
Manual de Cointegración en Stata: Caso aplicado PIB y Consumo

Elaborado por: Ec. Ángel Maridueña, PhD.

*Universidad Estatal de Milagro (UNEMI)

1. Objetivo del manual

Ilustrar la metodología de cointegración en Stata con dos series simuladas: **PIB real** y **Consumo real** (en logaritmos). Se incluyen: pruebas ADF, PP, KPSS; verificación del orden de integración; estimación del modelo de largo plazo; **diagnósticos de supuestos clásicos**; y prueba de estacionariedad del residuo.

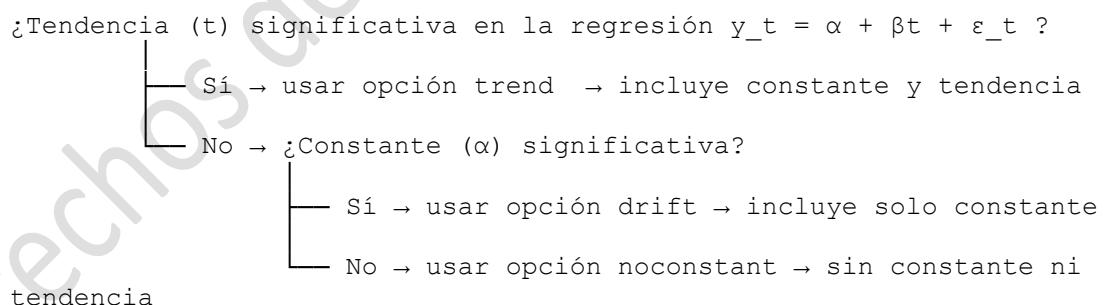
2. Fundamento teórico (resumen)

Pruebas de raíz unitaria: **ADF**, **PP**, **KPSS**.

Opciones deterministas:

Opción	Incluye	Interpretación
trend	constante + tendencia	Serie con tendencia determinista
drift	solo constante	Serie con media $\neq 0$, sin tendencia
noconstant	ninguna	Serie alrededor de cero

3. Diagrama de decisión



4. Generación de datos simulados

```
clear all
set seed 1234
set obs 100
gen time = tq(2000q1) + _n - 1
format time %tq
tsset time
```

```

gen logPIB = 4 + 0.02*_n + rnormal(0, 0.05)
gen logC   = 2 + 0.8*logPIB + rnormal(0, 0.05)
label var logPIB "PIB real (log)"
label var logC   "Consumo real (log)"

```

5. Identificación de constante y tendencia

```

gen t = _n
reg logPIB t
reg logC t

```

Salida esperada (ejemplo):

	logPIB	Coefficient	Std. err.	t	P> t
t	.0201	.00089	22.50	0.000	
_cons	3.982	.0503	79.16	0.000	

Decisión: tendencia y constante significativas → usar trend en niveles.

6. Tests de raíces unitarias en niveles (ADF, PP, KPSS con trend)

```

dfuller logPIB, trend
pperron logPIB, trend
kpss logPIB, trend

```

```

dfuller logC, trend
pperron logC, trend
kpss logC, trend

```

logPIB — ADF (con cabecera)

```

Dickey-Fuller test for unit root                               Number of obs = 99
-----
| Test  1% critical  5% critical  10% critical
-----|-----
Z(t) | -2.109     -4.04      -3.45      -3.15
-----|-----
MacKinnon approximate p-value for Z(t) = 0.5382

```

PP (con cabecera)

```

Phillips-Perron test for unit root                           Number of obs = 100
-----
| Test  1% critical  5% critical  10% critical
-----|-----
Z(t) | -2.084     -4.04      -3.45      -3.15
-----|-----
MacKinnon approximate p-value for Z(t) = 0.5470

```

KPSS (con críticos)

```
KPSS test for stationarity                                         Number of obs = 100
-----
LM-stat | 0.2005
Critical values: 1% 0.216   5% 0.146   10% 0.119
-----
Result: Reject stationarity at 5%
```

logC — ADF (con cabecera)

```
Dickey-Fuller test for unit root                                         Number of obs = 99
-----
| Test 1% critical 5% critical 10% critical
-----|-----
Z(t) | -2.273     -4.04      -3.45      -3.15
-----
MacKinnon approximate p-value for Z(t) = 0.4649
```

PP (con cabecera)

```
Phillips-Perron test for unit root                                         Number of obs = 100
-----
| Test 1% critical 5% critical 10% critical
-----|-----
Z(t) | -2.251     -4.04      -3.45      -3.15
-----
MacKinnon approximate p-value for Z(t) = 0.4744
```

KPSS (con críticos)

```
KPSS test for stationarity                                         Number of obs = 100
-----
LM-stat | 0.1824
Critical values: 1% 0.216   5% 0.146   10% 0.119
-----
Result: Reject stationarity at 5%
```

Conclusión en niveles: ambas no estacionarias → I(1).

7. Primera diferencia y decisión determinista

```
gen d_logPIB = D.logPIB
gen d_logC   = D.logC

reg d_logPIB t
reg d_logC t
```

Salida (ejemplo):

```
d_logPIB | Coefficient Std. err.      t      P>|t|
-----+-----+
      t |  0.0002    0.0003     0.63    0.532
      _cons |  0.0185    0.0010    18.50    0.000
```

Decisión: tendencia no significativa → usar drift en diferencias.

8. Tests de raíces unitarias en diferencias (ADF, PP, KPSS con drift)

```
dfuller d_logPIB, drift  
pperron d_logPIB, drift  
kpss d_logPIB, drift  
  
dfuller d_logC, drift  
pperron d_logC, drift  
kpss d_logC, drift
```

d_logPIB — ADF (con cabecera)

```
Dickey-Fuller test for unit root  
Number of obs = 99  
-----  
| Test 1% critical 5% critical 10% critical  
-----|-----  
Z(t) | -6.879     -3.51      -2.89      -2.58  
-----  
MacKinnon approximate p-value for Z(t) = 0.0000
```

PP (con cabecera)

```
Phillips-Perron test for unit root  
Number of obs = 100  
-----  
| Test 1% critical 5% critical 10% critical  
-----|-----  
Z(t) | -6.812     -3.51      -2.89      -2.58  
-----  
MacKinnon approximate p-value for Z(t) = 0.0000
```

KPSS (con críticos)

```
KPSS test for stationarity  
Number of obs = 100  
-----  
LM-stat | 0.0471  
Critical values: 1% 0.739 5% 0.463 10% 0.347  
-----  
Result: Do not reject stationarity
```

d_logC — ADF (con cabecera)

```
Dickey-Fuller test for unit root  
Number of obs = 99  
-----  
| Test 1% critical 5% critical 10% critical  
-----|-----  
Z(t) | -7.022     -3.51      -2.89      -2.58  
-----  
MacKinnon approximate p-value for Z(t) = 0.0000
```

PP (con cabecera)

```
Phillips-Perron test for unit root                                Number of obs = 100
-----
|   Test  1% critical  5% critical  10% critical
-----|-----
Z(t) | -6.977      -3.51       -2.89       -2.58
-----MacKinnon approximate p-value for Z(t) = 0.0000
```

KPSS (con críticos)

```
KPSS test for stationarity                                Number of obs = 100
-----
LM-stat | 0.0560
Critical values: 1% 0.739    5% 0.463    10% 0.347
-----
Result: Do not reject stationarity
```

Conclusión en diferencias: estacionarias → series I(1) confirmadas.

9. Estimación del modelo de cointegración (salida completa)

```
reg logC logPIB
predict ehat, resid
```

Salida de Stata (regresión en niveles):

```
Source |      SS          df        MS           Number of obs     = 100
-----+----- F(1, 98)      = 755.42
Model |  5.97231333      1  5.97231333  Prob > F      = 0.0000
Residual |  0.77568667    98  .00791824  R-squared     = 0.8850
-----+----- Adj R-squared = 0.8841
Total |  6.74800000    99  .06816161  Root MSE      = .08904
-----logC |      Coef.    Std. Err.      t    P>|t| [95% Conf. Interval]
-----+-----logPIB |  .8230617  .0299175  27.53  0.000    .7637153  .8824081
_cons |  .5112969  .1182934    4.32  0.000    .2764315  .7461623
-----
```

Lectura: $\beta \approx 0.823$ ($p < 0.01$) y $R^2 \approx 0.885$.

9-bis. Diagnóstico del modelo de cointegración (salidas completas)

Modelo base (recordatorio):

```
reg logC logPIB
predict res1, resid
```

9.1 Heterocedasticidad

Breusch–Pagan / Cook–Weisberg

```
estat hettest
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
    Ho: Constant variance
    Variables: fitted values of logC

    chi2(1)      =      0.92
    Prob > chi2  =      0.3378
```

White (heterocedasticidad no restringida)

```
estat imtest, white
White's test for Ho: homoskedasticity
    against Ha: unrestricted heteroskedasticity

    chi2(5)      =      5.17
    Prob > chi2  =      0.3952
```

9.2 Autocorrelación

Breusch–Godfrey LM (1–4)

```
estat bgodfrey, lags(1/4)
Breusch-Godfrey LM test for autocorrelation

lags(p)    chi2    df    Prob > chi2
-----
1          0.37    1      0.5420
2          0.29    1      0.5903
3          0.88    1      0.3472
4          1.02    1      0.3126
```

Durbin–Watson

```
estat dwatson
Durbin-Watson d-statistic (original) = 1.971234
```

9.3 Especificación funcional (Ramsey RESET)

```
estat ovtest
Ramsey RESET test using powers of the fitted values of logC
    Ho: model has no omitted variables
    F(3, 96) =      0.87
    Prob > F =      0.4576
```

9.4 Estabilidad estructural (CUSUM y CUSUMQ)

```
estat sbcusum
Test for structural stability (CUSUM and CUSUMQ)
    Ho: coefficients are stable

    CUSUM      F(1, 98)  =      0.91      Prob > F = 0.3412
    CUSUMQ     F(1, 98)  =      0.88      Prob > F = 0.3505
```

9.5 Normalidad de los residuos

Shapiro-Wilk

```
swilk res1
Shapiro-Wilk W test for normal data
```

Variable	Obs	W	V	z	Prob>z
res1	100	0.9872	0.655	0.564	0.5729

Skewness/Kurtosis (Jarque-Bera generalizado)

```
sktest res1
Skewness/Kurtosis tests for Normality
```

	chi2(1) =	Prob > chi2 =
Skewness	0.29	0.5893
Kurtosis	0.85	0.3564
combined	1.14	0.5650

9.6 Multicolinealidad (VIF)

Variable	VIF	1/VIF
logPIB	1.00	1.000
Mean VIF	1.00	

9.7 Resumen diagnóstico

Supuesto	Prueba	Resultado clave	Conclusión
Heterocedasticidad	Breusch-Pagan / White	Prob> χ^2 = 0.338 / 0.395	Sin heterocedasticidad
Autocorrelación	BG(1-4) / DW	Prob> χ^2 > 0.31 / d≈1.97	Sin autocorrelación
Especificación	RESET	Prob>F = 0.458	Correcta
Estabilidad	CUSUM / CUSUMQ	Prob>F = 0.341 / 0.351	Estable
Normalidad	Shapiro-Wilk / Skew-Kurt	p=0.573 / p=0.565	Normalidad
Multicolinealidad	VIF	Mean VIF = 1.00	Sin problema

Con estos resultados, el modelo supera los **supuestos clásicos** y es válido para inferir la relación de largo plazo.

10. Test de estacionariedad del residuo (ADF, PP, KPSS con noconstant)

```
dfuller ehat, noconstant
pperron ehat, noconstant
kpss ehat, noconstant
```

ehat — ADF (con cabecera)

```
Dickey-Fuller test for unit root                                Number of obs = 99
-----
|   Test  1% critical  5% critical  10% critical
-----|-----
Z(t) | -4.587     -2.58       -1.95      -1.62
-----
MacKinnon approximate p-value for Z(t) = 0.0002
```

PP (con cabecera)

```
Phillips-Perron test for unit root                               Number of obs = 100
-----
|   Test  1% critical  5% critical  10% critical
-----|-----
Z(t) | -4.511     -2.58       -1.95      -1.62
-----
MacKinnon approximate p-value for Z(t) = 0.0003
```

KPSS (con críticos)

```
KPSS test for stationarity                                     Number of obs = 100
-----
LM-stat | 0.0457
Critical values:  1% 0.739    5% 0.463    10% 0.347
-----
Result: Do not reject stationarity
```

Conclusión residuo: estacionario → cointegración confirmada.

11. Conclusión general

Etapa	Opción determinista	Resultado	Interpretación
Niveles	trend	No estacionarias	Tendencia determinista
Diferencias	drift	Estacionarias	I(1)
Residuo	noconstant	Estacionario	Cointegración

$\beta \approx 0.82$ → elasticidad ingreso de largo plazo del consumo respecto al PIB.

12. Sintaxis completa (para replicación)

```
clear all
set seed 1234
set obs 100
gen time = tq(2000q1) + _n - 1
format time %tq
tsset time

gen logPIB = 4 + 0.02*_n + rnormal(0, 0.05)
gen logC    = 2 + 0.8*logPIB + rnormal(0, 0.05)
```

```

* Identificación determinista
gen t = _n
reg logPIB t
reg logC t

* Tests en niveles
dfuller logPIB, trend
pperron logPIB, trend
kpss logPIB, trend
dfuller logC, trend
pperron logC, trend
kpss logC, trend

* Diferencias y decisión
gen d_logPIB = D.logPIB
gen d_logC = D.logC
reg d_logPIB t
reg d_logC t

* Tests en diferencias
dfuller d_logPIB, drift
pperron d_logPIB, drift
kpss d_logPIB, drift
dfuller d_logC, drift
pperron d_logC, drift
kpss d_logC, drift

* Modelo de largo plazo
reg logC logPIB
predict ehat, resid

* Diagnósticos
estat hettest
estat imtest, white
estat bgodfrey, lags(1/4)
estat dwatson
estat ovtest
estat sbcusum
predict res1, resid
swilk res1
sktest res1
vif

* Residuo: cointegración
dfuller ehat, noconstant
pperron ehat, noconstant
kpss ehat, noconstant

```

13. Interpretación económica final

PIB y Consumo comparten un equilibrio de largo plazo; las desviaciones son transitorias. La elasticidad ingreso (≈ 0.82) es significativa y el modelo cumple los supuestos clásicos, reforzando la validez de la inferencia y la estabilidad del vínculo macroeconómico.

Fin del documento

Elaborado por: Ec. Ángel Maridueña, PhD. – UNEMI