ANALISI DELLE SERIE TEMPORALI - SERIE STORICHE ECONOMICHE

Padova, 3 febbraio 2016

Nome Cognome N. matricola

Corso di Laurea

1 [2+3]	2 [1+1+3]	3 [2+2+3]	4 [5]	5 [4]	6 [4]	Totale

Domanda 1 Sia X_t un processo tale che $E[X_t] = 3t$ e $Cov(X_t, X_{t-k}) = \gamma_k$.

1. Si dica, giustificando la risposta, se X_t è stazionario.
2. Si consideri il processo $Y_t = 7 - 3t + X_t$. Si calcolino valore atteso e funzione di autocovarianza di Y_t e si dica, giustificando la risposta, se Y_t è stazionario.

Soluzione:

1. X_t non è un processo stazionario in quanto la sua media non è costante, ma varia con t.
2.
   - $E[Y_t] = 7 - 3t + E[X_t] = 7 - 3t + 3t = 7$ e non dipende da t.
   - Per quanto riguarda la funzione di autocovarianza si ha:

$$\begin{aligned} Cov(Y_t, Y_{t-k}) &= E[(Y_t - E[Y_t])(Y_{t-k} - E[Y_{t-k}])] = E[(Y_t - 7)(Y_{t-k} - 7)] = \\ &= E[(7 - 3t + X_t - 7)(7 - 3(t-k) + X_{t-k} - 7)] = \\ &= E[(X_t - 3t)(X_{t-k} - 3(t-k)] = Cov(X_t, X_{t-k}) = \gamma_k \end{aligned}$$

   che non dipende da t, ma solo da k, pertanto, il processo Y_t è stazionario.

Si poteva anche osservare che $Y_t = 7 + X_t - E[X_t]$ da cui discende immediatamente che $Cov(Y_t, Y_{t-k}) = Cov(X_t, X_{t-k}) = \gamma_k$, in virtù del fatto che $Cov(W, Z) = Cov(W + a, Z + a)$ con a costante qualsiasi.

Domanda 2 Si consideri il processo $Y_t = \epsilon_t - \epsilon_{t-12}$ con $\epsilon_t \sim WN(0, \sigma^2_\epsilon)$.

1. Si dica che tipo di modello SARIMA lo descrive.
2. Si dica, giustificando la risposta, se il processo è stazionario.
3. Se ne calcoli la funzione di autocorrelazione.

Soluzione:

1. SARIMA(0,0,0)(0,0,1)
2. Il processo è stazionario essendo un processo a media mobile (stagionale).
3. Per quanto riguarda la funzione di autocorrelazione, abbiamo:

$$\gamma_0 = Var(Y_t) = Var(\epsilon_t - \epsilon_{t-12}) = Var(\epsilon_t) + Var(\epsilon_{t-12}) = \sigma^2_\epsilon + \sigma^2_\epsilon = 2\sigma^2_\epsilon$$

$$\begin{aligned}
\gamma_k &= Cov(Y_t, Y_{t-k}) = Cov(\epsilon_t - \epsilon_{t-12}, \epsilon_{t-k} - \epsilon_{t-12-k}) = \\
&= E[(\epsilon_t - \epsilon_{t-k})(\epsilon_{t-k} - \epsilon_{t-12-k})] = \\
&= E[\epsilon_t \cdot \epsilon_{t-k} - \epsilon_t \cdot \epsilon_{t-12-k} - \epsilon_{t-k} \cdot \epsilon_{t-12} + \epsilon_{t-k} \cdot \epsilon_{t-12-k}] = \\
&= -\sigma^2_\epsilon \text{ solo per } k = 12
\end{aligned}$$

Pertanto la funzione di autocorrelazione, $\rho(k)$, vale 1 per $k = 0$ e $-1/2$ per $k = 12$.

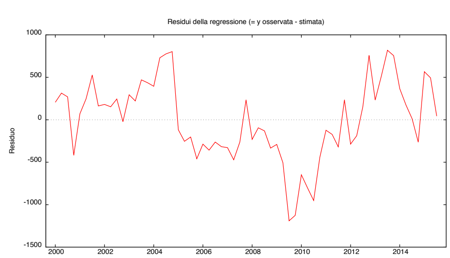
Domanda 3 Sia y_t, per $t = 1, 2, \ldots, 63$, la serie trimestrale relativa al Valore Aggiunto di agricoltura, silvicoltura e pesca rilevata fra il I trimestre del 2000 e il III trimestre del 2015. Poiché la serie mostra un chiaro andamento stagionale, per essa si ipotizza un modello additivo del tipo $y_t = \delta_1 + \delta_2 d_{2t} + \delta_3 d_{3t} + \delta_4 d_{4t} + \epsilon_t$, dove $\epsilon_t \sim WN(0, \sigma^2)$, dove $d_{it},\ i = 1, \ldots, 4$, sono le dummy stagionali. Il modello stimato con i minimi quadrati porta ai seguenti risultati:

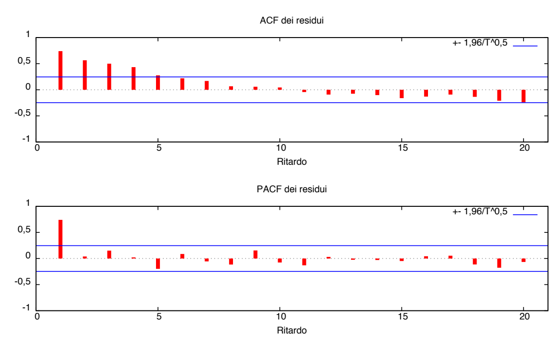
```
Modello 1: OLS, usando le osservazioni 2000:1-2015:3 (T = 63)
Variabile dipendente: y

             coefficiente   errore std.   rapporto t   p-value
  ------------------------------------------------------------
  const        2417,01        118,095       20,47      1,75e-28 ***
  dq2          3700,08        167,011       22,15      2,77e-30 ***
  dq3          9023,37        167,011       54,03      6,00e-52 ***
  dq4          8868,76        169,772       52,24      4,20e-51 ***

Media var. dipendente   7759,973   SQM var. dipendente     3844,355
Somma quadr. residui    13165334   E.S. della regressione  472,3783
R-quadro                0,985632   R-quadro corretto       0,984902
F(3, 59)                1349,125   P-value(F)              2,71e-54
Log-verosimiglianza    ?475,2670   Criterio di Akaike      958,5339
Criterio di Schwarz     967,1065   Hannan-Quinn            961,9055
rho                     0,740887   Durbin-Watson           0,515082
Note: SQM = scarto quadratico medio; E.S. = errore standard
```

Il grafico dei residui assieme alle funzioni di autocorrelazione (dei residui) sono i seguenti:





1. Si commenti l'adattamento del modello ai dati.
2. Si ipotizzi un modello SARIMA per i residui del modello.
3. Si calcolino i coefficienti di stagionalità ideali.

Soluzione (stesso esempio, con dati aggiornati, del libro pag. 57!):

1. I coefficienti stimati risultano tutti significativi a qualsiasi livello di confidenza. Il valore di R^2 è molto elevato: il modello spiega più del 98% della variabilità del fenomeno considerato. I residui del modello, tuttavia, pur non presentando alcun andamento stagionale, non sono incorrelati fra loro, come si può osservare dai grafici delle funzioni di autocorrelazione empiriche che presentano i primi coefficienti significativamente diversi da zero.

2. La funzione di autocorrelazione globale decresce verso zero esponenzialmente e si annulla dopo qualche ritardo, mentre la funzione di autocorrelazione parziale presenta solo il primo coefficiente significativamente diverso da zero. Questo ci fa ipotizzare che i residui possano essere modellati da un AR(1).

3. Il modello stima la costante δ_1 e i tre coefficienti associati alle variabili dummy del II, III e IV trimestre. Tra i parametri stimati e i coefficienti grezzi di stagionalità, $\gamma_i^*, \quad i = 1, \ldots, 4$, c'è una relazione ben precisa. In particolare si ha che $\delta_1 = \gamma_1^*$, $\delta_2 = \gamma_2^* - \gamma_1^*$, $\delta_3 = \gamma_3^* - \gamma_1^*$ e $\delta_4 = \gamma_4^* - \gamma_1^*$. Quindi la costante δ_1 corrisponde all'effetto del I trimestre, mentre δ_2, δ_3 e δ_4 rappresentano gli effetti differenziali rispetto al I trimestre. Per ottenere i coefficienti di stagionalià dobbiamo quindi calcolare prima i coefficienti grezzi $\hat{\gamma}_i^*$. Abbiamo:

$$\hat{\gamma}_1^* = 2417,01 \quad \hat{\gamma}_2^* = 3700,08 + 2417,01 = 6117,09$$

$$\hat{\gamma}_3^* = 9023,37 + 2417,01 = 11440,38 \quad \hat{\gamma}_4^* = 8868,76 + 2417,01 = 11285,77$$

Essendo $\bar{\gamma^*} = 7815,06$ i coefficienti ideali di stagionalità sono:

$$\hat{\gamma}_1 = 2417,01 - 7815,06 = -5398,05 \quad \hat{\gamma}_2 = 6117,09 - 7815,06 = -1697,97$$

$$\hat{\gamma}_3 = 11440,38 - 7815,06 = 3625,32 \quad \hat{\gamma}_4 = 11285,77 - 7815,06 = 3470,71.$$

Domanda 4 Dato un modello SARIMA(1,0,0)$\times$(0,1,1)$_4$, calcolare la previsione con orizzonte temporale $k = 5$ sapendo che:

- il parametro AR ϕ_1 è pari a -0,8;
- l'ultimo valore osservato della serie storica è $y_n = 3$;
- $\hat{y}_{n+1|n} = -2$ e $\hat{y}_{n+4|n} = 1$.

Soluzione: Esercizio 7.5 dell'eserciziario.

Il processo indicato si esprime nella forma:

$$(1 - \phi_1 B)(1 - B^4)Y_t = (1 - \Theta_1 B^4)\varepsilon_t.$$

Poiché

$$(1 - \phi_1 B)(1 - B^4) = 1 - \phi_1 B - B^4 + \phi_1 B^5,$$

il processo può essere riscritto nel modo seguente:

$$Y_t = \phi_1 Y_{t-1} + Y_{t-4} - \phi_1 Y_{t-5} + \varepsilon_t - \Theta_1 \varepsilon_{t-4}.$$

Il previsore di Y_{n+5} è dunque pari a:

$$\begin{aligned} \hat{Y}_{n+5|n} &= E_n[Y_{n+5}] \\ &= \phi_1 E_n[Y_{n+4}] + E_n[Y_{n+1}] - \phi_1 E_n[Y_n] + \\ &\quad + E_n[\varepsilon_{n+5}] - \Theta_1 E_n[\varepsilon_{n+1}]. \end{aligned}$$

Il calcolo dei valori attesi condizionati avviene secondo le seguenti espressioni:

$$\begin{aligned} E_n[Y_{n+j}] &= \begin{cases} y_{n+j} & j \leq 0 \\ \hat{y}_{n+j|n} & j > 0 \end{cases} \\ E_n[\varepsilon_{n+j}] &= \begin{cases} e_{n+j} & j \leq 0 \\ 0 & j > 0 \end{cases}. \end{aligned}$$

Pertanto, la previsione $\hat{y}_{n+5|n}$ è pari a

$$\hat{y}_{n+5|n} = \phi_1 \hat{y}_{n+4|n} + \hat{y}_{n+1|n} - \phi_1 y_n.$$

Sostituendo i valori noti nella precedente espressione otteniamo il risultato richiesto:

$$\begin{aligned} \hat{y}_{n+5|n} &= -0,8 - 2 - (-0,8) \cdot 3 \\ &= -0,8 - 2 + 2,4 = -0,4. \end{aligned}$$

Domanda 5 Si descriva, brevemente, la procedura di Box e Jenkins.

Domanda 6 Dopo aver definito cos'è una media mobile, M, dire quando una media mobile è:

1. semplice
2. simmetrica
3. centrata