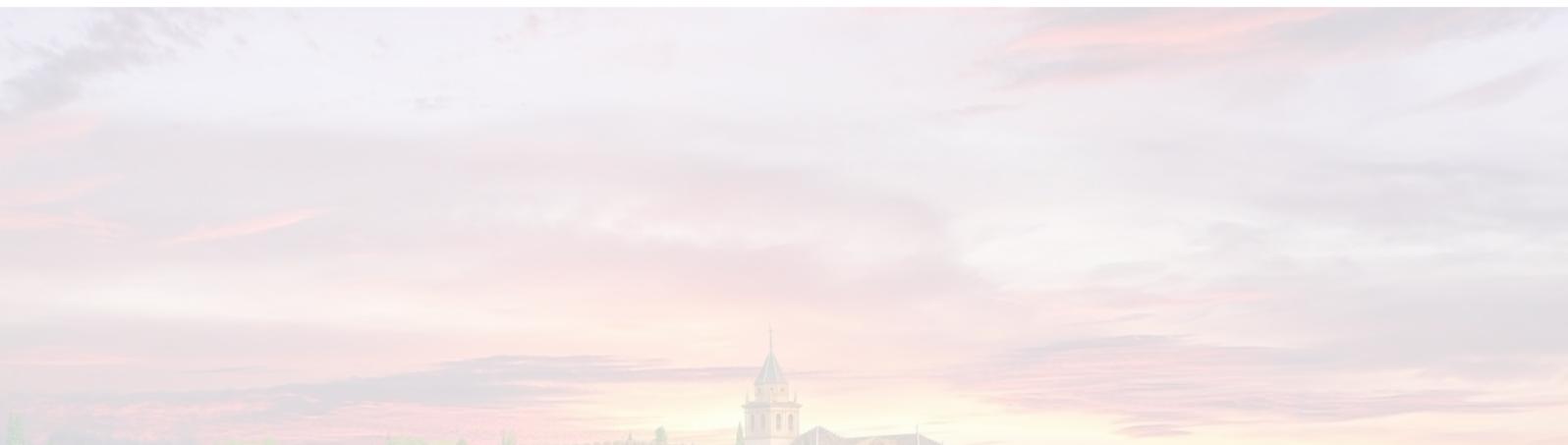


## 2.2.1. CONTRASTE DE SIGNIFICACIÓN GLOBAL

Realizando la partición de las matrices:

$$X^t X = \begin{bmatrix} \mathbf{1}^t \\ X_2^t \end{bmatrix} \begin{bmatrix} \mathbf{1}^t & X_2^t \end{bmatrix} = \begin{bmatrix} n & \mathbf{1}^t X_2 \\ X_2^t \mathbf{1} & X_2^t X_2 \end{bmatrix} \quad \vec{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \vec{\beta}_2 \end{bmatrix}$$

$$\frac{\left( R \vec{\beta} - \vec{r} \right)^t \cdot \left[ R \left( X^t X \right)^{-1} R^t \right]^{-1} \cdot \left( R \vec{\beta} - \vec{r} \right)}{\frac{m}{\vec{e}^t \vec{e}} \cdot \frac{1}{n-k}}$$



## 2.2.1. CONTRASTE DE SIGNIFICACIÓN GLOBAL

Realizando la partición de las matrices:

$$X^t X = \begin{bmatrix} \mathbf{1}^t \\ X_2^t \end{bmatrix} \begin{bmatrix} \mathbf{1}^t & X_2^t \end{bmatrix} = \begin{bmatrix} n & \mathbf{1}^t X_2 \\ X_2^t \mathbf{1} & X_2^t X_2 \end{bmatrix} \quad \vec{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \vec{\beta}_2 \end{bmatrix}$$

$$\frac{\left( R \vec{\beta} - \vec{r} \right)^t \cdot \left[ R \left( X^t X \right)^{-1} R^t \right]^{-1} \cdot \left( R \vec{\beta} - \vec{r} \right)}{\frac{m}{\vec{e}^t \vec{e}} \cdot \frac{1}{n-k}}$$

$$\frac{\vec{\beta}_2^t \cdot \left[ X_2^t X_2 - X_2^t \mathbf{1} \frac{1}{n} \mathbf{1}^t X_2 \right]^{-1} \cdot \vec{\beta}_2}{\frac{k-1}{\vec{e}^t \vec{e}} \cdot \frac{1}{n-k}} = \frac{\vec{\beta}_2^t X_2^t A X_2 \vec{\beta}_2}{\frac{k-1}{\vec{e}^t \vec{e}} \cdot \frac{1}{n-k}} \sim F_{k-1, n-k}$$

donde  $A = I_n - \frac{1}{n} \mathbf{1} \cdot \mathbf{1}^t$ .

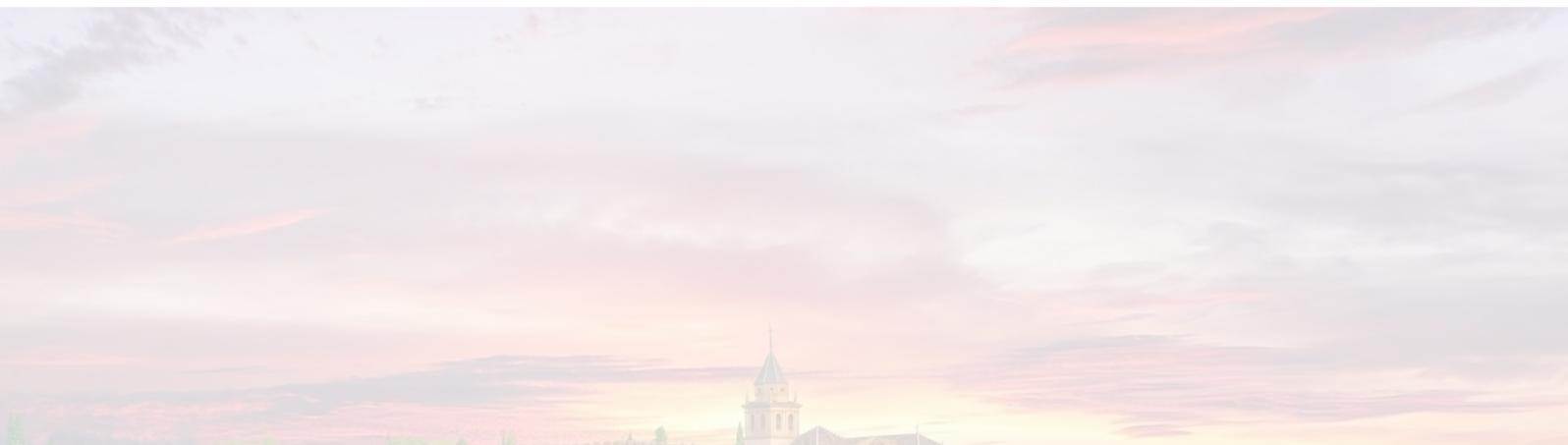
## 2.2.1. ANOVA

Considerando que se verifica

$$\begin{aligned} \vec{y}^t A \vec{y} &= \vec{\beta}_2^t X_2^t A X_2 \vec{\beta}_2 + \vec{e}^t \vec{e} \\ (SCT) &= (SCE) + (SCR) \end{aligned}$$

El estadístico experimental se puede escribir como:

$$F_{exp} = \frac{\frac{1}{k-1} SCE}{\frac{1}{n-k} SCR}$$



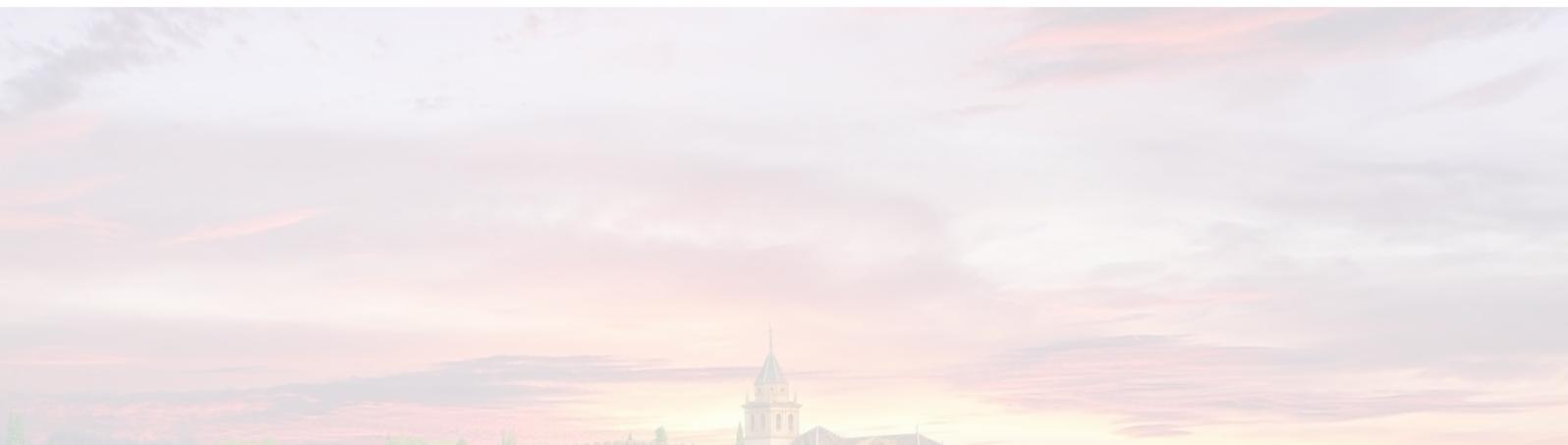
## 2.2.1. ANOVA

ANOVA El **Análisis de la Varianza** es el contraste que estudia la significación global del modelo donde:

- Hipótesis  $\begin{cases} H_0 : \beta_2 = \beta_3 = \dots = \beta_k = 0 \\ H_1 : \exists \beta_i \neq 0 \text{ con } i = 2, 3, \dots, k \end{cases}$
- Estadístico de contraste

$$F_{exp} = \frac{\frac{1}{k-1} SCE}{\frac{1}{n-k} SCR}$$

- Conclusión Se rechaza  $H_0$  si  $F_{exp} > F_{k-1, n-k, 1-\alpha}$ .



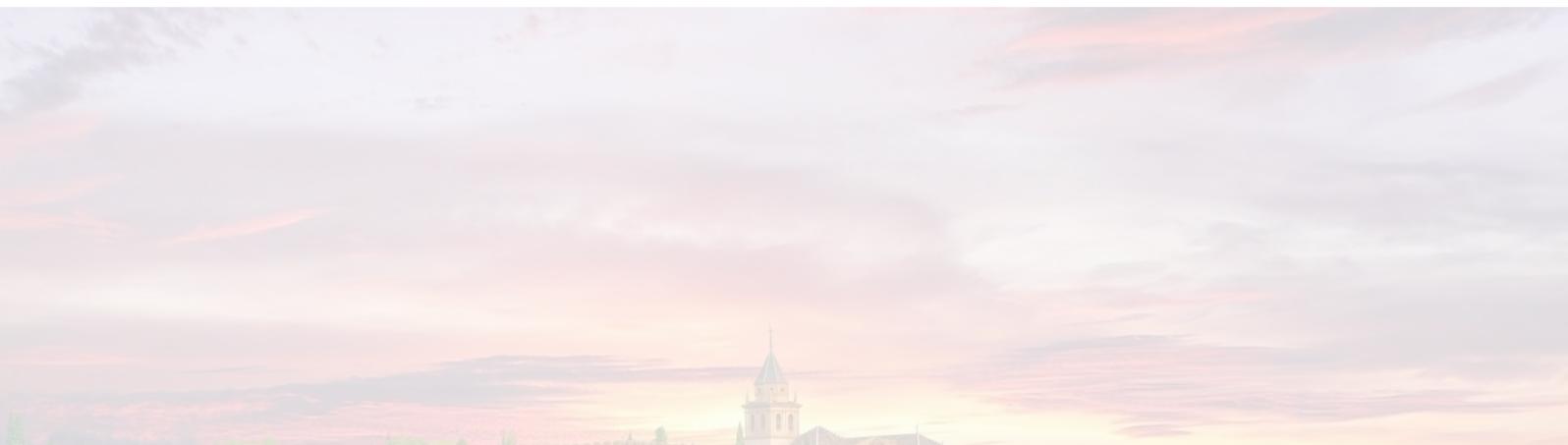
## TABLA ANOVA

Fuente de variación	Suma de Cuadrados	Grados de Libertad	Medias
Explicada	$SCE = \vec{\beta}^t X^t \vec{y} - n\bar{Y}^2$	$k - 1$	$\frac{SCE}{k-1}$
Residuos	$SCR = \vec{y}^t \vec{y} - \vec{\beta}^t X^t \vec{y}$	$n - k$	$\frac{SCR}{n-k}$
Total	$SCT = \vec{y}^t \vec{y} - n\bar{Y}^2$	$n - 1$	

Como

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT}$$

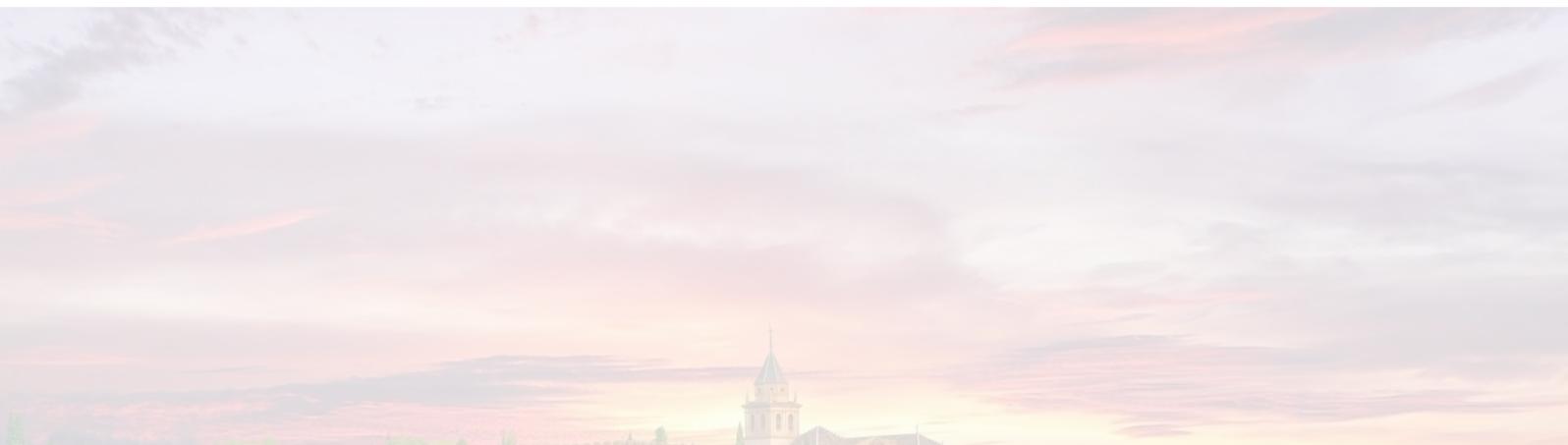
$$F_{exp} = \frac{\frac{1}{k-1} SCE}{\frac{1}{n-k} SCT(1 - R^2)} = \frac{\frac{1}{k-1} R^2}{\frac{1}{n-k} (1 - R^2)} = \frac{n - k}{k - 1} \frac{R^2}{(1 - R^2)}$$



## COTA $R^2$

Se rechaza  $H_0$  cuando  $F_{exp} > F_{k-1, n-k, 1-\alpha}$ , esto es si

$$\frac{n-k}{k-1} \frac{R^2}{(1-R^2)} > F_{k-1, n-k, 1-\alpha} \Rightarrow R^2 > \frac{(k-1) \cdot F_{k-1, n-k, 1-\alpha}}{(n-k) + (k-1) \cdot F_{k-1, n-k, 1-\alpha}}.$$



### 3. PREDICCIÓN PUNTUAL

Una vez estimado y validado el modelo  $Y = \vec{\beta} X + u_t$ :

- ▶ Si tenemos nuevos datos  $\vec{x}_0^t = (1 \ X_{20} \ X_{30} \ \dots \ X_{k0})$ ,
- ▶ Suponiendo que tenemos **permanencia estructural en la especificación del modelo** (el proceso de generación de los datos para la nueva observación  $\vec{x}_0$  es el mismo que ha generado la información muestral)
- ▶ Podemos tomar como **predictor para el valor medio y para el valor individual de la variable endógena**

$$\vec{x}_0^t \vec{\beta} = \hat{y}_0 = \mathbf{E}[\widehat{y_0} | \vec{x}_0]$$

(el predictor mínimo cuadrático  $\hat{y}_0$  es lineal, insesgado y óptimo)

#### Nota 1

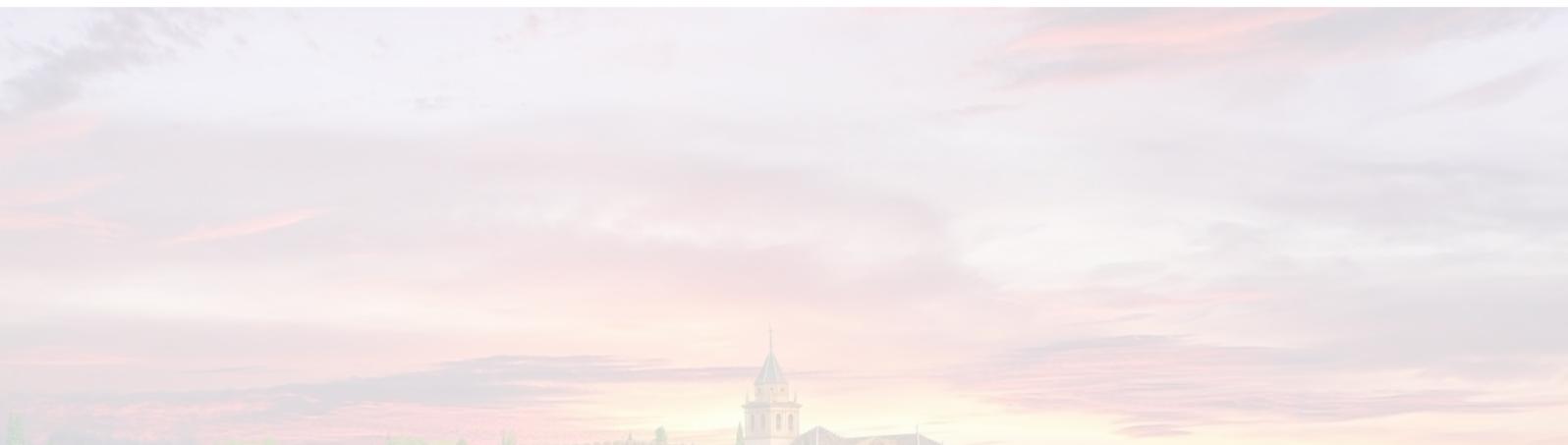
*Adviértase que, aunque se obtiene el mismo predictor para el valor individual  $y_0$  como para  $\mathbf{E}[\widehat{y_0} | \vec{x}_0]$ , los errores de predicción obtenidos con ambas variables no coinciden entre sí, y por consiguiente las varianzas asociadas a dichos errores también serán distintas.*

### 3. PREDICCIÓN PUNTUAL PARA EL VALOR INDIVIDUAL

Considerando el valor individual  $y_0 = \vec{x}_0^t \vec{\beta} + u_0$  se tiene que el error de predicción corresponde:

$$e_0 = y_0 - \hat{y}_0 = \vec{x}_0^t \vec{\beta} + u_0 - \vec{x}_0^t \vec{\tilde{\beta}} = u_0 - \vec{x}_0^t (\vec{\tilde{\beta}} - \vec{\beta}) \sim \mathcal{N}(E[e_0], \text{var}[e_0])$$

$$E[e_0] = E[u_0 - \vec{x}_0^t (\vec{\tilde{\beta}} - \vec{\beta})] = \vec{x}_0^t (E[\vec{\tilde{\beta}}] - \vec{\beta}) = 0.$$



### 3. PREDICCIÓN PUNTUAL PARA EL VALOR INDIVIDUAL

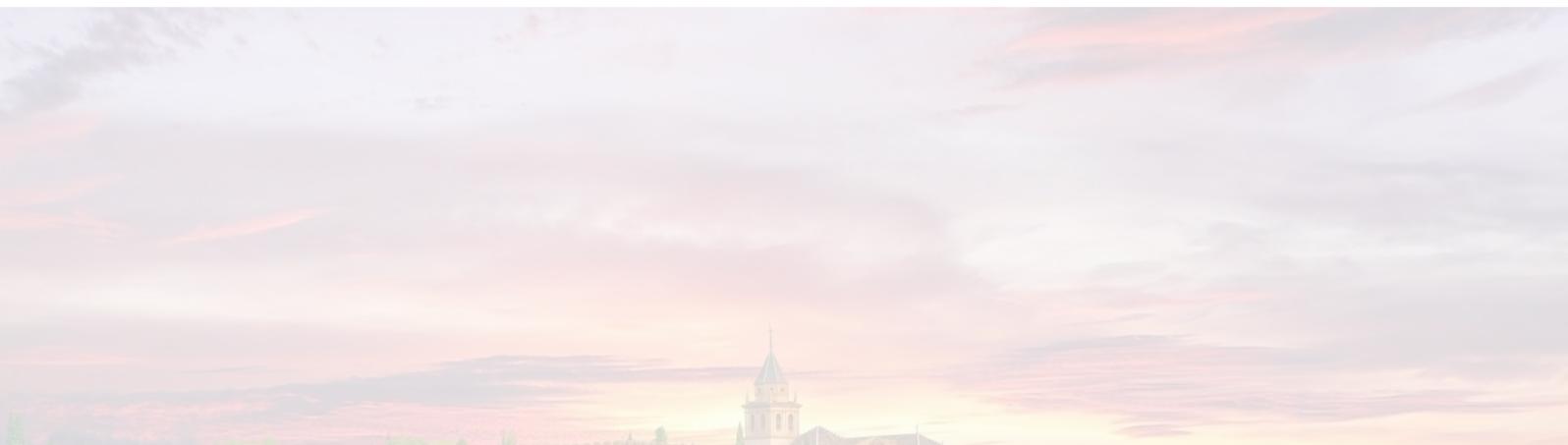
$$\begin{aligned}
 \text{var}[e_0] &= \text{var}[u_0 - \vec{x}_0^t(\vec{\beta} - \vec{\beta})] = \text{var}(u_0) + \text{var}(\vec{x}_0^t(\vec{\beta} - \vec{\beta})) - 2\text{cov}(u_0 \vec{x}_0^t(\vec{\beta} - \vec{\beta})) \\
 &= \text{var}(u_0) + \text{var}(\vec{x}_0^t(\vec{\beta} - \vec{\beta})) - 2\text{cov}(u_0 \vec{x}_0^t(X^t X)^{-1} X^t \vec{u}) = 1 = \\
 &= \text{var}(u_0) + \text{var}(\vec{x}_0^t(\vec{\beta} - \vec{\beta})) = \sigma^2 + E[\vec{x}_0^t(\vec{\beta} - \vec{\beta})(\vec{\beta} - \vec{\beta})^t \vec{x}_0] = \\
 &= \sigma^2 + \vec{x}_0^t E[(\vec{\beta} - \vec{\beta})(\vec{\beta} - \vec{\beta})^t] \vec{x}_0 = \sigma^2 + \vec{x}_0^t \sigma^2 (X^t X)^{-1} \vec{x}_0 = \\
 &=
 \end{aligned}$$

$$e_0 \sim N(0, \sigma^2(1 + \vec{x}_0^t (X^t X)^{-1} \vec{x}_0))$$

---

<sup>1</sup>  $\text{cov}(u_0 \vec{x}_0^t(\vec{\beta} - \vec{\beta})) = 0$  al no existir correlación entre  $u_0$  y  $\vec{u}$ .

<sup>1</sup>  $\text{cov}(u_0 \vec{x}_0^t(\vec{\beta} - \vec{\beta})) = 0$  al no existir correlación entre  $u_0$  y  $\vec{u}$ .



### 3. PREDICCIÓN PUNTUAL PARA EL VALOR ESPERADO

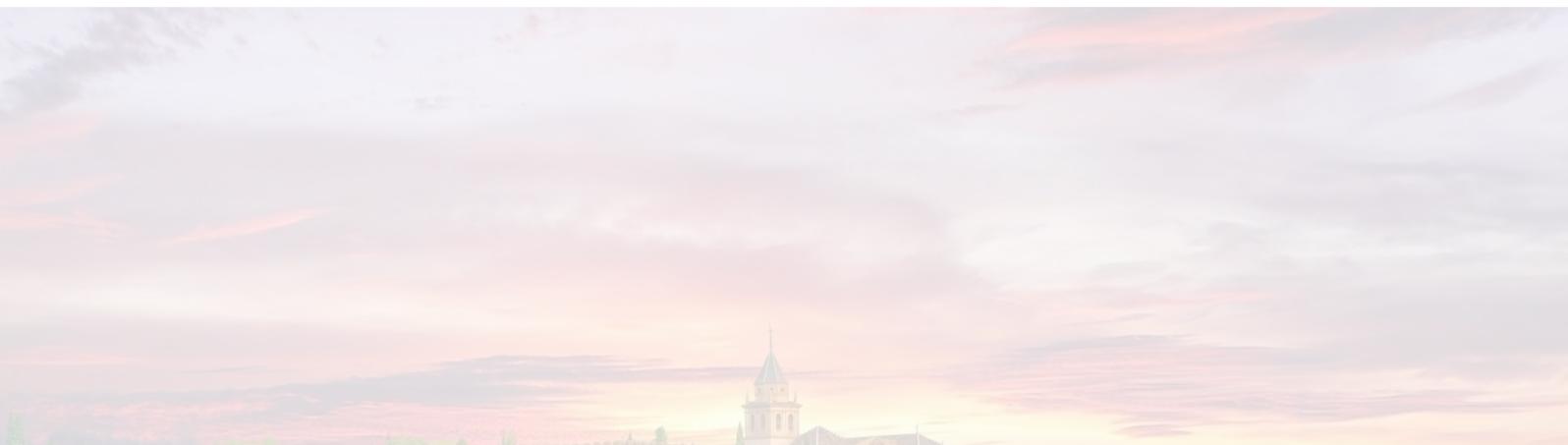
$E[y_0 | \vec{x}_0]$  tiene como residuo:

$$e_0^* = E[y_0 | \vec{x}_0] - \widehat{E[y_0 | \vec{x}_0]} = \vec{x}_0^t \vec{\beta} - \vec{x}_0^t \vec{\hat{\beta}} = -\vec{x}_0^t (\vec{\hat{\beta}} - \vec{\beta}) \sim \mathcal{N}(E[e_0^*], \text{var}[e_0^*])$$

$$E[e_0^*] = E[-\vec{x}_0^t (\vec{\hat{\beta}} - \vec{\beta})] = \vec{x}_0^t (E[\vec{\hat{\beta}}] - \vec{\beta}) = 0.$$

$$\begin{aligned} \text{var}[e_0^*] &= \text{var}[-\vec{x}_0^t (\vec{\hat{\beta}} - \vec{\beta})] = \vec{x}_0^t \text{var}(\vec{\hat{\beta}} - \vec{\beta}) \vec{x}_0 = \\ &= \vec{x}_0^t \text{var}(\vec{\hat{\beta}}) \vec{x}_0 = \vec{x}_0^t \sigma^2 (X^t X)^{-1} \vec{x}_0 = \\ &= \sigma^2 \vec{x}_0^t (X^t X)^{-1} \vec{x}_0. \end{aligned}$$

$$e_0^* = E[y_0 | \vec{x}_0] - \widehat{E[y_0 | \vec{x}_0]} \sim N(0, \sigma^2 \vec{x}_0^t (X^t X)^{-1} \vec{x}_0)$$



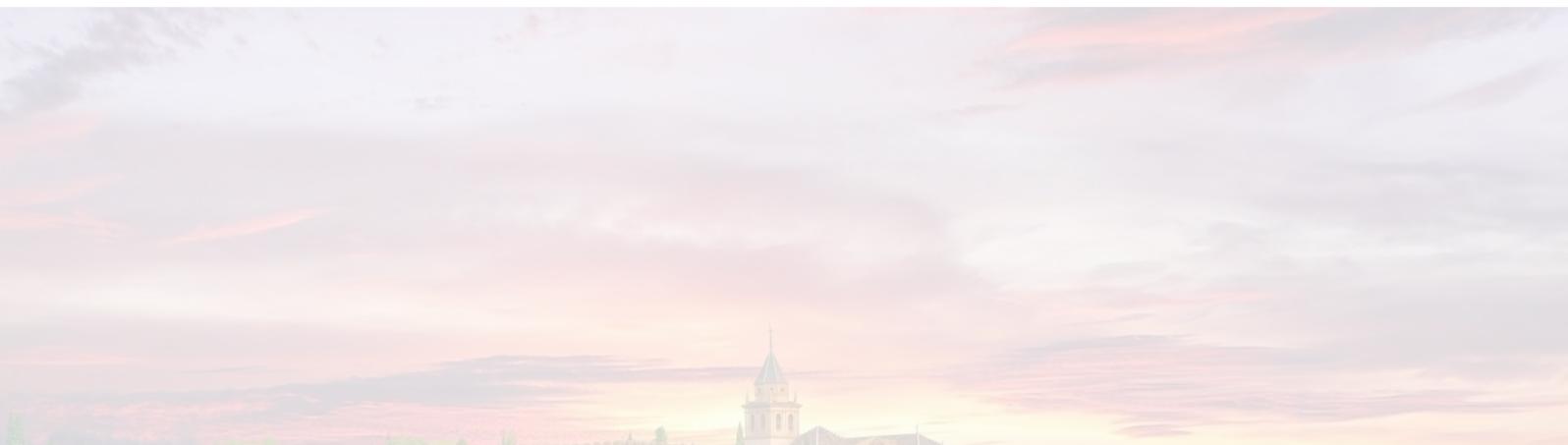
### 3. PREDICCIÓN POR INTERVALO

En las distribuciones de  $e_0$  y  $e_0^*$ ,  $\sigma^2$  es desconocida. Teniendo en cuenta que

$$\frac{(n-k)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-k}^2,$$

$$\frac{\frac{\hat{y}_0 - \vec{x}_0^t \vec{\beta}}{\sqrt{\sigma^2(1 + \vec{x}_0^t (X^t X)^{-1} \vec{x}_0)}}}{\sqrt{\frac{(n-k)\hat{\sigma}^2}{\sigma^2}}}{n-k} = \frac{\hat{y}_0 - \vec{x}_0^t \vec{\beta}}{\hat{\sigma} \sqrt{1 + \vec{x}_0^t (X^t X)^{-1} \vec{x}_0}} \sim t_{n-k}$$

$$\frac{\frac{E[\widehat{y_0 | \vec{x}_0}] - \vec{x}_0^t \vec{\beta}}{\sqrt{\sigma^2(\vec{x}_0^t (X^t X)^{-1} \vec{x}_0)}}}{\sqrt{\frac{(n-k)\hat{\sigma}^2}{\sigma^2}}}{n-k} = \frac{E[\widehat{y_0 | \vec{x}_0}] - \vec{x}_0^t \vec{\beta}}{\hat{\sigma} \sqrt{\vec{x}_0^t (X^t X)^{-1} \vec{x}_0}} \sim t_{n-k}$$



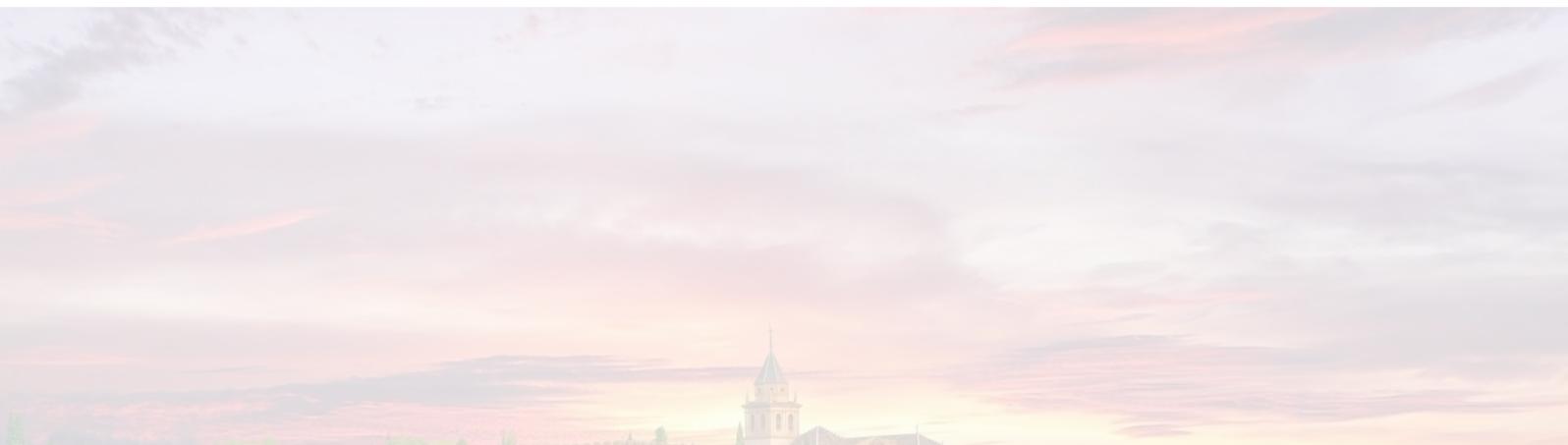
### 3. PREDICCIÓN POR INTERVALO

- Intervalo de confianza para el valor individual  $y_0$ .

$$\begin{aligned} IC_{y_0} &= \hat{y}_0 \pm t_{n-k, 1-\frac{\alpha}{2}} \cdot \hat{\sigma} \sqrt{1 + \vec{x}_0^t (X^t X)^{-1} \vec{x}_0} \\ &= \vec{x}_0^t \vec{\beta} \pm t_{n-k, 1-\frac{\alpha}{2}} \cdot \hat{\sigma} \sqrt{1 + \vec{x}_0^t (X^t X)^{-1} \vec{x}_0} \end{aligned}$$

- Intervalo de confianza para  $E[y_0 | \vec{x}_0]$ .

$$\begin{aligned} IC_{E[y_0 | \vec{x}_0]} &= \widehat{E[y_0 | \vec{x}_0]} \pm t_{n-k, 1-\frac{\alpha}{2}} \cdot \hat{\sigma} \sqrt{\vec{x}_0^t (X^t X)^{-1} \vec{x}_0} \\ &= \vec{x}_0^t \vec{\beta} \pm t_{n-k, 1-\frac{\alpha}{2}} \cdot \hat{\sigma} \sqrt{\vec{x}_0^t (X^t X)^{-1} \vec{x}_0} \end{aligned}$$



## DISTRIBUCIONES $\chi^2$ Y $t$ -STUDENT

- ▶ La distribución  $\chi^2$  de Pearson con  $n$  grados de libertad,  $\chi_n^2$ , se construye como la suma de los cuadrados de  $n$  variables aleatorias independientes e idénticamente distribuidas según una  $\mathcal{N}(0, 1)$ :

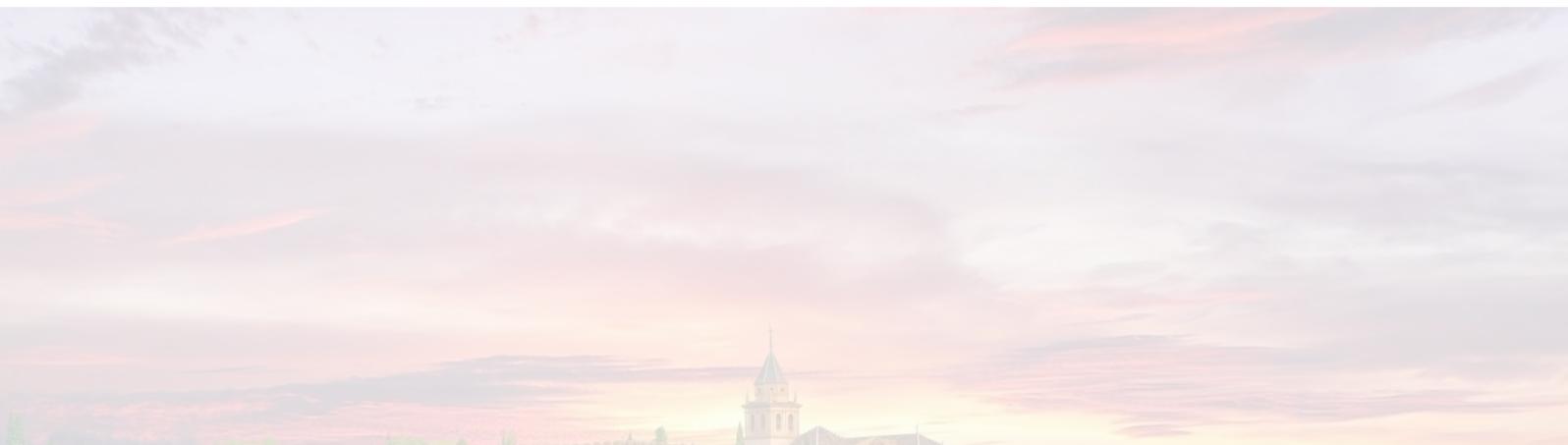
$$\chi_n^2 = \sum_{i=1}^n X_i^2, \quad X_i \sim N(0, 1), \quad \forall i.$$

- ▶ La distribución  $t$ -Student con  $n$  grados de libertad,  $t_n$ , se construye como el cociente entre una variable aleatoria  $\mathcal{N}(0, 1)$  y la raíz cuadrada de una  $\chi^2$ -cuadrado de  $n$  grados de libertad dividida entre sus grados de libertad, siendo ambas distribuciones independientes:

$$t_n = \frac{X}{\sqrt{\frac{Y}{n}}}, \quad X \sim N(0, 1), \quad Y \sim \chi_n^2.$$

### Propiedades:

- ▶  $\chi^2$  NO es simétrica,  $t$ -Student SÍ es simétrica.
- ▶ Para  $n > 30$ , la distribución  $t$ -Student se puede aproximar a una distribución Normal.



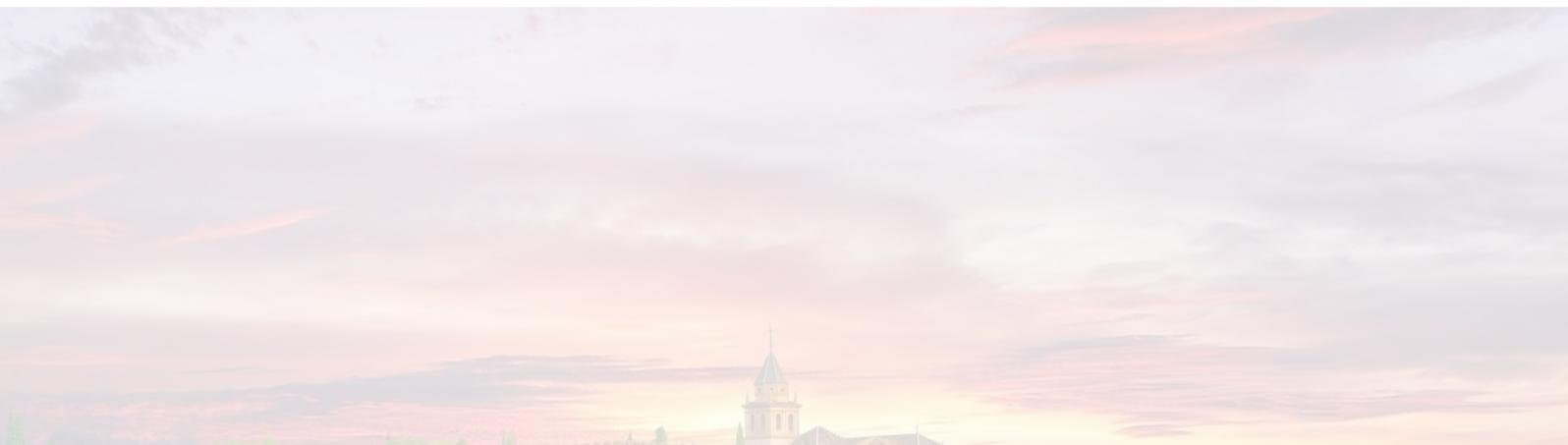
## DISTRIBUCIÓN F-SNEDECOR

- ▶ La distribución  $F$  de Snedecor con  $n$  y  $m$  grados de libertad, que denotaremos por  $F_{n,m}$ , se construye a partir del cociente de dos variables aleatorias independientes y distribuidas según chi-cuadrados con  $n$  y  $m$  grados de libertad, respectivamente, divididas entre sus correspondientes grados de libertad. Por tanto, la distribución  $F$  de Snedecor responde a la siguiente estructura:

$$F_{n,m} = \frac{\frac{X}{n}}{\frac{Y}{m}}, X \sim \chi_n^2, Y \sim \chi_m^2.$$

### Propiedades:

- ▶ La distribución  $F_{n,m}$  NO es simétrica.
- ▶  $F_{m,n,1-\alpha} = \frac{1}{F_{n,m,\alpha}}$ .
- ▶  $t_n^2 = F_{1,n}$ .



## NOTACIÓN MATRICIAL

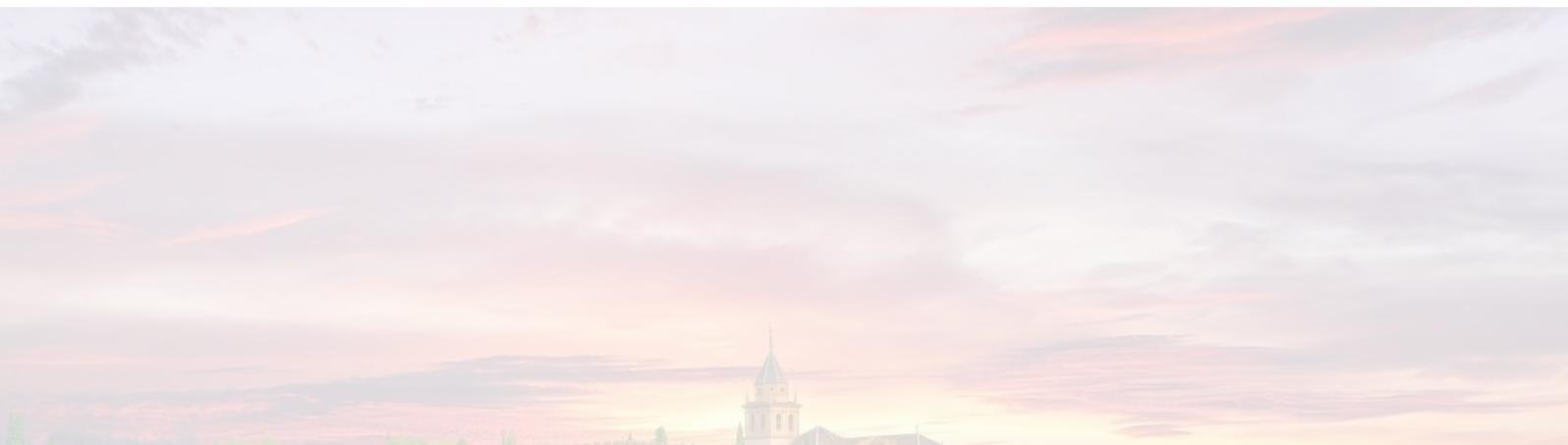
Se define la matriz  $A$  como

$$\begin{aligned} A &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} - \frac{1}{n} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} (1 \ 1 \ \dots \ 1) \\ &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} - \frac{1}{n} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix} \end{aligned}$$

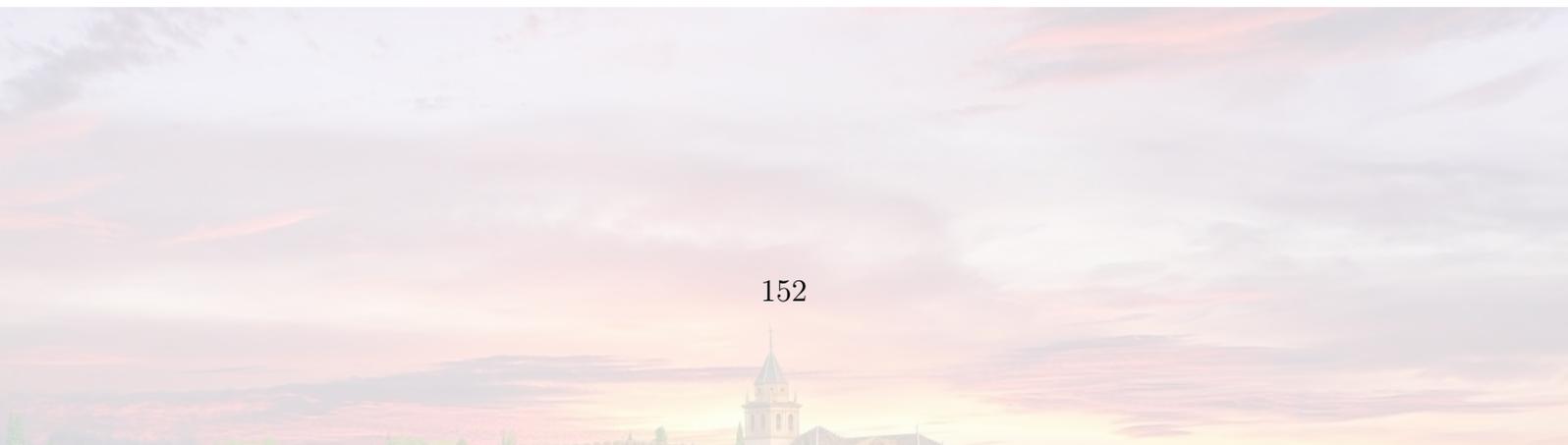
### Propiedades:

- ▶ La matriz  $A$  es simétrica e idempotente.
- ▶  $A(1, \dots, 1)^t = \vec{0}$  y  $A\vec{e} = \vec{e}$ .
- ▶  $A\vec{y}$  es:

$$A\vec{y} = \begin{pmatrix} Y_1 - \bar{Y} \\ Y_2 - \bar{Y} \\ \vdots \\ Y_n - \bar{Y} \end{pmatrix}$$



### 3.1. Apendice Variables Ficticias



## Apéndice A

### Uso de variables ficticias

Existe la posibilidad de introducir variables cualitativas en el modelo de regresión. Ese tipo de variables se consideran dicotómicas, tomando el valor 1 si la variable en estudio esta presente y el valor 0 si no lo esta. Por ejemplo, si analizamos el salario de distintos trabajadores se puede introducir la variable “poseer grado universitario” considerando que toma el valor 1 si lo posee y el valor 0 en caso contrario.

En este caso, dado el modelo:

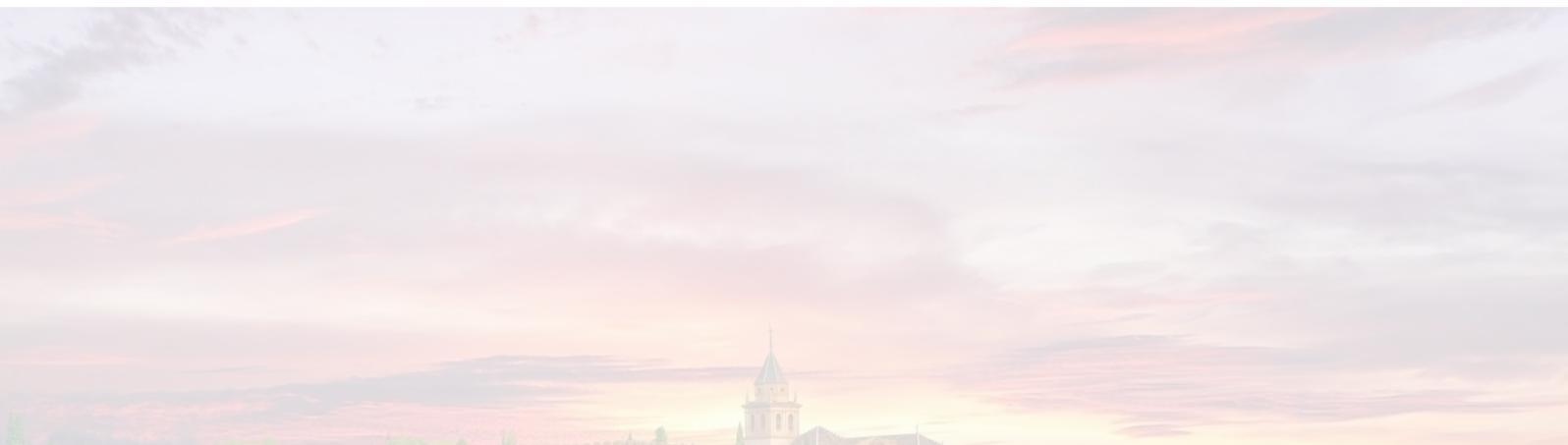
$$Y_t = \beta_0 + \beta_1 D_t + u$$

donde  $Y_t$  es la variable dependiente y la variable  $D_t$  es la variable independiente de tipo cualitativo.

De esta manera si:

- Si  $D_t = 1$  el modelo queda expresado como:

$$E(Y_t) = \beta_0 + \beta_1$$



- Si  $D_t = 0$  el modelo queda expresado como:

$$E(Y_t) = \beta_0$$

Por lo tanto, la interpretación de los coeficientes sería la siguiente:  $\beta_0$  sería el valor medio de la variable explicada cuando no se satisface la categoría ( $D_t = 0$ ) y  $\beta_1$  sería la diferencia en el valor medio de la variable explicada en el caso de que se satisfaga la categoría ( $D_t = 1$ ).

Si incluimos una variable cuantitativa en el modelo, por ejemplo, los años de experiencia ( $X_2$ ):

$$Y_t = \beta_0 + \beta_1 D_t + \beta_2 X_t + u$$

De manera que:

- Si  $D_t = 1$  el modelo quedaría expresado como:

$$E(Y_t) = \beta_0 + \beta_1 + \beta_2 X_t$$

- Si  $D_t = 0$  el modelo quedaría expresado como:

$$E(Y_t) = \beta_0 + \beta_2 X_t$$

Entonces, la interpretación de los coeficientes sería la siguiente:  $\beta_0$  es el valor medio de la variable explicada cuando no se satisface la categoría ( $D_t = 0$ ) y la variable cuantitativa toma el valor 0,  $\beta_1$  es la diferencia en el valor medio de la variable explicada en el caso de que se satisfaga la categoría ( $D_t = 1$ ) y la variable cuantitativa se mantenga constante.

**Ejemplo 18** Se tiene el siguiente modelo obtenido a partir de datos de 41 trabajadores en relación a su retribución mensual ( $Y$ ), si poseen o no grado universitario ( $X_1$ ) (toma el valor 1 si poseen) y sus años de experiencia ( $X_2$ ):

$$Y = 845,48 + 7122,08X_1 + 1301,31X_2$$
$$R^2 = 0,857$$

La estimación obtenida para los parámetros se interpreta de la siguiente forma:

- $\hat{\beta}_0 = 845,48$  se interpreta como la retribución de un empleado sin grado universitario y cero años de experiencia.
- $\hat{\beta}_1 = 7122,08$  se interpretan como la diferencia entre la retribución anual de un empleado con grado universitario y otro que no lo tenga, a igualdad del resto de variables explicativas.
- $\hat{\beta}_2 = 1301,31$  se interpreta como lo que aumenta la retribución anual de un empleado por cada año de experiencia, a igualdad del resto de variables explicativas.

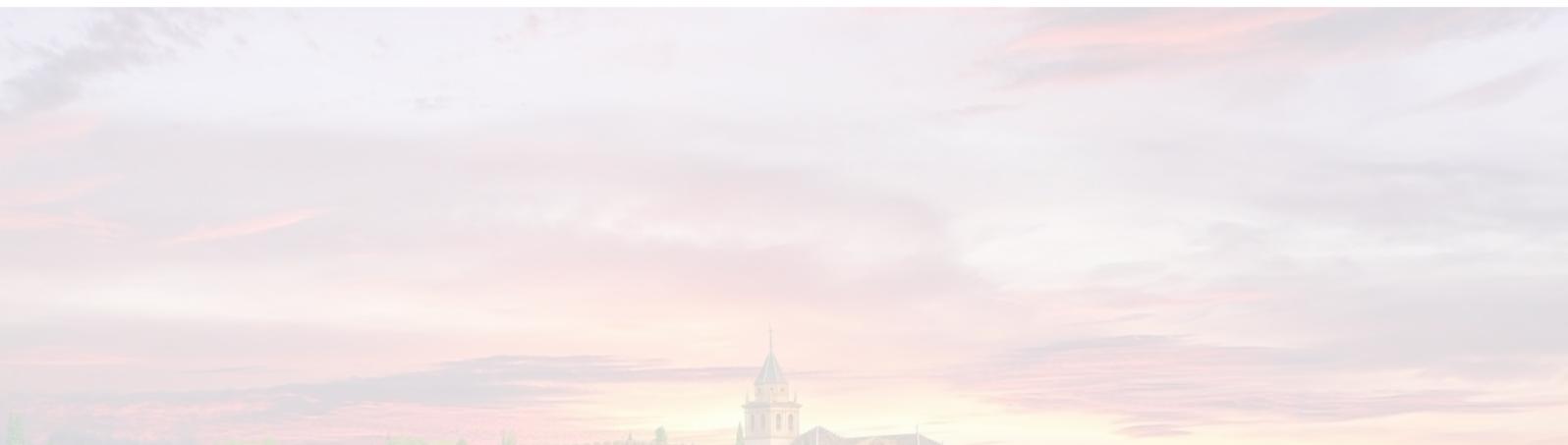
### **Trampa de las variables ficticias**

Al introducir la variable cualitativa en el modelo hay que tener presente el número de categorías que incluye, ya que una de las categorías debe quedarse como variable de control y, por tanto, se introducen tantas variables dicotómicas en el modelo como categorías tenga la variable menos una. De hecho, al describir la variable “poseer grado universitario” hemos introducido una única variable porque si incluyéramos ambas, la suma de las dos variables coincidiría con el término independiente dando lugar a un problema de colinealidad exacta que impediría la estimación del modelo. Este problema se conoce como trampa de las variables ficticias. Por

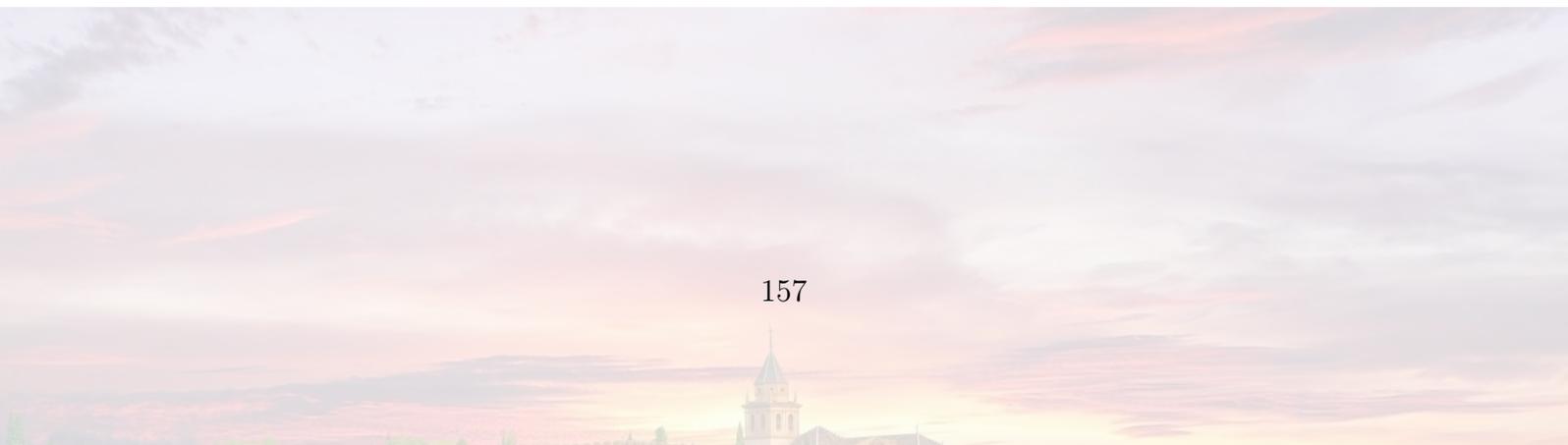
ejemplo, si queremos incorporar la variable “titulacion cursada en la FCCEE” que hemos descrito en cuatro categorías (GMIM, GADE, GFICO, GECO), se crearían tres variables dicotómicas:

- $D_1$  toma el valor 1 si la titulacion cursada es GMIM y cero si no lo es.
- $D_2$  toma el valor 1 si la titulacion cursada es GADE y cero si no lo es.
- $D_3$  toma el valor 1 si la titulacion cursada es GFICO y cero si no lo es.

quedando la variable GECO como variable de control.



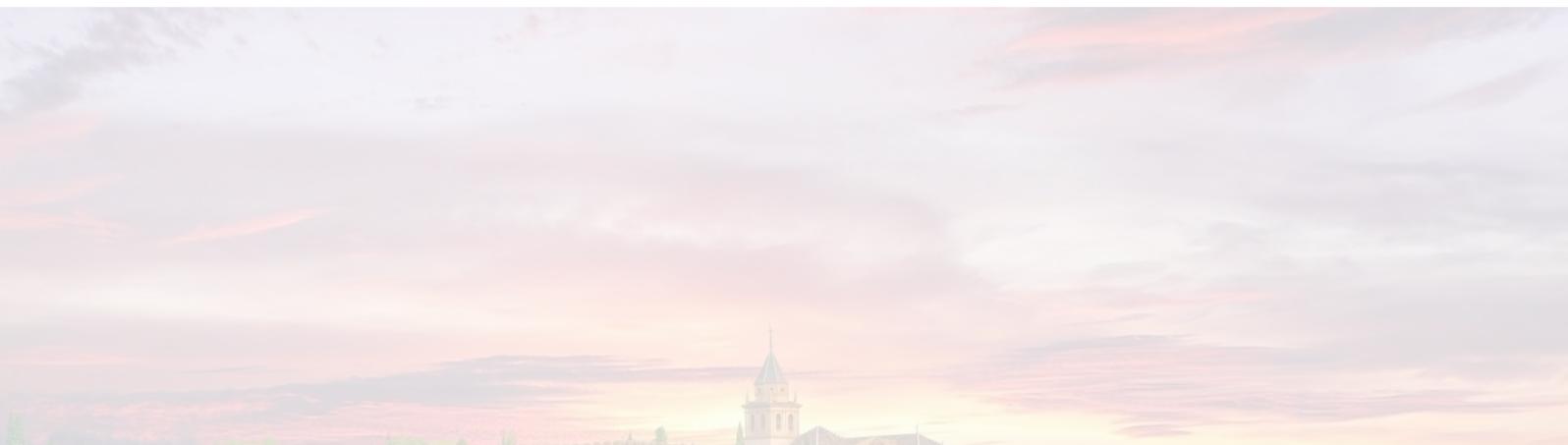
## 4 Tema 4



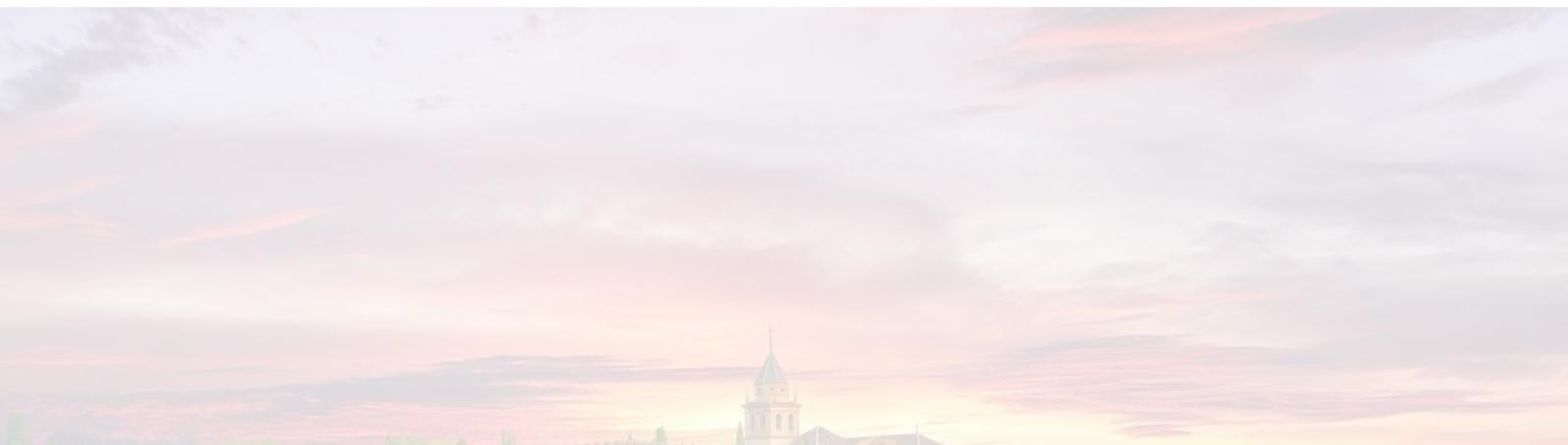
# ECONOMETRÍA

## TEMA 4: MULTICOLINEALIDAD

2024-2025

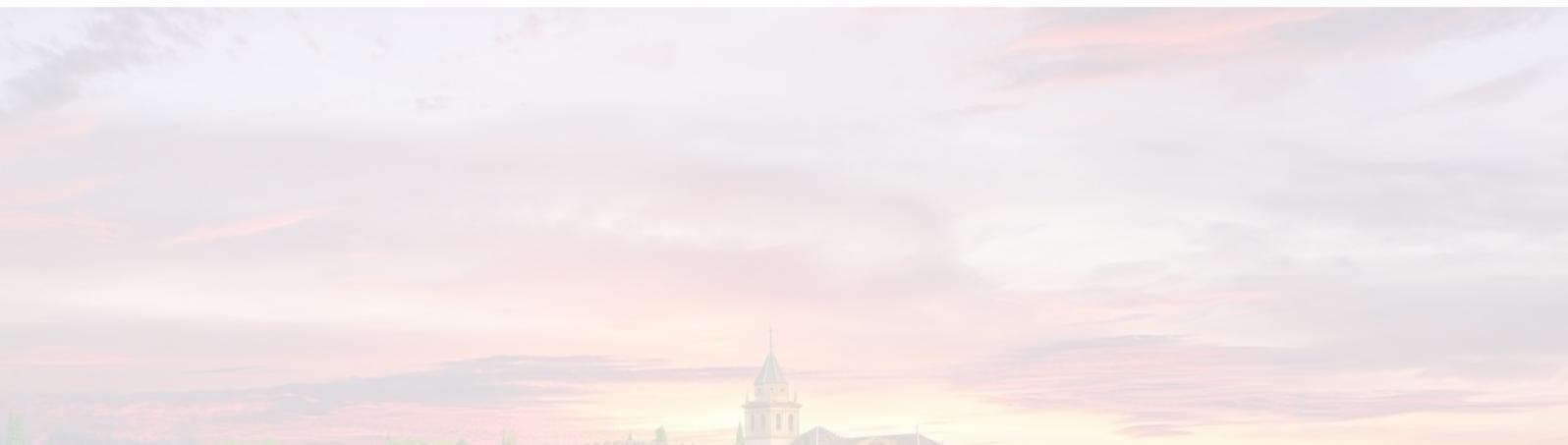


<b>1</b>	<b>Concepto, causas y consecuencias</b>	<b>2</b>
1.1	Multicolinealidad exacta	3
1.2	Multicolinealidad aproximada	4
1.3	Consecuencias	8
1.3.1	Consecuencias sobre el análisis estadístico del modelo	8
1.3.2	Efectos nocivos sobre el análisis numérico del modelo	10
<b>2</b>	<b>Detección de multicolinealidad aproximada</b>	<b>11</b>
2.1	Factor de Inflación de la Varianza	11
2.2	Número de Condición	12
2.3	Otras medidas	13
<b>3</b>	<b>Soluciones</b>	<b>14</b>



## CONCEPTO

- ▶ El problema de multicolinealidad consiste en la existencia de relaciones lineales entre dos o más variables independientes del modelo lineal uniecuacional múltiple.
- ▶ Dependiendo de cómo sea dicha relación lineal hablaremos de multicolinealidad perfecta (o exacta) o aproximada.
- ▶ Dentro de la multicolinealidad aproximada, distinguiremos entre multicolinealidad aproximada errática o sistemática y multicolinealidad aproximada esencial o no esencial.



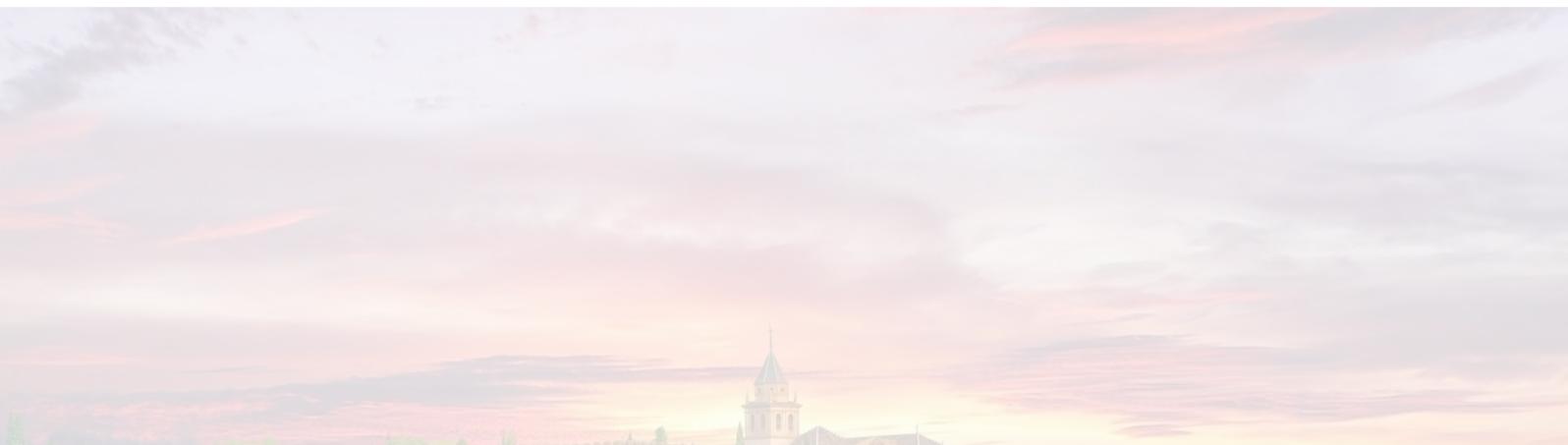
## MULTICOLINEALIDAD EXACTA

- ▶ **Concepto:** La multicolinealidad exacta o perfecta hace referencia a la existencia de una relación lineal exacta entre dos o más variables independientes.
- ▶ **Causas:** Dicho tipo de multicolinealidad se traduce en el incumplimiento de una de las hipótesis básicas del modelo uniecuacional múltiple: la matriz  $X$  no es de rango completo por columnas, esto es,  $rg(X) < k$ .
- ▶ **Consecuencias:** El incumplimiento de dicha hipótesis no permite invertir la matriz  $X^tX$ , por lo que el sistema normal

$$X^tX \cdot \beta = X^ty,$$

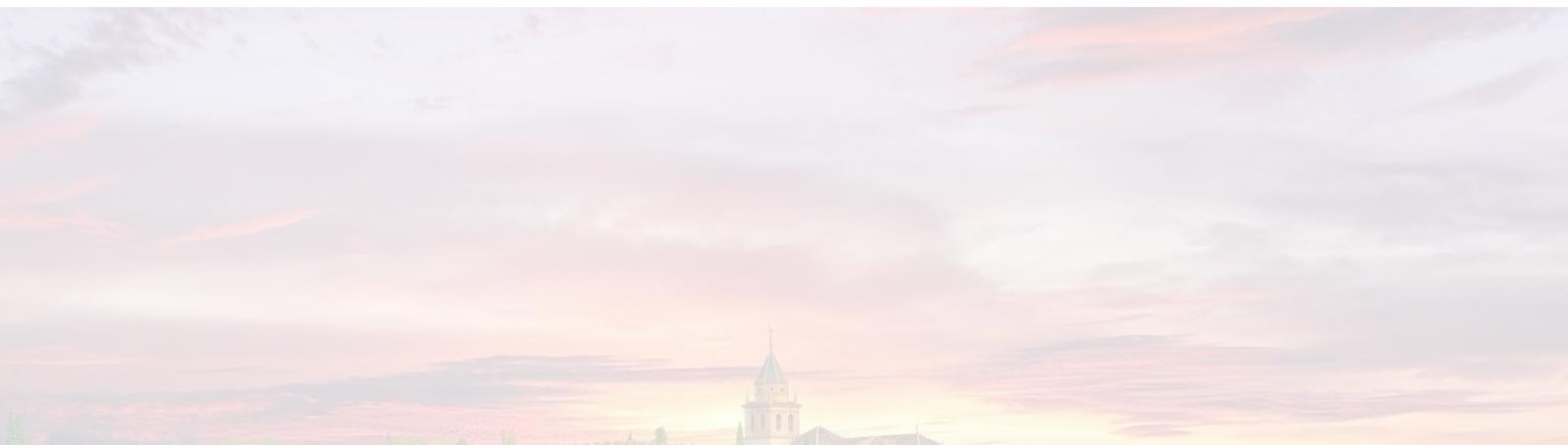
es compatible indeterminado, es decir, es imposible obtener una solución única para  $\hat{\beta}$  (hay infinitas).

- ▶ **Ejemplo:** Un ejemplo sería la trampa de las variables ficticias.



## MULTICOLINEALIDAD APROXIMADA

- ▶ **Concepto:** La multicolinealidad aproximada hace referencia a la existencia de una relación lineal aproximada entre dos o más variables independientes.
- ▶ **Causas:** Las principales causas que producen multicolinealidad aproximada en un modelo son:
  - relación causal entre variables explicativas del modelo.
  - escasa variabilidad en las observaciones de las variables independientes.
  - reducido tamaño de la muestra.



## TIPOS DE MULTICOLINEALIDAD APROXIMADA

En función de las causas se hace la siguiente diferenciación:

► Según Spanos y McGuirk:

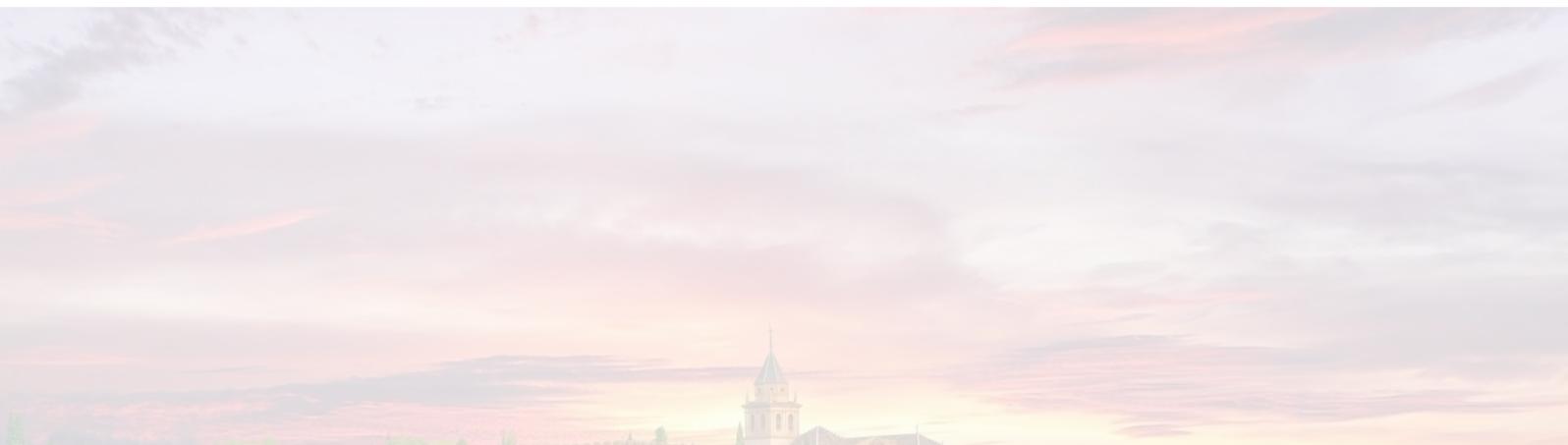
**Multicolinealidad sistemática:** debida a un problema estructural, es decir, a la alta correlación lineal de las variables exógenas consideradas.

**Multicolinelidad errática:** debido a un problema puramente numérico, es decir, a un mal condicionamiento de los datos considerados.

► Mientras que Marquandt y Snee:

**Multicolinealidad no esencial:** relación lineal de las variables exógenas con la constante (es sabido que se solventa centrando las variables).

**Multicolinelidad esencial:** relación lineal entre las variables exógenas (excluida la constante).

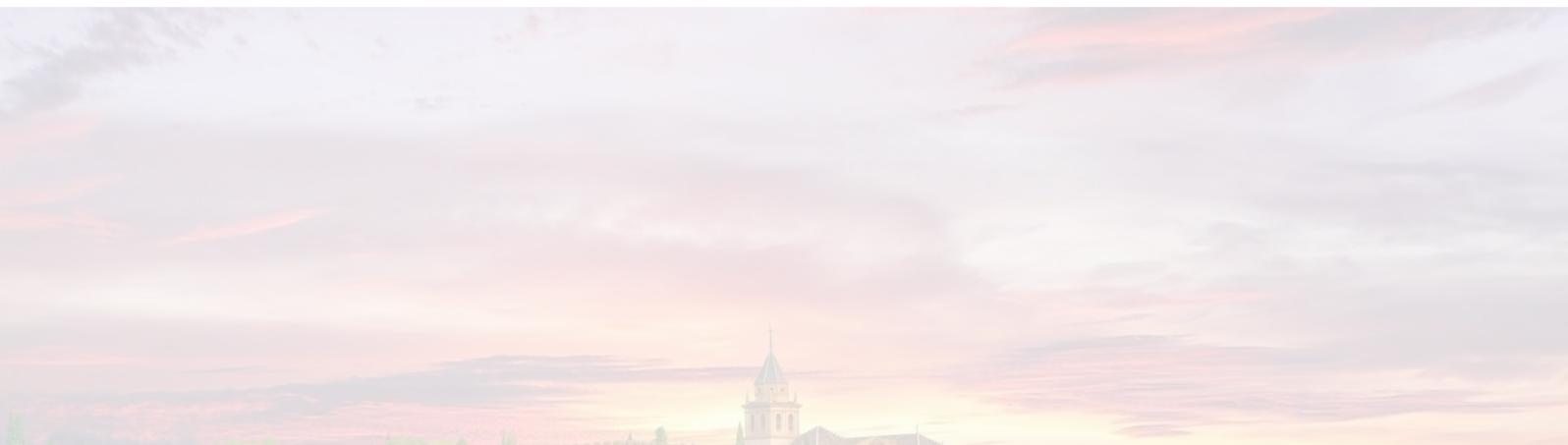


## TIPOS DE MULTICOLINEALIDAD APROXIMADA

Luego, se podrían distinguir los siguientes cuatro casos:

Multicolinealidad	Sistemática	Errática
No esencial	1	2
Esencial	3	4

¿Puede existir multicolinealidad aproximada preocupante en el modelo de regresión lineal simple  $\text{Salario} = \beta_1 + \beta_2 \cdot \text{Experiencia} + u$ ?



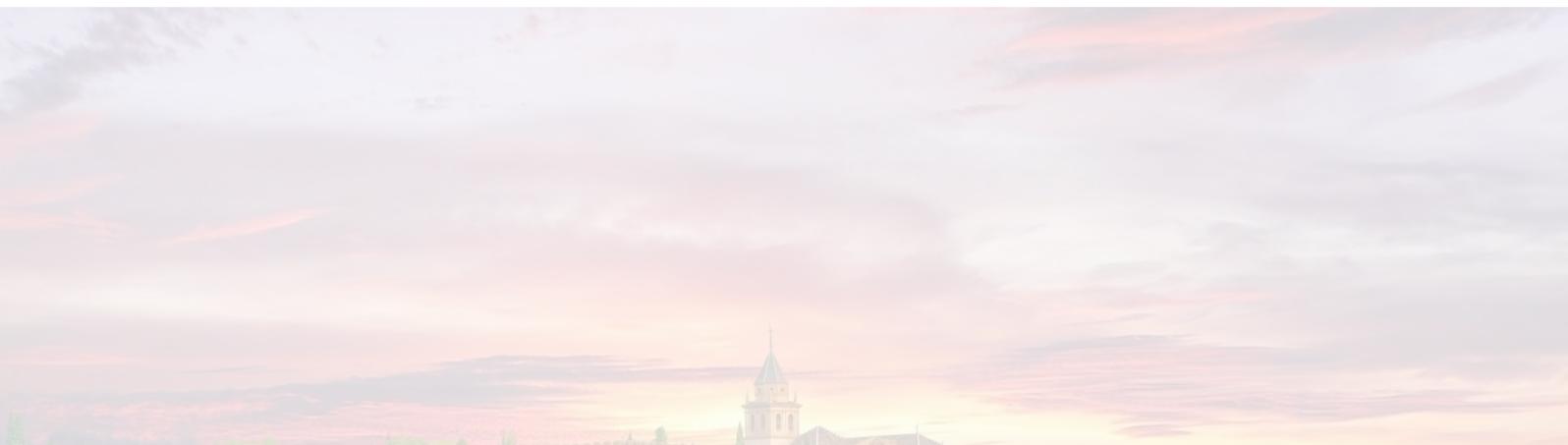
## TIPOS DE MULTICOLINEALIDAD APROXIMADA

Luego, se podrían distinguir los siguientes cuatro casos:

Multicolinealidad	Sistemática	Errática
No esencial	1	2
Esencial	3	4

¿Puede existir multicolinealidad aproximada preocupante en el modelo de regresión lineal simple  $\text{Salario} = \beta_1 + \beta_2 \cdot \text{Experiencia} + u$ ?

¿Y en el modelo  $\text{Salario} = \beta_1 + \beta_2 \cdot \text{Experiencia} + \beta_3 \cdot \text{Edad} + u$ ?



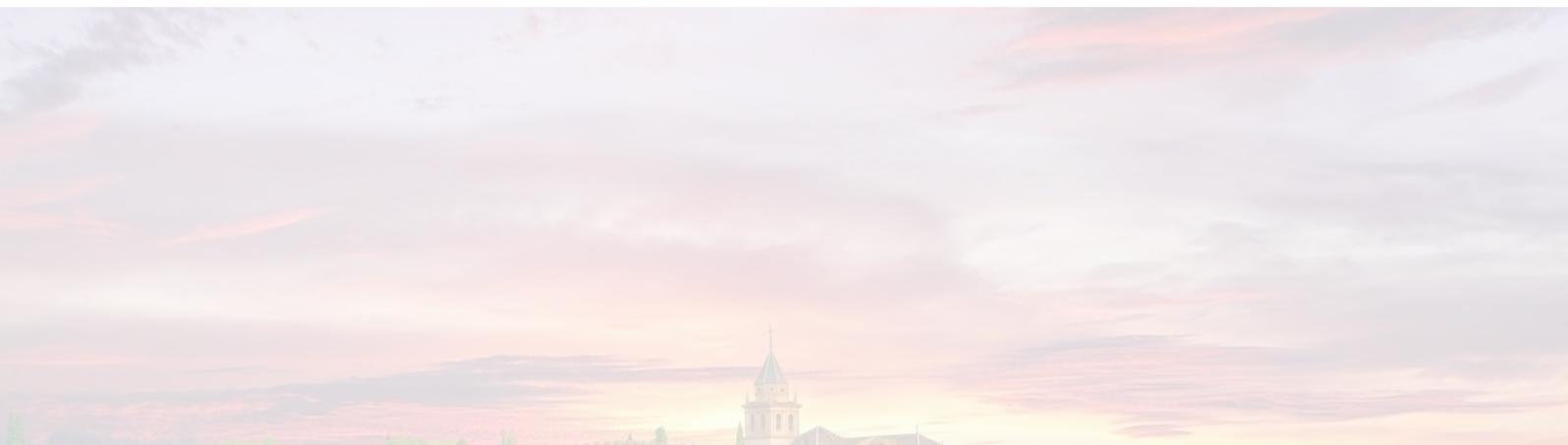
## TIPOS DE MULTICOLINEALIDAD APROXIMADA

Luego, se podrían distinguir los siguientes cuatro casos:

Multicolinealidad	Sistemática	Errática
No esencial	1	2
Esencial	3	4

¿Puede existir multicolinealidad aproximada preocupante en el modelo de regresión lineal simple **Salario** =  $\beta_1 + \beta_2 \cdot \text{Experiencia} + u$ ?

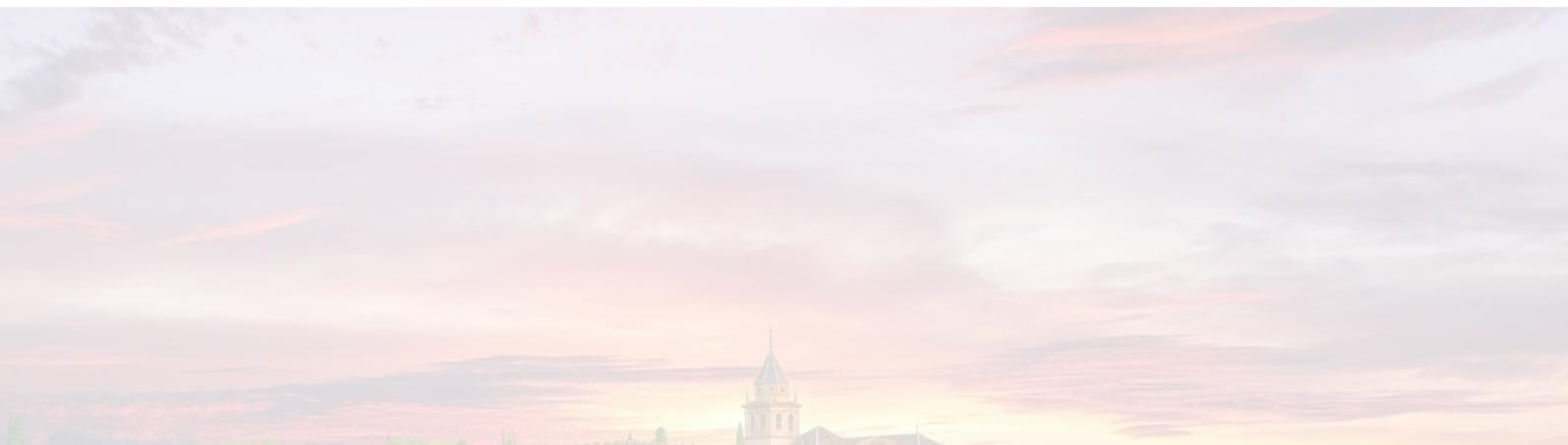
¿Y en el modelo **Salario** =  $\beta_1 + \beta_2 \cdot \text{Experiencia} + \beta_3 \cdot \text{Edad} + u$ ?



# CONCEPTO, CAUSAS Y CONSECUENCIAS

## MULTICOLINEALIDAD APROXIMADA

- ▶ **Consecuencias:** En este caso, no se incumplirá la hipótesis básica de que la matriz  $X$  sea completa por columnas ( $rg(X) = k$ ), por lo que se podrá invertir  $X^tX$  y obtener los estimadores por mínimos cuadrados ordinarios. Sin embargo, el determinante de  $X^tX$  será muy próximo a cero, por lo que  $(X^tX)^{-1}$  tenderá a tener valores altos y por tanto:
  - las varianzas de los estimadores son muy grandes.
  - al efectuar contrastes de significación individual no se rechazará la hipótesis nula, mientras que al realizar contrastes conjuntos si.
  - los coeficientes estimados serán muy sensibles ante pequeños cambios en los datos.
  - un coeficiente de determinación elevado.
- ▶ Ejemplo: La edad y la experiencia suelen presentar una alta relación ya que ambas evolucionan conjuntamente: a mayor edad se presupone mayor experiencia. Por tal motivo será difícil separar el efecto de cada una sobre la variable dependiente y que se produzca multicolinealidad debido a la relación causal existente entre dichas variables



## EFECTOS NOCIVOS SOBRE EL ANÁLISIS ESTADÍSTICO DEL MODELO

Wissel, J. (2009). *A new biased estimator for multivariate regression models with highly collinear variables*. Ph.D. thesis.

Variable	Estimación	Desviación típica
Constante	5.469264	13.016791
Consumo	-4.252429	5.135058
Ingresos	3.120395	2.035671
Crédito Pendiente	0.002879	0.005764
$R^2$		0.9235
$\hat{\sigma}^2$		0.8695563
$F_{3,13}$		52.3

Estimación por MCO del crédito en Estados Unidos

No se rechaza que  $\beta_i = 0, i = 1, 2, 3$ , y se rechaza que  $\beta_2 = \beta_3 = \beta_4 = 0$ .

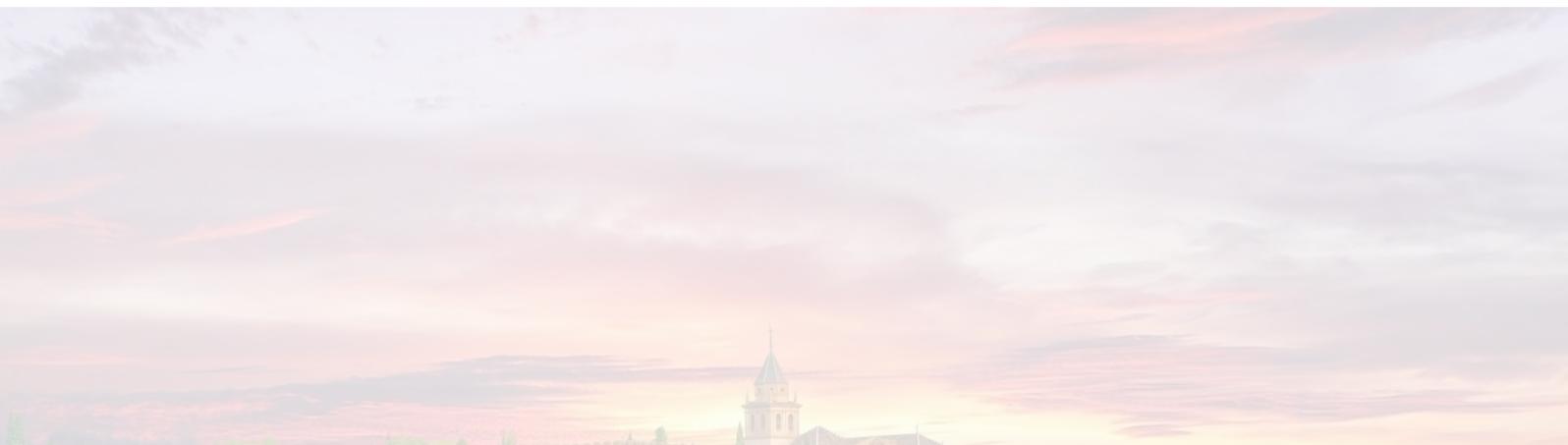
## EFFECTOS NOCIVOS SOBRE EL ANÁLISIS ESTADÍSTICO DEL MODELO

Wooldridge, J.M. (2000). *Introductory Econometrics: A modern approach*. South-Western, Cengage Learning.

Variable	Estimación	Desviación típica
Constante	0.4404	0.09556
Tipos de interés a 3 meses	<b>1.00569</b>	0.01343
$R^2 = 0.9787, \hat{\sigma}^2 = 0.2025, F_{2,122} = 5611$		
Estimación por MCO de los tipos de interés a 12 meses en función de los tipos de interés a 3 meses		

Variable	Estimación	Desviación típica
Constante	0.22471	0.0397
Tipos de interés a 3 meses	<b>-0.62891</b>	0.06582
Tipos de interés a 6 meses	1.59334	0.06394
$R^2 = 0.9965, \hat{\sigma}^2 = 0.03330625, F_{2,122} = 17371.66$		
Estimación por MCO de los tipos de interés a 12 meses en función de los tipos de interés a 3 y 6 meses		

Correlación entre los tipos de interés a 3 y 6 meses: 0.9893021.



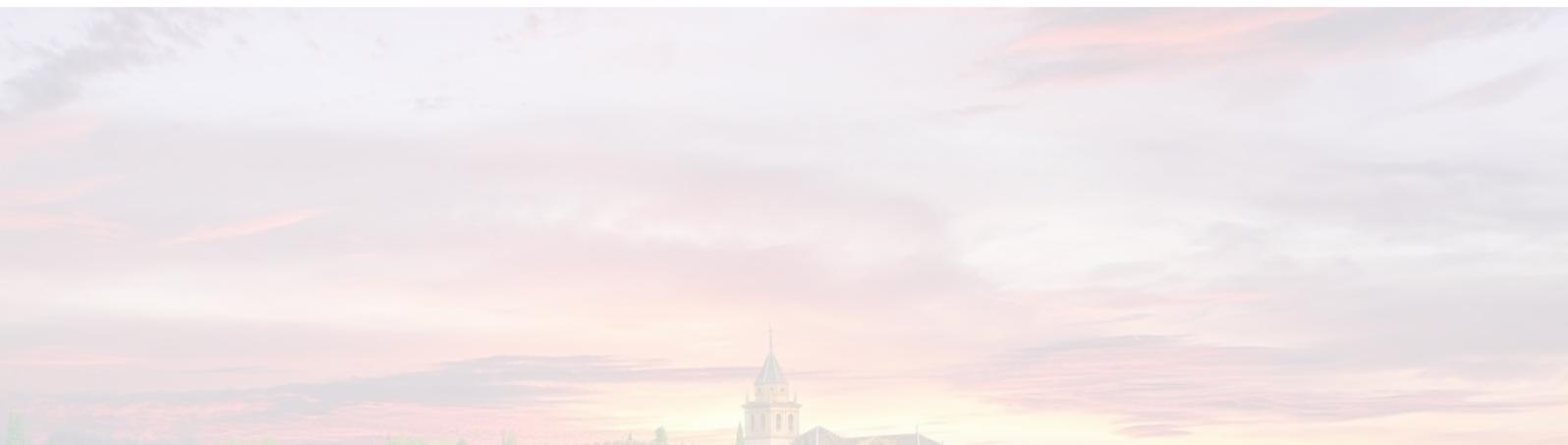
# EFECTOS NOCIVOS SOBRE EL ANÁLISIS NUMÉRICO DEL MODELO

## Wissel

C	C.p	I	I.p	CP	CP.p	$\beta$	$\beta_p$
4.7703	4.838464	4.8786	4.992323	808.23	839.7796	5.46926428	-12.844487481
4.7784	4.910650	5.0510	5.136480	798.03	789.1984	-4.252429358	3.324325322
4.9348	5.038435	5.3620	5.385545	806.12	836.3163	3.120395253	0.316260872
5.0998	5.044642	5.5585	5.635094	865.65	871.7818	0.002879118	-0.002397538
5.2907	5.350734	5.8425	5.838306	997.30	1006.9633		
5.4335	5.501856	6.1523	6.256959	1140.70	1148.8869		
⋮	⋮	⋮	⋮	⋮	⋮		

## Wooldridge

r3	r3.p	r6	r6.p	$\beta$	$\beta_p$
2.77	2.784788	3.02	3.027994	0.2247061	0.1845118
2.97	3.000442	3.43	3.463405	-0.6289143	-0.5209103
4.00	4.168316	4.32	4.423468	1.5933374	1.4875326
4.60	4.628050	4.68	4.721110		
4.16	4.148983	4.33	4.338835		
3.07	3.166463	3.50	3.588747		
⋮	⋮	⋮	⋮		



## 2.1. FACTOR DE INFLACIÓN DE LA VARIANZA

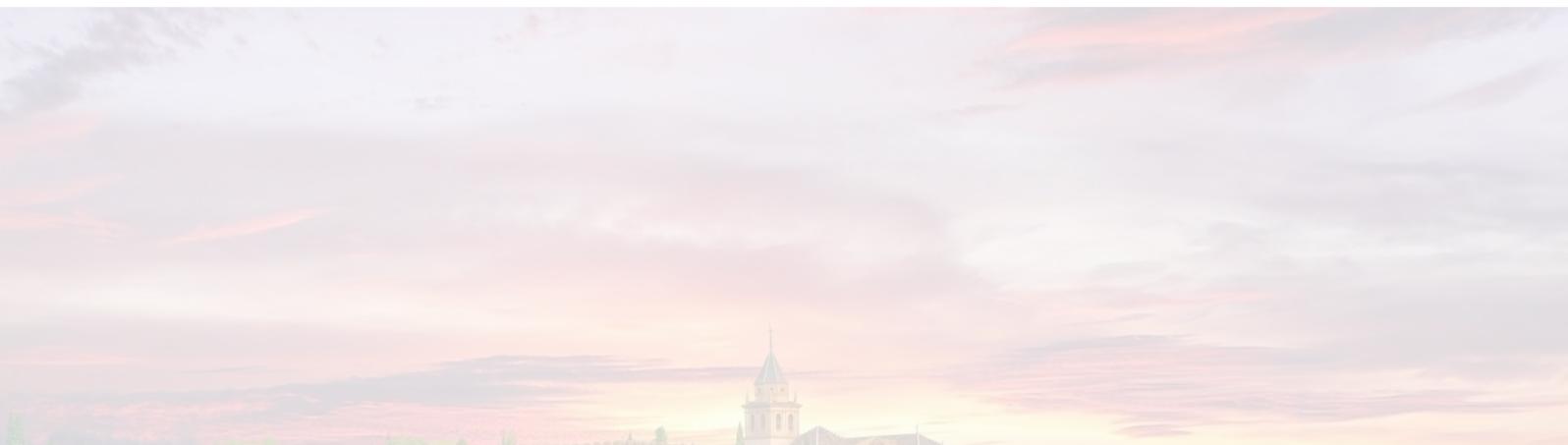
Una de las medidas más usadas para detectar el grado de multicolinealidad existente es el Factor de Inflación de la Varianza (FIV) dado por:

$$FIV(i) = \frac{1}{1 - R_i^2}, \quad i = 2, \dots, p, \quad (1)$$

donde  $R_i^2$  es el coeficiente de determinación de la regresión de  $X_i$  sobre el resto de variables independientes,  $X_{-i}$ .

Si esta medida es superior a 10 se supone que el grado de multicolinealidad presente en el modelo es preocupante.

El VIF no tiene en cuenta la relación de las variables exógenas del modelo,  $X_2 \dots X_p$ , con la constante,  $1$  Por tanto, no detecta la multicolinealidad no esencial.



## 2.2. NÚMERO DE CONDICIÓN

Otra medida muy extendida es el Número de Condición (NC), el cual viene dado por:

$$NC = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}, \quad (2)$$

donde  $\lambda_{max}$  y  $\lambda_{min}$  son, respectivamente, los autovalores máximo y mínimo de  $\tilde{\mathbf{X}}^t \tilde{\mathbf{X}}$  donde:

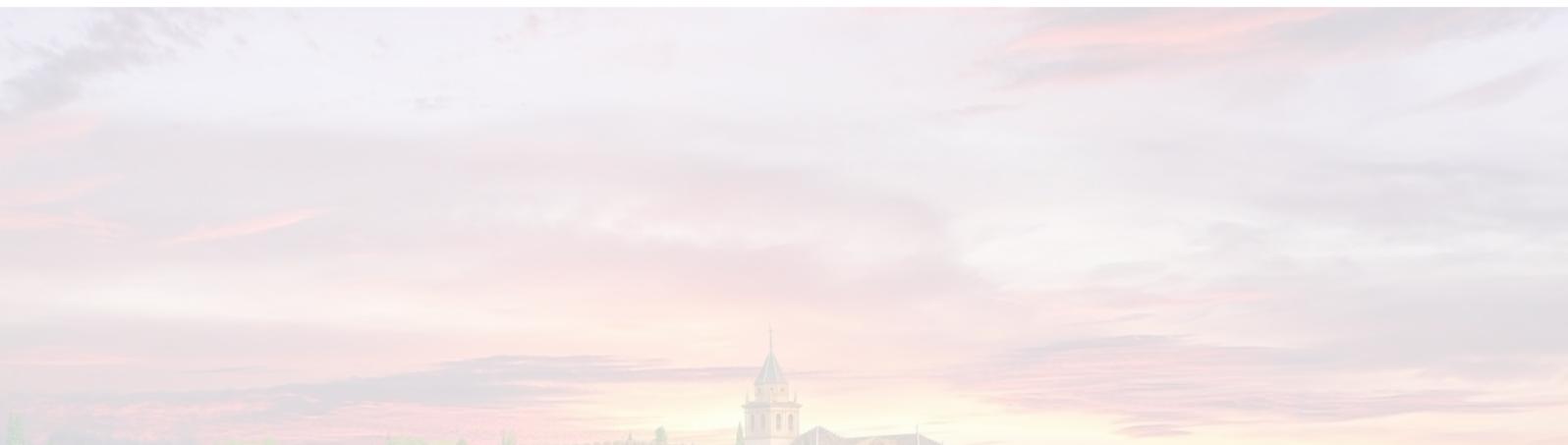
$$\tilde{\mathbf{X}} = [\tilde{\mathbf{1}} \tilde{\mathbf{X}}_2 \dots \tilde{\mathbf{X}}_p],$$
$$\tilde{\mathbf{1}} = \frac{\mathbf{1}}{\sqrt{n}}, \quad \tilde{\mathbf{X}}_i = \frac{\mathbf{X}_i}{\sqrt{\sum_{j=1}^n X_{ji}^2}}, \quad i = 2, \dots, p.$$

Si esta medida es superior a 20 se supone que el grado de multicolinealidad presente en el modelo es moderado y si es superior a 30 preocupante.

El NC tiene en cuenta la relación de las variables exógenas del modelo,  $\mathbf{X}_2 \dots \mathbf{X}_p$ , con la constante,  $\mathbf{1}$ .

## 2.3. OTRAS MEDIDAS

- ▶ **R**: Matriz de correlaciones simples de las variables independientes del modelo: Ignora por completo a la constante y proporciona información de las relaciones dos a dos, por lo tanto, sólo es capaz de detectar multicolinealidad aproximada del tipo esencial dos a dos.
- ▶  $\det(\mathbf{R})$ : Determinante de la matriz de correlaciones: Recoge estructuras más complejas aunque sigue ignorando la relación con la constante, por lo que detecta multicolinealidad aproxima del tipo esencial. En este caso, valores próximos a cero indica que este problema es grave.
- ▶  $CV(X_i)$ : Coeficiente de variación de las variables explicativas: Puede detectar una escasa variabilidad de las variables lo que podría provocar multicolinealidad aproximada de tipo no esencial.

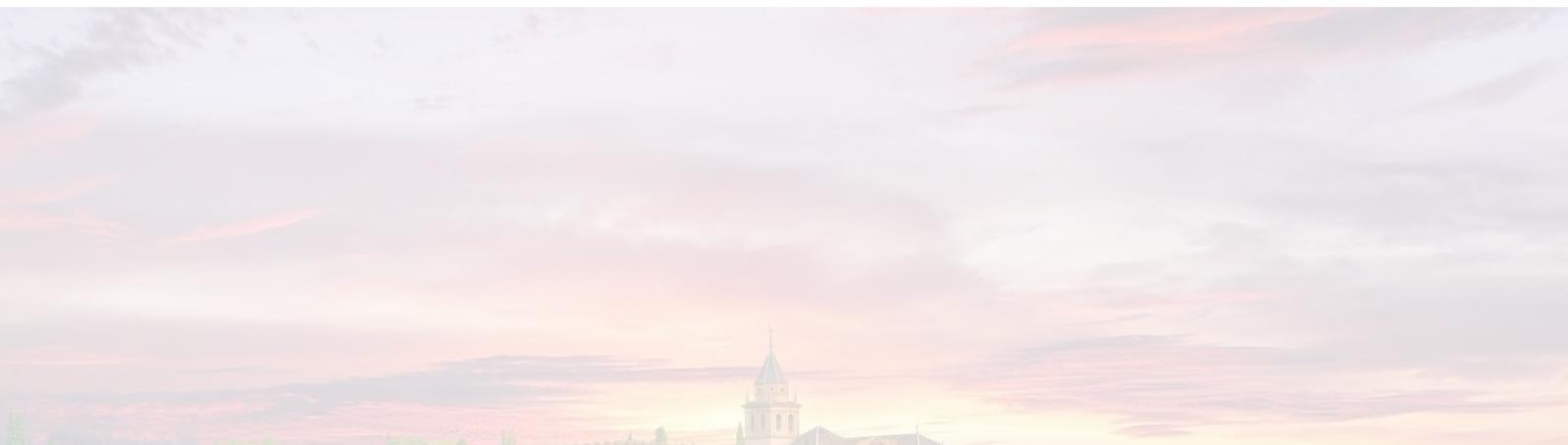


### 3. SOLUCIONES

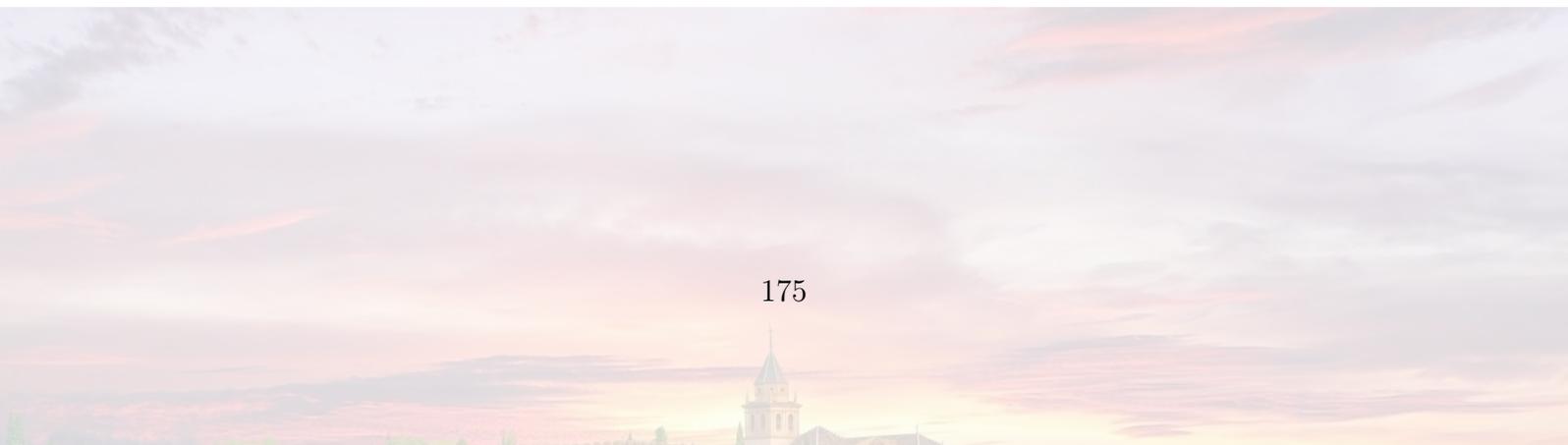
Algunas de las posibles soluciones al problema de multicolinealidad son las siguientes:

- ▶ mejora del diseño muestral extrayendo la información máxima de la variables observadas.
- ▶ eliminación de las variables que se sospechan son causantes de la multicolinealidad.
- ▶ en caso de disponer de pocas observaciones, aumentar el tamaño de la muestra.
- ▶ utilizar la relación extramuestral que permita realizar relaciones entre los parámetros (información a priori) que permita estimar el modelo por mínimos cuadrados restringidos.

Por otro lado, algunos autores sugieren tratar el problema de la multicolinealidad de forma mecánica y puramente numérica proponiendo métodos de estimación alternativos: estimación cresta, estimación alzada o residualización (entre otras).



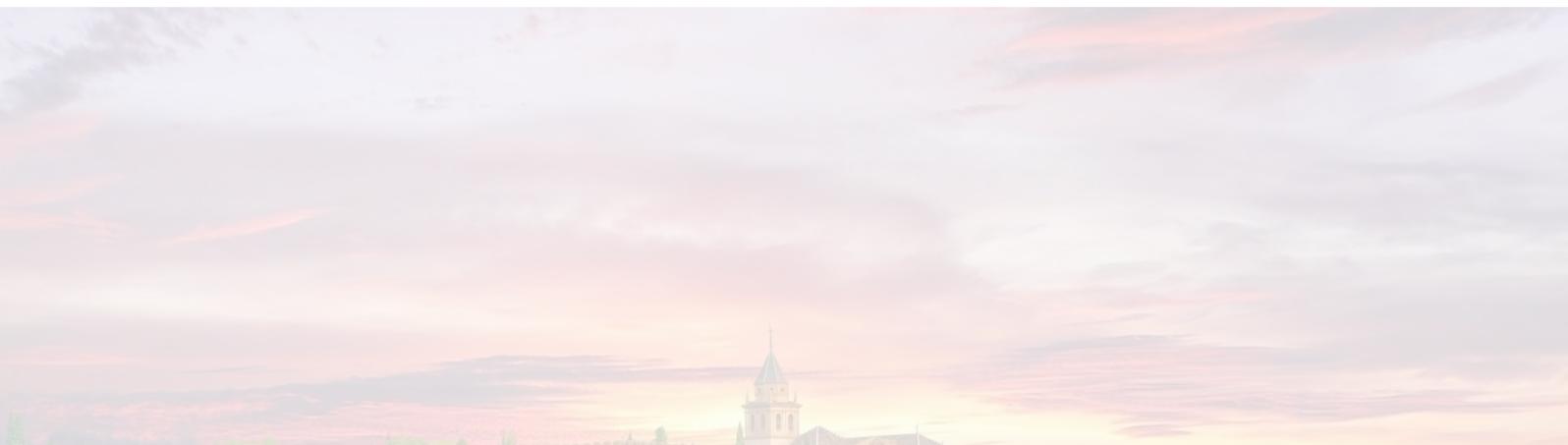
## 5 Tema 5



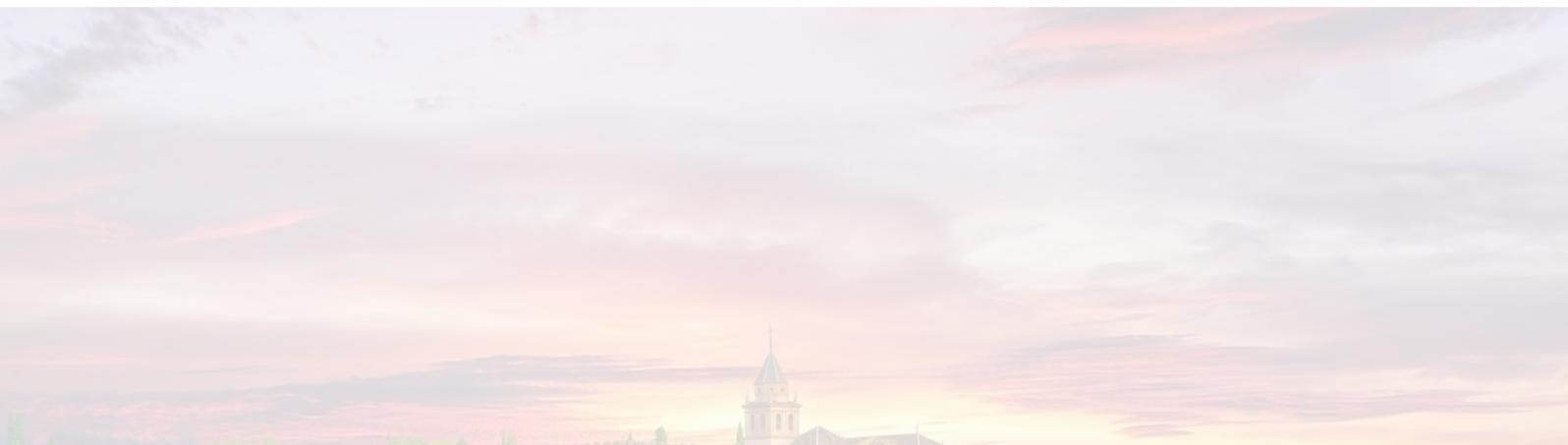
# ECONOMETRÍA

## TEMA 5: HETEROCEDASTICIDAD

2024-2025



<b>1</b>	<b>Concepto, causas y consecuencias</b>	<b>2</b>
1.1	Concepto	2
1.2	Causas	4
1.3	Consecuencias	7
<b>2</b>	<b>Procedimientos de Detección</b>	<b>8</b>
2.1	Métodos gráficos	9
2.2	Contrastes	11
<b>3</b>	<b>Estimación bajo heterocedasticidad</b>	<b>20</b>



## 1.1. CONCEPTO

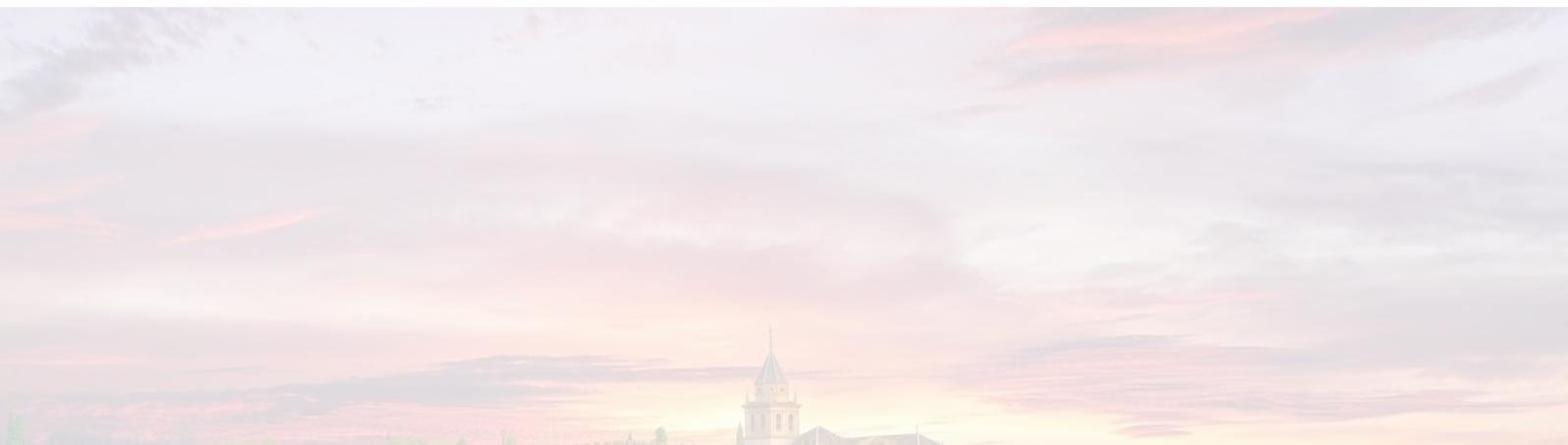
$$Y = X\beta + u$$

Hasta ahora suponíamos:

1.  $\mathbb{E}[u] = 0$ .
2.  $\text{var}[u_i] = \sigma^2, \forall i \in 1, 2, \dots, n$ .
3.  $\text{Cov}[u_i u_j] = 0, \forall i, j \in 1, 2, \dots, n$  con  $i \neq j$ .

Con lo que teníamos:

$$\text{var}[u] = \mathbb{E}[uu^t] = \sigma^2 I_n$$



## 1.1. CONCEPTO

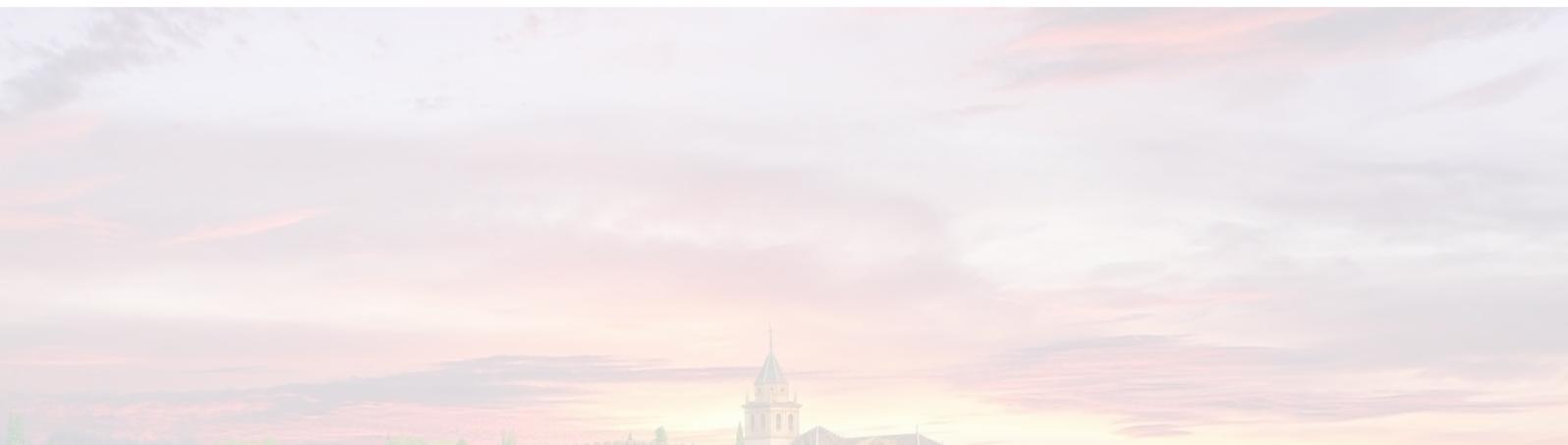
$$Y = X\beta + u$$

Hasta ahora suponíamos:

1.  $\mathbb{E}[u] = 0$ .
2.  $\text{var}[u_i] = \sigma^2, \forall i \in 1, 2, \dots, n$ .
3.  $\text{Cov}[u_i u_j] = 0, \forall i, j \in 1, 2, \dots, n$  con  $i \neq j$ .

Con lo que teníamos:

$$\text{var}[u] = \mathbb{E}[uu^t] = \sigma^2 I_n$$



## 1.1. CONCEPTO

$$Y = X\beta + u$$

Hasta ahora suponíamos:

1.  $\mathbb{E}[u] = 0$ .
2.  $\text{var}[u_i] = \sigma^2, \forall i \in 1, 2, \dots, n$ .
3.  $\text{Cov}[u_i u_j] = 0, \forall i, j \in 1, 2, \dots, n$  con  $i \neq j$ .

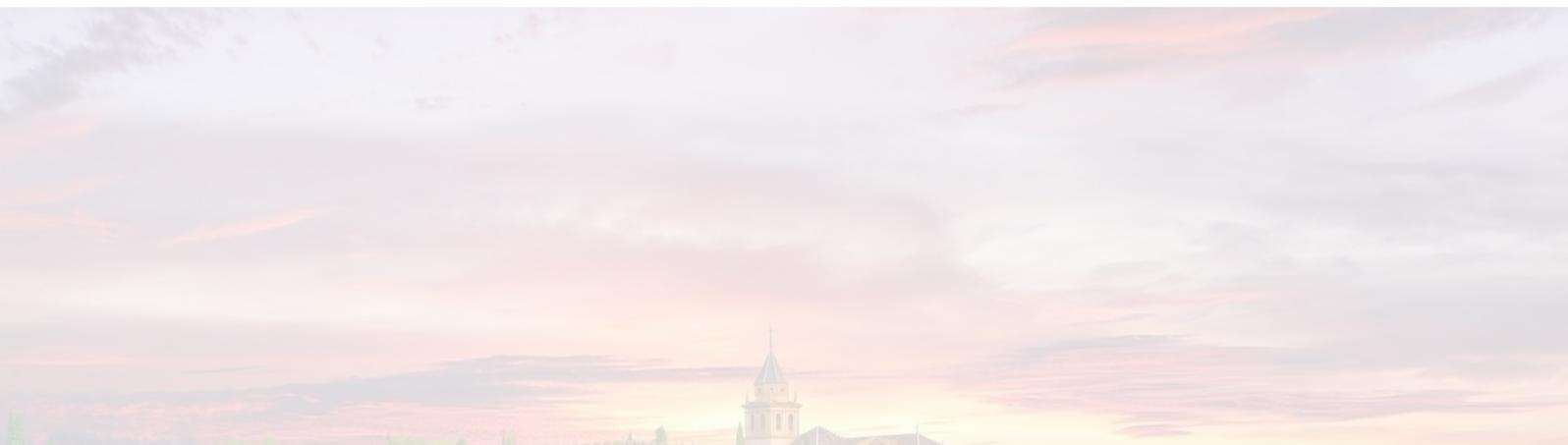
Con lo que teníamos:

$$\text{var}[u] = \mathbb{E}[uu^t] = \sigma^2 I_n$$

¿Y si la matriz de varianzas-covarianzas del vector de perturbaciones pudiera adoptar cualquier forma?

$$\text{VAR}[u] = \sigma^2 \Omega$$

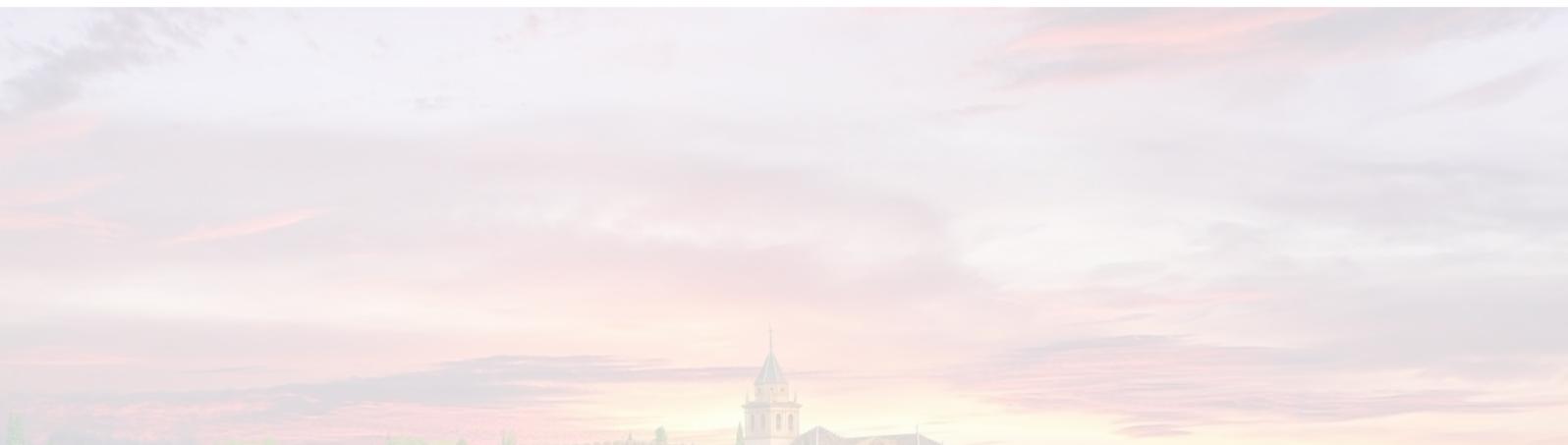
donde  $\Omega$  es una matriz simétrica y definida positiva.



## 1. CONCEPTO

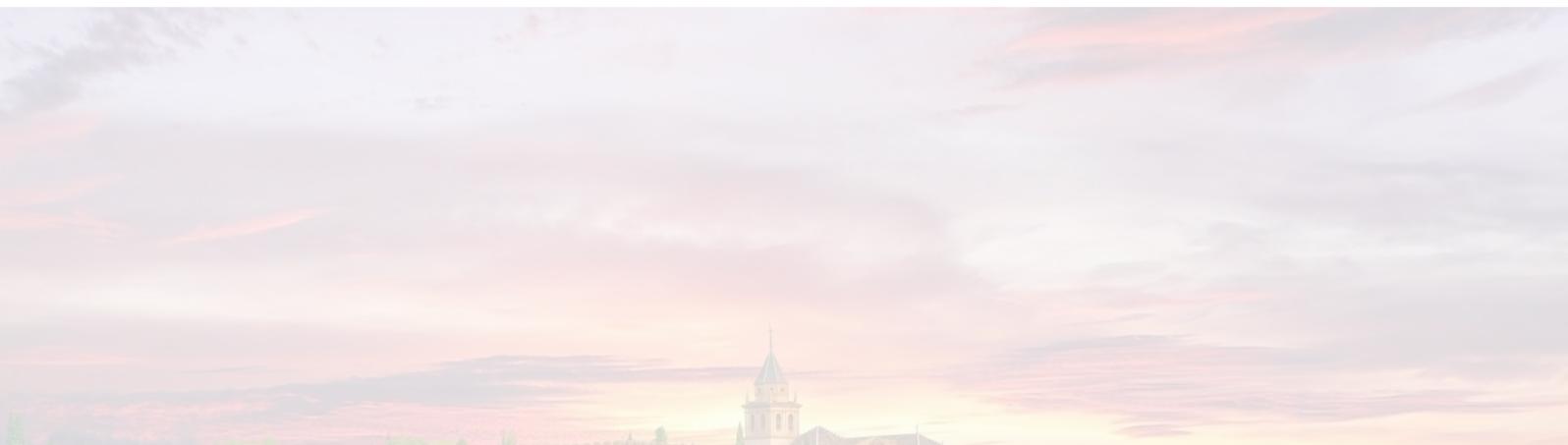
Se dice que un modelo de regresión presenta **heteroscedasticidad** cuando la varianza del término de perturbación no permanece constante a lo largo del tiempo.

$$\sigma^2\Omega = \begin{pmatrix} \sigma_1^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & 0 & \cdots & 0 \\ & & \ddots & & \\ 0 & 0 & \cdots & 0 & \sigma_n^2 \end{pmatrix}$$



## 1.2.CAUSAS DE LA HETEROSCEDASTICIDAD

- ▶ En datos de sección cruzada, donde la escala de la variable dependiente y el poder explicativo de la tendencia del modelo varía a lo largo de las observaciones.
- ▶ Cuando en el modelo se trabaja con datos agrupados, es decir, que las observaciones pudieran agruparse en  $m$  categorías, cada una de ellas con  $n_j$  observaciones,  $j = 1, 2, \dots, m$ . Una vez realizada la agrupación, en lugar de trabajar con los datos originales, los cuales son homoscedásticos, utilizamos las medias aritméticas de cada una de las categorías, las cuales resultan ser heteroscedásticas.
- ▶ Omisión de una variable relevante en la expresión del modelo de regresión. En esta situación es lógico pensar que el comportamiento de la perturbación dependa de dicha variable omitida, provocando que su varianza sea variable.



## 1.2.CAUSAS DE LA HETEROSCEDASTICIDAD: EJEMPLO

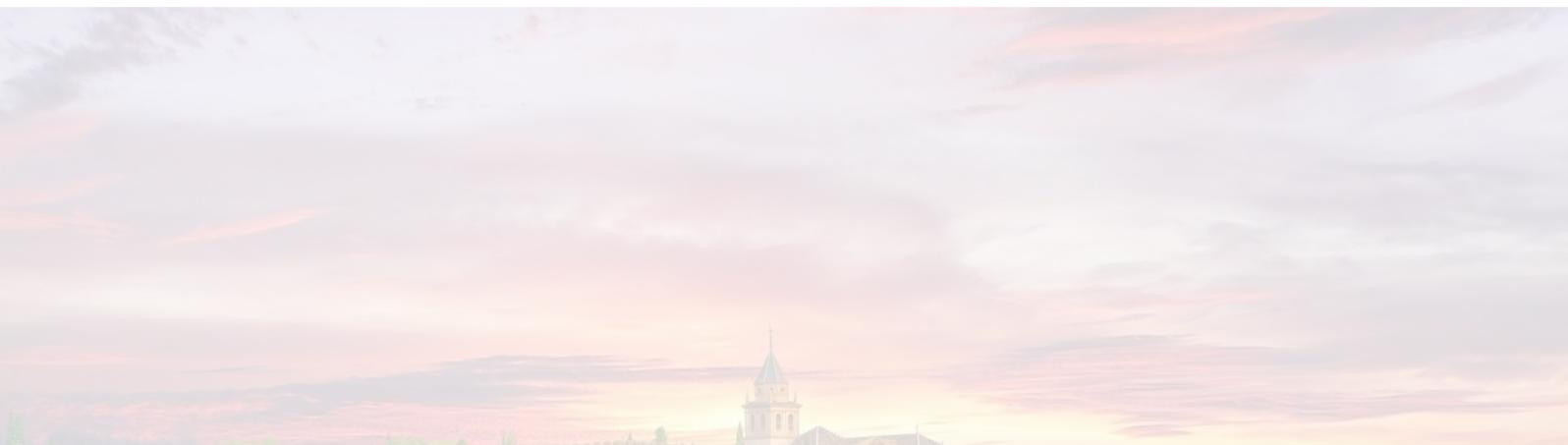
Supongamos que se quiere estudiar el consumo familiar en función de la renta familiar mediante el modelo de regresión

$$C_i = \alpha + \beta R_i + u_i \text{ con } i = 1, 2, \dots, n$$

donde se verifica que

$$\mathbb{E}[uu^t] = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix}$$

El hecho de que la varianza del término de perturbación sea distinta nos da la idea de que la dispersión que existe en los valores del consumo familiar entre los distintos hogares es distinta.



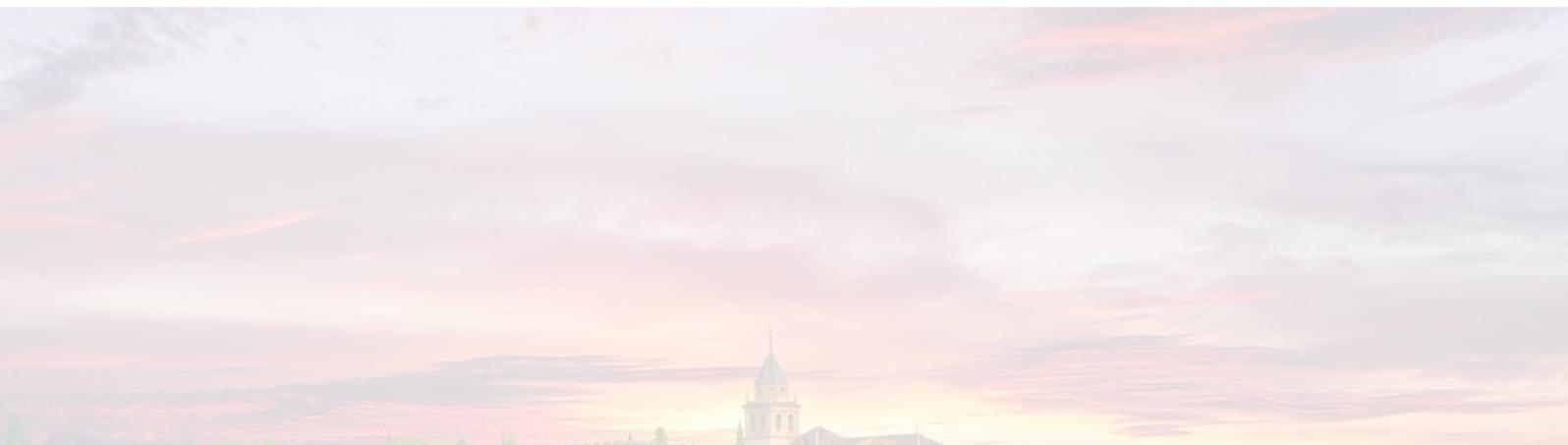
## 1.2.CAUSAS DE LA HETEROSCEDASTICIDAD: EJEMPLO

Supongamos que en lugar de considerar el modelo

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t,$$

se especifica este otro

$$Y_t = \alpha_1 + \alpha_2 X_{2t} + v_t.$$



## 1.2.CAUSAS DE LA HETEROSCEDASTICIDAD: EJEMPLO

Supongamos que en lugar de considerar el modelo

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t,$$

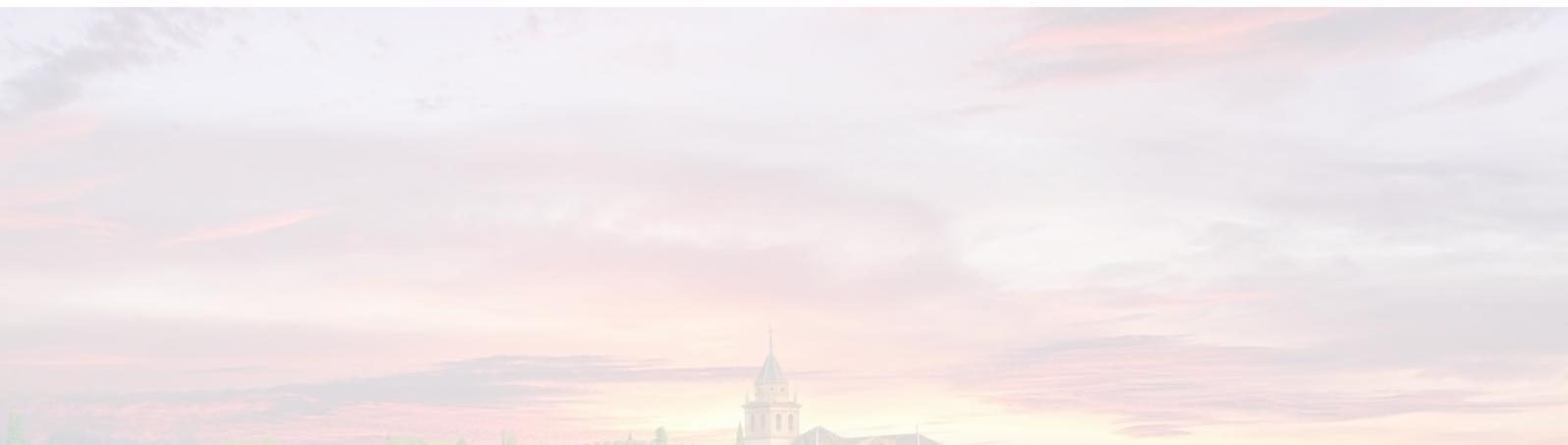
se especifica este otro

$$Y_t = \alpha_1 + \alpha_2 X_{2t} + v_t.$$

Si la variable omitida es realmente relevante, la perturbación aleatoria  $v_t$  dependerá de  $X_{3t}$  por lo que se puede establecer que  $v_t = u_t + \beta_3 X_{3t}$ . Y en tal caso:

$$\begin{aligned} \text{Var}(v_t) &= \mathbb{E}[v_t^2] = \mathbb{E}[u_t^2 + \beta_3^2 X_{3t}^2 + 2\beta_3 X_{3t} u_t] \\ &= \mathbb{E}[u_t^2] + \beta_3^2 X_{3t}^2 + 2\beta_3 X_{3t} \mathbb{E}[u_t] = \sigma^2 + \beta_3^2 X_{3t}^2 \neq \sigma^2. \end{aligned}$$

Como es evidente, la varianza de la perturbación aleatoria no es constante sino que varía con cada observación.



### 1.3. CONSECUENCIAS DE LA HETEROSCEDASTICIDAD

Puesto que en el método de estimación por MCO no influye la matriz de varianzas-covarianzas de la perturbación aleatoria es claro que el estimador por MCO será igualmente:

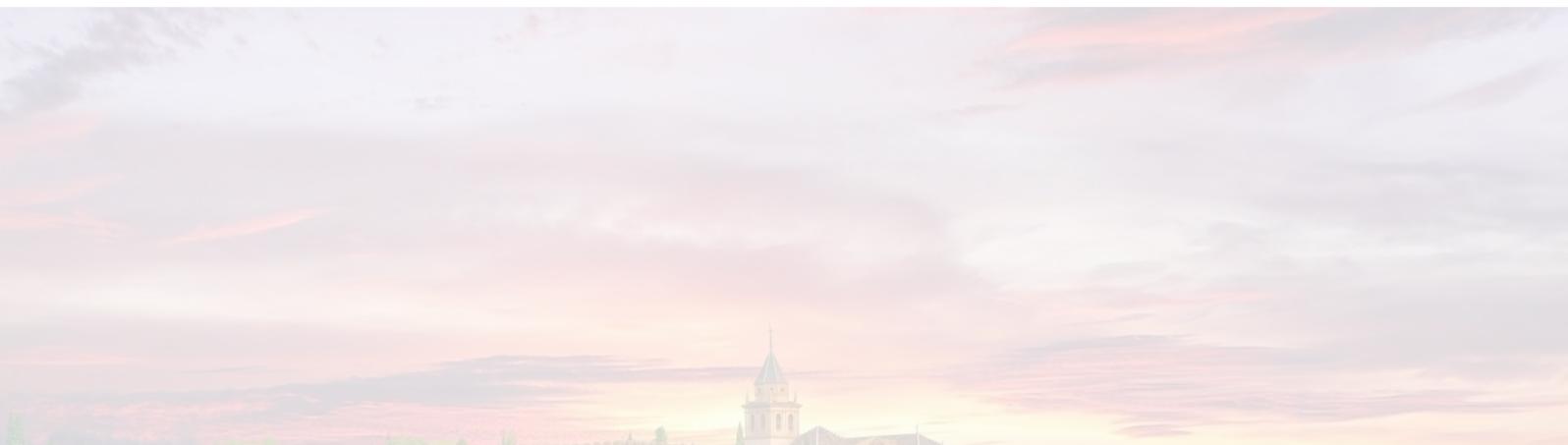
$$\hat{\beta}_{MCO} = (X^t X)^{-1} X^t y.$$

$\hat{\beta}_{MCO}$  es lineal e insesgado. Sin embargo, ya no se tiene asegurado que la varianza sea mínima:

$$\begin{aligned} \text{Var}(\hat{\beta}_{MCO}) &= (X^t X)^{-1} X^t \mathbb{E}[u \cdot u^t] X (X^t X)^{-1} \\ &= \sigma^2 (X^t X)^{-1} X^t \Omega X (X^t X)^{-1}, \end{aligned}$$

distinta a la del modelo con perturbaciones esféricas:  $\sigma^2 (X^t X)^{-1}$ .

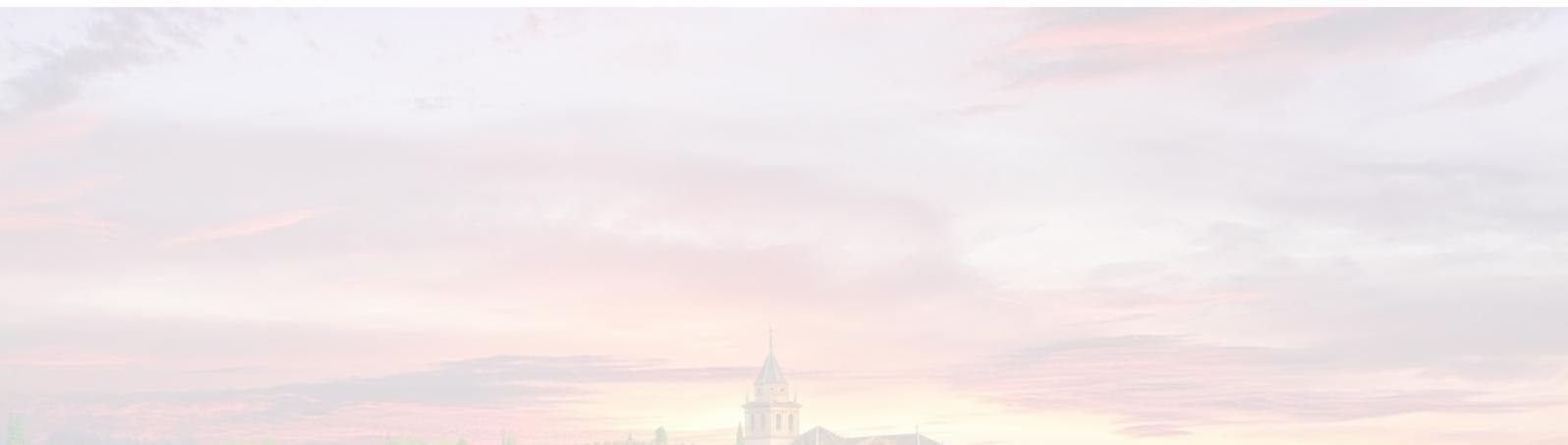
Por tanto, la consecuencia de la presencia de heteroscedasticidad en un modelo lineal es que los estimadores obtenidos, aunque serán lineales e insesgados, no serán óptimos.



## 2. PROCEDIMIENTOS DE DETECCIÓN

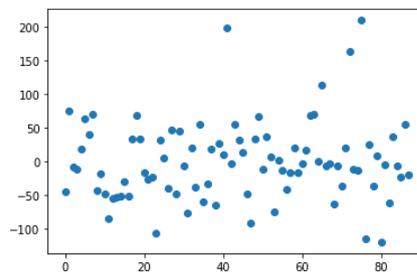
Para detectar la heteroscedasticidad en un modelo lineal múltiple disponemos de distintos procedimientos.

- ▶ **Métodos gráficos** a partir de los cuales intentaremos intuir cuáles son las variables que provocan la existencia de heteroscedasticidad en el modelo. Concretamente, estudiaremos los gráficos de los residuos y de dispersión.
- ▶ **Métodos analíticos** para determinar la presencia de heteroscedasticidad en el modelo: Tests de Glesjer, Goldfeld-Quandt, Breusch-Pagan y White. Los dos primeros se deben usar cuando la muestra es pequeña y una variable es la causa de la heteroscedasticidad, mientras que los otros dos cuando la muestra es grande y no se sabe la o las variables que provocan el problema. Además, la hipótesis nula de todos estos contrastes es siempre que el modelo es homocedástico.



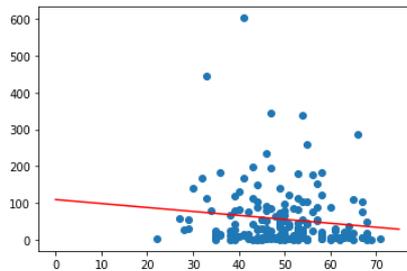
## MÉTODOS GRÁFICOS

- ▶ **Gráfico de los residuos:** es un gráfico de dispersión de los residuos o residuos al cuadrado,  $e_t$  ó  $e_t^2$ , frente a  $t$ . Si en dichos gráficos observamos grupos de observaciones con distinta varianza, podemos pensar en la presencia de heteroscedasticidad.



## MÉTODOS GRÁFICOS

- ▶ **Gráficos de dispersión:** consiste en el diagrama de dispersión de los residuos o residuos al cuadrado,  $e_t$  ó  $e_t^2$ , frente a la variable independiente que sospechamos que puede causar la heteroscedasticidad. Si la variabilidad de los residuos aumenta o disminuye conforme aumenta el valor de la variable independiente, entonces podemos pensar que la varianza de la perturbación aleatoria depende de dicha variable y, por tanto, habría presencia de heteroscedasticidad.



## CONTRASTE DE GOLDFELD QUANDT

Para muestras pequeñas, y si sospechamos que  $\sigma_t^2$  está relacionada positivamente con la variable  $X_i$ :

1. Ordenar la información muestral, según los valores crecientes de este regresor, de menor a mayor.
2. Omitir los  $m$  valores muestrales centrales, (por ejemplo  $m = \frac{n}{3}$ ).
3. Ajustar, mediante MCO, cada uno de los subgrupos obtenidos en el paso 2, una vez eliminados los valores centrales. Cada uno de estos subgrupos estarán formados por  $\frac{n-m}{2}$  observaciones. La estimación del modelo del primer subgrupo nos permitirá obtener la  $SCR_1$  y la del subgrupo 2, la  $SCR_2$ .
4. Bajo ausencia de heteroscedasticidad ( $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ):

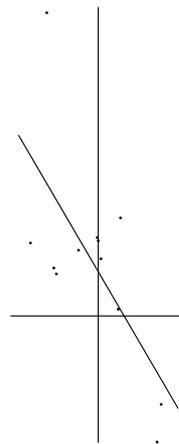
$$F_{exp} = \frac{SCR_2}{SCR_1} > F_{\frac{n-m}{2}-k, \frac{n-m}{2}-k, 1-\alpha}$$

se rechaza la hipótesis de homocedasticidad (hay heteroscedasticidad)

`sms.het_goldfeldquandt(mco.resid, mco.model.exog)`

## CONTRASTE DE GQ

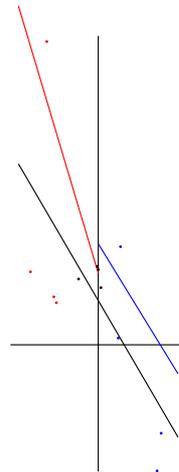
$y$	$x$
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198



$$\text{MCO: } y = 2.7941 - 1.7259x$$

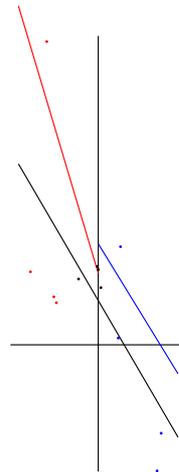
## EJEMPLO

$y$	$x$
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198

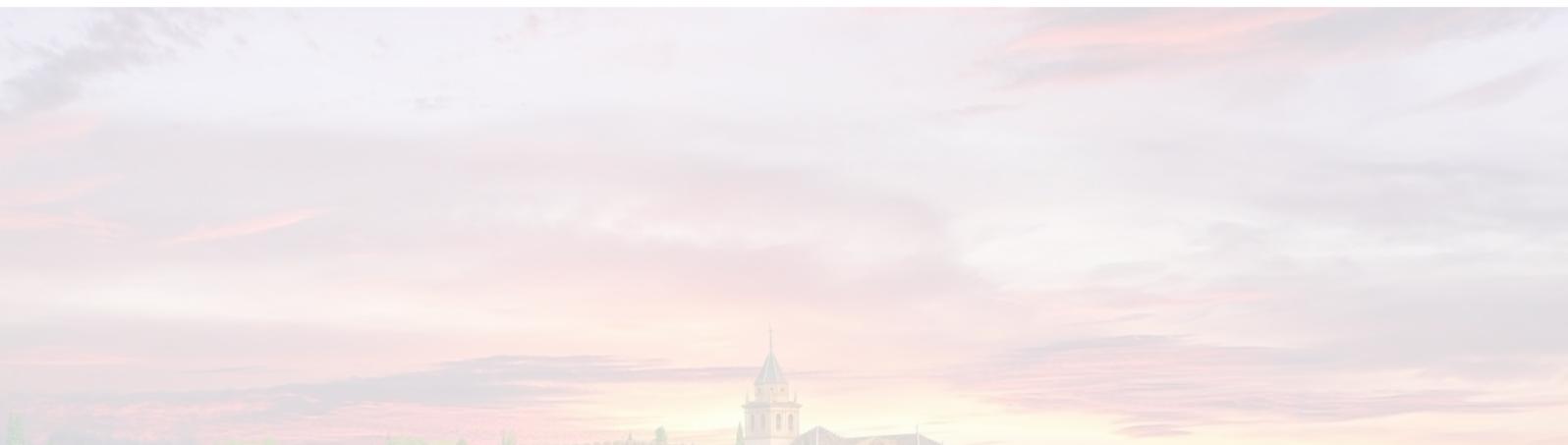


## EJEMPLO

$y$	$x$
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198

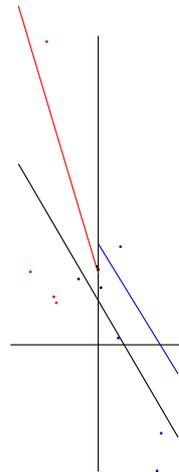


MCO1:  $y = 4.5144 - 3.3759x$ ,  $SCR_1 = 0.1918$ ,  $e^t = (-0.3415, 0.1379, 0.2349, -0.0313)$



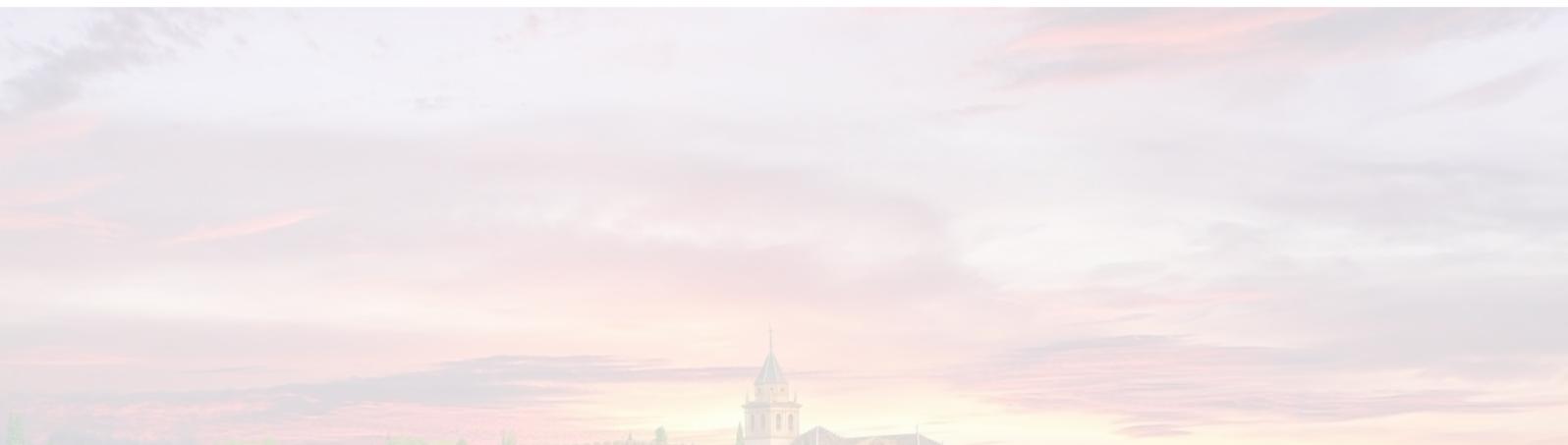
## EJEMPLO

$y$	$x$
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198



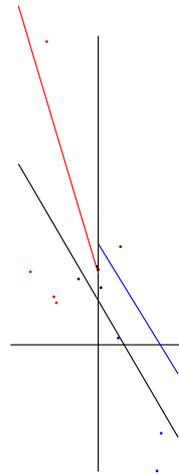
MCO1:  $y = 4.5144 - 3.3759x$ ,  $SCR_1 = 0.1918$ ,  $e^t = (-0.3415, 0.1379, 0.2349, -0.0313)$

MCO2:  $y = 6.380 - 1.642x$ ,  $SCR_2 = 127.0477$ ,  $e^t = (-1.5382, 7.4685, 2.1003, -8.0306)$



## EJEMPLO

$y$	$x$
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198



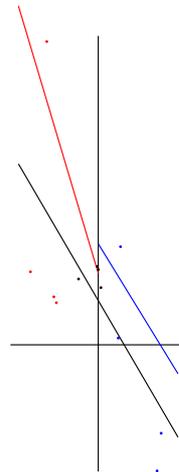
MCO1:  $y = 4.5144 - 3.3759x$ ,  $SCR_1 = 0.1918$ ,  $e^t = (-0.3415, 0.1379, 0.2349, -0.0313)$

MCO2:  $y = 6.380 - 1.642x$ ,  $SCR_2 = 127.0477$ ,  $e^t = (-1.5382, 7.4685, 2.1003, -8.0306)$

$$F_{exp} = \frac{SCR_2}{SCR_1} = 662.2987 > F_{2,2,0.975} = 39$$

## EJEMPLO

$y$	$x$
4.608895	-4.246639
19.1437	-3.225103
3.028715	-2.780972
2.654891	-2.622295
4.155246	-1.234524
4.948342	-0.06467424
4.750159	-0.0002376094
3.608508	0.1671859
0.4248052	1.252264
6.1918	1.394248
-7.961979	3.686417
-5.582933	3.942198



MCO1:  $y = 4.5144 - 3.3759x$ ,  $SCR_1 = 0.1918$ ,  $e^t = (-0.3415, 0.1379, 0.2349, -0.0313)$

MCO2:  $y = 6.380 - 1.642x$ ,  $SCR_2 = 127.0477$ ,  $e^t = (-1.5382, 7.4685, 2.1003, -8.0306)$

$F_{exp} = \frac{SCR_2}{SCR_1} = 662.2987 > F_{2,2,0.975} = 39$ : Se rechaza la hipótesis de homocedasticidad!! (hay heterocedasticidad).

## EJEMPLO

Utilizando una muestra de 25 observaciones anuales se estima el siguiente modelo de demanda  $D_t = \beta_1 + \beta_2 Y_t + \beta_3 PR_t + u_t$ .

Utilizando sólo las 10 primeras observaciones se obtiene la siguiente ecuación estimada:

$$\begin{array}{rcccl} \hat{D}_t & = & 80.50 & + & 0.93Y_t & - & 0.87PR_t \\ (d.t.estimadas) & & (86.17) & & (1.06) & & (1.9) \end{array}$$

donde se ha obtenido  $SCR_1 = 125.7$ . Del mismo modo, y utilizando las 10 últimas observaciones, se obtiene la siguiente ecuación estimada

$$\begin{array}{rcccl} \hat{D}_t & = & 20.61 & + & 0.53Y_t & - & 0.105PR_t \\ (d.t.estimadas) & & (221.44) & & (0.29) & & (2.41) \end{array}$$

con  $SCR_2 = 498.94$ . Detectar mediante el test de Goldfeld y Quandt la existencia o no de heteroscedasticidad. Hipótesis nula del contraste es  $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$ .

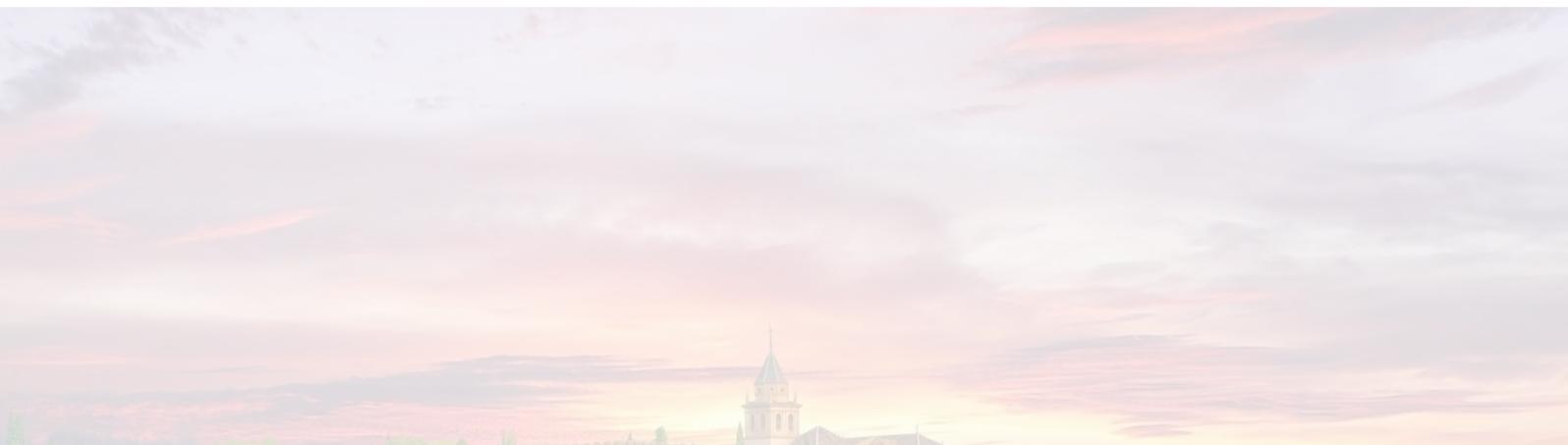
## EJEMPLO

En esta situación el estadístico de contraste es:

$$F_{exp} = \frac{SCR_2}{SCR_1} = \frac{498.94}{125.7} = 3.9693$$

y se verifica que  $F_{exp} \sim F_{\frac{n-m}{2}-k, \frac{n-m}{2}-k} = F_{\frac{25-5}{2}-3, \frac{25-5}{2}-3} = F_{7,7}$  ( $m = 5, k = 3$ ).

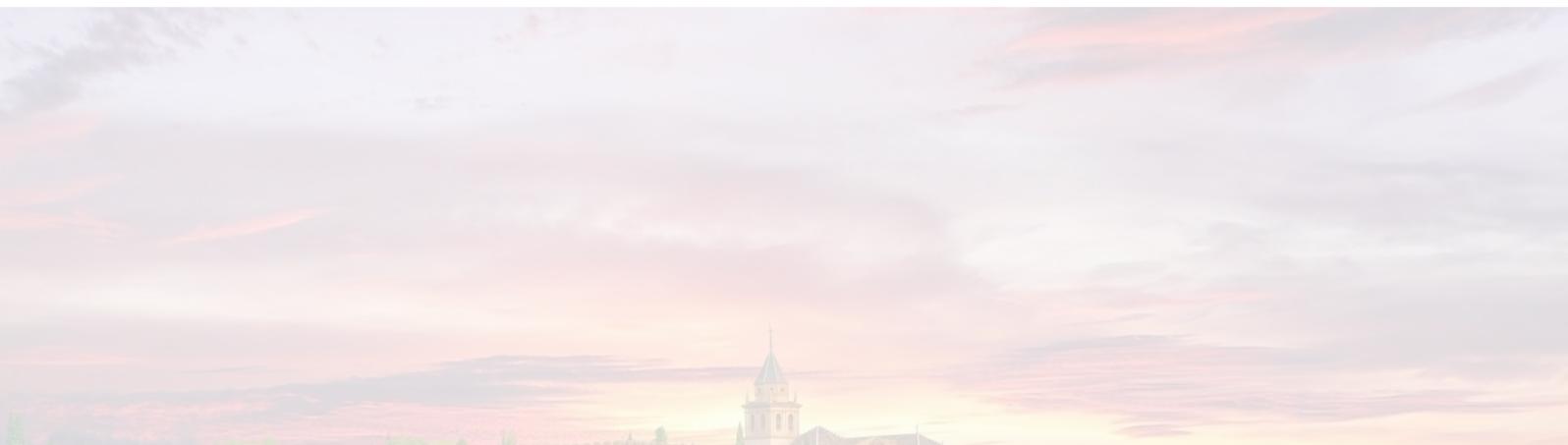
$F_{exp} > 3.79 \Rightarrow$  se rechaza la hipótesis nula a un nivel de significación del 5%.  
Por tanto, concluimos que existe heteroscedasticidad.



## CONTRASTES DE BREUSCH-PAGAN Y WHITE

Para número elevado de observaciones:

1. Estimamos el modelo  $y = X\beta + u$  y el vector correspondiente a los residuos de mínimos cuadrados ordinarios,  $e = y - \hat{y}$ .



## CONTRASTES DE BREUSCH-PAGAN Y WHITE

Para número elevado de observaciones:

1. Estimamos el modelo  $y = X\beta + u$  y el vector correspondiente a los residuos de mínimos cuadrados ordinarios,  $e = y - \hat{y}$ .

2. Breusch-Pagan:

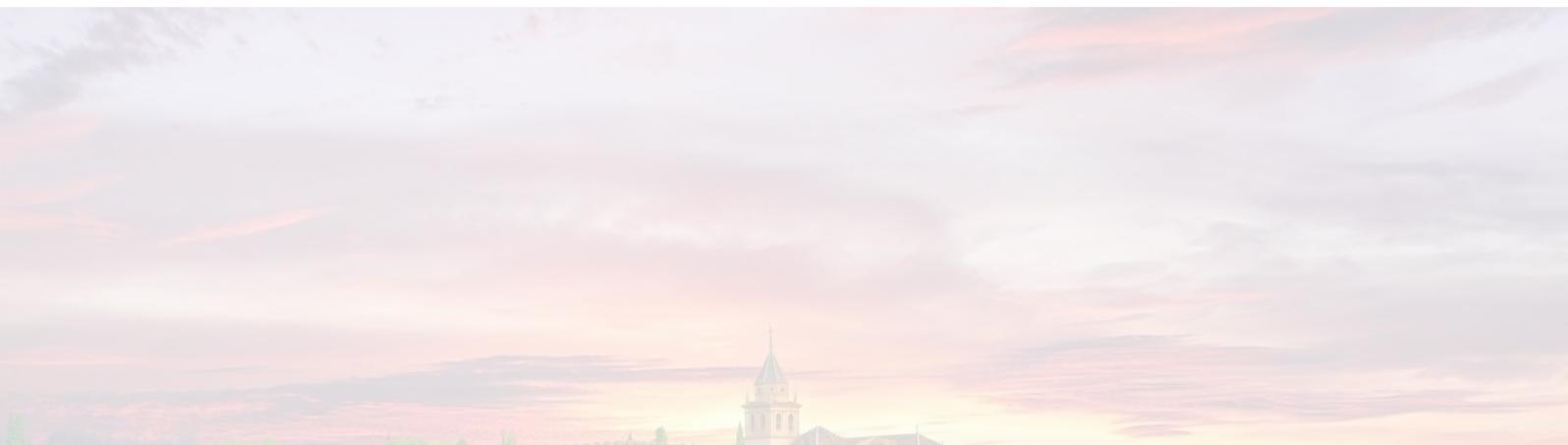
$$e^2 = \delta_0 + \delta_1 x_1 + \dots + \delta_p X_p + v$$

3. White:

$$e^2 = \delta_0 + \delta_1 x_1 + \dots + \delta_p X_p + \delta_{p+1} x_1^2 + \delta_{p+2} x_1 x_2 + \dots + \delta_{p+p^2} x_p^2 + v$$

4. Se contrasta  $H_0 : \delta_1 = \delta_2 = \dots = 0$  (HOMOCEDASTICIDAD)

5. Si se rechaza  $H_0$ : Existe heterocedasticidad!



## CONTRASTES DE BREUSCH-PAGAN Y WHITE

Para número elevado de observaciones:

1. Estimamos el modelo  $y = X\beta + u$  y el vector correspondiente a los residuos de mínimos cuadrados ordinarios,  $e = y - \hat{y}$ .

2. Breusch-Pagan:

$$e^2 = \delta_0 + \delta_1 x_1 + \dots + \delta_p X_p + v$$

3. White:

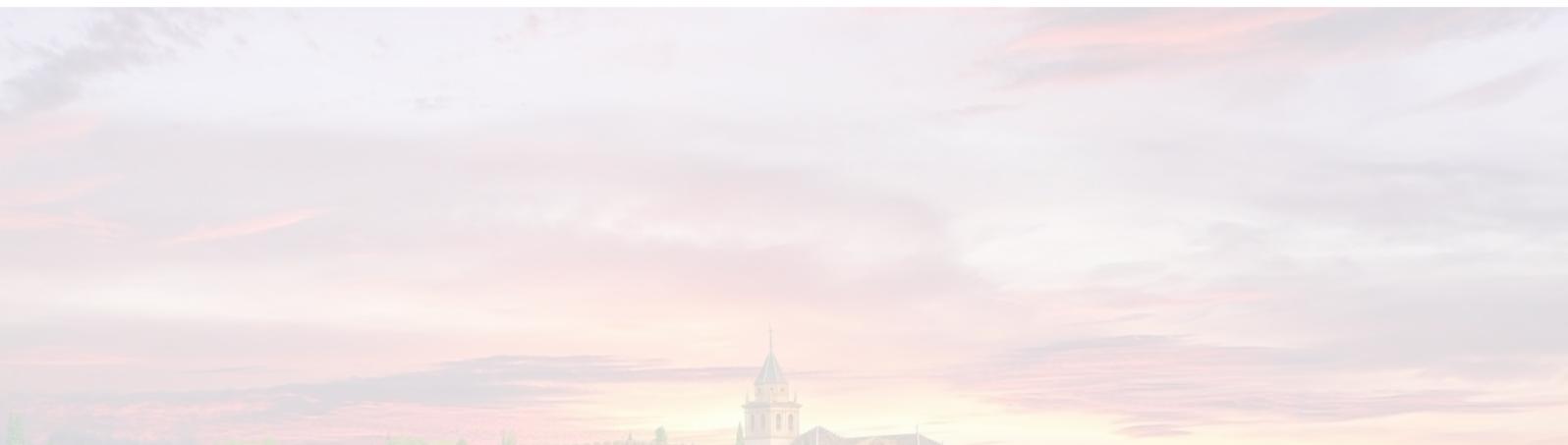
$$e^2 = \delta_0 + \delta_1 x_1 + \dots + \delta_p X_p + \delta_{p+1} x_1^2 + \delta_{p+2} x_1 x_2 + \dots + \delta_{p+p^2} x_p^2 + v$$

4. Se contrasta  $H_0 : \delta_1 = \delta_2 = \dots = 0$  (HOMOCEDASTICIDAD)

5. Si se rechaza  $H_0$ : Existe heterocedasticidad!

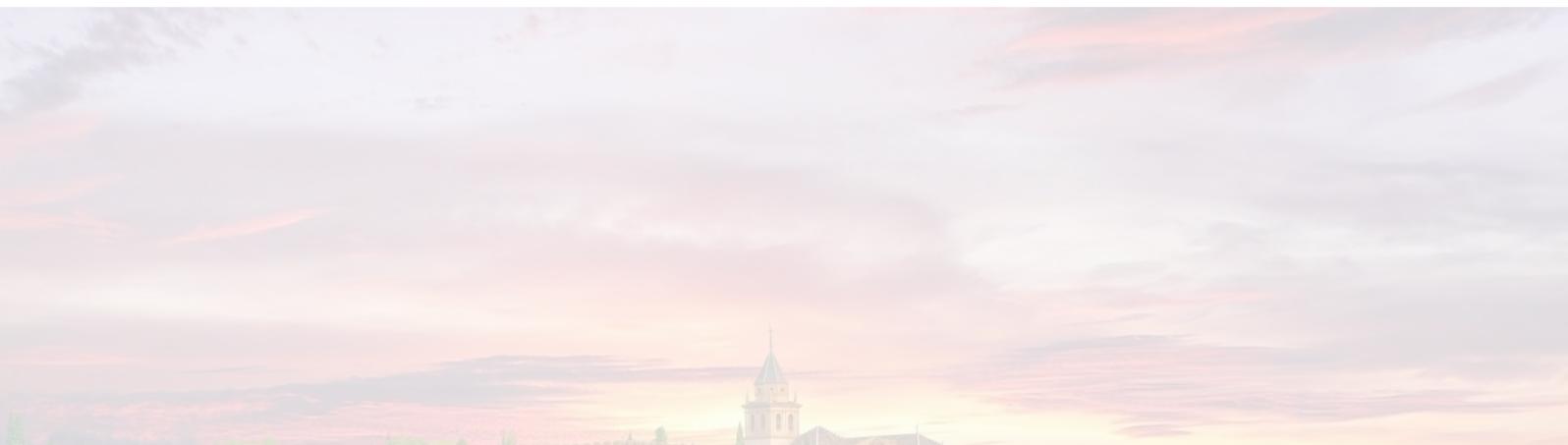
```
sms.het_breuschpagan(mco.resid, mco.model.exog)
```

```
sms.het_white(mco.resid, mco.model.exog)
```



## CONTRASTE DE GLEJSER

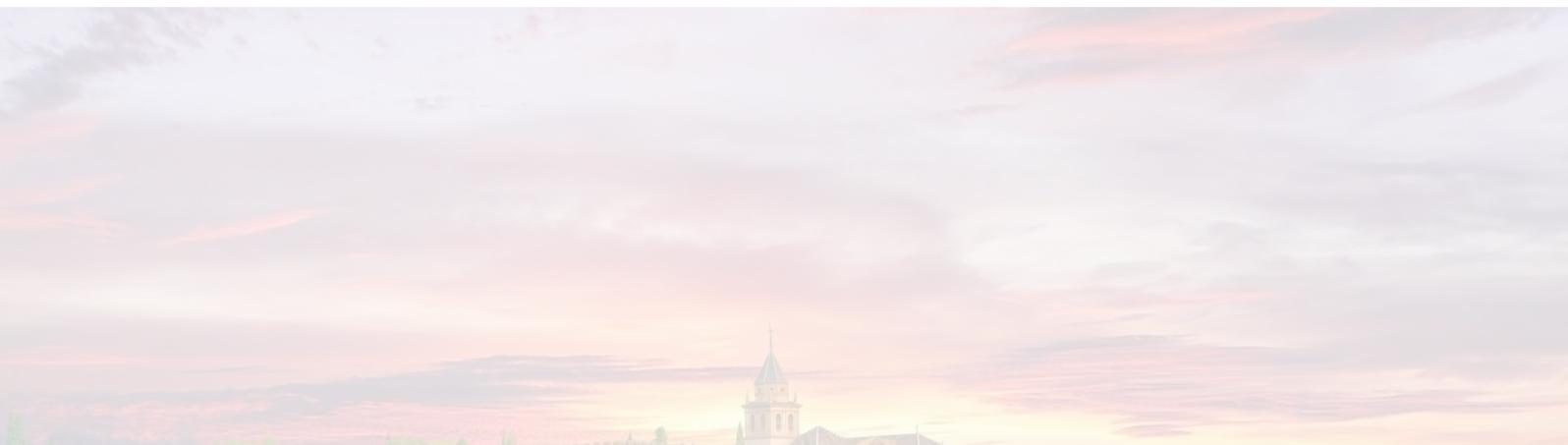
Si sospechamos de que la heteroscedasticidad está ligada a  $X_i$ :



## CONTRASTE DE GLEJSER

Si sospechamos de que la heteroscedasticidad está ligada a  $X_i$ :

1. Estimamos el modelo  $y = X\beta + u$  y  $e = y - \hat{y}$ .



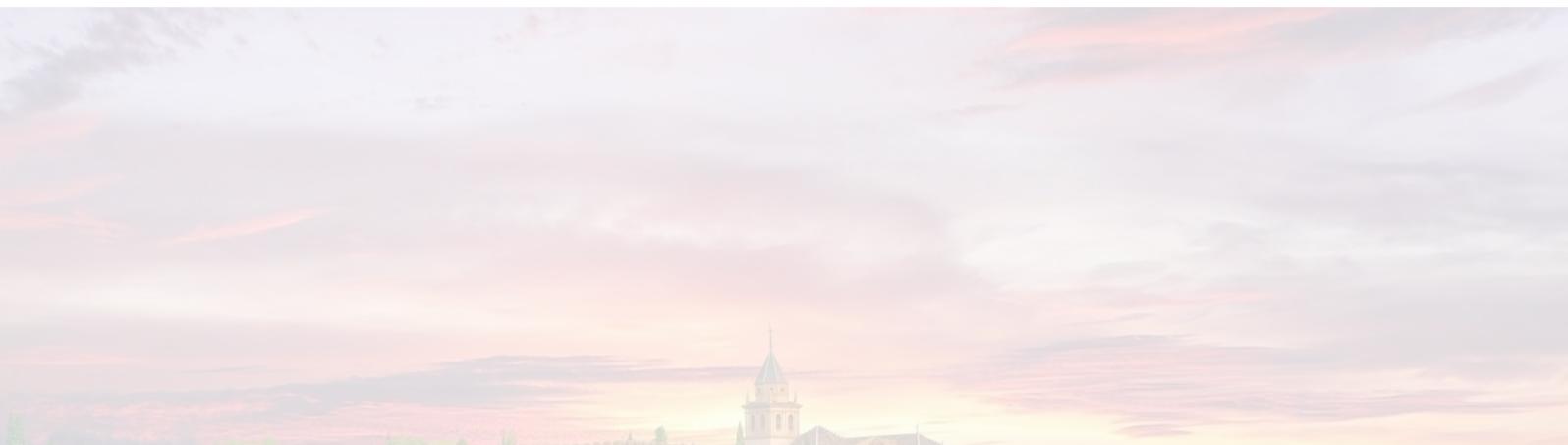
## CONTRASTE DE GLEJSER

Si sospechamos de que la heteroscedasticidad está ligada a  $X_i$ :

1. Estimamos el modelo  $y = X\beta + u$  y  $e = y - \hat{y}$ .
2. Se ajusta por MCO el modelo en el que la variable endógena es  $|e_t|$  y la variable exógena es la variable  $X_i$ , la cual pensamos que es la que provoca la heteroscedasticidad, que será denotada por  $z_t$ , es decir

$$|e_t| = \delta_0 + \delta_1 z_t^h + v_t$$

donde  $v_t$  es ruido blanco. La regresión se realiza para distintos valores de  $h = \pm 1, \pm 2, \pm \frac{1}{2}, \dots$



## CONTRASTE DE GLEJSER

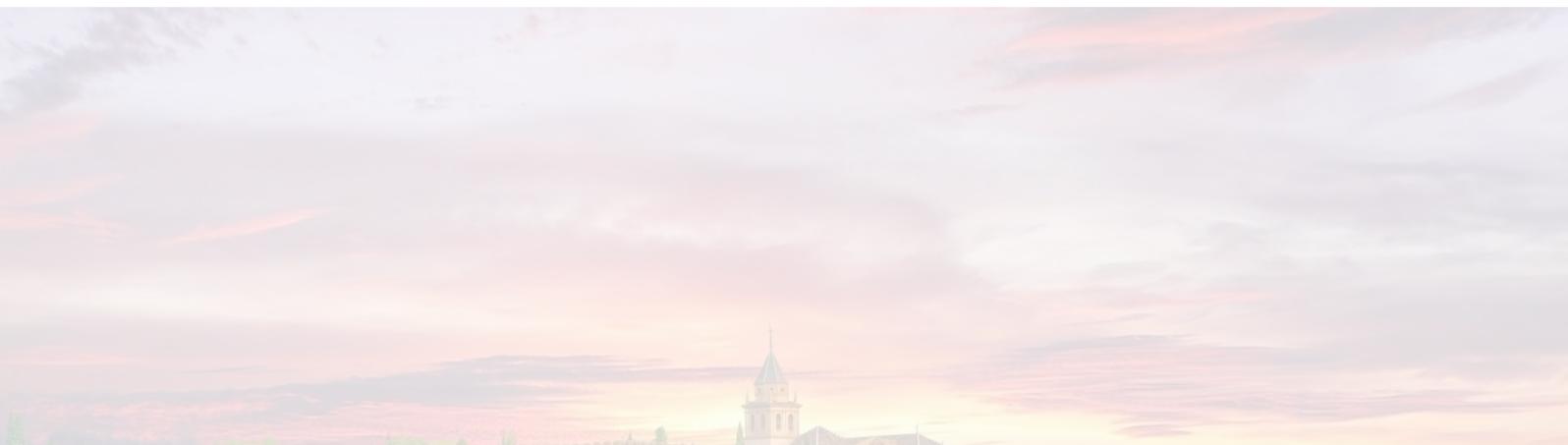
Si sospechamos de que la heteroscedasticidad está ligada a  $X_i$ :

1. Estimamos el modelo  $y = X\beta + u$  y  $e = y - \hat{y}$ .
2. Se ajusta por MCO el modelo en el que la variable endógena es  $|e_t|$  y la variable exógena es la variable  $X_i$ , la cual pensamos que es la que provoca la heteroscedasticidad, que será denotada por  $z_t$ , es decir

$$|e_t| = \delta_0 + \delta_1 z_t^h + v_t$$

donde  $v_t$  es ruido blanco. La regresión se realiza para distintos valores de  $h = \pm 1, \pm 2, \pm \frac{1}{2}, \dots$

3. Para cada  $h$ , contrastar  $H_0 : \delta_1 = 0$  ( $t$ -test). Si se rechaza  $H_0$ , entonces existe heteroscedasticidad en el modelo.



## EJEMPLO

$y_i$	55	65	70	80	79	84	...	178	191	137	189
$X_i$	80	100	85	110	120	115	...	265	270	230	250
$ e_i $	5.31	8.07	6.5	0.55	6.82	1.36	...	0.3	9.51	18.98	20.26

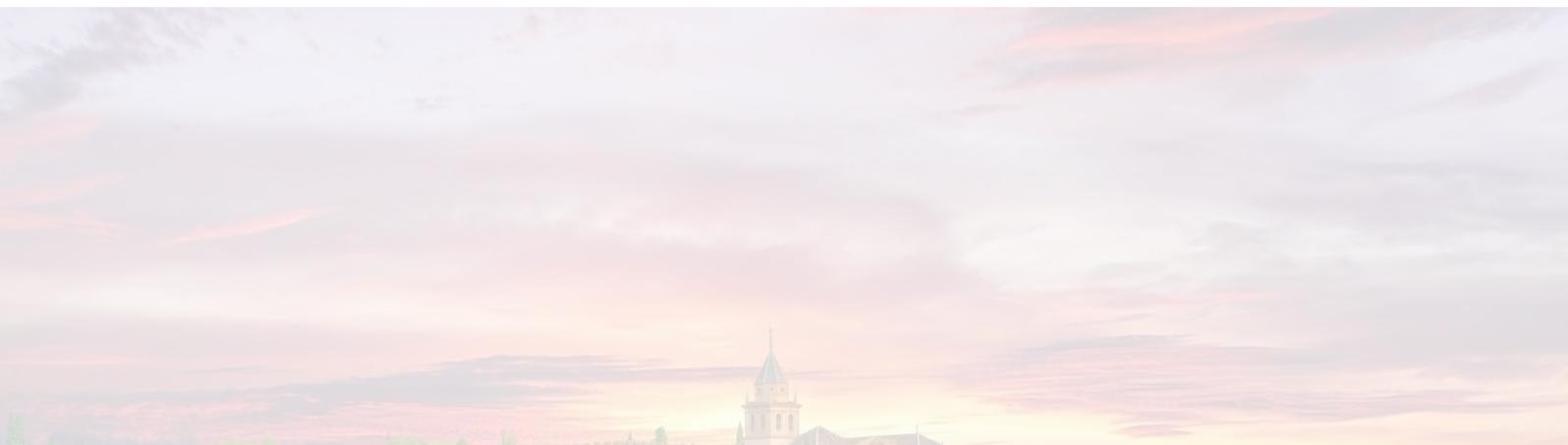
Para  $h = \pm 1, \pm 2, \pm \frac{1}{2}$  se estima el modelo de regresión en el cual la variable endógena es el valor absoluto de los errores de estimación,  $|e_t|$ , y la variable exógena es la variable que pensamos que puede ser la causante del problema de heteroscedasticidad,  $X_i$ .

$h = 1$	$\widehat{e}_t = 1.1126 + 0.0331 \cdot z_t$ (3.6348) (0.01988)	$R^2 = 0.1188$
$h = -1$	$\widehat{e}_t = 11.0619 - 639.379 \cdot z_t^{-1}$ (1.9876) (282.044)	$R^2 = 0.0814$
$h = 2$	$\widehat{e}_t = 3.6566 + 9.56 \cdot 10^{-5} \cdot z_t^2$ (1.3159) ( $3.34 \cdot 10^{-5}$ )	$R^2 = \mathbf{0.1233}$
$h = -2$	$\widehat{e}_t = 8.5926 - 35036.3 \cdot z_t^{-2}$ (1.1709) (18621.3)	$R^2 = 0.0575$
$h = 1/2$	$\widehat{e}_t = -3.8895 + 0.8289 \cdot z_t^{1/2}$ (4.0212) (0.3056)	$R^2 = 0.1126$
$h = -1/2$	$\widehat{e}_t = 16.0441 - 115.249 \cdot z_t^{-1/2}$ (3.8293) (47.1948)	$R^2 = 0.0932$

## EJEMPLO

Utilizando  $|t_{exp}| = \left| \frac{\hat{\delta}_1}{\sqrt{\text{var}[\hat{\delta}_1]}} \right|$  o el  $p$ -valor se estudia la significatividad del parámetro  $\delta_1$  de la regresión  $|e_t| = \delta_0 + \delta_1 z_t^h + v_t$ . Si se rechaza la hipótesis nula, existirá heteroscedasticidad.

$h = -2$	$ t_{exp}  = 1.8815$	$\not>$	$t_{29,0.975} = 2.045$	Se mantiene $H_0$
$h = -1$	$ t_{exp}  = 2.2669$	$>$	$t_{29,0.975} = 2.045$	Se <b>rechaza</b> $H_0$
$h = -\frac{1}{2}$	$ t_{exp}  = 2.4419$	$>$	$t_{29,0.975} = 2.045$	Se <b>rechaza</b> $H_0$
$h = 1$	$ t_{exp}  = 1.6649$	$\not>$	$t_{29,0.975} = 2.045$	Se mantiene $H_0$
$h = 2$	$ t_{exp}  = 2.8713$	$>$	$t_{29,0.975} = 2.045$	Se <b>rechaza</b> $H_0$
$h = \frac{1}{2}$	$ t_{exp}  = 2.7124$	$>$	$t_{29,0.975} = 2.045$	Se <b>rechaza</b> $H_0$

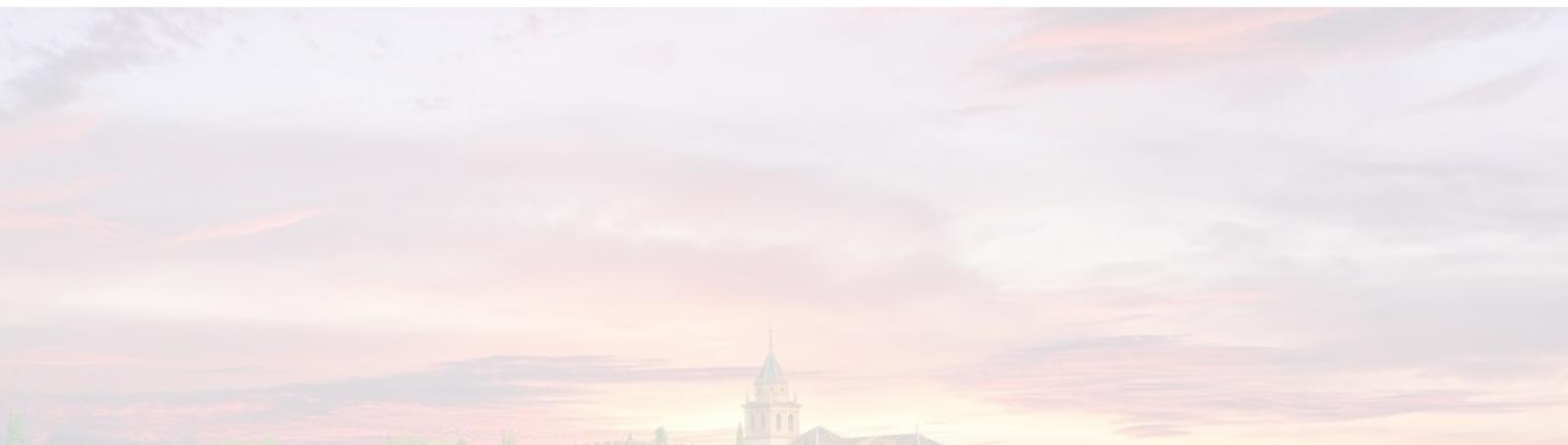


## EJEMPLO

Utilizando  $|t_{exp}| = \left| \frac{\hat{\delta}_1}{\sqrt{\text{var}[\hat{\delta}_1]}} \right|$  o el  $p$ -valor se estudia la significatividad del parámetro  $\delta_1$  de la regresión  $|e_t| = \delta_0 + \delta_1 z_t^h + v_t$ . Si se rechaza la hipótesis nula, existirá heteroscedasticidad.

$h = -2$	$ t_{exp}  = 1.8815$	$\not>$	$t_{29,0.975} = 2.045$	Se mantiene $H_0$
$h = -1$	$ t_{exp}  = 2.2669$	$>$	$t_{29,0.975} = 2.045$	Se <b>rechaza</b> $H_0$
$h = -\frac{1}{2}$	$ t_{exp}  = 2.4419$	$>$	$t_{29,0.975} = 2.045$	Se <b>rechaza</b> $H_0$
$h = 1$	$ t_{exp}  = 1.6649$	$\not>$	$t_{29,0.975} = 2.045$	Se mantiene $H_0$
$h = 2$	$ t_{exp}  = 2.8713$	$>$	$t_{29,0.975} = 2.045$	Se <b>rechaza</b> $H_0$
$h = \frac{1}{2}$	$ t_{exp}  = 2.7124$	$>$	$t_{29,0.975} = 2.045$	Se <b>rechaza</b> $H_0$

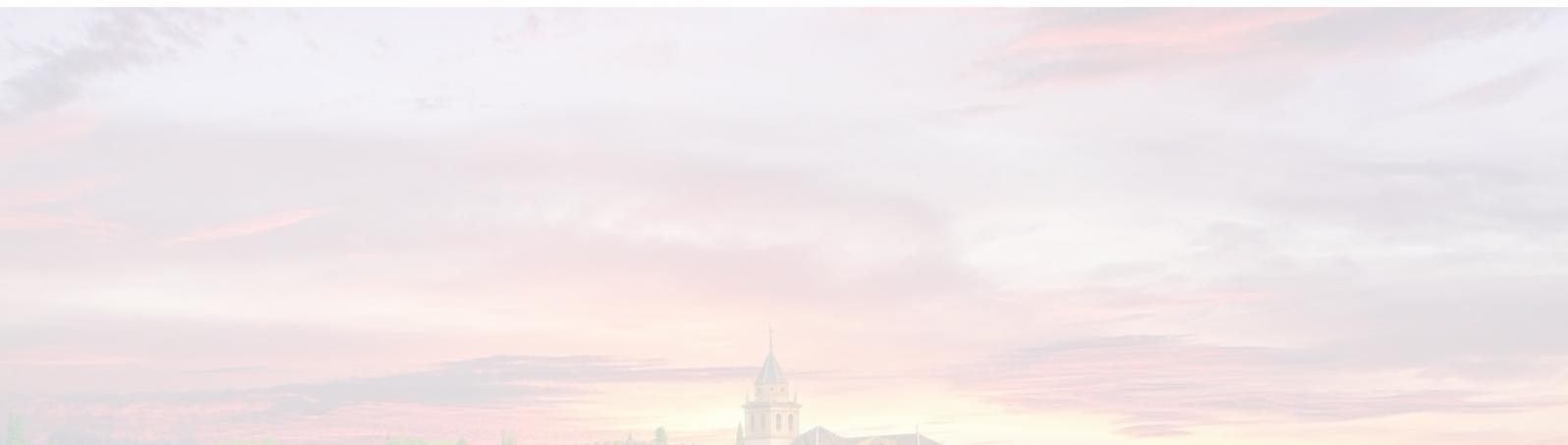
Para  $h = -1; 2; \frac{1}{2}; -\frac{1}{2}$ , existe heteroscedasticidad, y el modelo  $|\hat{e}_t| = 3.6566 + 9.56 \cdot 10^{-5} \cdot z_t^2$  el mejor (mayor  $R^2$ ). Podemos pensar que  $\mathbb{E}[u_t^2] = \sigma^2 z_t^2$ .



### 3. ESTIMACIÓN BAJO HETEROCEDASTICIDAD

Supongamos que queremos ajustar un modelo  $y = X\beta + u$ , donde sabemos que existe heterocedasticidad, es decir:

$$\mathbb{E}[uu^t] = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix}$$

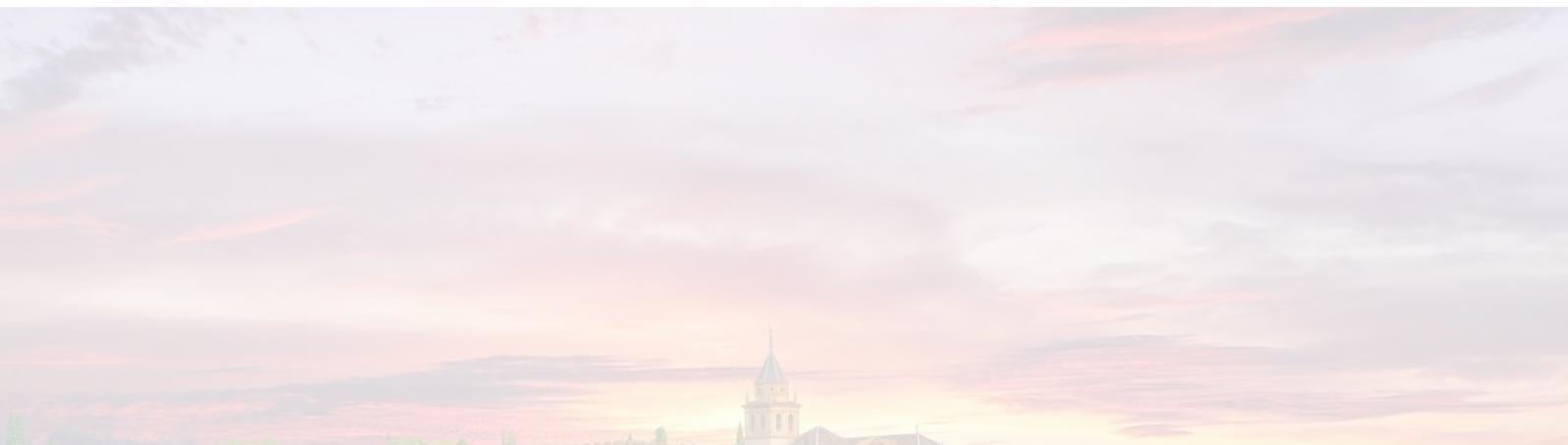


### 3. ESTIMACIÓN BAJO HETEROCEDASTICIDAD

Supongamos que queremos ajustar un modelo  $y = X\beta + u$ , donde sabemos que existe heterocedasticidad, es decir:

$$\mathbb{E}[uu^t] = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} \xrightarrow{\sigma^2 = \sigma_1^2} \mathbb{E}[uu^t] = \sigma^2 \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{\sigma_2^2}{\sigma_1^2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{\sigma_n^2}{\sigma_1^2} \end{pmatrix} = \sigma^2 \Omega$$

$$\Omega = \text{diag}(w_1, \dots, w_n), \text{ y } w_i = \frac{\sigma_i^2}{\sigma_1^2}$$

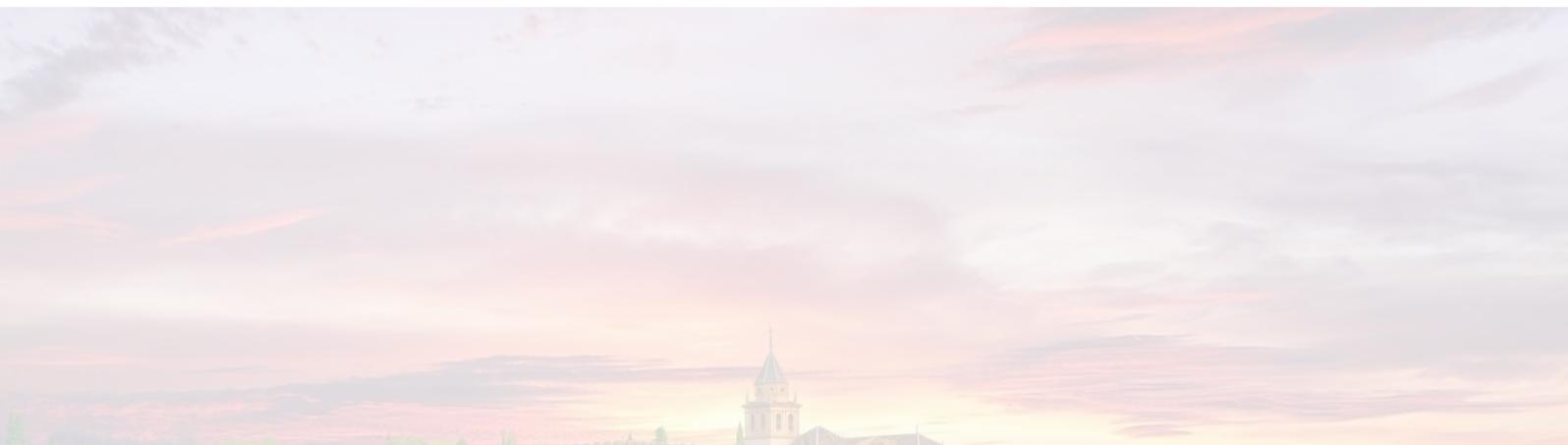


### 3. ESTIMACIÓN BAJO HETEROCEDASTICIDAD

Supongamos que queremos ajustar un modelo  $y = X\beta + u$ , donde sabemos que existe heterocedasticidad, es decir:

$$\mathbb{E}[uu^t] = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} \xrightarrow{\sigma^2 = \sigma_1^2} \mathbb{E}[uu^t] = \sigma^2 \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{\sigma_2^2}{\sigma_1^2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{\sigma_n^2}{\sigma_1^2} \end{pmatrix} = \sigma^2 \Omega$$

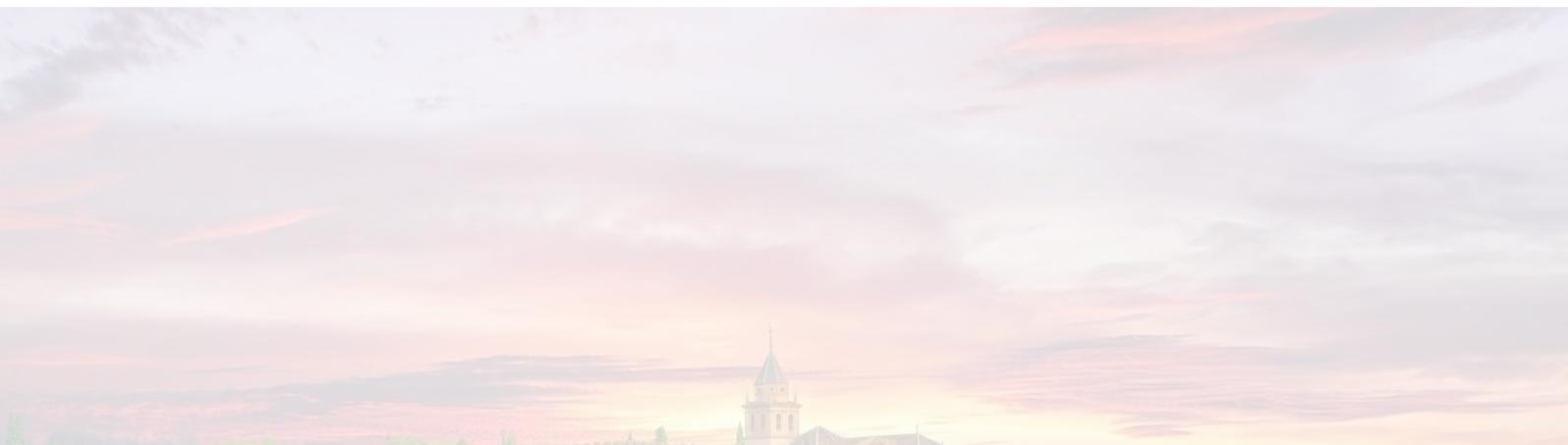
$$\Omega = \text{diag}(w_1, \dots, w_n), \text{ y } w_i = \frac{\sigma_i^2}{\sigma_1^2} \Rightarrow \text{Var}(u_i) = \mathbb{E}[u_i^2] = \sigma^2 w_i, \forall i.$$



## MÍNIMOS CUADRADOS GENERALIZADOS

Como  $\Omega \succ 0$  y simétrica, por Descomposición de Cholesky, existirá una matriz  $P(= Q^{-1})$ , no singular, tal que,

$$\Omega^{-1} = P^t P$$



## MÍNIMOS CUADRADOS GENERALIZADOS

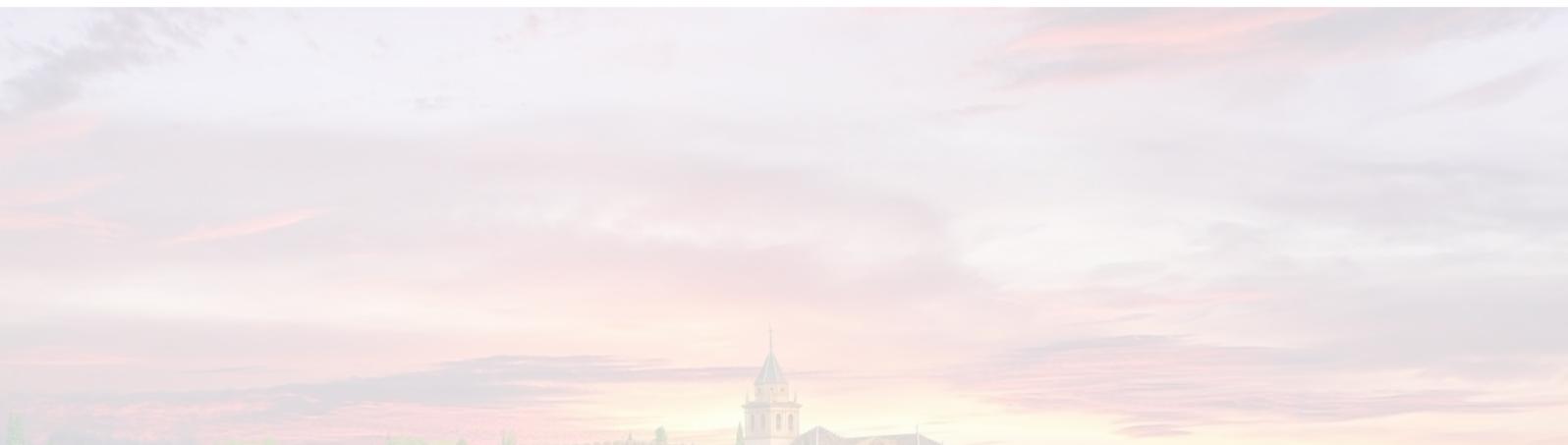
Como  $\Omega \succ 0$  y simétrica, por Descomposición de Cholesky, existirá una matriz  $P(= Q^{-1})$ , no singular, tal que,

$$\Omega^{-1} = P^t P$$

De manera que se podría transformar el modelo original para que la matriz  $\Omega$  se convierta en la matriz identidad y se pudiera aplicar MCO sobre el modelo transformado aplicando estas expresiones:

$$\begin{aligned}\widehat{\beta}_{MCG} &= [(X^*)^t X^*]^{-1} [(X^*)^t y^*] = [X^t P^t P X]^{-1} [X^t P^t P y] = \\ &= [X^t \Omega^{-1} X]^{-1} [X^t \Omega^{-1} y]\end{aligned}$$

$$\text{var} \left[ \widehat{\beta}_{MCG} \right] = \sigma^2 [(X^*)^t X^*]^{-1} = \sigma^2 [X^t P^t P X]^{-1} = \sigma^2 [X^t \Omega^{-1} X]^{-1}$$



## MÍNIMOS CUADRADOS GENERALIZADOS

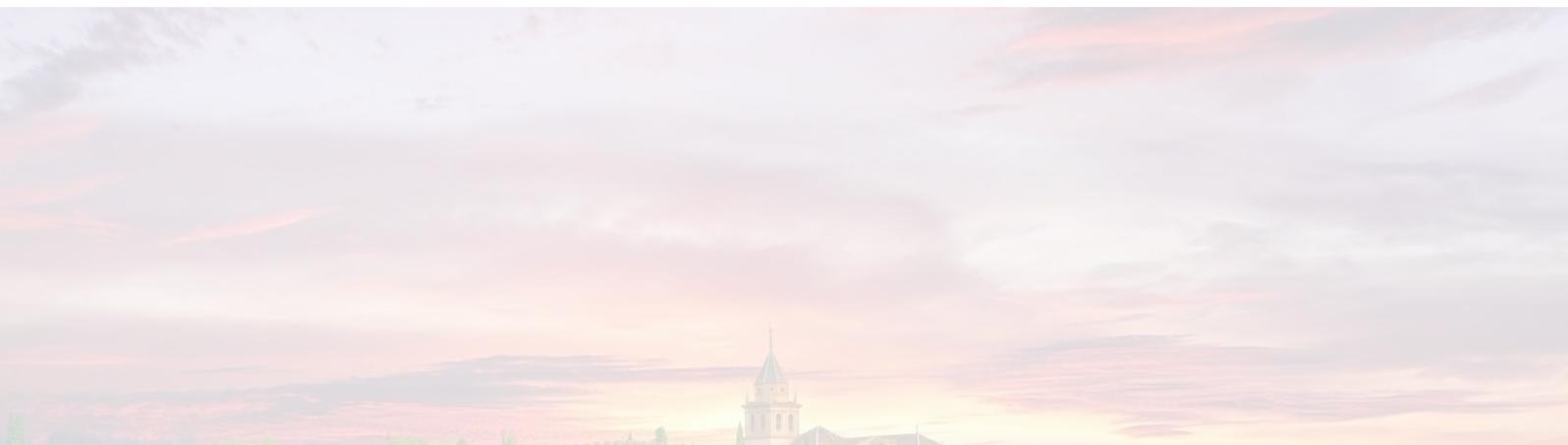
Como  $\Omega \succ 0$  y simétrica, por Descomposición de Cholesky, existirá una matriz  $P (= Q^{-1})$ , no singular, tal que,

$$\Omega^{-1} = P^t P$$

De manera que se podría transformar el modelo original para que la matriz  $\Omega$  se convierta en la matriz identidad y se pudiera aplicar MCO sobre el modelo transformado aplicando estas expresiones:

$$\begin{aligned}\widehat{\beta}_{MCG} &= [(X^*)^t X^*]^{-1} [(X^*)^t y^*] = [X^t P^t P X]^{-1} [X^t P^t P y] = \\ &= [X^t \Omega^{-1} X]^{-1} [X^t \Omega^{-1} y]\end{aligned}$$

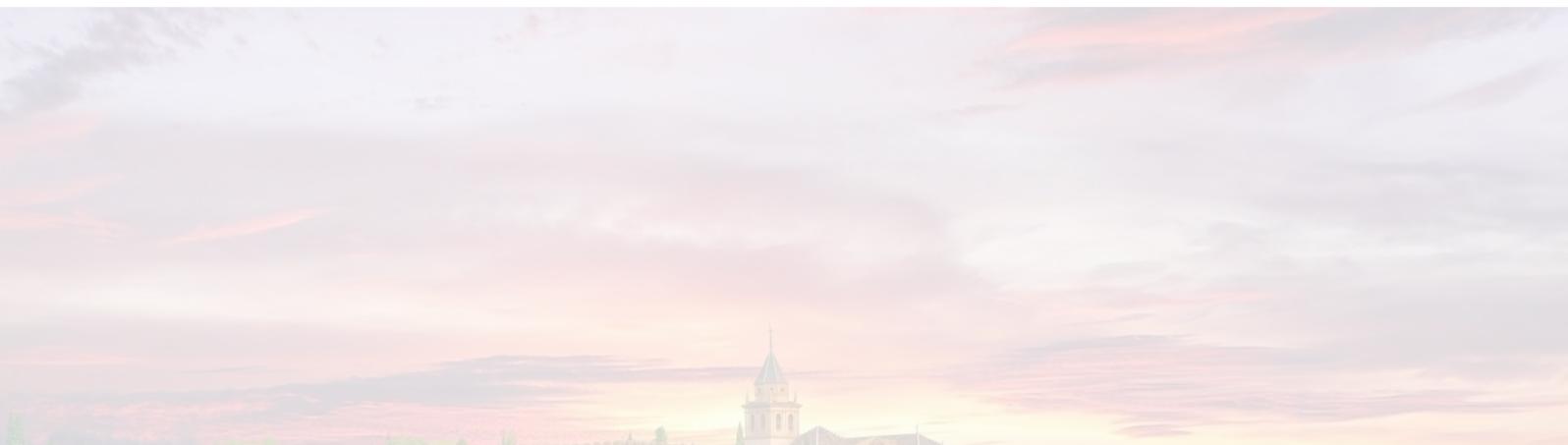
$$\text{var} [\widehat{\beta}_{MCG}] = \sigma^2 [(X^*)^t X^*]^{-1} = \sigma^2 [X^t P^t P X]^{-1} = \sigma^2 [X^t \Omega^{-1} X]^{-1}$$



## ESTIMADOR MÍNIMOS CUADRADOS PONDERADOS

En el caso particular de la existencia de heterocedastidad, ¿Cómo se descompone  $\Omega^{-1} = P^t P$ ?

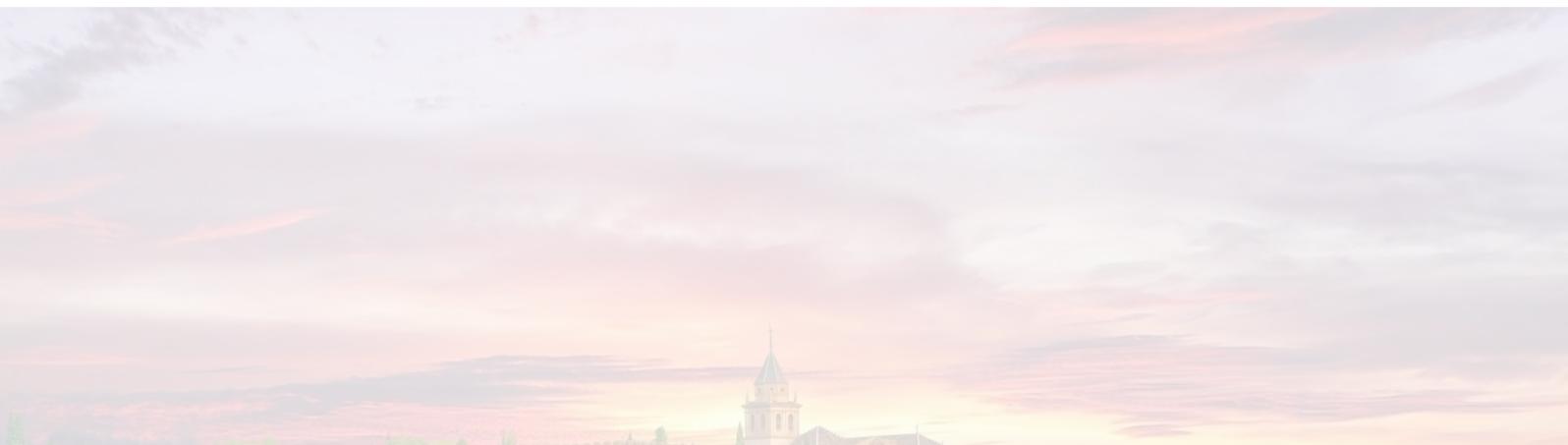
$$\Omega = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & w_n \end{pmatrix}$$



## ESTIMADOR MÍNIMOS CUADRADOS PONDERADOS

En el caso particular de la existencia de heterocedastidad, ¿Cómo se descompone  $\Omega^{-1} = P^t P$ ?

$$\Omega = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n \end{pmatrix} \Rightarrow \Omega^{-1} = \begin{pmatrix} \frac{1}{w_1} & 0 & \dots & 0 \\ 0 & \frac{1}{w_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{w_n} \end{pmatrix}$$



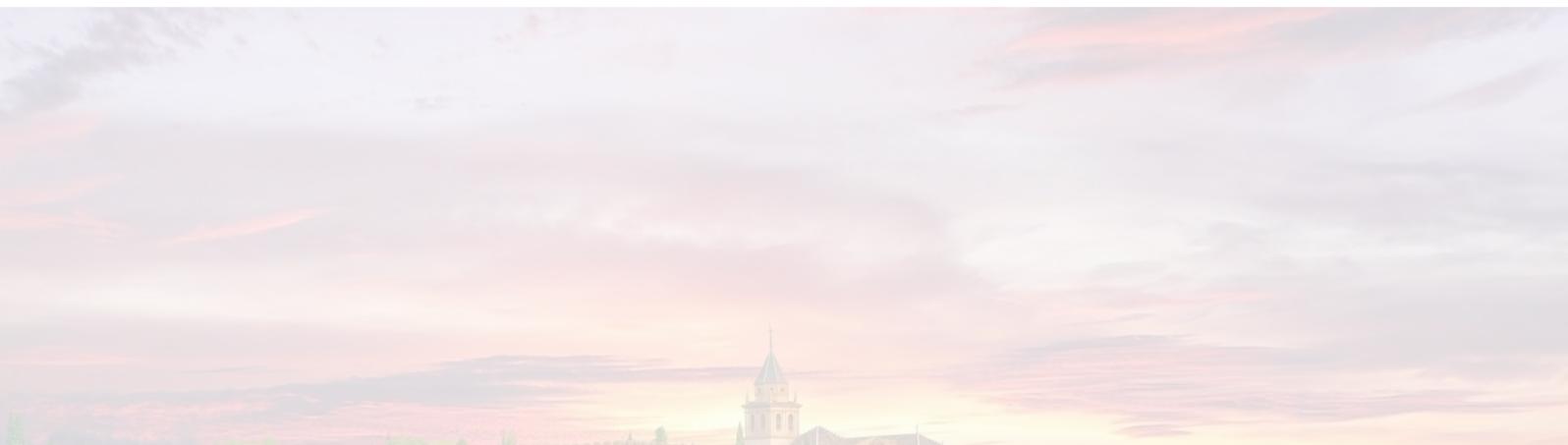
## ESTIMADOR MÍNIMOS CUADRADOS PONDERADOS

En el caso particular de la existencia de heterocedastidad, ¿Cómo se descompone  $\Omega^{-1} = P^t P$ ?

$$\Omega = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n \end{pmatrix} \Rightarrow \Omega^{-1} = \begin{pmatrix} \frac{1}{w_1} & 0 & \dots & 0 \\ 0 & \frac{1}{w_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{w_n} \end{pmatrix}$$

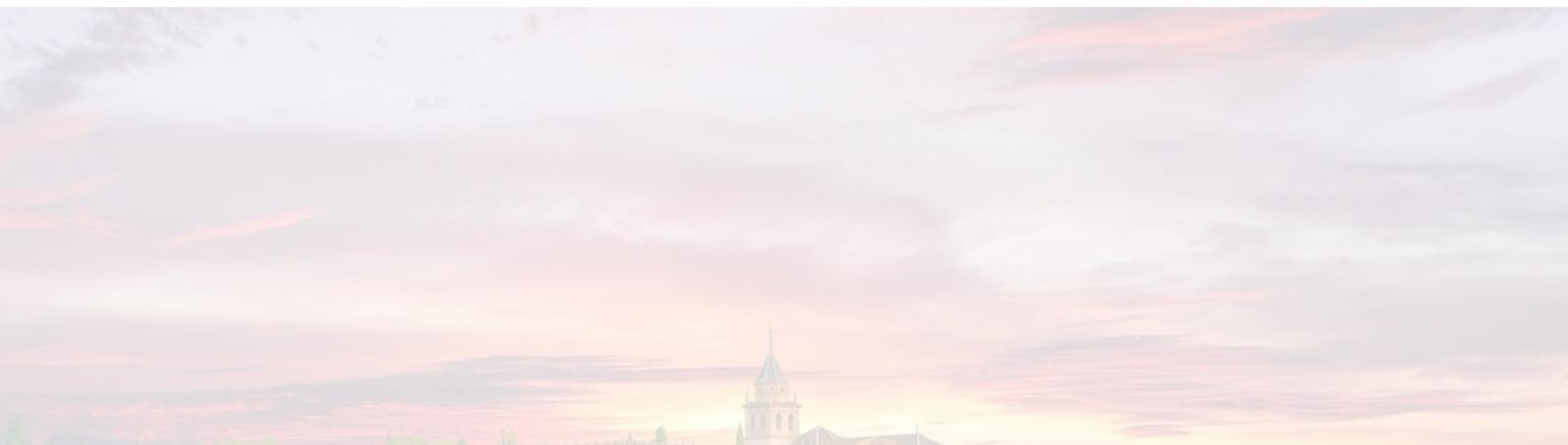
Considerando:

$$P = \begin{pmatrix} \frac{1}{\sqrt{w_1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{w_2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{w_n}} \end{pmatrix}$$



# MÍNIMOS CUADRADOS PONDERADOS

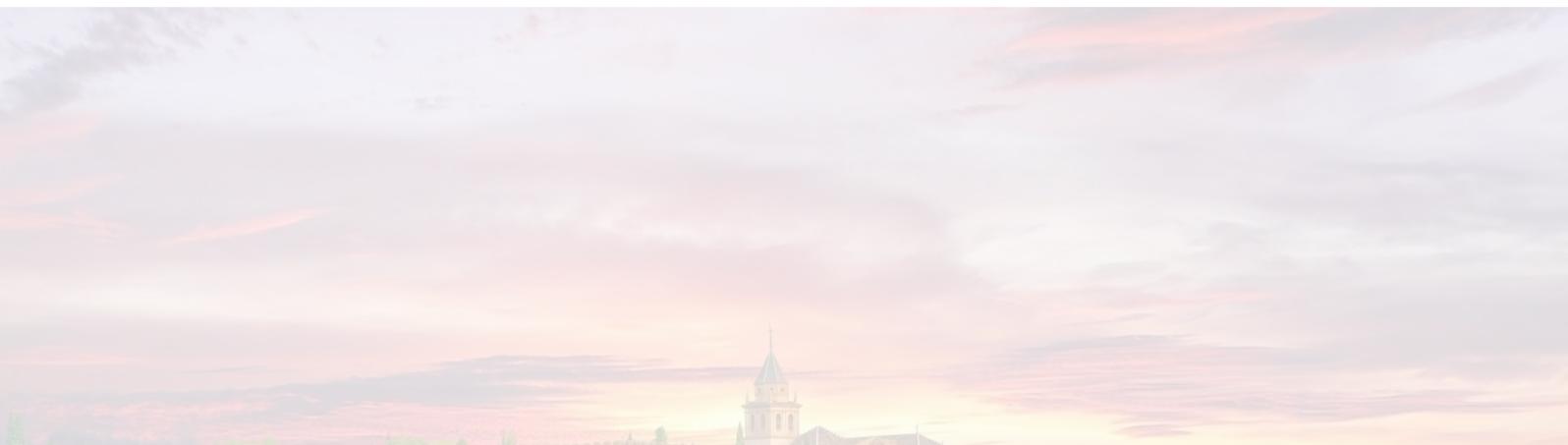
$$y^* = Py \quad X^* = PX \quad u^* = Pu$$



## MÍNIMOS CUADRADOS PONDERADOS

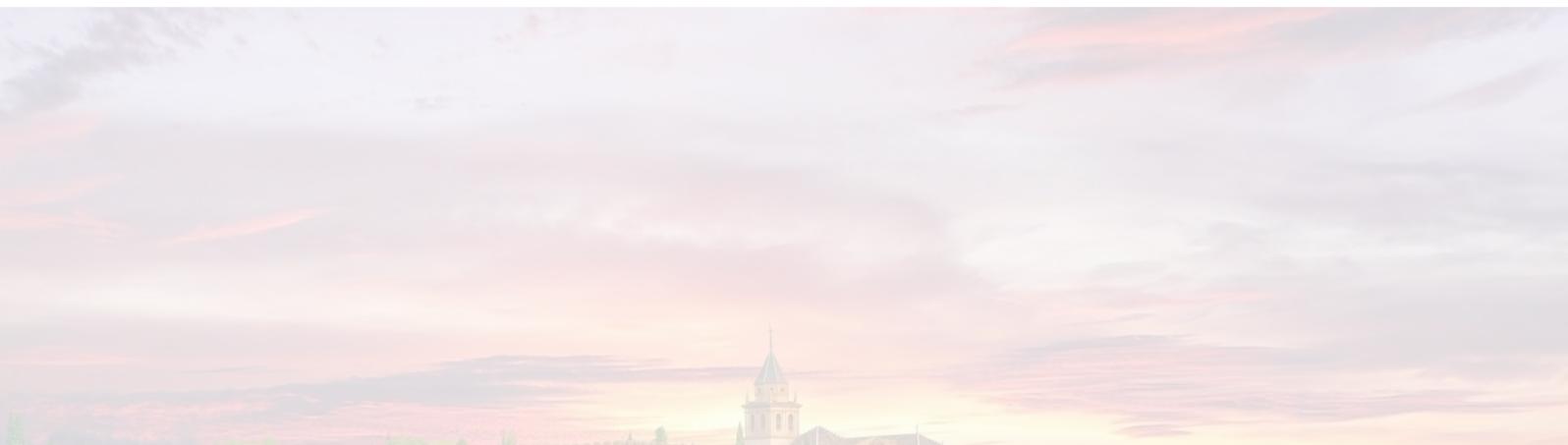
$$y^* = Py \quad X^* = PX \quad u^* = Pu$$

$$y^* = \begin{pmatrix} \frac{y_1}{\sqrt{w_1}} \\ \vdots \\ \frac{y_n}{\sqrt{w_n}} \end{pmatrix}, X^* = \begin{pmatrix} \frac{1}{\sqrt{w_1}} & \frac{x_{21}}{\sqrt{w_1}} & \cdots & \frac{x_{k1}}{\sqrt{w_1}} \\ \frac{1}{\sqrt{w_2}} & \frac{x_{22}}{\sqrt{w_2}} & \cdots & \frac{x_{k2}}{\sqrt{w_2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sqrt{w_n}} & \frac{x_{2n}}{\sqrt{w_n}} & \cdots & \frac{x_{kn}}{\sqrt{w_n}} \end{pmatrix}, u^* = \begin{pmatrix} \frac{u_1}{\sqrt{w_1}} \\ \vdots \\ \frac{u_n}{\sqrt{w_n}} \end{pmatrix}.$$



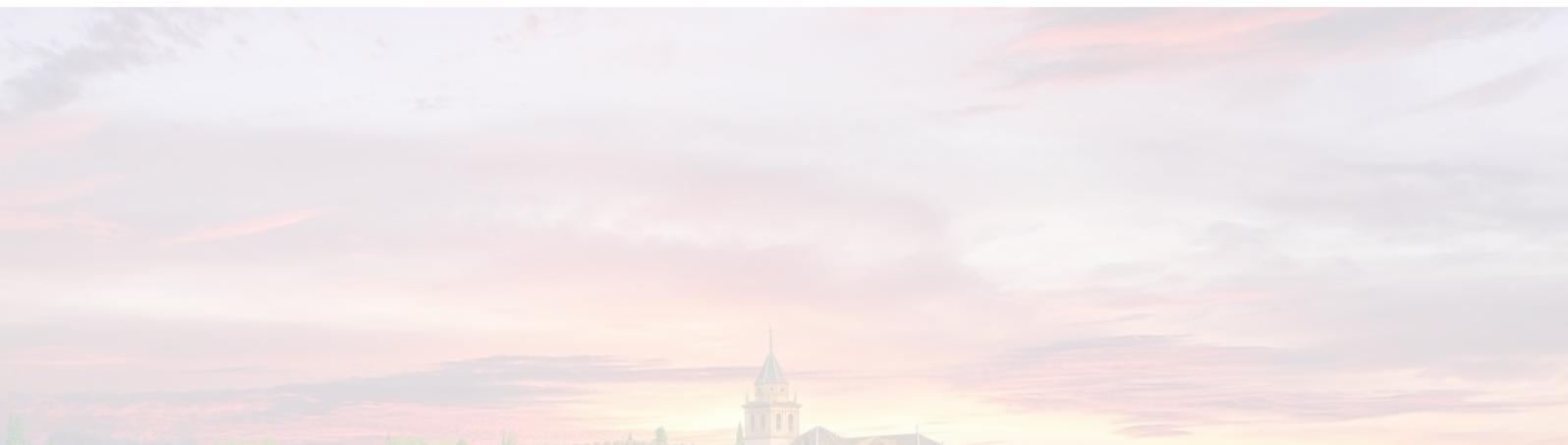
## MÍNIMOS CUADRADOS PONDERADOS

$$\blacktriangleright \mathbb{E}[u_i^*] = \mathbb{E}\left[\frac{u_i}{\sqrt{w_i}}\right] = \frac{1}{\sqrt{w_i}} \mathbb{E}[u_i] = 0.$$



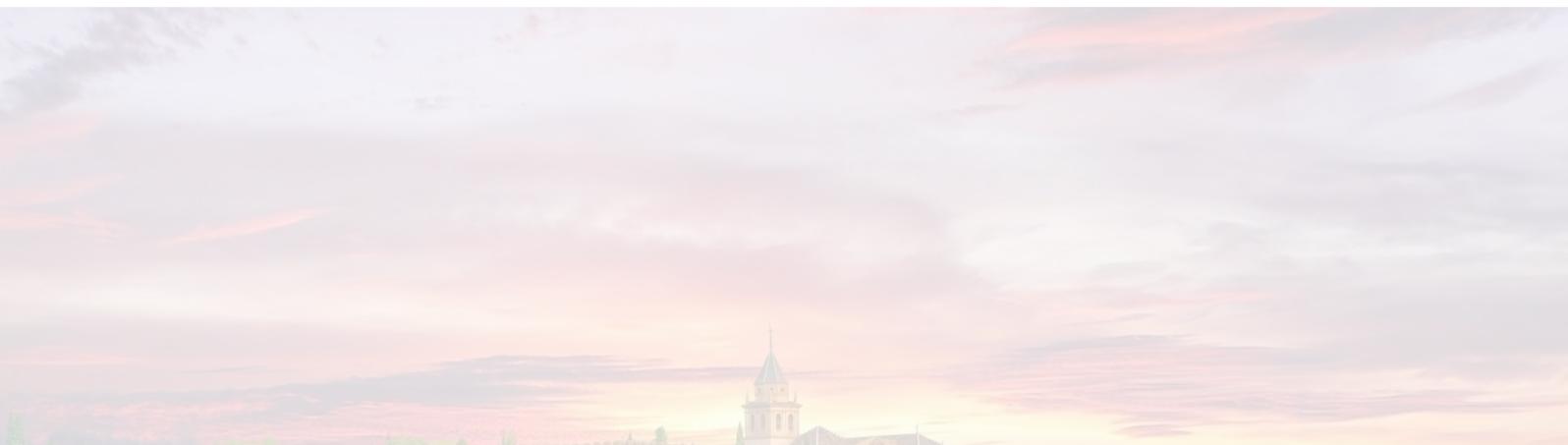
## MÍNIMOS CUADRADOS PONDERADOS

- ▶  $\mathbb{E}[u_i^*] = \mathbb{E}\left[\frac{u_i}{\sqrt{w_i}}\right] = \frac{1}{\sqrt{w_i}} \mathbb{E}[u_i] = 0.$
- ▶  $Var(u_i^*) = Var\left(\frac{u_i}{\sqrt{w_i}}\right) = \frac{1}{w_i} Var(u_i) = \frac{1}{w_i} \sigma^2 w_i = \sigma^2:$   
HOMOCEDEASTICIDAD!!!



## MÍNIMOS CUADRADOS PONDERADOS

- ▶  $\mathbb{E}[u_i^*] = \mathbb{E}\left[\frac{u_i}{\sqrt{w_i}}\right] = \frac{1}{\sqrt{w_i}} \mathbb{E}[u_i] = 0.$
- ▶  $Var(u_i^*) = Var\left(\frac{u_i}{\sqrt{w_i}}\right) = \frac{1}{w_i} Var(u_i) = \frac{1}{w_i} \sigma^2 w_i = \sigma^2:$   
HOMOCEDESTICIDAD!!!
- ▶  $\mathbb{E}[u_i^* u_j^*] = \mathbb{E}\left[\frac{u_i}{\sqrt{w_i}} \frac{u_j}{\sqrt{w_j}}\right] = \frac{1}{\sqrt{w_i w_j}} \mathbb{E}[u_i u_j] = 0:$  INCORRELACIÓN!!



## MÍNIMOS CUADRADOS PONDERADOS

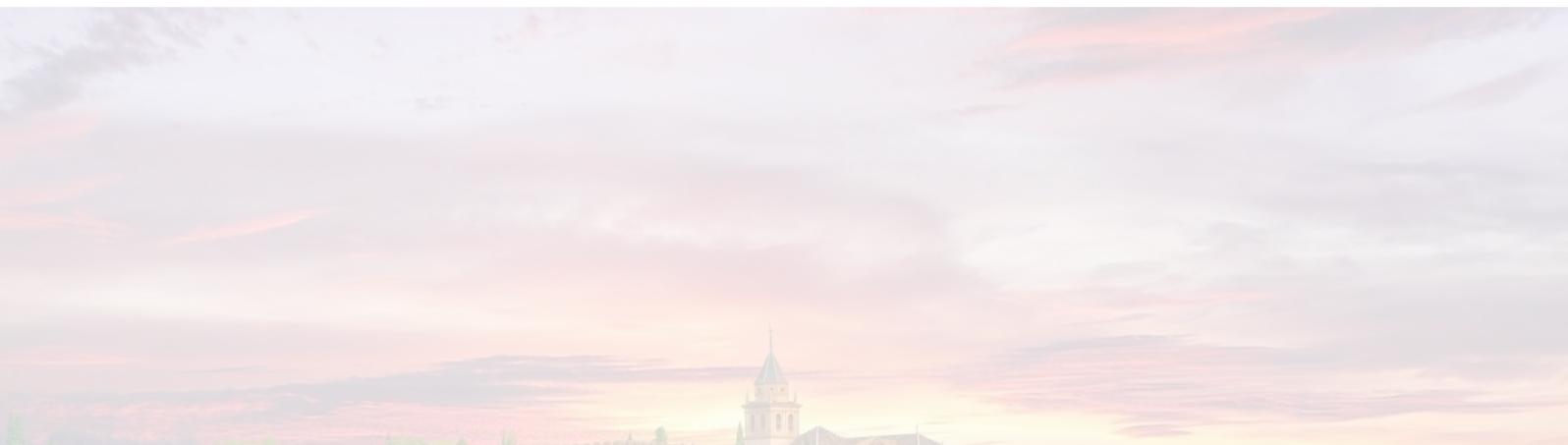
- ▶  $\mathbb{E}[u_i^*] = \mathbb{E}\left[\frac{u_i}{\sqrt{w_i}}\right] = \frac{1}{\sqrt{w_i}} \mathbb{E}[u_i] = 0.$
- ▶  $Var(u_i^*) = Var\left(\frac{u_i}{\sqrt{w_i}}\right) = \frac{1}{w_i} Var(u_i) = \frac{1}{w_i} \sigma^2 w_i = \sigma^2:$   
HOMOCEDESTICIDAD!!!
- ▶  $\mathbb{E}[u_i^* u_j^*] = \mathbb{E}\left[\frac{u_i}{\sqrt{w_i}} \frac{u_j}{\sqrt{w_j}}\right] = \frac{1}{\sqrt{w_i w_j}} \mathbb{E}[u_i u_j] = 0:$  INCORRELACIÓN!!

Podemos aplicar MCO para estimar  $y^* = X^* \beta + u^*$ .

## EJEMPLO

En el Ejemplo del Test de Glejser se tenía que para el modelo  $y = \beta_1 + \beta_2 x$  (solo una variable independiente):

$$\mathbb{E}[u_i^2] = \sigma^2 x_i^2$$



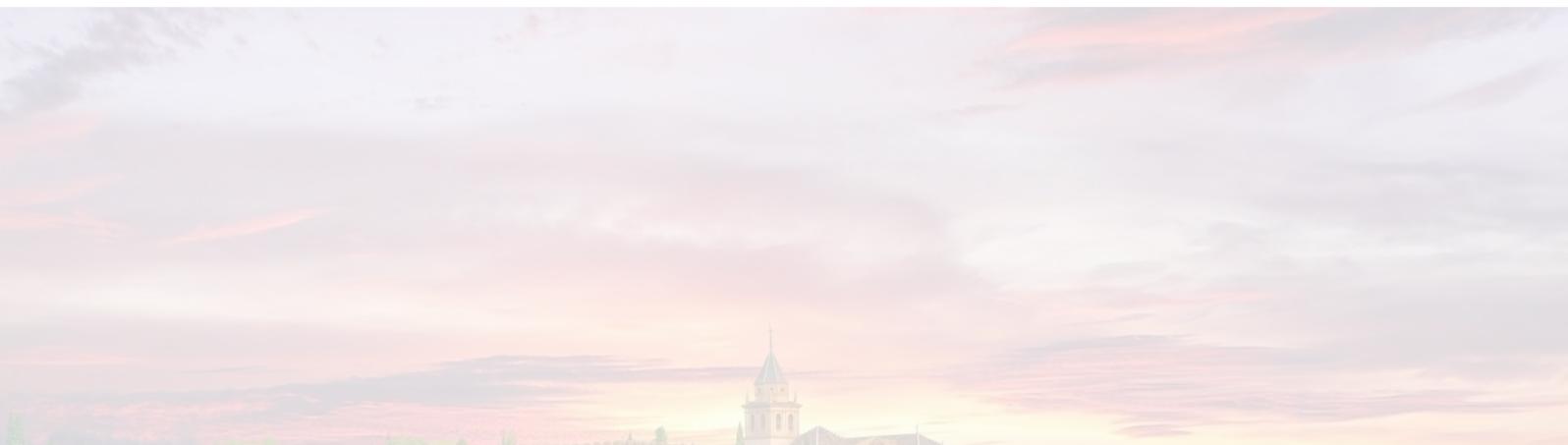
## EJEMPLO

En el Ejemplo del Test de Glejser se tenía que para el modelo  $y = \beta_1 + \beta_2 x$  (solo una variable independiente):

$$\mathbb{E}[u_i^2] = \sigma^2 x_i^2$$

En tal caso, la transformación que debemos hacer de nuestros datos es tomando  $w_i = x_i^2$ , esto es:

$$y^* = \begin{pmatrix} \frac{y_1}{x_1} \\ \vdots \\ \frac{y_n}{x_n} \end{pmatrix}, x^* = \begin{pmatrix} \frac{1}{x_1} & \frac{x_1}{x_1} \\ \frac{1}{x_2} & \frac{x_2}{x_2} \\ \vdots & \vdots \\ \frac{1}{x_n} & \frac{x_n}{x_n} \end{pmatrix} = \begin{pmatrix} \frac{1}{x_1} & 1 \\ \frac{1}{x_2} & 1 \\ \vdots & \vdots \\ \frac{1}{x_n} & 1 \end{pmatrix}.$$



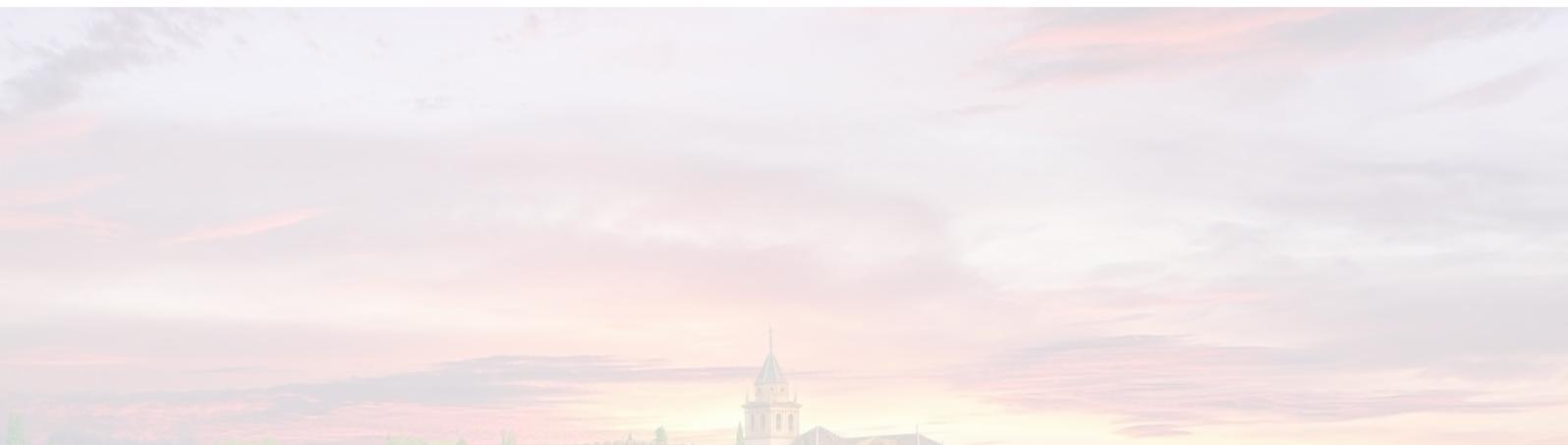
## EJEMPLO

Los estimadores se obtendrían como:

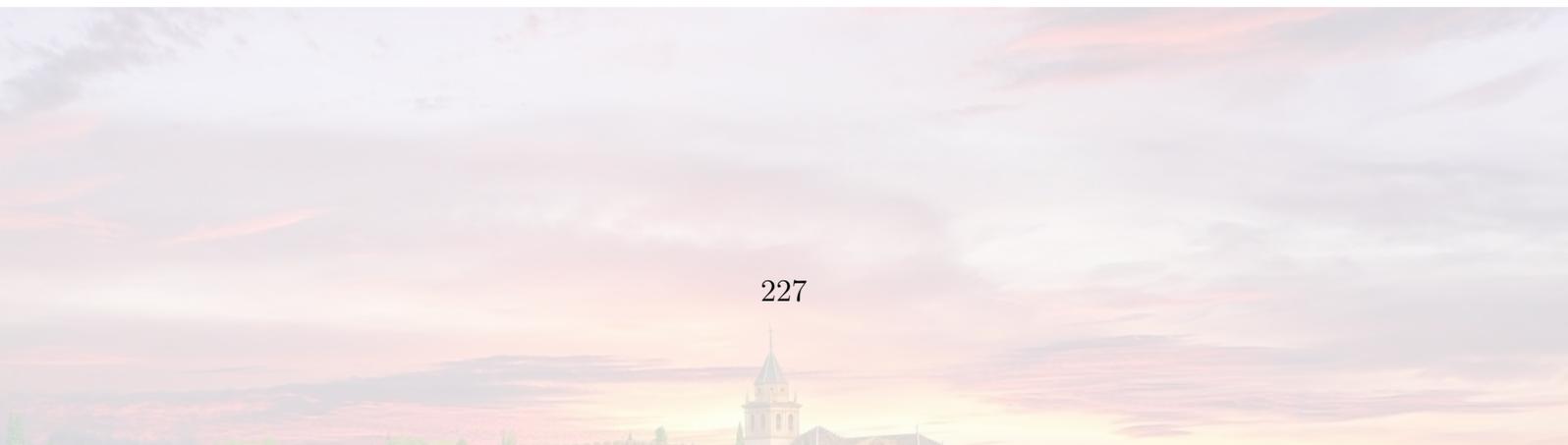
$$\begin{aligned}\widehat{\beta}_{MCG} &= [(X^*)^t X^*]^{-1} [(X^*)^t y^*] = [X^t P^t P X]^{-1} [X^t P^t P y] = \\ &= [X^t \Omega^{-1} X]^{-1} [X^t \Omega^{-1} y]\end{aligned}$$

donde  $\Omega^{-1} = \begin{pmatrix} \frac{1}{x_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{x_2^2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{x_n^2} \end{pmatrix}$

A partir de un patrón de comportamiento de  $\mathbb{E}[u_i u_j]$  en base a otra variable, realizamos la transformación de los datos y aplicamos MCO.



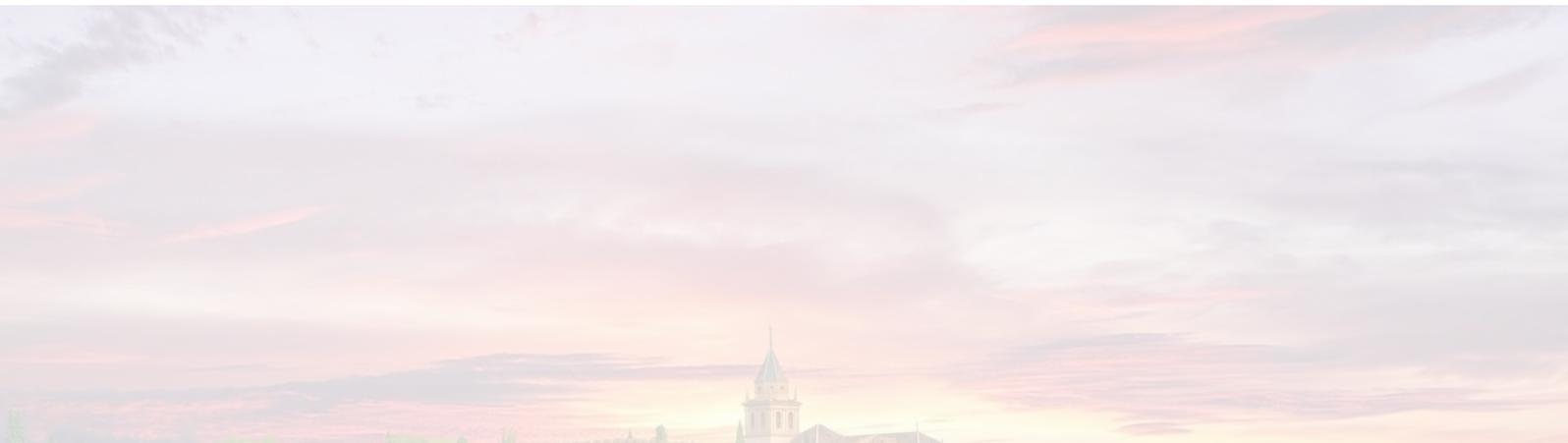
## 6 Tema 6



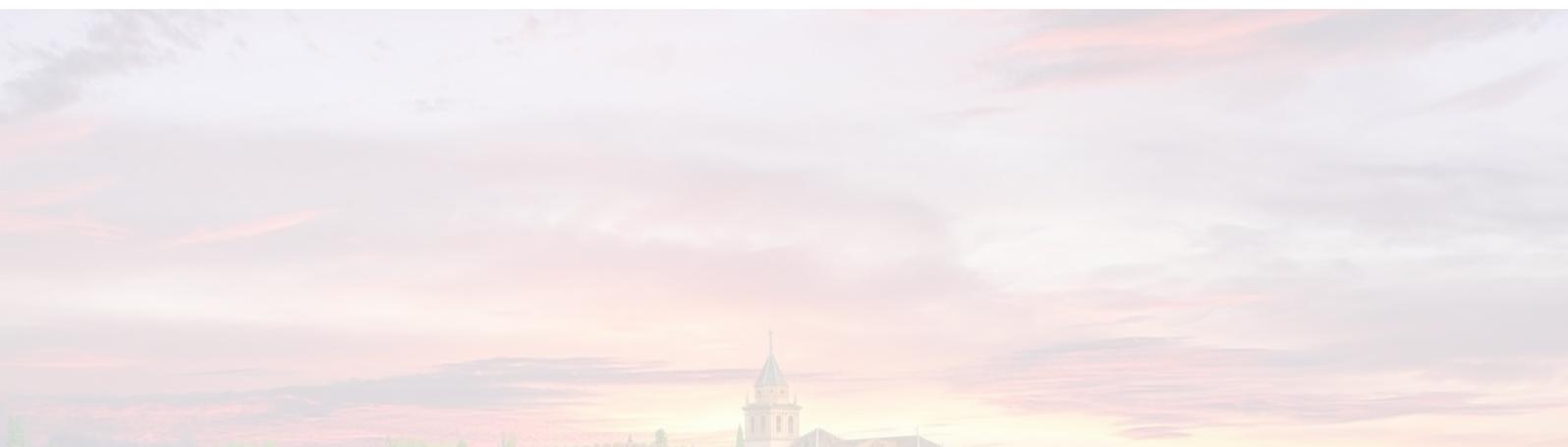
# ECONOMETRÍA

## TEMA 6: AUTOCORRELACIÓN

2024-2025



<b>1</b>	<b>Naturaleza, causas y consecuencias de la Autocorrelación . . . . .</b>	<b>2</b>
<b>2</b>	<b>Procedimientos para la detección de la Autocorrelación . . . . .</b>	<b>5</b>
2.1	Métodos Gráficos . . . . .	6
2.2	Contraste de Durbin-Watson . . . . .	10
2.3	Contraste H de Durbin . . . . .	14
2.4	Contraste de Ljung-Box . . . . .	15
<b>3</b>	<b>Estimación de modelos con perturbaciones autocorrelacionadas . .</b>	<b>20</b>
<b>4</b>	<b>Ejemplo . . . . .</b>	<b>26</b>



## NATURALEZA DE LA AUTOCORRELACIÓN

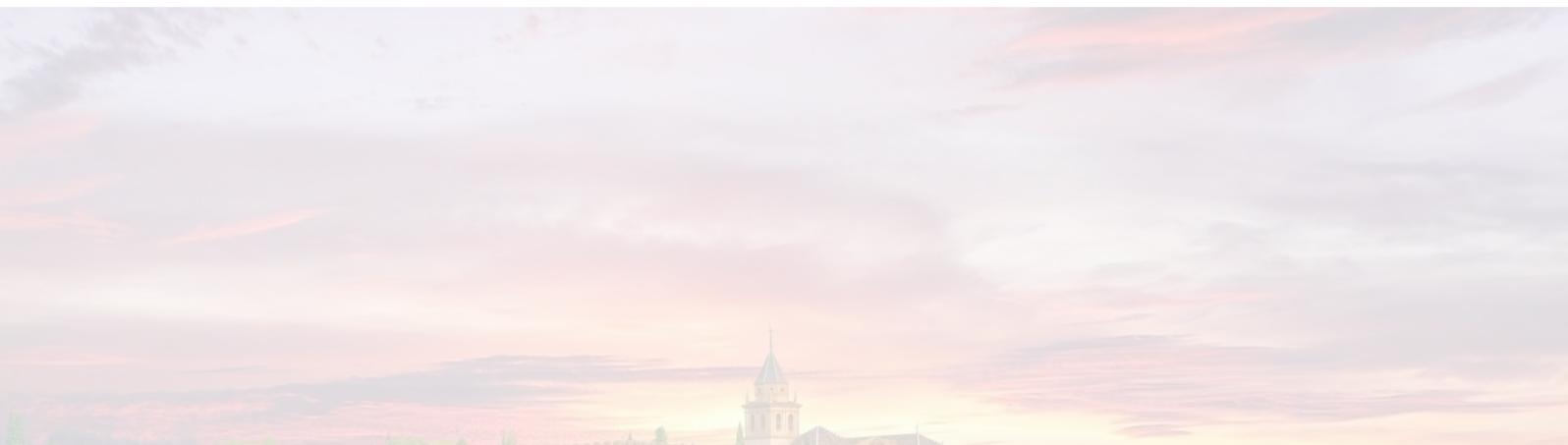
$$y = X\beta + u$$

con  $\mathbb{E}[u] = 0$  y  $\text{var}[u] = \mathbb{E}[uu^t] = \sigma^2 I_n$ , lo que implica:

- ▶  $\mathbb{E}[u_t] = 0, \forall t \in 1, \dots, n$
- ▶  $\mathbb{E}[u_t^2] = \sigma^2, \forall t \in 1, \dots, n$  (Homocedasticidad)
- ▶  $\mathbb{E}[u_i, u_j] = \text{Cov}[u_i, u_j] = 0, \forall i \neq j \in 1, \dots, n$  (Incorrelación).

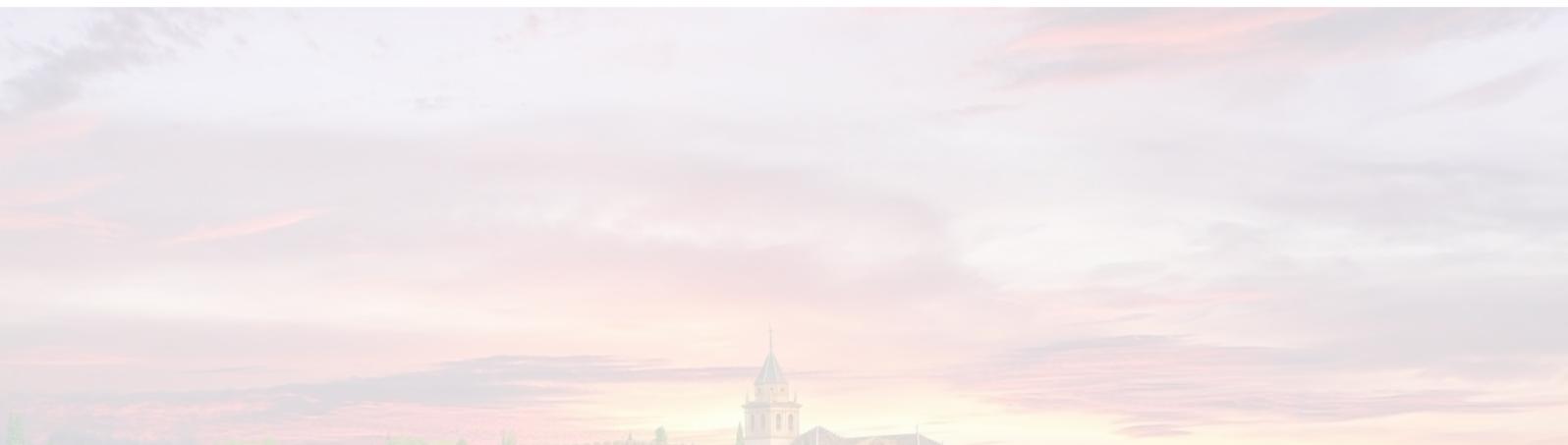
Si  $\text{Cov}[u_t, u_{t-k}] \neq 0 \forall k > 0 \rightarrow$  **Autocorrelación.**

Suponemos entonces que:  $u_t = \rho u_{t-1} + \varepsilon_t$  donde  $\varepsilon_t$  verifica que es un ruido blanco gaussiano.



## CAUSAS DE LA AUTOCORRELACIÓN

- ▶ Error en la especificación del modelo, es decir, se han omitido en la construcción del modelo variables exógenas que son relevantes con el fin de estudiar el comportamiento de la variable endógena.
- ▶ Existencia de ciclos y tendencias. Si la variable endógena presenta ciclos, y éstos no están correctamente explicados por el conjunto de las variables exógenas, entonces el término de perturbación presentará autocorrelación.
- ▶ Existencia de relaciones no lineales.
- ▶ Existencia de relaciones dinámicas.



## CONSECUENCIAS DE LA AUTOCORRELACIÓN

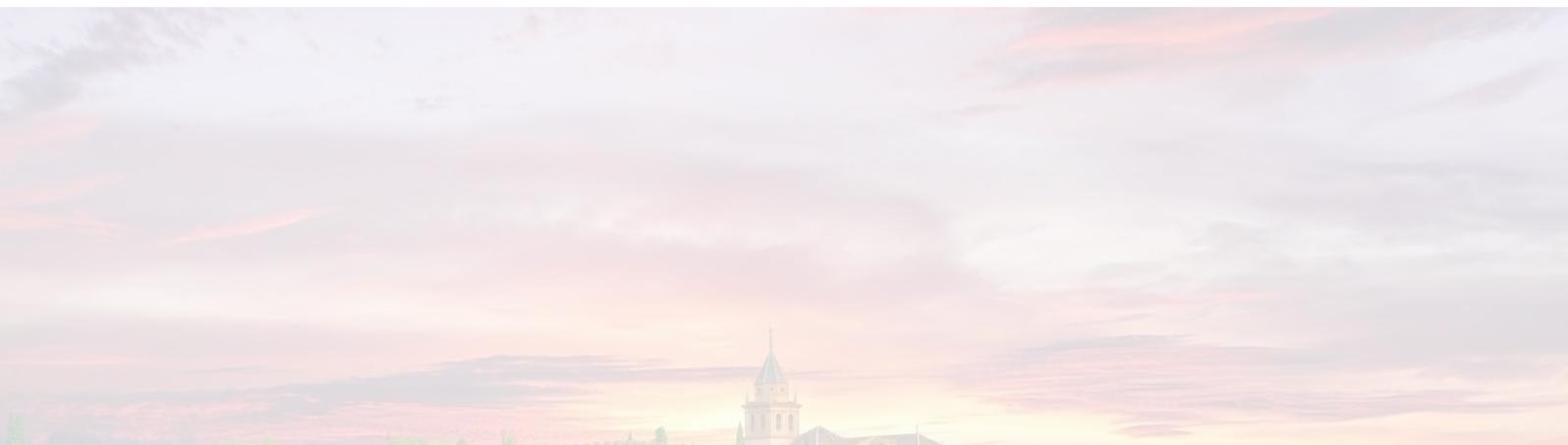
Los estimadores MCO son insesgados pero **no óptimos**.  
El estimador por Mínimos Cuadrados Generalizados (MCG)

$$\hat{\beta}_{MCG} = [X^t \Omega^{-1} X]^{-1} [X^t \Omega^{-1} y]$$

cuya matriz de varianzas-covarianzas viene dada por

$$\text{var} [\hat{\beta}_{MCG}] = \sigma^2 [X^t \Omega^{-1} X]^{-1}$$

Sí es **óptimo**.



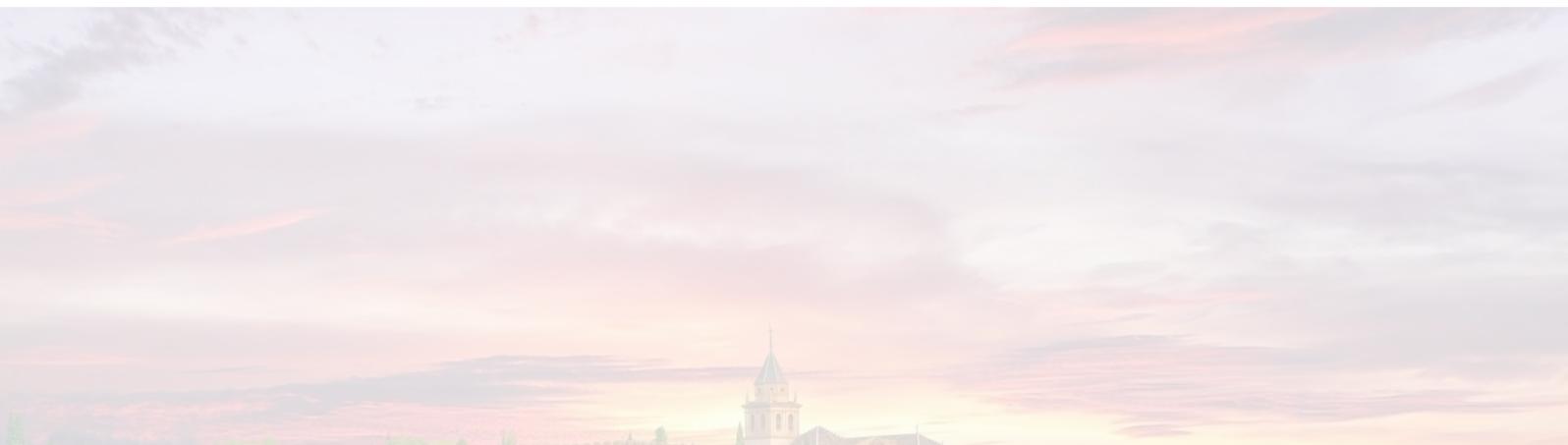
# DETECCIÓN DE LA AUTOCORRELACIÓN

## ▶ Gráficamente

- Gráfico temporal de los residuos obtenidos por MCO. Si los residuos están incorrelados deben distribuirse de forma aleatoria alrededor de cero.
- Gráfico de dispersión de los mismo frente a algún retardo suyo. Los residuos estarán autocorrelacionados cuando el gráfico presenta una tendencia creciente o decreciente.

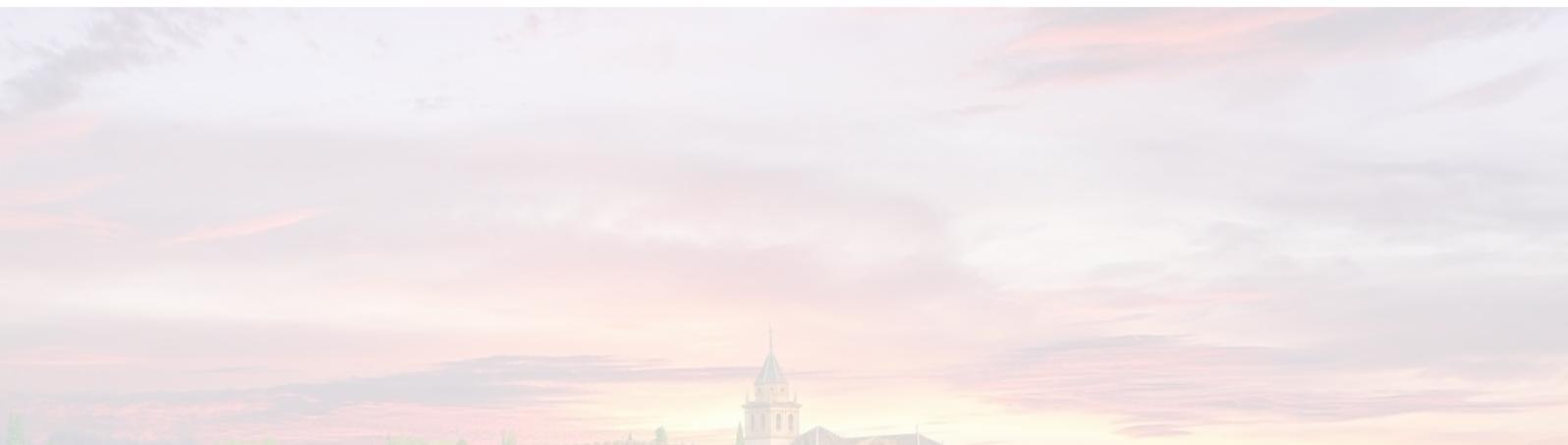
## ▶ Analíticamente

- Contraste de Durbin-Watson.
- Estadístico H de Durbin.
- Contraste de Ljung-Box



## MÉTODOS GRÁFICOS

- ▶ **Gráfico temporal de los residuos:** si los residuos están incorrelados deben distribuirse de forma aleatoria alrededor de cero. Sin embargo, si están correlacionados:
  - observaremos rachas de residuos por debajo y por encima de la media (correlación positiva).
  - observaremos una alternancia en el signo de los residuos (correlación negativa).
- ▶ **Gráficos de dispersión:** el gráfico de dispersión de los residuos,  $e_t$ , frente algún retardo suyo,  $e_{t-k}$  (normalmente se considera  $k = 1$ ), puede mostrar la existencia de autocorrelación positiva (tendencia creciente en el gráfico) o negativa (tendencia decreciente).



## EJEMPLO

C: Consumo de Energía Eléctrica  
(miles de TEP)

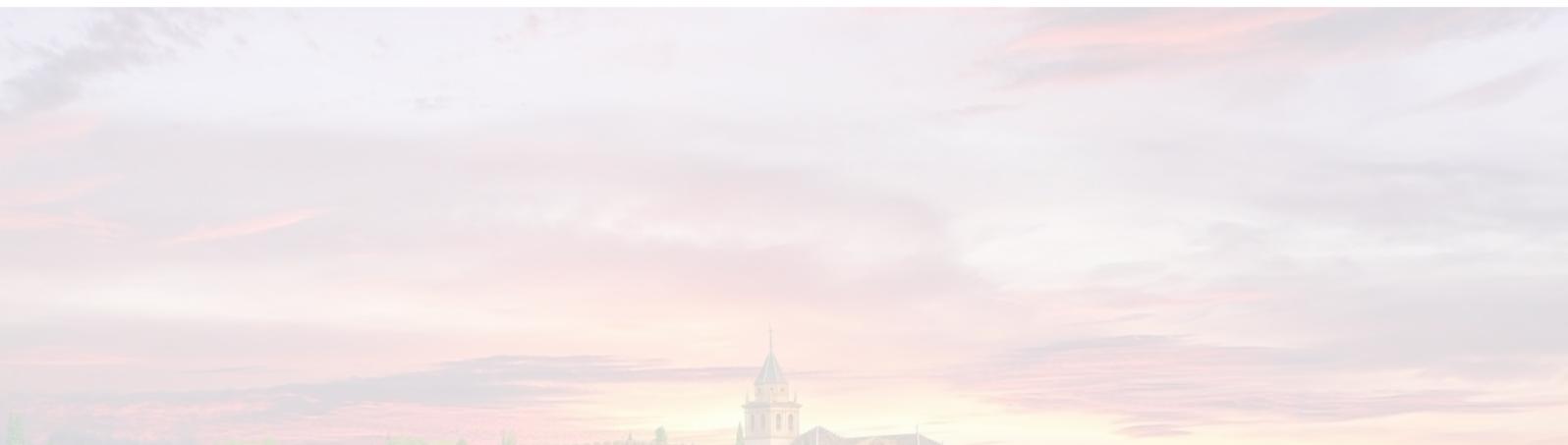
$$C = \beta_0 + \beta_1 \text{PIB} + u$$

↓

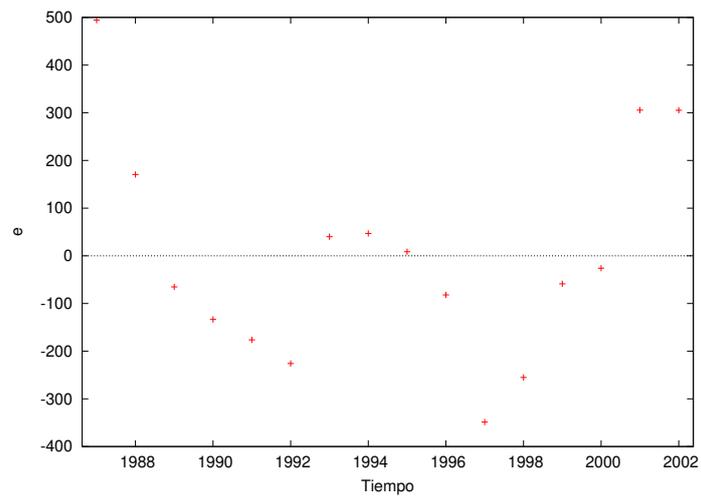
$$\hat{C}_t = -6234.54 + 0.0426873 \cdot \text{PIB}_t$$

$$R^2 = 0.992408$$

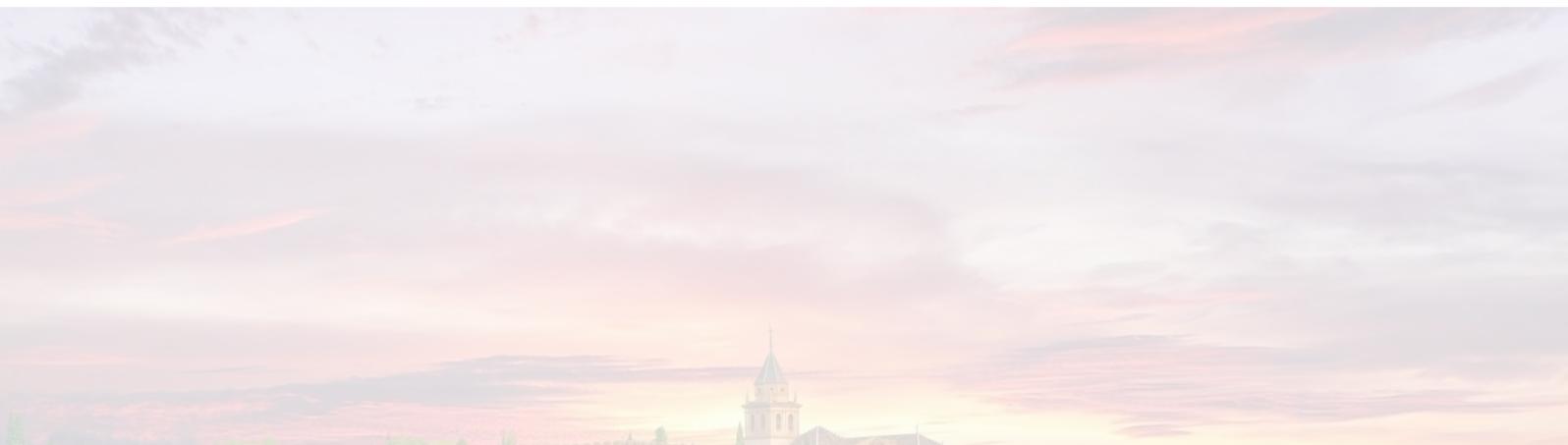
Año	C	PIB
1987	9427	355312
1988	9876	373412
1989	10410	391443
1990	10974	406252
1991	11372	416582
1992	11488	420462
1993	11569	416126
1994	11999	426041
1995	12462	437787
1996	12827	448457
1997	13331	466513
1998	14290	486785
1999	15364	507346
2000	16309	528714
2001	17282	543746
2002	17756	554852



# EJEMPLO



**Figure.** Gráfico temporal de los residuos



# EJEMPLO

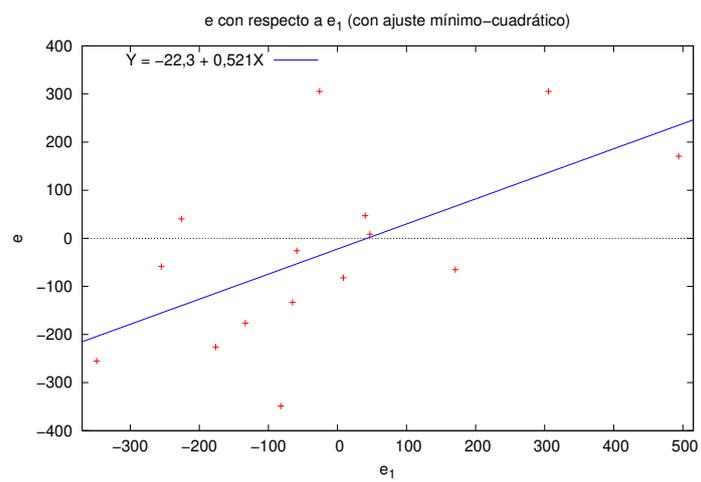
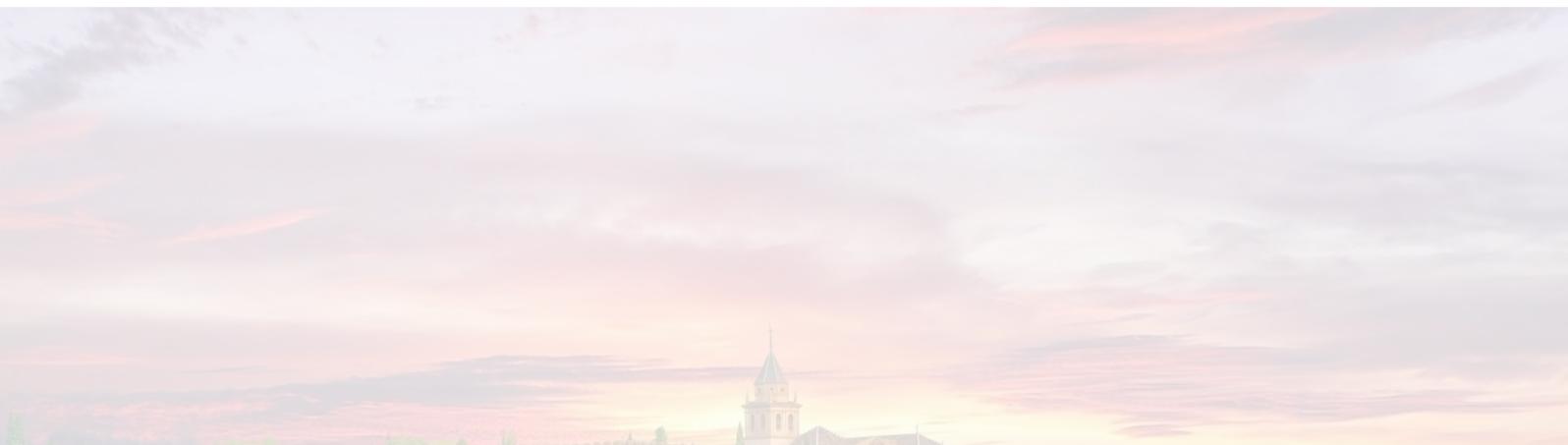


Figure. Gráfico de dispersión



## CONTRASTE DE DURBIN-WATSON

El contraste de Durbin-Watson resulta adecuado si suponemos que la autocorrelación que presenta el término de perturbación pudiera ser descrito mediante un proceso autorregresivo de orden uno:

$$u_t = \rho u_{t-1} + \varepsilon_t$$

donde  $\varepsilon_t$  es ruido blanco.

Estudiamos si el coeficiente  $\rho$  es significativo o no mediante los contrastes:

$$\begin{cases} H_0 : \rho = 0 & \text{(Incorrelación)} \\ H_1 : \rho < 0 & \text{(Correlación negativa)} \end{cases} \quad \begin{cases} H_0 : \rho = 0 & \text{(Incorrelación)} \\ H_1 : \rho > 0 & \text{(Correlación positiva)} \end{cases}$$

Estadístico de **Contraste de Durbin-Watson**:

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

siendo  $e_t$  los residuos MCO:  $e_t = y_t - \hat{y}_t = y_t - x_t \hat{\beta}$ .

## CONTRASTE DE DURBIN-WATSON

Si la muestra seleccionada es suficientemente grande, se puede considerar que  $\sum_{t=2}^n e_t^2 \simeq \sum_{t=2}^n e_{t-1}^2$ , ya que ambas sumatorias difieren únicamente en la primera de las observaciones muestrales:

$$DW = \frac{2 \sum_{t=2}^n e_t^2 - 2 \sum_{t=2}^n e_t e_{t-1}}{\sum_{t=1}^n e_t^2} = 2 \left( \frac{\sum_{t=2}^n e_t^2}{\sum_{t=1}^n e_t^2} - \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=1}^n e_t^2} \right) \simeq 2(1 - \hat{\rho})$$

donde  $\hat{\rho}$  es la estimación del parámetro  $\rho$ .

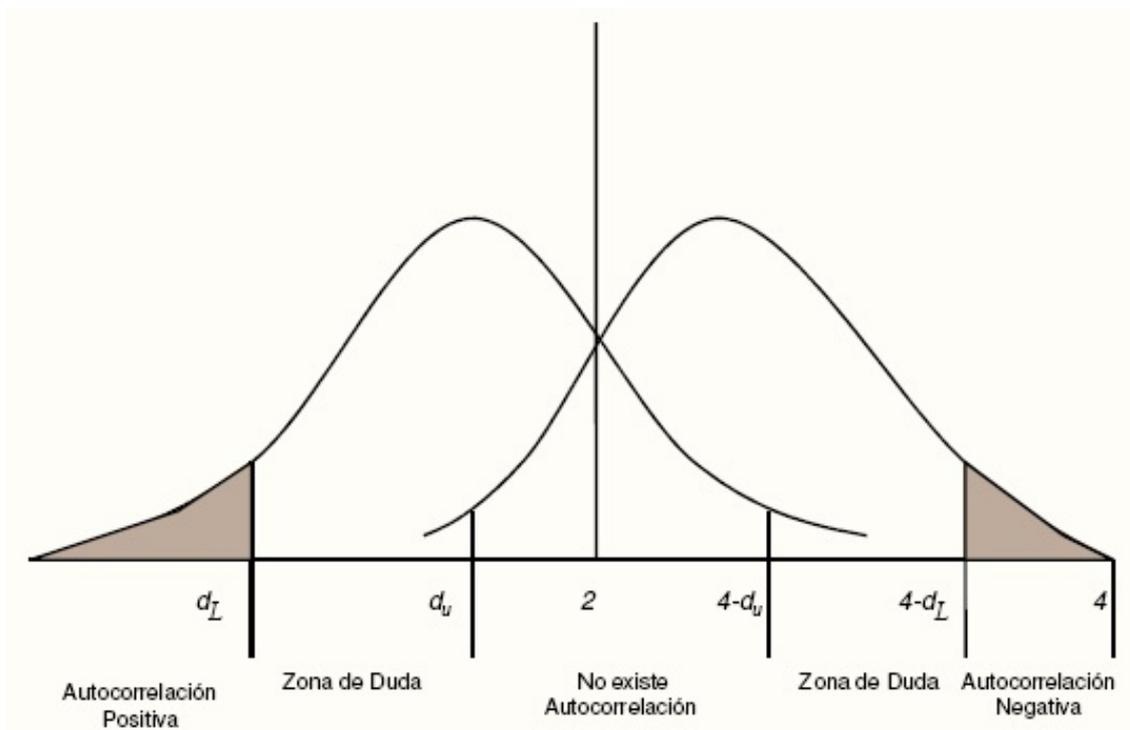
Como  $|\rho| < 1$ :

- ▶  $DW \simeq 0$  cuando  $\hat{\rho} \simeq 1$  (correlación positiva)
- ▶  $DW \simeq 2$  cuando  $\hat{\rho} \simeq 0$  (residuos incorrelacionados)
- ▶  $DW \simeq 4$  cuando  $\hat{\rho} \simeq -1$  (correlación es negativa)

```
from statsmodels.stats.stattools import durbin_watson
durbin_watson(mco.resid)
```

# CONTRASTES DE DURBIN-WATSON

Gráficamente se tiene:



## CONTRASTES DE DURBIN-WATSON

Las tablas estadísticas nos proporcionan una cota inferior,  $d_L$ , y otra cota superior,  $d_U$ . Si:

- ▶  $0 < DW < d_L$ , se mantiene la hipótesis  $\rho > 0$ , autocorrelación positiva.
- ▶  $d_L < DW < d_U$ , test no concluyente.
- ▶  $d_U < DW < 4 - d_U$ , se mantiene que  $\rho = 0$ , residuos incorrelados.
- ▶  $4 - d_U < DW < 4 - d_L$ , test no concluyente.
- ▶  $4 - d_L < DW < 4$ , mantenemos que  $\rho < 0$ , autocorrelación negativa.

Estadístico de Durbin-Watson - Puntos críticos de  $d_L$  y  $d_u$  al nivel de significación del 5%  
 $k^*$  corresponde al número de regresores del modelo excluido el término independiente (es decir,  $k^* = k - 1$ )

n	$k^* = 1$		$k^* = 2$		$k^* = 3$		$k^* = 4$		$k^* = 5$		$k^* = 6$	
	$d_L$	$d_u$										
6	0.610	1.400										
7	0.700	1.356	0.467	1.896								
8	0.763	1.332	0.559	1.777	0.368	2.287						
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588				
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822		
11	0.927	1.324	0.658	1.604	0.595	1.928	0.444	2.283	0.316	2.645	0.203	3.005
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.379	2.506	0.268	2.832
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390	0.328	2.692
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.472
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318

## H DE DURBIN

El contraste de Durbin-Watson es válido cuando la autocorrelación de los errores es autorregresiva de orden 1 y cuando la regresión no incluye entre las variables explicativas algún retardo de la variable dependiente.

En este segundo caso se recurre al contraste  $h$  de Durbin. Esto es, en modelos del tipo

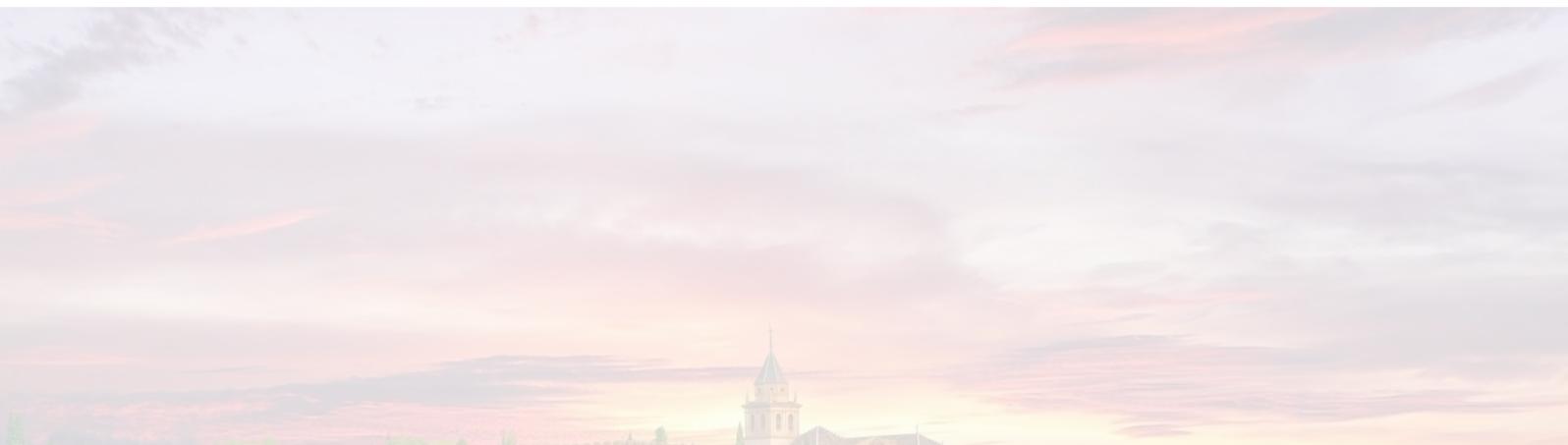
$$Y_t = \alpha Y_{t-1} + \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + u_t,$$

donde  $u_t = \rho u_{t-1} + v_t$ , para contrastar

$$H_0 : \rho = 0 \text{ (in correlación)}$$

$$H_1 : \rho \neq 0 \text{ (correlación)}$$

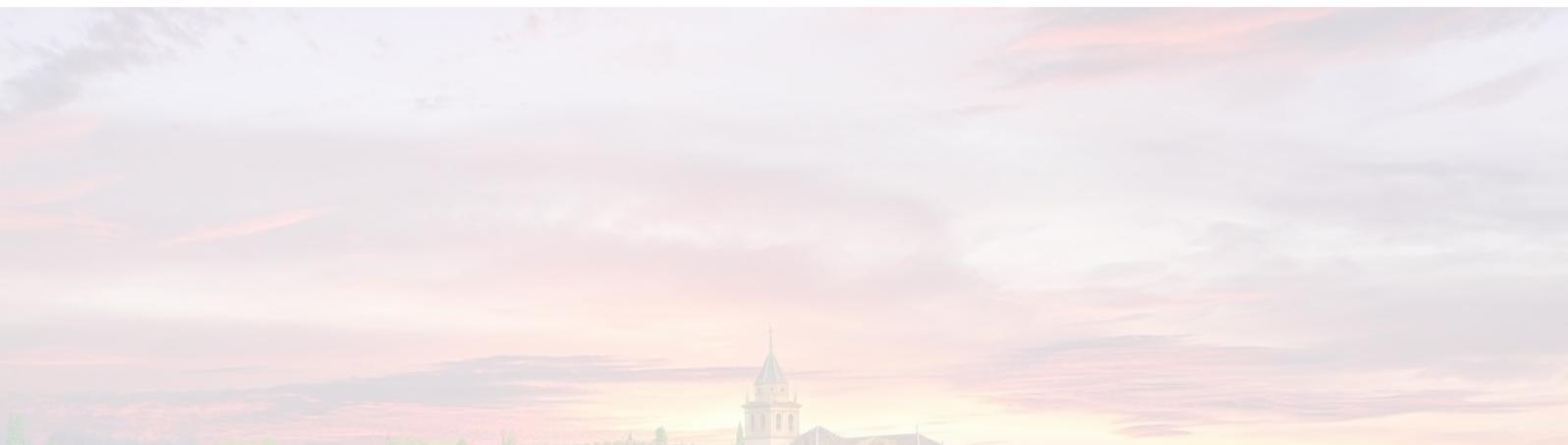
utilizaremos el estadístico  $h = \rho \cdot \sqrt{\frac{n}{1-n \cdot \widehat{\text{Var}}(\hat{\alpha})}} \sim N(0, 1)$ , de forma que se rechazará la hipótesis nula si  $|h| > Z_{1-\frac{\alpha}{2}}$ .



## CONTRASTE DE LJUNG-BOX

Si los residuos son independientes, sus primeras  $m$  autocorrelaciones son cero, para cualquier valor de  $m$ . El test de Ljung-Box contrasta la hipótesis nula de que las primeras  $m$  autocorrelaciones,  $\rho_m$ , son cero. Esto es:

$$\left. \begin{array}{l} H_0 : \rho_1 = \rho_2 = \dots = \rho_m = 0 \\ H_1 : \exists i \in \{1, 2, \dots, m\} \text{ tal que } \rho_i \neq 0 \end{array} \right\}$$



## CONTRASTE DE LJUNG-BOX

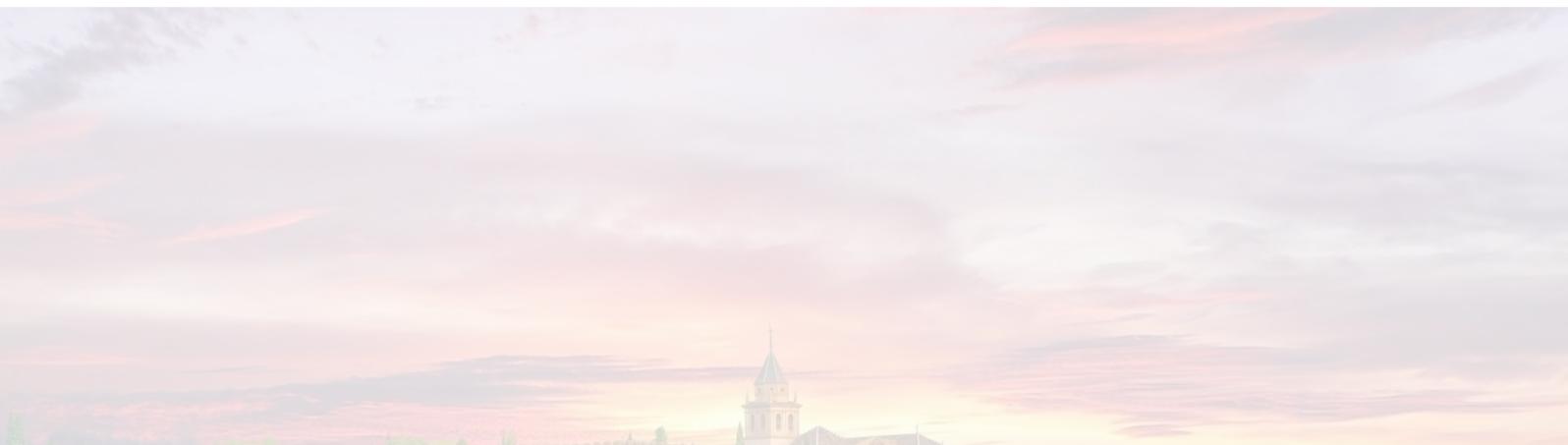
Se rechaza la hipótesis nula (de incorrelación) si  $Q_{LB} = n(n + 2) \sum_{s=1}^m \frac{r(s)^2}{n-s} > \chi_{m-1}^2(1 - \alpha)$  donde

$$r(s) = \frac{\sum_{t=s+1}^n e_t e_{t-s}}{\sum_{t=1}^n e_t^2},$$

es el coeficiente de autocorrelación muestral de orden  $k$ .

Si las observaciones son independientes (incorrelación), los coeficientes  $r(s)$  serán próximos a cero, por lo que no se rechazaría la hipótesis nula.

```
from statsmodels.stats.diagnostic import acorr_ljungbox
acorr_ljungbox(mco.resid, lags=3)
```



## EJEMPLO

Consideremos el modelo que ajusta el consumo energético en función del PIB para aplicar el anterior contraste para  $m = 3$ .

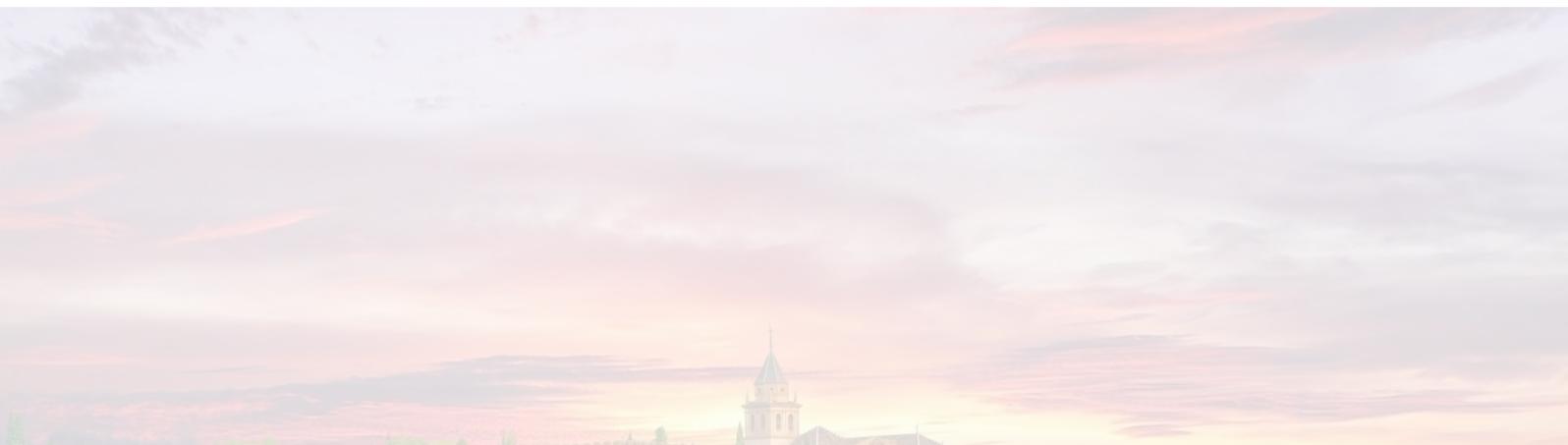
De este ejemplo ya sabemos que  $\sum_{t=1}^n e_t^2 = 765312'5$ . Además, a partir de los residuos podemos construir la siguiente tabla:

Año	$e_t$	$e_{t-1}$	$e_{t-2}$	$e_{t-3}$	$e_t \cdot e_{t-1}$	$e_t \cdot e_{t-2}$	$e_t \cdot e_{t-3}$
1987	494'1584						
1988	170'5190	494'1584			84263'40		
1989	-65'1749	170'5190	494'1584		-11113'56	-32206'724	
1990	-133'3305	-65'1749	170'5190	494'1584	8689'80	-22735'383	-65886'386
1991	-176'2899	-133'3305	-65'1749	170'5190	23504'82	11489'676	-30060'777
1992	-225'9164	-176'2899	-133'3305	-65'1749	39826'78	30121'546	14724'078
1993	40'1755	-225'9164	-176'2899	-133'3305	-9076'31	-7082'534	-5356'619
1994	46'9314	40'1755	-225'9164	-176'2899	1885'49	-10602'572	-8273'531
1995	8'5268	46'9314	40'1755	-225'9164	400'18	342'568	-1926'343
1996	-81'9462	8'5268	46'9314	40'1755	-698'74	-3845'849	-3292'229
1997	-348'7073	-81'9462	8'5268	46'9314	28575'23	-2973'357	-16365'321
1998	-255'0634	-348'7073	-81'9462	8'5268	88942'46	20901'476	-2174'874
1999	-58'7561	-255'0634	-348'7073	-81'9462	14986'52	20488'681	4814'839
2000	-25'8974	-58'7561	-255'0634	-348'7073	1521'63	6605'478	9030'612
2001	305'4278	-25'8974	-58'7561	-255'0634	-7909'78	-17945'746	-77903'453
2002	305'3431	305'4278	-25'8974	-58'7561	93260'26	-7907'592	-17940'769
Total					357073'5	-15350'3338	-200610'77

## EJEMPLO

A partir de la cual podemos obtener:

$$r(1) = \frac{\sum_{t=2}^n e_t \cdot e_{t-1}}{\sum_{t=1}^n e_t^2} = \frac{357073'5}{765312'5} = 0'4666,$$
$$r(2) = \frac{\sum_{t=3}^n e_t \cdot e_{t-2}}{\sum_{t=1}^n e_t^2} = \frac{-15350'3338}{765312'5} = -0'0201,$$
$$r(3) = \frac{\sum_{t=4}^n e_t \cdot e_{t-3}}{\sum_{t=1}^n e_t^2} = \frac{-200610'77}{765312'5} = -0'2622.$$



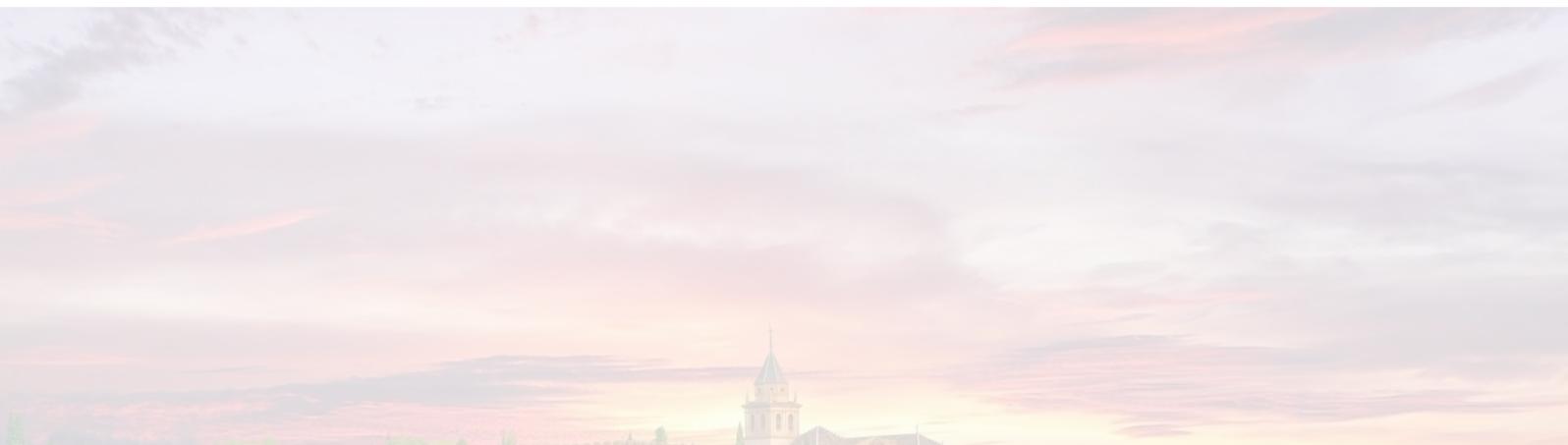
## EJEMPLO

De forma que:

$$\begin{aligned} Q_{LB} &= n(n+2) \sum_{s=1}^m \frac{r(s)^2}{n-s} = 16 \cdot 18 \cdot \left( \frac{r(1)^2}{15} + \frac{r(2)^2}{14} + \frac{r(3)^2}{13} \right) \\ &= 288 \cdot \left( \frac{0'21771556}{15} + \frac{0'00040401}{14} + \frac{0'06874884}{13} \right) \\ &= 288 \cdot (0'01452 + 0'000028 + 0'0053) = 288 \cdot 0'01983 \\ &= 5'7115. \end{aligned}$$

Como  $Q_{LB} = 5'7115 \not\geq 5'991 = \chi_2^2(0'95)$ , no rechazo la hipótesis nula de incorrelación.

Adviértase que para  $m = 2$ , se tiene que  $Q_{LB} = 4'1885 > 3'841 = \chi_1^2(0'95)$ . Por tanto, en tal caso si se rechazaría la hipótesis nula de incorrelación.



## ESTIMACIÓN BAJO AUTOCORRELACIÓN

$$y = X\beta + u$$

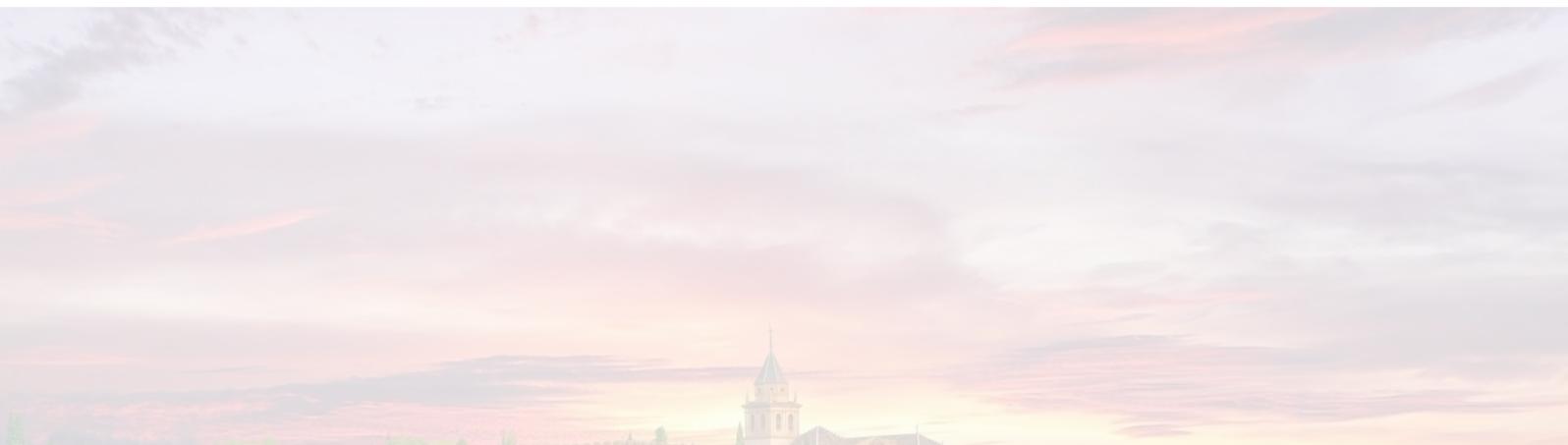
con:

- ▶  $\mathbb{E}[u] = 0$ ,
- ▶  $\mathbb{E}[u_i u_j] \neq 0$  ( $i \neq j$  con  $i, j \in \{1, \dots, n\}$ ) y
- ▶  $\text{var}[u] = \sigma^2 \Omega$ .

(no se verifican los supuestos de MCO)

Por el método de Mínimos Cuadrados Generalizado: necesitamos calcular  $P$  tal que  $\Omega^{-1} = P^t P$ .

Suponemos que  $u_t$  viene determinada por un proceso autorregresivo de orden 1,  $AR(1)$ :  $u_t = \rho u_{t-1} + \varepsilon_t$  ( $\varepsilon_t$  un ruido blanco) y  $\rho$  el coeficiente de autocorrelación con  $-1 < \rho < 1$ .



## ESTIMACIÓN BAJO AUTOCORRELACIÓN

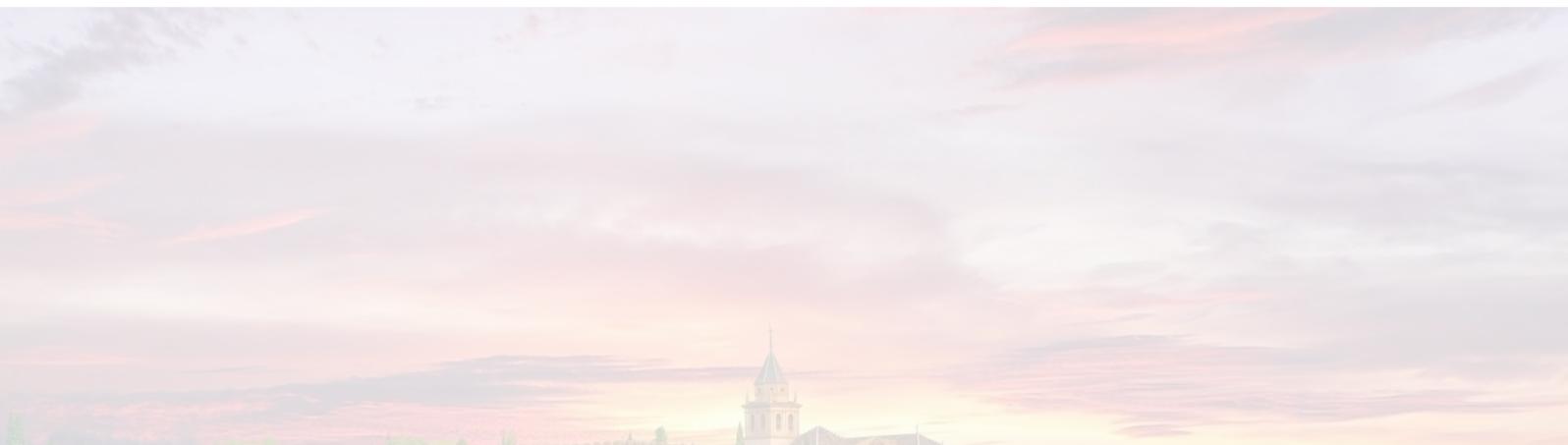
$$u_t \cdot u_{t-1} = (\rho u_{t-1} + \varepsilon_t) \cdot u_{t-1} = \rho u_{t-1}^2 + \varepsilon_t \cdot u_{t-1} \Rightarrow \mathbb{E}[u_t u_{t-1}] = \rho \sigma^2$$

De la misma forma:

$$\mathbb{E}[u_t u_{t-2}] = \mathbb{E}[(\rho u_{t-1} + \varepsilon_t) u_{t-2}] = \rho \mathbb{E}[u_{t-1} u_{t-2}] = \rho(\rho \sigma^2) = \rho^2 \sigma^2$$

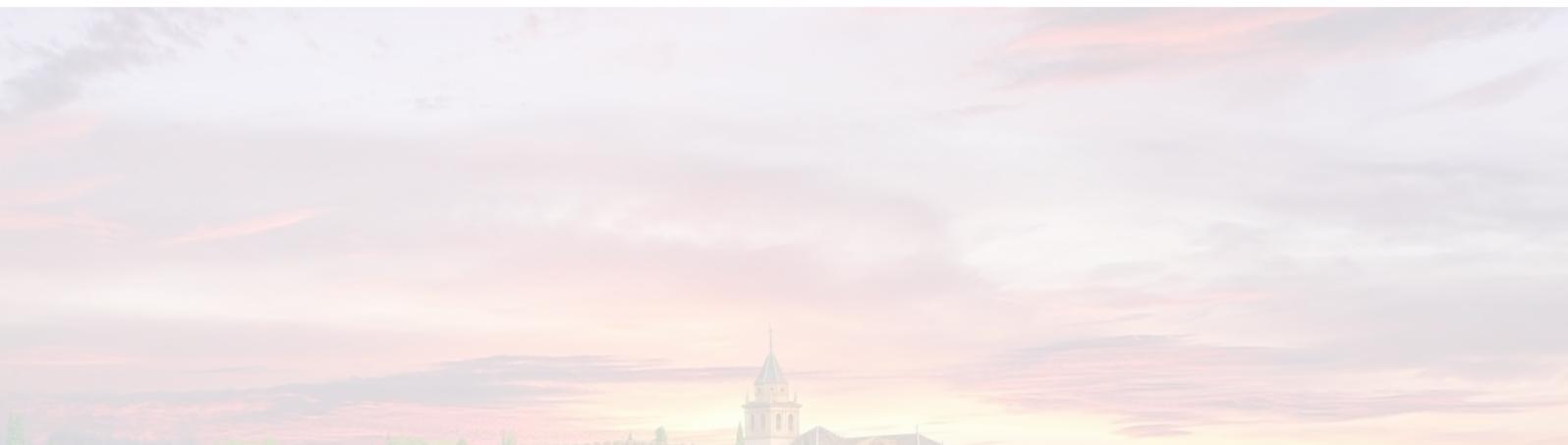
$$\mathbb{E}[u_t u_{t-3}] = \mathbb{E}[(\rho u_{t-1} + \varepsilon_t) u_{t-3}] = \rho \mathbb{E}[u_{t-1} u_{t-3}] = \rho(\rho^2 \sigma^2) = \rho^3 \sigma^2$$

$$\mathbb{E}[u_t u_{t-k}] = \rho^k \sigma^2$$
$$\Omega = \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \rho^2 & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \rho & \dots & \rho^{n-3} \\ & & \ddots & & & \\ \rho^{n-1} & \rho^{n-2} & \dots & \rho^2 & \rho & 1 \end{pmatrix}$$



## ESTIMACIÓN BAJO AUTOCORRELACIÓN: MCG

$$\Rightarrow \Omega^{-1} = \frac{1}{1 - \rho^2} \begin{pmatrix} 1 & -\rho & 0 & 0 & \dots & 0 \\ -\rho & 1 + \rho^2 & -\rho & 0 & \dots & 0 \\ 0 & -\rho & 1 + \rho^2 & -\rho & \dots & 0 \\ & & \ddots & & 1 + \rho^2 & -\rho \\ 0 & 0 & \dots & 0 & -\rho & 1 \end{pmatrix}$$



## ESTIMACIÓN BAJO AUTOCORRELACIÓN

Suponemos;

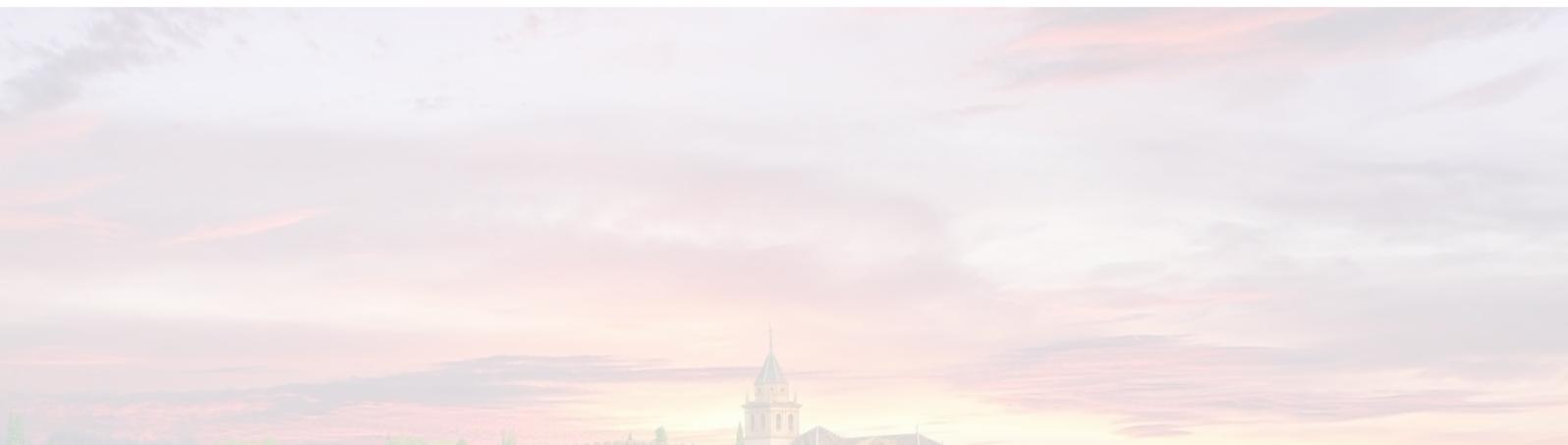
$$y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + u_t$$

y

$$y_{t-1} = \beta_1 + \beta_2 X_{2t-1} + \cdots + \beta_k X_{kt-1} + u_{t-1}$$

Luego:

$$y_t - \rho y_{t-1} = \beta_1(1 - \rho) + \beta_2(X_{2t} - \rho X_{2t-1}) + \cdots + \beta_k(X_{kt} - \rho X_{kt-1}) + u_t - \rho u_{t-1}$$



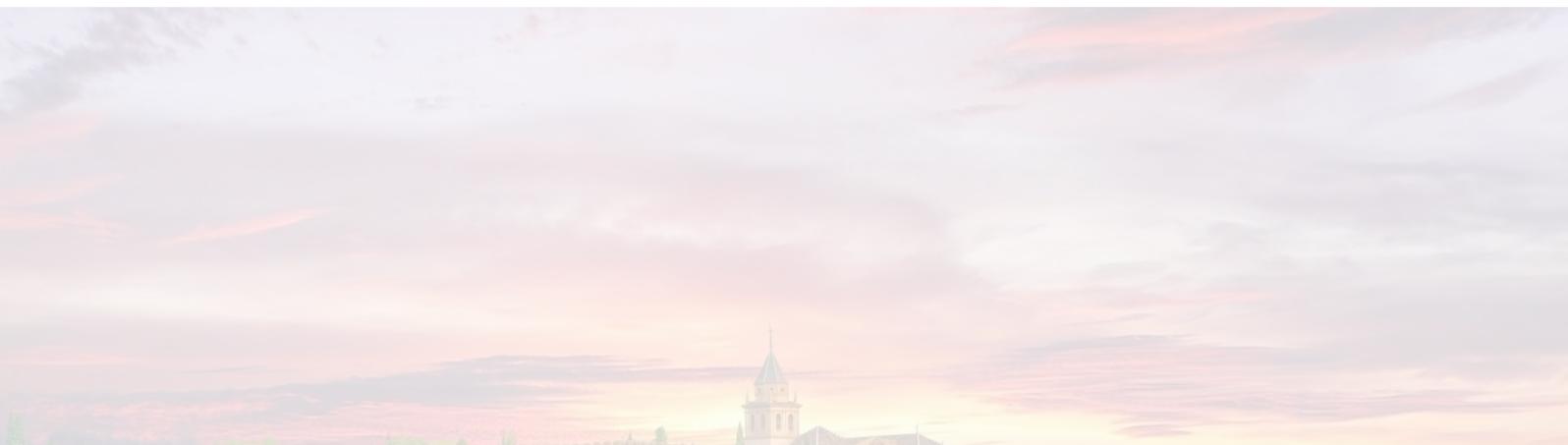
## ESTIMACIÓN BAJO AUTOCORRELACIÓN

Llamando  $y_t^* = y_t - \rho y_{t-1}$  y  $X_{it}^* = X_{it} - \rho X_{it-1}$ :

$$y_t^* = \beta_1 + \beta_2 X_{1t}^* + \dots + \beta_k X_{kt}^* + \varepsilon_t$$

donde ahora  $\varepsilon$  cumple las hipótesis para aplicar MCO.

**Nota:** Se perdería la primera observación, para ello Prais-Winsten proponen  $y_1^* = \sqrt{1 - \rho^2} y_1$ ,  $X_{i1}^* = \sqrt{1 - \rho^2} X_{i1} \dots$



## MODIFICACIÓN PRAIS-WINSTEN

1. Estimando por MCO y se obtienen los residuos  $e_t$ .
2. Se estima  $u_t = \rho u_{t-1} + \varepsilon_t$  el valor de  $\rho$  empleando los  $e_t$ :

$$\hat{\rho} = \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_t^2}$$

3. Se estima por MCO  $y^* = X^* \beta + u^*$ :  $e_t^*$ .
4. Estimamos  $\hat{\rho}$  de la regresión  $e_t^* = \hat{\rho} e_{t-1}^* + \varepsilon_t^*$ :  $\hat{\rho} = \frac{\sum_{t=2}^n e_t^* e_{t-1}^*}{\sum_{t=2}^n (e_t^*)^2}$ .
5. Repetimos... La estimación de  $\rho$  es el valor que estabiliza la secuencia  $\hat{\rho}, \hat{\hat{\rho}}, \hat{\hat{\hat{\rho}}}, \dots$  (hasta que la diferencia sea  $< 10^{-3}$ ).

**Proceso iterativo de Cochrane-Orcutt:** Obviamos la primera observación en las transformaciones.

```
sm.GLSAR(y, sm.add_constant(x), rho=rho).iterative_fit(maxiter
```

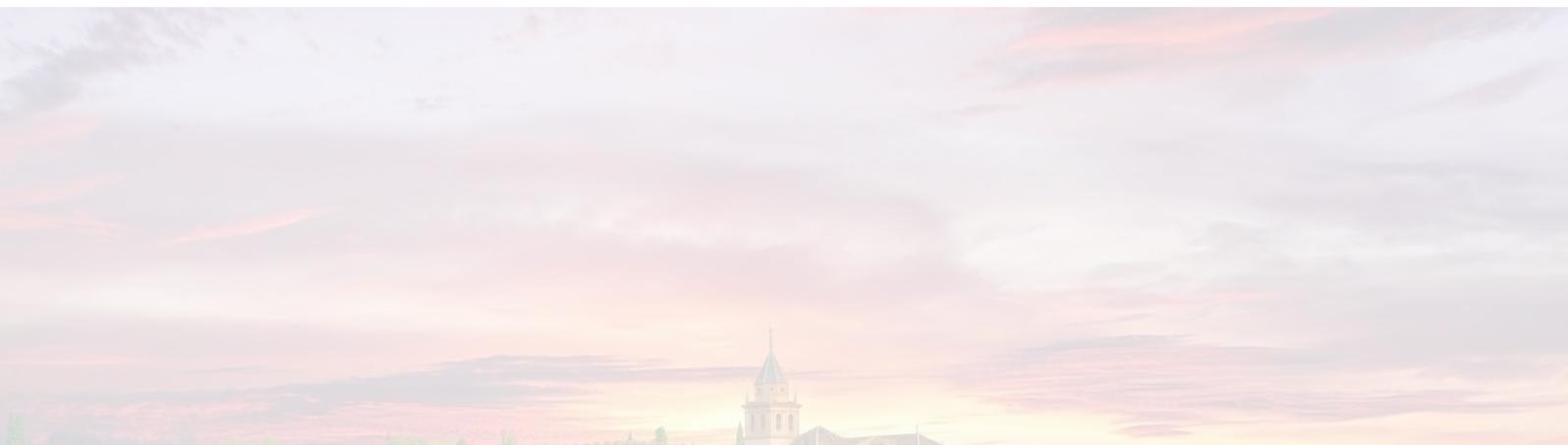
## EJEMPLO

El número de pequeños accidentes,  $Y$ , ocurridos en las calles de una ciudad y el número de coches que han sido matriculados,  $X$ , en la misma durante 10 años están recogidos en la siguiente tabla:

$Y$	$X$
25	510
27	520
28	528
32	540
33	590
36	650
38	700
40	760
41	800
45	870

Dado el modelo  $Y_t = \beta_1 + \beta_2 X_t + u_t$ :

- ▶ Contrastar la hipótesis de no autocorrelación por medio de Durbin-Watson.
- ▶ En caso de detectar problemas de autocorrelación, obtenga una estimación aplicando MCG.



## EJEMPLO

Se estima el modelo empleando MCO.

$$\hat{Y}_t = 2.56755 + 0.0493699X_t \quad (R^2 = 0.942095)$$

Y	$\hat{Y}$	$e_t$	$e_{t-1}$	$(e_t - e_{t-1})^2$	$e_t^2$
25	27.7462	-2.7462			7.5416
27	28.2399	-1.2399	-2.7462	2.2689	1.5374
28	28.6349	-0.6349	-1.2399	0.3661	0.4030
32	29.2273	2.7727	-0.6349	11.6115	7.6879
33	31.6958	1.3042	2.7727	2.1565	1.7010
36	34.6580	1.3420	1.3042	0.0014	1.8010
38	37.1265	0.8735	1.3420	0.2195	0.7630
40	40.0887	-0.0887	0.8735	0.9258	0.0079
41	42.0635	-1.0635	-0.0887	0.9502	1.1310
45	45.5194	-0.5194	-1.0635	0.2961	0.2697
				18.7960	22.8435

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} = \frac{18.7960}{22.8435} = 0.8228 \quad \begin{matrix} k=1, n=10 \\ \alpha=5\% \end{matrix} \left\{ \begin{array}{l} d_L = 0.879 \\ d_U = 1.320 \end{array} \right. \quad DW \leq d_L \left\{ \begin{array}{l} \text{Autocorrelaci3n} \\ \text{Positiva} \end{array} \right.$$

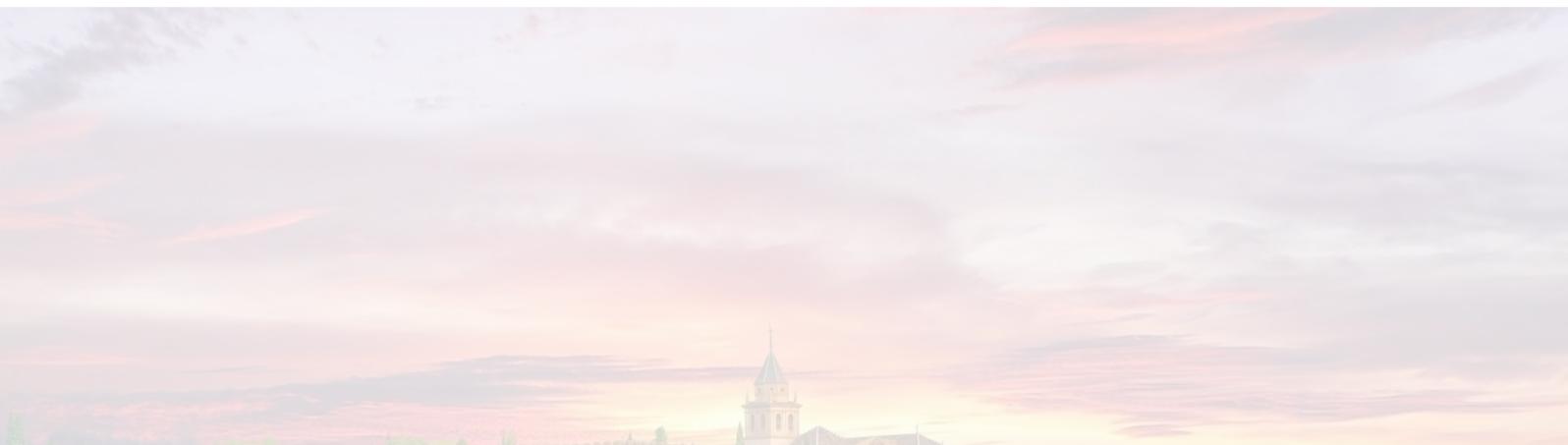
## EJEMPLO

Estima la regresión  $e_t = \rho e_{t-1} + \varepsilon_t$ :

$$\hat{e}_t = 0.4226e_{t-1} \Rightarrow \hat{\rho} = 0.4226$$

Como el  $n$  es pequeño usamos la Modificación Prais-Winsten (transformación):

$$y_t^* = \begin{cases} \sqrt{1 - (0.4226)^2} y_t & t = 1 \\ y_t - 0.4226 y_{t-1} & t > 1 \end{cases}$$
$$X_t^* = \begin{cases} \sqrt{1 - (0.4226)^2} X_t & t = 1 \\ X_t - 0.4226 X_{t-1} & t > 1 \end{cases}$$
$$cte_t^* = \begin{cases} \sqrt{1 - (0.4226)^2} cte_t & t = 1 \\ cte_t - 0.4226 cte_{t-1} & t > 1 \end{cases}$$



## EJEMPLO

$Y^*$	$cte^*$	$X^*$
22.61101004	0.904440402	461.2646048
16.335	0.5734	302.434
16.4818	0.5734	306.168
20.0552	0.5734	314.7552
19.3488	0.5734	359.636
21.9222	0.5734	398.306
22.6424	0.5734	422.71
23.7892	0.5734	461.38
23.936	0.5734	475.784
27.5094	0.5734	528.72

$$Y_t^* = \beta_3 + \beta_4 X_t^* + u_t^* \Rightarrow \hat{Y}^* = 2.1071 + 0.04975 X^*.$$

**Nota:** Como el valor de  $\rho$  ha sido estimado deberíamos realizar el proceso iterativo. Otra alternativa es calcular el estadístico de D-W y comprobar si se ha eliminado el problema de autocorrelación. En este caso  $DW = 1.5305$ , indicando pues que la autocorrelación ha desaparecido.

## 7 Referencias

- Diapositivas de clase.