Esercizio 1

La serie storica mensile y_t $t = 1, \ldots, 48$, si riferice al periodo 2018.1 - 2021.12. Per descriverla è stato identificato un modello dato dalla somma di un trend quadratico del tempo, di una componente stagionale, modellata con variabili *dummy*, e di una componente erratica. La tabella qui sotto riporta i valori della serie per l'anno 2021 e le stime dei coefficienti ideali di stagionalità:

Gen	Feb	Mar	Apr	Mag	Giu	Lug	Ago	Set	Ott	Nov	Dic
120.1	121.0	133.4	116.5	128,5	123.6	125.9	60.2	121.7	132.6	123.1	105.0
-0.431	2.273	11.487	0.804	7.474	6.587	6.252	-51.630	8.594	??	6.063	-7.998

Si ricavino il coefficiente ideale di stagionalità per ottobre 2021 ed i valori della serie destagionalizzata per i mesi di marzo, giugno, ottobre e dicembre 2021.

Soluzione

Nel caso di dati mensili, i coefficienti ideali di stagionalità godono, per definizione, della proprietà

$$\sum_{j=1}^{12} \hat{\gamma}_j = 0.$$

Quindi, poiché

$$\sum_{\substack{j=1 \\ j \neq 10}}^{12} \hat{\gamma}_j = -10.525$$

allora $\hat{\gamma}_{10} = 10.525$.

I valori destagionalizzati con il metodo della regressione si ottengono sottraendo ai valori originali i coefficienti ideali di stagionalità. Per l'anno 2021 questi sono calcolabili mediante l'espressione

$$\hat{y}^d_{2021,j} = y_{2021,j} - \hat{\gamma}_j, \quad j = 1, \ldots, 12.$$

Di seguito si riportano i valori destagionalizzati per i mesi richiesti:

$$\begin{array}{lll} \text{Marzo} & : & \hat{y}^d_{t_0,3} = 133.40 - 11.487 = 121,913 \\ \text{Giugno} & : & \hat{y}^d_{t_0,6} = 123.60 - 6,587 = 117,013 \\ \text{Ottobre} & : & \hat{y}^d_{t_0,10} = 132.60 - 10,525 = 122,075 \\ \text{Dicembre:} & & \hat{y}^d_{t_0,12} = 105.00 + 7.998 = 112,998 \end{array} .$$

Esercizio 2

Per la serie storica trimestrale $\{y_t\}_{t=1}^{21}$, che si riferisce ad una variabile economica osservata nel periodo 1997.2-2002.2, è stato stimato il seguente modello:

$$y_t = 5d_{1t} - 3,5d_{2t} + 2d_{3t} + 6d_{4t} + 0,8t + 0,02t^2 + \varepsilon_t, \;\; \varepsilon_t \sim WN(0,4)$$

dove d_{1t}, d_{2t}, d_{3t} e d_{4t} sono variabili *dummy* relative rispettivamente al primo, al secondo, al terzo ed al quarto trimestre.

1. Calcolare le previsioni per il periodo 2002.3-2003.2.
2. Sono stati osservati i valori $y_1 = 5,5$, $y_2 = 2,8$, $y_3 = 13$, $y_4 = 21,5$. Fornire i corrispondenti valori destagionalizzati per queste quattro osservazioni.
3. Riscrivere il modello inserendo la costante tra i regressori.

Soluzione

1. Le previsioni per il periodo 2002.3-2003.2 si ottengono applicando i coefficienti del modello ai valori assunti dalle variabili nei tempi corrispondenti ai quattro trimestri, ovvero per $t = 22, 23, 24, 25$:

$$\begin{array}{ll} t = 22: & \hat{y}_{22} = 2 + 0,8(22) + 0,02(22^2) = 29,28 \\ t = 23: & \hat{y}_{23} = 6 + 0,8(23) + 0,02(23^2) = 34,98 \\ t = 24: & \hat{y}_{24} = 5 + 0,8(24) + 0,02(24^2) = 35,72 \\ t = 25: & \hat{y}_{25} = -3,5 + 0,8(25) + 0,02(25^2) = 29,00. \end{array}$$

2. La serie destagionalizzata si calcola sulla base dei coefficienti ideali di stagionalità. I coefficienti riportati, γ_j^*, $j = 1, \ldots, 4$, sono in realtà grezzi perché

$$\sum_{j=1}^{4} \gamma_j^* = 9.5 \neq 0.$$

La loro media risulta essere:

$$\bar{\gamma} = \frac{9,5}{4} = 2,375$$

I coefficienti ideali sono quindi dati da:

$$\begin{array}{rcl} \hat{\gamma}_1 & = & \gamma_1^* - \bar{\gamma} = 5 - 2,375 = 2,625 \\ \hat{\gamma}_2 & = & \gamma_2^* - \bar{\gamma} = -3,5 - 2,375 = -5,875 \\ \hat{\gamma}_3 & = & \gamma_3^* - \bar{\gamma} = 2 - 2,375 = -0,375 \\ \hat{\gamma}_4 & = & \gamma_4^* - \bar{\gamma} = 6 - 2,375 = 3,625. \end{array}$$

Con questi coefficienti è possibile derivare i valori destagionalizzati richiesti (NB: la prima osservazione si riferisce al secondo trimestre):

$$\begin{array}{ll} t = 1: & \hat{y}_1^d = 5,5 - (-5,875) = 11,375 \\ t = 2: & \hat{y}_2^d = 2,8 - (-0,375) = 3,175 \\ t = 3: & \hat{y}_3^d = 13 - 3,625 = 9,375 \\ t = 4: & \hat{y}_4^d = 21,5 - 2,625 = 18,875. \end{array}$$

3. Il modello di regressione con variabili *dummy* stagionali può essere riscritto inserendo un termine costante ed eliminando una di queste a scelta.
   Scegliendo di eliminare d_{1t}. In questo caso si ha:

$$\begin{aligned}
\tilde{\alpha} &= \gamma_1^* = 5 \\
\tilde{\gamma}_2 &= \gamma_2^* - \gamma_1^* = -3,5 - 5 = -8,5 \\
\tilde{\gamma}_3 &= \gamma_3^* - \gamma_1^* = 2 - 5 = -3 \\
\tilde{\gamma}_4 &= \gamma_4^* - \gamma_1^* = 6 - 5 = 1.
\end{aligned}$$

   e il modello assume la seguente forma:

$$y_t = 5 - 8,5d_{2t} - 3d_{3t} + d_{4t} + 0,8t + 0,02t^2 + \varepsilon_t.$$

   In questo caso la costante $\tilde{\alpha} = 5$ ingloba anche l'effetto stagionale (grezzo) relativo al primo trimestre, mentre i coefficienti $\tilde{\gamma}_2, \tilde{\gamma}_3$ e $\tilde{\gamma}_4$, relativi a d_{2t}, d_{3t} e d_{4t}, rispettivamente, sono gli effetti stagionali differenziali rispetto al primo trimestre.

Esercizio 3

Si consideri la serie storica trimestrale generata dal seguente modello:

$$y_t = 10t + 0,5t^2 + 5d_{1t} - 10d_{2t} - 7d_{3t} + 12d_{4t} + \varepsilon_t$$

con $\varepsilon_t \sim N(0, 12.5)$, $t = 1, \ldots 12$ e d_{1t}, d_{2t}, d_{3t} e d_{4t} variabili *dummy* introdotte per rappresentare, rispettivamente, la prima, la seconda, la terza e la quarta stagione dell'anno.

1. Sulla base delle informazioni fornite è possibile calcolare i valori della serie destagionalizzata?
2. Si consideri il valore della serie ai tempi 1 e 2, corrispondenti al primo e al secondo trimestre di un anno. Quanto vale mediamente la serie in tale periodi? E quanto ci si aspetta sia variabile il valore della serie in tale periodo attorno alla loro media?
3. È possibile eliminare la componente di trend applicando alla serie l'operatore differenza prima?

Soluzione

1. No, per poterli calcolare è necessario disporre dei valori osservati (che non abbiamo!)
2. Per $t = 1$ e $d_{11} = 1$ abbiamo:

   $$y_1 = 5 + 10 \cdot 1 + 0.5 \cdot 1^2 + \varepsilon_1$$

   per cui

   $$E(y_1) = 5 + 10 \cdot 1 + 0.5 \cdot 1^2 + E(\varepsilon_1) = 15.5$$

   e

   $$Var(y_1) = Var(15.5 + \varepsilon_1) = Var(\varepsilon_1) = 12.5$$

   Per $t = 2$ e $d_{12} = 1$ abbiamo:

   $$y_2 = -10 + 10 \cdot 2 + 0.5 \cdot 2^2 + \varepsilon_2$$

   per cui

   $$E(y_2) = -10 + 10 \cdot 2 + 0.5 \cdot 2^2 + E(\varepsilon_2) = 12.0$$

   e

   $$Var(y_2) = Var(12.0 + \varepsilon_2) = Var(\varepsilon_2) = 12.5$$

3. No, perché il trend è di tipo quadratico e quindi per poterlo rimuovere è necessaria l'applicazione di una differenza di ordine 2.

Esercizio 4

Alla serie storica trimestrale y_t, per $t = 1, \ldots, n$, è stato adattato il seguente modello additivo:

$$y_t = 14 - 0.5t - 11d_{2t} - 19d_{3t} - 24d_{4t} + \varepsilon_t,$$

dove $Var\{\varepsilon_t\} = 4$ e $d_{2,t}$, d_{3t}, d_{4t} sono le variabili binarie indicatrici rispettivamente del secondo, del terzo e del quarto trimestre.

1. A quanto ammonta la variazione della serie nell'arco di un anno dovuta alla componente tendenziale? E quella dovuta alla componente stagionale?
2. Supponiamo che i primi quattro valori della serie, $t = 1, 2, 3, 4$, siano relativi ai quattro trimestri di un certo anno e supponiamo che in tali periodi i valori osservati della serie siano 16, 4, -2, -9. Si calcolino i corrispondenti valori destagionalizzati.
3. Si applichi alla serie l'operatore $(1 - B^4)$ ottenendo la serie

   $$y_t^* = (1 - B^4)y_t.$$

   Quanto vale mediamente la serie y_t^*? (Si ricordi che $E(\varepsilon_t) = 0$, $\forall t$).

Soluzione

1. La variazione della serie trimestrale nell'arco di un anno è data dalla differenza $y_t - y_{t-4}$. Se la serie risulta dalla somma di trend, stagionalità e componente accidentale, tale differenza risulta pari a

   $$\begin{aligned} y_t - y_{t-4} &= (T_t - T_{t-4}) + (S_t - S_{t-4}) \\ &\quad +(\varepsilon_t - \varepsilon_{t-4}). \end{aligned} \tag{1}$$

   Nel caso in esame la variazione della serie nell'arco di un anno dovuta alla componente tendenziale è pari a:

   $$T_t - T_{t-4} = 0,5t - 0,5(t-4) = 2.$$

   La variazione della serie nell'arco di un anno dovuta alla componente stagionale è zero trattandosi di una funzione periodica di periodo 4:

   $$S_t - S_{t-4} = 0$$

2. La serie destagionalizzata si ottiene sottraendo alla serie osservata la componente stagionale. Nel caso in questione, il modello presenta la costante e tre variabili dummy, quindi la costante va interpretata come il coefficiente grezzo relativo alla dummy mancante e gli altri come effetti differenziali rispetto a questo.
   Come prima cosa ricaviamo i coefficienti stagionali grezzi:

   $$\begin{aligned} \hat{\gamma}_1^* &= \tilde{\alpha} = 14 \\ \hat{\gamma}_2^* &= \tilde{\gamma}_2 + \tilde{\alpha} = -11 + 14 = 3 \\ \hat{\gamma}_3^* &= \tilde{\gamma}_3 + \tilde{\alpha} = -19 + 14 = -5 \\ \hat{\gamma}_4^* &= \tilde{\gamma}_4 + \tilde{\alpha} = -24 + 14 = -10 \end{aligned}$$

   Ora dobbiamo controllare se i coefficienti stagionali sommano a zero o, equivalentemente, se la loro media è nulla. In questo caso si ha

   $$\bar{\gamma} = \frac{14 + 3 - 5 - 10}{4} = 0.5 \neq 0.$$

Di conseguenza, per destagionalizzare bisogna fare riferimento ai coefficienti ideali, che si ottengono considerando gli scarti dei coefficienti grezzi dalla loro media:

$$\begin{aligned}\hat{\gamma}_1 &= \hat{\gamma}_1^* - \bar{\gamma} = 14 - 0.5 = 13.5\\ \hat{\gamma}_2 &= \hat{\gamma}_2^* - \bar{\gamma} = 3 - 0.5 = -3.5\\ \hat{\gamma}_3 &= \hat{\gamma}_3^* - \bar{\gamma} = -5 - 0.5 = -5.5\\ \hat{\gamma}_4 &= \hat{\gamma}_4^* - \bar{\gamma} = -10 - 0.5 = -10.5\end{aligned} \tag{2}$$

Ora possiamo utilizzare i coefficienti ideali per destagionalizzare i valori osservati: Avremo quindi:

$$\begin{aligned}t = 1&: \quad \hat{y}_1^d = 16 - 13.5,5 = 2.5\\ t = 2&: \quad \hat{y}_2^d = 4 - 3.5 = 1.5\\ t = 3&: \quad \hat{y}_3^d = -2 - (-5.5) = 3.5\\ t = 4&: \quad \hat{y}_4^d = -9 - (10.5) = 1.5.\end{aligned}$$

3. L'espressione per la serie y_t^* coincide con quella fornita precedentemente ovvero

$$\begin{aligned}y_t - y_{t-4} &= (T_t - T_{t-4}) + (S_t - S_{t-4})\\ &\quad +(\varepsilon_t - \varepsilon_{t-4}).\end{aligned} \tag{3}$$

Allora si era applicato implicitamente l'operatore differenza stagionale $(1 - B^4)$. Calcolando ora il valore atteso dei singoli termini otteniamo

$$\begin{aligned}E(y_t^*) &= E(T_t - T_{t-4}) + E(S_t - S_{t-4})\\ &\quad +E(\varepsilon_t - \varepsilon_{t-4}).\end{aligned}$$

Il secondo termine si annulla per quanto detto al punto (a). Pertanto il valore medio di y_t^* è pari a

$$E(y_t^*) = T_t - T_{t-4} - E(\varepsilon_t - \varepsilon_{t-4}).$$

Considerato il fatto che il trend è di tipo deterministico e quindi non è una variabile casuale, tenendo conto che il valore atteso di una somma è pari alla somma dei valori attesi e che il valore atteso della componente accidentale è nullo, si ottiene che in media la variazione annuale della serie y_t^* è pari a

$$E(y_t^*) = T_t - T_{t-4} = 0,5t - 0,5(t-4) = 2.$$