
Primera práctica de evaluación

Autor

Manuel DE LA LLAVE*

9 de Febrero de 2020

Índice general

Ejercicio 1	2
Ejercicio 2	3
Ejercicio 3	5
Ejercicio 4	5
Ejercicio 5	7
Ejercicio 6	9
Código	10

Considera el modelo de tres factores:

$$HiTec - R_f = \beta_1 + \beta_2(R_m - R_f) + \beta_3smb + \beta_4hml + u \quad (0.1)$$

Ejercicio 1

Estima la regresión lineal y muestra los resultados. Interpreta cada uno de los coeficientes del modelo.

Dependent Variable: HITEC-RF

Method: Least Squares

Date: 09/02/20 Time: 15:17

Sample: 1926M07 2015M12

Included observations: 1074

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.138453	0.063893	2.166956	0.0305
RMRF	0.986237	0.012658	77.91150	0.0000
SMB	0.041632	0.020747	2.006691	0.0450
HML	-0.319994	0.018407	-17.38404	0.0000
R-squared	0.864345	Mean dependent var	0.663892	
Adjusted R-squared	0.863965	S.D. dependent var	5.615646	
S.E. of regression	2.071217	Akaike info criterion	4.297867	
Sum squared resid	4590.235	Schwarz criterion	4.316411	
Log likelihood	-2303.955	Hannan-Quinn criter.	4.304891	
F-statistic	2272.552	Durbin-Watson stat	1.911409	
Prob(F-statistic)	0.000000			

Primero debemos considerar **qué** estamos midiendo y cómo lo estamos haciendo para interpretar correctamente los coeficientes. Nuestra variable dependiente es el exceso de rendimiento entre una cartera de bienes de equipo y comunicación frente al activo libre de riesgo, y está medida en términos porcentuales, así como el resto de variables explicativas de nuestro modelo.

Los coeficientes nos indican que un aumento del 1% del exceso de rendimiento del mercado supondría un aumento del 0.98%, +0.041% en el caso de que aumentara la diferencia de rendimiento entre las carteras pequeñas y las grandes y -0.31% si el rendimiento de las carteras de valor menos el de las carteras de crecimiento aumenta.

Todos los coeficientes son estadísticamente significativos (aunque β_3 por poco) y el R^2 es relativamente alto, por lo que a priori parece una buena estimación, en el ejercicio 2 comprobaremos si las hipótesis de partida del modelo MCO se cumplen o no.

Ejercicio 2

Haz una validación completa de las hipótesis de incorrelación, homocedasticidad y normalidad de los residuos.

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.492096	Prob. F(2,1068)	0.2254
Obs*R-squared	2.992594	Prob. Chi-Square(2)	0.2240

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 09/02/20 Time: 15:17

Sample: 1926M07 2015M12

Included observations: 1074

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	$-1.41E-05$	0.063864	-0.000220	0.9998
RMRF	-0.000632	0.012658	-0.049899	0.9602
SMB	0.003557	0.020840	0.170656	0.8645
HML	-0.001142	0.018421	-0.062015	0.9506
RESID(-1)	0.042788	0.030706	1.393450	0.1638
RESID(-2)	0.029612	0.030659	0.965846	0.3343
R-squared	0.002786	Mean dependent var	$1.04E-16$	
Adjusted R-squared	-0.001882	S.D. dependent var	2.068319	
S.E. of regression	2.070265	Akaike info criterion	4.298801	
Sum squared resid	4577.445	Schwarz criterion	4.326618	
Log likelihood	-2302.456	Hannan-Quinn criter.	4.309337	
F-statistic	0.596838	Durbin-Watson stat	2.000566	
Prob(F-statistic)	0.702420			

Dado que el *p-value* (Probability) es mayor que el 5% (medida generalmente aceptada para aceptar/rechazar este tipo de contrastes) no podemos rechazar la hipótesis nula, que en este test supone que no hay autocorrelación en los residuos. Esto significa que nuestro término de error no está correlacionado consigo mismo.

Heteroskedasticity Test: White

F-statistic	23.93199	Prob. F(3,1070)	0.0000
Obs*R-squared	67.53296	Prob. Chi-Square(3)	0.0000
Scaled explained SS	155.4673	Prob. Chi-Square(3)	0.0000

Test Equation:

Dependent Variable: RESID²

Method: Least Squares

Date: 09/02/20 Time: 15:17

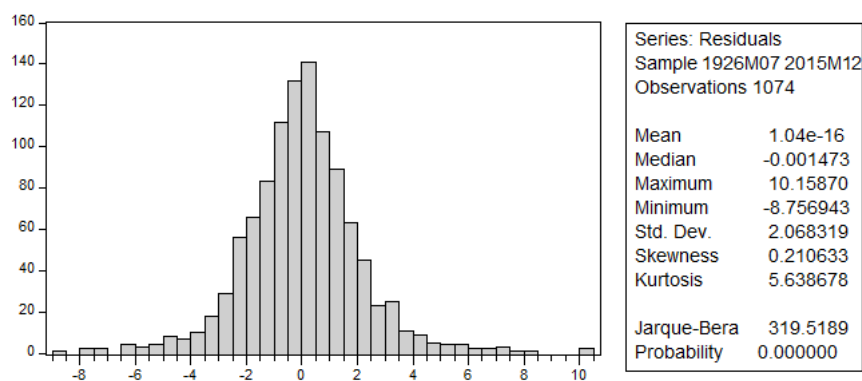
Sample: 1926M07 2015M12

Included observations: 1074

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.556009	0.288968	12.30587	0.0000
RMRF ²	0.016701	0.004055	4.118368	0.0000
SMB ²	0.002058	0.006015	0.342045	0.7324
HML ²	0.016281	0.006609	2.463255	0.0139
R-squared	0.062880	Mean dependent var	4.273962	
Adjusted R-squared	0.060252	S.D. dependent var	9.209372	
S.E. of regression	8.927619	Akaike info criterion	7.219894	
Sum squared resid	85281.54	Schwarz criterion	7.238438	
Log likelihood	-3873.083	Hannan-Quinn criter.	7.226918	
F-statistic	23.93199	Durbin-Watson stat	1.354465	
Prob(F-statistic)	0.000000			

En este caso, sí que rechazamos la hipótesis nula, por lo que la varianza de los residuos es heterocedástica (no es constante). Esto implica que los estimadores obtenidos previamente son ineficientes, ya que no podemos garantizar que la varianza es mínima.

Normalidad:



Aquí también rechazamos la hipótesis nula que plantea el estadístico Jarque-Bera sobre la normalidad de los residuos; es decir, nuestro término de error no se distribuye siguiendo una Ley normal.

Ejercicio 3

Contrasta mediante la prueba de Wald si el modelo con un factor es suficiente $H_0 : \beta_3 = \beta_4 = 0$

Wald Test:
Equation: EQ1

Test Statistic	Value	df	Probability
F-statistic	151.9302	(2, 1070)	0.0000
Chi-square	303.8604	2	0.0000

Null Hypothesis: $C(3) = C(4) = 0$
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(3)	0.041632	0.020747
C(4)	-0.319994	0.018407

Restrictions are linear in coefficients.

Podemos **rechazar** la hipótesis nula acerca de que ambos coeficientes son 0, y por tanto podemos decir que el modelo con un solo factor no es suficiente para explicar nuestra variable dependiente (necesitamos, al menos, los tres factores).

Ejercicio 4

Estamos interesados en analizar el posible efecto no lineal de *smb* sobre el exceso de rendimiento de la cartera de Bienes de equipo y Comunicación (*HiTec-Rf*). Añade en el modelo la variable *smb* al cuadrado e identifica el nuevo efecto marginal de *smb* sobre el consumo. Contrasta si efectivamente este término cuadrático es necesario y explica las consecuencias que se derivan del resultado del contraste.

Dependent Variable: HITEC-RF
Method: Least Squares
Date: 09/02/20 Time: 15:17
Sample: 1926M07 2015M12
Included observations: 1074

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.101179	0.064784	1.561796	0.1186
RMRF	0.986607	0.012609	78.24507	0.0000
SMB	0.012106	0.022784	0.531339	0.5953
HML	-0.330318	0.018639	-17.72158	0.0000
SMB^2	0.004536	0.001474	3.077388	0.0021
R-squared	0.865536	Mean dependent var		0.663892
Adjusted R-squared	0.865033	S.D. dependent var		5.615646
S.E. of regression	2.063067	Akaike info criterion		4.290909
Sum squared resid	4549.927	Schwarz criterion		4.314090
Log likelihood	-2299.218	Hannan-Quinn criter.		4.299689
F-statistic	1720.274	Durbin-Watson stat		1.885307
Prob(F-statistic)	0.000000			

Al añadir smb^2 podemos observar como la variable smb pierde significatividad, lo que quiere decir que el efecto que aportaba ya viene recogido por la nueva variable, que sí es que es significativa. Además, todos los criterios de información (Akaike, Schwarz y Hannan-Quinn) disminuyen mientras que el R^2 aumenta, lo que nos indica que la estimación ha mejorado.

Si realizamos el test de Wald para $H_0 : \beta_5 = 0$ veremos que debemos rechazar la hipótesis nula, confirmando lo que comentábamos en el apartado anterior.

Wald Test:
Equation: EQ2

Test Statistic	Value	df	Probability
t-statistic	3.077388	1069	0.0021
F-statistic	9.470314	(1, 1069)	0.0021
Chi-square	9.470314	1	0.0021

Null Hypothesis: $C(5) = 0$
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(5)	0.004536	0.001474

Restrictions are linear in coefficients.

Ejercicio 5

Partiendo del modelo inicial, se desea ahora analizar si la beta del exceso de rendimiento en el mercado (β_2) ha cambiado tras la segunda guerra mundial. Crea una variable ficticia, que podemos denominar $d1946$, que valga cero para los datos previos a 1946 (hasta 194512) y uno a partir de enero de 1946. Añade en el modelo la $d1946 * (R_m - R_f)$. ¿Cuál es ahora el efecto marginal del exceso de rendimiento en el mercado sobre el exceso de rendimiento en la cartera HiTec? Contrasta si efectivamente este efecto es diferente antes y después de la segunda guerra mundial.

Dependent Variable: HITEC-RF
Method: Least Squares
Date: 09/02/20 Time: 15:17
Sample: 1926M07 2015M12
Included observations: 1074

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.127978	0.064195	1.993583	0.0465
RMRF	0.963968	0.018982	50.78454	0.0000
SMB	0.041980	0.020734	2.024765	0.0431
HML	-0.305851	0.020473	-14.93923	0.0000
D1946*RMRF	0.040862	0.025967	1.573582	0.1159
R-squared	0.864659	Mean dependent var		0.663892
Adjusted R-squared	0.864152	S.D. dependent var		5.615646
S.E. of regression	2.069790	Akaike info criterion		4.297416
Sum squared resid	4579.627	Schwarz criterion		4.320596
Log likelihood	-2302.712	Hannan-Quinn criter.		4.306195
F-statistic	1707.385	Durbin-Watson stat		1.910821
Prob(F-statistic)	0.000000			

El efecto que tiene ahora el exceso de rendimiento es menor respecto a nuestra ecuación principal (0.1), debido a que parcialmente la nueva variable recoge los mismos datos, aunque esta no es significativa estadísticamente hablando.

Wald Test:
Equation: EQ3

Test Statistic	Value	df	Probability
t-statistic	-21.94391	1069	0.0000
F-statistic	481.5354	(1, 1069)	0.0000
Chi-square	481.5354	1	0.0000

Null Hypothesis: $C(5) = C(2)$
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
-C(2) + C(5)	-0.923106	0.042067

Restrictions are linear in coefficients.

Realizando el contraste para comprobar si el efecto es diferente antes y después de la segunda guerra mundial ($H_0 : \beta_2 = \beta_5$), debemos rechazar la hipótesis nula,

por lo que podemos determinar que los efectos son diferentes antes y después de la segunda guerra mundial.

Ejercicio 6

Estima de nuevo el modelo básico de tres factores solicitando una estimación robusta de la matriz de varianzas de los parámetros. Compara los resultados de esta estimación con la obtenida en el primer punto. Puedes probar diferentes fórmulas para la estimación de la matriz de varianzas para ver si los resultados varían.

Dependent Variable: HITEC-RF

Method: Least Squares

Date: 09/02/20 Time: 15:17

Sample: 1926M07 2015M12

Included observations: 1074

White-Hinkley (HC1) heteroskedasticity consistent standard errors and covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.138453	0.062816	2.204108	0.0277
RMRF	0.986237	0.019642	50.21179	0.0000
SMB	0.041632	0.027664	1.504943	0.1326
HML	-0.319994	0.032304	-9.905557	0.0000
R-squared	0.864345	Mean dependent var		0.663892
Adjusted R-squared	0.863965	S.D. dependent var		5.615646
S.E. of regression	2.071217	Akaike info criterion		4.297867
Sum squared resid	4590.235	Schwarz criterion		4.316411
Log likelihood	-2303.955	Hannan-Quinn criter.		4.304891
F-statistic	2272.552	Durbin-Watson stat		1.911409
Prob(F-statistic)	0.000000	Wald F-statistic		914.4917
Prob(Wald F-statistic)	0.000000			

Dependent Variable: HITEC-RF
Method: Least Squares
Date: 09/02/20 Time: 15:17
Sample: 1926M07 2015M12
Included observations: 1074
MacKinnon-White (HC2) heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.138453	0.062892	2.201430	0.0279
RMRF	0.986237	0.019916	49.51928	0.0000
SMB	0.041632	0.028066	1.483382	0.1383
HML	-0.319994	0.032995	-9.698145	0.0000
R-squared	0.864345	Mean dependent var		0.663892
Adjusted R-squared	0.863965	S.D. dependent var		5.615646
S.E. of regression	2.071217	Akaike info criterion		4.297867
Sum squared resid	4590.235	Schwarz criterion		4.316411
Log likelihood	-2303.955	Hannan-Quinn criter.		4.304891
F-statistic	2272.552	Durbin-Watson stat		1.911409
Prob(F-statistic)	0.000000	Wald F-statistic		888.5525
Prob(Wald F-statistic)	0.000000			

Aquí hemos empleado el método de MCO pero estimando la matriz de varianzas-covarianzas para que sea consistente (robusta) mediante dos métodos distintos: el primero es usando White y el segundo HC¹ (heteroskedasticity consistent).

Podemos observar que los coeficientes no cambian, ya que el estimador sigue siendo insesgado; sólo hemos cambiado la forma de calcular la matriz de varianzas-covarianzas, y la única diferencia es que el error típico de nuestros estimadores es distinto.

Código

A continuación dejo el código empleado para replicar los resultados obtenidos mediante un programa de Eviews, así como mi repositorio en [GitHub](#) donde también se puede ver el código empleado en L^AT_EX.²

```
' Creamos un fichero cuyo sample se adecue a nuestros datos
wfcreate(wf=entrega) m 1/07/1926 2015

' Importamos las series de Excel
read(t=xls,b2) hitec.xls hitec
read(t=xls,c2) hitec.xls rmrf
```

¹No usamos HAC porque ya hemos comprobado que los residuos no están autocorrelados.

²Colocar el fichero de datos en la misma carpeta que el programa de Eviews.

```

read(t=xls,d2) hitec.xls smb
read(t=xls,e2) hitec.xls hml
read(t=xls,f2) hitec.xls rf

' Ej.1:
output(x) ej1 ' Fichero de salida para los datos de interés.
equation eq1.ls(p) hitec-rf c rmrf smb hml ' Ajuste de regresion por MCO

' Ej. 2:
output(x) ej2
eq1.auto(2)(p) ' Contraste de autocorrelacion del modelo con 2 lags.
eq1.white(p) ' Test de heterocedasticidad.
eq1.hist(p) ' Realizamos el histograma para comprobar la normalidad.

' Ej. 3:
output(x) ej3
eq1.wald(p) c(3) = c(4) = 0' Test de Wald

' Ej. 4:
output(x) ej4
equation eq2.ls(p) hitec-rf c rmrf smb hml smb^2
eq2.wald(p) c(5) = 0

' Ej. 5:
output(x) ej5
series d1946=@date>@dateval("1945m12") ' Dummy variable que
' elimina los datos previos a 1946
equation eq3.ls(p) hitec-rf c rmrf smb hml d1946*rmrf
eq3.wald(p) c(5) = c(2)

' Ej. 6:
output(x) ej6
equation eq1_white.ls(cov=white, p) hitec-rf c rmrf smb hml
' Ajuste de regresion por MCO solicitando que la matriz de var-cov
' cumpla el test de White (sea robusta).
equation eq1_hc.ls(cov=hc, p) hitec-rf c rmrf smb hml
' Igual pero siguiendo el criterio de Heterokedasticity autocorrelation consistent
' (más usado).

```