SERIE STORICHE

Padova, 26 gennaio 2017

Nome ______ Cognome ______ N. matricola ______

Corso di Laurea ______

1.1 [2]	1.2 [3]	1.3 [2]	2 [6] (2+2+2)	3.1 [3]	3.2 [2]	4.1 [2]	4.2 [2]	5 [4]	6 [4]	Totale

Domanda 1 Alla serie delle vendite di un certo prodotto è stato adattato il seguente modello:

$$(1 - 1{,}4B + 0{,}4B^2)(1 - B)y_t = \varepsilon_t$$

dove $\hat{\sigma}^2_\varepsilon = 58.000$ e le ultime tre osservazioni sono $y_{198} = 640$, $y_{199} = 770$ e $y_{200} = 800$.

1. Che tipo di modello è stato adattato alla serie in questione?
2. Si calcolino le previsioni per $t = 201, 202$ e 203.
3. Si trovi l'intervallo di previsione al 95% per y_{201}.

Soluzione:

1. Il modello sembra essere un $ARIMA(2, 1, 0)$. In realtà, calcolando le radici dell'equazione $1 - 1{,}4B + 0{,}4B^2 = 0$ troviamo $B_1 = 5/2$ e $B_2 = 1$ e quindi il modello è un $ARIMA(1, 2, 0)$, ossia un modello non stazionario, e precisamente $(1 - 0{,}4B)(1 - B)^2 y_t = \varepsilon_t$.
2. Moltiplicando i due polinomi autoregressivi, possiamo riscrivere il modello come:

$$y_t - 2{,}4y_{t-1} + 1{,}8y_{t-2} - 0{,}4y_{t-3} = \varepsilon_t$$

da cui:

$$y_t = 2{,}4y_{t-1} - 1{,}8y_{t-2} + 0{,}4y_{t-3} + \varepsilon_t.$$

Le previsioni richieste, usando le regole del valore atteso condizionato, si ottengono nel modo seguente:

$$\begin{aligned}\hat{y}_{201} = \hat{y}_{200}(1) = E[Y_{201} \mid I_{200}] = E_{200}[Y_{201}] = \\ = 2{,}4E_{200}[Y_{200}] - 1{,}8E_{200}[Y_{199}] + 0{,}4E_{200}[Y_{198}] + E_{200}[\varepsilon_{201}] = \\ = 2{,}4 \cdot 800 - 1{,}8 \cdot 770 + 0{,}4 \cdot 640 = 790\end{aligned}$$

$$\begin{aligned}\hat{y}_{202} = \hat{y}_{200}(2) = E[Y_{202} \mid I_{200}] = E_{200}[Y_{202}] = \\ = 2{,}4E_{200}[Y_{201}] - 1{,}8E_{200}[Y_{200}] + 0{,}4E_{200}[Y_{199}] + E_{200}[\varepsilon_{202}] = \\ = 2{,}4 \cdot 790 - 1{,}8 \cdot 800 + 0{,}4 \cdot 770 = 764\end{aligned}$$

$$\begin{aligned}\hat{y}_{203} &= \hat{y}_{200}(3) = E[Y_{203} \mid I_{200}] = E_{200}[Y_{203}] = \\ &= 2{,}4E_{200}[Y_{202}] - 1{,}8E_{200}[Y_{201}] + 0{,}4E_{200}[Y_{200}] + E_{200}[\varepsilon_{203}] = \\ &= 2{,}4 \cdot 764 - 1{,}8 \cdot 790 - 0{,}7 \cdot 800 = 731{,}6\end{aligned}$$

3. L'intervallo di previsione al 95% per y_{201} è dato da:

$$IC(1) = \hat{y}_{201} \pm 1{,}96 \cdot \sqrt{Var(e(1))}$$

dove $e(1) = y_{201} - \hat{y}_{201} = \varepsilon_{201}$ è l'errore di previsione 1 passo in avanti.
Si ha $Var(e(1)) = Var(\varepsilon_{201}) = 58.000$ da cui:

$$IC(y_{201}) = (790 \pm 1,96 \cdot \sqrt{58.000}) = (317{,}9695, 1.262{,}0305)$$

Domanda 2 Si determini (giustificando la risposta) quale dei seguenti processi è stazionario e/o invertibile e si determini a che classe di processi ARIMA appartengono. Nel seguito si intende che $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$.

1. $X_t = 1{,}5X_{t-1} - 0{,}5X_{t-2} + \varepsilon_t$
2. $X_t = -1{,}7 + \varepsilon_t - 0{,}6\varepsilon_{t-1} - 0{,}3\varepsilon_{t-2}$
3. $X_t = 0{,}6X_{t-1} + \varepsilon_t - 1{,}2\varepsilon_{t-1} + 0{,}2\varepsilon_{t-2}$

Soluzione:

1. Il processo è un $AR(2)$, che è sempre invertibile. Vediamo ora se è stazionario. Osserviamo che possiamo riscrivere il processo in termini dell'operatore ritardo B come $(1-1{,}5B+0{,}5B^2)X_t = \varepsilon_t$. Le soluzioni dell'equazione caratteristica $1 - 1{,}5B + 0{,}5B^2 = 0$ sono $B_1 = 1$ e $B_2 = 2$. Poiché abbiamo una radice unitaria, il processo NON è stazionario ed è un $ARIMA(1,1,0)$ e, precisamente, $(1 - 0{,}5B)(1 - B)X_t = \varepsilon_t$.
2. Il processo è un MA(2) con costante (media) diversa da zero e, quindi, è sempre stazionario. Per vedere se è invertibile calcolo le radici dell'equazione caratteristica $1 - 0{,}6B - 0{,}3B^2 = 0$. Tali radici risultano pari a $B_1 = -3{,}083$ e $B_2 = 1{,}083$, quindi in modulo maggiori di 1, di conseguenza X_t è invertibile.
3. Il processo è un ARMA(1,2), stazionario, essendo il parametro autoregressivo minore di 1. Verifichiamo se è invertibile calcolando le radici dell'equazione caratteristica $1-1{,}2B+0{,}2B^2 = 0$. Tali radici risultano pari a $B_1 = 1$ e $B_2 = 5$ pertanto, essendo una delle due radici esattamente pari ad uno, il processo non è invertibile.

Domanda 3

Si dica, **giustificando la risposta**, se i seguenti processi sono o meno (debolmente) stazionari. (Si ricordi che se $X \sim \mathcal{N}(0, \sigma^2)$ allora $E[X^3] = 0$ e $E[X^4] = 3\sigma^4$)

1. $Y_t = \varepsilon_t - \theta\varepsilon_{t-1}^2$, dove $\varepsilon_t \sim WN\,\mathcal{N}(0, \sigma_\varepsilon^2)$.
2. $Y_t = (-1)^t Z$, dove Z è una variabile casuale a media nulla e varianza σ_Z^2.

Soluzione: Per verificare la stazionarietà di un processo bisogna valutarne i momenti primo e secondi. Abbiamo:

1. $$E[Y_t] = E[\varepsilon_t - \theta\varepsilon_{t-1}^2] = E[\varepsilon_t] - \theta E[\varepsilon_{t-1}^2] = -\theta\sigma_\varepsilon^2, \;\; \forall t$$

   Sfruttando il fatto che ε_t e ε_s per $t \neq s$ sono fra loro indipendenti, si ha:

   $$\begin{aligned} Var(Y_t) &= Var(\varepsilon_t - \theta\varepsilon_{t-1}^2) = Var(\varepsilon_t) + \theta^2 Var(\varepsilon_{t-1}^2) - 2Cov(\varepsilon_t, \varepsilon_{t-1}^2) = \\ &= \sigma_\varepsilon^2 + \theta^2\left(E[\varepsilon_{t-1}^4] - (E[\varepsilon_{t-1}^2])^2\right) = \sigma_\varepsilon^2 + \theta^2\left(3\sigma_\varepsilon^4 - \sigma_\varepsilon^4\right) = \sigma_\varepsilon^2 + 2\theta^2\sigma_\varepsilon^4, \;\; \forall t \end{aligned}$$

   Infine:

   $$\begin{aligned} Cov(Y_t, Y_{t-1}) &= E[(Y_t - E[Y_t])(Y_{t-1} - E[Y_{t-1}])] = \\ &= E[(\varepsilon_t - \theta\varepsilon_{t-1}^2 + \theta\sigma_\varepsilon^2)(\varepsilon_{t-1} - \theta\varepsilon_{t-2}^2 + \theta\sigma_\varepsilon^2) = \\ &= E[\varepsilon_t\varepsilon_{t-1} - \theta\varepsilon_t\varepsilon_{t-2}^2 + \theta\sigma_\varepsilon^2\varepsilon_t - \theta\varepsilon_{t-1}^2\varepsilon_{t-1} + \theta^2\varepsilon_{t-1}^2\varepsilon_{t-2}^2 \\ &\quad - \theta^2\sigma_\varepsilon^2\varepsilon_{t-1}^2 + \theta\sigma_\varepsilon^2\varepsilon_{t-1} - \theta^2\sigma_\varepsilon^2\varepsilon_{t-2}^2 + \theta^2\sigma_\varepsilon^4] = \\ &= 0 - 0 + 0 - 0 + 0 - \theta^2\sigma_\varepsilon^4 + 0 - \theta^2\sigma_\varepsilon^4 + \theta^2\sigma_\varepsilon^4 = -\theta^2\sigma_\varepsilon^4 \end{aligned}$$

   Allo stesso modo si può mostrare che $Cov(Y_t, Y_{t+1}) = -\theta^2\sigma_\varepsilon^4$ e, più in generale, $Cov(Y_t, Y_{t+k}) = -\theta^2\sigma_\varepsilon^4$ per $\mid k \mid > 1$.
   Pertanto il processo considerato è (debolmente) stazionario.

2. $$E[Y_t] = E[(-1)^t Z] = (-1)^t E[Z] = 0, \;\; \forall t$$

   $$Var(Y_t) = Var[(-1)^t Z] = (-1)^{2t} Var(Z) = Var(Z) = \sigma_Z^2, \;\; \forall t$$

   Infine:

   $$Cov(Y_t, Y_{t\pm k}) = Cov((-1)^t Z, (-1)^{t\pm k} Z) = (-1)^{2t\pm k} Var(Z) = (-1)^{\pm k}\sigma_Z^2$$

   Poiché la covarianza non dipende da t ma solo dal ritardo k, il processo considerato è (debolmente) stazionario.

Domanda 4

Si consideri la seguente serie storica:

$$x_t = (7, 11, 13, 11, 14, 13, 9, 9, 8, 8, 12, 13).$$

1. Dopo aver derivato i coefficienti della media mobile, M_3, che si ottiene componendo

$$M_1 = \left\{[3]; \frac{1}{3}[1, \mathbf{1}]\right\} \quad e \quad M_2 = \left\{[5]; \frac{1}{9}[1, 2, \mathbf{3}]\right\}$$

   applicare M_3 alla serie data.

2. Dire, **giustificando la risposta**, quale delle tre medie mobili liscia meglio la serie data.

Soluzione:

1. L'ordine di $M_3 = M_1M_2$ è 3+5-1=7. Possiamo scrivere

$$y^{**} = M_3y = M_2M_1y = M_2y^*$$

$$\begin{aligned} y_2^* &= \frac{1}{3}(y_1 + y_2 + y_3) \\ y_3^* &= \frac{1}{3}(y_2 + y_3 + y_4) \\ y_4^* &= \frac{1}{3}(y_3 + y_4 + y_5) \\ y_5^* &= \frac{1}{3}(y_4 + y_5 + y_6) \\ y_6^* &= \frac{1}{3}(y_5 + y_6 + y_7) \end{aligned}$$

   Applicando M_2 ai 5 termini calcolati otteniamo:

$$y_4^{**} = \frac{1}{9}(y_2^* + 2y_3^* + 3y_4^* + 2y_5^* + y_6^*).$$

   Ridisponendo i termini si ottiene

$$\begin{array}{rccccccccc} y_4^{**} = & \frac{1}{9}\frac{1}{3} & (y_1+ & y_2+ & y_3+ & & & & \\ & & & 2y_2+ & 2y_3+ & 2y_4+ & & & \\ & & & & 3y_3+ & 3y_4+ & 3y_5+ & & \\ & & & & & 2y_4+ & 2y_5+ & 2y_6+ & \\ & & & & & & y_5+ & y_6+ & y_7) \\ y_4^{**} = & \frac{1}{27} & (y_1+ & 3y_2+ & 6y_3+ & 7y_4+ & 6y_5+ & 3y_6+ & y_7) & . \end{array}$$

   Generalizzando, si ha

$$y_t^{**} = \frac{1}{27}(y_{t-3} + 3y_{t-2} + 6y_{t-1} + 7y_t + 6y_{t+1} + 3y_{t+2} + y_{t+3})$$

   con $M_3 = \left\{[7]; \frac{1}{27}[1, 3, 6, \mathbf{7}]\right\}$.

2. Per dire quale fra le tre medie mobili liscia meglio la serie dobbiamo calcolare il rapporto di riduzione della varianza residua. Per M_1 si ha:

$$\sum_{i=-1}^{1} \theta_i^2 = 3 \cdot \frac{1}{9} = \frac{1}{3} = 0.\bar{3}$$

Per M_2 si ha:

$$\sum_{i=-2}^{2} \theta_i^2 = \frac{1}{81}(1+4+9+4+1) = \frac{19}{81} = 0.235$$

Infine, per M_3 :

$$\sum_{i=-3}^{3} \theta_i^2 = \frac{1}{729}(1+9+36+49+36+9+1) = \frac{141}{729} = 0.193$$

Pertanto, la media mobile da preferire è M_3.

Domanda 5 Dare la definizione di stazionarietà in senso debole e fornire un esempio di processo stazionario (in senso debole).

Domanda 6 Descrivere brevemente come stimare un trend polinomiale secondo l'approccio classico all'analisi delle serie storiche.