Series de tiempo I

Procesos univariados y estacionarios

Gabriel V. Montes-Rojas

La estructura de las series temporales

- Series temporales Información para diferentes periodos indexados por el tiempo t, pero para el mismo individuo (país, firma)

  $\{y_t, x_t\}_{t=1}^{T}$, donde t indica tiempo.

  Ej: Inflación. Precio del barril de petróleo. PBI.
- La estructura temporal es importante: el pasado afecta el presente (y el futuro) pero no al revés.
- En un modelo de regresión:

$$y_t = \beta_0 + \beta_1 x_t + u_t, t = 1, 2, ..., T$$

  $\{u_t\}_{t=1}^{T}$ puede que no sea $i.i.d.$ (independiente e idénticamente distribuido). **Correlación serial** de los errores en regresión.
- Operador de rezagos o lags, L: $Ly_t = y_{t-1}$, $L^2 y_t = y_{t-2}$.
- Operador de diferencias, Δ: $\Delta y_t = y_t - y_{t-1} = (1 - L)y_t$.

Series temporales lineales

La serie temporal más simple se llama **ruido blanco**:
$\{e_t\}$ es ruido blanco si es una secuencia de variables aleatorias i.i.d. (independientes e idénticamente distribuidas) con

1. $E[e_t] = 0, \forall t$
2. $Var[e_t] = \sigma_e^2, \forall t$.

Una serie de tiempo es **lineal** si

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i e_{t-i},$$

donde $\psi_0 = 1$ y $\{e_t\}$ es ruido blanco. e_t es la nueva información que se adquiere en t, también llamada innovación o shock. ψ son los "pesos" de las innovaciones del pasado en el presente.
- **Ejercicio:** Calcular $E[y_t]$, $Var[y_t]$ y $Cov[y_t, y_{t-j}]$, $j = 0, 1, 2, ...$.

Procesos autorregresivos AR(1)

Consideremos un proceso autorregresivo de orden 1, AR(1)

$$y_t = \phi_0 + \phi_1 y_{t-1} + e_t, t = 1, 2, ...$$

donde $\{e_t : t = 0, 1, ...\}$ es una secuencia de errores ruido blanco.
Asumamos que $|\phi_1| < 1$.

Procesos autorregresivos AR(1)

Notar que estos procesos tienen memoria infinita, es decir, el valor actual depende de todos los shocks pasados. Veamos:

$$
\begin{aligned}
y_t &= \phi_0 + \phi_1 y_{t-1} + e_t = \phi_0 + \phi_1(\phi_0 + \phi_1 y_{t-2} + e_{t-1}) + e_t \\
&= \phi_0(1+\phi_1) + \phi_1^2 y_{t-2} + e_t + \phi_1 e_{t-1} \\
&= \phi_0(1+\phi_1) + \phi_1^2(\phi_0 + \phi_1 y_{t-3} + e_{t-2}) + e_t + \phi_1 e_{t-1} \\
&= \ldots = \sum_{j=0}^{\infty} \phi_1^j(\phi_0 + e_{t-j}) = \frac{\phi_0}{1-\phi_1} + \sum_{j=0}^{\infty} \phi_1^j e_{t-j}
\end{aligned}
$$

Nota: Dado que $|\phi_1| < 1$, entonces $\sum_{j=0}^{\infty} \phi_1^j \phi_0 = \frac{\phi_0}{1-\phi_1}$.

Procesos autorregresivos AR(1)

Probar que:

$$E[y_t] = \frac{\phi_0}{1-\phi_1}$$

$$Var[y_t] = \frac{\sigma_e^2}{1-\phi_1^2}$$

Hay varias formas de probarlo. (i) Una es usando la ecuación de memoria infinita, $y_t = \frac{\phi_0}{1-\phi_1} + \sum_{j=0}^{\infty} \phi_1^j e_{t-j}$, aplicando esperanza y varianza. (ii) Otra es usando el operador de lags L, y notando que $\frac{1}{(1-L)=1+L+L^2+L^3+...}$. (iii) Otra es asumir estacionariedad, que veremos más adelante, pero que implica que $E[y_t] = E[y_{t-1}] = E[y_{t-2}] = ...$ y que $Var[y_t] = Var[y_{t-1}] = Var[y_{t-2}] = ...$.

Procesos autorregresivos AR(1)

Definamos la **función de autocovarianzas de** k **rezagos** como

$$\gamma_k = Cov[y_t, y_{t-k}] = Cov[y_t, y_{t+k}].$$

(Nota: repasar las propiedades de varianzas y covarianzas vistos en la primera parte del curso.)

- $\gamma_0 = Cov[y_t, y_t] = Var[y_t] = \frac{\sigma_e^2}{1-\phi_1^2}$
- $\gamma_1 = Cov[y_{t-1}, y_t] = Cov[y_{t-1}, \phi_0 + \phi_1 y_{t-1} + e_t] = \phi_1\gamma_0 = \frac{\phi_1\sigma_e^2}{1-\phi_1^2}$
- $\gamma_2 = Cov[y_{t-2}, y_t] = Cov[y_{t-2}, \phi_0(1+\phi_1) + \phi_1^2 y_{t-2} + e_t + \phi e_{t-1}]$ $= \phi_1^2\gamma_0 = \frac{\phi_1^2\sigma_e^2}{1-\phi_1^2}$
- $\gamma_k = \phi_1^k\gamma_0 = \frac{\phi_1^k\sigma_e^2}{1-\phi_1^2}$
- Definir $x_t = y_t - E[y_t]$. Notar que el AR(1) se puede escribir como $x_t = \phi_1 x_{t-1} + e_t$. Entonces, tenemos que $E[x_t^2] = Var[y_t] = \gamma_0$. Por otro lado, si multiplicamos por $x_t t - 1$ la ecuación anterior, $x_t x_{t-1} = \phi_1 x_{t-1}^2 + x_{t-1}e_t$, $\gamma_1 = E[x_t x_{t-1}] = \phi_1 E[x_{t-1}^2] + E[x_{t-1}e_t] = \phi_1\gamma_0$. Esto lo podemos repetir para x_{t-2}, x_{t-3}, etc.

Definamos la **función de autocorrelación** como $\rho_k = \gamma_k/\gamma_0$. Para el AR(1) es $\rho_k = \phi_1^k$.

Procesos autorregresivos AR(2)

Procesos autorregresivos de orden 2, AR(2)

Supongamos ahora un proceso AR(2),

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + e_t,$$

donde $\{e_t : t = 0, 1, \ldots\}$ es una secuencia de errores ruido blanco.

Ejercicio: Resolver las fórmulas de $E[y_t], Var[y_t], \gamma_k, \rho_k$.

Prueba: Asumir estacionariedad y resolver $E[y_t] = \phi_0/(1 - \phi_1 - \phi_2)$. Definir $x_t = y_t - E[y_t]$, tal que $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + e_t$. Multiplicando por x_t, x_{t-1}, x_{t-2}, tenemos las ecuaciones de Yule-Walker:

$$\begin{aligned}
\gamma_0 &= \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma_e^2 \\
\gamma_1 &= \phi_1\gamma_0 + \phi_2\gamma_1 \\
\gamma_2 &= \phi_1\gamma_1 + \phi_2\gamma_0,
\end{aligned}$$

tres ecuaciones con tres incógnitas... fácil.. o sino

$$\begin{aligned}
\rho_1 &= \phi_1 + \phi_2\rho_1 \\
\rho_2 &= \phi_1\rho_1 + \phi_2.
\end{aligned}$$

Procesos autorregresivos AR(p)

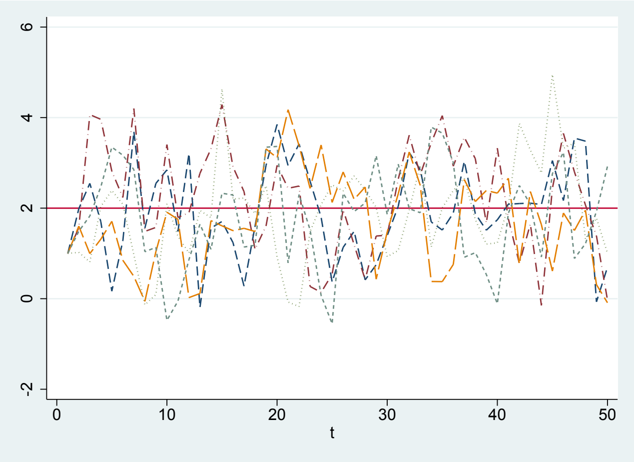
Procesos autorregresivos de orden p, AR(p)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + e_t, \, t = p, p+1, \ldots$$

donde $\{e_t : t = 0, 1, \ldots\}$ es una secuencia de errores ruido blanco.

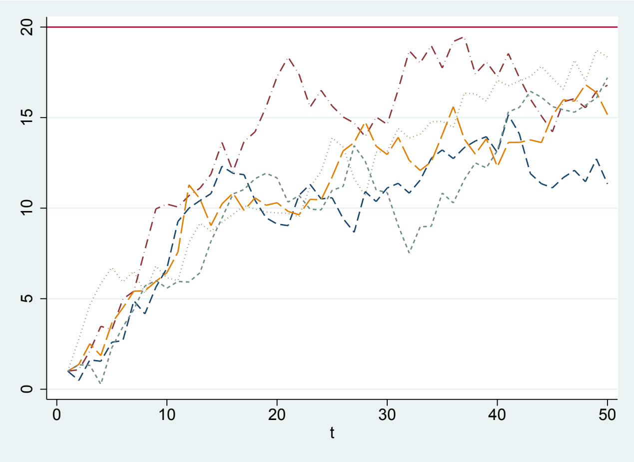
Simulación de 5 procesos AR(1)

$y_t = \phi_0 + \phi_1 y_{t-1} + e_t$, $t = 1, 2, ..., 50$, $\phi_0 = 1$, $\phi_1 = 0.5$, $y_0 = 1$, $e_t \sim \ iid \ N(0, 1)$



Simulación de 5 procesos AR(1)

$y_t = \phi_0 + \phi_1 y_{t-1} + e_t$, $t = 1, 2, ..., 50$, $\phi_0 = 1$, $\phi_1 = 0.95$, $y_0 = 1$, $e_t \sim$ *iid* $N(0, 1)$



Promedios móviles MA(1)

Promedios móviles de orden 1 MA(1) (MA:moving average)

$$y_t = \alpha_0 + e_t + \alpha_1 e_{t-1}, t = 1, 2, ...$$

donde $\{e_t : t = 0, 1, ...\}$ es ruido blanco.
Nota: y_t es un **promedio ponderado** de e_t y e_{t-1}.

Promedios móviles

Promedios móviles de orden q MA(q)

$$y_t = \alpha_0 + e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \alpha_3 e_{t-3} \ldots + \alpha_q e_{t-q}, t = q, q+1, \ldots,$$

donde $\{e_t : t = 0, 1, \ldots\}$ es ruido blanco.

Promedios móviles

Promedios móviles de orden MA(q)

$$E[y_t] = \alpha_0$$

$$\begin{aligned} Var[y_t] &= \sigma_e^2 + \alpha_1^2\sigma_e^2 + ... + \alpha_q^2\sigma_e^2 \\ &= \sigma_e^2(1 + \alpha_1^2 + ... + \alpha_q^2) \end{aligned}$$

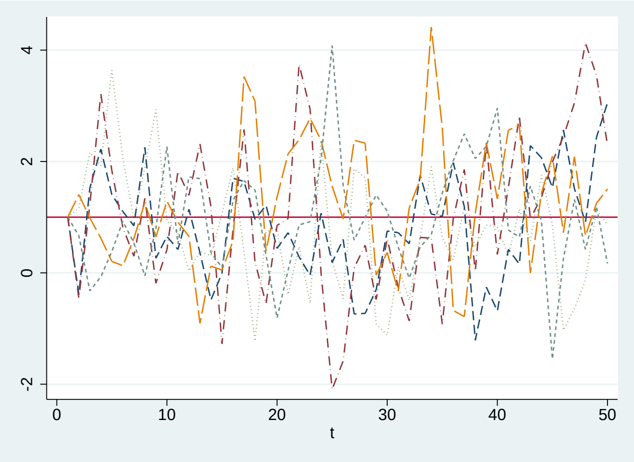
Propiedad de los procesos MA(q):

$$Cov[y_t, y_{t-k}] \neq 0, \text{ si } k \leq q$$

$$Cov[y_t, y_{t-k}] = 0, \text{ si } k > q$$

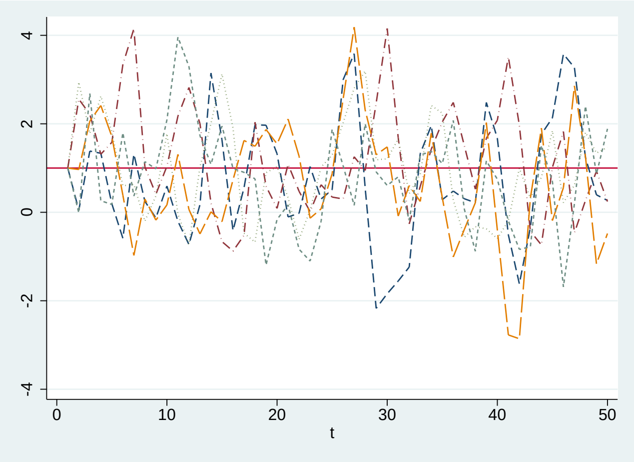
Simulación de 5 procesos MA(1)

$y_t = \alpha_0 + \alpha_1 e_{t-1} + e_t$, $t = 1, 2, ..., 50$, $\alpha_0 = 1$, $\alpha_1 = 0.5$, $y_0 = 1$, $e_t \sim$ *iid* $N(0, 1)$



Simulación de 5 procesos MA(1)

$y_t = \alpha_0 + \alpha_1 e_{t-1} + e_t$, $t = 1, 2, ..., 50$, $\alpha_0 = 1$, $\alpha_1 = 0.95$, $y_0 = 1$, $e_t \sim \ iid \ \ N(0, 1)$



Procesos ARMA

Un proceso $ARMA(p, q)$ se define como

$$y_t = \phi_0 + \sum_{j=1}^{p} \phi_j y_{t-j} + e_t + \sum_{j=1}^{q} \alpha_j e_{t-j}$$

donde $\{e_t\}$ es ruido blanco.

¿Cómo usar series temporales en STATA?

- Organizar los datos correctamente en Excel. Luego copiar & pegar a STATA.

TIME	YVAR	XVAR
1	y_1	x_1
2	y_2	x_2
3	y_3	x_3

- Tipear
  `tsset TIME` esto reconoce TIME como t, el indicador del tiempo
  [Ver http://www.stata.com/help.cgi?tsset]
- Para usar variables con rezagos/lags se debe usar L. para $t-1$, L2. para $t-2$, etc. (o sea Lj para j=1,2,3,...) Por ejemplo `L.YVAR` corresponde a $YVAR_{t-1}$.
- Para usar variables en (primeras) diferencias se debe usar D. Por ejemplo, `D.YVAR` corresponde a $\Delta YVAR_t = YVAR_t - YVAR_{t-1}$. [Se puede combinar D. y L.]
- Ejemplo: regresión para AR(2): `reg YVAR L.YVAR L2.YVAR`
- Modelo $ARMA(p,q)$: `arima YVAR, arima(p,0,q)`

Contrastes para ruido blanco

Supongamos el siguente contraste:

$$H_0 : y_t \text{ sigue un ruido blanco}$$

$$H_A : y_t \text{ no sigue un ruido blanco}$$

En STATA:
`wntestb YVAR` (Contraste de Bartlett)
`wntestq YVAR` (Contraste de Portmanteau)

Estacionariedad estricta

Procesos estrictamente estacionarios:: *El proceso* $\{y_t : t = 1, 2, ...\}$ *es estrictamente estacionario si para cada colección de índices de tiempo* $1 \leq t_1 \leq t_2 \leq ... \leq t_m$, *la distribución conjunta* $(y_{t_1}, y_{t_2}, ..., y_{t_m})$ *es la misma que la de* $(y_{t_1+h}, y_{t_2+h}, ..., y_{t_m+h})$ *para todo* $h \geq 1$.

- En otras palabras, la distribución es invariante en el tiempo no importa de que tiempo t estemos hablando.
- Es un supuesto fuerte.
- Ejercicio: Probar que los procesos AR(1) y MA(1) usados arriba satisfacen esta propiedad.

Estacionariedad débil

Procesos de estacionariedad débil: *El proceso* $\{y_t : t = 1, 2, ...\}$ *con momento segundo finito* $E[y_t^2] < \infty$ *es débilmente estacionario si (i)* $E[y_t]$ *es constante; (ii)* $Var[y_t]$ *es constante; y (iii) para cualquier* $t, h \geq 1$, $Cov[y_t, y_{t+h}]$ *depende solo de* h *pero no de* t.

- Nota: estacionariedad estricta implica estacionariedad débil.
- Ejercicio: Probar que los procesos AR(1) y MA(1) usados arriba satisfacen esta propiedad.

Procesos débilmente dependientes

Procesos débilmente dependientes: *El proceso estocástico* $\{y_t : t = 1, 2, ...\}$ *es débilmente dependiente si* y_t *y* y_{t+h} *se vuelven independientes (leer la covarianza es 0) cuando* $h \to \infty$*. O sea:* $Cov[y_t, y_{t+h}] \to 0$ *cuando* $h \to \infty$ *(no correlacionados asintóticamente)*

- Ejercicio: probar que los procesos AR(p) y MA(q) son débilmente dependientes.

'Todas las familias felices se parecen entre sí; las infelices son desgraciadas en su propia manera.'

Leo Tolstoy (1828-1910)

'Todas las series estacionarias se parecen entre sí; las no estacionarias son desgraciadas en su propia manera.'

¿Por qué es importante la estacionariedad?

- Para estimar y hacer inferencia sobre el proceso generador de los datos necesitamos del pasado. Necesitamos que el pasado se mantenga "parecido" al presente para que nuestras estimaciones tengan sentido hoy, y para predecir hacia el futuro.
- Antes de empezar a calcular medias, varianzas y covarianzas tenemos que asegurarnos que las serie es estacionaria. En general se busca estacionariedad débil.
- Los "sospechosos de siempre" para chequear estacionariedad
  - Análisis subjetivo en base a gráficos y funciones de correlación y correlación parcial (ver notas sobre caracterización de series estacionarias, próximo tópico)
  - Contrastes formales de raíces unitarias (ver notas sobre contrastes de raíces unitarias)
  - Tendencias (ver notas sobre tendencias y ciclos)
  - Cambios estructurales (ver notas sobre Contrastes de Chow)

¿Por qué es importante la estacionariedad?

Consideremos un proceso que sigue un **paseo aleatorio** un típico ejemplo de serie no estacionaria:

$$y_t = y_{t-1} + e_t, t = 1, 2, ..., e_t \sim i.i.d.(0, \sigma_e^2)$$

Este proceso tiene la propiedad

$$E[y_{t+h}|y_t] = y_t, \forall h \geq 1 \tag{1}$$

Esto significa que el proceso no revierte a la media y por lo tanto no es predecible. [En contraste con AR(1) $E[y_{t+h}|y_t] = \phi_1^h y_t \to 0$ cuando $h \to \infty$ (revierte a la media).]
Este es un caso especial de un proceso de **raíz unitaria** porque $\phi_1 = 1$ en un AR(1). (Ver notas sobre contrastes de raíces unitarias)

¿Por qué es importante la estacionariedad?

Un paseo aleatorio se puede escribir como

$$y_t = e_t + e_{t-1} + \ldots + e_1 + y_0$$

Entonces,

$$E[y_t] = E[e_t] + E[e_{t-1}] + \ldots + E[e_1] + E[y_0] = E[y_0]$$

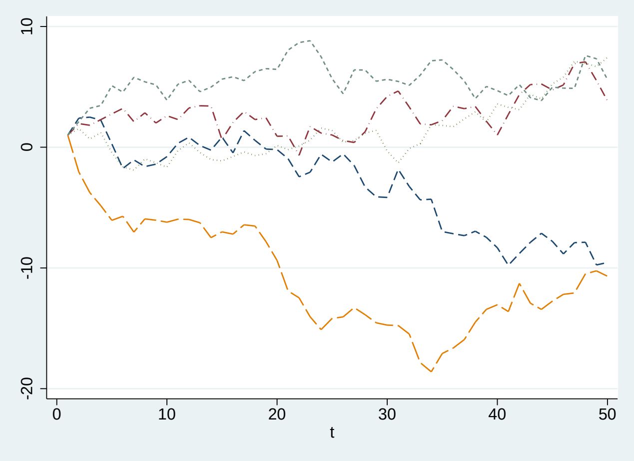
$$Var[y_t] = Var[e_t] + Var[e_{t-1}] + \ldots + Var[e_1] = t\sigma_e^2$$

(note que la varianza $\rightarrow \infty$ cuando $t \rightarrow \infty$)

$$Cov[y_t, y_{t-k}] = (t-k)\sigma_e^2$$

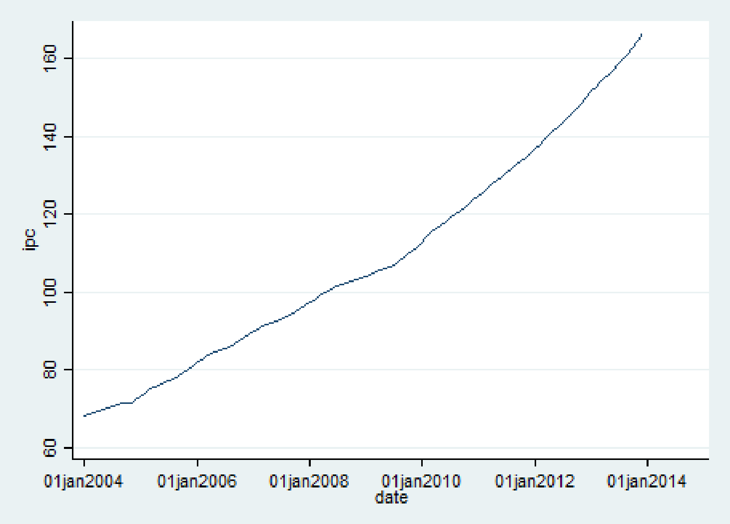
Simulación de 5 paseos aleatorios

$y_t = y_{t-1} + e_t$, $t = 1, 2, \ldots, 50$, $y_0 = 1$, $e_t \sim \ iid \ N(0, 1)$



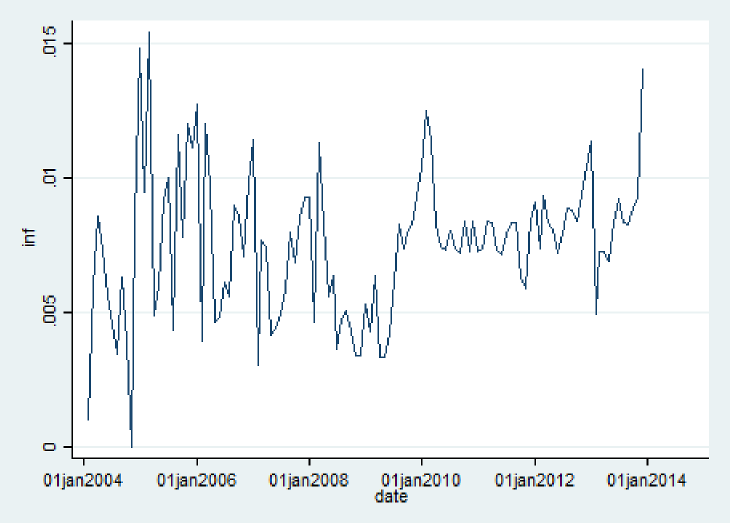
Indice de Precios al Consumidor GBA, base abril 2008=100

¿Es estacionaria o no estacionaria?



Inflación mensual, en base a IPC

¿Es estacionaria o no estacionaria?



Caracterización de series temporales

¿Cómo identificar el orden de una serie temporal?

- Para procesos MA usar la **función de autocorrelación**.
- Para procesos AR usar la **función de autocorrelación parcial**.
- Para procesos ARMA es más complicado... **función de autocorrelación extendida** (ver libro de Tsay, p.66)
- Criterios de información

La función de autocorrelación

$$\rho_k \equiv \frac{Cov[y_t, y_{t+k}]}{Var[y_t]} = \frac{\gamma_k}{\gamma_0}$$

(note que $\rho_0 = 1$)

$$\hat{\rho}_k = \frac{\sum_{t=1}^{T-k}(y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^{T-k}(y_t - \bar{y})^2}$$

La función de autocorrelación

- Considere la hipótesis $H_0 : \rho_k = 0$. Bajo la hipótesis nula $\hat{\rho}_k \sim N(0, 1/T)$ (resultado de Bartlett).
- Considere la hipótesis $H_0 : \rho_k = 0$ all $k = 1, 2, ..., K$. Bajo la hipótesis nula $Q = T \sum_{k=1}^{K} \hat{\rho}_k^2 \sim \chi_K^2$.

En `STATA`:

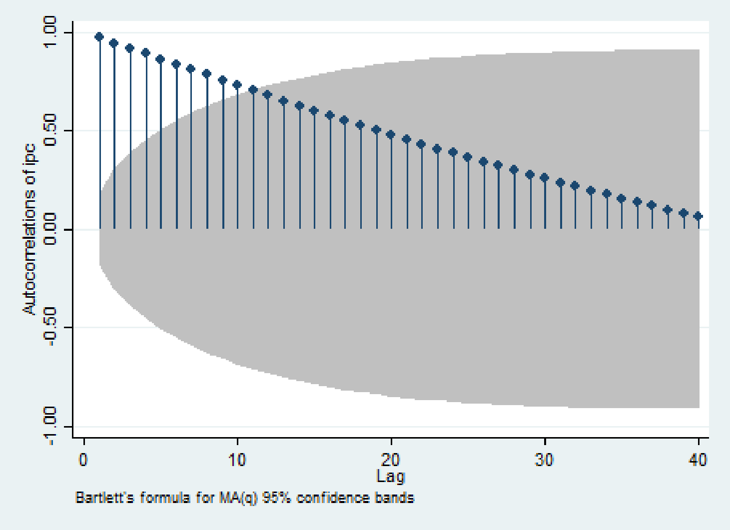
- `tsset TIEMPO`
- `ac YVAR`
- `corrgram YVAR`

(Nota: TIEMPO es la variable elegida para definir al tiempo. YVAR es una variable serie temporal.)

La función de autocorrelación

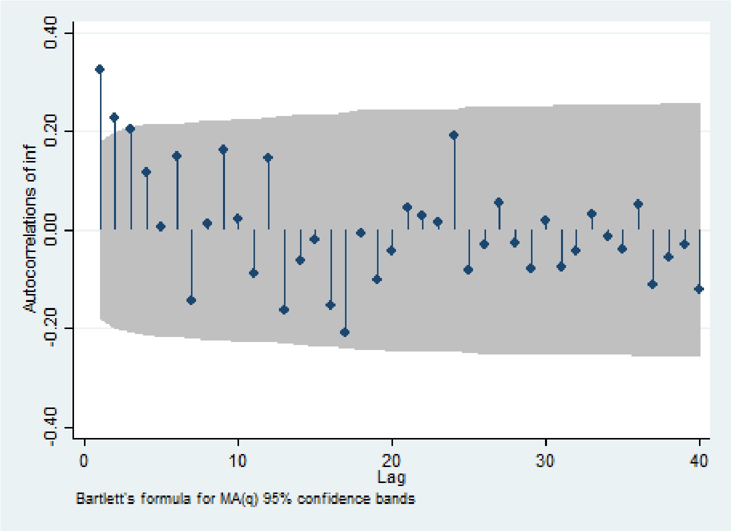
La función de autocorrelación es útil también para ver la existencia de ciclos. Por ejemplo, si la autocorrelación es alta cada ciclos de 12 meses, entonces la serie depende de lo que paso hace un año. Si esto fuera así habría alta correlación entre y_t con $y_{t-12}, y_{t+12}, y_{t-24}, y_{t+24}, y_{t-36}, y_{t+36} \cdots$.

La función de autocorrelación: IPC



(Ver http://www.indec.gov.ar/informacion-de-archivo.asp)

La función de autocorrelación: Inflación IPC



(Ver http://www.indec.gov.ar/informacion-de-archivo.asp)

La función de autocorrelación parcial

Consideremos un AR(p)

$$\gamma_k = Cov\left[y_{t-k}, \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p}\right]$$

Tomemos $k = 0, 1, \ldots, p$, para obtener $p + 1$ ecuaciones que se resuelven para $\gamma_0, \gamma_1, \ldots, \gamma_p$

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \ldots + \phi_p \gamma_p + \sigma_e^2$$

$$\gamma_1 = \phi_1 \gamma_0 + \phi_2 \gamma_1 + \ldots + \phi_{p-1} \gamma_{p-1}$$

$$\ldots\ldots\ldots$$

$$\gamma_p = \phi_1 \gamma_{p-1} + \phi_2 \gamma_{p-2} + \ldots + \phi_p \gamma_0$$

La función de autocorrelación parcial

Dividiendo por γ_0, obtenemos las ecuaciones de **Yule-Walker**

$$\rho_1 = \phi_1 + \phi_2 \rho_1 + ... + \phi_p \rho_{p-1}$$

$$.........$$

$$\rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + ... + \phi_p$$

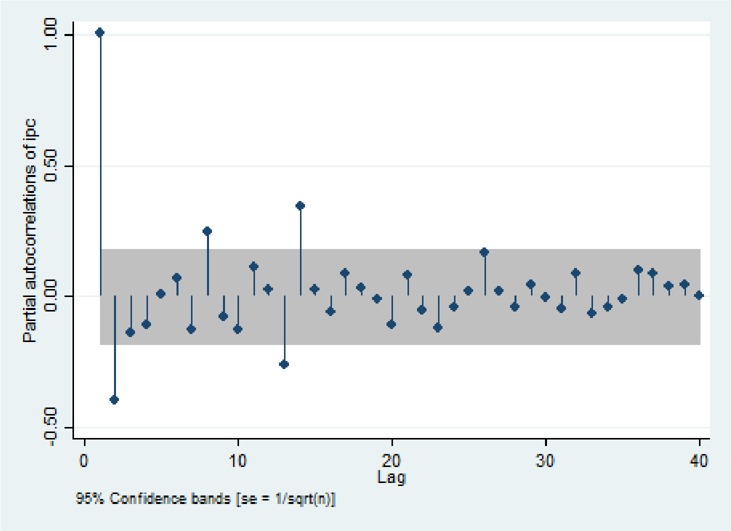
$$\rho_k = \phi_1 \rho_{p-k} + \phi_2 \rho_{p-k} + ... + \phi_p \rho_{k-p}, k > p$$

La función de autocorrelación parcial

El orden p no se conoce. Entonces se resuelve las ecuaciones de **Yule-Walker** *para valores sucesivos de p.*

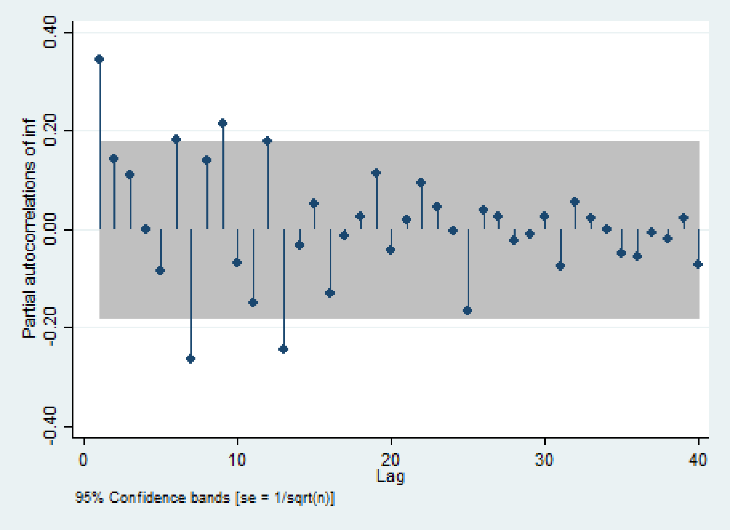
- *Asumiendo* $p = 1$ *obtenemos* $\hat{a}_1 = \hat{\phi}_1$
- *Asumiendo* $p = 2$ *obtenemos* $\hat{\phi}_1, \hat{\phi}_2$ *entonces* $\hat{a}_2 = \hat{\phi}_2$
- *Asumiendo* $p = 3$ *obtenemos* $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3$ *entonces* $\hat{a}_3 = \hat{\phi}_3$
- *Para graficar* $\hat{a}_1, \hat{a}_2, \hat{a}_3, ...$ *En* `STATA`*: pac YVAR*
- *Para test de hipótesis* $H_0 : a_k = 0$, *usamos* $\hat{a}_k \sim N(0, 1/T)$

La función de autocorrelación parcial: IPC



(Ver http://www.indec.gov.ar/informacion-de-archivo.asp)

La función de autocorrelación parcial: Inflación IPC



(Ver http://www.indec.gov.ar/informacion-de-archivo.asp)

Correlograma IPC

```
. corrgram ipc

                                          -1       0       1 -1       0       1
 LAG       AC       PAC      Q     Prob>Q  [Autocorrelation]  [Partial Autocor]
-------------------------------------------------------------------------------
1        0.9721   1.0101   116.27  0.0000          |-------           |--------
2        0.9448  -0.3957   227.02  0.0000          |-------        ---|
3        0.9176  -0.1396   332.38  0.0000          |-------          -|
4        0.8907  -0.1061   432.51  0.0000          |-------           |
5        0.8639   0.0108   527.51  0.0000          |------            |
6        0.8372   0.0677   617.52  0.0000          |------            |
7        0.8106  -0.1294   702.65  0.0000          |------           -|
8        0.7841   0.2505   783.02  0.0000          |------            |--
9        0.7576  -0.0785   858.73  0.0000          |------            |
10       0.7311  -0.1271   929.86  0.0000          |-----            -|
11       0.7044   0.1115    996.5  0.0000          |-----             |
12       0.6775   0.0250   1058.7  0.0000          |-----             |
13       0.6514  -0.2632   1116.8  0.0000          |-----           --|
14       0.6257   0.3457   1170.9  0.0000          |-----             |--
15       0.6007   0.0274   1221.2  0.0000          |----              |
16       0.5757  -0.0562   1267.8  0.0000          |----              |
17       0.5510   0.0879     1311  0.0000          |----              |
18       0.5266   0.0325   1350.8  0.0000          |----              |
19       0.5026  -0.0115   1387.4  0.0000          |----              |
20       0.4786  -0.1092   1420.9  0.0000          |---               |
21       0.4549   0.0799   1451.5  0.0000          |---               |
22       0.4315  -0.0534   1479.4  0.0000          |---               |
23       0.4087  -0.1202   1504.6  0.0000          |---               |
```

Correlograma Inflación IPC

```
. corrgram inf

                                          -1       0       1 -1       0       1
 LAG       AC       PAC      Q     Prob>Q  [Autocorrelation]  [Partial Autocor]
-------------------------------------------------------------------------------
1        0.3251   0.3437   12.899  0.0003          |--                |--
2        0.2280   0.1438   19.294  0.0001          |-                 |-
3        0.2042   0.1095   24.469  0.0000          |-                 |
4        0.1162  -0.0020   26.159  0.0000          |                  |
5        0.0067  -0.0848   26.165  0.0001          |                  |
6        0.1498   0.1818   29.022  0.0001          |-                 |-
7       -0.1435  -0.2651   31.671  0.0000         -|                --|
8        0.0127   0.1404   31.692  0.0001          |                  |-
9        0.1608   0.2143   35.079  0.0001          |-                 |-
10       0.0235  -0.0695   35.151  0.0001          |                  |
11      -0.0887  -0.1511   36.201  0.0002          |                 -|
12       0.1472   0.1778   39.119  0.0001          |-                 |-
13      -0.1625  -0.2452   42.705  0.0001         -|                 -|
14      -0.0620  -0.0336   43.233  0.0001          |                  |
15      -0.0191   0.0511   43.283  0.0001          |                  |
16      -0.1524  -0.1313   46.529  0.0001         -|                 -|
17      -0.2083  -0.0136   52.655  0.0000         -|                  |
18      -0.0077   0.0248   52.663  0.0000          |                  |
19      -0.1005   0.1139   54.117  0.0000          |                  |
20      -0.0433  -0.0430    54.39  0.0001          |                  |
21       0.0442   0.0180   54.677  0.0001          |                  |
22       0.0298   0.0926   54.809  0.0001          |                  |
23       0.0158   0.0465   54.846  0.0002          |                  |
```

Criterios de información para elegir rezagos/lags

Supongamos un modelo con K parámetros (ej., $AR(K-2)$ con una constante y un estimador de la varianza), T es el tamaño de la muestra y σ^2_{err} es el promedio de los errores al cuadrado. Los modelos de información cumplen un rol similar al R^2 ajustado, donde se penalizan la inclusión de parámetros (Parsimonia: es preferible un modelo con menos parámetros que con más.)

- Criterio de información de Akaike/Akaike information criterion (AIC): $ln(\sigma^2_{err}) + 2K/T$ (idea similar al R^2 ajustado)
- Criterio de información bayesiano/Bayesian information criterion (BIC) o Schwarz criterion (SBC, SBIC): $ln(\sigma^2_{err}) + 2Kln(T)/T$
- Rezagos/lags son seleccionados para minimizar AIC o BIC.
- Notar que $AIC \geq BIC$. Entonces en general AIC selecciona más parámetros que BIC.
- En general los criterios reemplazan $ln(\sigma^2_{err})$ por $ln(L)$ el logaritmo de la función de verosimilitud. Probar que para errores normales lo anterior es una igualdad.
- En STATA:
  - Luego de estimar el modelo (ej. `arima YVAR, arima(p,0,q)`): `estat ic`
  - En el caso de que se quiera modelar solo con procesos AR: `varsoc YVAR`

Criterios de información para elegir rezagos/lags: Inflación IPC

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
AR(1)	119	.	544.1034	3	-1082.207	**-1073.869**
AR(2)	119	.	545.3108	4	-1082.622	-1071.505
MA(1)	119	.	542.2383	3	-1078.477	-1070.139
MA(2)	119	.	543.5457	4	-1079.091	-1067.975
ARMA(1,1)	119	.	545.7594	4	**-1083.519**	-1072.402

Predicción (forecasting)

Uno de los usos fundamentales de las series temporales es predecir el futuro. La teoría tradicional de predicción se basa en dos supuestos:

1. El modelo econométrico es una buena representación del modelo económico (o de la realidad económica);
2. La estructura económica permanecerá relativamente estable en el futuro.

En base a estos supuestos entonces:

La mejor predicción se basa en el modelo que mejor "explica" los datos dentro de la muestra (in sample).

Predicción (forecasting)

- Supongamos que tenemos la serie $\{y_t\}_{t=0}^{T}$ y queremos predecir los valores futuros de y_t. Por ejemplo, la predicción un periodo hacia adelante (*one-step-ahead*) y_{T+1}, $\hat{y}_{T+1}$ (siempre le agregamos un sombrero a las variables/parámetros a predecir)
- Supongamos que Ξ_t es el conjunto de información en el momento t. Esto incluye y_t y todos sus valores anteriores $\{y_0, y_1, ..., y_{t-1}\}$ (y otras variables si la hubiere).[1]
- ¿Cuál es la mejor manera de predecir y_{T+1}?

[1] En general se usa la terminología campo sigma o en inglés sigma-field para denotar toda la información disponible.

Predicción (forecasting)

Consideremos el modelo

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

- En $T+1$, esta ecuación es $y_{T+1} = \beta_0 + \beta_1 z_{T+1} + u_{T+1}$. Si z_{T+1} es conocida y $E[u_{T+1}|\Xi_T] = 0$, entonces

$$E[y_{T+1}|\Xi_T] = \beta_0 + \beta_1 z_{T+1}$$

  Esta es la **predicción condicional** en z_{T+1}.
- Si consideramos

$$E[y_{T+1}|\Xi_T] = \beta_0 + \beta_1 E[z_{T+1}|\Xi_T]$$

  Esta es la **predicción sin condicionar**, o sea, también necesitamos ver que pasa con la variable z en $T+1$.

Predicción (forecasting)

Supogamos un modelo AR(1) $y_t = \phi_0 + \phi_1 y_{t-1} + e_t$, $E[e_t|\Xi_{t-1}] = 0$. Entonces,

- One-step-ahead forecast: $\hat{y}_T(1) = \hat{\phi}_0 + \hat{\phi}_1 y_T$. Esto es un estimador de $E[y_{T+1}|\Xi_T]$.
- Error de predicción (forecast error): $\hat{e}_T(1) = y_{T+1} - \hat{y}_T(1)$ (siempre es de esperar un error de predicción). Usando la propiedad anterior: $Var(\hat{e}_T(1)|\Xi_T) = \sigma_e^2$.
- 95% intervalo de predicción: $\hat{y}_T(1) \pm 1.96 * se(\hat{e}_T(1)|\Xi_T)$ (donde se es error estándar, o sea, la raíz cuadrada de la varianza del error)

Predicción (forecasting)

- Para predecir más periodos se debe hacer de forma iterativa. En general, cuanto más lejana es la predicción menos precisa va a ser: si $s < t$ entonces $Var\left[y_{t+1} - E[y_{t+1}|\Xi_t]\right] \leq Var[y_{t+1} - E[y_{t+1}|\Xi_s]]$.
- Entonces para construir $\hat{y}_{T+2}$ usamos

$$\hat{y}_T(2) = \hat{\phi}_0 + \hat{\phi}_1 \hat{y}_T(1) = (1 + \hat{\phi}_1)\hat{\phi}_0 + \hat{\phi}_1^2 y_T$$

$$Var(\hat{e}_T(2)) = Var(y_{T+2} - \hat{y}_T(2)) = (1 + \phi_1^2)\sigma_a^2$$

Notar que $Var(\hat{e}_T(2)|\Xi_T) \geq Var(\hat{e}_T(1)|\Xi_T)$.

- Notar que los modelos AR estacionarios satisfacen: $lim_{\ell \to \infty} \hat{y}_T(\ell) = E[y_T] = \frac{\phi_0}{1-\phi_1}$, o sea que la predicción a largo plazo debe acercarse a la media no condicional (la información por condicionar en el pasado deja de ser importante). Esta propiedad se conoce como *reversión a la media (mean reversion)*. Ver más adelante los procesos de raíces unitarias donde esto no se cumple.
- En general, el error se incrementa reflejando que sabemos menos del futuro: $Var(\hat{e}_T(\ell)) \geq Var(\hat{e}_T(h))$ si $\ell > h$. Para los procesos AR estacionarios la varianza condicional se vuelve la varianza sin condicionar. Ej: $lim_{\ell \to \infty} Var(\hat{e}(\ell)) = E[y_T] = \frac{\sigma_e^2}{1-\phi_1^2}$

Predicción (forecasting)

Supogamos un modeloMA(1) $y_t = \alpha_0 + e_t + \alpha_1 e_{t-1}$, $E[e_t|\Xi_{t-1}] = 0$. Entonces,

- One-step-ahead forecast: $\hat{y}_T(1) = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{e}_T$. Esto es un estimador de $E[y_{T+1}|\Xi_T]$. ¿De dónde sale $\hat{e}_T$? Se puede construir del modelo.
- Two-step-ahead forecast: $\hat{y}_T(2) = \hat{\alpha}_0$. Esto es un estimador de $E[y_{T+2}|\Xi_T]$. ¿Por qué?

STATA

Para modelos AR, usar el comando `var` y luego

- Para predicción sin condicionar se debe usar
  `fcast compute`
  `fcast graph`
- Para predicción condicional se debe usar (y las variables exógenas deben estar previamente especificadas)
  `predict`
- Para la predicción son tan importantes los estimadores puntuales como el intervalo de confianza. Para ello se debe especificar la opción `level(##)` donde ## es el nivel de significancia: por default es 95%.
- La predicción empieza por default en $T+1$. Pero se puede especificar el momento dentro de la muestra a partir del cual empieza la predicción usando la opción `dynamic(TIME)`, donde TIME es el tiempo a partir del cual se predice.

Para modelos MA, estimar usando el comando `arima` y luego `predict`.

Ejemplo: Predicción de inversión a largo plazo

- Consideremos un modelo de inversión:
  `webuse lutkepohl2, clear`
  `tsset qtr` /*los datos son trimestrales, la variable tiempo se llama qtr*/
  `var inv, lags(1/3) /*serie de inversión modelada con AR(3)*/`
  `fcast compute hat, step(8)` /*esto crea una nueva variable `hatinv`, 8 periodos hacia adelante*/
  `fcast graph hatinv`
- Veamos ahora un gráfico interesante, usando $qtr \leq 80$ para estimar y $qtr \in [81, 91]$ para predecir (in-sample forecast)
  `var inv if qtr<=80, lags(1/3)` /*serie de inversión modelada con AR(3), pero solo para qtr<=80*/
  `fcast compute hat80, step(8)` /*esto crea una nueva variable `hatinv80`, 8 periodos hacia adelante pero empezando desde $qtr = 81$*/
  `fcast graph hat80inv, observed`
- También se podría haber usado
  `fcast compute hat80b, step(8) dynamic(81)` /*¿Cuál es la diferencia con lo anterior?*/

Lecturas sugeridas

- Enders, W. "Applied Econometric Time Series", caps. 1,2.
- Johnston, J. y Dinardo, J. "Econometric Methods", cap. 7.