SERIE STORICHE

Padova, 18 febbraio 2026

Nome Cognome N. matricola

Domanda 1

Si consideri il processo X_t così definito:

$$X_t = \lambda^t + Z_t \qquad \text{con} \qquad Z_t \sim WN(0, \sigma^2).$$

1. Verificare per quali valori di λ il processo X_t è stazionario in senso debole.
2. Dire per quali valori di λ diversi da 0 e da 1 il processo Y_t così definito

$$Y_t = \frac{2}{9}X_{t-1} + \frac{5}{9}X_t + \frac{2}{9}X_{t+1}$$

è stazionario *per la media*.

Soluzione:

1. Perché il processo sia stazionario in senso debole i primi due momenti non devono dipendere dal tempo (la covarianza solo da k). La media risulta:

   $$E[X_t] = \lambda^t + E[Z_t] = \lambda^t$$

   Affinché la media non dipenda dal tempo deve essere $\lambda = 1$ o $\lambda = 0$.

   Per verificare la stazionarietà serve calcolare anche varianza e autocovarianza:

   $$Var[X_t] = E[X_t - E(X_t)]^2 = E[X_t - \lambda^t]^2 = E[Z_t^2] = \sigma^2$$
   $$Cov[X_t, X_{t-k}] = E[(X_t - \lambda^t)(X_{t-k} - \lambda^{t-k})] = E[Z_t Z_{t-k}] = 0 \quad k > 0$$

   I momenti secondi non dipendono dal tempo quindi, affinché il processo sia stazionario in senso debole, deve essere $\lambda = 1$ o $\lambda = 0$.

2. Il processo Y_t è stazionario in media se il suo valore atteso non dipende dal tempo.

   $$\begin{aligned} E[Y_t] &= E[\frac{2}{9}X_{t-1} + \frac{5}{9}X_t + \frac{2}{9}X_{t+1}] \\ &= \frac{2}{9}\lambda^{t-1} + \frac{5}{9}\lambda^t + \frac{2}{9}\lambda^{t+1} \\ &= \frac{1}{9}\lambda^{t-1}(2\lambda^2 + 5\lambda + 2) \end{aligned}$$

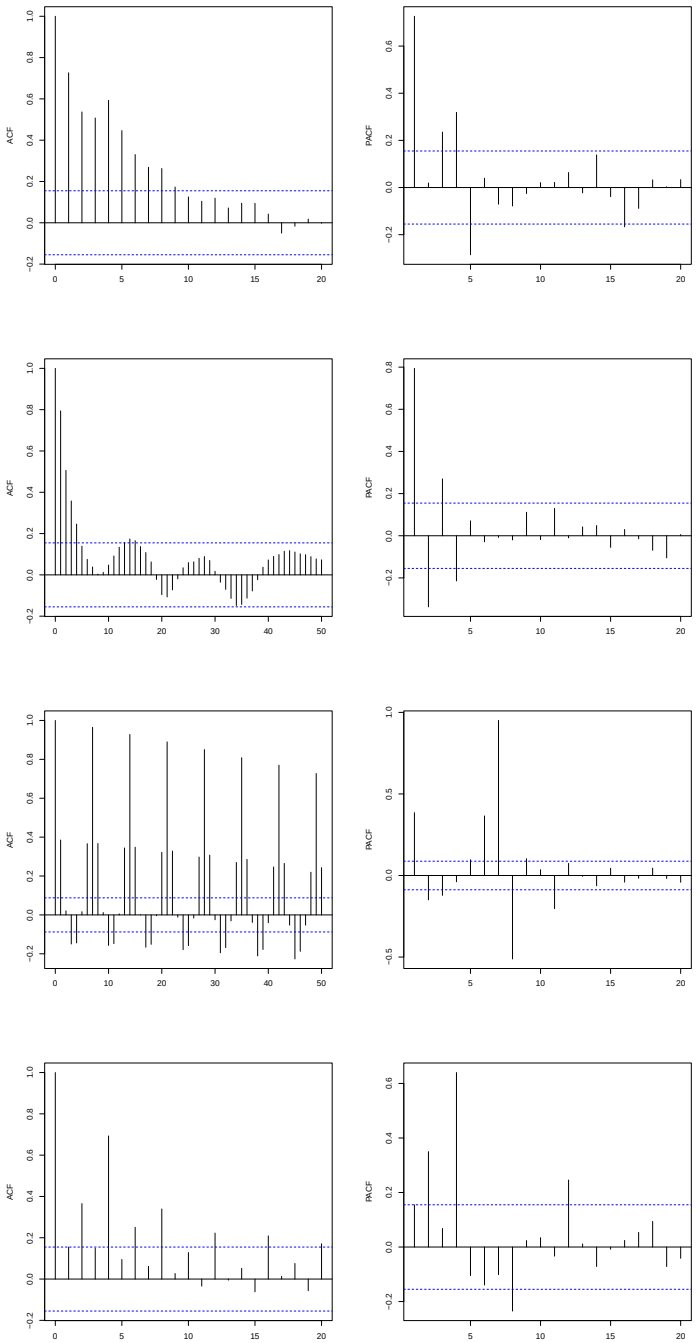
   Oltre che per $\lambda = 1$ o $\lambda = 0$, $E[Y_t]$ non dipende dal tempo anche per tutti i valori di λ che soddisfano l'equazione:

   $$(2\lambda^2 + 5\lambda + 2) = 0$$

   Tali valori sono $\lambda_1 = -2$ e $\lambda_2 = -1/2$.

Domanda 2

Identificare dei possibili processi generatori compatibili con le ACF e PACF riportate di seguito:



Soluzione:
SARIMA(1,0,0)(1,0,0)_4
ARIMA(1,0,1)
SARIMA(1,0,0)(0,1,0)_7
SARIMA(0,0,0)(1,0,1)_4

Domanda 3

Si consideri il processo

$$Y_t = 0.5\varepsilon_{t-1} + 0.3\varepsilon_{t-2} + \varepsilon_t.$$

1. *Derivare* la funzione di autocorrelazione del processo.
2. Fornire il valore dei coefficienti di autocorrelazione parziale ai ritardi 1 e 2.

Soluzione:

$$E[Y_t] = 0$$

$$V[Y_t] = \sigma_\varepsilon^2(1 + 0.5^2 + 0.3^2) = 1.34\ \sigma_\varepsilon^2$$

Funzione di autocovarianza

$$\begin{aligned} Cov[Y_t, Y_{t-1}] &= E[(0.5\ \varepsilon_{t-1} + 0.3\ \varepsilon_{t-2} + \varepsilon_t)(0.5\ \varepsilon_{t-2} + 0.3\ \varepsilon_{t-3} + \varepsilon_{t-1})] \\ &= E[(0.5\varepsilon_{t-1}^2 + 0.5 \cdot 0.3\varepsilon_{t-2}^2) \\ &= \sigma_\varepsilon^2(0.5 + 0.5 \cdot 0.3) \\ &= \sigma_\varepsilon^2 \cdot 0.65 \end{aligned}$$

$$\begin{aligned} Cov[Y_t, Y_{t-2}] &= E[(0.5\ \varepsilon_{t-1} + 0.3\ \varepsilon_{t-2} + \varepsilon_t)(0.5\ \varepsilon_{t-3} + 0.3\ \varepsilon_{t-4} + \varepsilon_{t-2})] \\ &= 0.3E[\varepsilon_{t-2}^2] = \sigma_\varepsilon^2 \cdot 0.3 \end{aligned}$$

$$Cov[Y_t, Y_{t-k}] = 0 \quad k > 2$$

Funzione di autocorrelazione

$$\rho_1 = \frac{Cov[Y_t, Y_{t-1}]}{V[Y_t]} = \frac{0.65\ \sigma_\varepsilon^2}{1.34\ \sigma_\varepsilon^2} = 0.485$$

$$\rho_2 = \frac{Cov[Y_t, Y_{t-2}]}{V[Y_t]} \frac{0.3\ \sigma_\varepsilon^2}{1.34\ \sigma_\varepsilon^2} = 0.224$$

$$\rho_k = 0 \quad k > 2$$

Autocorrelazioni parziali per $k = 1, 2$

$$P_1 = \rho_1 = 0.485$$

$$P_2 = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{0.224 - 0.485^2}{1 - 0.485^2} = -0.015$$

Domanda 4

Per modellare la serie storica trimestrale y_t di una variabile economica per il periodo 2020:1 - 2024:4 è stato utilizzato un modello additivo del tipo $y_t = T_t + S_t + \varepsilon_t$, dove T_t rappresenta un trend lineare, S_t la componente stagionale e ε_t la componente erratica.
I valori dei coefficienti specifici, ottenuti utilizzando un'opportuna media mobile, sono i seguenti:

Anno	I trim	II trim	III trim	IV trim
2020			1.6	5.1
2021	-1.1	-5.1	1.4	5.3
2022	-2.0	-5.5	1.8	6.0
2023	-2.3	-4.4	0.8	4.8
2024	-1.0	-4.2		

Gli ultimi 4 valori osservati sono stati: $y_{2024,1} = 21.5$ $y_{2024,2} = 18.8$ $y_{2024,3} = 23.5$ $y_{2024,4} = 29.7$.
Sulla base di questi dati, determinare:

1. i valori assunti dalla componente stagionale;
2. i valori destagionalizzati degli ultimi 4 trimestri (2024.1-2024.4).

Soluzione:

1. La tabella mostra i coefficienti specifici di stagionalità ottenuti come:

$$CS_t = y_t - mm_t,$$

dove mm_t indica la media mobile a 5 termini, centratura di due medie mobili semplici a 4 termini.

Per calcolare i coefficienti grezzi di stagionalità, si devono calcolare le medie stagionali dei coefficienti specifici:

$$\hat{CG}_j = \frac{1}{N}\sum_{T=1}^{N} CS_{T,j} \qquad j = 1, \cdots, 4 \tag{1}$$

dove N è il numero di anni e $CS_{T,j}$ è il coefficiente specifico corrispondente al periodo j dell'anno T. In questo caso abbiamo

$$\begin{aligned} \hat{CG}_1 &= -1.6 \\ \hat{CG}_2 &= -4.8 \\ \hat{CG}_3 &= 1.4 \\ \hat{CG}_4 &= 5.3 \end{aligned}$$

Calcoliamo la media dei coefficienti grezzi:

$$\frac{1}{4}\sum_{j=1}^{4} \hat{CG}_j = \bar{CG} = 0.075 \neq 0$$

I coefficienti *ideali* si ottengono sottraendo a quelli grezzi la loro media $\bar{CG}$. I coefficienti ideali (valori assunti dalla stagionalità) sono dunque pari a:

$$
\begin{aligned}
\hat{CI}_1 &= -1.6 - 0.075 = -1.675 \\
\hat{CI}_2 &= -4.8 - 0.075 = -4.875 \\
\hat{CI}_3 &= 1.4 - 0.075 = 1.325 \\
\hat{CI}_4 &= 5.3 - 0.075 = 5.225
\end{aligned}
$$

2. I valori destagionalizzati mediante i coefficienti ideali di stagionalità si ottengono sottraendo ai valori $y_{T,j}$ i coefficienti ideali $\hat{CI}_j$:

$$
\begin{aligned}
\hat{y}^d_{2024,1} &= y_{2024,1} - \hat{CI}_1 = 21.5 + 1.675 = 23.175 \\
\hat{y}^d_{2024,2} &= y_{2024,2} - \hat{CI}_2 = 18.8 + 4.875 = 23.675 \\
\hat{y}^d_{2024,3} &= y_{2024,3} - \hat{CI}_3 = 23.5 - 1.325 = 22.175 \\
\hat{y}^d_{2024,4} &= y_{2024,4} - \hat{CI}_4 = 29.7 - 5.225 = 24.475
\end{aligned}
$$

Domanda 5

In relazione ad un generico processo stocastico, si definiscano, spiegandone anche il senso, i concetti di:
1. stazionarietà forte;
2. stazionarietà debole;
3. invertibilità.

Soluzione:

1. Un processo stocastico viene detto stazionario in senso stretto se:

$$F(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}) = F(Y_{t_1+\tau}, Y_{t_2+\tau}, \ldots, Y_{t_n+\tau}) \qquad \forall (t_1, t_2, \ldots, t_n),\ \forall \tau \text{ e } \forall n.$$

   Ciò implica che tutte le relazioni probabilistiche tra le componenti del processo rimangono invariate rispetto a traslazioni temporali.

2. Un processo stocastico viene detto stazionario in senso debole se:

$$E(Y_t) = \mu \quad \forall t$$
$$V(Y_t) = \sigma^2 \quad \forall t$$
$$Cov(Y_t, Y_{t-k}) = \gamma(k) \quad \forall t, \quad k = 0, \pm 1, \pm 2, \ldots$$

   Ciò implica l'esistenza (finita) dei momenti primo e secondi, cosa che non è necessariamente vera per un processo stazionario in senso stretto (ma lo è nel caso Gaussiano).

3. Un processo Y_t è invertibile se può essere scritto come una funzione del suo passato e di un WN, ovvero se si può scrivere

$$Y_t = h(Y_{t-1}, Y_{t-2}, ...) + \varepsilon_t. \tag{2}$$

   La sua rilevanza risiede nel fatto che solo per modelli stazionari e invertibili è possibile identificare univocamente il processo generatore dei dati a partire dalla funzione di autocorrelazione.

Domanda 6

Spiegare in cosa consiste la cosiddetta "trappola delle dummy" in un modello in cui è presente una componente stagionale deterministica e quali sono le strategie per evitarla.

Soluzione

Consideriamo un modello per serie storiche in cui sia presente solo una componente stagionale (ad es. di periodo 4), oltre alla componente erratica

$$y_t = \gamma_0 + \gamma_1 d_{1t} + \gamma_2 d_{2t} + \gamma_3 d_{3t} + \gamma_4 d_{4t} + \varepsilon_t, \quad t = 1, 2, \ldots, n.$$

Così scritto, cioè includendo sia la costante sia le 4 variabili dummy, il modello assume la seguente forma

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ \vdots \\ y_{12} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \vdots \\ \varepsilon_{12} \end{bmatrix} \tag{3}$$

E' chiaro che, in questo caso, la matrice dei regressori non ha rango pieno dato che la prima colonna è una combinazione lineare delle altre 4 colonne. Questo implica che, nel calcolo delle stime LS (minimi quadrati) la matrice $(X'X)^{-1}$ non è più invertibile. Questa circostanza è quella che viene chiamata trappola delle dummy.

Ci sono due modi per evitare la trappola delle dummy:
1) evitare di includere la costante;
2) inserire costante ed eliminare una dummy (ad es. la prima). In questo caso il modello diventa:

$$y_t = \gamma_0 + \gamma_2^* d_{2t} + \gamma_3^* d_{3t} + \gamma_4^* d_{4t} + \varepsilon_t, \quad t = 1, 2, \ldots, n.$$

dove

$$\begin{aligned} \gamma_0 &= \gamma_1 \\ \gamma_2^* &= \gamma_2 - \gamma_0 \\ \gamma_3^* &= \gamma_3 - \gamma_0 \\ \gamma_4^* &= \gamma_4 - \gamma_0 \end{aligned}$$

rappresentano degli effetti differenziali rispetto al livello tipico del primo periodo (in questo caso).