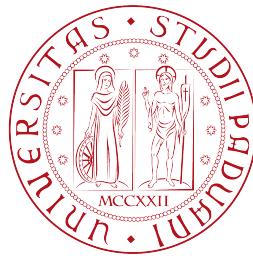


UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE STATISTICHE

CORSO DI LAUREA MAGISTRALE IN
SCIENZE STATISTICHE



Modelli mSARIMA per l'analisi di serie storiche con stagionalità multipla

Relatore:
PROF. FRANCESCO LISI

Laureando:
DIEGO FAGGIAN
1241873

Anno Accademico 2022/2023

Indice

Introduzione	1
1 Modelli per serie storiche multiciclo	3
1.1 Metodo MSTL	6
1.2 Modello Holt-Winters	6
1.3 Modelli BATS	7
1.4 Modelli TBATS	8
2 Modelli mSARIMA	9
2.1 Stima	11
2.2 Diagnostica	13
2.2.1 Distribuzione della autocorrelazione nei residui	14
2.2.2 Test per la stagionalità e multistagionalità	15
2.3 Previsione	16
3 Studio di simulazione	19
3.1 Diagnostica	37
3.2 Confronto tra modelli	38
3.2.1 Analisi dei residui	39
3.2.2 Previsione	41
Conclusioni	47
A Tabelle	49
A.1 Stime e varianze campionarie	49
A.2 Test Jarque-Bera distribuzione stimatori	55
Bibliografia	59

Introduzione

L'argomento esposto in questa tesi deriva dalle necessità di trattare la stagionalità multiciclo, caratteristica osservata in molte serie storiche, adottando i modelli MSARIMA.

La stagionalità, nelle serie temporali, si riferisce a un pattern regolare e prevedibile di fluttuazioni o variazioni che si verificano a intervalli specifici. Questi pattern spesso si ripetono in un periodo fisso, ad esempio giornaliero, settimanale, mensile o annuale. I pattern stagionali possono essere osservati in vari tipi di dati, come per esempio gli indicatori economici, i dati meteorologici, i prezzi delle azioni, il livello delle vendite al dettaglio ed altro ancora.

L'obiettivo di questa tesi è, dunque, quello di introdurre i modelli MSARIMA, sviluppare un test per la diagnostica dei residui in presenza di serie multiciclo e confrontare questa classe di modelli, rispetto ai principali modelli presenti in letteratura.

A supporto di questo obiettivo, è stato sviluppato un software R per trattare questi modelli.

Il primo capitolo introduce il problema della stagionalità multiciclo, le implicazioni e i principali modelli per trattarla, evidenziandone gli aspetti positivi e negativi.

Nel secondo capitolo vengono, poi, introdotti i modelli MSARIMA, gli aspetti inferenziali e la relativa diagnostica con l'utilizzo di un test statistico per verificare congiuntamente la stagionalità multipla nei residui del modello e le problematiche riscontrate nella realizzazione del software R.

Il capitolo tre contiene uno studio di simulazione, al fine di verificare il corretto comportamento del software con le proprietà asintotiche del modello, la diagnostica dei residui ed il confronto con altri modelli.

Infine, nel quarto capitolo, vengono applicati i modelli MSARIMA e i principali

modelli per la stagionalità multipla nelle serie storiche reali.

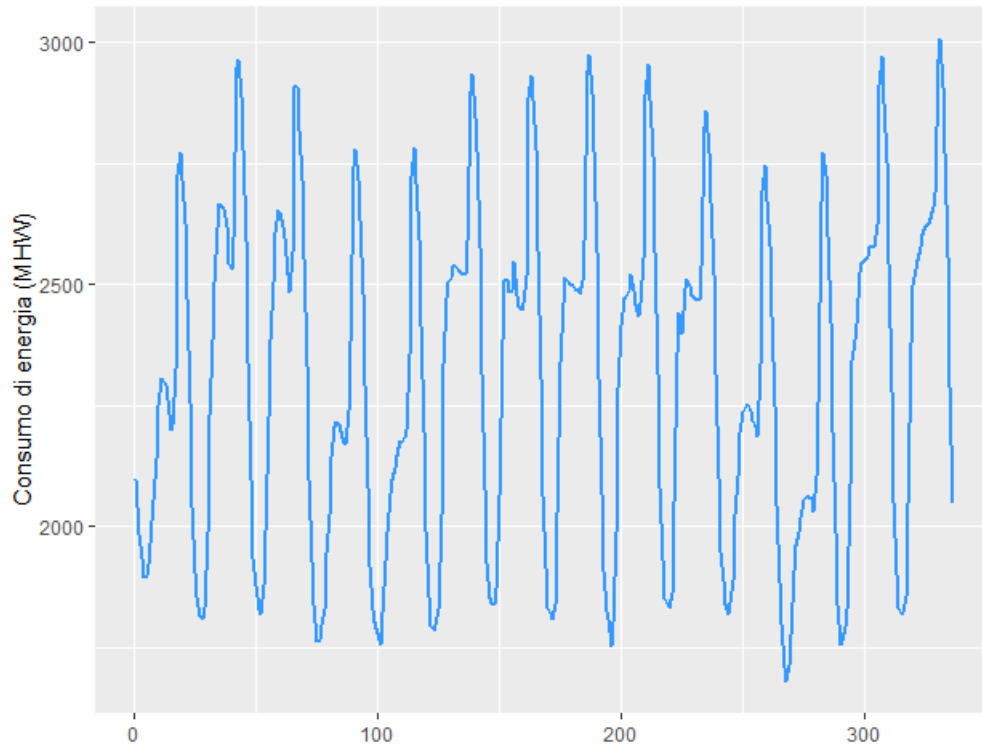
Capitolo 1

Modelli per serie storiche multiciclo

Per stagionalità multipla si intende la presenza di più cicli stagionali all'interno di una serie storica. A differenza della stagionalità singola, dove variazioni regolari si ripetono in modo simile all'interno del periodo considerato, la stagionalità multipla (o multiciclo) prevede la presenza di più pattern ciclici di periodi diversi. Ad esempio, si può osservare un ciclo mensile legato ad eventi con cadenza ricorrente ogni mese, di conseguenza nello stesso tempo è possibile osservare un ciclo annuale.

Un primo esempio in cui è riscontrabile la stagionalità multipla riguarda la commodity dell'energia.

Come si può osservare in figura 1.1, la serie della domanda energetica oraria mostra un ciclo giornaliero ed un ciclo settimanale. Ciò vuol dire che le variazioni della domanda di energia variano in maniera regolare sia all'interno della giornata sia tra i giorni della settimana.



(a) $(1, 0)(1, 0)_4(0, 0)_7$

Figura 1.1: Consumo orario (MHW) San Francisco.

Altri esempi in cui è possibile osservare stagionalità multipla sono riferibili alle serie storiche delle chiamate al call centers, delle richieste di prelievo contanti presso gli sportelli ATM, degli accessi al pronto soccorso ed altri eventi assimilabili.

Al giorno d'oggi, con la rapida crescita della disponibilità di sensori e capacità di archiviazione dei dati, si osserva un aumento significativo delle serie temporali con frequenze di campionamento più elevate (sub-oraria, oraria, quotidiano). I dati delle serie temporali con frequenza più elevata possono mostrare proprietà più complesse, come appunto cicli stagionali multipli e stagioni non intere.

Tornando al primo esempio, la previsione della domanda di energia elettrica è fondamentale per le aziende che operano in questo mercato.

La previsione può riferirsi a diversi orizzonti temporali: lungo periodo, medio periodo e breve periodo; Nel lungo periodo le previsioni sono necessarie per pianificare la capacità di carico della rete e della relativa manutenzione; Nel medio periodo, invece, le previsioni sono fondamentali per l'operatività del sistema; Infine, nel breve periodo occorrono per mantenere l'equilibrio sulla rete.

Infatti, il buon funzionamento della rete prevede che la quantità di energia immessa in rete sia in equilibrio. E' fondamentale che la domanda non superi l'offerta per evitare che non ne venga erogata in quantità sufficiente, al contrario è ugualmente importante che l'offerta non superi la domanda per evitare dispersioni.

Pertanto, avere previsioni non affidabili ha un impatto significativo per chi opera in questo mercato ed è quindi essenziale ricorrere a modelli statistici che siano in grado sia di spiegare la stagionalità multipla, sia di produrre previsioni affidabili.

Diversi approcci e modelli statistici sono stati presentati in letteratura nel corso degli anni: medie mobili, metodi di liscio esponenziale, modelli ARIMA, modelli State Space e più recentemente le reti neurali.

Di seguito vengono presentati i principali modelli statistici per trattare la stagionalità multipla.

1.1 Metodo MSTL

Il Multiple Seasonal-Trend decomposition using Loess (MSTL) è una estensione del tradizionale Seasonal-Trend decomposition using Loess (STL) e consente la decomposizione di serie storiche con più andamenti stagionali.

Il modello MSTL assume che le serie storiche possono essere espresse come una decomposizione additiva della forma

$$y_t = \hat{T}_t + \hat{S}_t^{(1)} + \dots + \hat{S}_t^{(N)} + \hat{R}_t \quad (1.1)$$

dove ogni componente stagionale rappresenta una stagionalità differente (ad esempio, giornaliera, settimanale, mensile, annuale).

Nello specifico MSTL è un algoritmo per la scomposizione additiva delle serie storiche con lo scopo di gestire più stagionalità, tra le particolarità di questo metodo è particolarmente veloce e computazionalmente efficiente. Semplificando, come primo passo vengono identificate le frequenze dei pattern stagionali e questi ultimi vengono ordinati in ordine crescente. Le frequenze che sono minori della metà della lunghezza della serie vengono ignorate. Se specificato, viene applicata una trasformazione Box-Cox alle osservazioni di conseguenza per ogni ciclo stagionale individuato, viene eseguita la decomposizione STL.

1.2 Modello Holt-Winters

Il metodo Holt-Winters, estensione del metodo di Holt, consiste in un sistema di equazioni ricorsivo in grado di decomporre la serie y_t in diverse componenti: il livello, il trend e la componente stagionale. Il metodo può essere visto come un modello state space.

La formulazione più generica del modello Holt-Winters additivo è la seguente

$$\begin{aligned} y_t &= l_{t-1} + b_{t-1} + s_t^{(1)} + s_t^{(2)} + d_t, \\ l_t &= l_{t-1} + b_{t-1} + \alpha d_t, \\ b_t &= b_{t-1} + \beta d_t, \\ s_t^{(1)} &= s_{t-m_1}^{(1)} + \gamma_1 d_t, \\ s_t^{(2)} &= s_{t-m_2}^{(2)} + \gamma_2 d_t \end{aligned} \quad (1.2)$$

dove $d_t \sim NID(0, \sigma^2)$ e α, β, γ_w sono parametri di lisciamiento per il livello, il trend e la stagionalità. I parametri di lisciamiento riflettono quanto velocemente la rispettiva componente si adatta a nuova informazione.

Mentre $l_0, b_0, \{s_{1-m_1}^{(1)}, \dots, s_0^{(1)}\}$, e $\{s_{1-m_2}^{(2)}, \dots, s_0^{(2)}\}$ sono i valori iniziali degli stati. I parametri e i valori iniziali devono essere stimati, ma questo può essere complicato quando il numero di componenti stagionali è elevato.

Inoltre questo modello non permette di gestire pattern stagionali complessi come stagionalità non intera ed effetti di calendario.

1.3 Modelli BATS

Il modello BATS è una variazione dei modelli state space per il lisciamiento esponenziale in grado di gestire una varietà più ampia di pattern stagionali.

Questa classe di modelli inoltre permette qualche forma di linearità tramite l'utilizzo della trasformazione Box-Cox, anche se questo limita a solo serie positive. La notazione $y_t^{(\omega)}$ è usata per rappresentare la trasformazione Box-Cox dei dati con parametro ω .

Quindi è possibile estendere il modello di Holt-Winters (1.2) includendo la trasformazione Box-Cox, errori ARMA e T pattern stagionali:

$$\begin{aligned}
 y_t^\omega &= \begin{cases} \frac{y_t^\omega - 1}{\omega}, & \omega \neq 0 \\ \log(y_t) & \omega = 0 \end{cases} \\
 y_t^{(\omega)} &= l_{t-1} + \phi b_{t-1} + \sum_{i=1}^T s_{t-m_i}^{(i)} + d_t, \\
 l_t &= l_{t-1} + \phi b_{t-1} + \alpha d_t, \\
 b_t &= (1 - \phi)b + \phi b_{t-1} + \beta d_t, \\
 s_t^{(i)} &= s_{t-m_i}^{(i)} + \gamma_i d_t, \\
 d_t &= \sum_{i=1}^p \psi d_{t-i} + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon_t
 \end{aligned} \tag{1.3}$$

dove m_1, \dots, m_T indicano i periodi stagionali, l_t è il livello locale al periodo t , b rappresente il trend di lungo periodo, b_t il trend di breve periodo, $s_t^{(i)}$ è la componente stagionale i -esima al tempo t , d_t è un processo $ARMA(p, q)$, e ϵ_t è un processo White Noise Gaussiano con media zero e varianza costante σ^2 .

I parametri di lisciamiento sono α, β e γ_i per $i = 1, \dots, T$.

Questa classe di modelli sono menzionati con il nome di modelli BATS (Box-Cox transform, ARMA errors, Trend, and Seasonal components).

Anche questi modelli non sono in grado di gestire la stagionalità non intera, e possono avere un alto numero di stati; la sola componente stagionale contiene m_T stati non zero. Questo comporta un alto numero di valori per i pattern stagionali con periodi elevati.

1.4 Modelli TBATS

Alla ricerca di un approccio più flessibile e parsimonioso, si è introdotto una rappresentazione trigonometrica per le componenti stagionali basate sulle serie di Fourier:

$$\begin{aligned}
 S_t^{(i)} &= \sum_{j=1}^{k_i} s_{j,t}^{(i)}, \\
 s_{j,t}^{(i)} &= s_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin \lambda_j^{(i)} + \gamma_1^{(i)} d_t, \\
 s_{j,t}^{*(i)} &= -s_{j,t-1}^{(i)} \sin \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos \lambda_j^{(i)} + \gamma_2^{(i)} d_t
 \end{aligned} \tag{1.4}$$

dove $\gamma_1^{(i)}, \gamma_2^{(i)}$ sono parametri di lisciamiento e $\lambda_j^{(i)} = \frac{2\pi j}{m_i}$.

Questi modelli prendono il nome di modelli TBATS (Trigonometric BATS). Un modello TBATS richiede la stima di $2(k_1 + k_2 + \dots + k_T)$ valori stagionali iniziali, un numero molto minore rispetto il modello BATS.

In aggiunta, l'utilizzo di funzioni trigonometriche, permette di usare frequenze stagionali non intere.

Capitolo 2

Modelli mSARIMA

I modelli SARIMA permettono di modellare andamenti periodici di tipo stagionale, ma trattano solo una stagionalità, mentre come si spiegato precedentemente, molte serie storiche presentano stagionalità multiciclo.

I modelli mSARIMA estendono i modelli SARIMA permettendo la modellazione di m stagionalità, in particolare, l'utilizzo di questi modelli ha lo scopo di trattare le componenti congiuntamente.

L'idea di fondo di questi modelli scaturisce dunque dalla necessità di descrivere più relazioni all'interno della serie osservata. La relazione tra valori consecutivi viene modellata dalla parte ARIMA e le m relazioni che sussistono tra osservazioni che distano tra loro di un multiplo del periodo S_i con $i = 1, \dots, m$.

Definizione 1 Sia $\epsilon_t \sim WN(0, \sigma^2)$ con $\sigma_2 < \infty$.

Si indica con $mSARIMA(p, q)(P_1, Q_1)_{S_1} \cdots (P_m, Q_m)_{S_m}$, il processo stazionario Y_t

$$\phi(B) \prod_{i=1}^m \Phi(B^{S_i}) Y_t = \theta(B) \prod_{i=1}^m \Theta(B^{S_i}) \epsilon_t \quad (2.1)$$

dove $\phi(B)$ e $\theta(B)$ sono gli usuali polinomi AR e MA in B di grado p e q , e B è l'operatore ritardo definito come $(1 - B)y_t = y_t - y_{t-1}$, mentre

$$\Phi(B^{S_i}) = (1 - \Phi_{i,1} B^{S_i} - \cdots - \Phi_{i,P_i} B^{P_i S_i}) \quad (2.2)$$

è il polinomio AR per la componente stagionale i -esima di grado P_i e periodo

S_i e

$$\Theta(B^{S_i}) = (1 - \Theta_{i,1}B^{S_i} - \dots - \Theta_{i,P_i}B^{P_i S_i}) \quad (2.3)$$

è il corrispondente polinomio MA per la componente stagionale i -esima di grado Q_i e periodo S_i .

I polinomi $\phi(B), \theta(B)$ descrivono la dipendenza seriale non stagionale della serie storica, mentre i polinomi $\Phi(B^{S_i}), \Theta(B^{S_i})$ modellano la correlazione periodica per le m componenti stagionali di periodo S_i con $(i = 1, \dots, m)$.

Ad esempio un modello $2SARIMA(1,0)(1,0)_4(1,0)_7$ ha i seguenti polinomi

$$(1 - \phi B)(1 - \Phi_{1,1}B^4)(1 - \Phi_{2,1}B^7)Y_t = \epsilon_t \quad (2.4)$$

e calcolando il prodotto tra polinomi si ottiene la seguente forma esplicita

$$Y_t = \phi_1 Y_{t-1} + \Phi_{1,1} Y_{t-4} - \phi_1 \Phi_{1,1} Y_{t-5} + \Phi_{2,1} Y_{t-8} - \Phi_{1,1} \Phi_{2,1} Y_{t-11} - \phi_1 \Phi_{1,1} \Phi_{2,1} Y_{t-12} + \epsilon_t \quad (2.5)$$

Quindi sono presenti 6 differenti lags e solo 3 parametri da stimare.

Le condizioni di stazionarietà per il modello sono le stesse di un modello SARMA, infatti esplicitati i vincoli, un modello mSARIMA può essere visto come un SARIMA vincolato. Lo stesso vale per le condizioni di invertibilità.

Il modello (2.1) può essere generalizzato per il caso non stazionario includendo dei polinomi per le radici unitarie

$$\phi(B)(1 - B)^d \prod_{i=1}^m \Phi(B^{S_i})(1 - B^{S_i})^{d_i} Y_t = \theta(B) \prod_{i=1}^m \Theta(B^{S_i}) \epsilon_t \quad (2.6)$$

dove $(1 - B^S)^d = (Y_t - Y_{t-S})^d$ è la differenza d -esimo stagionale di Y_t .

In questa tesi si è scelto di trattare solo la parte autoregressiva del modello, quindi se non specificato diversamente, si farà sempre riferimento a modelli 2SARI, dunque con due stagionalità per la parte autoregressiva.

$$\phi(B)\Phi_1(B^{S_1})\Phi_2(B^{S_2})(1 - B)^d(1 - B^{S_1})^{d_1}(1 - B^{S_2})^{d_2} Y_t = \epsilon_t \quad (2.7)$$

2.1 Stima

Per la stima dei parametri di un modello 2SARI si può ricorrere alla stima di massima verosimiglianza. Come si è appurato in precedenza, il modello 2SARI è un SARIMA vincolato, pertanto è necessario definire la verosimiglianza specifica per ogni ordine e ordine di integrazione prima di procedere alla massimizzazione della verosimiglianza.

Il vettore dei parametri è dato da

$$\delta = (\phi_0, \phi_1, \dots, \phi_p, \Phi_{1,1}, \dots, \Phi_{1,P_1}, \Phi_{2,1}, \dots, \Phi_{2,P_2}) \quad (2.8)$$

La stima di massima verosimiglianza consiste nel massimizzare la funzione di verosimiglianza, ovvero la funzione di densità congiunta delle variabili casuali che definiscono il processo stocastico

$$L(\delta) = f(Y_1, \dots, Y_n | \delta) \quad (2.9)$$

ed è definita come quel valore $\hat{\delta}$ tale che

$$L(\hat{\delta}) = \max_{\delta} L(\delta) \quad (2.10)$$

Riuscire a ricavare la funzione di densità congiunta dell'intero vettore delle osservazioni è generalmente complicato.

Sotto ipotesi di indipendenza tra le variabili casuali del processo Y_1, \dots, Y_n , la verosimiglianza può essere scritta come

$$L(\delta) = f(y_1, \dots, y_n | \delta) = \prod_{t=1}^n f(y_t | \delta) \quad (2.11)$$

Tuttavia le serie storiche, spesso, evidenziano una autocorrelazione, anche se debole, e quindi l'ipotesi di indipendenza viene violata.

E' possibile quindi utilizzare la formula della probabilità composta e fattorizzare la densità congiunta nel prodotto di densità condizionate più facilmente trattabili.

Nello specifico, si può scomporre la funzione di densità congiunta dell'intero campione come il prodotto delle densità delle prime m variabili e di $n - m$ densità

condizionate ai valori assunti dalle variabili casuali precedenti.

$$L(\delta) = f(y_1, \dots, y_n | \delta) = f(y_1, \dots, y_m | \delta) \cdot \prod_{t=m+1}^n f(y_t | I_{t-1}; \delta) \quad (2.12)$$

dove $I_{t-1} = y_{t-1}, y_{t-2}, \dots$ è la storia passata del processo fino al tempo $t - 1$ compreso.

Anzichè massimizzare $L(\delta)$, per convenienza computazionale, si può massimizzare il logaritmo della verosimiglianza, $l(\delta) = \log(L(\delta))$.

Quando si passa al logaritmo la (2.12) diventa

$$l(\delta) = \log f(y_1, \dots, y_m | \delta) + \sum_{t=m+1}^n \log f(y_t | I_{t-1}; \delta) \quad (2.13)$$

Per massimizzare $l(\delta)$, sotto opportune ipotesi sulla distribuzione delle ϵ_t le densità condizionate $f(y_t | I_{t-1}; \delta)$ si riescono a calcolare.

La distribuzione congiunta $f(y_1, \dots, y_m | \delta)$, invece, è complicata da ricavare e si riesce a calcolarla solo per valori di m molto piccoli, $m = 1$ o $m = 2$. Nel contesto delle serie storiche, in cui la lunghezza del campione è generalmente elevata, il contributo dato dal primo termine di (2.13) al valore totale della log-verosimiglianza è trascurabile.

Di conseguenza è possibile massimizzare una espressione approssimata della log-verosimiglianza:

$$l(\delta) \approx \sum_{t=m+1}^n \log f(y_t | I_{t-1}; \delta) \quad (2.14)$$

che trascura la densità congiunta delle prime m variabili ed assume m condizioni iniziali ottenute, ad esempio, ponendo pari a zero i primi valori mancanti.

E' possibile mostrare che gli stimatori di massima verosimiglianza esatta e condizionata hanno la stessa distribuzione asintotica (Hamilton, 1994).

Ad esempio per un $2SARI(1, 0)(1, 0)_4(1, 0)_7$ con ipotesi di gaussianità delle in-

novazioni ϵ_t la log-verosimiglianza condizionata può essere scritta come

$$\begin{aligned}
 l(\delta) &\approx \sum_{t=p+P_1+P_2+1}^n \log(f(Y_t|I_{t-1}; \delta)) = \\
 &= \log\left[\left(\frac{1}{\sqrt{2\phi\sigma_\epsilon^2}}\right)^{n-(p+P_1+P_2)} \exp\left\{-\frac{1}{2\sigma_\epsilon^2} \sum_{t=p+P_1+P_2+1}^n (y_t - m(t))^2\right\}\right] \quad (2.15) \\
 &\approx -\frac{n - (p + P_1 + P_2)}{2} \log(\sigma_\epsilon^2) - \frac{1}{2\sigma_\epsilon^2} \sum_{t=p+1}^n (y_t - m(t))^2
 \end{aligned}$$

dove $m(t) = (1 - \phi B)(1 - \Phi_1 B^4)(1 - \Phi_2 B^7)$.

2.2 Diagnostica

Successivamente alla stima del modello, il passo finale è quello di controllarne l'adeguatezza tramite delle analisi e test diagnostici.

Se il modello è correttamente identificato e stimato allora i residui, definiti come $\hat{\epsilon}_t = (y_t - \hat{y}_t)$, devono soddisfare le ipotesi fatte a priori sulle innovazioni del modello ϵ_t .

In particolare è importante verificare che i residui siano incorrelati, pertanto è necessario studiare l'autocorrelazione, cioè il grado di dipendenza lineare tra valori ritardati dei residui.

E' definita nel seguente modo

$$\hat{r}_k = \frac{\sum_{i=1}^n \hat{\epsilon}_i \hat{\epsilon}_{i-k}}{\sum_{i=1}^n \hat{\epsilon}_i^2} \quad (2.16)$$

Box, G. E. e Pierce, D. A. ([1] 1970) dimostrarono che per un modello ARMA(p,q) se il modello è correttamente specificato e la serie dei residui viene calcolata usando i veri parametri, allora ϵ_t sono variabili casuali incorrelate, e per le prime m autocorrelazioni $\mathbf{r} = (r_1, r_2, \dots, r_m)^T$, quando m è relativamente piccolo rispetto a n , allora al crescere di n , il vettore ha distribuzione normale multivariata.

$$\mathbf{r} \sim \mathbf{N}(\mathbf{0}, (\mathbf{1}/n)\mathbf{I}) \quad (2.17)$$

Inoltre

$$Var(r_k) = \frac{n - k}{n(n + 2)} \approx \frac{1}{n} \quad (2.18)$$

Da questo risultato segue che la quantità $n(n+2) \sum_{k=1}^n (n-k)^{-1} r_k^2$ al crescere di n ha una distribuzione χ_m^2 .

2.2.1 Distribuzione della autocorrelazione nei residui

Sempre Box, G. E. e Pierce, D. A. ([1] 1970), dimostrano che è possibile ricavare la distribuzione approssimata delle autocorrelazioni dei residui empirici $\hat{\mathbf{r}}_k$ usando l'espansione di Taylor al primo ordine nel punto $\hat{\phi} = \phi$.

Quindi è possibile esprimere \hat{r}_k in funzione delle autocorrelazioni dei residui teorici tramite

$$\hat{r}_k = r_k + \sum_{j=1}^p (\phi_j - \hat{\phi}_j) \hat{\delta}_{jk} + O_p(1/n) \quad (2.19)$$

dove $\hat{\delta}_{jk}$ rappresenta la derivata del residuo k rispetto al parametro j , cambiata di segno. Dopo ulteriori passaggi algebrici ed esplicitando la relazione $\delta_{jk} = \psi_{k-j}$, dove ψ_{k-j} rappresenta il coefficiente $k-j$ della rappresentazione $MA(\infty)$ del processo, si ottiene

$$\hat{r}_k = r_k + \sum_{j=1}^p (\phi_j - \hat{\phi}_j) \psi_{k-j} + O_p(1/n) \quad (2.20)$$

Oppure scritta in forma matriciale

$$\hat{\mathbf{r}} = \mathbf{r} + \mathbf{X}(\phi - \hat{\phi}) \quad (2.21)$$

Usando la proprietà di ortogonalità dei residui $\hat{\mathbf{r}}^T \mathbf{X} = \mathbf{0}$, si ottiene

$$\hat{\mathbf{r}} = (\mathbf{I} - \mathbf{Q})\mathbf{r} \quad (2.22)$$

con $\mathbf{Q} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$.

Conoscendo la distribuzione dei residui teorici (2.17), si ricava

$$\hat{\mathbf{r}} \sim \mathbf{N}(\mathbf{0}, (1/n)[\mathbf{I} - \mathbf{Q}]) \quad (2.23)$$

Con $\mathbf{I} - \mathbf{Q}$ idempotente e di rango $m - p$.

Da questi risultati deriva che la statistica test

$$\hat{Q}(m) = n(n+2) \sum_{k=1}^m (n-k)^{-1} \hat{r}_k^2 \sim \chi_{m-p-q}^2 \quad (2.24)$$

può essere utilizzata per testare l'ipotesi nulla di incorrelazione dei residui.

2.2.2 Test per la stagionalità e multistagionalità

Questo test però non è sufficiente per verificare la stagionalità in quanto considera tutte le autocorrelazioni corrispondenti a ritardi fino a m lag di distanza e ignora la possibilità di autocorrelazioni a ritardi stagionali multipli del periodo s .

McLeod, A. I. ([3] 1978) dimostra che la varianza dei residui campionari può essere espressa come

$$\mathbf{Var}(\hat{\mathbf{r}}) = (\mathbf{1} - \mathbf{X}\mathbf{I}^{-1}\mathbf{X}^T)/n \quad (2.25)$$

In un modello SARMA, le autocorrelazioni dei residui ($\hat{r}_{1s}, \dots, \hat{r}_{ms}$) hanno approssimativamente la stessa matrice di varianze/covarianze delle prime m autocorrelazioni dei residui in modello ARMA con i polinomi caratteristici del modello SARMA.

$$\Phi(B)y_t = \Theta(B)\epsilon_t \quad (2.26)$$

A fronte di questo risultato propone la statistica test

$$\hat{Q}_m(s) = n(n+2) \sum_{k=1}^m (n-kS)^{-1} \hat{r}_{kS}^2 \sim \chi_{m-P-Q}^2 \quad (2.27)$$

dove P, Q sono gli ordini della parte AR e MA stagionale del modello SARMA, per testare l'incorrelazione dei residui nei ritardi stagionali.

Per i modelli 2SARI è necessario testare congiuntamente l'incorrelazione dei residui per entrambe le stagionalità.

Assumendo che valga il risultato di McLeod anche per la struttura della matrice di varianza/covarianza del vettore dei residui di un modello 2SARI, la statistica

test proposta è la seguente

$$\hat{Q}(m_1, m_2) = n(n+2) \left[\sum_{k=1}^{m_1} (n - kS_1)^{-1} \hat{r}_{kS_1}^2 + \sum_{k=1}^{m_2} (n - kS_2)^{-1} \hat{r}_{kS_2}^2 \right] \sim \chi_{m_1+m_2-P_1-P_2}^2 \quad (2.28)$$

Si ipotizza che valga il risultato di McLeod and per modelli mSARIMA, in particolare la struttura della matrice di varianza/covarianza del vettore dei residui.

Si dimostra facilmente che trattandosi della somma di due variabili casuali con distribuzione rispettivamente $\chi_{m_1-P_1}$, $\chi_{m_2-P_2}$, la statistica test ha distribuzione

$$\chi_{m_1+m_2-P_1-P_2}.$$

Questo risultato verrà dimostrato per simulazione nel capitolo 3.

2.3 Previsione

La previsione è uno degli obiettivi dell'analisi delle serie storiche. La previsione è intesa come stima di un valore di una variabile casuale tramite un prefissato insieme di informazione, chiamato *informazione al tempo t* e indicato con I_t . Questo insieme può contenere diverse informazioni e variabili, ma in questo caso ci si limita all'insieme di tutte le variabili che compongono un processo fino al tempo t .

Prevedere il processo al tempo $t+k$ significa determinare il valore della variabile casuale Y_{t+k} *condizionatamente* ad I_t , cioè alla storia passata del processo.

Si indica con \hat{Y}_{t+k} il previsore di Y_{t+k} cioè la variabile casuale che genera la previsione per un orizzonte temporale pari a k periodi, questa variabile è una funzione dell'informazione passata, $g(I_t)$.

La strategia seguita nella classe dei modelli ARIMA e loro varianti, è quella di selezionare la funzione $g(\cdot)$ che minimizza una funzione perdita, in questo caso l'errore di previsione, definito come

$$e_{t+k} = Y_{t+k} - \hat{Y}_{t+k} \quad (2.29)$$

È possibile dimostrare (Hamilton 1994, p. 73) che la funzione $g(\cdot)$ che minimizza l'errore di previsione è il *valore atteso condizionato*. Quindi è possibile definire il previsore ottimo come

$$\hat{Y}_{t+k} = \mathbb{E}[Y_{t+k}|I_t] \quad (2.30)$$

Ci sono ulteriori aspetti da tenere in considerazione per i modelli della classe ARIMA, ma quanto n è sufficientemente elevato rispetto al numero di parametri

da stimare, si può calcolare il previsore in questo modo

$$\begin{aligned}\hat{Y}_{t+k} &= \mathbb{E}[Y_{t+k}|I_t] \\ &\approx \mathbb{E}[\phi_1 Y_{t+k-1} + \dots + \phi_p Y_{t+k-p} + \epsilon_{t+k} - \dots - \theta_q \epsilon_{t+k-q}]\end{aligned}\tag{2.31}$$

e per il calcolo del valore atteso si applicano le seguenti regole

$$\mathbb{E}[Y_{t+j}|I_t] = \begin{cases} y_{t+j} & j \leq 0 \\ \hat{y}_{t+j} & j > 0 \end{cases}\tag{2.32}$$

$$\mathbb{E}[\epsilon_{t+j}|I_t] = \begin{cases} e_{t+j} & j \leq 0 \\ 0 & j > 0 \end{cases}\tag{2.33}$$

Capitolo 3

Studio di simulazione

Uno studio di simulazione permette di testare, sotto condizioni controllate, l'affidabilità di metodi statistici senza affidarsi completamente ai dati del mondo reale. In quanto è possibile simulare dati con caratteristiche specifiche, e quindi di esplorare scenari rari o difficili da ottenere nella realtà. Inoltre le simulazioni possono venire utilizzate per stimare la potenza statistica dei test, cioè la probabilità di rilevare un effetto se esiste veramente.

Lo studio di simulazione realizzato in questa tesi ha lo scopo di verificare le proprietà di stima del modello 2SARI e di confrontare il modello rispetto ai principali modelli trattati in letteratura sotto varie tipologie di multistagionalità.

A questo scopo è stato realizzato un software R per la generazione, stima e previsione di un modello 2SARI, quindi una modello della classe mSARIMA ma in grado di modellare solo la parte autoregressiva e due stagionalità.

Il modello 2SARI ha la seguente formulazione

$$\phi(B)\Phi_1(B^{S_1})\Phi_2(B^{S_2})(1-B)^d(1-B^{S_1})^{d_1}(1-B^{S_2})^{d_2}Y_t = \epsilon_t \quad (3.1)$$

Il software stima i parametri di un modello 2SARI tramite stima di massima verosimiglianza, quindi è necessario comprovare che la procedura di stima produca stime che siano conformi con le proprietà asintotiche degli stimatori di massima verosimiglianza.

Infatti la stima di massima verosimiglianza, sotto opportune condizioni di regolarità, ha le seguenti proprietà:

- **Consistenza:** Lo stimatore converge in probabilità al suo vero valore, $\hat{\theta}_{mle} \rightarrow \theta_0$;

- **Asintoticamente normale:** La distribuzione dello stimatore converge in distribuzione ad una distribuzione normale $\sqrt{(n)}(\hat{\theta}_{mle} - \theta_0) \rightarrow N(0, J^{-1})$;
- **Efficienza.**

Dove J è la matrice di informazione osservata così definita:

$$J = E\left[-\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta^T}\right] \quad (3.2)$$

Una stima della matrice di informazione osservata può essere ottenuta tramite

$$\hat{J} = -\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t(\hat{\theta})}{\partial \hat{\theta} \partial \hat{\theta}^T} \quad (3.3)$$

Per ogni modello indicato nelle tabelle seguenti, si sono effettuate $M = 5000$ simulazioni per diversi valori della numerosità campionaria, $n = 100, n = 300, n = 700, n = 1500$; Ogni simulazione prevede la generazione di una serie dal modello e la successiva fase di stima.

Modello	Parametri
$2SARI(1, 0)(1, 0)_4(0, 0)_7$	$\phi = 0.6, \Phi_1 = 0.5$
$2SARI(1, 1)(1, 0)_4(0, 0)_7$	$\phi = 0.6, \Phi_1 = 0.5$
$2SARI(1, 0)(1, 1)_4(0, 0)_7$	$\phi = 0.6, \Phi_1 = 0.5$
$2SARI(1, 1)(1, 1)_4(0, 0)_7$	$\phi = 0.6, \Phi_1 = 0.5$
$2SARI(1, 0)(1, 0)_4(0, 0)_7$	$\phi = 0.5, \Phi_1 = 0.1$
$2SARI(1, 1)(1, 0)_4(0, 0)_7$	$\phi = 0.5, \Phi_1 = 0.1$
$2SARI(1, 0)(1, 1)_4(0, 0)_7$	$\phi = 0.5, \Phi_1 = 0.1$
$2SARI(1, 1)(1, 1)_4(0, 0)_7$	$\phi = 0.5, \Phi_1 = 0.1$
$2SARI(1, 0)(1, 0)_4(0, 0)_7$	$\phi = 0.5, \Phi_1 = 0.9$
$2SARI(1, 1)(1, 0)_4(0, 0)_7$	$\phi = 0.5, \Phi_1 = 0.9$
$2SARI(1, 0)(1, 1)_4(0, 0)_7$	$\phi = 0.5, \Phi_1 = 0.9$
$2SARI(1, 1)(1, 1)_4(0, 0)_7$	$\phi = 0.5, \Phi_1 = 0.9$

Modello	Parametri
$2SARI(1, 0)(1, 0)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$
$2SARI(1, 1)(1, 0)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$
$2SARI(1, 0)(1, 1)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$
$2SARI(1, 1)(1, 0)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$
$2SARI(1, 1)(1, 1)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$
$2SARI(1, 1)(1, 0)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$
$2SARI(1, 0)(1, 1)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$
$2SARI(1, 1)(1, 1)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$

Modello	Parametri
$2SARI(1, 0)(1, 0)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$
$2SARI(1, 1)(1, 0)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$
$2SARI(1, 0)(1, 1)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$
$2SARI(1, 1)(1, 0)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$
$2SARI(1, 1)(1, 1)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$
$2SARI(1, 1)(1, 0)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$
$2SARI(1, 0)(1, 1)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$
$2SARI(1, 1)(1, 1)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$

Modello	Parametri
$2SARI(1, 0)(1, 0)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$
$2SARI(1, 1)(1, 0)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$
$2SARI(1, 0)(1, 1)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$
$2SARI(1, 1)(1, 0)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$
$2SARI(1, 1)(1, 1)_4(1, 0)_7$	$\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$
$2SARI(1, 1)(1, 0)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$
$2SARI(1, 0)(1, 1)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$
$2SARI(1, 1)(1, 1)_4(1, 1)_7$	$\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$

Queste stime sono caratterizzate dalla proprietà di consistenza, poichè la distorsione, definita come la differenza tra il valore atteso dello stimatore ed il vero valore, tende a zero al crescere della numerosità campionaria.

$$\mathbb{E}[\hat{\theta}_{mte}] - \theta_0 \rightarrow 0 \quad (3.4)$$

Nei grafici seguenti viene mostrato il valore della distorsione al variare della lunghezza della serie generata, per tutti i modelli presi in considerazione dallo studio di simulazione.

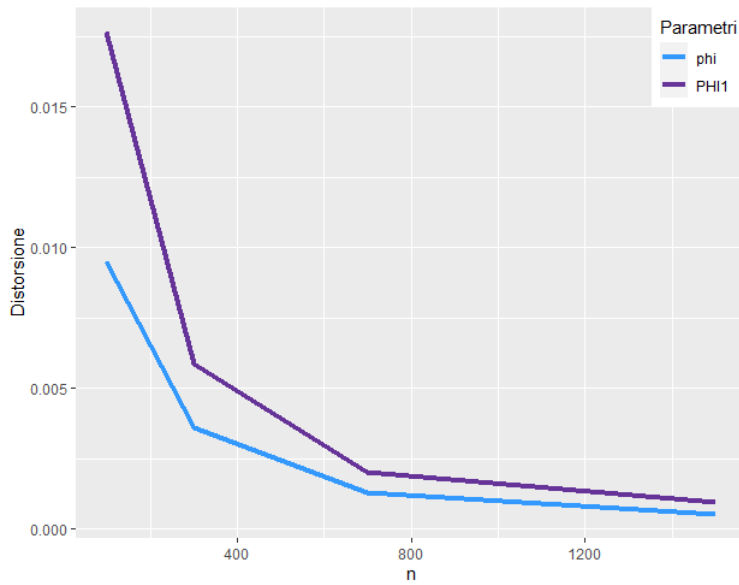


Figura 3.1: Distorsione delle stime per il modello $(1,0)(1,0)_4(0,0)_7$ con $\phi = 0.6$, $\Phi_1 = 0.5$

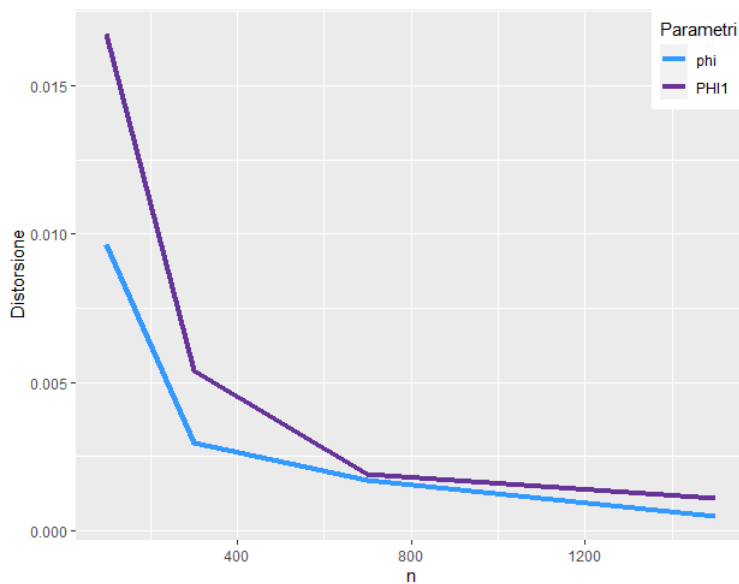


Figura 3.2: Distorsione delle stime per il modello $(1,1)(1,1)_4(0,0)_7$ con $\phi = 0.6$, $\Phi_1 = 0.5$

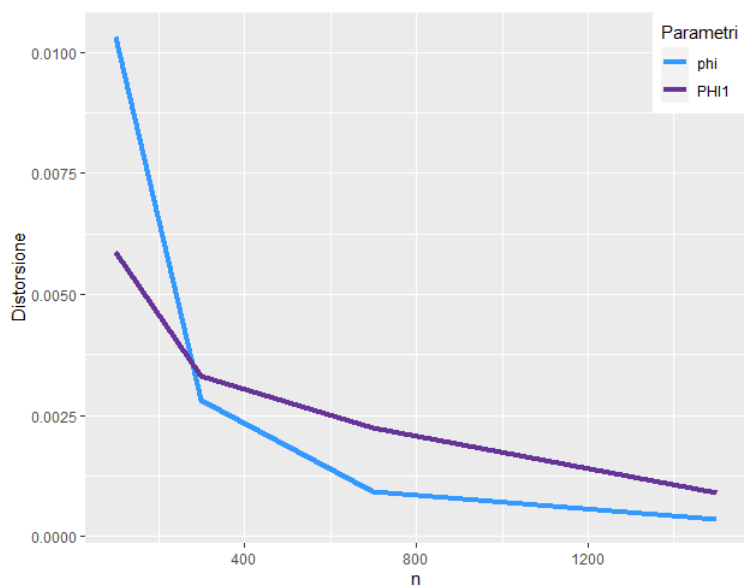


Figura 3.3: Distorsione delle stime per il modello $(1, 0)(1, 0)_4(0, 0)_7\phi = 0.5, \Phi_1 = 0.1$

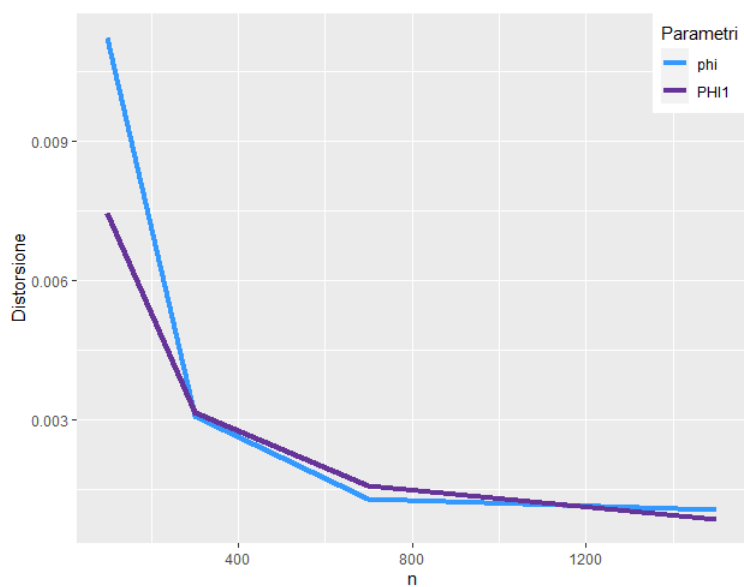


Figura 3.4: Distorsione delle stime per il modello $(1, 1)(1, 1)_4(0, 0)_7\phi = 0.5, \Phi_1 = 0.1$

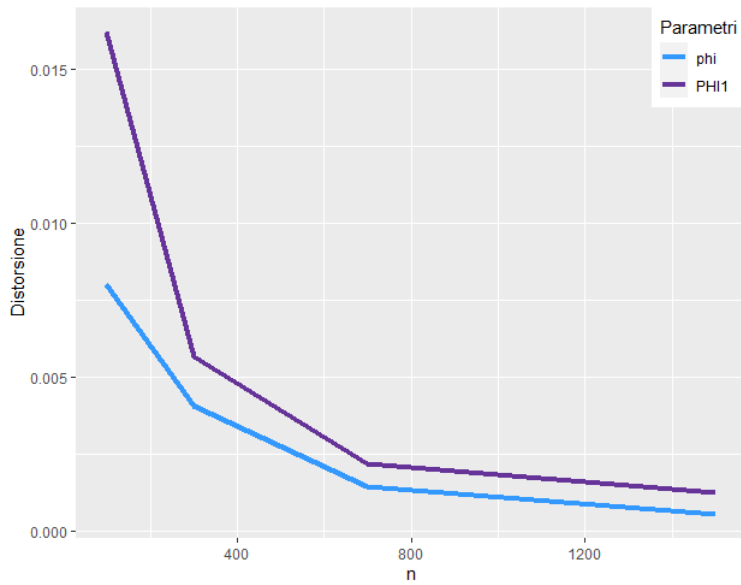


Figura 3.5: Distorsione delle stime per il modello $(1, 0)(1, 0)_4(0, 0)_7$ con $\phi = 0.5$, $\Phi_1 = 0.9$

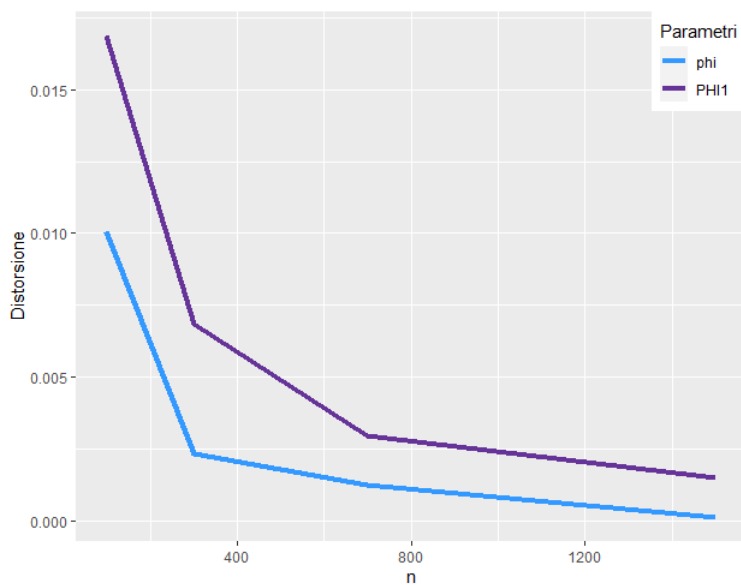


Figura 3.6: Distorsione delle stime per il modello $(1, 1)(1, 1)_4(0, 0)_7$ con $\phi = 0.5$, $\Phi_1 = 0.9$

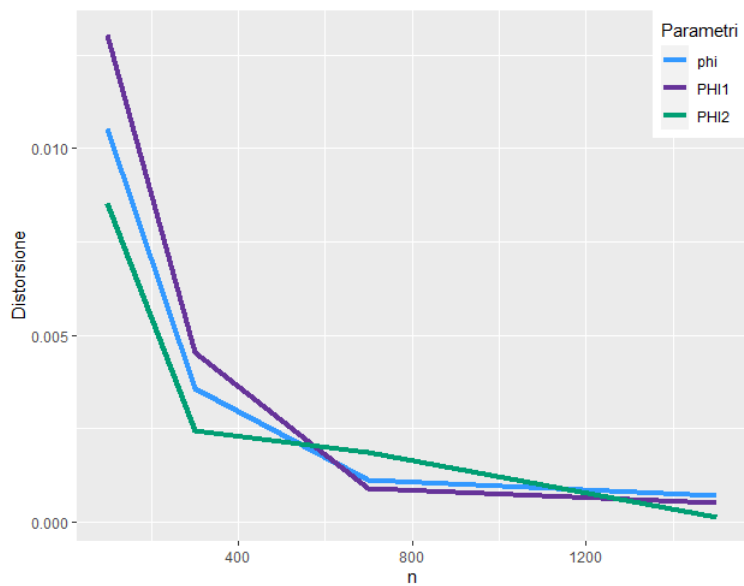


Figura 3.7: Distorsione delle stime per il modello $(1,0)(1,0)_4(1,0)_7\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$

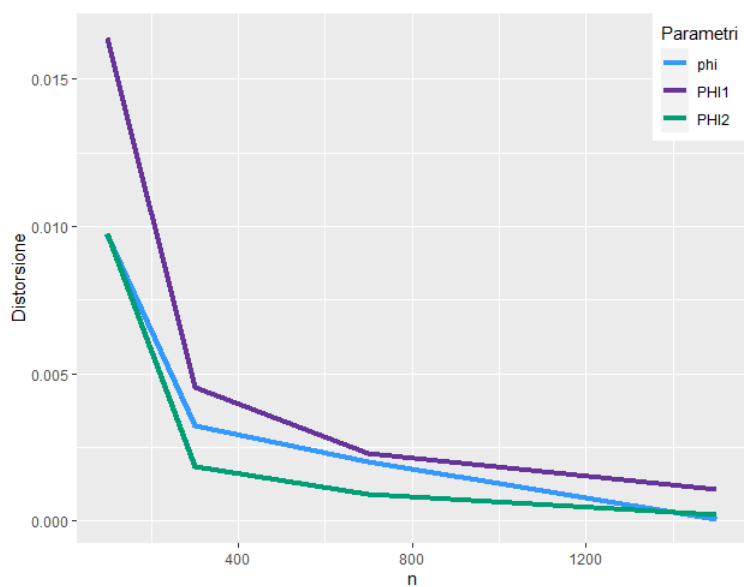


Figura 3.8: Distorsione delle stime per il modello $(1,0)(1,1)_4(1,1)_7\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$

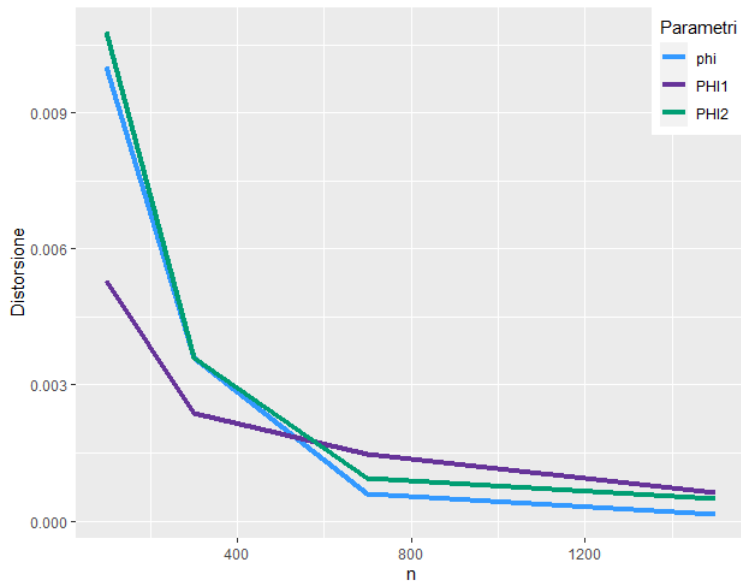


Figura 3.9: Distorsione delle stime per il modello $(1,0)(1,0)_4(1,0)_7 \phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$

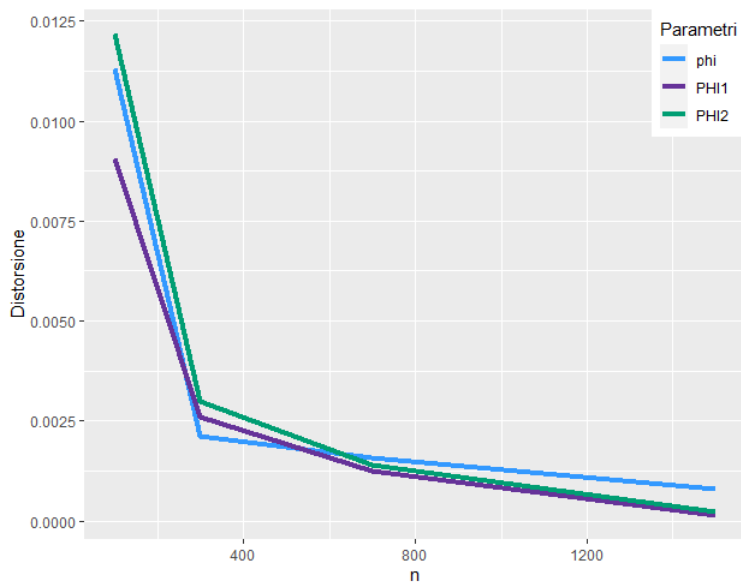


Figura 3.10: Distorsione delle stime per il modello $(1,0)(1,1)_4(1,1)_7 \phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$

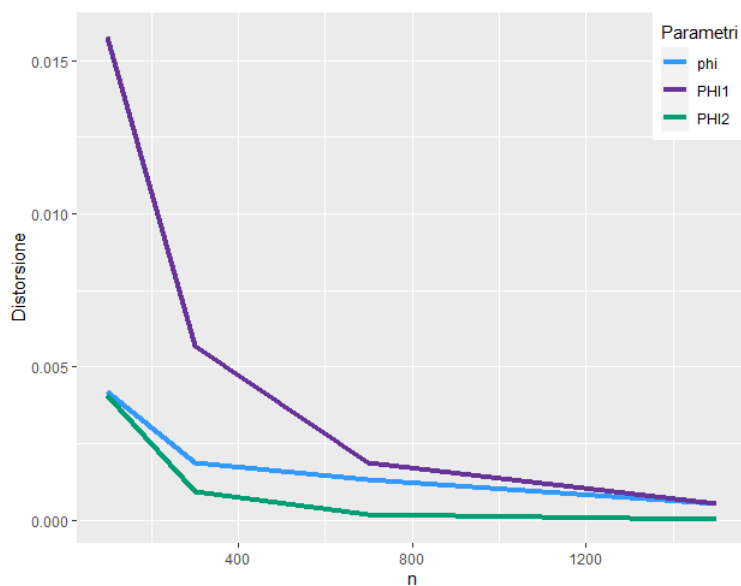


Figura 3.11: Distorsione delle stime per il modello $(1,0)(1,0)_4(1,0)_7\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$

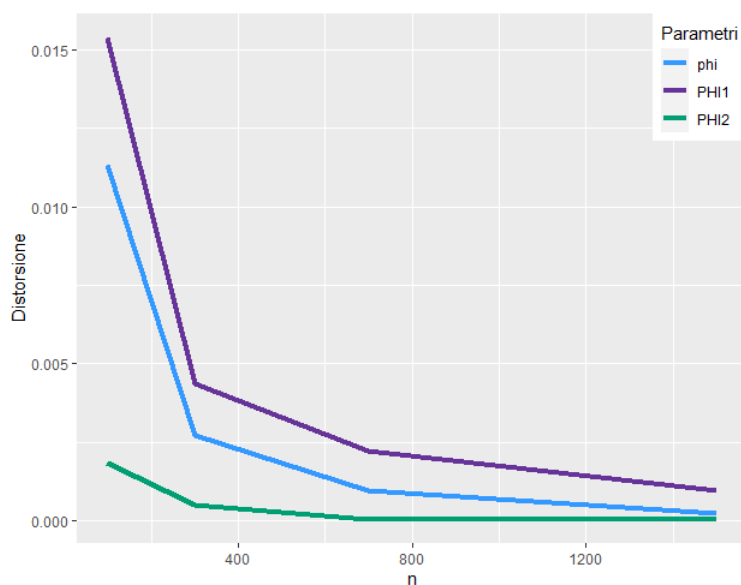


Figura 3.12: Distorsione delle stime per il modello $(1,0)(1,1)_4(1,1)_7\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$

In tutti i grafici la distorsione degli stimatori risulta decrescente al crescere della numerosità campionaria.

Lo stimatore di massima verosimiglianza ha distribuzione asintotica $\sqrt{(n)}(\hat{\theta}_{mle} -$

$\theta_0) \rightarrow N(0, J^{-1})$ e quindi la quantità

$$\frac{\hat{\theta}_{mle} - \theta_0}{\sqrt{\hat{Var}(\hat{\theta}_{mle})}} \sim N(0, 1) \quad (3.5)$$

I grafici seguenti riportano la stima kernel della quantità (3.5) per ogni parametro del modello e la densità di una normale standard.

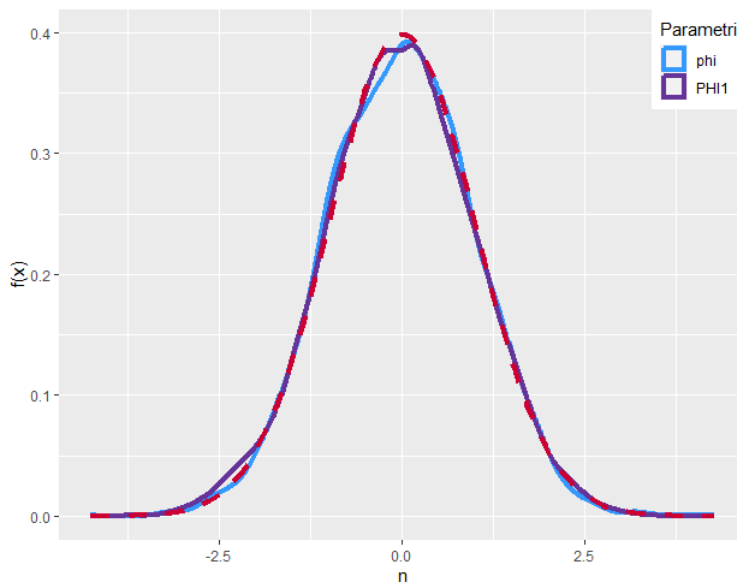


Figura 3.13: Densità empirica stima parametri modello $(1, 0)(1, 0)_4(0, 0)_7 \phi = 0.6, \Phi_1 = 0.5$

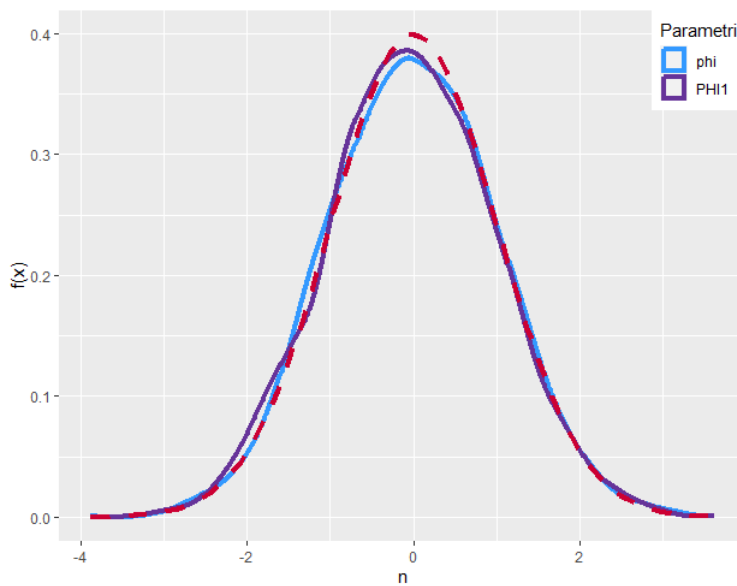


Figura 3.14: Densità empirica stima parametri modello $(1, 1)(1, 1)_4(0, 0)_7 \phi = 0.6, \Phi_1 = 0.5$

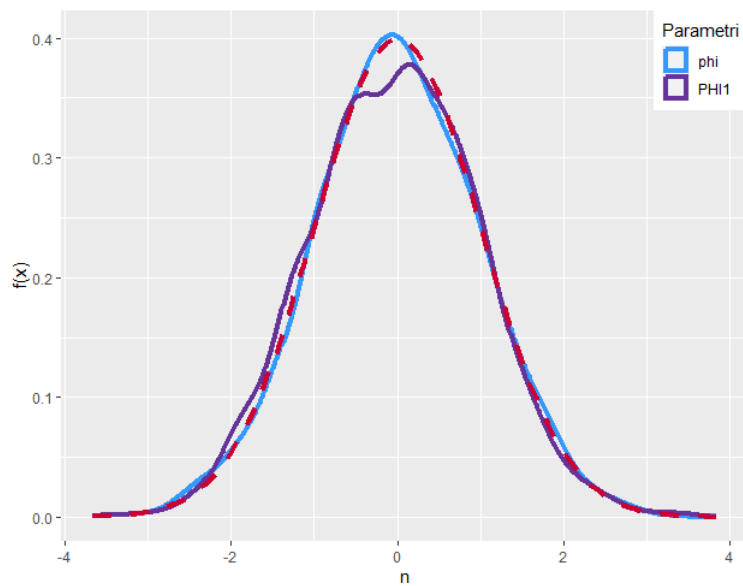


Figura 3.15: Densità empirica stima parametri modello $(1, 0)(1, 0)_4(0, 0)_7\phi = 0.5, \Phi_1 = 0.1$

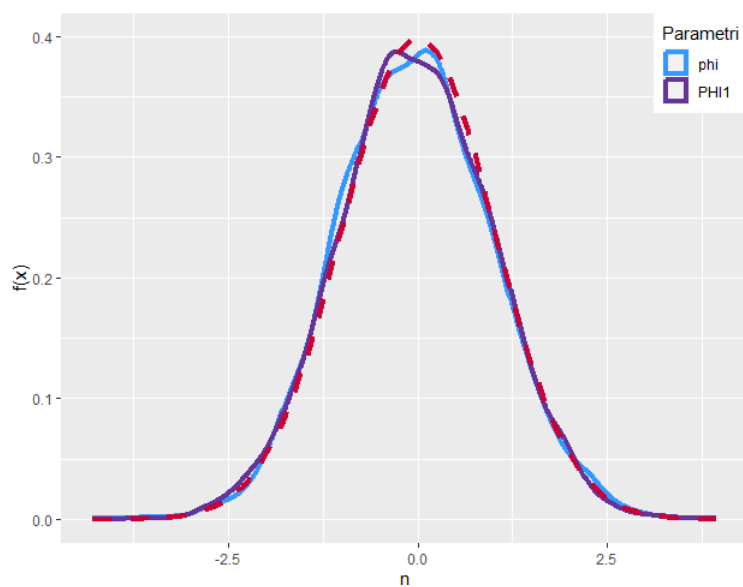


Figura 3.16: Densità empirica stima parametri modello $(1, 1)(1, 1)_4(0, 0)_7\phi = 0.5, \Phi_1 = 0.1$

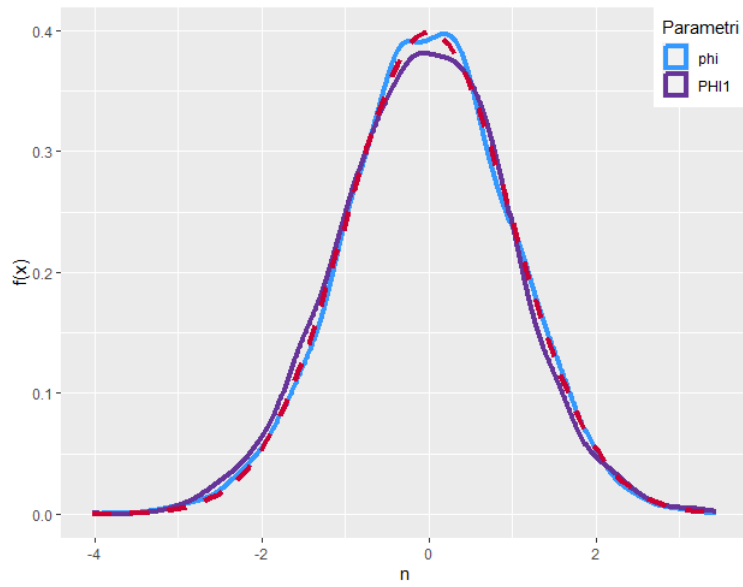


Figura 3.17: Densità empirica stima parametri modello $(1, 0)(1, 0)_4(0, 0)_7\phi = 0.5, \Phi_1 = 0.9$

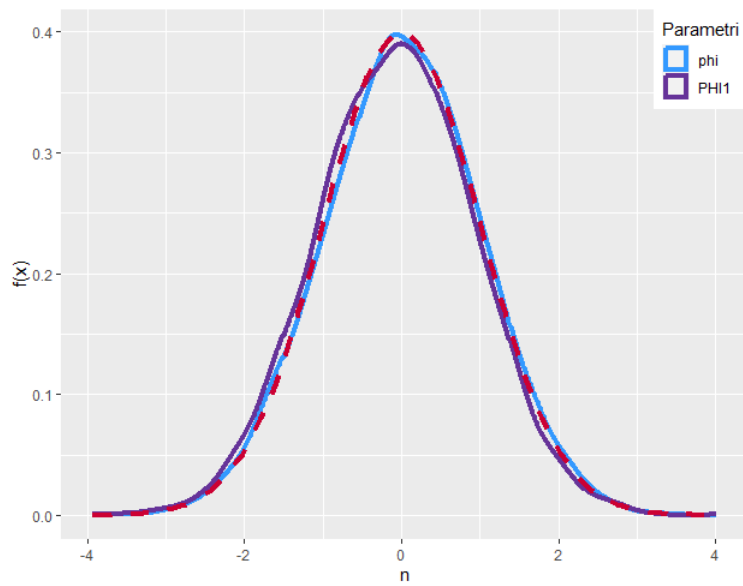


Figura 3.18: Densità empirica stima parametri modello $(1, 1)(1, 1)_4(0, 0)_7\phi = 0.5, \Phi_1 = 0.9$

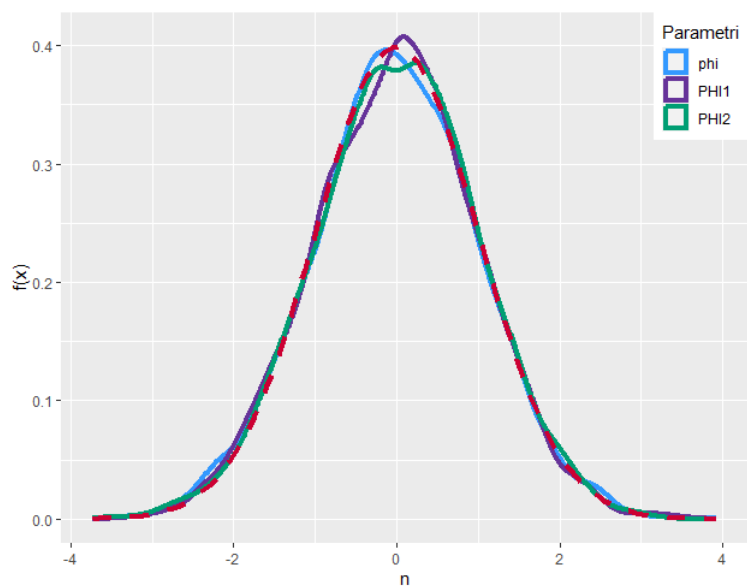


Figura 3.19: Densità empirica stima parametri modello $(1,0)(1,0)_4(1,0)_7\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$

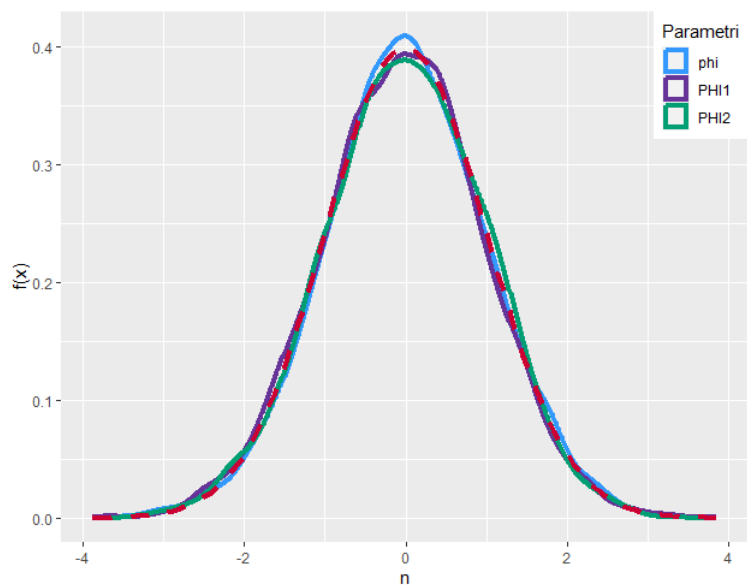


Figura 3.20: Densità empirica stima parametri modello $(1,0)(1,1)_4(1,1)_7\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$

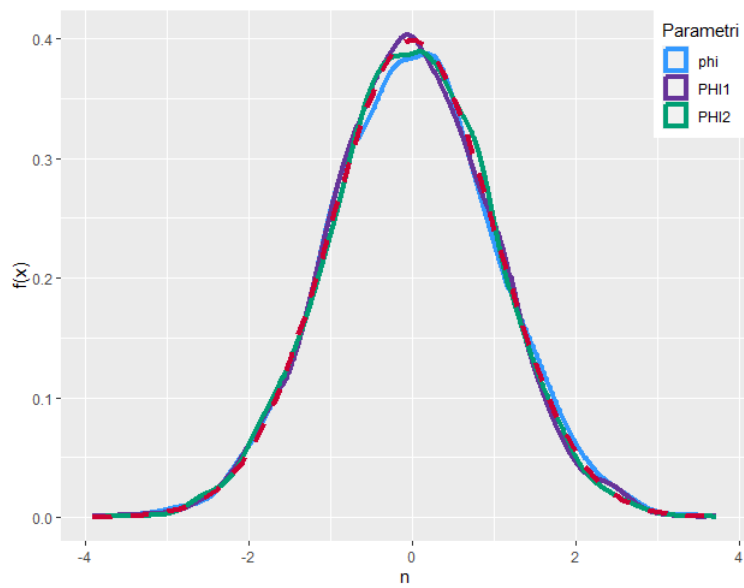


Figura 3.21: Densità empirica stima parametri modello $(1, 0)(1, 0)_4(1, 0)_7\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$

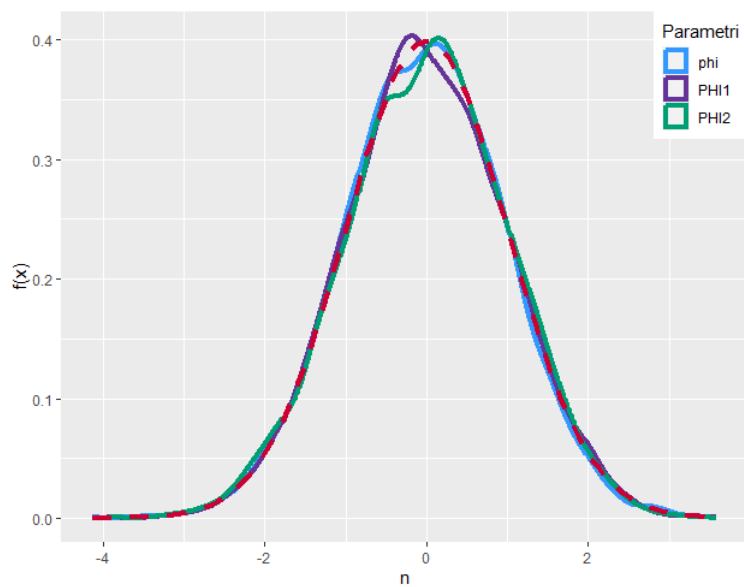


Figura 3.22: Densità empirica stima parametri modello $(1, 0)(1, 1)_4(1, 1)_7\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$

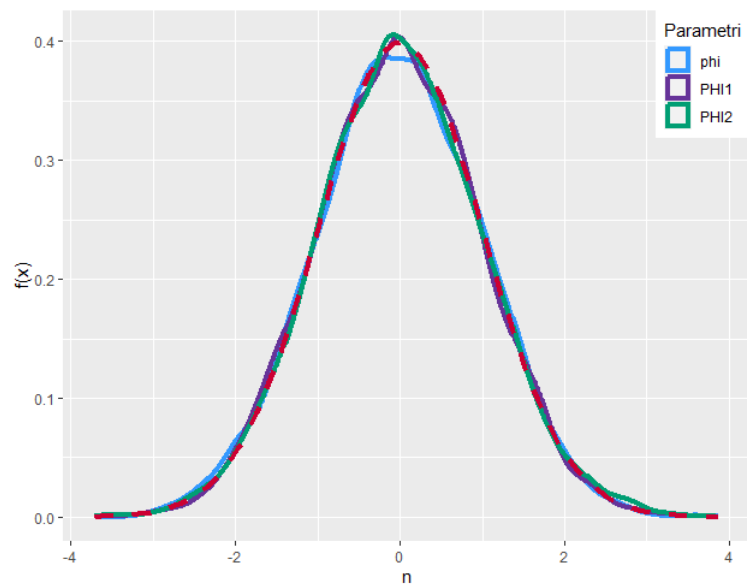


Figura 3.23: Densità empirica stima parametri modello $(1,0)(1,0)_4(1,0)_7$ $\phi = 0.4$, $\Phi_1 = 0.5$, $\Phi_2 = 0.1$

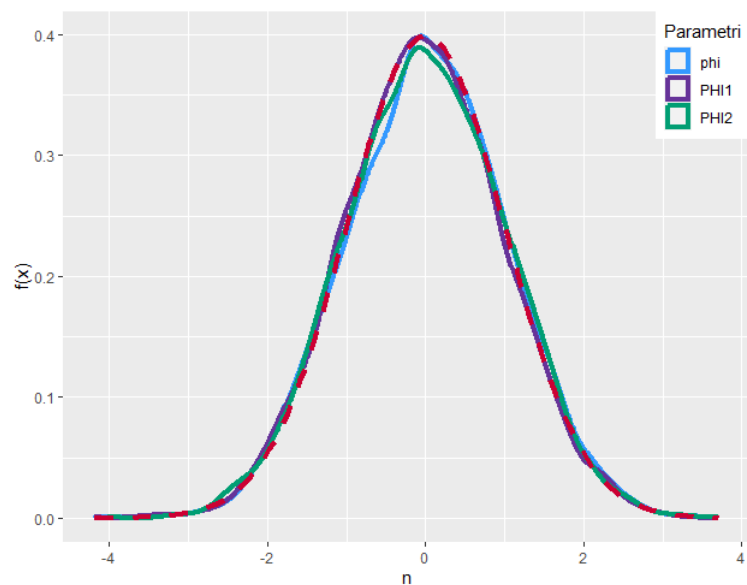


Figura 3.24: Densità empirica stima parametri modello $(1,0)(1,1)_4(1,1)_7$ $\phi = 0.4$, $\Phi_1 = 0.5$, $\Phi_2 = 0.1$

Le stime kernel graficamente sono molto simili ad una normale standard. Per verificare se sussiste evidenza statistica riguardo la normalità degli stimatori di massima verosimiglianza ottenuti dal software è stato calcolato il test di Jarque-Bera sulle stime ottenute dalle simulazioni, al variare della numerosità campionaria.

Il test di Jarque-Bera è un test statistico per verificare l'ipotesi nulla di normalità

$$JB = \frac{n}{6} \left(S^2 + \frac{K - 3}{4} \right)^2 \sim \chi_2 \quad (3.6)$$

dove S è l'indice di asimmetria e K la curtosi.

Nei grafici seguenti vengono riportati i livelli di significatività osservati del test per i vari parametri al crescere della numerosità campionaria. La linea tratteggiata rossa rappresenta il livello di significatività del 5%..

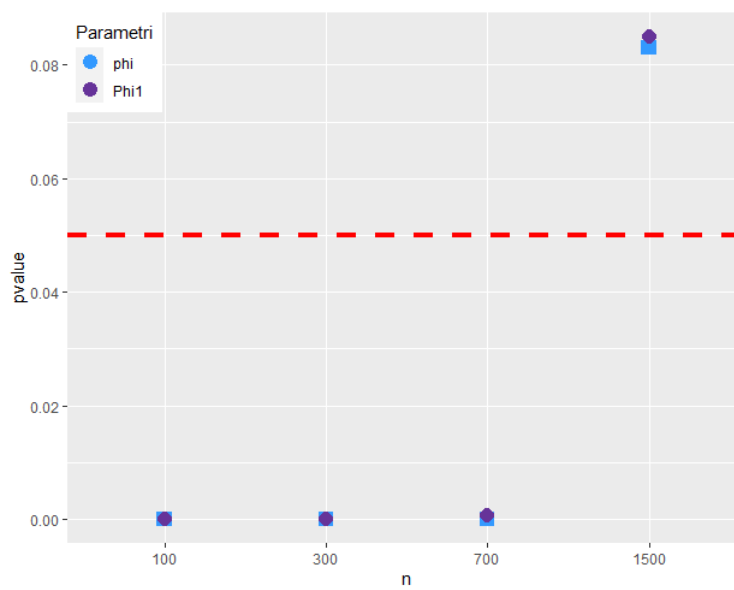


Figura 3.25: P-value test Jargue-Bera stime parametri, modello $(1, 0)(1, 0)_4(0, 0)_7\phi = 0.6, \Phi_1 = 0.5$

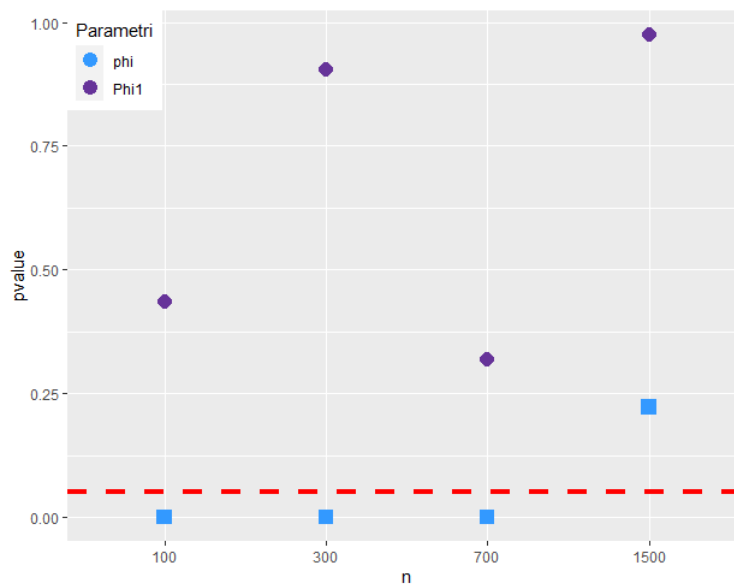


Figura 3.26: P-value test Jargue-Bera stime parametri, modello $(1, 0)(1, 0)_4(0, 0)_7\phi = 0.5, \Phi_1 = 0.1$

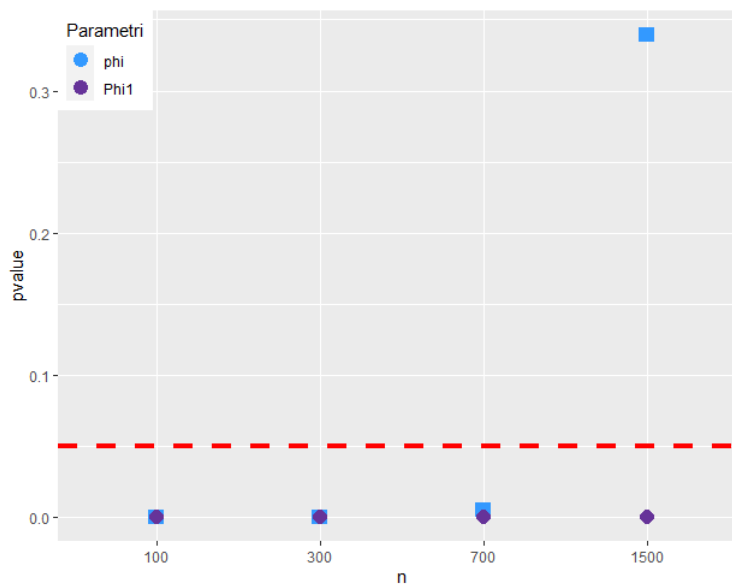


Figura 3.27: P-value test Jargue-Bera stime parametri, modello $(1,0)(1,0)_4(0,0)_7\phi = 0.5, \Phi_1 = 0.9$

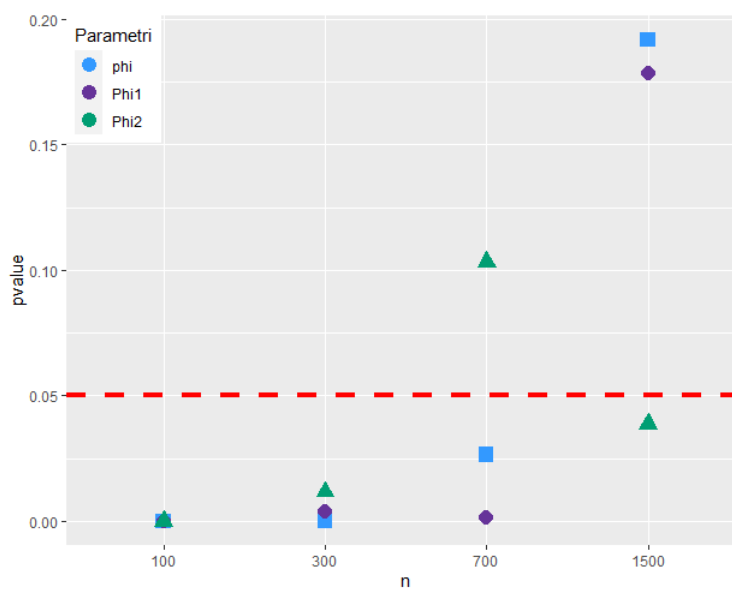


Figura 3.28: P-value test Jargue-Bera stime parametri, modello $(1,0)(1,0)_4(1,0)_7\phi = 0.4, \Phi_1 = 0.4, \Phi_2 = 0.4$

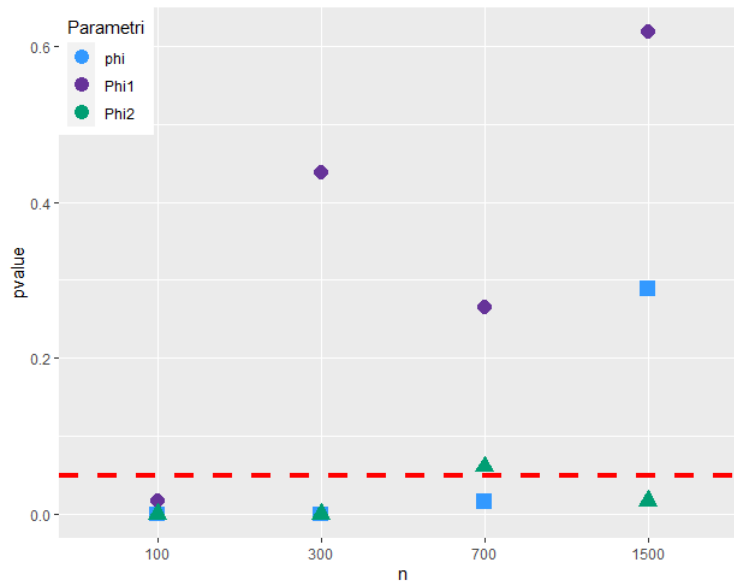


Figura 3.29: P-value test Jargue-Bera stime parametri, modello $(1,0)(1,0)_4(1,0)_7\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$

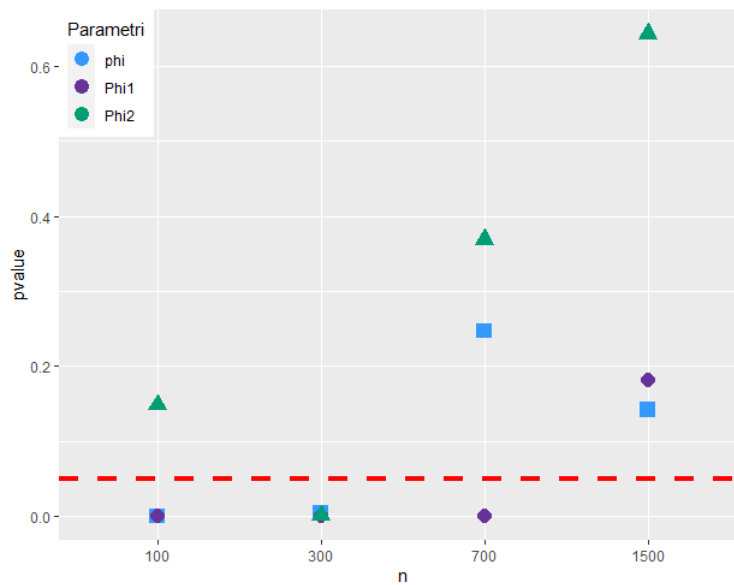


Figura 3.30: P-value test Jargue-Bera stime parametri, modello $(1,0)(1,0)_4(1,0)_7\phi = 0.4, \Phi_1 = 0.5, \Phi_2 = 0.1$

Al crescere della numerosità campionaria l'ipotesi nulla di normalità non viene rifiutata per molti modelli. In appendice sono riportate tutte le tabelle con i livelli di significatività di tutti i test per tutti i modelli.

3.1 Diagnostica

Dopo aver verificato la correttezza della procedura di stima, si prosegue con la diagnostica, per controllare che i residui del modello soddisfino le assunzioni fatte a priori sulle innovazioni.

In particolare, nel contesto della stagionalità e multistagionalità, si vuole verificare che il modello sia stato in grado di cogliere oltre alla dipendenza contemporanea, anche le dipendenze stagionali.

Come primo passo si verifica che il modello sia in grado di cogliere una singola stagionalità.

Si è simulato $M = 1000$ serie da un processo con stagionalità di periodo 4 e per ogni serie si è stimato un SARIMA e un 2SARI e calcolati i test di Ljung-Box (ritardo 20) e Ljung-Box stagionale (ritardo 4). In tabella 3.1 sono riportati i livelli di copertura empirici dei test, e come si può notare, entrambi i modelli raggiungono la copertura teorica.

$2SARI(1, 0)(1, 0)_4(0, 0)_7$		
α	LB	LBS(4)
$SARIMA(1, 0, 0)(1, 0, 0)_4$		
0.10	0.0612	0.0922
0.05	0.0322	0.0470
0.01	0.0096	0.0114
$2SARI$		
0.10	0.0678	0.1010
0.05	0.0366	0.0528
0.01	0.0114	0.0122

Dopo aver verificato che il modello sia in grado di cogliere la stagionalità singola, si vuole appurare la diagnostica multistagionale.

Il test proposto in (2.28) assume che la distribuzione della statistica test abbia distribuzione approssimata $\chi_{m_1+m_2+P_1+P_2}$. Per dimostrarlo si sono simulate $M = 5000$ serie da un processo con stagionalità multipla, si è stimato un 2SARI e calcolato il test Ljung-Box multistagionale. Se la distribuzione approssimata è corretta il livello di copertura empirico (il numero di volte in cui viene rifiutata l'ipotesi nulla) non deve essere superiore al livello di copertura teorico.

In figura 3.31 sono riportati la copertura empirica e quella teorica, si evince che la copertura empirica coincide con quella teorica e quindi l'assunzione sulla distribuzione è verificata.

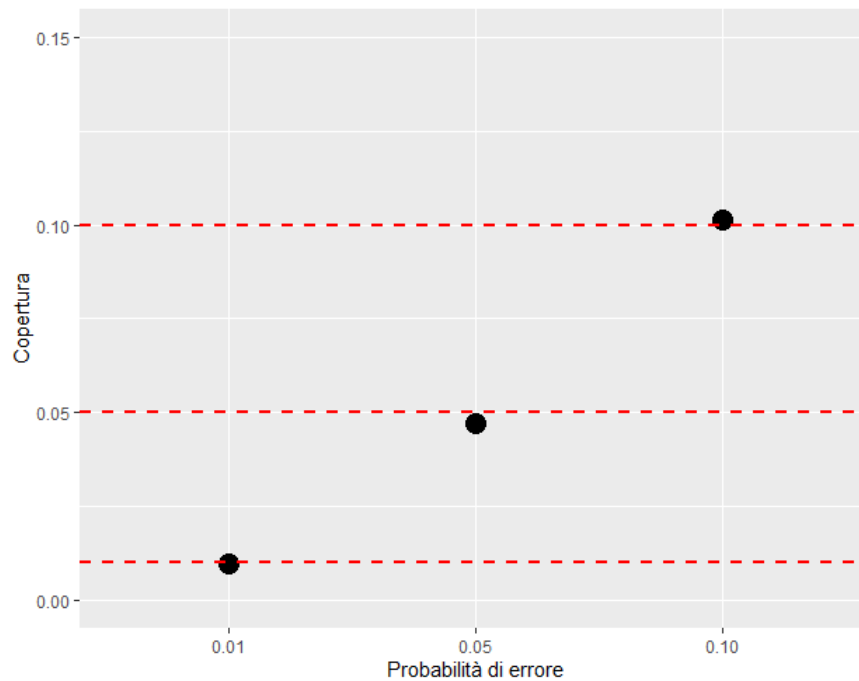


Figura 3.31: Copertura empirica del test Ljung-Box multistagionale

A questo punto, è stato verificato che un modello 2SARI riesce a cogliere la multistagionalità.

Tramite simulazione con processo generatore multistagionale, si è stimato un modello SARIMA e un modello 2SARI e calcolati il test multistagionale sui residui di entrambi i modelli.

In tabella 3.1 è riportata il livello di copertura raggiunto. Il modello SARIMA, non riesce a cogliere la seconda componente stagionale, infatti il livello di copertura dei test ci indica che l'ipotesi nulla di indipendenza dei residui stagionali viene rifiutata frequentemente, mentre per il modello 2SARI il livello di copertura empirico del test è molto vicino a quello teorico.

$2SARI(1,0)(1,0)_4(1,0)_7$		
α	MLBS.Sarima	MLBS.2SARI
0.10	0.6448	0.1012
0.05	0.5426	0.0472
0.01	0.3486	0.0098

3.2 Confronto tra modelli

E' ora necessario confrontare l'adeguatezza dei modelli mSARIMA rispetto ai principali modelli trattati nel capitolo 1.

Nello specifico si vuole confrontare:

- Adeguatezza dei residui
- Errore di previsione

Il confronto dei modelli viene fatto con diverse tipologie di stagionalità multipla: stocastica, deterministica e mista (stocastica e deterministica).

Nella stagionalità stocastica il processo generatore è un $2SARI(1, 0)(1, 0)_4(1, 0)_7$ di parametri $\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$.

Per la stagionalità deterministica, le serie sono generate dal seguente modello

$$Y_t = \sum_{i=1}^{m_1} \gamma_i d_{it} + \sum_{i=1}^{m_2} \tau_i d_{it} + \epsilon_t \quad (3.7)$$

dove d_{it} rappresentano delle variabili dummy per il periodo e γ_i, τ_i rappresentano i livelli stagionali nel periodo i ; $m_1 = 4, m_2 = 7$ sono i periodi stagionali.

Infine, il modello generatore per la stagionalità mista (stocastica e deterministica) è il seguente

$$Y_t = \phi y_{t-1} + \Phi y_{t-j} + \sum_{i=1}^m \gamma_i d_{it} + \epsilon_t \quad (3.8)$$

dove y_{t-j} è il ritardato stagionale.

I modelli confrontati sono i seguenti: 2SARI, TBATS e MSTL.

3.2.1 Analisi dei residui

I residui di un modello devono essere incorrelati, e nel contesto della multistagionalità si vuole che i modelli riescano a spiegare la multistagionalità. Si sono eseguite delle simulazioni e calcolato il livello di copertura dei test Ljung-Box e Ljung-Box multistagionale per ogni tipologia di multistagionalità.

In tutti gli scenari, vengono generate $M = 500$ serie da modelli con la rispettiva multistagionalità, stimato dei modelli e calcolati i test Ljung-Box e Ljung-Box multistagionali per ogni modello. Successivamente è stato calcolato il livello di copertura empirica ottenuta.

Per serie con multistagionalità stocastica si sono ottenuti i seguenti livelli di copertura empirica

α	MSTL	TBATS	2SARI
0.10	1.0000	1.0000	0.0467
0.05	1.0000	0.9967	0.0267
0.01	1.0000	0.9867	0.0067

Tabella 3.1: Copertura test Ljung-Box, stagionalità stocastica

α	MSTL	TBATS	2SARI
0.10	1.0000	1.0000	0.13
0.05	1.0000	1.0000	0.06
0.01	1.0000	0.9933	0.01

Tabella 3.2: Copertura test Ljung-Box multistagionale, stagionalità stocastica

I modelli MSTL e TBATS non riescono a spiegare la correlazione presente nella serie, infatti entrambi i test rifiutano sempre l'ipotesi di incorrelazione dei residui. Mentre il modello 2SARI è in grado di cogliere la correlazione seriale e multistagionale.

Per l'analisi dei residui in un contesto di stagionalità deterministica, si ottengono i seguenti livelli di copertura empirica dei test

α	MSTL	TBATS	2SARI
0.10	1.0000	0.1867	1.0000
0.05	1.0000	0.1233	1.0000
0.01	1.0000	0.0267	1.0000

Tabella 3.3: Copertura test Ljung-Box, stagionalità deterministica

α	MSTL	TBATS	2SARI
0.10	1.0000	0.3533	1.0000
0.05	1.0000	0.2400	1.0000
0.01	1.0000	0.0733	1.0000

Tabella 3.4: Copertura test Ljung-Box multistagionale, stagionalità deterministica

In questo caso il modello che riesce a spiegare meglio le relazioni temporali e stagionali presenti nelle serie è il modello TBATS, anche se i livelli di copertura empirica dei test sono distanti da quelli attesi.

Mentre per la stagionalità mista i livelli di copertura empirica ottenuti sono

i seguenti

α	MSTL	TBATS	2SARI
0.10	1.0000	0.0333	1.0000
0.05	1.0000	0.0267	1.0000
0.01	1.0000	0.0067	1.0000

Tabella 3.5: Copertura test Ljung-Box, stagionalità mista

α	MSTL	TBATS	2SARI
0.10	1.0000	0.1767	1.0000
0.05	1.0000	0.0833	1.0000
0.01	1.0000	0.0167	1.0000

Tabella 3.6: Copertura test Ljung-Box multistagionale, stagionalità mista

Anche in questo caso i residui del modello TBATS soddisfano le assunzioni di base il maggior numero di volte.

3.2.2 Previsione

L'obiettivo principale è come si possono confrontare diversi modelli concorrenti per determinare quale sia il migliore. Nel contesto di previsione lo scopo è quello di valutare la capacità previsiva e misurarne l'accuratezza.

Si è valutata la previsione con diverse tipologie di stagionalità multipla, in particolare stagionalità stocastica, stagionalità deterministica e stagionalità stocastica e deterministica.

Per ogni tipo di stagionalità si sono generate $M = 200$ serie e si sono stimati sui dati in-sample un modello 2SARI, un modello TBATS e il modello MSTL, successivamente è stata fatta previsione con orizzonte temporale $h = 21$ per ogni modello. Infine sono stati calcolati l'errore quadratico medio per ogni passo previsivo e per ogni modello.

In figura 3.32 viene riportato l'RMSE di previsione per un orizzontale temporale di $h = 21$.

Come è possibile osservare la stima di un modello 2SARI ottiene l'errore di previsione minore rispetto il modello TBATS e il modello MSTL.

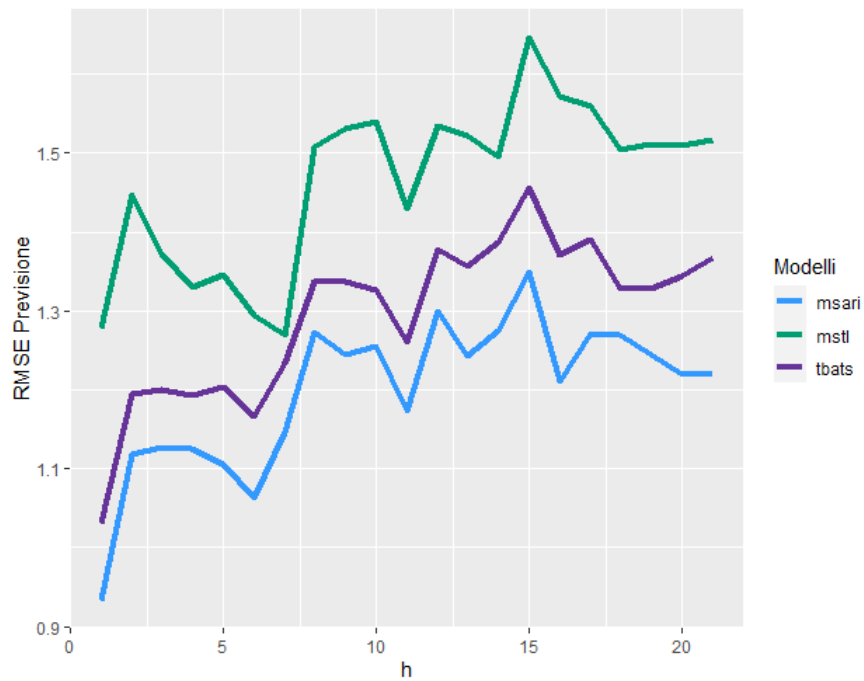


Figura 3.32: RMSE previsione per $h = 1, \dots, 21$
 Stagionalità stocastica, parametri processo generatore $\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$

Questo avviene anche in figura 3.33 dove il processo generatore è sempre a stagionalità stocastica ma con una radice unitaria $2SARI(1, 0)(1, 1)_4(1, 0)_7$ sempre di parametri $\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$.

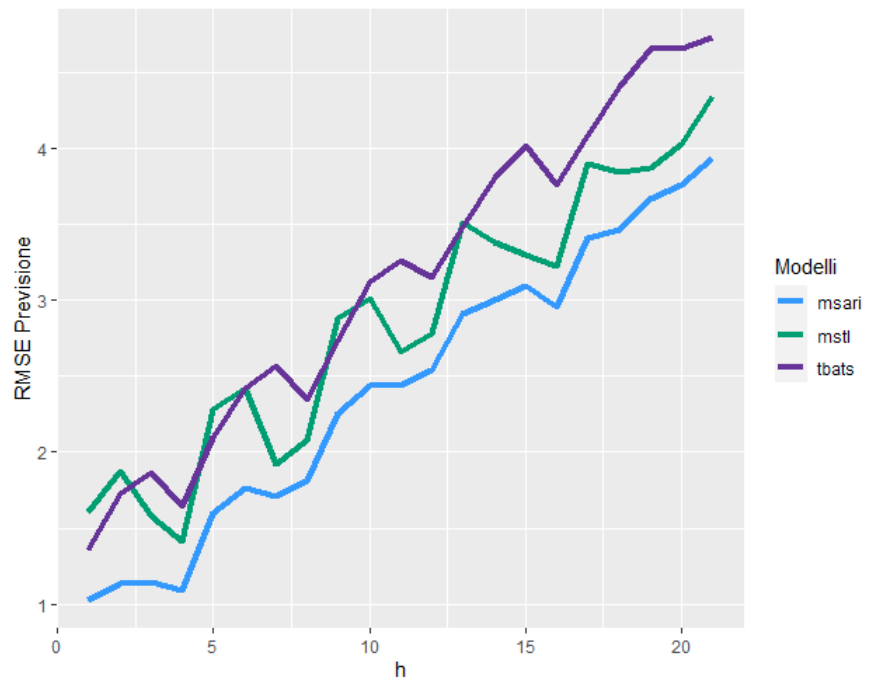


Figura 3.33: RMSE previsione per $h = 1, \dots, 21$
 Stagionalità stocastica con processo integrato periodo 4, parametri processo generatore
 $\phi = 0.4, \Phi_1 = 0.1, \Phi_2 = 0.5$

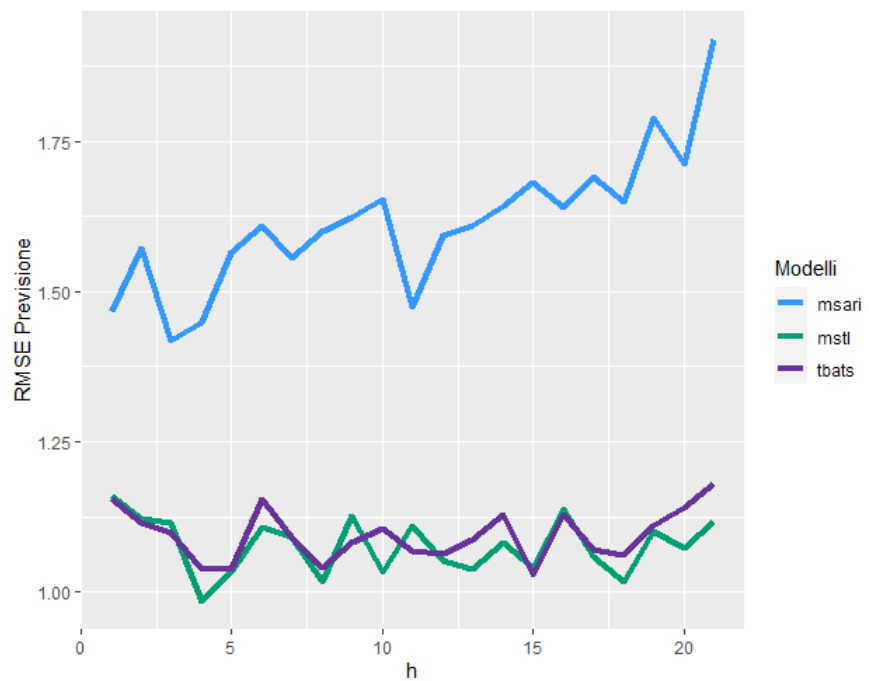


Figura 3.34: RMSE previsione per $h = 1, \dots, 21$
 Stagionalità deterministica

In questo caso, il modello 2SARI ottiene un errore maggiore rispetto ai modelli TBATS e MSTL. Questo è dovuto alla presenza di una componente MA che viene

introdotto dalla differenziazione

$$(1 - B^4)(1 - B^7)Y_t = \epsilon_t - \epsilon_{t-4} - \epsilon_{t-7} + \epsilon_{t-11} \quad (3.9)$$

Il modello 2SARI trattato in questo tesi modella solo la parte autoregressiva, e quindi la componente MA presente nella stagionalità deterministica non viene spiegata.

In figura 3.35 è riportato l'RMSE di previsione con processo generatore con stagionalità deterministica di periodo 7 e stocastica di periodo 4, mentre in figura 3.36 il processo generato ha stagionalità stocastica di periodo 7 e deterministica di periodo 4.

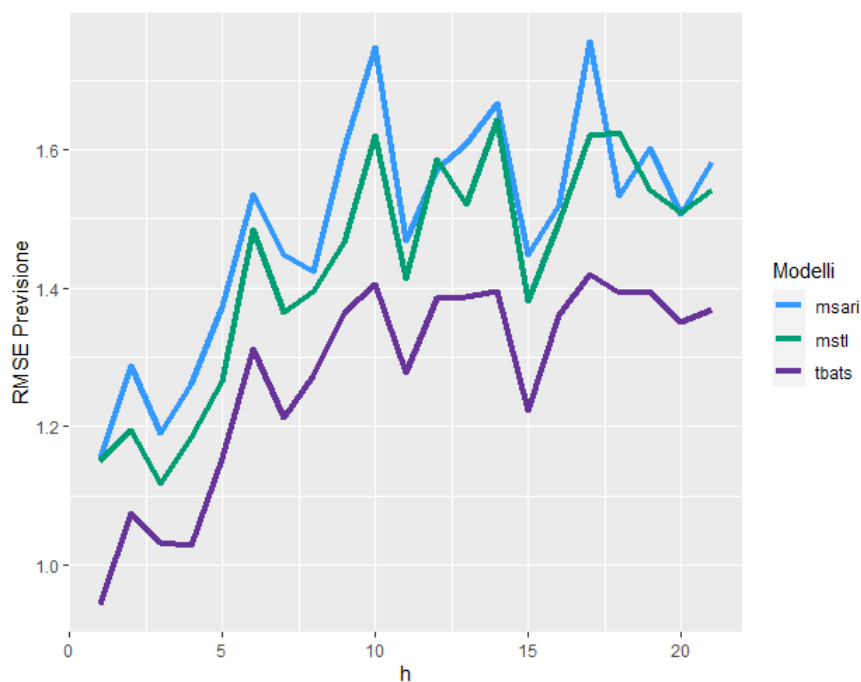


Figura 3.35: RMSE previsione per $h = 1, \dots, 21$
Stagionalità deterministica (periodo 7) e stocastica (periodo 4)

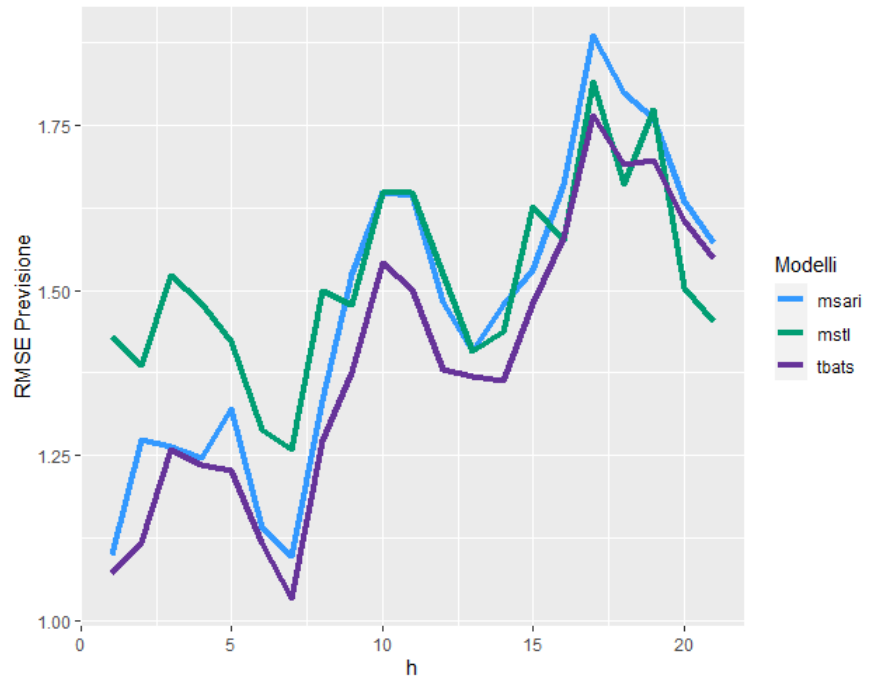


Figura 3.36: RMSE previsione per $h = 1, \dots, 21$
Stagionalità deterministica (periodo 4) e stocastica (periodo 7)

Conclusioni

In questa tesi si sono introdotti i modelli mSARIMA e si è effettuato uno studio di simulazioni sui modelli 2SARI con lo scopo di trattare le serie storiche che contengono stagionalità multipla con un modello parametrico.

In un contesto parametrico si applica la teoria della verosimiglianza con lo scopo di stimare i valori ottimali dei parametri, inoltre si possiede un modo rigoroso ed efficace per valutare la significatività dei parametri stimati.

Si è inoltre sviluppato un test per verificare l'ipotesi congiunta di incorrelazione dei residui ai ritardi multistagionali e si è verificato, tramite simulazione, la distribuzione asintotica del test.

Per quanto riguarda il confronto con i modelli MSTL e TBATS, in riferimento ad un contesto di previsione, il modello 2SARI ottiene i risultati migliori con la stagionalità stocastica e risultati equivalenti agli altri modelli in presenza di stagionalità mista. Se la multistagionalità è deterministica il modello 2SARI ottiene un RMSE di previsione più elevato rispetto ai modelli confrontati, anche se come dimostrato in (3.9) la differenziazione della serie con stagionalità deterministica evidenzia una componente MA la quale non viene trattata dal modello 2SARI, in quanto coglie solo la componente autoregressiva.

Un ulteriore sviluppo potrebbe essere quello di introdurre la componente a media mobile ed effettuare uno studio di simulazione del processo 2SARIMA.

Appendice A

Tabelle

A.1 Stime e varianze campionarie

Di seguito vengono riportare le stime di massima verosimiglianza ottenute delle simulazioni e le relative deviazioni standard.

$(1, 0)(1, 0)_4(0, 0)_7 \quad \phi = 0.6, \Phi_1 = 0.5$				
n	$\hat{\phi}$	$\hat{\Phi}_1$	$se(\hat{\phi})$	$se(\hat{\Phi}_1)$
100	0.5905	0.4823	0.0841	0.0915
300	0.5964	0.4941	0.0474	0.0514
700	0.5987	0.4980	0.0308	0.0334
1500	0.5995	0.4991	0.0210	0.0227

$(1, 1)(1, 0)_4(0, 0)_7 \quad \phi = 0.6, \Phi_1 = 0.5$				
n	$\hat{\phi}$	$\hat{\Phi}_1$	$se(\hat{\phi})$	$se(\hat{\Phi}_1)$
100	0.5892	0.4824	0.0846	0.0920
300	0.5967	0.4941	0.0475	0.0515
700	0.5987	0.4982	0.0308	0.0334
1500	0.5988	0.4988	0.0210	0.0227

$(1, 0)(1, 1)_4(0, 0)_7 \quad \phi = 0.6, \Phi_1 = 0.5$				
n	$\hat{\phi}$	$\hat{\Phi}_1$	$se(\hat{\phi})$	$se(\hat{\Phi}_1)$
100	0.5883	0.4855	0.0861	0.0933
300	0.5960	0.4972	0.0477	0.0516
700	0.5988	0.4977	0.0309	0.0335
1500	0.5992	0.4983	0.0210	0.0228

$(1, 1)(1, 1)_4(0, 0)_7 \quad \phi = 0.6, \Phi_1 = 0.5$				
n	$\hat{\phi}$	$\hat{\Phi}_1$	$se(\hat{\phi})$	$se(\hat{\Phi}_1)$
100	0.5903	0.4833	0.0864	0.0940
300	0.5970	0.4946	0.0478	0.0518
700	0.5983	0.4981	0.0309	0.0335
1500	0.5995	0.4989	0.0210	0.0227

$$(1, 0)(1, 0)_4(0, 0)_7 \quad \phi = 0.5, \Phi_1 = 0.1$$

n	$\hat{\phi}$	$\hat{\Phi}_1$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)
100	0.4897	0.0941	0.0902	0.1031
300	0.4972	0.0967	0.0509	0.0584
700	0.4991	0.0978	0.0331	0.0380
1500	0.4996	0.0991	0.0225	0.0259

$$(1, 1)(1, 0)_4(0, 0)_7 \quad \phi = 0.5, \Phi_1 = 0.1$$

n	$\hat{\phi}$	$\hat{\Phi}_1$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)
100	0.4891	0.0918	0.0906	0.1036
300	0.4961	0.0976	0.0510	0.0585
700	0.4983	0.0991	0.0331	0.0380
1500	0.4999	0.0992	0.0225	0.0259

$$(1, 0)(1, 1)_4(0, 0)_7 \quad \phi = 0.5, \Phi_1 = 0.1$$

n	$\hat{\phi}$	$\hat{\Phi}_1$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)
100	0.4919	0.0924	0.0920	0.1054
300	0.4975	0.0962	0.0512	0.0588
700	0.4980	0.0988	0.0332	0.0381
1500	0.4996	0.0996	0.0226	0.0259

$$(1, 1)(1, 1)_4(0, 0)_7 \quad \phi = 0.5, \Phi_1 = 0.1$$

n	$\hat{\phi}$	$\hat{\Phi}_1$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)
100	0.4888	0.0926	0.0927	0.1060
300	0.4969	0.0968	0.0513	0.0589
700	0.4987	0.0984	0.0332	0.0381
1500	0.4989	0.0991	0.0226	0.0259

$$(1, 0)(1, 0)_4(0, 0)_7 \quad \phi = 0.5, \Phi_1 = 0.9$$

n	$\hat{\phi}$	$\hat{\Phi}_1$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)
100	0.4920	0.8838	0.0901	0.0478
300	0.4959	0.8944	0.0508	0.0260
700	0.4986	0.8978	0.0330	0.0167
1500	0.4995	0.8988	0.0224	0.0113

$$(1, 1)(1, 0)_4(0, 0)_7 \quad \phi = 0.5, \Phi_1 = 0.9$$

n	$\hat{\phi}$	$\hat{\Phi}_1$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)
100	0.4880	0.8845	0.0907	0.0477
300	0.4959	0.8948	0.0508	0.0260
700	0.4982	0.8971	0.0330	0.0168
1500	0.4993	0.8986	0.0225	0.0114

$$(1, 0)(1, 1)_4(0, 0)_7 \quad \phi = 0.5, \Phi_1 = 0.9$$

n	$\hat{\phi}$	$\hat{\Phi}_1$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)
100	0.4899	0.8829	0.0922	0.0489
300	0.4980	0.8938	0.0510	0.0262
700	0.4996	0.8972	0.0330	0.0168
1500	0.4991	0.8984	0.0225	0.0114

$$(1, 1)(1, 1)_4(0, 0)_7 \quad \phi = 0.5, \Phi_1 = 0.9$$

n	$\hat{\phi}$	$\hat{\Phi}_1$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)
100	0.4899	0.8831	0.0927	0.0490
300	0.4977	0.8931	0.0511	0.0264
700	0.4987	0.8970	0.0331	0.0168
1500	0.4999	0.8985	0.0225	0.0114

n	$\hat{\phi}$	$\hat{\Phi}_1$	$\hat{\Phi}_2$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)	se($\hat{\Phi}_2$)
$(1, 0)(1, 0)_4(1, 0)_7$						
100	0.3920	0.3873	0.3927	0.0996	0.0997	0.0992
300	0.3971	0.3955	0.3967	0.0544	0.0544	0.0543
700	0.3984	0.3974	0.4002	0.0351	0.0351	0.0350
1500	0.3998	0.3987	0.3997	0.0238	0.0238	0.0238
$(1, 1)(1, 0)_4(1, 0)_7$						
100	0.3916	0.3882	0.3909	0.1002	0.1002	0.0998
300	0.3986	0.3963	0.3971	0.0545	0.0545	0.0544
700	0.3990	0.3994	0.3986	0.0351	0.0351	0.0351
1500	0.3999	0.3992	0.3991	0.0238	0.0238	0.0238
$(1, 0)(1, 1)_4(1, 0)_7$						
100	0.3932	0.3879	0.3940	0.1019	0.1020	0.1015
300	0.3981	0.3970	0.3974	0.0548	0.0548	0.0547
700	0.3992	0.3982	0.3983	0.0352	0.0352	0.0351
1500	0.3992	0.3992	0.3994	0.0239	0.0239	0.0238
$(1, 0)(1, 0)_4(1, 1)_7$						
100	0.3920	0.3847	0.3905	0.1038	0.1041	0.1034
300	0.3971	0.3944	0.3977	0.0551	0.0551	0.0550
700	0.3996	0.3984	0.3989	0.0352	0.0353	0.0352
1500	0.3998	0.3990	0.4000	0.0239	0.0239	0.0238
$(1, 1)(1, 1)_4(1, 0)_7$						
100	0.3906	0.3882	0.3898	0.1027	0.1028	0.1023
300	0.3963	0.3965	0.3990	0.0549	0.0549	0.0547
700	0.3995	0.3976	0.3998	0.0352	0.0352	0.0351
1500	0.3995	0.3992	0.3998	0.0239	0.0239	0.0238
$(1, 1)(1, 0)_4(1, 1)_7$						
100	0.3917	0.3856	0.3931	0.1045	0.1047	0.1039
300	0.3969	0.3951	0.3959	0.0552	0.0552	0.0551
700	0.3989	0.3987	0.3991	0.0353	0.0353	0.0352
1500	0.3991	0.4002	0.3999	0.0239	0.0239	0.0238
$(1, 0)(1, 1)_4(1, 1)_7$						
100	0.3903	0.3836	0.3903	0.1067	0.1070	0.1062
300	0.3967	0.3954	0.3981	0.0555	0.0555	0.0553
700	0.3980	0.3977	0.3991	0.0354	0.0354	0.0353
1500	0.3999	0.3989	0.3998	0.0239	0.0239	0.0239
$(1, 1)(1, 1)_4(1, 1)_7$						
100	0.3907	0.3892	0.3917	0.1074	0.1075	0.1067
300	0.3977	0.3947	0.3978	0.0556	0.0556	0.0554
700	0.3985	0.3975	0.3997	0.0354	0.0354	0.0353
1500	0.3998	0.3990	0.3991	0.0239	0.0239	0.0239

Tabella A.1: $\phi = 0.4$, $\Phi_1 = 0.4$, $\phi_2 = 0.4$

n	$\hat{\phi}$	$\hat{\Phi}_1$	$\hat{\Phi}_2$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)	se($\hat{\Phi}_2$)
$(1, 0)(1, 0)_4(1, 0)_7$						
100	0.3900	0.0947	0.4892	0.0997	0.1079	0.0941
300	0.3964	0.0976	0.4964	0.0544	0.0590	0.0514
700	0.3994	0.0985	0.4991	0.0351	0.0381	0.0331
1500	0.3998	0.0994	0.4995	0.0238	0.0259	0.0225
$(1, 1)(1, 0)_4(1, 0)_7$						
100	0.3908	0.0919	0.4888	0.1003	0.1085	0.0946
300	0.3995	0.0985	0.4943	0.0545	0.0591	0.0515
700	0.3983	0.1000	0.4984	0.0351	0.0381	0.0331
1500	0.4001	0.0992	0.4988	0.0238	0.0259	0.0225
$(1, 0)(1, 1)_4(1, 0)_7$						
100	0.3951	0.0941	0.4884	0.1017	0.1104	0.0963
300	0.3969	0.0986	0.4972	0.0548	0.0594	0.0517
700	0.3995	0.0980	0.4982	0.0352	0.0382	0.0332
1500	0.3997	0.0995	0.4991	0.0239	0.0259	0.0225
$(1, 0)(1, 0)_4(1, 1)_7$						
100	0.3906	0.0923	0.4884	0.1040	0.1125	0.0981
300	0.3977	0.0980	0.4963	0.0551	0.0598	0.0520
700	0.3988	0.0985	0.4983	0.0353	0.0383	0.0333
1500	0.3992	0.0993	0.4996	0.0239	0.0259	0.0225
$(1, 1)(1, 1)_4(1, 0)_7$						
100	0.3908	0.0931	0.4893	0.1028	0.1112	0.0969
300	0.3982	0.0971	0.4951	0.0549	0.0596	0.0519
700	0.3992	0.0992	0.4983	0.0352	0.0382	0.0332
1500	0.4000	0.0999	0.4993	0.0239	0.0259	0.0225
$(1, 1)(1, 0)_4(1, 1)_7$						
100	0.3916	0.0920	0.4884	0.1046	0.1132	0.0987
300	0.3971	0.0979	0.4962	0.0552	0.0599	0.0521
700	0.3986	0.1002	0.4989	0.0353	0.0383	0.0333
1500	0.4008	0.0998	0.4992	0.0239	0.0259	0.0225
$(1, 0)(1, 1)_4(1, 1)_7$						
100	0.3887	0.0910	0.4878	0.1068	0.1155	0.1008
300	0.3979	0.0974	0.4970	0.0555	0.0602	0.0524
700	0.3984	0.0987	0.4986	0.0354	0.0384	0.0334
1500	0.3992	0.0999	0.4998	0.0239	0.0260	0.0226
$(1, 1)(1, 1)_4(1, 1)_7$						
100	0.3898	0.0929	0.4893	0.1075	0.1162	0.1014
300	0.3973	0.0981	0.4970	0.0556	0.0603	0.0525
700	0.3985	0.0981	0.4992	0.0354	0.0384	0.0334
1500	0.3996	0.0996	0.4991	0.0239	0.0260	0.0226

Tabella A.2: $\phi = 0.4$, $\Phi_1 = 0.1$, $\phi_2 = 0.5$

n	$\hat{\phi}$	$\hat{\Phi}_1$	$\hat{\Phi}_2$	se($\hat{\phi}$)	se($\hat{\Phi}_1$)	se($\hat{\Phi}_2$)
$(1, 0)(1, 0)_4(1, 0)_7$						
100	0.3958	0.4842	0.0959	0.0994	0.0945	0.1074
300	0.3981	0.4943	0.1009	0.0544	0.0515	0.0589
700	0.3987	0.4981	0.0998	0.0351	0.0332	0.0380
1500	0.3995	0.4995	0.1000	0.0238	0.0225	0.0258
$(1, 1)(1, 0)_4(1, 0)_7$						
100	0.3897	0.4869	0.0977	0.1003	0.0950	0.1080
300	0.3970	0.4965	0.0990	0.0545	0.0515	0.0590
700	0.3997	0.4976	0.0995	0.0351	0.0332	0.0380
1500	0.3989	0.4993	0.0995	0.0238	0.0225	0.0258
$(1, 0)(1, 1)_4(1, 0)_7$						
100	0.3910	0.4856	0.0996	0.1020	0.0967	0.1098
300	0.3978	0.4951	0.0992	0.0548	0.0518	0.0593
700	0.3986	0.4972	0.0995	0.0352	0.0333	0.0381
1500	0.3998	0.4989	0.1000	0.0238	0.0225	0.0258
$(1, 0)(1, 0)_4(1, 1)_7$						
100	0.3927	0.4870	0.0975	0.1038	0.0985	0.1119
300	0.3964	0.4961	0.0997	0.0551	0.0521	0.0596
700	0.3992	0.4979	0.0992	0.0352	0.0333	0.0382
1500	0.3995	0.4998	0.0997	0.0239	0.0226	0.0259
$(1, 1)(1, 1)_4(1, 0)_7$						
100	0.3934	0.4838	0.0984	0.1025	0.0974	0.1106
300	0.3968	0.4966	0.0997	0.0549	0.0519	0.0594
700	0.3993	0.4975	0.0995	0.0352	0.0333	0.0381
1500	0.3996	0.4994	0.0994	0.0239	0.0225	0.0259
$(1, 1)(1, 0)_4(1, 1)_7$						
100	0.3921	0.4881	0.0977	0.1045	0.0990	0.1126
300	0.3974	0.4944	0.0997	0.0552	0.0522	0.0597
700	0.3988	0.4985	0.0989	0.0353	0.0333	0.0382
1500	0.3991	0.4989	0.0993	0.0239	0.0226	0.0259
$(1, 0)(1, 1)_4(1, 1)_7$						
100	0.3929	0.4838	0.0990	0.1066	0.1012	0.1148
300	0.3968	0.4950	0.0994	0.0555	0.0525	0.0600
700	0.3981	0.4978	0.0988	0.0354	0.0334	0.0383
1500	0.3995	0.4993	0.1000	0.0239	0.0226	0.0259
$(1, 1)(1, 1)_4(1, 1)_7$						
100	0.3897	0.4850	0.0980	0.1073	0.1017	0.1156
300	0.3976	0.4945	0.1010	0.0555	0.0526	0.0601
700	0.3996	0.4970	0.0991	0.0354	0.0335	0.0383
1500	0.3998	0.4992	0.1000	0.0239	0.0226	0.0259

A.2 Test Jarque-Bera distribuzione stimatori

Di seguito vengono riportati i test Jarque-Bera per i vari modelli al variare della lunghezza della serie.

n	$\hat{\phi}$	$\hat{\Phi}_1$
$(1, 0)(1, 0)_4(0, 0)_7$		
300	0.0000	0.0000
700	0.0000	0.0035
1500	0.0213	0.0007
$(1, 1)(1, 0)_4(0, 0)_7$		
300	0.0000	0.0000
700	0.0055	0.0220
1500	0.0000	0.0708
$(1, 0)(1, 1)_4(0, 0)_7$		
300	0.0000	0.0000
700	0.0010	0.0004
1500	0.0026	0.0156
$(1, 1)(1, 1)_4(0, 0)_7$		
300	0.0000	0.0000
700	0.0000	0.0007
1500	0.0831	0.0850

Tabella A.4: Test Jarque-Bera (pvalue) per $\phi = 0.6, \Phi_1 = 0.5$

n	$\hat{\phi}$	$\hat{\Phi}_1$
$(1, 0)(1, 0)_4(0, 0)_7$		
300	0.0011	0.6192
700	0.0128	0.0422
1500	0.0263	0.7233
$(1, 1)(1, 0)_4(0, 0)_7$		
300	0.0000	0.2873
700	0.1859	0.3718
1500	0.0271	0.4191
$(1, 0)(1, 1)_4(0, 0)_7$		
300	0.0001	0.6233
700	0.0050	0.1969
1500	0.3702	0.1510
$(1, 1)(1, 1)_4(0, 0)_7$		
300	0.0000	0.9054
700	0.0000	0.3195
1500	0.2222	0.9752

Tabella A.5: Test Jarque-Bera (pvalue) per $\phi = 0.5, \Phi_1 = 0.1$

n	$\hat{\phi}$	$\hat{\Phi}_1$
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{0})_4(\mathbf{0}, \mathbf{0})_7$		
300	0.0000	0.0000
700	0.0000	0.0000
1500	0.0000	0.0000
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{0})_4(\mathbf{0}, \mathbf{0})_7$		
300	0.0000	0.0000
700	0.0000	0.0000
1500	0.0000	0.0000
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{1})_4(\mathbf{0}, \mathbf{0})_7$		
300	0.0000	0.0000
700	0.0050	0.0000
1500	0.3395	0.0000
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{1})_4(\mathbf{0}, \mathbf{0})_7$		
300	0.0000	0.0000
700	0.0000	0.0000
1500	0.0025	0.0000

Tabella A.6: Test Jarque-Bera (pvalue) per $\phi = 0.5$, $\Phi_1 = 0.9$

Modello	ϕ		Φ_1		Φ_2	
	300	1500	300	1500	300	1500
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{0})_7$	0.0005	0.0478	0.0000	0.1327	0.0490	0.0004
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{0})_7$	0.0000	0.1919	0.0037	0.1786	0.0117	0.0388
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{0})_7$	0.0003	0.0012	0.0006	0.1103	0.0002	0.1977
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{1})_7$	0.0021	0.7057	0.0014	0.1112	0.0000	0.0361
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{0})_7$	0.0000	0.0217	0.0003	0.2296	0.4493	0.1408
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{1})_7$	0.0076	0.0225	0.0000	0.1141	0.0001	0.1631
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{1})_7$	0.0000	0.0053	0.0009	0.0104	0.0017	0.0002
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{1})_7$	0.0000	0.9802	0.0002	0.0510	0.0138	0.3073

Tabella A.7: Test Jarque-Bera (pvalue) per $\phi = 0.4$, $\Phi_1 = 0.4$, $\Phi_2 = 0.4$

Modello	ϕ		Φ_1		Φ_2	
	300	1500	300	1500	300	1500
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{0})_7$	0.0014	0.5376	0.0529	0.7873	0.0009	0.0084
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{0})_7$	0.0004	0.1923	0.0726	0.0490	0.0008	0.0008
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{0})_7$	0.0166	0.0052	0.9101	0.0622	0.0000	0.0143
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{1})_7$	0.0000	0.2891	0.4375	0.6187	0.0612	0.0170
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{0})_7$	0.0002	0.8218	0.2247	0.0805	0.0003	0.0026
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{1})_7$	0.1259	0.1092	0.3952	0.6410	0.0007	0.0036
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{1})_7$	0.0042	0.0541	0.7617	0.4327	0.0000	0.0000
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{1})_7$	0.0030	0.5578	0.6398	0.9592	0.0006	0.0000

Tabella A.8: Test Jarque-Bera (pvalue) per $\phi = 0.4$, $\Phi_1 = 0.1$, $\Phi_2 = 0.5$

Modello	ϕ		Φ_1		Φ_2	
	300	1500	300	1500	300	1500
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{0})_7$	0.0000	0.0774	0.000	0.2621	0.4078	0.2284
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{0})_7$	0.0001	0.0713	0.000	0.2815	0.5427	0.3821
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{0})_7$	0.0000	0.0192	0.000	0.2100	0.2904	0.7408
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{1})_7$	0.0043	0.1417	0.000	0.1819	0.3687	0.6433
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{0})_7$	0.0008	0.1976	0.000	0.0267	0.8669	0.8467
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{0})_4(\mathbf{1}, \mathbf{1})_7$	0.0000	0.1412	0.000	0.0781	0.9111	0.96296
$(\mathbf{1}, \mathbf{0})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{1})_7$	0.0002	0.0422	0.000	0.0397	0.4265	0.2219
$(\mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{1})_4(\mathbf{1}, \mathbf{1})_7$	0.0002	0.3590	0.000	0.0171	0.0536	0.5066

Tabella A.9: Test Jarque-Bera (pvalue) per $\phi = 0.4$, $\Phi_1 = 0.5$, $\Phi_2 = 0.1$

Bibliografia

- [1] Box, G. E. and P., Pierce, D. A. (1970) 'Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models', *Journal of the American Statistical Association*, 65(332), pp. 1509-1526.
- [2] Ljung, G. M and Box, G E. P. (1978) 'On a Measure of Lack of Fit in Time Series Models', *Biometrika*, 65(2), pp. 297-303.
- [3] McLeod, A. I. (1978), 'On the Distribution of Residual Autocorrelations in Box-Jenkins Models', *Journal of the Royal Statistical Society. Series B (Methodological)*, 40(3), pp 296-302.
- [4] McLeod, A. I. (1979), 'Distribution of the Residual Cross-Correlation in Univariate ARMA Time Series Models', *Journal of the American Statistical Association*, 74(368), pp-849-855.
- [5] De Livera A. M. and Hyndman R. J. and Snyder R. D. (2011), 'Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing', *Journal of the American Statistical Association*, 106(496), pp 1513-1527.
- [6] Di Fonzo T. and Lisi F. (2012) *Serie storiche economiche, Analisi statistiche e applicazioni*. 1rd edn. Carocci editore.
- [7] pratha19 *Hourly Energy Consumption Prediction Available at https://github.com/pratha19/Hourly_Energy_Consumption_Prediction* (Data ultima visita: 08/10/2023)