Serie Storiche
I processi ARIMA

Corso di Laurea in Statistica per le Tecnologie e le Scienze
Corso di Laurea in Statistica per l'Economia e l'Impresa
matricole dispari

Francesco Lisi

a.a. 2026/2027

Alcuni (utili) processi stocastici

- Alcuni processi stocastici *notevoli* saranno caratterizzati mediante il calcolo dei loro primi due momenti:
  - $E[Y_t]$
  - $Var(Y_t)$
  - $\gamma_k = Cov(Y_t, Y_{t-k})$
  - $\rho_k = Corr(Y_t, Y_{t-k})$
  - $P_k = Corr(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1})$
- Si studieranno, quindi, le condizioni di stazionarietà e di invertibilità.

Processo puramente casuale o *white noise*

- $\varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$ se:
  - $E[\varepsilon_t] = 0 \quad \forall t$
  - $Var[\varepsilon_t] = \sigma^2_\varepsilon \quad \forall t$
  - $Cov(\varepsilon_t, \varepsilon_s) = 0 \quad \forall t \neq s$
- le funzioni di autocorrelazione globale e parziale sono:

$$\rho_k = \begin{cases} 1 & k = 0 \\ 0 & k = \pm 1, \pm 2, \ldots \end{cases} \qquad P_k = \rho_k$$

- Approssimazione di Bartlett: $Var(\hat{\rho}_k) = \frac{1}{n}$
- Processi di questo tipo sono utili in molte situazioni, in particolare servono come *blocchi* per costruire processi più complicati come, p.e. il processi a media mobile.

Processo *Random Walk*

- Un processo Y_t è detto *random walk* se: $Y_t = Y_{t-1} + \varepsilon_t$, con $\varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$
- Per $t = 0$ solitamente si assume $Y_0 = c$ (e, talvolta, $c = 0$) così che $Y_1 = c + \varepsilon_1$ e

$$Y_t = c + \sum_{i=1}^{t} \varepsilon_i$$

- Valore atteso

$$E[Y_t] = E[c + \sum_{i=1}^{t} \varepsilon_i] = c$$

- Varianza

$$Var[Y_t] = Var[c + \sum_{i=1}^{t} \varepsilon_i] = t\sigma^2_\varepsilon$$

- Dato che la varianza dipende dal tempo si tratta di un processo non stazionario (in varianza)

Processo *Random Walk*

- Calcoliamo anche la covarianza

$$\begin{aligned}\gamma_k &= Cov[Y_t, Y_{t-k}] \\ &= E\left[\left(c+\sum_{i=1}^{t}\varepsilon_i\right)\left(c+\sum_{i=1}^{t-k}\varepsilon_i\right)\right] - E[Y_t]\cdot E[Y_{t-k}] \\ &= c^2+0+0+E\left[\left(\sum_{i=1}^{t}\varepsilon_i\right)\left(\sum_{i=1}^{t-k}\varepsilon_i\right)\right] - c^2 \\ &= E\left[\left(\sum_{i=1}^{t}\varepsilon_i\right)\left(\sum_{i=1}^{t-k}\varepsilon_i\right)\right] = (t-k)\sigma^2\end{aligned}$$

perché ci sono solo (t-k) termini contemporanei.

Processo *Random Walk*

- Dato che la covarianza dipende dal tempo, il processo RW non è stazionario neppure in covarianza.
- Si noti, inoltre, che

$$Cov[Y_t, Y_{t+k}] = t\sigma^2 \neq Cov[Y_t, Y_{t-k}]$$

perché in questo caso ci sono solo t termini contemporanei

- Correlazione:

$$Corr[Y_t, Y_{t-k}] = \frac{(t-k)\sigma^2}{\sqrt{t}\,\sigma\cdot\sqrt{t-k}\,\sigma} = \frac{\sqrt{t-k}}{\sqrt{t}}$$

$$Corr[Y_t, Y_{t+k}] = \frac{t\sigma^2}{\sqrt{t}\,\sigma\cdot\sqrt{t+k}\,\sigma} = \frac{\sqrt{t}}{\sqrt{t+k}}$$

Processo *Random Walk*

- Se, ad esempio, $t = 30$ e $k = 5$

$$\rho_{t-k} = 0.9128$$

$$\rho_{t+k} = 0.9258$$

- Tuttavia, $\forall k$ fissato

$$\lim_{t\to\infty} \rho_{t\pm k} = 1$$

Processo *Random Walk con drift*

- Si ha un *random walk* con costante (drift) se:

$$Y_t = \mu + Y_{t-1} + \varepsilon_t \qquad \text{con } \mu \neq 0$$

- In questo caso, assumendo $Y_0 = 0$, si ha

$$\begin{aligned}Y_1 &= \mu + 0 + \varepsilon_1 \\ Y_2 &= \mu + Y_1 + \varepsilon_2 = \mu + (\mu+\varepsilon_1) + \varepsilon_2 \\ Y_3 &= \mu + Y_2 + \varepsilon_3 = \mu + (2\mu+\varepsilon_1+\varepsilon_2) + \varepsilon_3 \\ &= 3\mu + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 \\ &\vdots \\ Y_t &= t\mu + \sum_{i=1}^{t}\varepsilon_i\end{aligned}$$

- Dunque $E[Y_t] = t\mu$ e $Var(Y_t) = t\sigma_\varepsilon^2$.

Processo *Random Walk* con drift

- Si noti che un RW con drift produce una dinamica che contiene un parte deterministica ($t\mu$) ed una stocastica($\sum_{i=1}^{t} \varepsilon_i$)
- $t\mu$ rappresenta un trend deterministico lineare
- $\sum_{i=1}^{t} \varepsilon_i$ rappresenta invece un trend stocastico.

$$\Delta Y_t = (1 - B) Y_t = Y_t - Y_{t-1} = \varepsilon_t$$

- Analogamente per un RW con drift si ha:

$$\Delta Y_t = (1 - B) Y_t = Y_t - Y_{t-1} = \mu + \varepsilon_t,$$

Processo a media mobile, $MA(q)$

- Sia $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. Un processo a media mobile di ordine q è così definito:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

  dove i θ_i sono parametri.
- Definito l'operatore ritardo come:

$$B^k Y_t = Y_{t-k} \qquad B^0 Y_t = Y_t$$

  possiamo riscrivere il processo in termini dell'operatore polinomiale a media mobile:
- $\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)$ e quindi:

$$Y_t = \theta(B) \varepsilon_t$$

Momenti di un processo $MA(q)$

- Valore atteso

$$E[Y_t] = E[\varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}] = 0$$

- Covarianza (e Varianza)

$$\begin{aligned} \gamma_k &= E[Y_t Y_{t-k}] \\ &= E[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}) \cdot \\ &\qquad (\varepsilon_{t-k} - \theta_1 \varepsilon_{t-k-1} - \ldots - \theta_q \varepsilon_{t-k-q})] \end{aligned}$$

$$\gamma_k = \begin{cases} \sigma_\varepsilon^2 \left(1 + \sum_{i=1}^{q} \theta_i^2\right) & k = 0 \\ \sigma_\varepsilon^2 (-\theta_k + \theta_1 \theta_{k+1} + \ldots + \theta_{q-k} \theta_q) & k = 1, \ldots, q \\ 0 & k > q \end{cases}$$

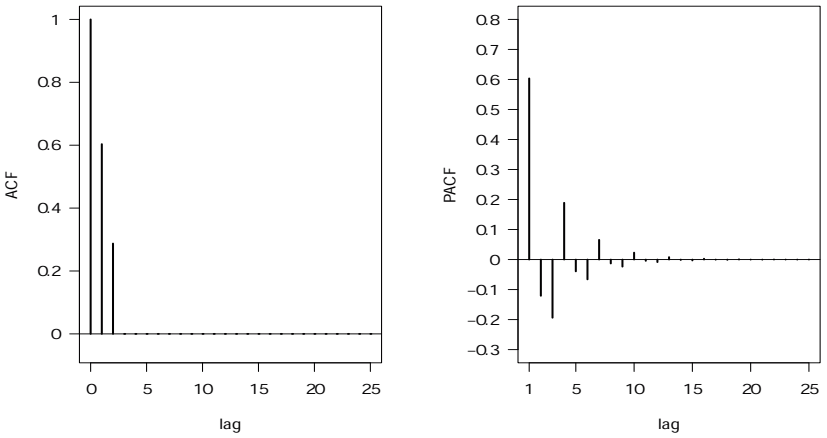
Momenti di un processo $MA(q)$

- Funzione di autocorrelazione globale (ACF)

$$\rho_k = \begin{cases} 1 & k = 0 \\ \frac{-\theta_k + \theta_1 \theta_{k+1} + \ldots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \ldots + \theta_q^2} & k = 1, \ldots, q \\ 0 & k > q \end{cases}$$

  ρ_k è diversa da zero solo per i primi q ritardi, mentre per $k > q$ vale zero.
- Per quanto riguarda la funzione di autocorrelazione parziale (PACF), P_k tende ad annullarsi al divergere di k, con comportamento misto tra l'esponenziale e lo pseudo-periodico

Funzioni di autocorrelazione teoriche di un processo MA(2)



Stazionarietà ed invertibilità di un processo MA(q)

- Un processo $MA(q)$ è sempre stazionario.
- Tuttavia, questo non è sufficiente e dobbiamo richiedere ai coefficienti θ_i che soddisfino un'ulteriore condizione detta invertibilità.
- La condizione di invertibilità è soddisfatta, e quindi il processo $MA(q)$ si dice invertibile se, data l'equazione caratteristica

$$\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q) = 0$$

tutte le radici B_i, $i = 1, 2, \ldots, q$ sono in modulo maggiori di uno o, equivalentemente se tutte le radici, B_i, dell'equazione:

$$B^q - \theta_1 B^{q-1} - \theta_2 B^{q-2} - \ldots - \theta_{q-1} B - \theta_q = 0$$

sono in modulo minori di uno.

Ancora sulla condizione di invertibilità

- Consideriamo i seguenti due processi $MA(1)$:
  a. $Y_t = \varepsilon_t + \theta\varepsilon_{t-1}$ $\quad \varepsilon_t \sim N(0, \sigma^2)$ $\quad | \theta | < 1$
  b. $Z_t = \eta_t + \frac{1}{\theta}\eta_{t-1}$ $\quad \eta_t \sim N(0, \theta^2\sigma^2)$ $\quad | \theta | < 1$
- Si può facilmente dimostrare che questi due processi hanno la stessa funzione di autocorrelazione (provare!).
- Lo stesso vale per
  a. $Y_t = \varepsilon_t + \theta\varepsilon_{t-1}$ $\quad \varepsilon_t \sim N(0, \sigma^2)$ $\quad | \theta | < 1$
  b. $Z_t = \eta_t + \frac{1}{\theta}\eta_{t-1}$ $\quad \eta_t \sim N(0, \sigma^2)$ $\quad | \theta | < 1$
- Questo significa che non è possibile identificare in modo univoco un processo MA a partire dalla sua ACF.
- Come fare?

Ancora sulla condizione di invertibilità

- Se riscriviamo gli ultimi due modelli ponendo ε_t in funzione di $Y_t, Y_{t-1}, \ldots,$ otteniamo:
  a. $\varepsilon_t = Y_t - \theta Y_{t-1} + \theta^2 Y_{t-2} - \ldots$
  b. $\varepsilon_t = Z_t - \frac{1}{\theta} Z_{t-1} + \frac{1}{\theta^2} Z_{t-2} - \ldots$
- Se $| \theta | < 1$ i coefficienti del primo modello convergono a zero, mentre nel secondo caso divergono.
- Quindi una procedura di stima che coinvolge la stima dei residui porta, naturalmente, a prediligere il modello *a*.
- Così se $| \theta | < 1$ il modello *a*. è detto essere invertibile, mentre non lo è il modello *b*.
- L'imposizione della condizione di invertibilità assicura che esiste un unico processo MA per una data funzione di autocorrelazione.

Processo $MA(1)$

- Quando $\theta(B) = (1 - \theta B)$ si parla di processo $MA(1)$:

$$Y_t = \varepsilon_t - \theta\varepsilon_{t-1} = (1 - \theta B)\varepsilon_t$$

- Il processo è invertibile quando le radici dell'equazione caratteristica $\theta(B) = 0$ sono in modulo maggiori di 1, quindi

$$1 - \theta B = 0$$

ha come soluzione:

$$B = \frac{1}{\theta} \;\Rightarrow\; \left|\frac{1}{\theta}\right| > 1 \;\Rightarrow\; |\theta| < 1$$

- **Attenzione!** questo risultato è valido solo per processi $MA(1)$

Momenti di un processo $MA(1)$

- Valore atteso:

$$E[Y_t] = E[\varepsilon_t - \theta\varepsilon_{t-1}] = E[\varepsilon_t] - \theta E[\varepsilon_{t-1}] = 0$$

- Autocovarianza

$$\gamma_k = \begin{cases} \sigma_\varepsilon^2(1+\theta^2) & k = 0 \\ -\theta\sigma_\varepsilon^2 & k = 1 \\ 0 & k \geq 2 \end{cases}$$

Autocovarianza

$$\begin{aligned}
\gamma_0 &= Var(Y_t) = E[Y_t^2] = E[(\varepsilon_t - \theta\varepsilon_{t-1})^2] = \\
&= E[\varepsilon_t^2 + \theta^2\varepsilon_{t-1}^2 - 2\theta\varepsilon_t\varepsilon_{t-1}] = E[\varepsilon_t^2] + \theta^2 E[\varepsilon_{t-1}^2] - 2\theta E[\varepsilon_t\varepsilon_{t-1}] \\
&= \sigma_\varepsilon^2 + \theta^2\sigma_\varepsilon^2 = \sigma_\varepsilon^2(1+\theta^2)
\end{aligned}$$

$$\begin{aligned}
\gamma_1 &= Cov(Y_t, Y_{t-1}) = E[(\varepsilon_t - \theta\varepsilon_{t-1})(\varepsilon_{t-1} - \theta\varepsilon_{t-2})] = \\
&= E[\varepsilon_t\varepsilon_{t-1} - \theta\varepsilon_t\varepsilon_{t-2} - \theta\varepsilon_{t-1}^2 + \theta^2\varepsilon_{t-1}\varepsilon_{t-2}] = \\
&= E[\varepsilon_t\varepsilon_{t-1}] - \theta E[\varepsilon_t\varepsilon_{t-2}] - \theta E[\varepsilon_{t-1}^2] + \theta^2 E[\varepsilon_{t-1}\varepsilon_{t-2}] = -\theta\sigma_\varepsilon^2
\end{aligned}$$

$$\begin{aligned}
\gamma_2 &= Cov(Y_t, Y_{t-2}) = E[(\varepsilon_t - \theta\varepsilon_{t-1})(\varepsilon_{t-2} - \theta\varepsilon_{t-3})] = \\
&= E[\varepsilon_t\varepsilon_{t-2} - \theta\varepsilon_t\varepsilon_{t-3} - \theta\varepsilon_{t-1}\varepsilon_{t-2} + \theta^2\varepsilon_{t-1}\varepsilon_{t-3}] = \\
&= E[\varepsilon_t\varepsilon_{t-2}] - \theta E[\varepsilon_t\varepsilon_{t-3}] - \theta E[\varepsilon_{t-1}\varepsilon_{t-2}] + \theta^2 E[\varepsilon_{t-1}\varepsilon_{t-3}] = 0
\end{aligned}$$

Funzioni di autocorrelazione

- Autocorrelazione globale

$$\rho_k = \begin{cases} 1 & k = 0 \\ -\dfrac{\theta}{1+\theta^2} & k = 1 \\ 0 & k \geq 2 \end{cases}$$

- Autocorrelazione parziale

$$P_k = \frac{-\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}} \quad k \geq 1$$

PACF di un processo MA(1)

- $k = 1$

$$P_1 = \frac{-\theta(1-\theta^2)}{(1-\theta^4)} = \frac{-\theta(1-\theta^2)}{(1+\theta^2)(1-\theta^2)} = -\frac{\theta}{1+\theta^2}$$

- $k = 2$

$$P_2 = \frac{-\theta^2(1-\theta^2)}{(1-\theta^6)} = \frac{-\theta^2+\theta^4}{1-\theta^6}$$

- $k = 3$

$$P_3 = \frac{-\theta^3(1-\theta^2)}{(1-\theta^8)} = \frac{-\theta^3+\theta^5}{1-\theta^8}$$

PACF di un processo MA(1)

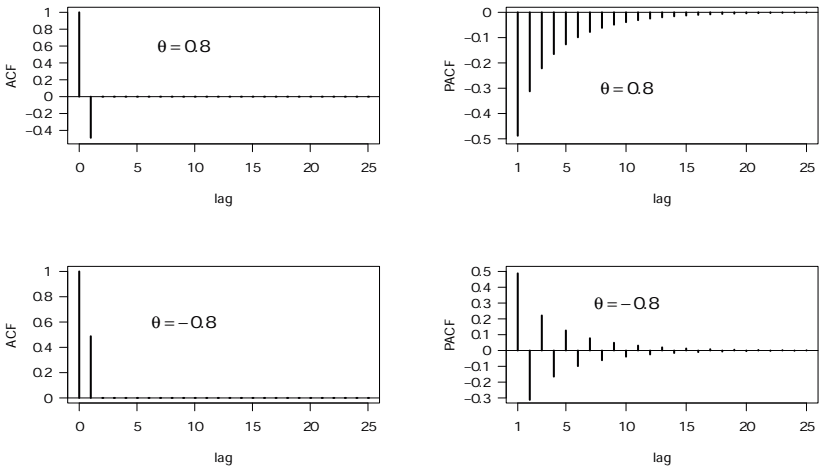
- Dato che $|\theta| < 1$

$$\lim_{k\to\infty} \theta^k = 0$$

e questo implica che, al crescere di k, il denominatore di ρ_k converge a 1 e il numeratore è dominato da θ^k, che converge a zero come $-(\theta^k)$

- se $\theta > 0$, allora tutte le autocorrelazioni parziali sono negative, perché $-\theta^k < 0$;
- se $\theta < 0$, allora
  $-(\theta^k) > 0$ per valori dispari di k e
  $-(\theta^k) < 0$ per valori pari di k.

Esempi di ACF e PACF



ACF e PACF per $\theta = 0.8$ e $\theta = -0.8$

Processo AutoRegressivo $AR(p)$

- Sia $\varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$. Un processo autoregressivo di ordine p è così definito:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2+} \ldots + \phi_p Y_{t-p} + \varepsilon_t$$

dove i ϕ_i, $i = 0, 1, \ldots, p$ sono parametri.

- Utlizzando l'operatore ritardo possiamo riscrivere il processo in termini dell'operatore polinomiale autoregressivo:

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)$$

e quindi:

$$\phi(B) Y_t = \phi_0 + \varepsilon_t$$

- Un processo autoregressivo è un particolare modello di regressione multipla in cui la variabile dipendente, Y_t, è regredita non su altre variabili indipendenti, bensì sui suoi valori passati, da cui il prefisso *auto*.

Valore atteso di un $AR(p)$

- Valore atteso:

$$E[Y_t] = \frac{\phi_0}{1 - \phi_1 - \ldots - \phi_p} \equiv \mu$$

- $\phi_0 = 0 \quad \Longleftrightarrow \quad E[Y_t] = \mu = 0$
- Talvolta è preferibile ipotizzare di avere media nulla. Questa condizione si può sempre ottenere lavorando con la serie degli scarti dalla media:

$$y_t - \bar{y}$$

Autocovarianza di un $AR(p)$

- Assumiamo $\mu = 0$ ($\phi_0 = 0$) $\Rightarrow E[y_t] = 0$.
-

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t.$$

moltiplicando tutto per y_{t-k} si ha:

$$y_t y_{t-k} = \phi_1 y_{t-1} y_{t-k} + \phi_2 y_{t-2} y_{t-k} + \cdots + \phi_p y_{t-p} y_{t-k} + \varepsilon_t y_{t-k}.$$

Prendiamo ora il valore atteso; dato che $E[y_t] = 0$, si ha che $\gamma_k = E[y_t y_{t-k}]$ e:

$$\begin{aligned} E[y_t y_{t-k}] &= \phi_1 E[y_{t-1} y_{t-k}] + \cdots + \phi_p E[y_{t-p} y_{t-k}] + E(\varepsilon_t y_{t-k-p}) \\ &= \phi_1 E[y_{t-1} y_{t-k}] + \cdots + \phi_p E[y_{t-p} y_{t-k}] \\ &\quad + E\left[\varepsilon_t \left(\phi_1 y_{t-k-1} + \cdots + \phi_p y_{t-k-p} + \varepsilon_{t-k}\right)\right] \end{aligned}$$

L'ultimo termine vale 0 per $k \neq 0$ e vale σ_ε^2 per $k = 0$, quindi Quindi:

$$\gamma_k = \begin{cases} \phi_1 \gamma_1 + \phi_2 \gamma_2 + \cdots + \phi_p \gamma_p + \sigma_\varepsilon^2, & k = 0, \\ \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p}, & k > 0. \end{cases}$$

- Ora permettiamo che $\phi_0 \neq 0$; allora

$$\mu = \frac{\phi_0}{1 - \sum_i \phi_i}.$$

- Possiamo riscrivere il processo come

$$(y_t - \mu) = \phi_1 (y_{t-1} - \mu) + \cdots + \phi_p (y_{t-p} - \mu) + \varepsilon_t.$$

- Moltiplicando tutto per $(y_{t-k} - \mu)$ sia ha

$$\begin{aligned} (y_t - \mu)(y_{t-k} - \mu) &= \phi_1 (y_{t-1} - \mu)(y_{t-k} - \mu) + \cdots \\ &\quad + \phi_p (y_{t-p} - \mu)(y_{t-k} - \mu) + \varepsilon_t (y_{t-k} - \mu). \end{aligned}$$

Se ora ne prendiamo i valori attesi si torna esattamente alla covarianza appena vista.
Quindi lavorare nel processo y_t o sui suoi scarti dalla media è lo stesso.

Autocorrelazione di un $AR(p)$

- Assumiamo $\phi_0 = 0$
- Autocovarianza

$$\gamma_k = \begin{cases} \phi_1\gamma_1 + \phi_2\gamma_2 + \ldots + \phi_p\gamma_p + \sigma_\varepsilon^2 & k = 0 \\ \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} + \ldots + \phi_p\gamma_{k-p} & k > 0 \end{cases}$$

- Autocorrelazione

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \ldots + \phi_p\rho_{k-p} \quad k > 0$$

- Equazioni di Yule-Walker

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \ldots + \phi_p\rho_{k-p} \quad k = 1, \ldots, p$$

ACF e PACF di un AR(p)

- L'ACF ρ_k tende ad annullarsi al divergere di k, con comportamento misto tra l'esponenziale e lo pseudo-periodico
- La PACF P_k è diversa da zero per $k \leq p$ e si annulla per $k > p$.
- Questa proprietà è utile nell'identificare un modello $AR(p)$ come PGD per una data serie storica.

PACF di un processo $AR(1)$

- $\rho_k = \phi^k$
- k=1 $\qquad P_1 = \rho_1 = \phi$
- k=2

$$P_2 = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{\phi^2 - \phi^2}{1 - \phi^2} = 0$$

- k=3

$$\begin{aligned} P_3 &= \frac{\rho_3 - \rho_2\rho_1 - \rho_1^2\rho_3 + \rho_1\rho_2^2 + \rho_1\rho_2 + \rho_2^2}{1 - 2\rho_2^2 + 2\rho_1^2\rho_2 - \rho_2^2} \\ &= \frac{\phi^3 - \phi^2\phi - \phi^2\phi^3 + \phi\phi^4 + \phi^2\phi^2 - \phi^4}{1 - 2\phi^4 + 2\phi^2\phi^2 - \phi^4} = 0 \end{aligned}$$

Stazionarietà e invertibilità di un processo $AR(p)$

- Un processo $AR(p)$ è sempre invertibile, ma
- è stazionario solo se tutte le p radici, B_i, dell'equazione caratteristica $\phi(B) = 0$ sono in modulo maggiori di uno
- Esempio: consideriamo un processo $AR(1)$, cioè $Y_t = \phi Y_{t-1} + \varepsilon_t$ o $(1 - \phi B)Y_t = \varepsilon_t$. L'equazione caratteristica è, quindi:

$$\phi(B) = 1 - \phi B = 0 \;\Rightarrow\; B = \frac{1}{\phi} \;\Rightarrow\; \left|\frac{1}{\phi}\right| > 1 \;\Rightarrow\; |\phi| < 1$$

- **Attenzione!** questo risultato è valido solo per processi $AR(1)$. Per $AR(2)$ vedi allegato: *La stazionarietà di un processo $AR(2)$*.

Momenti di un processo $AR(1)$: $Y_t = \phi Y_{t-1} + \varepsilon_t$

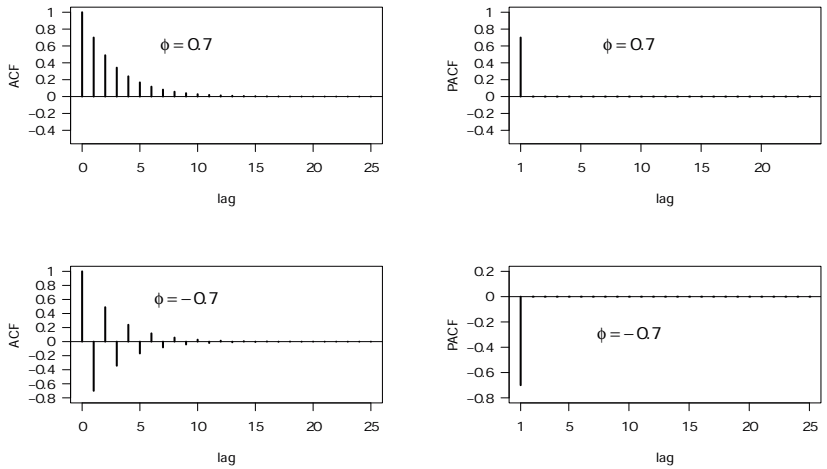
- Valore atteso: $E[Y_t] = 0$
- Autocovarianza

$$\gamma_k = \begin{cases} \dfrac{\sigma_\varepsilon^2}{1-\phi^2} & k = 0 \\ \phi_1 \gamma_{k-1} = \phi^k \gamma_0 & k > 0 \end{cases}$$

- Autocorrelazione globale e parziale

$$\rho_k = \phi^k \quad k > 0, \qquad P_k = \begin{cases} \phi & k = 1 \\ 0 & k > 1 \end{cases}$$

Esempi di ACF e PACF



ACF e PACF per $\phi = 0.7$ e $\phi = -0.7$

La relazione duale fra processi $AR(p)$ e $MA(q)$

- Un processo $AR(p)$ stazionario $\phi(B)Y_t = \varepsilon_t$ può sempre scriversi come un processo $MA(\infty)$:

$$Y_t = \frac{1}{\phi(B)}\varepsilon_t = \psi(B)\varepsilon_t$$

  con $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots$ tale che $\phi(B)\psi(B) = 1$.
- In modo duale, un processo $MA(q)$ invertibile $Y_t = \theta(B)\varepsilon_t$ può sempre scriversi come un processo $AR(\infty)$:

$$\pi(B)Y_t = \frac{1}{\theta(B)}Y_t = \varepsilon_t$$

  dove $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 \ldots$ è tale che $\theta(B)\pi(B) = 1$.
- I pesi ψ_i e π_i si trovano eguagliando i coefficienti di B^j nelle equazioni $\phi(B)\psi(B) = 1$ e $\theta(B)\pi(B) = 1$.

Esempio

- Per un processo $AR(1)$ si ha:

$$Y_t = \phi Y_{t-1} + \varepsilon_t = \phi(\phi Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \ldots = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$$

  In questo caso $\psi_j = \phi^j$ con$\psi_0 = 1$.
- Per un processo $MA(1)$ si ha:

$$Y_t = \theta\varepsilon_{t-1} + \varepsilon_t$$

  da cui

$$\varepsilon_t = Y_t - \theta\varepsilon_{t-1} = Y_t - \theta(Y_{t-1} - \theta\varepsilon_{t-2}) = \ldots = \sum_{j=0}^{\infty} -\theta^j Y_{t-j}$$

  con $\pi_j = -\theta^j$ e $\pi_0 = 1$.

Distribuzioni condizionate e loro momenti

- Caso AR(1)

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t$$

- Momenti non condizionati di un AR(1):

$$E[Y_t] = \frac{\phi_0}{1-\phi_1} \qquad V[Y_t] = \frac{\sigma_\varepsilon^2}{1-\phi_1^2}$$

- Momenti condizionati di un AR(1):

$$E[Y_t|I_{t-1}] = \phi_0 + \phi_1 y_{t-1} \qquad V[Y_t|I_{t-1}] = \sigma_\varepsilon^2$$

- Se poi gli ε_t sono gaussiani anche le distribuzioni non condizionate e condizionate degli Y_t sono gaussiane

Distribuzioni condizionate e loro momenti

- Caso MA(1)

$$Y_t = \varepsilon_t - \theta\varepsilon_{t-1}$$

- Momenti non condizionati di un MA(1):

$$E[Y_t] = 0 \qquad V[Y_t] = \sigma_\varepsilon^2(1+\theta^2)$$

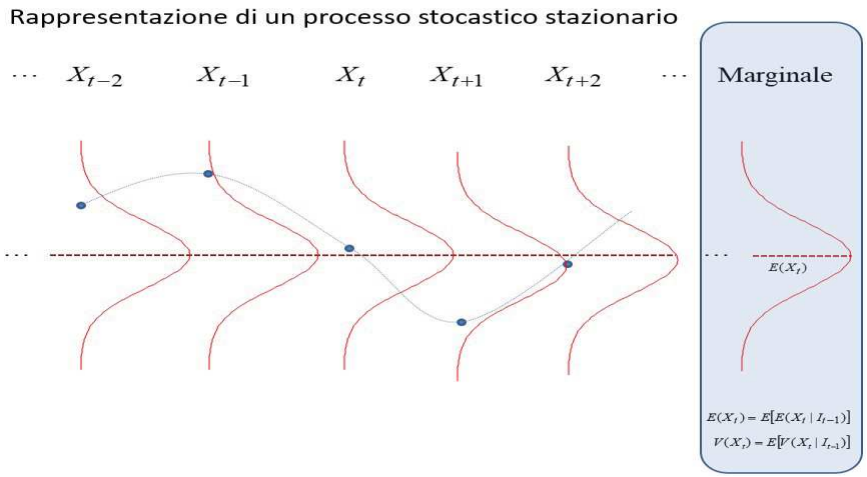
- Momenti condizionati di un MA(1):

$$E[Y_t|I_{t-1}] = -\theta\hat{\varepsilon}_{t-1} \qquad V[Y_t|I_{t-1}] = \sigma_\varepsilon^2$$

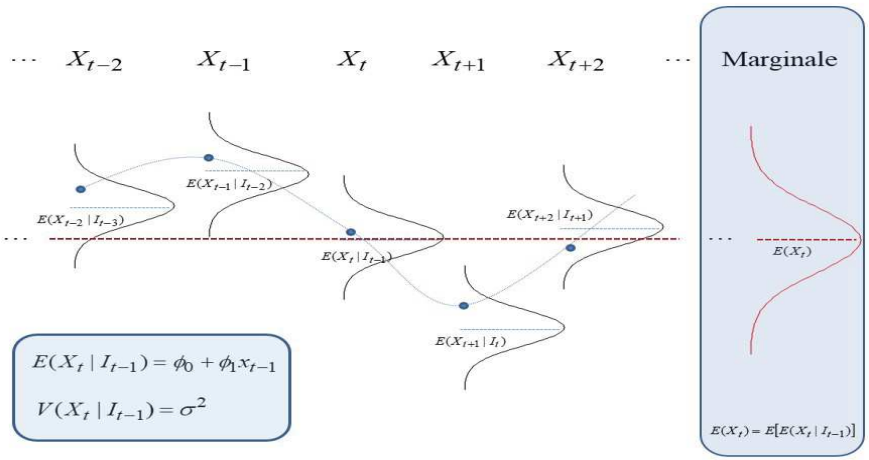
con $\hat{\varepsilon}_t$ scritto in funzione delle y_t passate.

- Ancora, se gli ε_t sono gaussiani anche le distribuzioni non condizionate e condizionate degli Y_t sono gaussiane

Distribuzioni non condizionate (marginali)



Distribuzioni condizionate



Esercizio

Si consideri il processo $AR(2)$ dato da:

$$Y_t = Y_{t-1} - 0.5Y_{t-2} + \varepsilon_t$$

i. Si dica se il processo è stazionario.
ii. Si calcolino i primi due coefficienti di autocorrelazione.

Soluzione:

i. Per rispondere a questa domanda dobbiamo cercare le radici dell'equazione caratteristica:

$$\phi(B) = 1 - B + 0.5B^2 = 0$$

e vedere se sono in modulo maggiori di uno. Si ha:

$$B_{1,2} = \frac{1 \pm \sqrt{1-2}}{2 \cdot 0.5} = 1 \pm i$$

Poichè $\mid B_{1,2} \mid > 1$ le radici sono fuori il cerchio di raggio unitario e quindi il processo è stazionario.

ii. Per calcolare i coefficienti di autocorrelazione possiamo utilizzare le equazioni di Yule-Walker, dalla prima abbiamo:

$$\rho_1 = \rho_0 - 0.5\rho_{-1} = 1 - 0.5\rho_1$$

che porta a $\rho_1 = 2/3$. Per $k = 2$ si ha:

$$\rho_2 = \rho_1 - 0.5\rho_0 = \frac{2}{3} - 0.5 = \frac{1}{6}$$

Processi AutoRegressivi a Media Mobile - $ARMA(p, q)$

- Sia $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. Un processo $ARMA(p, q)$ è così definito:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

dove i $\phi_i,\ i = 0, 1, \ldots, p$ e i $\theta_i,\ i = 1, \ldots, q$ sono parametri.

- Utlizzando l'operatore ritardo possiamo riscrivere il processo come:

$$\phi(B)Y_t = \phi_0 + \theta(B)\varepsilon_t$$

- Un processo $ARMA(p, q)$ è quindi un modello *misto*.
- Se $p = 0$ ritroviamo il processo $MA(q)$ se, invece, $q = 0$ ritroviamo il processo $AR(p)$.

Stazionarietà e invertibilità di un $ARMA(p, q)$

- Come per un processo AR, affinchè il processo sia stazionario le radici dell'equazione caratteristica:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p = 0$$

devono essere in modulo maggiori di uno.

- Come per un processo MA, affinchè il processo sia invertibile le radici dell'equazione caratteristica:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q = 0$$

devono essere in modulo maggiori di uno.

- Si assume, inoltre, che $\phi(B) = 0$ e $\theta(B) = 0$ non abbiano radici in comune.

Momenti

- Valore atteso

$$E[Y_t] = \frac{\phi_0}{1 - \phi_1 - \ldots - \phi_p} \equiv \mu$$

- Autocovarianza

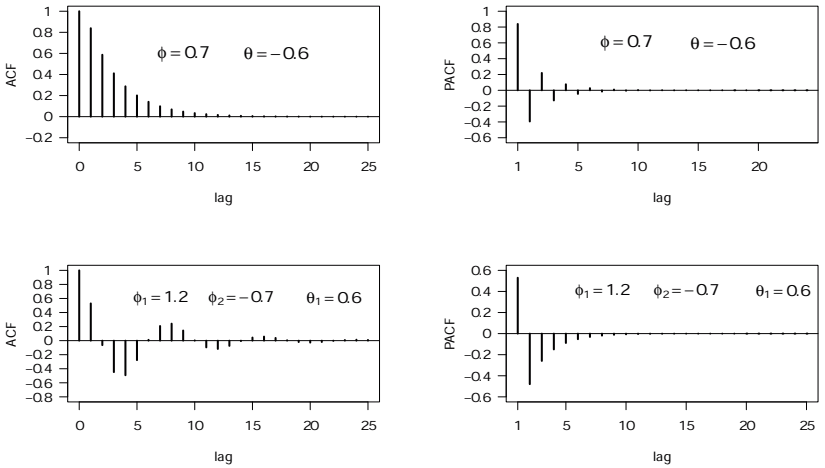
$$\gamma_k = \begin{cases} \phi_1\gamma_{k-1} + \ldots + \phi_p\gamma_{k-p} - \theta_1 E[Y_{t-k}\varepsilon_{t-1}] - \ldots - \theta_q E[Y_{t-k}\varepsilon_{t-q}] & \\ & k = 0, 1, \ldots, q \\ \phi_1\gamma_{k-1} + \ldots + \phi_p\gamma_{k-p} & k \geq q+1 \end{cases}$$

ACF e PACF

- ACF

$$\rho_k = \begin{cases} \frac{\gamma_k}{\gamma_0} & k = 0, 1, \ldots, q \\ \phi_1\rho_{k-1} + \ldots + \phi_p\rho_{k-p} & k \geq q+1 \end{cases}$$

- Fino al ritardo q l'ACF ha un comportamento *misto*. Da $q+1$ in poi il comportamento è quello proprio di un $AR(p)$
- PACF: Comportamento simmetrico per la PACF. Andamento *misto* fino al ritardo p. Per $k > p$ comportamento analogo alla PACF di un processo $MA(q)$.



ACF e PACF per un ARMA(1,1) con $\phi = 0.7$ e $\theta = -0.6$, e per un ARMA(2,1) con $\phi_1 = 1.2$, $\phi_2 = -0.7$ e $\theta = 0.6$

Osservazioni

- In generale è più difficile riconoscere le funzioni di autocorrelazione di un $ARMA(p, q)$
- L'importanza dei processi ARMA sta nel fatto che tali modelli possono descrivere una serie stazionaria in modo più parsimonioso di un modello AR puro o MA puro.
- Un modello $ARMA(p, q)$ stazionario e invertibile ammette una rappresentazione $AR(\infty)$:

$$\pi(B)Y_t = \varepsilon_t, \quad \pi(B) = 1 - \sum_{i=1}^{\infty} \pi_i B^i$$

e una rappresentazione $MA(\infty)$

$$Y_t = \psi(B)\varepsilon_t, \quad \psi(B) = \sum_{i=0}^{\infty} \psi_i B^i$$

Processo $ARMA(1,1)$

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

- Il processo è stazionario se $|\phi_1| < 1$
- Il processo è invertibile se $|\theta_1| < 1$
- Valore atteso
$$E[Y_t] = \frac{\phi_0}{1-\phi_1}$$
- Varianza
$$Var(Y_t) = \gamma_0 = \frac{1+\theta_1^2-2\theta_1\phi_1}{1-\phi_1^2}\sigma_\varepsilon^2$$
- Nel seguito, per semplicità, ipotizzeremo processi a media nulla ossia $\phi_0 = 0$

Momenti di un processo $ARMA(1,1)$

- Autocovarianza
$$\gamma_k = \begin{cases} \frac{(\phi_1-\theta_1)(1-\theta_1\phi_1)}{1-\phi_1^2}\sigma_\varepsilon^2 & k=1 \\ \phi_1\gamma_{k-1} = \phi_1^{k-1}\gamma_1 & k \geq 2 \end{cases}$$
- Autocorrelazione
$$\rho_k = \begin{cases} 1 & k=0 \\ \frac{(\phi_1-\theta_1)(1-\theta_1\phi_1)}{1+\theta_1^2-2\phi_1\theta_1} & k=1 \\ \phi_1^{k-1}\rho_1 & k \geq 2 \end{cases}$$

ACF e PACF per processi ARMA

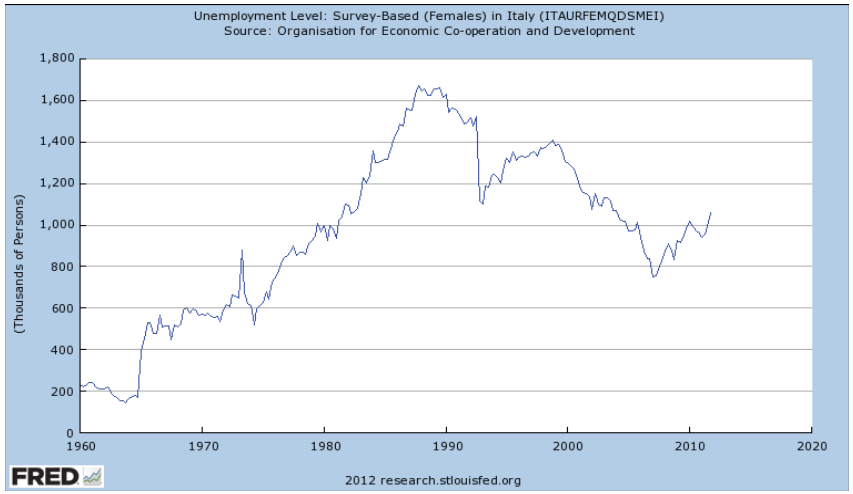
- L'andamento dell'ACF è *governato* dalla parte autoregressiva, per cui tende ad annullarsi al crescere di k.
- L'andamento della PACF è *governato* dalla parte a media mobile, per cui tende ad annullarsi al crescere di k
- Schema dei comportamenti ARMA (non stagionali)

Processo	ACF	PACF
AR(p)	Tende ad annullarsi	Nulla dopo il ritardo p
MA(q)	Nulla dopo il ritardo q	Tende ad annullarsi
ARMA(p,q)	Tende ad annullarsi	Tende ad annullarsi

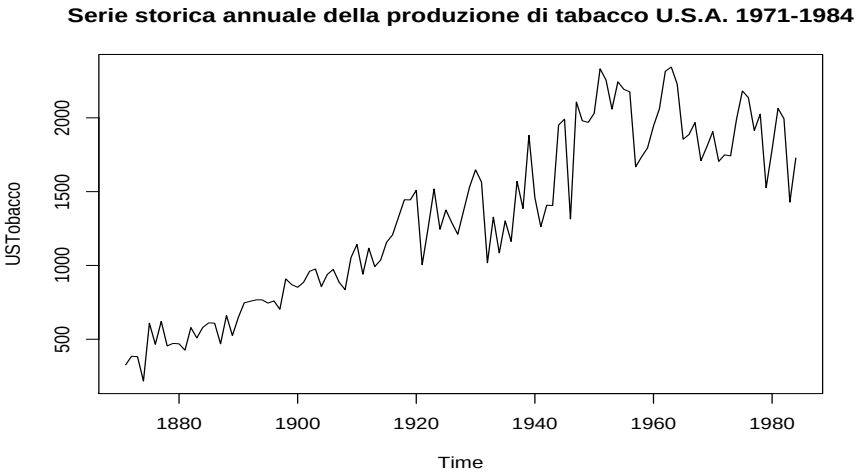
Processi stocastici non stazionari

- I processi per serie storiche discussi fino ad ora sono tutti stazionari. Tuttavia, molte serie storiche reali, soprattutto in ambito economico e finanziario (ma non solo) sono non stazionarie.
- Rispetto alla classe dei processi stazionari in covarianza, la non stazionarietà può rivelarsi in modi diversi:
  - non stazionarietà in media, ossia μ_t che varia nel tempo
  - non stazionarietà in varianza, ossia σ_t^2 che varia nel tempo
  - non stazionarietà in covarianza
  - ovviamente più non stazionarietà possono coesistere
- Differenza fra trend stocastico e trend deterministico

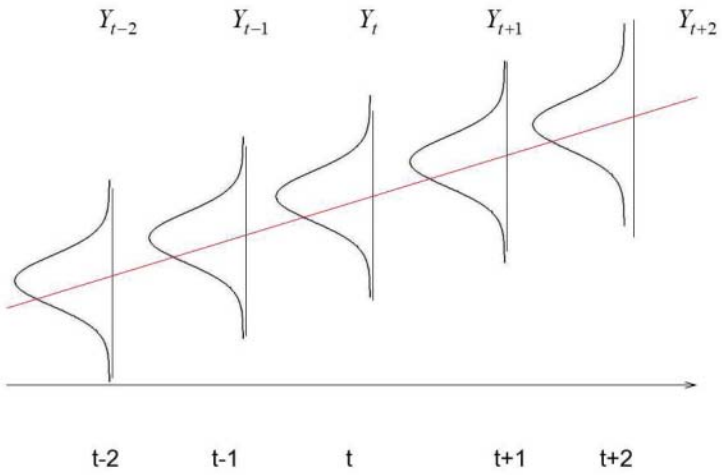
Non stazionarietà in media



Non stazionarietà in varianza



Non stazionarietà in media: esempio

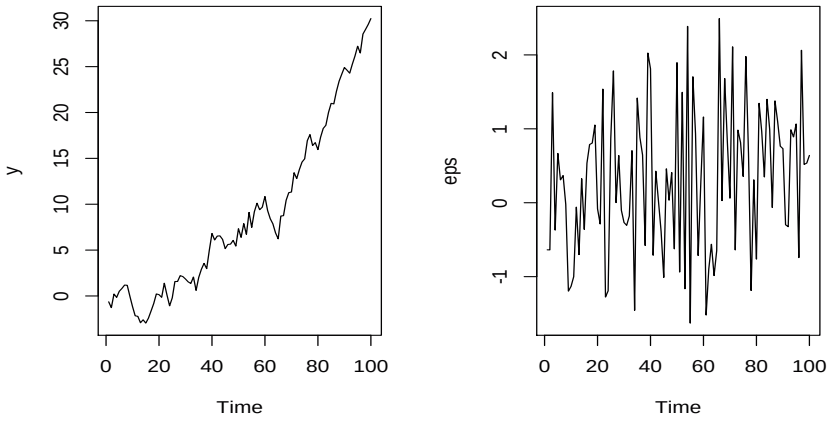


Processi non stazionari in media

- Alcuni tipi di non stazionarietà possono essere eliminati mediante opportune trasformazioni
- Un processo non stazionario che diventa stazionario a seguito di d differenziazioni successive si dice processo non stazionario omogeneo di grado d. Per esempio, sia $Y_t = \beta_0 + \beta_1 t + \varepsilon_t$, una volta differenziato abbiamo

$$(1 - B)Y_t = Y_t - Y_{t-1} = \beta_1 + \varepsilon_t - \varepsilon_{t-1}$$

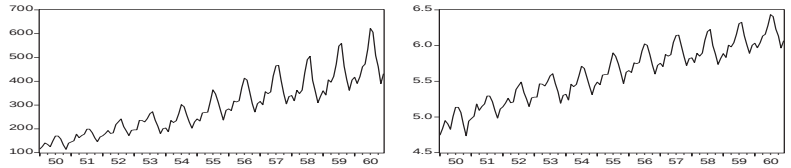
che è stazionario.



Processi non stazionari in varianza

Eventuali non stazionarietà in varianza possono essere trattate mediante una classe di trasformazioni detta trasformata di Box-Cox

$$T(Y_t) = \begin{cases} \frac{Y_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ lnY_t & \lambda = 0 \end{cases}$$



Ancora sul processo Random Walk

- Il processo RW $Y_t = Y_{t-1} + \varepsilon_t,\ Y_0 = c$ dove $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$ può essere visto come un particolare processo $AR(1)$ con $\phi_1 = 1$ e pertanto è un processo non stazionario.
- Operando per sostituzioni successive si può scrivere:

$$Y_t = Y_0 + \sum_{j=0}^{t-1} \varepsilon_{t-j} = c + \sum_{j=1}^{t} \varepsilon_j$$

- Valore atteso:

$$E[Y_t] = E\left[c + \sum_{j=1}^{t} \varepsilon_j\right] = c$$

Random Walk

- Varianza

$$Var(Y_t) = Var\left(c + \sum_{j=1}^{t} \varepsilon_j\right) = t\sigma_\varepsilon^2 \Rightarrow \infty$$

- Covarianza

$$\gamma(t, s) = \min(t, s)\sigma_\varepsilon^2$$

per cui

$$\gamma(t, t-k) = (t-k)\sigma_\varepsilon^2$$

$$\gamma(t, t+k) = t\sigma_\varepsilon^2$$

- Correlazione

$$\lim_{t\to\infty} \rho_k = 1$$

- Si dice anche che il processo ha memoria infinita

Random Walk con costante o *drift*

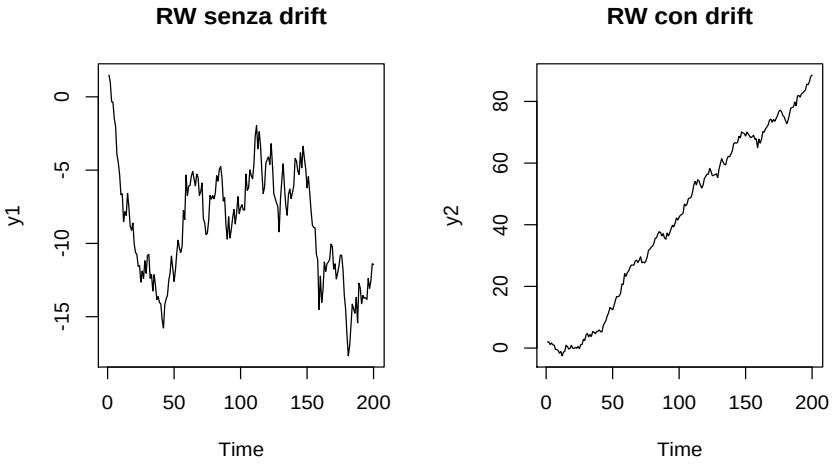
- $$Y_t = \mu + Y_{t-1} + \varepsilon_t, \quad Y_0 = c$$
- Mediante sostituzioni successive si ha:

$$Y_t = c + \mu t + \sum_{j=1}^{t} \varepsilon_j$$

dove compare il trend deterministico lineare $c + \mu t$ indotto dalla presenta della costante μ. Oltre al trend deterministico, è presente anche la componente $\sum_{j=1}^{t} \varepsilon_j$ che viene generalmente indicata come trend stocastico in quanto costituisce la parte non stazionaria e non deterministica del processo.

- Valore atteso

$$E[Y_t] = E\left[c + \mu t + \sum_{j=1}^{t} \varepsilon_j\right] = c + \mu t$$



Differenza prima di un Random Walk

- processo differenziato

$$Y_t = \mu + Y_{t-1} + \varepsilon_t, \quad Y_0 = c$$

quindi

$$(1 - B)Y_t = Y_t - Y_{t-1} = \mu + \varepsilon_t$$

- Il processo ottenuto dalla differenza prima è un WN più un termine costante, pertanto è un processo stazionario
- Il processo RW è un processo omogeneo di grado 1

Processo AutoRegressivo Integrato a Media Mobile *ARIMA*(p, d, q)

- Sia $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$, il processo Y_t si dice *ARIMA*(p, d, q) se:
  - i $(1 - B)^d Y_t = X_t$
  - ii $X_t = \phi_0 + \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$
- Alternativamente possiamo scrivere:

$$\phi(B)(1 - B)^d Y_t = \theta(B)\varepsilon_t$$

dove gli operatori $\phi(B)$ e $\theta(B)$ non hanno zeri in comune.

- Il parametro ϕ_0 gioca un ruolo diverso a seconda che $d = 0$ o $d \neq 0$:
  - i. se $d = 0$ Y_t è stazionario e ϕ_0 è legato alla media del processo
  - ii. se $d > 0$ ϕ_0 è il *drift* e, quindi, è legato al trend deterministico del processo

- Se $d > 0$ allora il processo è, evidentemente non stazionario, dal momento che l'equazione caratteristica:

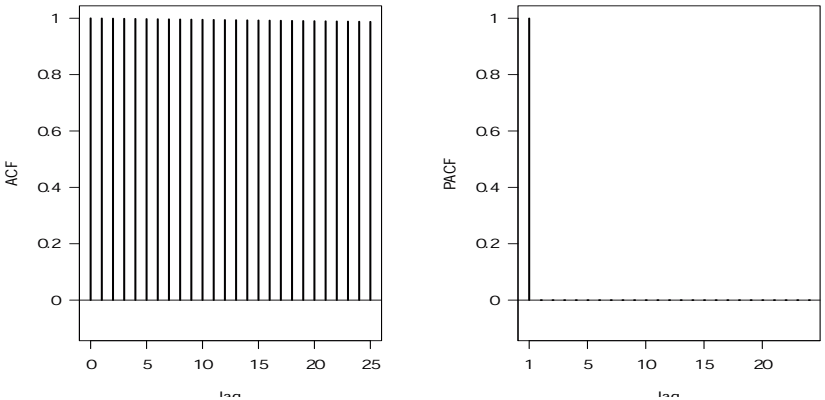
$$(1 - B)^d \phi(B) = 0$$

contiene d radici unitarie.
- In pratica il valore di d è spesso 1, più raramente 2.
- Il processo RW può essere visto come un $ARIMA(0, 1, 0)$

Casi particolari di un $ARIMA(p, d, q)$

- $p = d = q = 0 \quad \Rightarrow \quad WN(0, \sigma_\varepsilon^2)$
- $p = 1,\ d = q = 0 \quad \Rightarrow \quad AR(1)$
- $p = d = 0,\ q = 1 \quad \Rightarrow \quad MA(1)$
- $d = 0,\ p = q = 1 \quad \Rightarrow \quad ARMA(1, 1)$
- $d = 1,\ p = q = 0 \quad \Rightarrow \quad RW = ARIMA(0, 1, 0) = I(1)$

Funzioni di autocorrelazione teoriche per un processo Integrato, I(1)



Esempio: $ARIMA(1, 1, 1)$

- Consideriamo un processo $ARIMA(1, 1, 1)$ senza costante:

$$(1 - \phi B)(1 - B) Y_t = (1 - \theta B)\varepsilon_t$$

- Se poniamo $(1 - B) Y_t = X_t$ abbiamo:

$$X_t = \phi X_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}$$

che un processo $ARMA(1, 1)$ stazionario posto che $|\phi| < 1$ e $|\theta| < 1$

Realizzazione di un $ARIMA(1,1,1)$ con $\phi = 0.7$ e $\theta = 0.5$

