

## Soluciones a la práctica 5.2 con R

En esta práctica vamos a utilizar los operadores `abline`, `head`, `linearHypothesis`, `lm`, `str` y `summary`. Para utilizar el operador `linearHypothesis` debemos instalar y cargar el paquete `car`. Con `lm` utilizamos `subset`.

En el modelo de Solow la dinámica de la transición implica que cuanto más alejada esté la renta per cápita de su propio estado estacionario más rápido crecerá y viceversa. Esto implica que si el estado estacionario es el mismo, los países pobres crecerán más rápido que los más ricos.

Para contrastar este hecho se propone analizar, a partir de datos de la Penn World Table 8.1, la relación entre la renta real per cápita del año 1960 (`pibrpc_60`) y la tasa de crecimiento media de la renta per cápita del período 1960-2007 (`pibrpc_tcm`), para dos bloques de países con grado de desarrollo homogéneo para cada bloque pero muy diferente entre bloques (países ricos de la OCDE y países pobres a nivel global, según clasificación del Banco Mundial). El fichero `practica52.RData` contiene observaciones de las variables `pibrpc_60` y `pibrpc_tcm`. También incluye una variable ficticia que informa si el país es rico o pobre (variable `estado`, que toma valor 1 para países ricos y 0 para países pobres).

El comportamiento de la dinámica de la transición lo podemos representar mediante la siguiente ecuación econométrica:

$$pibrpc\_tcm = \beta_0 + \beta_1 pibrpc\_60 + \varepsilon \quad (1)$$

Antes de resolver esta práctica examinamos el tipo de datos que contiene el fichero `Practica52.RData`. Para ello ejecutamos el código `str(Practica52)` y obtenemos la salida de la Figura 5.2.1:

Figura 5.2.1: `str(practica52)`

```
Classes 'tbl_df', 'tbl' and 'data.frame':   36 obs. of  4 variables:
 $ COUNTRY   : chr  "Japan" "Portugal" "Greece" "Israel" ...
 $ ESTADO    : num  1 1 1 1 1 1 1 1 1 1 ...
 $ PIBRPC_60 : num  3256 3623 4123 4937 5025 ...
 $ PIBRPC_TCM: num  5.28 3.93 3.82 3.58 3.98 ...
```

Podemos ver que la variable `COUNTRY` es una variable cualitativa que nos informa de los países que hay en nuestra muestras, mientras `ESTADO`, `PIBRPC_60` y `PIBRPC_TCM` son variables numéricas. `ESTADO` es una variable ficticia que toma el valor 1 si un país es rico y 0 si es pobre. En la Figura 5.2.1 también podemos ver las primeras observaciones de las variables. Otra manera de obtener las primeras (seis) observaciones es con el código `head(practica52)`, como podemos ver en la Figura 5.2.2.

Figura 5.2.2: `head(practica52)`

	COUNTRY	ESTADO	PIBRPC_60	PIBRPC_TCM
1	Japan	1	3256.183	5.276010
2	Portugal	1	3623.160	3.934464
3	Greece	1	4122.978	3.815365
4	Israel	1	4937.398	3.578726
5	Spain	1	5025.129	3.979855
6	Ireland	1	5958.408	4.493003

**1) Estime la ecuación (1) para el conjunto de la muestra. ¿Se cumple la dinámica de la transición?**

La dinámica de la transición del modelo de Solow implica que  $\beta_1 < 0$ . Es decir, cuanto más alta es la renta inicial de un país más cerca estará de su estado estacionario y, por tanto, más lento crecerá, es decir, más pequeña será su tasa de crecimiento media. La estimación con toda la muestra aparece en la Figura 5.2.3 y se realiza con el siguiente código:

```
mod.1 <- lm(PIBRPC_TCM~PIBRPC_60, data = practica52)
```

Figura 5.2.3: summary(mod.1)

```
Call:
lm(formula = PIBRPC_TCM ~ PIBRPC_60, data = practica52)

Residuals:
    Min       1Q   Median       3Q      Max
-2.9311 -0.7958 -0.2063  0.5836  3.3882

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.602e+00  3.341e-01  4.795 3.16e-05 ***
PIBRPC_60   8.780e-05  3.941e-05  2.228  0.0326 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.287 on 34 degrees of freedom
Multiple R-squared:  0.1274,    Adjusted R-squared:  0.1017
F-statistic: 4.963 on 1 and 34 DF, p-value: 0.03262
```

Observando los resultados, no parece cumplirse la dinámica de transición; la pendiente de la ecuación es prácticamente 0, y es positiva y significativa al 5%. Por otro lado, el  $R^2$  de la estimación es bastante bajo.

Pero, ¿por qué ocurre esto? Es posible que sea debido a que el modelo propuesto no es adecuado para estimar la dinámica de transición cuando el grupo de países no es homogéneo. Para poder contrastar el modelo de Solow con la ecuación propuesta necesitamos trabajar con países que tengan un estado estacionario similar. Eso ocurre cuando trabajamos sólo con los países ricos o sólo con los países pobres.

**2) Estime ahora la ecuación (1) por separado para países ricos y países pobres. ¿Existen grandes diferencias entre las dos ecuaciones estimadas? Interprete los coeficientes de cada uno de los dos modelos. ¿Se cumple ahora la hipótesis de la dinámica de la transición?**

Para hacer una regresión para países ricos y otra para países pobres vamos a hacer uso de la información contenida en la variable *estado* que nos indica el tipo de país. Primero, estimamos el modelo con los datos de los países ricos. Para ello podemos ejecutar uno de los dos códigos siguientes:

```
mod.ricos <- lm(PIBRPC_TCM~PIBRPC_60, data =
subset(practica52,ESTADO=="I"))
```

```
mod.ricos <- lm(PIBRPC_TCM~PIBRPC_60, data = practica52, subset =
(ESTADO == "I"))
```

Los dos código anteriores guardan la estimación del modelo para países ricos en el objeto *mod.ricos*. El resultado de la estimación se muestra en la siguiente figura.

Figura 5.2.4: summary(mod.ricos)

```
Call:
lm(formula = PIBRPC_TCM ~ PIBRPC_60, data = practica52, subset = (ESTADO ==
"1"))

Residuals:
    Min       1Q   Median       3Q      Max
-0.7730 -0.2817 -0.1208  0.1263  1.3665

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.649e+00  2.908e-01  15.985 3.14e-13 ***
PIBRPC_60   -1.729e-04  2.747e-05  -6.294 3.05e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5423 on 21 degrees of freedom
Multiple R-squared:  0.6535,    Adjusted R-squared:  0.637
F-statistic: 39.61 on 1 and 21 DF,  p-value: 3.053e-06
```

El código para estimar el modelo para los países pobres es similar a los códigos anteriores porque en este caso hay que asignar a la variable *estado* el valor 0. En concreto, el código para estimar este modelo y guardar los resultados en el objeto *mod.pobres* es el siguiente:

```
mod.pobres <- lm(PIBRPC_TCM~PIBRPC_60, data =
subset(practica52,ESTADO=="0"))
```

Figura 5.2.5: summary(mod.pobres)

```
Call:
lm(formula = PIBRPC_TCM ~ PIBRPC_60, data = subset(practica52,
ESTADO == "0"))

Residuals:
    Min       1Q   Median       3Q      Max
-1.02822 -0.17056  0.09689  0.15420  0.56180

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.0383912  0.2738201   7.444 1.29e-05 ***
PIBRPC_60   -0.0017234  0.0003356  -5.135 0.000326 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4326 on 11 degrees of freedom
Multiple R-squared:  0.7057,    Adjusted R-squared:  0.6789
F-statistic: 26.37 on 1 and 11 DF,  p-value: 0.0003255
```

La primera conclusión que podemos extraer cuando se analizan las dos ecuaciones es el cumplimiento de la dinámica de la transición, ya que *la pendiente de cada una de las ecuaciones es negativa y altamente significativa*.

Como podemos ver en las estimaciones, la diferencia en la tasa de crecimiento media entre dos países ricos con un diferencial unitario de renta per cápita es igual a 0.0001729 puntos porcentuales, mientras que en el grupo de países pobres es de 0.0017234. La disparidad que se observa entre las dos ecuaciones estimadas es bastante grande. Este hecho se explica por la gran desigualdad de renta per cápita existente entre los dos grupos de países. En los países pobres el mismo diferencial positivo de renta per cápita te acerca más rápido al estado estacionario y, por tanto, la caída en la tasa de crecimiento media debe ser mucho mayor que en los países ricos.

**3) Especifique un modelo que permita contrastar la hipótesis de Solow utilizando la muestra completa.**

Si queremos trabajar con toda la muestra al mismo tiempo tenemos que incluir alguna variable que nos permita discriminar entre un grupo de países y el otro. En nuestro caso, esta variable es la ficticia *estado*.

El modelo con la variable ficticia *estado* aditiva y multiplicativa permitiría contrastar la hipótesis de Solow con la muestra completa de países:

$$pibrpc\_tcm = \beta_0 + \delta_0 estado + \beta_1 pibrpc\_60 + \delta_1 pibrpc\_60 \cdot estado + \varepsilon$$

**4) Estime el modelo propuesto en el apartado anterior y contraste la igualdad de los parámetros de la ecuación para países ricos y para países pobres ¿Se debería rechazar la hipótesis nula?**

La estimación de este modelo, que guardamos en el objeto *mod.2*, aparece en la Figura 5.2.6 y se obtiene con el código:

```
mod.2 <- lm(PIBRPC_TCM~ESTADO + PIBRPC_60 + PIBRPC_60*ESTADO,
data = practica52)
```

Figura 5.2.6: summary(mod.2)

```
Call:
lm(formula = PIBRPC_TCM ~ ESTADO + PIBRPC_60 + PIBRPC_60 * ESTADO,
    data = practica52)

Residuals:
    Min       1Q   Median       3Q      Max
-1.02822 -0.27845 -0.07297  0.16557  1.36647

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.0383912   0.3210734   6.349 3.97e-07 ***
ESTADO       2.6106590   0.4208398   6.203 6.03e-07 ***
PIBRPC_60   -0.0017234   0.0003935  -4.380 0.000119 ***
ESTADO:PIBRPC_60  0.0015505   0.0003943   3.932 0.000424 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5073 on 32 degrees of freedom
Multiple R-squared:  0.8725,    Adjusted R-squared:  0.8606
F-statistic:  73 on 3 and 32 DF,  p-value: 2.112e-14
```

En el modelo con ficticias, los coeficientes  $\delta_0$  y  $\delta_1$  nos miden los diferenciales de la constante y la pendiente, respectivamente, de la ecuación de los países ricos respecto a la de los países pobres. Más concretamente,  $\hat{\delta}_1$  mide el diferencial del coeficiente de pendiente de la ecuación de los países ricos respecto de la de los países pobres. Así, el valor 0.0015505 indica que la diferencia entre los dos coeficientes de pendiente es de 0.0015505 puntos porcentuales. De hecho, el valor del coeficiente  $\hat{\beta}_1$  de la ecuación de países pobres, categoría de referencia, es -0.0017234, siendo el de la ecuación de países ricos igual a  $-0.0001729 = -0.0017234 + 0.0015505$ . Obsérvese que los resultados coinciden con las estimaciones por submuestras realizadas en el apartado 2).

De esta manera, para contrastar la igualdad de los parámetros de las dos ecuaciones bastará con contrastar la igualdad a cero de los diferenciales. En concreto,  $H_0: \delta_0 = \delta_1 = 0$ ,  $H_1: \text{no } H_0$ .

Este contraste se conoce como **contraste de cambio estructural**, y para llevarlo a cabo utilizaremos un estadístico F. El modelo restringido del contraste es aquel en el que se supone la ausencia de cambio estructural, es decir, el modelo que se estimó en el apartado 1) y cuyo  $R^2$  es 0.1274. El modelo no restringido es el que se ha estimado en este apartado, cuyo  $R^2$  es igual a 0.8725. Dado que  $SCE = (1 - R^2)STC$ , podemos obtener que  $SCE_R = 56.35937$  y  $SCE_{NR} = 8.234492$ . Introduciendo estos valores en el estadístico de contraste, tenemos:

$$F = \frac{\frac{SCE_R - SCE_{NR}}{SCE_{NR}}}{\frac{q}{N - 2K}} = \frac{\frac{56.35937 - 8.234492}{8.234492}}{\frac{2}{36 - 4}} = 93.509$$

Otra forma de obtener el valor del estadístico F es a través de la siguiente formula:

$$F = \frac{\frac{R_{NR}^2 - R_R^2}{1 - R_{NR}^2}}{\frac{2}{N - 2K}} = \frac{\frac{0.8725 - 0.1274}{1 - 0.8725}}{\frac{2}{36 - 4}} = 93.509$$

Para llevar a cabo este contraste directamente con R utilizamos el operador *linearHypothesis* y obtenemos los resultados de la siguiente figura:

Figura 5.2.7: `linearHypothesis(mod.2,c("ESTADO=0","ESTADO:PIBRPC_60=0"),test="F")`  
Linear hypothesis test

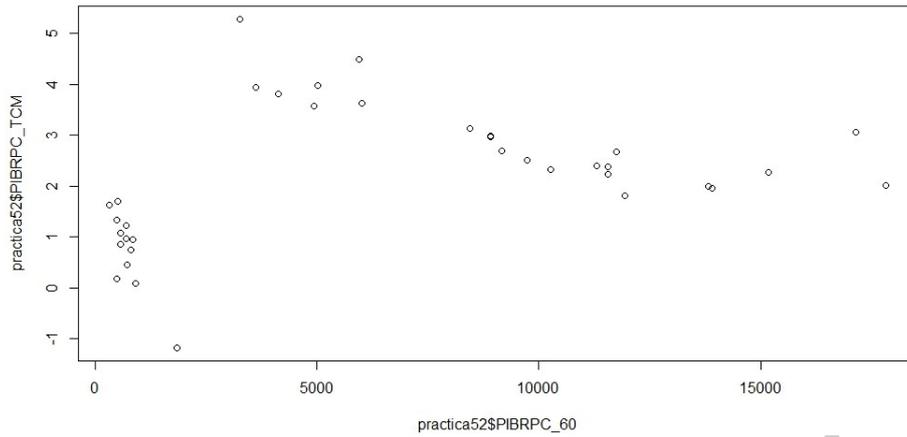
```
Hypothesis:
ESTADO = 0
ESTADO:PIBRPC_60 = 0

Model 1: restricted model
Model 2: PIBRPC_TCM ~ ESTADO + PIBRPC_60 + PIBRPC_60 * ESTADO

Res.Df  RSS Df Sum of Sq    F    Pr(>F)
1      34 56.359
2      32  8.234  2    48.125 93.509 4.313e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Como se obtiene un p-valor de 0.0000, se concluye que la hipótesis nula debe ser rechazada a favor de la alternativa; las variables ficticias del modelo son conjuntamente significativas, es decir, al menos una de las variables ficticias incluidas en el modelo es significativa. Como se aprecia en la ecuación estimada, también son individualmente significativas. Las ecuaciones de países ricos y países pobres son, por tanto, diferentes.

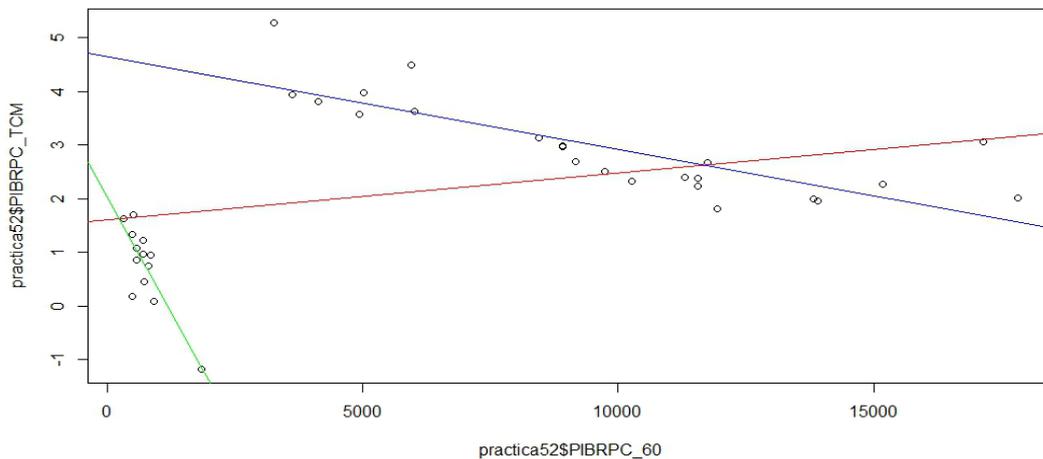
**Nota:** Una manera gráfica de ver el resultado tan dispar que observamos entre estimar con toda la muestra o separando entre países ricos y pobres es obtener el diagrama de dispersión entre las variables *pibrpc\_60* y *pibrpc\_tcm*. Dicho diagrama de dispersión aparece en la Figura 5.2.8.

Figura 5.2.8: `plot(practica52$PIBRPC_60, practica52$PIBRPC_TCM)`

A continuación utilizamos el operador *abline* para añadir la representación gráfica de las rectas de regresión obtenidas para toda la muestra, objeto *mod.1*, para los países pobres, objeto *mod.pobres*, y para los países ricos, objeto *mod.ricos*. Para ejecutarlos los siguientes códigos:

```
abline(mod.1, col="red")
abline(mod.ricos, col="blue")
abline(mod.pobres, col="green")
```

A medida que se ejecute cada línea de código se añadirá al gráfico de dispersión una recta de regresión. En rojo se representa la recta de regresión obtenida para toda la muestra, en azul se representa la recta de regresión obtenida con los datos de los países ricos y en verde aquella obtenida con los datos de los países pobres.

Figura 5.2.9: `abline`

Vemos, por tanto, como la recta de regresión cuando usamos toda la muestra tiene pendiente positiva, tal y como hemos estimado en el apartado 1). Sin embargo, cuando utilizamos de forma separada sólo la información referente a los países ricos y a los países pobres, el resultado cambia completamente, tal y como hemos visto en las estimaciones del apartado 2). En estos casos, las pendientes son negativas.