

# Hoja de Referencia Series Temporales

Por Marcelo Moreno Porras - Universidad Rey Juan Carlos  
The Econometrics Cheat Sheet Project

## Conceptos básicos

### Definiciones

**Serie temporal** - sucesión de observaciones ordenadas en el tiempo con una frecuencia fija.

Dado el formato de una serie temporal:

- **Punto en el tiempo (stock)** - se registra un único valor para cada período.
- **Agregado (flujo)** - los valores representan totales o promedios a lo largo del período.
- **Rango/intervalo (OHLC)** - cada período registra múltiples estadísticas, como *min*, *max*, *open*, *close*.

**Proceso estocástico** - es una secuencia de variables aleatorias que están indexadas en el tiempo.

### Componentes de una serie temporal

- **Tendencia** - movimiento general a l/p de una serie.
- **Variaciones estacionales** - oscilaciones periódicas producidas en un período igual o inferior al año, y pueden ser fácilmente identificadas en diferentes años (usualmente resultado de la climatología).
- **Ciclo** - oscilaciones periódicas producidas en un periodo mayor al año (son resultado del ciclo económico).
- **Variaciones residuales** - movimientos que no siguen una oscilación periódica identificable (eventos irreg.)

### Tipos de modelos de series temporales

- **Modelos estáticos** - la relación entre  $y$  y  $x$  es contemporánea. Conceptualmente:

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

- **Modelos de rezagos distribuidos** - la relación entre  $y$  y  $x$  no es contemporánea. Conceptualmente:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \dots + \beta_s x_{t-(s-1)} + u_t$$

El efecto acumulado a largo plazo en  $y$  cuando  $\Delta x$  es:

$$\beta_1 + \beta_2 + \dots + \beta_s$$

- **Modelos dinámicos** - rezagos de la variable dependiente (endogeneidad). Conceptualmente:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_s y_{t-s} + u_t$$

- Combinación de lo anterior, como modelos de rezagos distribuidos racionales (rezagos distribuidos + dinámicos).

## Supuestos y propiedades

### Supuestos MCO bajo series temporales

Bajo estos supuestos, los estimadores de los parámetros MCO presentarán buenas propiedades. **Supuestos Gauss-Markov** extendidos para series temporales:

- t1. **Linealidad de parámetros y dependencia débil.**
  - a.  $y_t$  debe ser una función lineal de  $\beta$ .
  - b. El proceso estocástico  $\{(x_t, y_t) : t = 1, 2, \dots, T\}$  es estacionario y débilmente dependiente.
- t2. **No colinealidad perfecta.**
  - No hay variables independientes que sean constantes:  $\text{Var}(x_j) \neq 0, \forall j = 1, \dots, k$
  - No hay una relación lineal exacta entre variables independientes.
- t3. **Media condicional cero y correlación cero.**
  - a. No hay errores sistemáticos:  $E(u | x_1, \dots, x_k) = E(u) = 0 \rightarrow$  **exogeneidad fuerte** (a implica b).
  - b. No hay variables relevantes no incluidas en el modelo:  $\text{Cov}(x_j, u) = 0, \forall j = 1, \dots, k \rightarrow$  **exogeneidad débil**.
- t4. **Homocedasticidad.** La variabilidad de los resid. es igual para cualquier nivel de  $x$ :  $\text{Var}(u | x_1, \dots, x_k) = \sigma_u^2$
- t5. **No autocorrelación.** Los residuos no contienen información sobre otros residuos:  $\text{Corr}(u_t, u_s | x_1, \dots, x_k) = 0, \forall t \neq s$
- t6. **Normalidad.** Los residuos son independientes e idénticamente distribuidos (**i.i.d.**):  $u \sim \mathcal{N}(0, \sigma_u^2)$
- t7. **Tamaño de datos.** El número de observaciones disponibles debe ser mayor a  $(k+1)$  parámetros a estimar. (Ya satisfecho bajo situaciones asintóticas)

### Propiedades asintóticas de MCO

Bajo los supuestos del modelo econométrico y el Teorema Central del Límite:

- De t1 a t3a: MCO es **insesgado**.  $E(\hat{\beta}_j) = \beta_j$
- De t1 a t3: MCO es **consistente**.  $\text{plim}(\hat{\beta}_j) = \beta_j$  (a t3b sin t3a, exogeneidad débil, insesg. y consistente).
- De t1 a t5: **normalidad asintótica** de MCO (entonces, t6 es necesariamente satisfecho):  $u \sim \mathcal{N}(0, \sigma_u^2)$
- De t1 a t5: **estimador insesgado** de  $\sigma_u^2$ .  $E(\hat{\sigma}_u^2) = \sigma_u^2$
- De t1 a t5: MCO es MELI (Mejor Estimador Lineal Insesgado, **BLUE** en inglés) or **eficiente**.
- De t1 a t6: contrastes de hipótesis e intervalos de confianza son fiables.

## Tendencia y estacionalidad

**Regresión espuria** - es cuando la relación entre  $y$  y  $x$  es debida a factores que afectan a  $y$  y que tienen correlación con  $x$ ,  $\text{Corr}(x_j, u) \neq 0$ . Es el **incumplimiento de t3**.

### Tendencia

Dos series temporales pueden tener la misma (o contraria) tendencia, lo que lleva a altos niveles de correlación. Esto provoca una falsa apariencia de causalidad; el problema es **regresión espuria**. Dado el modelo:

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

donde:

$$y_t = \alpha_0 + \alpha_1 \text{Tendencia} + v_t$$

$$x_t = \gamma_0 + \gamma_1 \text{Tendencia} + v_t$$

Añadir una tendencia al modelo puede resolver el problema:

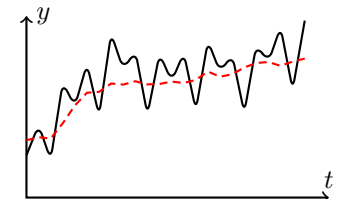
$$y_t = \beta_0 + \beta_1 x_t + \beta_2 \text{Tendencia} + u_t$$

Una tendencia puede ser lineal o no lineal (cuadrática, cúbica, exponencial, etc.)

Otra manera, es hacer uso del **filtro Hodrick-Prescott** y extraer la tendencia y el componente cíclico.

### Estacionalidad

Una serie temporal puede manifestar estacionalidad. Esto es, que la serie está sujeta a variaciones estacionales o patrones usualmente relacionados al clima.



Por ejemplo, el PIB es mayor en verano y menor en invierno (serie ajustada estacionalmente en **rojo discontinuo**).

- La regresión de series temporales que presentan estacionalidad puede conducir a **resultados espurios**.

Existen diferentes métodos de **ajuste estacional**:

- a. Incluir variables estacionales binarias en el modelo. Por ejemplo, serie trimestral ( $Sq_t$  son variables binarias):
$$y_t = \beta_0 + \beta_1 S2_t + \beta_2 S3_t + \beta_3 S4_t + \beta_4 x_{1t} + \dots + \beta_k x_{kt} + u_t$$
- b. Ajustar estacionalmente las variables y luego realizar la regresión con las variables ajustadas.
- c. Tomar diferencias estacionales, lo que también puede ayudar a eliminar la tendencia ( $s$  es el período estacional):

$$\nabla_s y_t = y_t - y_{t-s}$$

- d. Aplicar **X-13ARIMA-SEATS** (método mejor y más complicado que lo anterior).

## Autocorrelación

El residuo de cualquier observación,  $u_t$ , está correlacionado con el residuo de cualquier otra observación. Las observaciones no son independientes. **Incumplimiento de t5.**

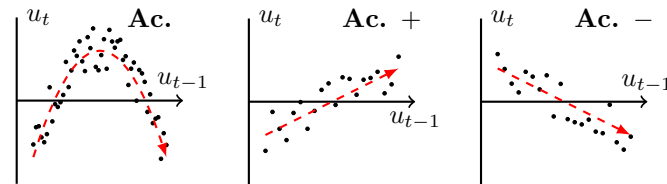
$$\text{Corr}(u_t, u_s \mid x_1, \dots, x_k) = \text{Corr}(u_t, u_s) \neq 0, \forall t \neq s$$

### Consecuencias

- Estimadores MCO son insesgados.
- Estimadores MCO son consistentes.
- MCO ya **no es eficiente**, pero sigue siendo ELI (Estimador Lineal Insesgado).
- La **estimación de la varianza** de los estimadores es **sesgada**: la construcción de intervalos de confianza y contraste de hipótesis no son fiables.

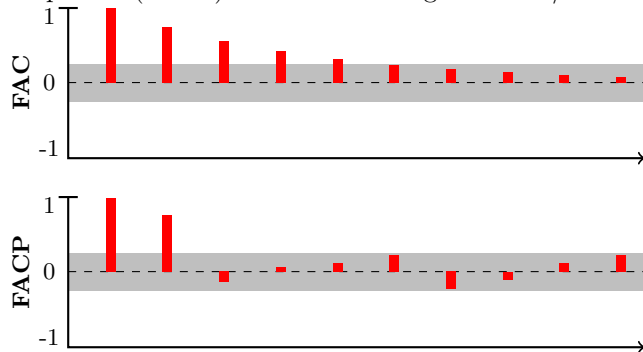
### Detección

**Gráficos de dispersión** - patrones en  $u_{t-1}$  vs.  $u_t$ .



**Correlograma** - función de autocorrelación (FAC) y el FAC parcial (FACP).

- Eje Y: correlación.
- Eje X: núm. de retardo.
- Área gris:  $\pm 1.96/T^{0.5}$



- **Proceso MA(q)**. **FAC**: sólo los  $q$  primeros coeficientes son significativos, el resto se anulan bruscamente. **FACP**: decrecimiento rápido exponencial atenuado u ondas sinusoidales.
- **Proceso AR(p)**. **FAC**: decrecimiento rápido exponencial atenuado u ondas sinusoidales. **FACP**: sólo los  $p$  primeros coeficientes son significativos, el resto se anulan bruscamente.

- **Proceso ARMA(p, q)**. **FAC** y **FACP**: los coeficientes no se anulan bruscamente y presentan un decrecimiento rápido.

Si los coeficientes de la FAC no decaen rápidamente, hay claro indicio de falta de estacionariedad en media.

**Contrastes** - Generalmente,  $H_0$ : No autocorrelación.

Suponiendo que  $u_t$  sigue un proceso AR(1):

$$u_t = \rho_1 u_{t-1} + \varepsilon_t$$

donde  $\varepsilon_t$  es ruido blanco.

- **Prueba t AR(1)** (regresores exógenos):

$$t = \frac{\hat{\rho}_1}{\text{ee}(\hat{\rho}_1)} \sim t_{T-k-1, \alpha/2}$$

$H_1$ : Autocorrelación de orden uno, AR(1).

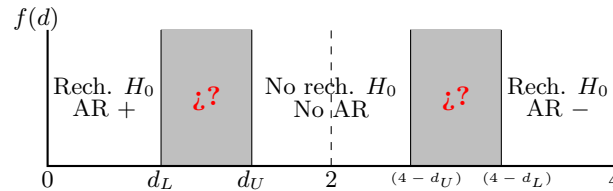
- **Estadístico Durbin-Watson** (regresores exógenos y normalidad de residuos):

$$d = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2} \approx 2 \cdot (1 - \hat{\rho}_1)$$

Donde  $0 \leq d \leq 4$

$H_1$ : Autocorrelación de orden uno, AR(1).

$d =$	0	2	4
$\rho \approx$	1	0	-1



- **h de Durbin** (regresores endógenos):

$$h = \hat{\rho} \cdot \sqrt{\frac{T}{1 - T \cdot v}}$$

donde  $v$  es la varianza estimada del coeficiente asociado a la variable endógena.

$H_1$ : Autocorrelación de orden uno, AR(1).

- **Prueba Breusch-Godfrey** (regresores endógenos): puede detectar procesos MA(q) y AR(p) ( $\varepsilon_t$  ruido b.):

– MA(q):  $u_t = \varepsilon_t - m_1 u_{t-1} - \dots - m_q u_{t-q}$

– AR(p):  $u_t = \rho_1 u_{t-1} + \dots + \rho_p u_{t-p} + \varepsilon_t$

Bajo  $H_0$ : No autocorrelación:

$$T \cdot R_{\hat{u}_t}^2 \underset{a}{\sim} \chi_q^2 \quad \text{or} \quad T \cdot R_{\hat{u}_t}^2 \underset{a}{\sim} \chi_p^2$$

$H_1$ : Autocorrelación de orden  $q$  (ó  $p$ ).

- **Prueba Ljung-Box Q**:

$H_1$ : Autocorrelación hasta orden  $h$ .

## Corrección

- Usar MCO con un estimador de la matriz de varianzas-covarianzas **robusto a la heterocedasticidad y autocorrelación** (HAC), por ejemplo, la propuesta de **Newey-West**.

- Usar **Mínimos Cuadrados Generalizados** (MCG). Suponiendo  $y_t = \beta_0 + \beta_1 x_t + u_t$ , con  $u_t = \rho u_{t-1} + \varepsilon_t$ , donde  $|\rho| < 1$  y  $\varepsilon_t$  es **ruido blanco**.

- Si  $\rho$  es **conocido**, usar **modelo cuasi-diferenciado**:

$$y_t - \rho y_{t-1} = \beta_0(1 - \rho) + \beta_1(x_t - \rho x_{t-1}) + u_t - \rho u_{t-1}$$

$$y_t^* = \beta_0^* + \beta_1^* x_t^* + \varepsilon_t$$

donde  $\beta_1' = \beta_1$ ; y estimarlo por MCO.

- Si  $\rho$  es **desconocido**, estimarlo -por ejemplo- el **método iterativo de Cochrane-Orcutt** (el método de Prais-Winsten también es bueno):

1. Obtener  $\hat{u}_t$  del modelo original.

2. Estimar  $\hat{u}_t = \rho \hat{u}_{t-1} + \varepsilon_t$  y obtener  $\hat{\rho}$ .

3. Crear un modelo cuasi-diferenciado:

$$y_t - \hat{\rho} y_{t-1} = \beta_0(1 - \hat{\rho}) + \beta_1(x_t - \hat{\rho} x_{t-1}) + u_t - \hat{\rho} u_{t-1}$$

$$y_t^* = \beta_0^* + \beta_1^* x_t^* + \varepsilon_t$$

donde  $\beta_1' = \beta_1$ ; y estimarlo por MCO.

4. Obtener  $\hat{u}_t^* = y_t - (\hat{\beta}_0^* + \hat{\beta}_1^* x_t) \neq y_t - (\hat{\beta}_0 + \hat{\beta}_1 x_t^*)$ .

5. Repetir desde el paso 2. El algoritmo termina cuando los parámetros estimados varían muy poco entre iteraciones.

- Si no se arregla, buscar **fuerte dependencia** en la serie.

## Suavizado exponencial

Dado  $\{y_t\}$ , la serie suavizada  $\{f_t\}$ :

$$f_t = \alpha y_t + (1 - \alpha) f_{t-1}$$

donde  $0 < \alpha < 1$  es el factor de suavizado y  $f_0 = y_0$ .

## Predicciones

Dos tipos de predicciones:

- Valor medio de  $y$  para un valor específico de  $x$ .
- Valor individual de  $y$  para un valor específico de  $x$ .

**U de Theil** - compara los resultados pronosticados con los de la previsión con datos históricos mínimos.

$$U = \sqrt{\frac{\sum_{t=1}^{T-1} \left( \frac{\hat{y}_{t+1} - y_{t+1}}{y_t} \right)^2}{\sum_{t=1}^{T-1} \left( \frac{y_{t+1} - y_t}{y_t} \right)^2}}$$

- $< 1$ : La predicción es mejor que adivinar.
- $= 1$ : La predicción es tan buena como adivinar.
- $> 1$ : La predicción es peor que adivinar.

# Estacionariedad

La estacionariedad permite reconocer relaciones – inalteradas en el tiempo– entre variables.

- **Proceso estacionario** (estacionariedad estricta) - la distribución de probabilidad conjunta del proceso permanece inalterada al desplazarse  $h$  periodos.
- **Proceso no estacionario** - por ejemplo, una serie con tendencia, donde al menos la media cambia con el tiempo.
- **Proceso estacionario en covarianza** - es una forma más débil de estacionariedad:
  - $E(x_t)$  es constante.      –  $\text{Var}(x_t)$  es constante.
  - Para cualquier  $t, h \geq 1$ ,  $\text{Cov}(x_t, x_{t+h})$  depende sólo de  $h$ , no de  $t$ .

# Dependencia débil

La dependencia débil reemplaza el supuesto de muestreo aleatorio en series temporales.

- Un proceso estacionario  $\{x_t\}$  es **débilmente dependiente** cuando  $x_t$  y  $x_{t+h}$  son casi independientes a medida que  $h$  aumenta sin límite.
- Un proceso estacionario en covarianza es **débilmente dependiente** si la correlación entre  $x_t$  y  $x_{t+h}$  tiende a 0 lo suficientemente rápido cuando  $h \rightarrow \infty$  (no están asintóticamente correlacionados).

Los procesos débilmente dependientes se llaman **integrados de orden cero**,  $I(0)$ . Algunos ejemplos:

- **Media móvil** -  $\{x_t\}$  es una media móvil de orden  $q$ ,  $MA(q)$ :

$$x_t = e_t + m_1e_{t-1} + \dots + m_qe_{t-q}$$

donde  $\{e_t : t = 0, 1, \dots, T\}$  es una secuencia *i.i.d.* con media cero y varianza  $\sigma_e^2$ .

- **Proceso autorregresivo** -  $\{x_t\}$  es un proceso autorregresivo de orden  $p$ ,  $AR(p)$ :

$$x_t = \rho_1x_{t-1} + \dots + \rho_px_{t-p} + e_t$$

donde  $\{e_t : t = 1, 2, \dots, T\}$  es una secuencia *i.i.d.* con media cero y varianza  $\sigma_e^2$ .

**Condición de estabilidad:** si  $1 - \rho_1z - \dots - \rho_pz^p = 0$  para  $|z| > 1$ , entonces  $\{x_t\}$  es un proceso  $AR(p)$  débilmente dependiente. Para  $AR(1)$ , la condición es:  $|\rho_1| < 1$ .

- **Proceso ARMA** - es una combinación de los dos anteriores;  $\{x_t\}$  es un  $ARMA(p, q)$ :

$$x_t = e_t + m_1e_{t-1} + \dots + m_qe_{t-q} + \rho_1x_{t-1} + \dots + \rho_px_{t-p}$$

# Raíces unitarias

Un proceso es integrado de orden  $d$ ,  $I(d)$ , si al aplicar diferencias  $d$  veces hace al proceso estacionario.

Cuando  $d \geq 1$ , se dice que el proceso tiene **raíz unitaria**. Un proceso tiene una raíz unitaria cuando no se cumple la condición de estabilidad (hay raíces en el círculo unitario).

# Dependencia fuerte

Generalmente, las series económicas son fuertemente persistentes. Algunos ejemplos de **raíz unitaria**  $I(1)$ :

- **Paseo aleatorio** - un proceso  $AR(1)$  con  $\rho_1 = 1$ .

$$y_t = y_{t-1} + e_t$$

donde  $\{e_t : t = 1, 2, \dots, T\}$  es una secuencia *i.i.d.* con media cero y varianza  $\sigma_e^2$ .

- **Paseo aleatorio con deriva** - un proceso  $AR(1)$  con  $\rho_1 = 1$  y una constante.

$$y_t = \beta_0 + y_{t-1} + e_t$$

donde  $\{e_t : t = 1, 2, \dots, T\}$  es una secuencia *i.i.d.* con media cero y varianza  $\sigma_e^2$ .

# Contrastes de raíz unitaria

Contraste	$H_0$	Rechazar $H_0$
ADF	$I(1)$	$\text{tau} < \text{Valor crítico}$
KPSS	$I(0)$ nivel	$\mu > \text{Valor crítico}$
	$I(0)$ tendencia	$\text{tau} > \text{Valor crítico}$
Phillips-Perron	$I(1)$	$Z\text{-tau} < \text{Val. crítico}$
Zivot-Andrews	$I(1)$	$\text{tau} < \text{Valor crítico}$

# De raíz unitaria a dependencia débil

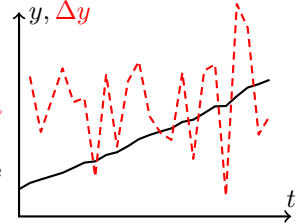
Integrado de **orden uno**,  $I(1)$ , significa que la **primera diferencia** del proceso es **débilmente dependiente** ó  $I(0)$  (y usualmente, estacionaria). Sea  $\{y_t\}$  un paseo aleatorio:

$$\Delta y_t = y_t - y_{t-1} = e_t$$

donde  $\{e_t\} = \{\Delta y_t\}$  es *i.i.d.*

Nota:

- La **primera diferencia** elimina su tendencia.
- El logaritmo de una serie estabiliza su varianza.



# De raíz unitaria a cambio porcentual

Cuando una serie  $I(1)$  es estrictamente positiva, a menudo se utilizan logaritmos antes de diferenciar para aproximar cambios porcentuales:

$$\Delta \log(y_t) = \log(y_t) - \log(y_{t-1}) \approx \frac{y_t - y_{t-1}}{y_{t-1}}$$

# Ergodicidad

Un proceso estrictamente estacionario  $\{y_t\}$  es **ergódico** si los promedios temporales convergen a sus promedios muestrales (esperanzas). Esto suele garantizarse mediante una **mezcla fuerte** (strong mixing), que implica independencia asintótica de eventos lejanos.

$$\frac{1}{T} \sum_{t=1}^T y_t \xrightarrow{a} E(y_t)$$

Sin ello, los momentos muestrales pueden no reflejar los momentos poblacionales. Los estimadores son inconsistentes.

# Cointegración

Dos series  $I(1)$  están **cointegradas** si una combinación lineal de ellas es  $I(0)$ . Una regresión entre ellas no es espuria, sino que refleja una relación de **largo plazo**. Las variables cointegradas comparten una tendencia estocástica común. Por ejemplo,  $\{x_t\}$  y  $\{y_t\}$  son  $I(1)$ , pero  $y_t - \beta x_t = u_t$  donde  $\{u_t\}$  es  $I(0)$ . ( $\beta$  es el parámetro cointegrador).

# Contraste de cointegración

1. Estimar  $y_t = \alpha + \beta x_t + u_t$  y obtener  $\hat{u}_t$ .
  2. Realizar el ADF sobre  $\hat{u}_t$  con una distribución especial.
- El resultado del contraste es equivalente a:
- $H_0: \beta = 0$  (no cointegración)
  - $H_1: \beta \neq 0$  (cointegración)
- si el estadístico de contraste  $>$  valor crítico, rechazar  $H_0$ .

# Heterocedasticidad en series temp.

Afecta al **supuesto t4**, lo que lleva a que **MCO no sea eficiente**.

Usar contrastes como el Breusch-Pagan o White, donde  $H_0$ : No heterocedasticidad. Para que los contrastes funcionen, no ha de existir **autocorrelación**.

# ARCH

Un modelo de heteroced. condicional autorreg. (ARCH) se usa para analizar una forma de heteroced. dinámica, donde la varianza del error sigue un proceso  $AR(p)$ .

Dado el modelo:  $y_t = \beta_0 + \beta_1 z_t + u_t$  donde, hay  $AR(1)$  y heterocedasticidad:

$$E(u_t^2 | u_{t-1}) = \alpha_0 + \alpha_1 u_{t-1}^2$$

# GARCH

Un modelo ARCH general (GARCH) es similar a ARCH, pero la varianza del error sigue un proceso  $ARMA(p, q)$ .