

Lección 3

Marcos Bujosa

26 de julio de 2024

Índice

1. Correlación espuria entre la incidencia de melanoma en el estado de Connecticut y el PNB de EEUU	1
1.1. Actividad 1 - Dibujar ambas series en un mismo gráfico	1
1.2. Actividad 2 - Dibujar el diagrama de dispersión	2
1.3. Actividad 3 - Calcular la correlación entre ambas series	2
1.4. Actividad 4 - Regresar GNP sobre Melanoma y constatar que el ajuste es bueno	2
1.5. Actividad 5 - Añadir la primera diferencia de los datos	2
1.6. Actividad 6 - Regresar d_GNP sobre d_Melanoma y constatar que el ajuste es horrible	3
2. Tipos de interés a corto y largo plazo	4
2.1. Actividad 1 - Dibujar ambas series en un mismo gráfico	4
2.2. Actividad 2 - Dibujar el diagrama de dispersión y calcule la correlación	4
2.3. Actividad 3 - Regrese la primera diferencia de los tipos a corto sobre la diferencia de los tipos a largo	4
2.4. Actividad 4 - Ajuste los tipos corto plazo en función de los tipos a largo	5

1. Correlación espuria entre la incidencia de melanoma en el estado de Connecticut y el PNB de EEUU

Guión: [GNPvsMelanoma.inp](#)

- Objetivo Comprobar cómo la aparente relación entre ambas series temporales se desvanece al tomar primeras diferencias
- Carga de datos *Archivo -->Abrir datos -->Archivo de usuario* y buscar en el disco el fichero `GNPvsMelanoma.gdt`
o bien teclee en línea de comandos:

```
open datos/GNPvsMelanoma.gdt
```

1.1. Actividad 1 - Dibujar ambas series en un mismo gráfico

Marque las series `GNP` y `Melanoma`. Pulse sobre ellas con el botón derecho del ratón. En el menú desplegable seleccione *Gráfico de Series Temporales* (indique representar en un único gráfico).

Guarde el gráfico en la sesión como un icono.

o bien teclee en línea de comandos:

```
GraficoSeriesEnNiveles <- gnuplot GNP Melanoma --time-series --with-lines
```

1.2. Actividad 2 - Dibujar el diagrama de dispersión

Marque las series **GNP** y **Melanoma**. Pulse sobre ellas con el botón derecho del ratón. En el menú desplegable seleccione *Gráfico de dispersión XY* (elijas como variable del eje X **Melanoma** y marque suprimir la recta estimada).

Guarde el gráfico en la sesión como un icono.

o bien teclee en línea de comandos:

```
DiagramDispersion <- gnuplot GNP Melanoma --fit=none
```

1. ¿Tienen tendencia estas series temporales?
2. ¿Parece ser una tendencia común a ambas series?
3. Pero... ¿lo podemos saber con seguridad solo mirando el gráfico?

1.3. Actividad 3 - Calcular la correlación entre ambas series

Marque las series **GNP** y **Melanoma**. Pulse sobre ellas con el botón derecho del ratón. En el menú desplegable seleccione *Matriz de correlación*

o bien teclee en línea de comandos:

```
corr GNP Melanoma
```

1.4. Actividad 4 - Regresar GNP sobre Melanoma y constatar que el ajuste es bueno

Estime el modelo mediante los menús desplegables: *Modelo -> Mínimos Cuadrados Ordinarios*; indique a **Gretl** el regresando y regresor y pulse *Aceptar*.

o bien teclee en línea de comandos:

```
AjusteEnNiveles <- ols GNP 0 Melanoma
```

Aunque el coeficiente de determinación es muy elevado y los parámetros muy significativos, el modelo "*no vale una castaña*". Una forma de constatarlo es darse cuenta de que si fuera cierto que

$$\mathbf{y} = \beta_1 \mathbf{1} + \beta_2 \mathbf{x} + \mathbf{u}$$

Entonces también sería cierto que (y nótese que $\nabla \mathbf{1} = \mathbf{0}$)

$$\nabla \mathbf{y} = \beta_2 \nabla \mathbf{x} + \nabla \mathbf{u}$$

Consecuentemente, si \mathbf{y} corresponde al **GNP** y \mathbf{x} a **Melanoma**, al regresar la primera diferencia de **GNP** sobre la primera diferencia de **Melanoma** el ajuste debería indicar que el parámetro de la constante (β_1) no es significativo, pero si la pendiente (β_2). Veamos que ocurre justo lo contrario...

1.5. Actividad 5 - Añadir la primera diferencia de los datos

Selecione con el ratón la variable **GNP** y **Melanoma**. Luego pulse en el menú desplegable *Añadir* que aparece arriba, en el centro de la ventana principal de **Gretl**.

■ *Añadir -> Primeras diferencias de las variables seleccionadas*

Haga un gráfico con ambas series (verá que la tendencia ha desaparecido y que ya no se parecen entre sí).

Calcule también la correlación entre ambas series diferenciadas (recuerde que en un modelo lineal simple el cuadrado de dicha correlación es el coeficiente de determinación).

o bien teclee en línea de comandos:

```
diff GNP Melanoma
GraficoSeriesEnDiferencias <- gnuplot d_GNP d_Melanoma --time-series --with-lines
corr d_GNP d_Melanoma
```

1.6. Actividad 6 - Regresar d_GNP sobre d_Melanoma y constatar que el ajuste es horrible

Estime el modelo mediante los menús desplegables: *Modelo* -> *Mínimos Cuadrados Ordinarios*; indique a *Gretl* el regresando y regresor y pulse *Aceptar*.

o bien teclee en línea de comandos:

```
AjusteEnDiferencias <- ols d_GNP 0 d_Melanoma
```

Conclusión. Las variables GNP y Melanoma muestran una tendencia creciente, lo que conduce a un elevado coeficiente de correlación entre ellas; pero la tendencia ni es común, ni la correlación se puede atribuir a ninguna relación de causalidad entre ellas. La correlación en este caso es espuria (es decir, carece de sentido tratar de interpretarla).

Código completo de la práctica GNPvsMelanoma.inp

```
open datos/GNPvsMelanoma.gdt

GraficoSeriesEnNiveles <- gnuplot GNP Melanoma --time-series --with-lines

DiagramDispersion <- gnuplot GNP Melanoma --fit=none

corr GNP Melanoma

AjusteEnNiveles <- ols GNP 0 Melanoma

diff GNP Melanoma
GraficoSeriesEnDiferencias <- gnuplot d_GNP d_Melanoma --time-series --with-lines
corr d_GNP d_Melanoma

AjusteEnDiferencias <- ols d_GNP 0 d_Melanoma
```

2. Tipos de interés a corto y largo plazo

Guión: [UKinterestRates.inp](#)

1. Objetivo Ver que probablemente la correlación entre los tipos a corto y largo plazo no es espuria, y que ambas series probablemente están cointegrados.

2. Carga de datos *Archivo -->Abrir datos -->Archivo de usuario* y buscar en el disco el fichero `UKinterestRates`.
o bien teclee en línea de comandos:

```
open datos/UKinterestRates.gdt
```

2.1. Actividad 1 - Dibujar ambas series en un mismo gráfico

Marque las series `Long` y `Short`. Pulse sobre ellas con el botón derecho del ratón. En el menú desplegable seleccione *Gráfico de Series Temporales* (indique representar en un único gráfico).

Guarde el gráfico en la sesión como un icono.

o bien teclee en línea de comandos:

```
GraficoSeriesEnNiveles <- gnuplot Long Short --time-series --with-lines
```

1. ¿Tienen tendencia estas series temporales?
2. ¿Parece ser una tendencia común a ambas series?
3. Pero... ¿lo podemos saber con seguridad solo mirando el gráfico?

2.2. Actividad 2 - Dibujar el diagrama de dispersión y calcule la correlación

Marque las series `Long` y `Short`. Pulse sobre ellas con el botón derecho del ratón. En el menú desplegable seleccione *Gráfico de dispersión XY* (elija como variable del eje X `Long` y marque suprimir la recta estimada).

Guarde el gráfico en la sesión como un icono.

Calcule la correlación entre ambas series.

o bien teclee en línea de comandos:

```
DiagramDispersion <- gnuplot Short Long --fit=none  
corr Long Short
```

2.3. Actividad 3 - Regrese la primera diferencia de los tipos a corto sobre la diferencia de los tipos a largo

1. Incluya las primeras diferencias de `Short` y `Long`
2. Dibuje ambas series diferenciadas. ¿Parecen "ser estacionarias en media"? ¿Son `Short` y `Long` aparentemente $I(1)$?
3. ¿Están correladas?
4. Regrese `d_Short` sobre `d_Long`
5. Observe los resultados de la regresión.
 - ¿Son significativos los parámetros? ¿cuales sí y cuales no? (compare esto con lo que pasaba en el ejemplo anterior)
 - ¿Reproduce el modelo parte de la varianza de `d_Short`?

Realice los pasos con la interfaz gráfica y los menús desplegables, *o bien teclee en línea de comandos:*

```
diff Short Long  
GraficoSeriesEnDiferencias <- gnuplot d_Short d_Long --time-series --with-lines  
corr d_Short d_Long  
AjusteEnDiferencias <- ols d_Short 0 d_Long
```

2.4. Actividad 4 - Ajuste los tipos corto plazo en función de los tipos a largo

Veamos si las series en niveles pueden estar cointegradas. Para ello debemos analizar los residuos de la regresión de **Short** sobre **Long**.

1. Regrese **Short** sobre **Long**
2. Observe los resultados de la regresión.
 - ¿Son significativos los parámetros?
 - ¿Reproduce el modelo parte de la varianza de **Short**?
3. Dibuje los residuos de la regresión. ¿Parecen .estacionarios en media? ¿dicho de otra forma ¿muestran alguna tendencia?

Realice los pasos con la interfáz gráfica y los menús desplegables, o bien teclee en línea de comandos:

```
AjusteEnNiveles <- ols Short 0 Long
residuos = $uhat
GraficoResiduos <- gnuplot residuos --time-series --with-lines
```

Conclusión. Las variables **Short** y **Long** son no estacionarias (tienen tendencia), lo que conduce a un elevado coeficiente de correlación entre ellas; sus primeras diferencias parecen .estacionarias"lo que sugiere que ambas series son $I(1)$.

La regresión de las series en diferencias y los residuos de la regresión en niveles parecen compatibles con que **Short** y **Long** estén cointegradas, es decir, que tengan una tendencia común.

En este caso la correlación no parece ser espuria, por lo que cabe interpretar dicha correlación y pensar que hay relación entre los tipos a corto y a largo plazo.

A este análisis le falta la realización de contrastes estadísticos que confirmen la estacionariedad de las series diferenciadas y de los residuos de la última regresión.

Código completo de la práctica UKinterestRates.inp

```
open datos/UKinterestRates.gdt

GraficoSeriesEnNiveles <- gnuplot Long Short --time-series --with-lines

DiagramDispersion <- gnuplot Short Long --fit=none
corr Long Short

diff Short Long
GraficoSeriesEnDiferencias <- gnuplot d_Short d_Long --time-series --with-lines
corr d_Short d_Long
AjusteEnDiferencias <- ols d_Short 0 d_Long

AjusteEnNiveles <- ols Short 0 Long
residuos = $uhat
GraficoResiduos <- gnuplot residuos --time-series --with-lines
```