

UNIVERSITÀ DEGLI STUDI DI PAVIA
DIPARTIMENTO DI MATEMATICA
CORSO DI LAUREA MAGISTRALE IN MATEMATICA



UNIVERSITÀ
DI PAVIA

**ANALISI DELLE SERIE STORICHE PER LE QUOTE
DI CARBONIO DELL'EMISSIONS TRADING SYSTEM
EUROPEO**

Tesi di Laurea Magistrale in Matematica

Relatore:

Prof.ssa Raffaella Carbone

Correlatore:

Prof. Andrea Martinelli

Tesi di Laurea di:

Anna Guarneri

Matricola 504517

ANNO ACCADEMICO 2022 - 2023

Abstract

Time series analysis is a highly significant field in statistics as it allows us to examine the historical evolution of a phenomenon over time. The purpose of this analysis is to identify a stochastic process that provides a suitable model for the observations, which can also be used for predictive purposes.

In the course of the paper, the details of various classes of processes, both stationary and non-stationary, are presented. In particular, the ARMA processes are introduced, which are used for stationary time series and consist of an autoregressive component (AR) and a moving average component (MA). Subsequently, several classes of processes are introduced for handling heteroskedastic time series. These include the ARCH, GARCH, and EGARCH processes, which provide suitable models for volatility.

These processes are then applied to assess and analyze the volatility of daily prices of European Carbon Allowances under the European Union Emissions Trading System (EU ETS), with a focus on Phase 4 (2021-2023). The EU ETS is a system introduced in Europe with the aim of achieving decarbonization goals, to which major emitter categories must adhere.

Different models are estimated and the joint ARMA(5,1)-GARCH(1,1) model for the timeserie without the trend component, exhibits a superior ability to describe the price volatility. In this model, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are lower than other models, and all its coefficients are significant ($p < 0.05$). Moreover, the standardized residuals are uncorrelated and their distribution is adequate.

Indice

Introduzione	v
1 Introduzione dell'Emissions Trading System europeo	1
1.1 Cambiamento climatico	1
1.1.1 Il protocollo di Kyoto	2
1.1.2 L'Accordo di Parigi	3
1.1.3 Conferenze delle Parti 26 e 27	3
1.1.4 Green Deal Europeo	4
1.2 Sistemi di Carbon Pricing	4
1.2.1 European Union Emissions Trading System (EU ETS)	5
2 Analisi di serie storiche	9
2.1 Probabilità e variabili aleatorie	9
2.2 Processi stazionari e funzioni caratterizzanti	13
2.3 Serie storiche	16
2.4 Analisi di serie storiche	19
2.4.1 Decomposizione di serie storiche	19
2.4.2 Identificazione parametrica	26
2.4.3 Criteri di scelta del modello	27
2.4.4 Verifica del modello	28
3 Modelli autoregressivi e a media mobile	29
3.1 White noise	29
3.2 Processi a media mobile (MA)	30
3.3 Processi autoregressivi (AR)	31
3.4 Processi autoregressivi a media mobile (ARMA)	33
3.4.1 Causalità e invertibilità	35
3.4.2 Metodi per calcolo della funzione di autocovarianza	36
3.5 Generalizzazioni dei processi ARMA	41
4 Modelli a volatilità stocastica	43
4.1 Processi ARCH	44

4.2	Processi GARCH	49
4.3	Processi EGARCH	53
5	Studio della serie storica dell'EU ETS	55
5.1	Analisi qualitativa	55
5.2	Identificazione dei modelli	57
5.2.1	Considerazioni iniziali	58
5.2.2	Metodo 1	61
5.2.3	Metodo 2	70
5.2.4	Metodo 3	70
5.2.5	Conclusioni	79
	Acronimi	83
	Bibliografia	84

Introduzione

L'idea di questa tesi nasce a seguito della mia partecipazione al progetto LM+, esperienza che consente agli studenti di svolgere un periodo in azienda per integrare le proprie conoscenze con il mondo del lavoro. Tale progetto mi ha dato l'opportunità di svolgere un tirocinio in Eni S.p.A., in particolare nel dipartimento di Monitoraggio e Asset Valorization della direzione CCUSFA, acronimo di Carbon Capture Utilization & Storage, Forestry & Agros-Feedstock.

Argomento centrante dell'elaborato è l'analisi delle serie storiche, con la presentazione di alcuni processi stocastici che possono essere utilizzati come modelli per i dati. L'obiettivo finale è poi quello di utilizzare i diversi modelli introdotti per analizzare la serie storica dei prezzi spot dell'Emissions Trading System europeo, un sistema che regola le emissioni di CO₂ e che influenza particolarmente i progetti sviluppati dall'unità in cui ho svolto il tirocinio.

Nello specifico vengono svolte diverse attività, con l'obiettivo comune di portare un contributo nella transizione energetica della società, al fine di raggiungere l'obiettivo di neutralità carbonica entro il 2050. Con tale intento si stanno sviluppando diverse progettualità.

- Progetti innovativi e ambiziosi volti alla cattura e allo stoccaggio dell'anidride carbonica, sfruttando giacimenti di gas esauriti in cui stoccare la CO₂ in modo permanente;
- Progetti che mirano alla compensazione delle emissioni e alla generazione di crediti carbonici. Queste attività si basano sul presupposto che, non essendo possibile allo stato attuale eliminare completamente le emissioni di carbonio della società, l'obiettivo diventa quello di compensarle finanziando progetti che riducono o rimuovono una quantità equivalente di gas serra dall'atmosfera;
- Progetti di agrobusiness, volti alla produzione di oli vegetali dalla spremitura di semi oleosi, da fornire poi alle bioraffinerie come biocombustibili.

Tutti questi progetti volti alla riduzione delle emissioni, sono influenzati dal mercato del carbonio, i primi in modo diretto mentre gli altri in modo indiretto.

Il prezzo del carbonio è uno degli strumenti fondamentali e driver delle politiche adottate dalle imprese per raggiungere gli obiettivi di decarbonizzazione.

Differenti sono le strategie ed i sistemi adottati dai vari Paesi per raggiungere gli obiettivi di riduzione delle emissioni e, come anticipato, quello che ho deciso di analizzare in questo elaborato è l'European Union Emissions Trading System (EU ETS). Si tratta del primo sistema internazionale, gestito dall'Unione Europea, per lo scambio di quote di emissione in cui viene fissato un numero massimo di quote che le diverse società possono acquistare o scambiare.

Nel primo capitolo viene presentato il percorso che ha portato all'introduzione di questo sistema, il suo funzionamento e le successive fasi evolutive con le rispettive caratteristiche.

Al fine di analizzare la serie storica dei prezzi spot dell'EU ETS, nel secondo capitolo viene presentato un quadro introduttivo sull'analisi di serie storiche, che ha come scopo quello di individuare un processo stocastico che fornisca un modello opportuno per i dati da utilizzare anche a fine predittivo. Si focalizza l'attenzione sui processi stazionari e si analizzano le definizioni delle funzioni di autocorrelazione e autocovarianza con le rispettive proprietà.

Con l'ausilio del modello di decomposizione classica di un processo $X = (X_t)_{t \in \mathbb{Z}}$ per il quale

$$X_t = m_t + s_t + Y_t, \quad t \in \mathbb{Z},$$

viene mostrato come eliminare eventuali fenomeni di trend (m_t) e stagionalità (s_t) presenti nelle serie storiche, con l'obiettivo di ottenere un processo stazionario (Y_t). Vengono poi presentati diversi criteri di scelta per valutare il modello più opportuno tra i vari candidati.

Il terzo capitolo affronta l'analisi dei processi a media mobile (MA(q)) e autoregressivi (AR(p)) per giungere alla definizione della classe dei processi ARMA(p,q), che comprende sia i processi AR che i processi MA come caso particolare. Come mostrato nell'equazione seguente, il processo ARMA, indicato con $X = (X_t)_{t \in \mathbb{Z}}$, viene visto come somma di una parte a media mobile ed una parte autoregressiva:

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q}, \quad t \in \mathbb{Z},$$

dove $(Z_t)_{t \in \mathbb{Z}}$ è un *white noise*, ovvero una famiglia di variabili aleatorie a media nulla, varianza costante e incorrelate l'una all'altra.

Dopo aver presentato le proprietà principali, i diversi metodi per il calcolo della funzione di autocovarianza e le nozioni di invertibilità e causalità, si analizza come caratterizzare un processo ARMA(1,1). Viene poi presentato un cenno sulle possibili generalizzazioni dei processi ARMA, in particolare i processi ARIMA e SARIMA utili qualora la serie storica in esame non mostri un comportamento stazionario.

Il quarto capitolo è dedicato alla descrizione di classi di processi che forniscono modelli per le serie storiche eteroschedastiche, ossia per cui la varianza non è costante. In questi

casi, indicato con $Y = (Y_t)_{t \in \mathbb{Z}}$ il processo stocastico, ciascuna variabile aleatoria Y_t è ottenuta come prodotto di un white noise $\{e_t\}$ e della radice della varianza condizionale σ_t^2 :

$$Y_t = e_t \sigma_t,$$

dove σ_t^2 è una v.a. con diverse forme di dipendenza dal passato del processo a seconda della famiglia di modelli prescelta.

Precisamente vediamo i processi ARCH, GARCH ed EGARCH, per ciascuno dei quali viene definita una diversa espressione per la varianza condizionale. Questi modelli sono in grado di cogliere in parte o totalmente le caratteristiche peculiari delle serie storiche finanziarie, quali la leptocurtosi, l'effetto leverage, la persistenza della volatilità, l'aggregational gaussianity e la volatility clustering. Così come nel capitolo precedente, vengono analizzate le caratteristiche principali, con alcuni dettagli nei casi di processi ARCH(1) e GARCH(1,1).

Il quinto capitolo prende in esame l'evoluzione del prezzo dell'EU ETS dal 2018 ad oggi, e nello specifico analizza numericamente l'andamento in Fase 4 (in notazione indicheremo con $P = P_1, \dots, P_n$ la successione dei prezzi e con $\mathcal{T} = t_1, \dots, t_n$ gli istanti di tempo relativi). In particolare, si osserva che questa serie storica non ha un comportamento stazionario e viene quindi considerata la successione dei ritorni logaritmici, definiti da:

$$Y_{t_k} = \log \frac{P_{t_k}}{P_{t_{k-1}}}, \quad \text{per ogni } k = 2, \dots, n.$$

Questa infatti, mostra un andamento stazionario e, dopo averne valutato le caratteristiche principali, tra cui il comportamento eteroschedastico tramite il test ARCH, vengono stimati diversi modelli, al fine di individuare il miglior candidato. In particolare, rimuovendo la componente di trend tramite una media mobile, si trova che il modello congiunto ARMA(5,1) - GARCH(1,1) è la scelta più opportuna. La volatilità stimata mostra infatti il comportamento atteso e la verifica del modello tramite l'analisi dei residui ha mostrato buoni risultati.

Capitolo 1

Introduzione dell'Emissions Trading System europeo

1.1 Cambiamento climatico

La scoperta che l'utilizzo di combustibili fossili e la deforestazione del pianeta producesse il surriscaldamento dell'atmosfera, risale alla fine del XIX secolo, ma il prevalere degli obiettivi di sviluppo economico, mise in secondo piano il problema. Solo alla fine degli anni Settanta, nel corso della prima World Meteorological Conference, i pericoli posti dal cambiamento climatico causato dall'aumento della temperatura del pianeta, divengono oggetto di esame e discussione a livello internazionale.

L'Intergovernmental Panel on Climate Change (IPCC), istituito nel 1988, è formato da scienziati provenienti da vari paesi che stipulano rapporti in cui vengono illustrati i progressi scientifici, gli impatti e le azioni di contenimento del cambiamento climatico. Nel primo report, elaborato nel 1990, sono state formulate ipotesi sugli effetti delle attività dell'uomo sul clima. Sulla base di questo rapporto, nel 1992 è stata adottata la Convenzione quadro delle Nazioni Unite sui cambiamenti climatici (UNFCCC), il cui obiettivo era la stabilizzazione della concentrazione di gas serra nell'atmosfera in modo da impedire pericolose interferenze antropogeniche con il clima, mediante processi di mitigazione del cambiamento climatico.

Un aspetto importante trattato nella convenzione è il *principio della responsabilità comune ma differenziata*, che prevede l'adozione di norme più permissive agli Stati meno sviluppati rispetto ai Paesi industrializzati. È in parte legato anche ad un aspetto della *giustizia climatica*, che riguarda la diseguale responsabilità storica che i paesi hanno in relazione alla crisi climatica. A tal proposito, i Paesi che si sono arricchiti grazie ad attività anche hanno emesso la maggior parte delle emissioni di gas serra, hanno la responsabilità di contribuire a mitigare gli impatti dei cambiamenti climatici su Paesi e comunità più colpiti.

La convenzione è stata ratificata da 195 paesi, che da allora partecipano alle annuali Conferenze delle Parti (COP).

1.1.1 Il protocollo di Kyoto

Il protocollo di Kyoto, adottato nel 1997 a conclusione della COP 3, è uno dei più importanti strumenti giuridici internazionali volti a combattere i cambiamenti climatici. È infatti il primo accordo che contiene gli impegni dei paesi industrializzati a ridurre le emissioni di alcuni Greenhouse Gases (GHG), ovvero gas ad effetto serra, tra cui anidride carbonica (CO_2), metano (CH_4), protossido di azoto (N_2O) e i gas fluoruranti. Questi gas assorbono radiazioni solari e intrappolano calore nell'atmosfera, causando il riscaldamento globale.

Inoltre, a differenza della Convenzione quadro definita come un accordo legalmente non vincolante, il protocollo di Kyoto stabilisce obiettivi vincolati e quantificati dei gas ad effetto serra. I Paesi industrializzati o con economie in transizione avevano l'obbligo di ridurre le loro emissioni nel periodo 2008 - 2012, di almeno il 5% rispetto al livello del 1990.

Il protocollo di Kyoto prevede due tipi di strumenti per conseguire gli obiettivi di riduzione:

1. Politiche e misure: includono gli interventi previsti dagli Stati attraverso programmi nazionali specifici;
2. Meccanismi flessibili: sono strumenti basati sul mercato, il cui scopo è quello di permettere ai paesi industrializzati di raggiungere i loro obiettivi grazie allo scambio di quote di emissioni e ottenere crediti realizzando progetti per la riduzione delle emissioni in altri paesi. Si distinguono tre tipi di meccanismi flessibili:
 - International Emissions Trading (IET): consiste nello scambio di permessi di emissione tra Paesi industrializzati e ad economia in transizione. Il meccanismo è simile a quello dell'EU ETS, presentato in 1.2.1, con la differenza che in questo caso la vendita di crediti di emissione è a livello internazionale.
 - Clean Development Mechanism (CDM): consente ai Paesi sviluppati di realizzare progetti a tecnologia pulita nei Paesi in via di sviluppo, ricevendo crediti di emissione pari alla riduzione ottenuta rispetto ai livelli che si sarebbero avuti senza il progetto. Tali crediti vengono chiamati Certified Emissions Reduction (CER);
 - Joint Implementation (JI): consente la collaborazione tra Paesi industrializzati e quelli ad economia in transizione, nello sviluppo di progetti per il raggiungimento dei rispettivi obiettivi di riduzione delle emissioni. I crediti prodotti vengono chiamati Emissions Reduction Unit (ERU) e vanno utilizzati congiuntamente con il paese ospite.

1.1.2 L'Accordo di Parigi

A seguito della COP 21 del 2015, si è raggiunto l'Accordo di Parigi, il primo piano d'azione globale e giuridicamente vincolante sul clima mondiale. Questo accordo prevede tre fronti diversi su cui agire.

L'aspetto più sfidante riguarda i processi di *mitigazione*, ovvero le azioni intraprese da governi, aziende e persone per ridurre, sequestrare o prevenire le emissioni di gas serra. Le Parti si pongono come obiettivo limitare l'innalzamento della temperatura globale a $+1.5^{\circ}\text{C}$ rispetto al periodo pre-industriale. Questa soglia è anche detta *tipping point*, ovvero punto di svolta, oltre al quale alcune delle alterazioni causate dai cambiamenti climatici diventano irreversibili.

Ad oggi, anche gli scenari migliori indicano una significativa possibilità di superare questo limite, seppur temporaneamente. Con *climate overshoot* ci si riferisce al periodo durante il quale il riscaldamento sarà aumentato oltre $+1.5^{\circ}\text{C}$, prima di riscendere; questo periodo è atteso intorno alla metà di questo secolo.

Il secondo punto su cui si focalizza l'accordo riguarda i processi di *adattamento*, ovvero l'insieme di azioni che aiutano a ridurre la vulnerabilità agli impatti attuali o previsti dei cambiamenti climatici.

L'aspetto più innovativo di questo piano riguarda l'organizzazione degli investimenti e dei flussi finanziari in modo da favorire uno sviluppo basato su ridotte emissioni di gas serra.

Ciascun Paese si è impegnato a creare un piano nazionale (Nationally Determined Contribution (NDC) o "contributo determinato a livello nazionale") che delinea le priorità di mitigazione e adattamento perseguite per ridurre le proprie emissioni. Si tratta di piani a breve e medio termine che vengono aggiornati ogni cinque anni.

1.1.3 Conferenze delle Parti 26 e 27

Nel 2020 la COP 26 si è tenuta a Glasgow e ha portato un contributo notevole nella sfida di ridurre gli effetti del cambiamento climatico. È stato confermato l'obiettivo di mantenere la temperatura entro $+1.5^{\circ}\text{C}$ rispetto ai livelli pre-industriali. Inoltre il documento fissa l'obiettivo minimo di decarbonizzazione per tutti gli stati firmatari, che dovranno ridurre del 45% le emissioni di anidride carbonica al 2030 rispetto al 2010, e zero emissioni nette intorno al 2050.

Altri progressi che si sono raggiunti con la COP 26 sono stati il mercato globale delle emissioni di carbonio e il reporting format con le norme con cui gli stati comunicano i loro risultati nella decarbonizzazione, favorendone la trasparenza.

Lo scorso anno si è tenuta la COP 27, in cui si è affrontato il tema riguardante sia i finanziamenti da reperire per la transizione energetica, sia il ritmo di riduzione dei combustibili fossili in uno scenario in cui di fatto non si sono ancora prese misure politiche nazionali concrete per raggiungere gli obiettivi prefissati. Sono stati stanziati fondi per

l'eliminazione del carbone in Indonesia, Sud Africa e Vietnam. Si sono presi impegni per ridurre la deforestazione e promuovere iniziative di agricoltura sostenibile. Si è definito inoltre che un ruolo importante per accelerare la transizione energetica è il mercato dei crediti carbonici, fondamentale per compensare le emissioni.

1.1.4 Green Deal Europeo

In linea con quanto previsto dall'Accordo di Parigi e dalle COP più recenti, diversi Paesi hanno presentato i rispettivi piani strategici per raggiungere l'obiettivo di neutralità carbonica entro il 2050 e riuscire quindi a mantenere la temperatura entro 1.5°C rispetto al periodo pre-industriale. Le leve su cui agire sono molteplici; alcuni esempi possono essere le tecnologie di Carbon Capture Utilization & Storage (CCUS), che catturano la CO₂ dall'atmosfera stoccandola in modo permanente o utilizzandola in altri processi, la produzione di energia da fonti rinnovabili, quali solare e eolico, l'elettrificazione di veicoli e l'utilizzo di biocarburanti.

Per quanto riguarda le proposte legislative adottate in Europa, nel 2021 è stato formulato il Green Deal europeo, ovvero un insieme di strategie e piani d'azione proposti e adottati dalla commissione europea per ridurre le emissioni di gas serra del 55% entro il 2030 (Fit for 55) e raggiungere la neutralità carbonica entro il 2050.

In particolare, a febbraio del 2023 è stato presentato il *Green Deal Industrial Plan*, che ha come obiettivo fare dell'Unione Europea il leader nelle tecnologie carbon neutral. Questo piano punta a semplificare le norme sugli aiuti di Stato per l'introduzione di energie rinnovabili e per la decarbonizzazione dei processi industriali. Un primo ambito in cui si è focalizzato il piano è proprio lo snellimento del quadro regolatorio, l'incremento e la semplificazione dei fondi nazionali ed europei e l'aumento della cooperazione internazionale. Un obiettivo importante da perseguire è lo sviluppo di nuove tecnologie e competenze. A tal proposito gli aiuti di maggior rilievo sono stati stanziati per la produzione di batterie, pannelli solari, turbine eoliche, pompe di calore, elettrolizzatori e sistemi CCUS e relative materie prime critiche.

1.2 Sistemi di Carbon Pricing

Il prezzo del carbonio è uno degli strumenti fondamentali per raggiungere gli obiettivi di decarbonizzazione, poiché crea un incentivo finanziario per mitigare le emissioni attraverso l'imposizione di prezzi. Ci sono diversi sistemi di carbon pricing adottati nei vari paesi.

L'*Emissions Trading System (ETS)* è un sistema dove gli emettitori possono commerciare le unità di emissione (tipicamente una tonnellata), tenendo conto del loro target di emissioni definito dal tipo di settore. Ci sono due tipologie differenti: il Cap and trade system, in cui alle società è applicato un massimale alle emissioni e le unità vengono distribuite gratuitamente o tramite aste. Se le emissioni non hanno superato il cap assegnato, l'impresa avrà disponibilità di quote da vendere (trade) sul mercato. Viceversa,

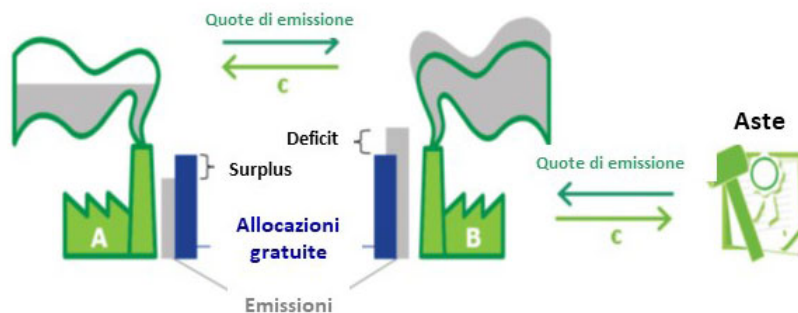


Figura 1.1: European Emissions Trading System (EU ETS)

se l'azienda avrà emesso oltre il cap assegnato, dovrà reperire le quote mancanti acquistandole dal mercato. Nel *Baseline and credit system* invece a ciascuna società vengono definiti dei livelli di emissione di riferimento e, qualora un'attività si impegnasse a ridurre le sue emissioni, le vengono concessi dei crediti che possono essere venduti ad altri emettitori che invece superano i livelli concessi.

Un altro sistema possibile è la *Carbon Tax*, una tassa che stabilisce direttamente il prezzo del carbonio, definendo un'aliquota fiscale esplicita sulle emissioni di GHG o su tonnellata di CO₂ emessa.

L'*Offset Mechanism* o meccanismo di compensazione è un sistema di compensazioni a seguito di riduzioni delle emissioni di gas serra derivanti da attività basate su progetti e programmi. Questi emettono crediti di carbonio, che possono essere utilizzati da una società per compensare emissioni non evitabili di altri progetti. Eventualmente questi crediti possono anche essere venduti.

Un approccio diverso è il *Result - Based Climate Finance (RBCF)*. Si tratta di un sistema di finanziamenti rilasciati a seguito di risultati relativi alla gestione del cambiamento climatico. In questa categoria è importante che le riduzioni di emissioni di GHG siano verificate e allo stesso tempo abbiano anche un impatto sociale, ovvero siano volte a ridurre la povertà, migliorare l'accesso all'energia pulita e offrire benefici per la salute e la comunità.

Infine vi è l'*Internal carbon pricing*, uno strumento utilizzato all'interno di un'organizzazione, volto a guidare il proprio processo decisionale in relazione a impatti, rischi e opportunità per il cambiamento climatico. Questo sistema è spesso adottato nel settore dell'aviazione.

1.2.1 European Union Emissions Trading System (EU ETS)

Il principale strumento di carbon pricing adottato in Europa è il sistema degli ETS, l'**European Union Emissions Trading System (EU ETS)**, che opera secondo il

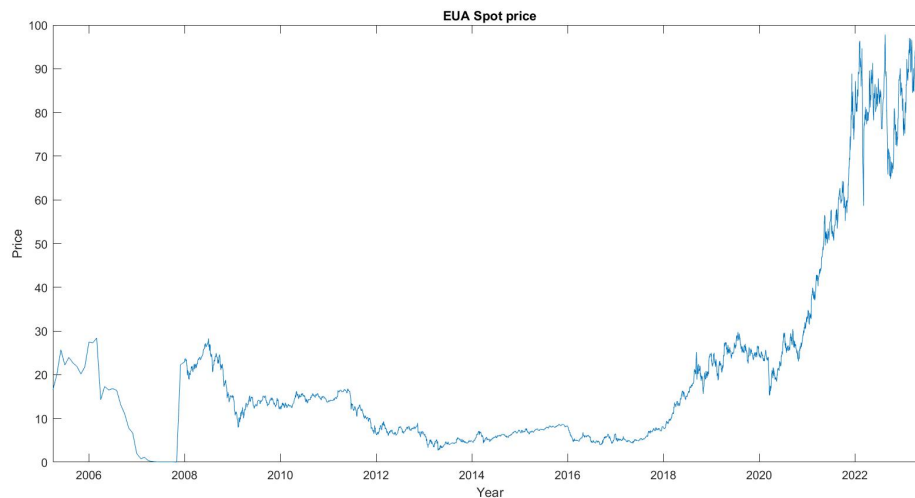


Figura 1.2: Andamento del prezzo spot delle quote dell'EU ETS (2005 - 2023)

principio del Cap and Trade. Viene quindi fissato un massimale ovvero un numero prefissato di quote (**European Union Allowances (EUA)**, pari a 1 tonnellata di CO₂ equivalente) rilasciate nel sistema, che stabilisce la quantità massima che può essere emessa dagli impianti che partecipanti. Entro questo limite, le imprese possono acquistare o vendere quote in base alle loro esigenze (Figura 1.1). Le quote destinate al settore dell'aviazione sono distinte dalle EUA, e vengono chiamate European Union Allowances Aviation (EUA A).

L'EU ETS è il primo sistema internazionale per lo scambio di quote di emissione al mondo, è stato istituito nel 2005 ed è attivo in 30 paesi (i 27 dell'UE, a cui si aggiungono l'Islanda, il Liechtenstein e la Norvegia).

Dalla sua introduzione ha subito numerosi cambiamenti, diventando gradualmente sempre più restrittivo. Queste variazioni sono state una delle cause della grande instabilità del prezzo di una quota, come mostrato in Figura 1.2.

Nell'intervallo temporale 2008 - 2023, sono stati utilizzati i prezzi giornalieri della quota, mentre, in assenza di tali dati per gli anni 2005 - 2007, sono esposte le medie mensili.

La Fase 1 (2005 - 2007), chiamata anche fase pilota, copriva solamente le emissioni di CO₂ dei produttori di energia elettrica e delle industrie ad alta intensità energetica. Al tempo non vi erano dati affidabili sulle emissioni e di conseguenza il massimale di quote rilasciate è stato deciso sulla base di stime. La quantità totale di quote rilasciate (quasi tutte concesse a titolo gratuito) ha quindi superato le emissioni effettive, e, con l'offerta notevolmente superiore alla domanda, nel 2007 il prezzo delle quote è sceso a zero (infatti le quote della Fase 1 non potevano essere accumulate per l'utilizzo nella Fase

2). In questo primo periodo la sanzione per il mancato rispetto delle emissioni concesse era di 40 € per ogni tonnellata di CO₂ emessa in eccesso.

Durante la Fase 2 (2008 - 2012) il numero di quote allocate diminuì del 6.5% rispetto al 2005. Tuttavia, la crisi economica del 2008 ha portato a riduzioni delle emissioni superiori alle attese, provocando un notevole surplus di quote (2.2 miliardi), che ha gravato pesantemente sul prezzo del carbonio durante tutta la Fase 2. La percentuale di assegnazione gratuita fu pari al 90%, leggermente inferiore rispetto alla fase precedente, mentre la multa per il mancato rispetto delle emissioni concesse è stata aumentata a 100 € per tonnellata.

La Fase 3 (2013 - 2020) introdusse cambiamenti significativi, volti a potenziare il funzionamento del sistema. In particolare, sono stati compresi nel sistema EU ETS nuovi settori, quali gli impianti di produzione di alluminio e di ammoniaca, nonché nuovi gas, come i fluorocarburi e il protossido di azoto. Inoltre, al posto dei singoli cap nazionali, è stato previsto un tetto massimo di emissioni a livello europeo da ridurre progressivamente dell'1.74% all'anno fino al 2020. È stato apportato un cambiamento sostanziale anche per quanto riguarda la distribuzione delle quote, per la quale il metodo più diffuso fu la messa all'asta, mentre le quote distribuite gratuitamente, erano basate su benchmark prestabiliti a livello europeo e sulla produzione storica. Le aste si sono tenute su un'unica piattaforma unica europea, la European Energy Exchange (EEX). Inoltre, per prevenire il calo del prezzo dei permessi come registrato nella seconda fase, è stata prevista la possibilità di *banking*, ossia la conservazione dei permessi non utilizzati e l'utilizzo nel periodo successivo.

All'inizio della terza fase ci si rese anche conto dell'eccedenza di quote (circa 2 miliardi nel 2013), con conseguente riduzione del prezzo e minore incentivo a ridurre le emissioni. Una misura adottata nel breve periodo fu posticipare la messa all'asta di 900 milioni di quote tra il 2014 e il 2016. Accanto a questa misura è stata prevista l'istituzione di una Riserva Stabilizzatrice del Mercato (MSR), con l'obiettivo di bilanciare l'attuale eccedenza di quote e adeguare l'offerta di quote da mettere all'asta sul mercato.

La Fase 4 (2021 - 2030) segue il regolamento proposto nella Direttiva 2018/410/UE pubblicata a marzo 2018. La riforma è stata necessaria per poter adempiere agli impegni assunti nell'ambito dell'Accordo di Parigi. Infatti, il Quadro per il clima e l'energia 2030 prevede l'obiettivo vincolante di ridurre entro il 2030 le emissioni nel territorio dell'Unione Europea di almeno il 40% rispetto ai livelli del 1990, mentre i settori interessati dal sistema ETS dovranno ridurre le emissioni del 43%.

Le novità principali introdotte dalla direttiva si concentrano su quattro aspetti.

Il numero complessivo di quote rilasciate diminuirà a un tasso annuo del 2.2% invece che dell'1.74%.

Per quanto riguarda le modalità di assegnazione delle quote, la percentuale da mettere all'asta viene fissata pari al 57% del totale e vengono introdotte nuove norme per il calcolo dell'assegnazione gratuita in funzione del livello di attività. In particolare vengono

assegnate quote gratuitamente ai settori rientranti nella lista del “carbon leakage” diretto, ossia esposti ad un elevato rischio di rilocalizzazione delle emissioni. Eventuali quote rimaste dalla Fase 3, rimarranno automaticamente valide in Fase 4. Non è invece permesso il *borrowing* per le quote di emissioni di quarta fase; questo significa che le quote emesse in Fase 4 non potranno essere utilizzate per raggiungere lo stato di conformità in Fase 3.

Viene rafforzato il meccanismo della riserva stabilizzatrice del mercato, istituito dall’UE nel 2015 per ridurre l’eccedenza di quote nel mercato del carbonio e migliorare la resilienza dell’EU ETS a shock futuri.

Infine, attraverso meccanismi di finanziamento, mira ad aiutare i settori industriale ed energetico a far fronte alle sfide dell’innovazione e degli investimenti della transizione, verso un’economia a basse emissioni di carbonio.

Ad oggi il sistema coinvolge oltre 11 000 operatori a livello europeo, tra cui operatori del settore aviazione, impianti termoelettrici industriali, manifatture e impianti di produzione, stoccaggio e trasporto di diverso tipo. In Italia sono circa 1200 gli impianti coinvolti. Sono esclusi dall’EU ETS gli ospedali e i “Piccoli Emettitori”, ovvero impianti con emissioni inferiori a 25 000 tonnellate di CO₂ equivalente e, nel caso di impianti di combustione, con potenza termica nominale inferiore a 35 MW, escluse le emissioni da biomassa.

Il 16 maggio 2023 sono stati poi pubblicati alcuni strumenti normativi inerenti il pacchetto “Fit for 55” che prevede la riduzione di emissioni di GHG del 55% entro il 2030, tra cui l’aggiornamento della direttiva ETS e l’introduzione del Carbon Border Adjustment Mechanism (CBAM), una tassa che gli importatori europei pagheranno a partire dal 2026 sul contenuto carbonico dei prodotti importati dai Paesi non-UE.

Per quanto riguarda la revisione della normativa dell’EU ETS sono state introdotte importanti modifiche al fattore di riduzione. Per raggiungere l’obiettivo di riduzione di almeno 62% delle emissioni di GHG entro il 2030, il fattore di riduzione viene aumentato dal 2.2% al 4.3% dal 2024 al 2027 e a 4.4% dal 2028 al 2030.

La panoramica sull’introduzione e l’evoluzione dell’EU ETS sarà utile nel Capitolo 5, dove verrà fatto un’analisi dettagliata di questa serie storica nella Fase 4.

Capitolo 2

Analisi di serie storiche

In questo capitolo viene introdotta la procedura di analisi di serie storiche, poi approfondita nei capitoli successivi. Dopo aver richiamato alcuni concetti base di probabilità, variabili aleatorie e processi stocastici, definiremo la classe dei processi stazionari e le sue funzioni caratterizzanti, quali la funzione di autocovarianza e autocorrelazione. In seguito viene presentata la definizione formale di serie storica e le varie fasi del processo di analisi di una serie storica, tra cui l'eliminazione di eventuali fenomeni di trend e/o stagionalità per rendere la serie storica in esame stazionaria. A questo punto è possibile identificare alcuni modelli adeguati per tali osservazioni, andando quindi a stimare i parametri, valutare i residui ottenuti e scegliere il candidato migliore da utilizzare anche a scopo predittivo.

2.1 Probabilità e variabili aleatorie

In questa sezione richiamiamo velocemente alcune nozioni elementari di probabilità con il principale scopo di fissare una terminologia e definizioni in modo chiaro, ma senza approfondire tali concetti.

Definizione 2.1. Uno spazio di probabilità è una terna (Ω, \mathcal{A}, P) dove:

1. Ω è l'insieme degli eventi elementari;
2. \mathcal{A} è una σ -algebra su Ω , ovvero una famiglia di sottoinsiemi di Ω tale che:
 - (a) $\emptyset, \Omega \in \mathcal{A}$;
 - (b) $A \in \mathcal{A} \implies A^c \in \mathcal{A}$;
 - (c) $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A} \implies \bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A}$;
3. $P : \mathcal{A} \rightarrow [0, 1]$ è una misura di probabilità, ovvero:

(a) $P(\Omega) = 1$;

(b) $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$, $A_i \cap A_j = \emptyset$ per ogni $i \neq j \implies P\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} P(A_i)$.

Definizione 2.2. Siano (Ω, \mathcal{A}, P) uno spazio di probabilità e (E, \mathcal{E}) uno spazio misurabile. $X : (\Omega, \mathcal{A}, P) \rightarrow (E, \mathcal{E})$ si dice **variabile aleatoria (v.a.)** se X è una funzione misurabile, ossia per ogni $B \in \mathcal{E}$ risulta $X^{-1}(B) \in \mathcal{A}$.

Se una v.a. assume un insieme finito o al più numerabile di valori, si dice *discreta* e come σ -algebra di riferimento viene comunemente utilizzata quella delle parti. Tipicamente però considereremo v.a. *continue*, a valori reali, dove $E = \mathbb{R}$ oppure, nel caso di v.a. multidimensionali, $E = \mathbb{R}^n$. In questo caso, se non diversamente specificato, verrà utilizzata la σ -algebra dei boreliani di \mathbb{R} e \mathbb{R}^n rispettivamente (in notazione $\mathcal{B}(\mathbb{R})$ e $\mathcal{B}(\mathbb{R}^n)$).

Definizione 2.3. Sia $X : (\Omega, \mathcal{A}, P) \rightarrow (E, \mathcal{E})$ una variabile aleatoria. Si dice **legge** o **distribuzione** di X la funzione $\mu_X : \mathcal{E} \rightarrow [0, 1]$ definita da:

$$\mu_X(B) := P(X \in B) = P(\{\omega \in \Omega : X(\omega) \in B\}) \quad \text{per ogni } B \in \mathcal{E}.$$

Si osserva che la legge di una v.a. è una misura di probabilità.

Ricordiamo che per le v.a. a valori in \mathbb{R} o in \mathbb{R}^n , una delle quantità caratterizzanti è la **funzione di ripartizione**, che indicheremo con F_X . Nel caso di una v.a. multidimensionale $X = (X_1, \dots, X_n)$, $F_X : \mathbb{R}^n \rightarrow [0, 1]$ è definita da:

$$F_X(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n) \quad \text{per ogni } (x_1, \dots, x_n) \in \mathbb{R}^n.$$

A partire dalla funzione di ripartizione è talvolta possibile definire la funzione di densità associata alla v.a., che indicheremo con f_X . Quando F_X è assolutamente continua, data X v.a. multidimensionale, si definisce **funzione di densità** di X , la funzione $f_X : \mathbb{R}^n \rightarrow [0, 1]$ integrabile, tale che per ogni $(x_1, \dots, x_n) \in \mathbb{R}^n$ valga:

$$F_X(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_X(t) dt.$$

Nel caso invece in cui la v.a. sia puramente discreta, la funzione di densità è definita come:

$$f_X(x_1, \dots, x_n) = P(X_1 = x_1, \dots, X_n = x_n).$$

Definizione 2.4. Sia X una variabile aleatoria integrabile definita sullo spazio di probabilità (Ω, \mathcal{A}, P) . Chiamiamo **valore atteso** o **speranza** il numero $E[X]$ definito da:

$$E[X] = \int_{\Omega} X dP.$$

Nel caso particolare in cui la v.a. abbia densità f_X , il valore atteso assume le seguenti forme:

- $E[X] = \sum_{j=1}^{\infty} x_j f_X(x_j)$ se X è v.a. discreta;
- $E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$ se X è v.a. continua.

Definizione 2.5. Sia X una variabile aleatoria di quadrato integrabile definita sullo spazio di probabilità (Ω, \mathcal{A}, P) con valore atteso $E[X]$. Chiamiamo **varianza** di X il numero reale positivo:

$$\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

La quantità $\sqrt{\text{Var}(X)}$ si chiama **deviazione standard** o **scarto quadratico medio** di X .

Definizione 2.6. Sia (Ω, \mathcal{A}, P) uno spazio di probabilità, X e Y due variabili aleatorie di quadrato integrabile. Chiamiamo **covarianza** di X e Y la quantità:

$$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y].$$

Enunciamo ora il seguente teorema che garantisce l'esistenza e unicità della speranza condizionale ad una σ -algebra.

Teorema 2.1. Sia (Ω, \mathcal{A}, P) uno spazio di probabilità, $X : \Omega \rightarrow \mathbb{R}$ una variabile aleatoria integrabile e \mathcal{G} una sotto σ -algebra di \mathcal{A} . Allora esiste Y variabile aleatoria integrabile che verifica:

- Y è \mathcal{G} -misurabile;
- $\int_G Y dP = \int_G X dP$ per ogni $G \in \mathcal{G}$.

$Y = E[X|\mathcal{G}]$ si dice *valore atteso di X condizionato a \mathcal{G}* .

Proposizione 2.1. (Proprietà del valore atteso condizionato) Siano X, Y variabili aleatorie integrabili, \mathcal{G} σ -algebra, $\mathcal{G} \subset \mathcal{A}$ e $\alpha, \beta \in \mathbb{R}$, allora:

1. $E[X|X] = X$;
2. $E[\alpha X_1 + \beta X_2|\mathcal{G}] = \alpha E[X_1|\mathcal{G}] + \beta E[X_2|\mathcal{G}]$;
3. $E[E[X|\mathcal{G}]] = E[X]$;
4. se X è \mathcal{G} -misurabile, allora $E[X|\mathcal{G}] = X$;
5. se X è indipendente da \mathcal{G} , allora $E[X|\mathcal{G}] = E[X]$;
6. (Proprietà della torre) se \mathcal{G}_1 e \mathcal{G}_2 sono σ -algre, con $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{A}$, allora $E[E[X|\mathcal{G}_2]|\mathcal{G}_1] = E[X|\mathcal{G}_1]$;
7. se Y è \mathcal{G} -misurabile e limitata, allora $E[XY|\mathcal{G}] = Y E[X|\mathcal{G}]$.

Oltre al valore atteso e alla varianza, ci sono altre quantità che possono indicare delle caratteristiche di una variabile aleatoria.

Definizione 2.7. Sia X una variabile aleatoria definita sullo spazio di probabilità (Ω, \mathcal{A}, P) . Dato $p \geq 1$, si chiama (se esiste, ovvero se $|X|^p$ è integrabile) **momento di ordine p** la quantità $E[X^p]$.

Si noti che se X è una v.a. centrata, il momento di secondo ordine è proprio la varianza ed è una misura della dispersione della distribuzione attorno alla propria media.

Altre due proprietà della distribuzione di una variabile aleatoria sono la simmetria e la curtosi. Per determinarle vengono utilizzati due indici: l'indice di *simmetria* o *asimmetria* e l'indice di *curtosi* di Pearson.

Il primo rileva se e quanto una distribuzione sia disposta simmetricamente attorno alla sua media. Con asimmetria negativa si intende che la coda più lunga è quella di sinistra, mentre, se prevalgono valori con forti scarti positivi e pochi valori con deboli scarti negativi, la coda più lunga sarà quella di destra, e si parlerà di asimmetria positiva. Se non è presente differenza tra gli scarti, si parla di distribuzione simmetrica.

L'**indice di asimmetria** di una v.a. X è dato da:

$$A := \frac{E[X^3]}{E[X^2]^{3/2}}.$$

L'indice di curtosi di Pearson invece, rileva l'allontanamento della distribuzione da quella Normale. Se la distribuzione è più appuntita rispetto alla distribuzione Normale si dice *leptocurtica*, mentre se è più piatta si dice *platicurtica*.

L'**indice di curtosi** di una v.a. X è dato da:

$$K := \frac{E[X^4]}{E[X^2]^2}.$$

L'indice di curtosi corrispondente alla distribuzione normale è 3; spesso quindi, per valutare la presenza di curtosi si preferisce utilizzare $\frac{E[X^4]}{E[X^2]^2} - 3$, dove un valore negativo indica una distribuzione platicurtica, mentre uno positivo una distribuzione leptocurtica.

Definizione 2.8. Preso (Ω, \mathcal{A}, P) uno spazio di probabilità e $\mathcal{T} \subseteq \mathbb{R}^+$ l'insieme dei tempi, si definisce **processo stocastico** reale una famiglia $X = (X_t)_{t \in \mathcal{T}}$ di variabili aleatorie $X_t : (\Omega, \mathcal{A}, P) \rightarrow \mathbb{R}$, $t \in \mathcal{T}$, definite sullo stesso spazio di probabilità.

Un processo stocastico $(X_t)_{t \in \mathcal{T}}$, quindi può essere visto come una funzione

$$\begin{aligned} X : \Omega \times \mathcal{T} &\longrightarrow E \\ (\omega, t) &\longmapsto X_t(\omega) \end{aligned}$$

che, per ogni $t \in \mathcal{T}$ fissato, $X(\cdot, t)$ è una variabile aleatoria. Fissato invece $\omega \in \Omega$, la funzione $t \in \mathcal{T} \mapsto X_t(\omega)$ si definisce *realizzazione* o *traiettoria* di ω per il processo X .

A seconda della tipologia dell'insieme \mathcal{T} , è possibile classificare i processi stocastici. Chiameremo *processo stocastico a tempo discreto* (o semplicemente *discreto*) un processo per il quale \mathcal{T} è un insieme discreto (generalmente \mathbb{N}, \mathbb{Z} o $\{0, 1, \dots, T\}$); indicheremo invece con *processo stocastico a tempo continuo*, (o semplicemente *continuo*) un processo per cui \mathcal{T} è un intervallo di \mathbb{R} (generalmente $[0, +\infty)$ o $[0, T]$). È bene sottolineare che con la terminologia *discreto* e *continuo* ci riferiremo quindi all'insieme dei tempi \mathcal{T} , e non all'insieme degli stati in cui prende valori il processo (E).

Grazie al teorema di Kolmogorov, la legge del processo è univocamente determinata una volta che si conoscono le leggi finito-dimensionali. Per descrivere completamente un processo stocastico $(X_t)_{t \in \mathcal{T}}$ è quindi necessario conoscere le sue leggi finito-dimensionali, ossia le leggi dei vettori del tipo $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ per ogni insieme dei tempi (t_1, t_2, \dots, t_n) . Nel caso di processi a valori reali, queste leggi sono descritte, ad esempio, dalla funzione di ripartizione:

$$F_X(x_1, \dots, x_n; t_1, \dots, t_n) = P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n).$$

2.2 Processi stazionari e funzioni caratterizzanti

In questa sezione poniamo l'attenzione su una classe particolare di processi, i *processi stazionari*, a partire dai quali è possibile sviluppare tecniche efficienti per l'analisi di serie storiche. Dato che questo è l'argomento centrale dell'elaborato e le serie che tratteremo saranno a tempo discreto, consideriamo l'insieme dei tempi $\mathcal{T} = \mathbb{Z}$.

Notazione: Preso un processo stocastico $(X_t)_{t \in \mathbb{Z}}$ tale che ogni X_t sia di quadrato integrabile, indichiamo con $\mu_X(t) = E[X_t]$ la media del processo e con $\gamma_X(r, s) = \text{Cov}(X_r, X_s)$ la funzione di covarianza.

Intuitivamente, un processo stocastico è stazionario se la sua struttura probabilistica è invariante nel tempo. Vediamo ora la definizione formale.

Definizione 2.9. Un processo stocastico $(X_t)_{t \in \mathbb{Z}}$ è detto **stazionario** o **debolmente stazionario** se le variabili aleatorie X_t sono di quadrato integrabile e soddisfa le seguenti condizioni:

- i) $E[X_t] = E[X_s]$ per ogni $s, t \in \mathbb{Z}$, ovvero il processo è costante in media;
- ii) $\gamma_X(r + t, s + t) = \gamma_X(r, s)$ per ogni $r, s, t \in \mathbb{Z}$, ovvero la funzione di covarianza dipende solo dalla differenza dei tempi in argomento.

Grazie alla condizione ii), per un processo stazionario è conveniente definire la funzione di covarianza in funzione di una sola variabile:

$$\gamma_X(h) := \gamma_X(h, 0) = \gamma_X(t + h, t).$$

In questo caso verrà detta anche *funzione di autocovarianza*. Ponendo $h = 0$ si ottiene la varianza del processo:

$$\gamma_X(0) = \text{Var}(X_t) = E[(X_t - E[X_t])^2], \quad \text{per ogni } t.$$

Definizione 2.10. Un processo stocastico $(X_t)_{t \in \mathbb{Z}}$ è detto **strettamente (o fortemente) stazionario** se le distribuzioni congiunte di (X_1, \dots, X_n) e di $(X_{1+h}, \dots, X_{n+h})$ sono uguali per ogni $h \in \mathbb{Z}$, $n > 0$, ossia se

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_{1+h}, \dots, X_{n+h}}(x_1, \dots, x_n).$$

Osservazione 2.2.1. Un processo stocastico fortemente stazionario con momenti secondi finiti, è anche stazionario. Il viceversa, invece, non è sempre vero; infatti la stazionarietà in senso stretto coinvolge tutti i momenti della distribuzione, la cui esistenza, per ordini superiori al secondo, non è garantita dalla semplice condizione di stazionarietà.

Un esempio di processi stocastici stazionari che sono anche fortemente stazionari sono i processi gaussiani.

Definizione 2.11. Un processo stocastico $(X_t)_{t \in \mathbb{Z}}$ si dice **processo gaussiano** se per ogni $n \in \mathbb{N}^+$ e per ogni $(t_1, t_2, \dots, t_n) \in \mathbb{Z}^n$, il vettore aleatorio n -dimensionale $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ ha distribuzione gaussiana multivariata.

Nel caso particolare in cui la matrice di covarianza \mathbf{K} sia definita positiva, il processo ha una densità di probabilità pari a

$$f_{(X_{t_1}, X_{t_2}, \dots, X_{t_n})}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(\mathbf{K})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{b})^T \mathbf{K}^{-1}(\mathbf{x} - \mathbf{b})\right),$$

dove $b_i = E[X_{t_i}]$ per ogni $i = 1, \dots, n$, $\mathbf{K} = (\text{Cov}(X_{t_i}, X_{t_j}))_{i,j=1, \dots, n}$ e dipendono entrambi da (t_1, t_2, \dots, t_n) .

È chiaro che se $(X_t)_{t \in \mathbb{Z}}$ è un processo gaussiano stazionario, allora è anche fortemente stazionario. Infatti, per ogni $n \geq 1$ e per ogni $h \in \mathbb{Z}$, i vettori (X_1, \dots, X_n) e $(X_{1+h}, \dots, X_{n+h})$ sono gaussiani con stessa media e matrice di covarianza, dunque hanno la stessa distribuzione.

Definizione 2.12. Dati un processo stocastico stazionario $(X_t)_{t \in \mathbb{Z}}$ e $h \in \mathbb{Z}$, si definiscono:

- **funzione di autocovarianza (ACVF)** di $(X_t)_{t \in \mathbb{T}}$ di passo (o lag) h

$$\gamma_X(h) = \text{Cov}(X_{t+h}, X_t); \quad (2.1)$$

- **funzione di autocorrelazione (ACF)** di $(X_t)_{t \in \mathbb{T}}$ di passo (o lag) h

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \text{Corr}(X_{t+h}, X_t). \quad (2.2)$$

Per ogni h , $\gamma_X(h)$ misura la covarianza tra due variabili del processo separate da un intervallo di tempo di ampiezza h , mentre $\rho_X(h)$ rappresenta il coefficiente di correlazione tra tali variabili. Intuitivamente, l'autocorrelazione esprime la dipendenza che esiste tra il processo al tempo t e al tempo $t+h$.

Vediamo ora alcune proprietà della funzione di autocovarianza.

Proposizione 2.2. *Dati un processo stocastico stazionario $(X_t)_{t \in \mathbb{Z}}$ e $\gamma_X(\cdot)$ la sua funzione di autocovarianza, valgono le seguenti proprietà:*

1. $\gamma_X(0) \geq 0$;
2. $|\gamma_X(h)| \leq \gamma_X(0)$ per ogni h ;
3. $\gamma_X(h) = \gamma_X(-h)$ per ogni h ;
4. per ogni $n \in \mathbb{N}^+$ fissato, la matrice $\Gamma_n = \{\gamma_X(i-j)\}_{i,j \in \{1, \dots, n\}}$, è semidefinita positiva.

Dimostrazione. La prima proprietà si verifica immediatamente; infatti

$$\gamma_X(0) = \text{Cov}(X_t, X_t) = \text{Var}(X_t) \geq 0.$$

La seconda si ottiene come conseguenza di $|\text{Corr}(X_t, X_{t+h})| \leq 1$ e $\gamma_X(0) \geq 0$. Per la terza proprietà si osserva che

$$\gamma_X(h) = \text{Cov}(X_t, X_{t+h}) = \text{Cov}(X_{t-h}, X_t) = \gamma_X(-h).$$

Infine, presi a_1, \dots, a_n coefficienti reali, $\mathbf{X}_n = (X_n, \dots, X_1)'$, dalla non negatività della varianza, si ottiene che la matrice di autocovarianza è semidefinita positiva:

$$0 \leq \text{Var}(\mathbf{a}'\mathbf{X}_n) = \mathbf{a}'\Gamma_n\mathbf{a} = \sum_{i,j=1}^n a_i \gamma(i-j) a_j.$$

□

Osservazione 2.2.2. La funzione di autocorrelazione $\rho(\cdot)$ ha le stesse proprietà della funzione di autocovarianza, con l'aggiunta di $\rho(0) = 1$.

Un'altra funzione importante dei processi stazionari è la *funzione di autocorrelazione parziale* che, così come la funzione di autocorrelazione, contiene informazioni sulla dipendenza fra le variabili aleatorie di un processo stazionario.

Definizione 2.13. La **funzione di autocorrelazione parziale (ACF)** $\alpha_X(\cdot)$ di un processo stazionario $X = (X_t)_{t \in \mathbb{Z}}$ è definita come:

$$\alpha_X(1) = \text{Corr}(X_2, X_1) = \rho_X(1)$$

$$\alpha_X(k) = \text{Corr}(X_{k+1} - P_{\bar{sp}\{1, X_2, \dots, X_k\}} X_{k+1}, X_1 - P_{\bar{sp}\{1, X_2, \dots, X_k\}} X_1) \quad k \geq 2,$$

dove $P_{\bar{sp}\{1, X_2, \dots, X_k\}} X_{k+1}$ e $P_{\bar{sp}\{1, X_2, \dots, X_k\}} X_1$ sono rispettivamente le proiezioni ortogonali di X_{k+1} e X_1 sul sottospazio chiuso generato da $\{1, X_2, \dots, X_k\}$ e il prodotto scalare utilizzato è dato dalla speranza del prodotto.

Il valore $\alpha_X(k)$ è detto autocorrelazione parziale al passo k .

L'autocorrelazione parziale $\alpha(k)$ al passo k può essere quindi vista come la correlazione tra X_1 e X_{k+1} , tenendo in considerazione le osservazioni intermedie X_2, \dots, X_k .

2.3 Serie storiche

Una *serie storica* (o *temporale*) $\{x_t, t \in T_0\}$, è una sequenza di dati osservati in un intervallo di tempo, in cui ciascuna osservazione è associata ad un istante di tempo specifico.

A seconda della tipologia dell'insieme dei tempi T_0 , le serie storiche si possono dividere in due categorie: si diranno *discrete* quando T_0 è un insieme discreto (ad esempio quando le osservazioni vengono fatte ad intervalli temporali fissi); si diranno *continue* quando le osservazioni vengono fatte in un intervallo dei tempi continuo (ad esempio $T_0 = [0, 1]$).

Si possono poi categorizzare come *deterministiche*, quando i valori futuri sono esattamente stabiliti da una funzione matematica dei valori noti, mentre *stocastiche* quando questi possono essere descritti esclusivamente tramite una distribuzione probabilistica. Data la natura incerta dei valori futuri nei casi di cui ci vorremo interessare, ci concentriamo sulle serie storiche stocastiche.

Per fare inferenze, è quindi necessario determinare un modello (o una classe di modelli) per i dati. Chiamiamo *modello della serie storica* il processo stocastico $X = (X_t)_{t \in \mathcal{T}}$ con $T_0 \subseteq \mathcal{T}$, per il quale ciascuna osservazione x_t è la realizzazione di una variabile aleatoria X_t del processo.

In accordo con quanto detto sui processi stocastici, tale modello è *completo* se vengono definite tutte le distribuzioni finito-dimensionali.

Vediamo quindi la definizione formale di serie storica (che spesso chiameremo semplicemente *serie*) da considerare nel corso della discussione.

Definizione 2.14. Una **serie storica** x_1, \dots, x_n è una realizzazione del processo stocastico $X = \{X_t\}_{t \in \mathbb{Z}}$ in una quantità di finita di tempi $(t_1, \dots, t_n) \in \mathbb{Z}^n$, in cui ciascuna x_t è un'osservazione della variabile aleatoria X_t .

Vediamo ora come definire la funzione di autocovarianza e autocorrelazione campionarie a partire dalla serie storica osservata. Queste sono molto utili nei casi reali, in quanto spesso si ha a disposizione una sequenza di osservazioni, ma non il modello sottostante.

Definizione 2.15. Date $\{x_1, \dots, x_n\}$ una sequenza di osservazioni, definiscono:

- **media campionaria** di x_1, \dots, x_n

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t;$$

- **funzione di autocovarianza campionaria**

$$\hat{\gamma}(h) := \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}), \quad -n < h < n;$$

- **funzione di autocorrelazione campionaria**

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}, \quad -n < h < n.$$

Osserviamo che nella definizione precedente non è richiesta la stazionarietà della serie storica; le funzioni di autocovarianza e autocorrelazione campionarie infatti, possono essere calcolate per ogni insieme di osservazioni $\{x_1, \dots, x_n\}$.

Quando si considerano però serie stazionarie, determinare i momenti di primo e secondo ordine campionari è di cruciale importanza per costruire un appropriato modello dei dati. Se infatti $\{x_1, \dots, x_n\}$ è una realizzazione di un processo stazionario X , allora le funzioni di media campionaria e autocorrelazione campionaria sono rispettivamente stimatori per le funzioni di media e di autocorrelazione del processo X . Questi stimatori permettono di identificare quale, tra i possibili processi stazionari è il miglior modello per la rappresentazione dei dati.

Vediamo quindi alcune proprietà di \bar{x} , $\hat{\gamma}(\cdot)$ e $\hat{\rho}(\cdot)$ nel caso di processi stazionari.

Proprietà di \bar{x} : Indichiamo lo stimatore della media μ di un processo stazionario come:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

Innanzitutto è uno stimatore non distorto per μ ; infatti, grazie alla stazionarietà, si ottiene che

$$E[\bar{X}_n] = \frac{1}{n} \sum_{i=1}^n E[X_i] = \mu.$$

Inoltre, l'errore quadratico medio di \bar{X}_n è pari a

$$\begin{aligned} E[(\bar{X}_n - \mu)^2] &= \text{Var}(\bar{X}_n) = \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(X_i, X_j) = \\ &= \frac{1}{n^2} \sum_{i-j=-n}^n (n - |i - j|) \gamma(i - j) = \\ &= \frac{1}{n} \sum_{h=-n}^n \left(1 - \frac{|h|}{n}\right) \gamma(h). \end{aligned}$$

Dai risultati precedenti si deducono alcune conseguenze presentate nella seguente proposizione.

Proposizione 2.3. *Sia $(X_t)_{t \in \mathbb{Z}}$ è un processo stazionario di media μ e funzione di autocovarianza $\gamma(\cdot)$. Allora si ottiene che:*

- $Var(\bar{X}_n) = E[(\bar{X}_n - \mu)^2] \xrightarrow{n \rightarrow \infty} 0$ se $\gamma(h) \rightarrow 0$;
- $nE[(\bar{X}_n - \mu)^2] \xrightarrow{n \rightarrow \infty} \sum_{|h| < \infty} \gamma(h)$ se $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$.

Per poter fare inferenza su μ utilizzando la media campionaria \bar{X}_n , è necessario conoscere la distribuzione (o eventualmente un'approssimazione della distribuzione) di \bar{X}_n .

Proprietà di $\hat{\gamma}(\cdot)$ e $\hat{\rho}(\cdot)$: A differenza dello stimatore della media, gli stimatori $\hat{\gamma}(\cdot)$ e $\hat{\rho}(\cdot)$ di $\gamma(\cdot)$ e $\rho(\cdot)$ sono distorti; ciò nonostante, si assume che $\hat{\gamma}(h)$ abbia distribuzione asintotica con media $\gamma(h)$, ovvero, quando il numero di osservazioni $n \rightarrow \infty$, lo stimatore $\hat{\gamma}(\cdot)$ è non distorto. Un discorso analogo vale per $\hat{\rho}(\cdot)$.

Proposizione 2.4. *La matrice di autocovarianza campionaria*

$$\hat{\Gamma}_n := [\hat{\gamma}(i-j)]_{i,j=1}^n = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \dots & \hat{\gamma}(n-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \dots & \hat{\gamma}(n-2) \\ \vdots & \vdots & \dots & \vdots \\ \hat{\gamma}(n-1) & \hat{\gamma}(n-2) & \dots & \hat{\gamma}(0) \end{bmatrix} \quad (2.3)$$

è semidefinita positiva.

Dimostrazione. Per mostrare questo risultato, scriviamo la matrice $\hat{\Gamma}_n$ come

$$\hat{\Gamma}_n = n^{-1}TT',$$

dove T è la matrice $n \times 2n$

$$T = \begin{bmatrix} 0 & \dots & 0 & 0 & Y_1 & Y_2 & \dots & Y_n \\ 0 & \dots & 0 & Y_1 & Y_2 & \dots & Y_n & 0 \\ \vdots & & & & & & & \vdots \\ 0 & Y_1 & Y_2 & \dots & Y_n & 0 & \dots & 0 \end{bmatrix}$$

e $Y_i = X_i - \bar{X}_n$, $i = 1, \dots, n$.

Preso poi $\mathbf{a} = (a_1, \dots, a_n)'$, con $a_i \in \mathbb{R}$ per $i = 1, \dots, n$, si ottiene che

$$\mathbf{a}'\hat{\Gamma}_n\mathbf{a} = n^{-1}(\mathbf{a}'T)(T'\mathbf{a}) \geq 0.$$

Segue quindi che la matrice $\hat{\Gamma}_n$ è semidefinita positiva. □

Corollario 2.5. *La matrice di autocorrelazione campionaria*

$$\hat{R}_n := [\hat{\rho}(i-j)]_{i,j=1}^n = \frac{\hat{\Gamma}_n}{\gamma(0)}$$

è semidefinita positiva e ciascun elemento diagonale è pari a 1.

Dimostrazione. Sfruttando la dimostrazione della Proposizione 2.4 e $\hat{R}_n = \hat{\Gamma}_n/\gamma(0)$, si ottiene la tesi. \square

Ancora una volta, per poter fare inferenza su $\gamma(h)$ o $\rho(h)$, è necessario conoscere la distribuzione di $\hat{\gamma}(h)$ e $\hat{\rho}(h)$ rispettivamente. Nonostante queste distribuzioni siano particolarmente complicate anche per modelli di serie storiche semplici, si possono approssimare a distribuzioni normali quando il numero n di osservazioni è grande.

La funzione di autocorrelazione campionaria ha un ruolo importante nell'analisi delle serie storiche. Può essere infatti utile per diversi scopi, come per valutare la stazionarietà, le relazioni di dipendenza tra le osservazioni o per selezionare i parametri dei modelli statistici. Maggiori dettagli verranno specificati nel seguito della discussione.

2.4 Analisi di serie storiche

Con il termine *analisi di serie storiche*, si intende la procedura di studio della serie con lo scopo di individuare un processo stocastico che fornisca un modello opportuno per i dati da utilizzare, ad esempio, per fare previsioni.

A tal proposito è bene fare alcune considerazioni preliminari sui dati osservati, per valutarne la stazionarietà e, in qual caso questa non sia verificata, sarà necessario fare opportune trasformazioni per renderli stazionari.

Il primo passo è fare un grafico dei dati per determinare l'eventuale presenza di discontinuità nella serie, di trend, di componenti stagionali, di outliers e di cambi di comportamento improvvisi.

L'obiettivo diventa poi quello di analizzare separatamente i diversi segmenti qualora fosse presente una discontinuità e di rimuovere le eventuali componenti di trend e stagionalità per ottenere un processo stazionario, i cui valori verranno chiamati residui. Questo è possibile grazie alla *decomposizione classica* di una serie storica, che verrà illustrata nella sezione successiva.

Viene quindi scelto un modello per i residui, in modo tale da poter fare previsioni. Infine, applicando la trasformazione inversa a quella fatta inizialmente, ovvero aggiungendo nuovamente le componenti di trend e/o stagionalità che erano state rimosse, si otterrà la previsione per la serie storica di partenza.

2.4.1 Decomposizione di serie storiche

In questa sezione vediamo la prima parte del processo di analisi della serie storica, quindi la valutazione e l'eliminazione di eventuali fenomeni di trend e stagionalità.

Lo studio del grafico dei dati di una serie storica suggerisce la rappresentazione dei dati come la realizzazione di un processo (modello di **decomposizione classica**)

$$X_t = m_t + s_t + Y_t, \quad t \in \mathbb{Z}, \quad (2.4)$$

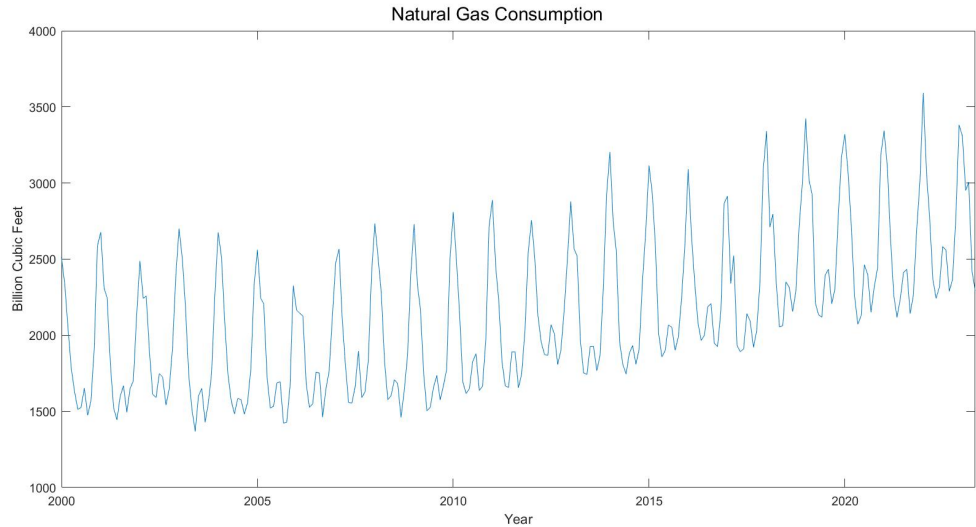


Figura 2.1: Consumo storico mensile di gas naturale

dove, Y_t è un processo stazionario, mentre m_t ed s_t sono funzioni del tempo a valori reali. Più precisamente, m_t è detta **componente di trend** e indica quanto i valori cambiano in media, s_t è la **componente di stagionalità** e descrive il comportamento periodico di periodo d , e Y_t è un processo stazionario chiamato **rumore**.

Prendiamo ad esempio la serie storica del consumo del gas naturale, i cui valori sono stati raccolti dall' U.S. Energy Information Administration (EIA). Dalla Figura 2.1 si osserva che sono presenti sia un trend che una componente stagionale.

L'obiettivo è quello di stimare ed estrarre dalla serie storica le componenti deterministiche m_t e s_t , nella speranza che la componente di rumore Y_t sia un processo stazionario. In questo modo è possibile utilizzare la teoria dei processi stazionari per trovare un modello probabilistico del processo Y_t , analizzarne le proprietà, ed usarlo insieme al trend m_t e alla stagionalità s_t per fare previsioni e simulazioni di X_t .

Eliminazione di trend in assenza di stagionalità

In assenza di stagionalità, il modello di decomposizione classica diventa

$$X_t = m_t + Y_t, \quad t \in \mathbb{Z} \tag{2.5}$$

con $E[Y_t] = 0$.

Si noti che la condizione $E[Y_t] = 0$ non è restrittiva. Infatti, qualora $E[Y_t] \neq 0$, è sufficiente sostituire m_t e Y_t nell'equazione (2.5), rispettivamente con $m_t + E[Y_t]$ e $Y_t - E[Y_t]$.

Vediamo ora alcuni dei possibili metodi per stimare la componente di trend per poi sottrarla al modello e renderlo stazionario.

Metodo T1: (Minimi quadrati)

Prendiamo una serie storica $\{x_1, \dots, x_n\}$ che presenta un trend. Generalmente si possono considerare i trend di forma polinomiale:

$$m_t = \sum_{j=0}^k a_j t^j \quad \text{per ogni } a_j \in \mathbb{R}, \quad j = 1, \dots, k.$$

In questo caso è possibile stimare la componente di trend attraverso la stima ai minimi quadrati, che permette di determinare i valori dei parametri a_j , $j = 0, \dots, k$, tali da minimizzare

$$\sum_{t=1}^n (x_t - m_t)^2.$$

Metodo T2: (Media mobile)

Prendiamo una serie storica $\{x_1, \dots, x_n\}$ che presenta un fenomeno di trend e sia q un intero non negativo. Consideriamo la media mobile del processo X definito da (2.5):

$$W_t = \frac{1}{(2q+1)} \sum_{j=-q}^q X_{t+j}. \quad (2.6)$$

Per $q+1 \leq t \leq n-q$, si ottiene

$$W_t = \frac{1}{(2q+1)} \sum_{j=-q}^q m_{t+j} + \frac{1}{(2q+1)} \sum_{j=-q}^q Y_{t+j} \simeq m_t,$$

assumendo che m_t possa essere approssimata con una funzione lineare su $[t-q, t+q]$ e che, su questo intervallo, la media del processo Y sia vicina a zero. In questo modo, la media mobile permette di stimare la componente di trend:

$$\hat{m}_t = \frac{1}{(2q+1)} \sum_{j=-q}^q X_{t+j}, \quad q+1 \leq t \leq n-q. \quad (2.7)$$

Per $t < q+1$ e $t > n-q$, non è possibile utilizzare (2.7) per stimare la componente di trend, poiché non si hanno osservazioni del processo per $t < 1$ e $t > n$. Una possibilità per ovviare questo problema è definire $X_t := X_1$ per $t < 1$ e $X_t := X_n$ per $t > n$.

Metodo T3: (Differenze di ordine k)

Questo metodo permette di rimuovere la componente di trend grazie alle differenze. Viene quindi introdotto l'operatore alle differenze ∇ che agisce sul processo X tale che

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t, \quad (2.8)$$

dove B è l'**operatore di ritardo**, anch'esso un operatore che agisce sul processo X , che trasforma una sequenza in un'altra, che avrà gli stessi valori, ma sfalsati di un periodo:

$$BX_t = X_{t-1}. \quad (2.9)$$

Si possono poi definire le potenze degli operatori B e ∇ , dove

$$B^j(X_t) = X_{t-j},$$

$$\nabla^j(X_t) = \nabla(\nabla^{j-1}(X_t)) = (1 - B)^j X_t,$$

per $j \geq 1$, con $B^0 = 1$.

Se l'operatore ∇ è applicato a una funzione lineare che stima il trend $m_t = a_0 + a_1 t$, si ottiene una funzione costante $\nabla m_t = a_1$. Nello stesso modo, qualora il trend risulti stimato da un polinomio di grado k , si può applicare l'operatore ∇^k per renderlo una costante.

Ad esempio, se in (2.5) $m_t = \sum_{j=0}^k a_j t^j$ e Y_t è stazionario di media nulla, applicando l'operatore ∇^k , si ottiene

$$\nabla^k X_t = \nabla^k m_t + \nabla^k Y_t.$$

Induttivamente, si può dimostrare che $\nabla^k m_t = k! a_k$. Infatti, come mostrato nelle equazioni successive, ∇m_t è un polinomio di grado $k - 1$ con coefficiente del termine di grado massimo $k a_k$; $\nabla^2 m_t$ è un polinomio di grado $k - 2$ con coefficiente del termine di grado massimo $k(k - 1) a_k$. Proseguendo fino alla derivata k -esima e componendo questi risultati, si verifica che $\nabla^k m_t$ è la costante $k! a_k$.

$$\nabla m_t = a_1 + 2a_2 t + 3a_3 t^2 + \dots + k a_k t^{k-1}$$

$$\nabla^2 m_t = 2a_2 + 6a_3 t + \dots + k(k - 1) a_k t^{k-2}$$

Si ottiene quindi

$$\nabla^k X_t = k! a_k + \nabla^k Y_t, \quad (2.10)$$

che è un processo stazionario di media $k! a_k$.

Questo metodo suggerisce quindi che, data una sequenza di osservazioni $\{x_1, \dots, x_n\}$ si può applicare ripetutamente l'operatore ∇ finché non si ottiene un insieme di osservazioni $\nabla^k x_t$ che si può interpretare come una realizzazione di un processo stazionario. Questa operazione può essere fatta qualsiasi sia la forma del trend, e non solo nel caso polinomiale. Solitamente l'ordine k dell'operatore alle differenze è piuttosto basso, il più delle volte 1 o 2.

Eliminazione di trend e stagionalità

I metodi descritti nella sezione precedente, possono essere adattati nel caso in cui la serie storica presenti anche una componente stagionale.

Riprendiamo il modello di decomposizione classica di una serie storica:

$$X_t = m_t + s_t + Y_t, \quad t \in \mathbb{Z} \quad (2.11)$$

dove $E[Y_t] = 0$, $s_{t+d} = s_t$ e $\sum_{j=1}^d s_j = 0$.

Metodo S1: (Media mobile)

Supponiamo di avere una serie storica $\{x_1, \dots, x_n\}$ che presenta sia un trend che una componente stagionale. Inizialmente utilizziamo la media mobile per stimare il trend: se il periodo stagionale d è dispari la media mobile è centrata e si utilizza (2.7) con $d = 2q + 1$; se invece d è pari, quindi $d = 2q$, la stima del trend si ottiene da:

$$\hat{m}_t = \frac{0.5x_{t-q} + x_{t-q+1} + \dots + x_{t+q-1} + 0.5x_{t+q}}{d}, \quad q < t \leq n - q. \quad (2.12)$$

Il secondo passaggio è quello di stimare la componente stagionale. Per ogni $k = 1, \dots, d$, si calcola la media delle deviazioni

$$w_k = \frac{1}{p} \sum_{j=1}^p (x_{k+d(j-1)} - \hat{m}_{k+d(j-1)}),$$

dove $p = n \div d$ è l'ampiezza temporale di considerazione dei dati.

Dato che non è detto che $\sum_{i=1}^d w_i = 0$, lo stimatore della componente stagionale è dato da

$$\hat{s}_k = w_k - d^{-1} \sum_{i=1}^d w_i, \quad k = 1, \dots, d \quad (2.13)$$

con $\hat{s}_k = \hat{s}_{k-d}$ per $k > d$.

Rimuovendo poi dalla sequenza di osservazioni la componente stagionale, si ottengono dati destagionalizzati, ai quali è possibile applicare uno dei metodi descritti in precedenza (T1, T2 o T3) per determinare lo stimatore del trend. Questa operazione è necessaria per avere una forma parametrica per la componente di trend che possa essere estrapolata dal modello per fare eventuali previsioni o simulazioni.

A questo punto, la stima del processo di rumore è data da

$$\hat{Y}_t = x_t - \hat{m}_t - \hat{s}_t, \quad t = 1, \dots, n \quad (2.14)$$

dove \hat{m}_t è la nuova stima del trend, ottenuta a partire dai dati destagionalizzati.

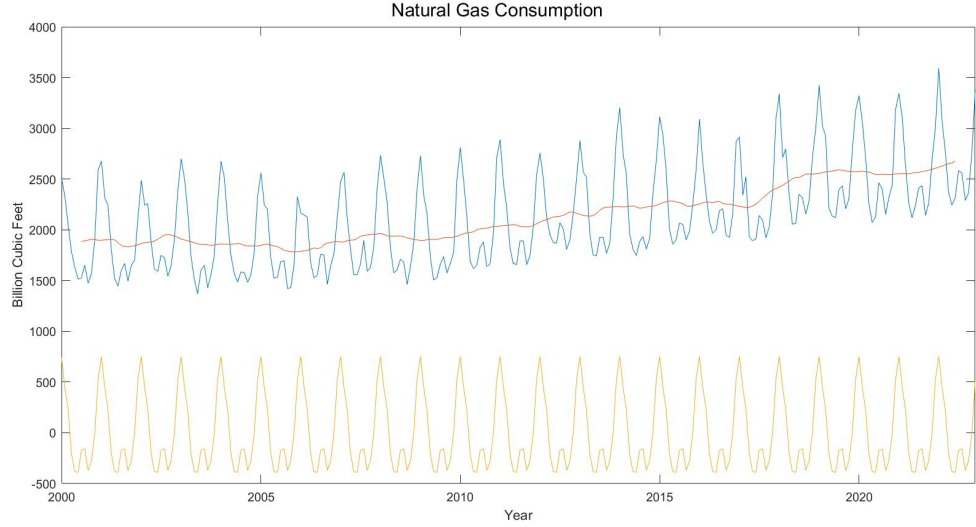


Figura 2.2: Stima delle componenti di trend e stagionalità della serie storica raffigurante il consumo di gas naturale nel periodo 2020-2022

Metodo S2: (Differenze di ordine d)

Il Metodo T3 descritto nella sezione precedente per dati che non presentano stagionalità, si può adattare a dati che presentano una componente stagionale di periodo d , introducendo l'operatore alle differenze ∇_d definito da

$$\nabla_d X_t = X_t - X_{t-d} = (1 - B^d)X_t. \quad (2.15)$$

Applicando l'operatore ∇_d al modello $X_t = m_t + s_t + Y_t$, si ottiene

$$\nabla_d X_t = m_t - m_{t-d} + Y_t - Y_{t-d}.$$

Grazie a questa operazione si ottiene un processo destagionalizzato, dove $(m_t - m_{t-d})$ è la componente di trend, e $(Y_t - Y_{t-d})$ è il processo rumore. Per ottenere un processo stazionario, si può quindi eliminare la componente di trend utilizzando una delle tecniche precedenti, come le potenze dell'operatore ∇ .

Esempio 1. Prendiamo ad esempio il consumo del gas naturale descritto in Figura 2.1. Come già osservato in precedenza, dall'analisi grafica si evidenziano sia una componente stagionale, che un fenomeno di trend, più evidente dal 2010 in avanti. La periodicità è annuale ($d = 12$), dove i picchi di maggior consumo corrispondono ai mesi di dicembre e gennaio.

Vediamo ora come eliminare le componenti di trend e stagionalità seguendo i due metodi illustrati sopra. Indichiamo con x_1, \dots, x_n la sequenza di osservazioni in $t = (t_1, \dots, t_n)$.

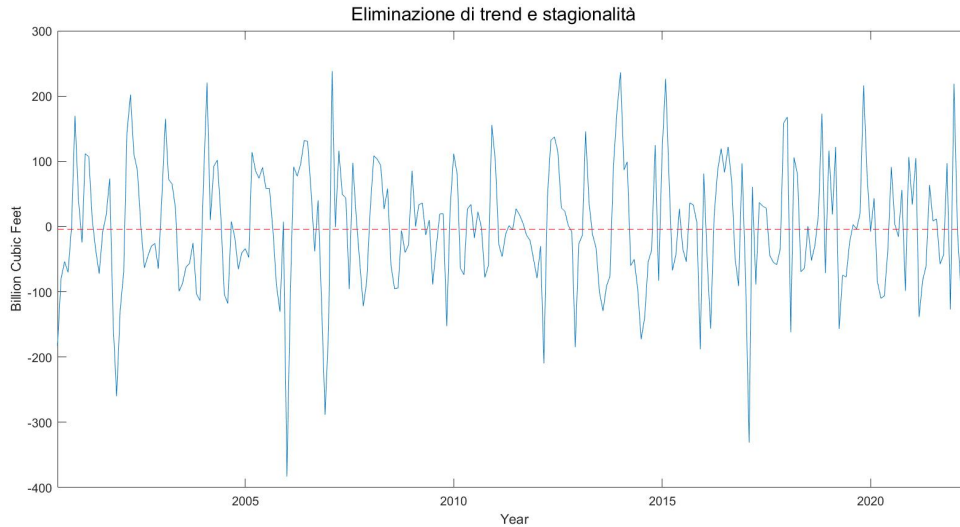


Figura 2.3: Serie del consumo di gas naturale a cui sono state tolte le componenti di trend e periodicità tramite il Metodo S1

Seguendo la tecnica illustrata nel Metodo S1, in particolare le equazioni (2.12) e (2.13), sono stati stimati i due contributi di trend e stagionalità, come mostrato in Figura 2.2. L'ampiezza temporale di considerazione dei dati è $p = 23$ anni.

Dopo aver rimosso la componente periodica e aver ottenuto dati destagionalizzati,

$$d_t = x_t - \hat{s}_t, \quad t = 1, \dots, n$$

è stata stimata nuovamente la componente di trend utilizzando la media mobile per le serie storiche in assenza di stagionalità (Metodo T2). La nuova componente di trend e quella stagionale sono state poi sottratte ai dati iniziali, ottenendo così i residui. La loro rappresentazione in Figura 2.3 non evidenzia più comportamenti stagionali o trend, indice di una buona stima di m_t ed s_t .

In alternativa è possibile utilizzare il Metodo S2. In un primo momento è stato applicato l'operatore ∇_{12} alla serie storica per eliminare la componente stagionale. Seguendo poi il Metodo T3, alle osservazioni destagionalizzate è stato applicato l'operatore ∇ , rimuovendo così anche il trend.

La Figura 2.4 mostra i residui ottenuti applicando gli operatori alle differenze

$$\nabla \nabla_{12} x_t, \quad t = 14, \dots, n$$

e non sono più presenti né componenti stagionali né trend.

I risultati ottenuti con i Metodi S1 ed S2 sono molto simili, ma il Metodo S2 sembra aver rimosso più efficacemente le componenti stagionale e di trend. Avendo quindi ottenuto serie stazionarie, si può poi proseguire con l'analisi della serie storica.

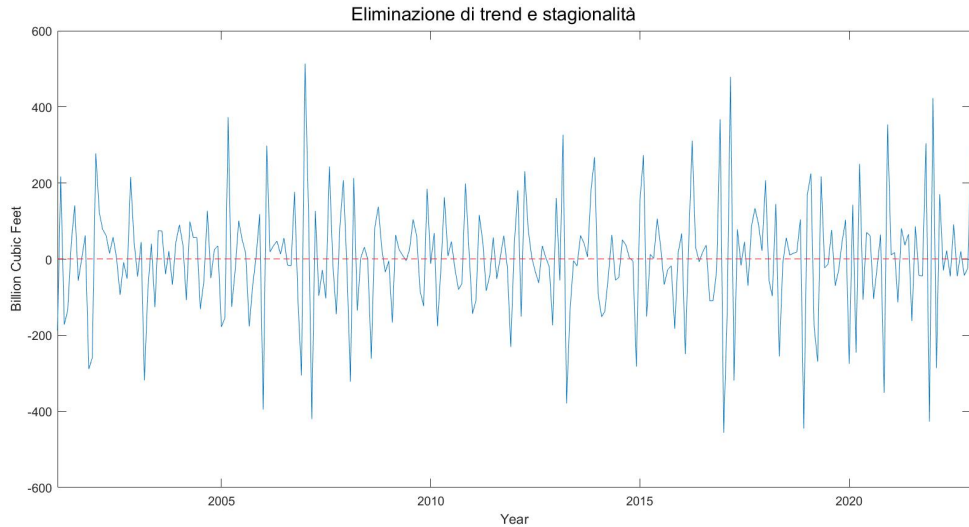


Figura 2.4: Serie del consumo di gas naturale a cui sono state tolte le componenti di trend e periodicità tramite il Metodo S2

2.4.2 Identificazione parametrica

Nel processo di analisi di serie storiche, dopo aver operato un'opportuna trasformazione dei dati in modo da produrre una serie stazionaria senza fenomeni di trend o stagionalità, è bene andare a fare un'analisi sui residui ottenuti ed identificare un modello adeguato.

A seconda della complessità del modello scelto si avranno più o meno parametri da stimare. Qualora questi valori trasformati siano osservazioni di v.a. indipendenti e identicamente distribuite, è sufficiente stimare la loro media e la loro varianza. Il più delle volte questo non accade ed è necessario adottare modelli più complessi, con un numero maggiore di parametri da determinare.

In generale, una delle tecniche principali per la stima parametrica è la *stima di massima verosimiglianza*.

Sia $\{x_1, \dots, x_n\}$ una serie storica e $f(x_1, \dots, x_n, \theta)$ la loro funzione di densità congiunta dove θ è il vettore dei parametri da stimare che varia in un opportuno insieme dei parametri Θ .

Definizione 2.16. Si definisce **funzione di verosimiglianza** la seguente quantità:

$$\mathcal{L}(\theta) := f(x_1, \dots, x_n, \theta). \quad (2.16)$$

Osserviamo che \mathcal{L} per definizione è anche una funzione di (x_1, \dots, x_n) . Per non appesantire la notazione, questa dipendenza non verrà specificata.

Definizione 2.17. Si definisce **stimatore di massima verosimiglianza** $\hat{\theta}_n$ il vettore dei parametri che massimizza la funzione di verosimiglianza:

$$\hat{\theta}_n = \operatorname{argmax}_{\theta \in \Theta} \mathcal{L}(\theta). \quad (2.17)$$

Spesso, invece della funzione di verosimiglianza è vantaggioso utilizzare, e quindi massimizzare, la log-verosimiglianza:

$$l(\theta) := \ln \mathcal{L}(\theta). \quad (2.18)$$

Nel caso di osservazioni indipendenti e identicamente distribuite, lo stimatore di massima verosimiglianza gode di importanti proprietà asintotiche. Innanzitutto, se si conosce la vera funzione di densità $f(\cdot, \theta_0)$, lo stimatore $\hat{\theta}$ converge in probabilità a θ_0 . Inoltre, la distribuzione di $\sqrt{n}(\hat{\theta} - \theta_0)$ si può approssimare con una distribuzione normale di media 0 e matrice di covarianza $I(\theta_0)^{-1}$, dove $I(\theta_0)$ è la matrice di informazione di Fisher, con componente (i, j) pari a

$$(I(\theta))_{i,j} := I_{i,j}(\theta) = E \left[\frac{\partial}{\partial \theta_i} \ln \mathcal{L}_\theta(X) \frac{\partial}{\partial \theta_j} \ln \mathcal{L}_\theta(X) \right]$$

e rappresenta l'informazione fornita dal campione.

2.4.3 Criteri di scelta del modello

Vediamo ora come affrontare la scelta del modello più adeguato ad un dataset. Questa scelta si basa su un compromesso: da una parte è bene scegliere un modello sufficientemente ampio, che si avvicini il più possibile a quello che ha generato le osservazioni; dall'altra il numero di parametri che definiscono il modello non deve essere troppo elevato, per evitare che il modello si adatti eccessivamente ai dati, memorizzando anche dettagli rumori e perdendo la capacità di generalizzare. Questo fenomeno viene chiamato *overfitting*.

Vediamo quindi tre possibili criteri per la scelta del modello.

Dato $X = (X_1, \dots, X_n)$ un vettore di variabili aleatorie, indichiamo con G la funzione di ripartizione congiunta e con g la densità congiunta. Consideriamo poi la famiglia di funzioni di ripartizione $\{F(\cdot; \theta) : \theta \in \Theta\}$ che contiene la vera distribuzione G e supponiamo che per ogni $\theta \in \Theta$ esista la funzione di densità $f(\cdot; \theta)$ relativa ad $F(\cdot; \theta)$.

Il primo approccio possibile per la scelta del modello è l'utilizzo dell'*informazione di Kullback-Leibler* indicata con

$$KLIC_\theta(g : f) := E \left[\ln \frac{g(X)}{f(X; \theta)} \right] = E_g[\ln g(X)] - E_g[\ln f(X; \theta)], \quad (2.19)$$

che misura la discrepanza tra la vera distribuzione G e la distribuzione $F(X; \theta)$. Il criterio è quello di minimizzare questa quantità, trovando quindi θ^* che massimizzi $E_g[\ln f(X; \theta)]$.

Utilizzando la verosimiglianza e cercando di risolvere l'eventuale problema di un eccessivo numero di parametri, è possibile definire un criterio alternativo per la selezione del modello. Viene utilizzata l'*informazione di Akaike (AIC)*, pari a

$$AIC := -2 \ln \mathcal{L}(\hat{\theta}) + 2 * k, \quad (2.20)$$

dove \mathcal{L} è la funzione di verosimiglianza del modello che ha generato i dati, $\hat{\theta}$ è il vettore dei parametri stimati e k è il numero dei parametri del modello.

Il criterio *AIC*, minimizzando la quantità calcolata con (2.20) al variare di k , ha come obiettivo la selezione del modello che massimizzi la verosimiglianza ai dati, ma al contempo penalizzi i modelli che hanno un elevato numero di parametri.

Un altro approccio è dato dal criterio di *informazione Bayesiana (BIC)*. Questo è molto simile al criterio *AIC*, ma ha una penalizzazione più severa per i modelli con un elevato numero di parametri. Questo comporta che l'*AIC* selezioni modelli leggermente più complessi rispetto al *BIC*,

$$BIC := -2 \ln \mathcal{L}(\hat{\theta}) + k * \ln(n) \quad (2.21)$$

dove n che indica il numero delle osservazioni del campione.

2.4.4 Verifica del modello

La procedura di verifica del modello consiste nel valutare la bontà e l'adeguatezza ai dati osservati del modello stimato. Per far ciò bisogna analizzare le stime dei parametri, le stime dei residui e verificare che il numero di parametri utilizzato sia il minore possibile.

Le stime dei parametri devono essere significativamente diverse da zero e talvolta è bene accertarsi che i valori dei coefficienti ottenuti siano coerenti con le aspettative teoriche.

Per valutare la bontà del modello è poi importante andare a studiare il comportamento dei residui in modo opportuno. Dipenderà dalla tipologia del modello in esame e dalle proprietà attese dei residui. Per questa analisi verranno utilizzati diversi metodi, sia quantitativi che grafici. Maggiori dettagli verranno approfonditi con lo studio dei dati nel Capitolo 5.

Infine, per quanto riguarda la parsimonia del modello, ovvero la scelta del modello con un numero non elevato di parametri, vengono utilizzati i criteri di informazione *AIC* e *BIC* presentati in precedenza.

Qualora un modello superi questa fase di verifica, può poi essere utilizzato anche a scopo predittivo.

Capitolo 3

Modelli autoregressivi e a media mobile

In questo capitolo presentiamo alcuni processi elementari per l'analisi di serie storiche stazionarie. L'obiettivo è quello di introdurre i processi ARMA, un'importante famiglia di processi stocastici, spesso utilizzati in econometria.

Introduciamo inizialmente il white noise, uno tra i processi elementari. Segue poi la definizione di due classi di processi: i processi *autoregressivi* (AR) e i processi *a media mobile* (MA), che sono due casi particolari dei processi ARMA. Per entrambi verranno specificate le proprietà principali e le funzioni caratteristiche.

Per quanto riguarda invece i processi ARMA, dopo aver presentato le proprietà principali, i diversi metodi per il calcolo della funzione di autocovarianza e le nozioni di invertibilità e causalità, analizzeremo come caratterizzare un processo ARMA(1,1).

Nel corso del capitolo viene proposto anche un accenno delle possibili generalizzazioni dei processi ARMA, quali i processi ARIMA e i SARIMA.

3.1 White noise

Definizione 3.1. Un processo $(X_t)_{t \in \mathcal{T}}$ si dice **white noise** (o *rumore bianco*) se soddisfa le seguenti condizioni:

- $E[X_t] = 0$ per ogni $t \in \mathcal{T}$;
- $\text{Var}[X_t] = \sigma^2$ per ogni $t \in \mathcal{T}$;
- $\text{Cov}(X_t, X_s) = 0$ per ogni $t, s \in \mathcal{T}$, con $t \neq s$.

Un white noise è quindi un processo composto da un numero infinito di variabili casuali a media zero, varianza costante e tutte incorrelate l'una all'altra.

Segue dalla definizione che questi processi sono stazionari. In particolare, la funzione di autocovarianza è pari a

$$\gamma_X(h) = \begin{cases} \sigma^2 & \text{se } h = 0 \\ 0 & \text{se } h \neq 0, \end{cases} \quad (3.1)$$

mentre la funzione di autocorrelazione è pari a

$$\rho_X(h) = \begin{cases} 1 & \text{se } h = 0 \\ 0 & \text{se } h \neq 0. \end{cases} \quad (3.2)$$

Quindi, per un processo rumore bianco, tutte le funzioni di autocorrelazione (ACF) di lag $h > 0$ sono nulle. In pratica, nel processo di analisi di serie storica, se tutte le ACF campionarie sono vicine a 0, la serie storica è un white noise.

Nel corso della discussione indicheremo un processo white noise $(X_t)_{t \in \mathcal{T}}$ con l'espressione $WN(0, \sigma^2)$.

3.2 Processi a media mobile (MA)

Definizione 3.2. Dato $q \in \mathbb{N} \cup \{\infty\}$, un processo $(X_t)_{t \in \mathbb{Z}}$ si dice **a media mobile di ordine q** o **MA(q)** se è della forma:

$$X_t = \theta_0 Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q} = \sum_{i=0}^q \theta_i Z_{t-i}, \quad t \in \mathbb{Z} \quad (3.3)$$

dove $(Z_t)_{t \in \mathbb{Z}} \sim WN(0, \sigma^2)$, e $\theta_i, i = 0, \dots, q$, sono costanti reali.

Utilizzando l'operatore di ritardo B introdotto nel capitolo precedente ($B^j(X_t) = X_{t-j}$), si può riscrivere la definizione precedente come

$$X_t = \theta(B)Z_t, \quad (3.4)$$

dove θ è un polinomio di grado q dato da

$$\theta(z) = \theta_0 + \theta_1 z + \dots + \theta_q z^q \quad (3.5)$$

e θ_i sono le costanti reali introdotte sopra. Senza perdere di generalità si può prendere $\theta_0 = 1$.

Dalla definizione si può calcolare il valore atteso del processo:

$$E[X_t] = E\left[\sum_{i=0}^q \theta_i Z_{t-i}\right] = \sum_{i=0}^q \theta_i E[Z_{t-i}] = 0.$$

Si verifica quindi che un processo MA ha media nulla. Sfruttando questo risultato e le proprietà del rumore bianco, segue che il processo ha anche momenti secondi finiti.

$$\text{Var}(X_t) = E[X_t^2] = E\left[\left(\sum_{i=0}^q \theta_i Z_{t-i}\right)^2\right] = E\left[\sum_{i=0}^q \theta_i^2 Z_{t-i}^2\right] = \sigma^2 \sum_{i=0}^q \theta_i^2 < +\infty$$

Vediamo ora come calcolare la funzione di autocovarianza:

$$\begin{aligned}
\gamma_X(t, t+h) &= \text{Cov}(X_t, X_{t+h}) = \\
&= \text{Cov}\left(\sum_{j=0}^q \theta_j Z_{t-j}, \sum_{i=0}^q \theta_i Z_{t+h-i}\right) = \\
&= E\left[\left(\sum_{j=0}^q \theta_j Z_{t-j}\right)\left(\sum_{i=0}^q \theta_i Z_{t+h-i}\right)\right] - E\left[\sum_{j=0}^q \theta_j Z_{t-j}\right]E\left[\sum_{i=0}^q \theta_i Z_{t+h-i}\right] = \\
&= E\left[\left(\sum_{j=0}^q \theta_j Z_{t-j}\right)\left(\sum_{i=0}^q \theta_i Z_{t+h-i}\right)\right] = \\
&= \sum_{j=0}^q \sum_{i=0}^q \theta_j \theta_i E[Z_{t-j} Z_{t+h-i}] = \\
&= \begin{cases} \sigma^2 \sum_{j=0}^{q-|h|} \theta_j \theta_{j+|h|} & \text{se } |h| < q \\ 0 & \text{se } |h| > q. \end{cases}
\end{aligned}$$

Dalla forma precedente, si nota che la funzione di autocovarianza è indipendente dal tempo e che quindi un processo MA(q) risulta sempre stazionario.

Per completezza vediamo la forma della funzione di autocorrelazione:

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \begin{cases} \frac{\sum_{j=0}^{q-|h|} \theta_j \theta_{j+|h|}}{\sum_{j=0}^q \theta_j^2} & \text{se } |h| < q \\ 0 & \text{se } |h| > q. \end{cases}$$

Proposizione 3.1. *Se $(X_t)_{t \in \mathbb{Z}}$ è un processo stazionario con media zero e funzione di autocovarianza $\gamma_X(h)$ tale che $\gamma_X(h) = 0$ per $|h| > q$, allora $(X_t)_{t \in \mathbb{Z}}$ è un processo MA(q), ovvero esiste un rumore bianco $(Z_t)_{t \in \mathbb{Z}}$ tale che*

$$X_t = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q}$$

Una dimostrazione del risultato viene proposta in [2, Proposizione 3.2.1].

3.3 Processi autoregressivi (AR)

Definizione 3.3. Dato $p \in \mathbb{N} \cup \{\infty\}$, un processo $(X_t)_{t \in \mathbb{Z}}$ si dice **autoregressivo di ordine p** o **AR(p)** se può essere scritto nella forma seguente:

$$\phi_0 X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + Z_t = \sum_{i=1}^p \phi_i X_{t-i} + Z_t, \quad t \in \mathbb{Z} \quad (3.6)$$

dove $(Z_t)_{t \in \mathbb{Z}} \sim WN(0, \sigma^2)$ e $\phi_i, i = 0, \dots, p$, sono costanti reali.

Come per i processi a media mobile, utilizzando l'operatore di ritardo, si può riscrivere la definizione precedente:

$$Z_t = \phi_0 X_t - \sum_{i=1}^p \phi_i X_{t-i} = \phi(B) X_t,$$

dove ϕ è il polinomio di grado p dato da

$$\phi(z) = \phi_0 - \phi_1 z - \dots - \phi_p z^p. \quad (3.7)$$

Senza perdere di generalità si può prendere $\phi_0 = 1$.

Questi processi forniscono una rappresentazione intuitiva di una serie storica; infatti un valore osservato al tempo t , è una funzione lineare dei propri valori passati più un white noise.

Si può notare come i processi AR siano complementari ai processi MA. Un processo MA infatti è definito dall'applicazione di un polinomio nell'operatore B ad un white noise, mentre un processo AR ha la proprietà che l'applicazione ad esso di un polinomio nell'operatore B , produce un white noise.

È bene evidenziare che, a differenza dei processi MA, che sono sempre stazionari, i processi AR lo sono solamente sotto determinate condizioni.

Proposizione 3.2. *Un processo autoregressivo $AR(p)$, $p \in \mathbb{N}$, è stazionario se e solo se*

$$|\phi_i| < 1 \quad \text{per ogni } i = 1, \dots, p.$$

La dimostrazione nel caso particolare $AR(1)$ si può trovare in [2, Esempio 3.1.2].

Seguendo il testo [2], per evitare passaggi particolarmente complessi, presentiamo ora il calcolo dei momenti e delle funzioni di autocovarianza e autocorrelazione nel caso $AR(1)$ stazionario, sufficiente per il seguito della discussione:

$$X_t = \phi_1 X_{t-1} + Z_t. \quad (3.8)$$

Innanzitutto, facendo il valore atteso da entrambi i membri dell'equazione (3.8), tenendo in considerazione che $E[Z_t] = 0$ perché white noise e che $(X_t)_{t \in \mathbb{Z}}$ è un processo stazionario, si verifica che

$$E[X_t] = 0.$$

Vediamo ora il calcolo della varianza:

$$\text{Var}(X_t) = E[X_t^2] = E[(\phi_1 X_{t-1} + Z_t)^2] = \phi_1^2 \text{Var}(X_t) + \sigma^2 + 2\phi_1 E[X_{t-1} Z_t],$$

da cui

$$\text{Var}(X_t) = \frac{\sigma^2}{1 - \phi_1^2}.$$

Per calcolare la funzione di autocovarianza moltiplichiamo entrambi i membri dell'equazione (3.8) per X_{t-h} e prendiamone i valori attesi. Si ottiene quindi

$$\begin{aligned}
\gamma_X(h) &= \text{Cov}(X_t, X_{t-h}) = \\
&= \text{Cov}(\phi_1 X_{t-1}, X_{t-h}) + \text{Cov}(Z_t, X_{t-h}) = \\
&= E[\phi_1 X_{t-1} X_{t-h}] + 0 = \\
&= \phi_1 \gamma_X(h-1) = \\
&= \phi_1^h \gamma_X(0).
\end{aligned} \tag{3.9}$$

Dato che $\gamma_X(0) = \text{Var}(X_t)$, sostituendo in (3.9) si ottiene:

$$\gamma_X(h) = \phi_1^h \frac{\sigma^2}{1 - \phi_1^2}. \tag{3.10}$$

Avendo richiesto la stazionarietà del processo, risulta semplice il calcolo della funzione di autocorrelazione:

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \phi_1^h.$$

3.4 Processi autoregressivi a media mobile (ARMA)

In questa sezione introduciamo la famiglia dei processi ARMA che, come anticipato, viene spesso utilizzata per l'analisi di serie storiche stazionarie. Infatti, per un'ampia classe di funzioni di autocovarianza $\gamma(\cdot)$, è possibile trovare un processo ARMA $(X_t)_{t \in \mathbb{Z}}$ con funzione di autocovarianza $\gamma_X(\cdot)$ che ben approssima $\gamma(\cdot)$.

Prima di analizzarne le caratteristiche principali sono necessarie alcune definizioni e proprietà di base.

Introduciamo la classe dei processi lineari, di cui fanno parte anche i processi ARMA. Questi sono utili più in generale per lo studio dei processi stazionari; infatti ciascun processo stazionario di second'ordine è un processo lineare o può essere trasformato in un processo lineare sottraendo la componente deterministica (nota come la *decomposizione di Wold* di un processo stazionario (3.11)).

Teorema 3.1. (*Decomposizione di Wold*) Sia $(X_t)_{t \in \mathbb{Z}}$ un processo stazionario non deterministico, allora

$$X_t = \sum_{j=1}^{\infty} \psi_j Z_{t-j} + V_t, \tag{3.11}$$

dove $\psi_0 = 1$, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, $(Z_t)_{t \in \mathbb{Z}} \sim WN(0, \sigma^2)$, $(V_t)_{t \in \mathbb{Z}}$ è la componente deterministica e $\text{Cov}(Z_s, V_t) = 0$ per ogni $s, t \in \mathbb{Z}$.

La dimostrazione di questo risultato si può trovare in [2, Teorema 5.7.1].

Definizione 3.4. Un processo $(X_t)_{t \in \mathbb{Z}}$ si dice **processo lineare** se può essere rappresentato nella forma seguente:

$$X_t = \sum_{j=-\infty}^{+\infty} \psi_j Z_{t-j}, \quad (3.12)$$

dove $(Z_t)_{t \in \mathbb{Z}} \sim WN(0, \sigma^2)$ e ψ_j costante per ogni j con $\sum_{j=-\infty}^{+\infty} |\psi_j| < \infty$.

Osserviamo che la condizione $\sum_{j=-\infty}^{+\infty} |\psi_j| < \infty$ è necessaria per garantire la convergenza in media quadratica di (3.12).

Utilizzando l'operatore di ritardo, l'equazione (3.12) può essere scritta come

$$X_t = \psi(B)Z_t, \quad (3.13)$$

dove $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$.

Osservazione 3.4.1. Un processo lineare è a **media mobile** o **MA**(∞) se $\psi_j = 0$ per ogni $j < 0$, ovvero se

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

Si osserva poi che l'operatore $\psi(B)$, quando applicato ad un processo stazionario, produce un processo ancora stazionario. Formalmente vediamo questo risultato nella proposizione seguente.

Proposizione 3.3. Sia $(Y_t)_{t \in \mathbb{Z}}$ un processo stazionario, con media 0 e funzione di covarianza γ_Y . Se $(\psi_j)_{j \in \mathbb{Z}}$ è una successione di numeri reali tali che $\sum_{j=-\infty}^{+\infty} |\psi_j| < \infty$, allora la serie

$$X_t = \sum_{j=-\infty}^{+\infty} \psi_j Y_{t-j} = \psi(B)Y_t \quad (3.14)$$

è stazionaria con media 0 e funzione di autocovarianza pari a

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_Y(h+k-j). \quad (3.15)$$

Se inoltre $(X_t)_{t \in \mathbb{Z}}$ è un processo lineare, $\gamma_X(h)$ risulta semplicemente

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+h} \sigma^2. \quad (3.16)$$

La dimostrazione di questo risultato si può trovare in [1, Proposizione 2.2.1].

Questo risultato può essere utilizzato per calcolare in modo alternativo la funzione di autocovarianza di un processo lineare.

Definizione 3.5. Presi $p, q \in \mathbb{N}$, un processo $(X_t)_{t \in \mathbb{Z}}$ si dice **autoregressivo a media mobile di ordine (p, q)** o **ARMA (p, q)** se è stazionario ed esistono dei coefficienti $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ tali che per ogni $t \in \mathbb{Z}$ soddisfi:

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q} \quad (3.17)$$

dove $(Z_t)_{t \in \mathbb{Z}} \sim WN(0, \sigma^2)$ e i polinomi ϕ e θ , introdotti dopo le Definizioni 3.2 e 3.3, non hanno zeri in comune.

Utilizzando l'operatore B, l'equazione (3.17) può essere scritta in modo compatto come:

$$\phi(B)X_t = \theta(B)Z_t, \quad \text{per ogni } t \in \mathbb{Z}, \quad (3.18)$$

dove ϕ e θ sono i polinomi di grado p e q dati rispettivamente dalle equazioni (3.7) e (3.5).

Si osserva che i processi MA (q) e AR (p) sono casi elementari del modello ARMA (p, q) , ottenuti rispettivamente se $\phi(z) = 1$ e $\theta(z) = 1$.

Proposizione 3.4. *Un processo ARMA (p, q) definito dall'equazione $\phi(B)X_t = \theta(B)Z_t$, ammette un'unica soluzione stazionaria se e solo se:*

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \neq 0, \quad \text{per ogni } |z| \leq 1. \quad (3.19)$$

Una dimostrazione di questo risultato si può trovare in [2, Remark 3.2].

3.4.1 Causalità e invertibilità

Definizione 3.6. Un processo ARMA (p, q) definito dall'equazione $\phi(B)X_t = \theta(B)Z_t$, si dice **causale** o **funzione causale di $(Z_t)_{t \in \mathbb{Z}}$** , se esistono delle costanti ψ_j tali che $\sum_{j=0}^{\infty} |\psi_j| < \infty$ e

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad \text{per ogni } t \in \mathbb{Z}. \quad (3.20)$$

Teorema 3.2. *Sia $(X_t)_{t \in \mathbb{Z}}$ un processo ARMA (p, q) . Allora $(X_t)_{t \in \mathbb{Z}}$ è causale se e solo se $\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \neq 0$ per ogni $z \in \mathbb{C}$ con $|z| \leq 1$. I coefficienti ψ_j di (3.20) sono determinati da*

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1. \quad (3.21)$$

Una dimostrazione di questo risultato si può trovare in [2, Teorema 3.1.1].

Il concetto duale di causalità è l'*invertibilità*. Infatti un processo $(X_t)_{t \in \mathbb{Z}}$ è causale se X_t può essere espresso in funzione di Z_s con $s \leq t$, mentre è invertibile se Z_t può essere espresso in funzione di X_s con $s \leq t$.

Definizione 3.7. Un processo ARMA(p,q) definito dall'equazione $\phi(B)X_t = \theta(B)Z_t$, si dice **invertibile** se esistono delle costanti π_j tali che $\sum_{j=0}^{\infty} |\pi_j| < \infty$ e

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad \text{per ogni } t \in \mathbb{Z}. \quad (3.22)$$

Teorema 3.3. Sia $(X_t)_{t \in \mathbb{Z}}$ un processo ARMA(p,q). Allora $(X_t)_{t \in \mathbb{Z}}$ è invertibile se e solo se $\theta(z) = 1 - \theta_1 z - \dots - \theta_p z^p \neq 0$ per ogni $z \in \mathbb{C}$ con $|z| \leq 1$. I coefficienti π_j di (3.22) sono determinati da

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1. \quad (3.23)$$

La dimostrazione di questo risultato si può trovare in [2, Teorema 3.1.2].

Si nota che le proprietà di causalità e invertibilità non riguardano solamente il processo $(X_t)_{t \in \mathbb{Z}}$, ma la relazione tra i processi $(X_t)_{t \in \mathbb{Z}}$ e $(Z_t)_{t \in \mathbb{Z}}$ del processo ARMA(p,q).

3.4.2 Metodi per calcolo della funzione di autocovarianza

In questa sezione vediamo due possibili modi per determinare un'espressione della funzione di autocovarianza nel caso di processi ARMA(p,q) causali.

Modo 1. Il primo modo sfrutta l'assunzione di causalità del processo ARMA(p,q), che ne garantisce la linearità e la possibilità di applicare la Proposizione 3.3. Risulta quindi

$$\gamma(h) = E[X_{t+h}X_t] = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|h|}, \quad (3.24)$$

dove

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1. \quad (3.25)$$

Per determinare i coefficienti ψ_j , scriviamo (3.25) nella seguente forma:

$$\psi(z)\phi(z) = \theta(z).$$

Ponendo $\theta_0 = 1$, $\theta_j = 0$ per $j > q$ e $\phi_j = 0$ per $j > p$, uguagliamo i coefficienti di z^j , ottenendo così:

$$\psi_j - \sum_{0 < k \leq j} \phi_k \psi_{j-k} = \theta_j, \quad 0 \leq j < m \quad (3.26)$$

$$\psi_j - \sum_{0 < k \leq p} \phi_k \psi_{j-k} = 0, \quad j \geq m \quad (3.27)$$

dove $m = \max\{p, q + 1\}$.

Da queste equazioni si possono ricavare ricorsivamente ψ_0, ψ_1, \dots . In alternativa è possibile scrivere una soluzione generale per (3.27) come

$$\psi_n = \sum_{i=1}^k \sum_{j=0}^{r_i-1} \alpha_{ij} n^j \xi_i^{-n}, \quad n \geq m - p,$$

dove ξ_i , con $i = 1, \dots, k$ sono gli zeri distinti di $\psi(z)$, mentre r_i è la molteplicità di ξ_i , in modo che $\sum_{i=1}^k r_i = p$. Le p costanti α_{ij} e i coefficienti ψ_j , $0 \leq j < m - p$, sono univocamente determinati da (3.26).

Una volta calcolati tutti i coefficienti ψ_j , si riesce a determinare la funzione di autocovarianza γ .

Maggiori dettagli sui passaggi per arrivare a questa formulazione si possono trovare in [2, Sezione 3.3].

Modo 2. Il secondo modo si basa sulle equazioni alle differenze per $\gamma_X(k)$, con $k = 0, 1, 2, \dots$. Queste sono ottenute moltiplicando entrambi i membri dell'equazione (3.18) per X_{t-k} , $k = 0, 1, \dots$ e considerarne poi il valore atteso:

$$X_t X_{t-k} - \phi_1 X_{t-1} X_{t-k} - \dots - \phi_p X_{t-p} X_{t-k} = Z_t X_{t-k} + \theta_1 Z_{t-1} X_{t-k} + \dots + \theta_q Z_{t-q} X_{t-k}$$

$$\gamma(k) - \phi_1 \gamma(k-1) - \dots - \phi_p \gamma(k-p) = E[Z_t X_{t-k}] + \theta_1 E[Z_{t-1} X_{t-k}] + \dots + \theta_q E[Z_{t-q} X_{t-k}]$$

Dall'ipotesi di causalità e la corrispondente rappresentazione del processo $(X_t)_{t \in \mathbb{Z}}$ (3.20), si ottengono due equazioni:

$$\gamma(k) - \phi_1 \gamma(k-1) - \dots - \phi_p \gamma(k-p) = \sigma^2 \sum_{j=0}^{\infty} \theta_{k+j} \psi_j, \quad 0 \leq k < m, \quad (3.28)$$

$$\gamma(k) - \phi_1 \gamma(k-1) - \dots - \phi_p \gamma(k-p) = 0, \quad k \geq m, \quad (3.29)$$

dove $m = \max\{p, q + 1\}$, $\psi_j = 0$ per ogni $j < 0$, $\theta_0 := 1$ e $\theta_j := 0$ per ogni $j \notin \{0, \dots, q\}$.

Si può poi scrivere la soluzione della (3.29) come:

$$\gamma(h) = \sum_{i=1}^k \sum_{j=0}^{r_i-1} \beta_{ij} h^j \xi_i^{-h}, \quad h \geq m - p, \quad (3.30)$$

dove le p costanti β_{ij} e le covarianze $\gamma_X(j)$, con $0 \leq j < m - p$, sono univocamente determinate dell'equazione (3.28), dopo aver calcolato $\psi_0, \psi_1, \dots, \psi_q$ come mostrato nel Modo 1.

A partire dalla funzione di autocovarianza è poi possibile determinare anche altre due funzioni importanti per un processo ARMA(p,q), ovvero la funzione di autocorrelazione e di autocorrelazione parziale.

Seguendo la Definizione 2.13 la funzione di autocorrelazione parziale di un processo ARMA(p,q), è pari a:

$$\alpha(h) = \begin{cases} 1 & h = 0 \\ \phi_{hh} & h \geq 1, \end{cases} \quad (3.31)$$

dove ϕ_{hh} è l'ultima componente di $\phi_h = \Gamma_h^{-1}\gamma_h$, con $\Gamma_h = [\gamma(i-j)]_{i,j=1}^h$ e $\gamma_h = [\gamma(1), \gamma(2), \dots, \gamma(h)]'$.

Maggiori dettagli si possono trovare in [2, Sezione 3.4].

In parallelo con quanto fatto per i processi AR(1), vediamo qualche dettaglio in più per il caso elementare di processi ARMA(1,1). In particolare mostriamo le condizioni affinché esista una soluzione stazionaria, e affinché il processo sia causale ed invertibile. Infine, vediamo il calcolo dei momenti per il processo ARMA(1,1).

ARMA(1,1) Presi $p = 1$ e $q = 1$, questa classe di processi soddisfa l'equazione:

$$X_t - \phi X_{t-1} = Z_t + \theta Z_{t-1} \text{ per ogni } t \in \mathbb{Z}, \quad (3.32)$$

con $\phi_1 = \phi$ e $\theta_1 = \theta$.

Come sopra, utilizzando l'operatore B, l'equazione (3.32) si può riscrivere in modo compatto:

$$\phi(B)X_t = \theta(B)Z_t, \quad (3.33)$$

dove $\phi(B) = 1 - \phi B$ e $\theta(B) = 1 + \theta B$.

Il processo ARMA(1,1) nasce dall'unione dei processi MA(1) e AR(1). Proprio per questa ragione le proprietà dei modelli ARMA(1,1) sono molto simili a quelle dei modelli AR(1), con qualche differenza dovuta dalla componente MA(1).

Si verifica che esiste un'unica soluzione stazionaria dell'equazione ARMA(1,1) se e solo se $\phi \neq \pm 1$.

- Se $|\phi| < 1$, l'unica soluzione stazionaria dell'equazione è data da:

$$X_t = Z_t + (\phi + \theta) \sum_{j=1}^{\infty} \phi^{j-1} Z_{t-j}. \quad (3.34)$$

Dimostrazione. Sia $\chi(z)$ la serie di potenze $\chi(z) = \sum_{j=0}^{\infty} \phi^j z^j$. Ricordando che $\phi(B) = 1 - \phi B$, si ottiene che $\chi(B)\phi(B) = 1$. Applicando poi $\chi(B)$ a (3.33) e ricordando che $\theta(B) = 1 + \theta B$, risulta

$$X_t = \chi(B)\theta(B)Z_t = \psi(B)Z_t,$$

dove

$$\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j = (1 + \theta B) \sum_{j=0}^{\infty} \phi^j B^j.$$

Osservando che

$$\psi_0 = 1 \quad e \quad \psi_j = (\phi + \theta)\phi^{j-1} \quad \text{per } j \geq 1, \quad (3.35)$$

si ottiene la tesi. \square

In questo caso si evince che il processo $(X_t)_{t \in \mathbb{Z}}$ è causale, dato che X_t può essere espressa in funzione di valori presenti o passati di Z_s , per ogni $s < t$.

- Se $|\phi| > 1$, l'unica soluzione stazionaria dell'equazione è data da:

$$X_t = -\theta\phi^{-1}Z_t - (\phi + \theta) \sum_{j=1}^{\infty} \phi^{-j-1}Z_{t+j}. \quad (3.36)$$

Dimostrazione. Sia $\chi(z)$ la serie di potenze $\chi(z) = -\sum_{j=1}^{\infty} \phi^{-j}z^{-j}$. Utilizzando lo stesso argomento del punto precedente, e applicando $\chi(B)$ a (3.33), si ottiene la tesi. \square

In questo caso il processo $(X_t)_{t \in \mathbb{Z}}$ è non causale, dato che X_t è funzione di Z_s , per ogni $s \geq t$.

Vediamo ora quali sono le condizioni affinché un processo ARMA(1,1) sia invertibile.

- Se $\theta < 1$, il processo ARMA(1,1) è invertibile, dato che Z_t si può esprimere in funzione di X_s , $s \geq t$ come mostrato nella seguente equazione:

$$Z_t = X_t - (\phi + \theta) \sum_{j=1}^{\infty} (-\theta)^{j-1} X_{t-j}. \quad (3.37)$$

- Se $\theta > 1$, il processo ARMA(1,1) è non invertibile, dato che Z_t si può esprimere in funzione di X_s , $s \leq t$ come mostrato di seguito:

$$Z_t = -\phi\theta^{-1}X_t + (\phi + \theta) \sum_{j=1}^{\infty} (-\theta)^{-j-1} X_{t+j}. \quad (3.38)$$

Vediamo ora il calcolo dei momenti per un processo ARMA(1,1).

$$E[X_t] - \phi E[X_{t-1}] = E[Z_t] + \theta E[Z_{t-1}]$$

Sapendo che $E[Z_i] = 0$ per ogni i , la media di $E[X_t]$ è pari a quella del modello AR(1), dunque $E[X_t] = 0$.

Per il calcolo della varianza si sfrutta il valore atteso estratto dall'equazione (3.32) moltiplicata per Z_t

$$E[X_t Z_t] - \phi E[X_{t-1} Z_t] = E[Z_t^2] + \theta E[Z_{t-1} Z_t],$$

da cui $E[X_t Z_t] = \sigma^2$. Dopo aver riscritto il modello ARMA(1,1) come

$$X_t = \phi X_{t-1} + Z_t + \theta Z_{t-1},$$

si calcola la varianza, tenendo conto che $E[X_{t-1} Z_t] = 0$:

$$\text{Var}(X_t) = \phi^2 \text{Var}(X_{t-1}) + \sigma^2 + \theta^2 \sigma^2 + 2\phi\theta E[Z_{t-1} X_{t-1}].$$

Sapendo poi che $E[X_t Z_t] = \sigma^2$ e che il processo è stazionario, si ottiene che

$$\text{Var}(X_t) = \frac{(1 + 2\phi\theta + \theta^2)\sigma^2}{1 - \phi^2}. \quad (3.39)$$

Dal risultato precedente, affinché la varianza sia positiva, segue che $1 - \phi^2 > 0$, quindi $|\phi| < 1$.

Vediamo ora come calcolare la funzione di autocovarianza utilizzando il Modo 1. Questo sfrutta l'assunzione di causalità, che per il processo ARMA(1,1) vale quando $|\phi| < 1$. Sapendo poi che per questo caso valgono (3.34) e (3.35), si ottengono i seguenti risultati:

$$\begin{aligned} \gamma(0) &= \sigma^2 \sum_{j=0}^{\infty} \psi_j^2 = \\ &= \sigma^2 \left[1 + (\theta + \phi)^2 \sum_{j=0}^{\infty} \phi^{2j} \right] = \\ &= \sigma^2 \left[1 + \frac{(\theta + \phi)^2}{1 - \phi^2} \right], \end{aligned}$$

$$\begin{aligned} \gamma(1) &= \sigma^2 \sum_{j=0}^{\infty} \psi_{j+1} \psi_j = \\ &= \sigma^2 \left[\theta + \phi + (\theta + \phi)^2 \phi \sum_{j=0}^{\infty} \phi^{2j} \right] = \\ &= \sigma^2 \left[\theta + \phi + \frac{(\theta + \phi)^2 \phi}{1 - \phi^2} \right]. \end{aligned}$$

Con ulteriori passaggi, sostituendo (3.35) in (3.24), si può verificare che in generale vale

$$\gamma(h) = \phi^{h-1} \gamma(1), \quad h \geq 2.$$

3.5 Generalizzazioni dei processi ARMA

I processi introdotti fino ad ora, vengono per lo più utilizzati per generare serie storiche stazionarie. Tuttavia, la maggior parte delle serie storiche economiche, mostrano un comportamento non stazionario.

Vediamo quindi due classi di processi da poter utilizzare in questi casi. Si tratta di generalizzazioni dei processi ARMA, ovvero i processi ARIMA e SARIMA, che riescono a cogliere la non stazionarietà delle serie storiche.

Definizione 3.8. Presi $p, q, d \in \mathbb{N}$, un processo $(X_t)_{t \in \mathbb{Z}}$ si dice **autoregressivo integrato a media mobile** o **ARIMA(p,d,q)**, se il processo

$$Y_t := (1 - B)^d X_t$$

è un processo ARMA(p,q) causale, dove B è l'operatore di ritardo.

L'attributo *integrato*, segue dal fatto che, se $Y_t = (1 - B)^d X_t$, allora il processo $(X_t)_{t \in \mathbb{Z}}$ si può ottenere dall'espressione inversa $X_t = (1 - B)^{-d} Y_t$, ossia come somme ripetute di $(Y_t)_{t \in \mathbb{Z}}$, che, per processi a tempo continuo, corrisponde ad un integrale.

Osservazione 3.5.1. Un processo ARIMA è stazionario se e solo se $d = 0$, caso in cui si riduce ad un processo ARMA(p,q).

Una caratteristica distintiva delle osservazioni di una serie storica provenienti da un processo ARIMA è che la loro funzione di autocorrelazione presenta un lento decadimento positivo.

Una classe più generale dei processi ARIMA, sono i processi SARIMA, utili nel caso in cui una serie storica presenti anche una componente stagionale.

Definizione 3.9. Presi $d, D, p, P, q, Q \in \mathbb{N}$, un processo $(X_t)_{t \in \mathbb{Z}}$ si dice **autoregressivo integrato a media mobile stagionale** o **SARIMA(p,d,q) × (P,D,Q)_s** di periodo (o stagionalità) s , se $Y_t := (1 - B)^d (1 - B^s)^D X_t$ è un processo ARMA causale definito da:

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)Z_t \quad (Z_t)_{t \in \mathbb{Z}} \sim WN(0, \sigma^2),$$

dove $\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$, $\Phi(z) = 1 - \Phi_1 z - \dots - \Phi_P z^P$, $\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$ e $\Theta(z) = 1 + \Theta_1 z + \dots + \Theta_Q z^Q$.

Come si evince dalla definizione questa classe di processi sono rappresentabili come processi ARIMA, ma con ritardi pari alla stagionalità.

Capitolo 4

Modelli a volatilità stocastica

In questo capitolo presentiamo alcuni modelli per l'analisi di serie storiche finanziarie e, più in generale, per serie temporali caratterizzate da varianza non costante nel tempo. Si tratta di modelli per la *volatilità*, ovvero la variabilità di un valore o di un indice finanziario in un determinato intervallo di tempo. Analiticamente la volatilità si identifica nel concetto di varianza condizionale, ovvero nel modo in cui la varianza varia nel tempo.

Si è osservato che i dati finanziari, come ad esempio i prezzi di un bene nel tempo (in notazione $P = (P_t)_{t \in \mathbb{N}}$), non hanno un comportamento stazionario. Tuttavia, la successione dei ritorni logaritmici (chiamati anche rendimenti) definiti come differenza logaritmica del prezzo del bene al tempo t e al tempo $t - 1$, $\{Y_t := \log P_t - \log P_{t-1}\}$, mostra un andamento che si può modellizzare con processi stazionari. In questo capitolo indicheremo generalmente con $Y = (Y_t)_{t \in \mathbb{N}}$ il processo dei rendimenti e con \mathcal{F} la filtrazione naturale associata ($\mathcal{F}_t = \sigma(Y_t, Y_{t-1}, \dots)$).

Empiricamente, si è poi visto che le serie storiche di tipo prezzo presentano usualmente alcune caratteristiche peculiari, note come *fatti stilizzati*.

- **Leptocurtosi:** Si è osservato che le distribuzioni di probabilità delle serie finanziarie sono leptocurtiche, ovvero hanno la particolarità di assegnare una probabilità maggiore ad eventi lontani dal valor medio rispetto alle probabilità che verrebbero assegnate a tali eventi da una distribuzione Normale. Spesso queste distribuzioni vengono dette *distribuzioni con code spesse* o *with heavy tails* il che riflette la concentrazione maggiore della massa di probabilità sulle code rispetto ad una distribuzione Normale.
- **Effetto leverage:** Questa caratteristica riflette il comportamento asimmetrico dei prezzi, in base al quale variazioni di prezzi negative incrementano la volatilità più di quanto non facciano variazioni positive.

- **Aggregational Gaussianity:** Aumentando la scala dei tempi Δt su cui i ritorni logaritmici di serie storiche finanziarie sono calcolati, la loro distribuzione mostra un comportamento Normale. In particolare, la forma della distribuzione cambia a seconda della scala temporale.
- **Persistenza della volatilità:** La volatilità dei rendimenti di una serie finanziaria è fortemente dipendente dai valori passati. Questo significa che i valori assunti dalla varianza condizionale sono in relazione con i valori assunti in periodi di osservazione precedenti. Numericamente questa caratteristica si vede dal fatto che i quadrati dei rendimenti non sono indipendenti.
- **Volatility clustering:** Questo fenomeno spiega come periodi di elevata volatilità tendono a permanere nel tempo e sono seguiti da periodi di relativa stabilità che a loro volta manifestano una certa persistenza.

Queste ultime due caratteristiche, riflettono il comportamento **eteroschedastico** delle serie storiche. Queste infatti si possono dividere in due classi in base al comportamento della varianza. Nei capitoli precedenti si sono considerati solamente modelli *omoschedastici*, che presentano varianza costante nel tempo. Questa assunzione è piuttosto restrittiva. Infatti, la maggior parte delle serie storiche mostra un comportamento *eteroschedastico*, ossia con varianza caratterizzata da un comportamento non costante nel tempo.

Negli anni sono stati quindi sviluppati diversi modelli a tempo discreto per riflettere queste caratteristiche delle serie storiche finanziarie. In particolare, in questo capitolo presentiamo il modello autoregressivo ad eteroschedasticità condizionata (ARCH) di Engle (1982), il modello ARCH generalizzato (GARCH) di Bollerslev (1986) ed infine il modello GARCH esponenziale (EGARCH) di Nelson (1991).

4.1 Processi ARCH

L'economista Robert F. Engle [4], nel 1982, introdusse i processi ARCH, la prima classe di processi che generano serie storiche eteroschedastiche. Questi processi sono definiti a media nulla, serialmente incorrelati, con una varianza condizionale non costante basata sui valori passati e una varianza non condizionale costante.

Vengono spesso utilizzati come modelli per la serie dei ritorni logaritmici di serie finanziarie. Infatti, una delle caratteristiche principali della classe dei processi ARCH, è la capacità di cogliere il fenomeno del volatility clustering. Questo è dovuto al fatto che, nel modello proposto da Engle, la varianza condizionata dipende dal quadrato dei valori precedenti.

Notazione: Nel corso del capitolo viene utilizzata spesso la notazione $\{e_t\} \sim \text{IID}(0, 1)$ per indicare un processo stocastico di v.a. indipendenti, identicamente distribuite, con media 0 e varianza 1.

Vediamo ora la definizione formale di un processo ARCH:

Definizione 4.1. Un processo $(Y_t)_{t \in \mathbb{Z}}$ si dice **autoregressivo a eteroschedasticità condizionata di ordine q** o in breve **ARCH(q)**, con $q \in \mathbb{N}^*$, se soddisfa la seguente equazione:

$$Y_t = \sigma_t e_t \quad (4.1)$$

dove

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i Y_{t-i}^2 \quad (4.2)$$

con $\alpha_0 > 0$, $\alpha_j \geq 0$ per $j = 1, \dots, q$ e $\{e_t\} \sim \text{IID}(0, 1)$.

Utilizzando l'operatore di ritardo B , l'espressione per σ_t^2 può essere riscritta nel seguente modo:

$$\sigma_t^2 = \alpha_0 + \alpha(B)Y_t^2 \quad (4.3)$$

dove α è il polinomio $\alpha(x) = \alpha_1 x + \dots + \alpha_q x^q$.

Notiamo che la positività dei parametri $\alpha_0 > 0$ e $\alpha_j \geq 0$ per $j = 1, \dots, q$, è necessaria per garantire la non negatività di σ_t^2 , che vedremo poi essere la varianza condizionale del processo.

Per quanto riguarda invece il processo $\{e_t\}$, viene scelto tipicamente un processo gaussiano o un processo con distribuzione di Student con ν gradi di libertà, più precisamente

$$\sqrt{\frac{\nu}{\nu-2}} e_t \sim t_\nu, \quad \text{con } \nu > 2.$$

Dalla definizione si deducono le principali proprietà di un processo ARCH(q).

- La media condizionale di Y_t rispetto a \mathcal{F}_{t-1} e la media non condizionale di Y_t sono nulle per ogni t :

$$E[Y_t | \mathcal{F}_{t-1}] = E[\sigma_t e_t | \mathcal{F}_{t-1}] = \sigma_t E[e_t | \mathcal{F}_{t-1}] = \sigma_t E[e_t] = 0; \quad (4.4)$$

$$E[Y_t] = E[E[Y_t | \mathcal{F}_{t-1}]] = 0. \quad (4.5)$$

- Il processo è serialmente incorrelato condizionatamente ad un set informativo \mathcal{F}_{t-h} , ovvero per ogni $h \geq 1$, $k \geq 1$,

$$\text{Cov}(Y_t Y_{t+k} | \mathcal{F}_{t-h}) = E[Y_t Y_{t+k} | \mathcal{F}_{t-h}] - E[Y_t | \mathcal{F}_{t-h}] E[Y_{t+k} | \mathcal{F}_{t-h}],$$

e, grazie alla proprietà della torre e a (4.4), si ottiene:

$$E[Y_t | \mathcal{F}_{t-h}] = E[E[Y_t | \mathcal{F}_{t-1}] | \mathcal{F}_{t-h}] = 0,$$

da cui

$$\begin{aligned} \text{Cov}(Y_t Y_{t+k} | \mathcal{F}_{t-h}) &= E[Y_t Y_{t+k} | \mathcal{F}_{t-h}] = \\ &= E[E[Y_t Y_{t+k} | \mathcal{F}_{t+k-1}] | \mathcal{F}_{t-h}] = \\ &= E[Y_t E[Y_{t+k} | \mathcal{F}_{t+k-1}] | \mathcal{F}_{t-h}] = \\ &= 0. \end{aligned} \quad (4.6)$$

- σ_t^2 è l'espressione analitica della varianza condizionata del processo Y . Infatti,

$$\text{Var}(Y_t|\mathcal{F}_{t-1}) = E[Y_t^2|\mathcal{F}_{t-1}] = E[\sigma_t^2 e_t^2|\mathcal{F}_{t-1}] = \sigma_t^2 E[e_t^2|\mathcal{F}_{t-1}] = \sigma_t^2 E[e_t^2] = \sigma_t^2. \quad (4.7)$$

Affinché questa classe di processi sia utile per spiegare il comportamento eteroschedastico di una serie storica, deve esistere almeno un $i \in (1, \dots, q)$ tale che $\alpha_i > 0$. In questo modo, la varianza condizionata non è costante nel tempo, ma dipende dalla storia passata.

Da queste osservazioni si evince che, qualora il processo $\{e_t\}$ fosse gaussiano, la distribuzione condizionata di un processo di tipo ARCH è Normale:

$$Y_t|\mathcal{F}_{t-1} \sim N(0, \sigma_t^2). \quad (4.8)$$

- I momenti dispari della distribuzione di processi ARCH sono nulli:

$$E[Y_t^m] = E[E[Y_t^m|\mathcal{F}_{t-1}]] = 0 \quad \text{per ogni } m \text{ dispari}. \quad (4.9)$$

- Y_t è leptocurtica ([3, Sezione 3.4.1]). Si può dimostrare infatti che la distribuzione non condizionata del processo $\{Y_t\}$ non è Normale, ma ha un indice di curtosi maggiore di 3:

$$K = \frac{E[Y_t^4]}{[E[Y_t^2]]^2} \geq 3. \quad (4.10)$$

Da questa proprietà segue che una caratteristica delle serie finanziarie catturata dai processi ARCH è proprio la leptocurtosi.

Una verifica di questo risultato nel caso particolare di un processo ARCH(1) è mostrata nell'*Esempio 2*.

- I quadrati delle osservazioni di un processo ARCH(q) sono un processo AR(q) ([1, Sezione 7.2]). Definendo infatti $v_t := Y_t^2 - \sigma_t^2$, ossia come la differenza tra Y_t^2 e la sua media condizionata, si ottiene:

$$Y_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i Y_{t-i}^2 + v_t, \quad (4.11)$$

dove si può dimostrare che v_t è un white noise.

Una verifica di questo risultato nel caso particolare di un processo ARCH(1) è mostrata nell'*Esempio 2*.

- Un'altra proprietà di un processo ARCH(q) è la non linearità, che si riferisce al modo in cui Y_t dipende dalle osservazioni precedenti dello stesso processo. Come si osserva chiaramente sostituendo (4.2) in (4.1), si ottiene:

$$Y_t = e_t \sqrt{\alpha_0 + \sum_{i=1}^q \alpha_i Y_{t-i}^2}.$$

Teorema 4.1. *Un processo ARCH(q) è debolmente stazionario se e solo se*

$$\sum_{i=1}^q \alpha_i < 1. \quad (4.12)$$

Se la condizione di stazionarietà è verificata, la varianza non condizionale di Y_t è data da:

$$\text{Var}(Y_t) = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i}. \quad (4.13)$$

La dimostrazione del Teorema si può trovare in [4, Teorema 2].

Vediamo ora quanto detto nel caso elementare di processi ARCH(1) stazionari. Poniamo particolare attenzione a questa classe di processi poiché, nel seguito dell'elaborato utilizzeremo processi ARCH(1) per l'analisi della serie storica.

Esempio 2. ARCH(1) Stazionari

Questa classe di processi ha la varianza condizionale definita da

$$\text{Var}(Y_t | \mathcal{F}_{t-1}) = \sigma_t^2 = \alpha_0 + \alpha_1 Y_{t-1}^2, \quad (4.14)$$

in cui $\alpha_0 > 0$ e $0 \leq \alpha_1 < 1$ sono le restrizioni sui parametri necessarie per garantire la positività della varianza e la stazionarietà.

Dal Teorema 4.1 la varianza non condizionale è pari a

$$E[Y_t^2] = \frac{\alpha_0}{1 - \alpha_1}. \quad (4.15)$$

Vediamo ora come trovare l'indice di curtosi e dimostrare (4.10). Prima di tutto calcoliamo il momento di quarto ordine:

$$\begin{aligned} E[Y_t^4] &= E[E[Y_t^4 | \mathcal{F}_{t-1}]] = \\ &= E[E[\sigma_t^4 e_t^4 | \mathcal{F}_{t-1}]] = \\ &= E[\sigma_t^4 E[e_t^4]] = \\ &= E[3\sigma_t^4] = \\ &= 3E[(\alpha_0 + \alpha_1 Y_{t-1}^2)^2] = \\ &= 3(E[\alpha_0^2] + E[\alpha_1^2 Y_{t-1}^4] + 2E[\alpha_0 \alpha_1 Y_{t-1}^2]) = \\ &= 3(\alpha_0^2 + \alpha_1^2 E[Y_{t-1}^4] + 2\alpha_0 \alpha_1 E[Y_{t-1}^2]). \end{aligned}$$

Sostituendo poi il valore della varianza non condizionata e tenendo conto dell'uguaglianza dei momenti di quarto ordine grazie alla stazionarietà, si ottiene

$$(1 - 3\alpha_1^2)E[Y_t^4] = 3\alpha_0^2 + 6\frac{\alpha_0^2 \alpha_1}{1 - \alpha_1} = \frac{3\alpha_0^2(1 + \alpha_1)}{1 - \alpha_1},$$

da cui

$$E[Y_t^4] = \frac{3\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}.$$

L'indice di curtosi è quindi pari a

$$K = \frac{E[Y_t^4]}{(E[Y_t^2])^2} = \frac{3(1 - \alpha_1^2)}{1 - 3\alpha_1^2}. \quad (4.16)$$

Notiamo che è necessario richiedere $3\alpha_1^2 < 1$ per garantire che il momento di quarto ordine sia finito. Sotto questa condizione, l'indice di curtosi (4.16) esiste, ed è maggiore di 3, il che rispecchia la caratteristica di leptocurtosi delle serie finanziarie.

Come mostrato in (4.11), si deduce che il quadrato del processo ARCH(1) può essere rappresentato come un processo AR(1), dato dalla forma:

$$Y_t^2 = \alpha_0 + \alpha_1 Y_{t-1}^2 + v_t. \quad (4.17)$$

Verifichiamo quindi che v_t è effettivamente un white noise. Innanzitutto:

$$\begin{aligned} E[v_t] &= E[Y_t^2 - \sigma_t^2] = \\ &= E[\sigma_t^2(e_t^2 - 1)] = \\ &= E[E[\sigma_t^2(e_t^2 - 1)|\mathcal{F}_{t-1}]] = \\ &= E[\sigma_t^2 E[e_t^2 - 1]] = \\ &= 0. \end{aligned}$$

Poi, sfruttando il calcolo del momento di quarto ordine svolto precedentemente, si ricava che la varianza di v_t è costante; infatti:

$$\text{Var}(v_t) = E[v_t^2] = E[(Y_t^2 - \sigma_t^2)^2] = E[Y_t^4] + E[\sigma_t^4] - 2E[Y_t^2 \sigma_t^2] = \frac{2\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}.$$

Infine si verifica che $\text{Cov}(v_t, v_{t-h}) = 0$, che conclude la dimostrazione che v_t è un white noise.

$$\begin{aligned} \text{Cov}(v_t, v_{t-h}) &= E[v_t v_{t-h}] - E[v_t]E[v_{t-h}] = \\ &= E[v_t v_{t-h}] = \\ &= E[(Y_t^2 - \sigma_t^2)v_{t-h}] = \\ &= E[E[(\sigma_t^2 e_t^2 - \sigma_t^2)v_{t-h}|\mathcal{F}_{t-1}]] = \\ &= E[E[\sigma_t^2(e_t^2 - 1)v_{t-h}|\mathcal{F}_{t-1}]] = \\ &= E[\sigma_t^2 v_{t-h} E[e_t^2 - 1|\mathcal{F}_{t-1}]] = \\ &= E[\sigma_t^2 v_{t-h}(1 - 1)] = \\ &= 0. \end{aligned}$$

4.2 Processi GARCH

I processi ARCH(q) presentati nella sezione precedente, spesso richiedono un ordine q alto per descrivere la struttura di persistenza della volatilità. Questo comporta che il numero di parametri $\alpha_0, \dots, \alpha_q$ da stimare sia piuttosto elevato, cosa non opportuna quando si vuole scegliere un modello econometrico.

A tal proposito, una soluzione è stata proposta da Tim Bollerslev [5] che, nel 1986, introduce i processi GARCH. Questa classe di processi, pur essendo basata su un numero limitato di parametri, permette di riprodurre situazioni di lunga memoria per la volatilità.

Definizione 4.2. Un processo $(Y_t)_{t \in \mathbb{Z}}$ si dice **autoregressivo a eteroschedasticità condizionata generalizzato di ordine p e q** o in breve **GARCH(p, q)**, con $p, q \in \mathbb{N}^*$, se soddisfa la seguente equazione:

$$Y_t = \sigma_t e_t \quad (4.18)$$

dove

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i Y_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (4.19)$$

con $\alpha_0 > 0$, $\alpha_j, \beta_j \geq 0$ per $j = 1, 2, \dots$ e $\{e_t\} \sim \text{IID}(0, 1)$.

Utilizzando l'operatore di ritardo B , l'espressione per σ_t^2 può essere riscritta nel seguente modo:

$$\sigma_t^2 = \alpha_0 + \alpha(B)Y_t^2 + \beta(B)\sigma_t^2 \quad (4.20)$$

dove α e β sono i polinomi $\alpha(x) = \alpha_1 x + \dots + \alpha_q x^q$ e $\beta(x) = \beta_1 x + \dots + \beta_p x^p$.

Analogamente al caso dei processi ARCH, le condizioni di positività dei parametri garantiscono la non negatività di σ_t^2 e il processo $\{e_t\}$ è tipicamente gaussiano o distribuito come una Student.

Questa classe di processi garantisce una maggior flessibilità rispetto ad un processo ARCH(q), dato che la formulazione di σ_t^2 , che mostreremo essere proprio la varianza condizionata, comprende anche una componente autoregressiva data da p ritardi della varianza condizionata stessa.

Osservazione 4.2.1. Un processo ARCH(q) si ottiene come caso particolare di un processo GARCH(p, q), prendendo $p = 0$. Inoltre, se $p = q = 0$, il processo che si ottiene è un semplice white noise.

Vediamo ora le principali proprietà di un processo GARCH(p, q):

- Come nel caso ARCH(q), la media condizionata di Y_t è nulla e la varianza condizionata coincide con σ_t^2 :

$$E[Y_t | \mathcal{F}_{t-1}] = 0; \quad (4.21)$$

$$\text{Var}(Y_t | \mathcal{F}_{t-1}) = \sigma_t^2. \quad (4.22)$$

Inoltre, per definizione, risulta che nel caso in cui il processo $\{e_t\}$ fosse gaussiano, la distribuzione condizionale di un processo di tipo GARCH è Normale, quindi:

$$Y_t | \mathcal{F}_{t-1} \sim N(0, \sigma_t^2) \quad (4.23)$$

- Y_t è leptocurtica ([3, Sezione 3.5]). Analogamente al caso precedente, la distribuzione non condizionata di un processo GARCH(p,q) non è Normale. Si può infatti verificare che l'indice di curtosi K è maggiore di 3:

$$K = \frac{E[Y_t^4]}{[E[Y_t^2]]^2} \geq 3. \quad (4.24)$$

La verifica di questo risultato nel caso particolare di un processo GARCH(1,1) è mostrata nell'*Esempio 3*.

- I quadrati delle osservazioni di un processo GARCH(p,q) sono un processo ARMA(m,p), dove $m = \max\{p, q\}$ ([3, Sezione 3.5]). Definiamo il processo $v_t := Y_t^2 - \sigma_t^2$, che si può dimostrare essere un white noise. Grazie a questa espressione, si ottiene la seguente forma per i quadrati delle osservazioni del processo Y :

$$\begin{aligned} Y_t^2 &= \sigma_t^2 + v_t = \\ &= \alpha_0 + \sum_{i=1}^q \alpha_i Y_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + v_t = \\ &= \alpha_0 + \sum_{i=1}^q \alpha_i Y_{t-i}^2 + \sum_{j=1}^p \beta_j Y_{t-j}^2 - \sum_{j=1}^p \beta_j v_{t-j} + v_t = \\ &= \alpha_0 + \sum_{i=1}^m (\alpha_i + \beta_i) Y_{t-i}^2 - \sum_{j=1}^p \beta_j v_{t-j} + v_t. \end{aligned} \quad (4.25)$$

La verifica di questo risultato nel caso particolare di un processo GARCH(1,1) è mostrata nell'*Esempio 3*.

Vediamo ora come caratterizzare i processi GARCH(p,q) stazionari.

Teorema 4.2. *Un processo GARCH(p,q) è stazionario se e solo se*

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1. \quad (4.26)$$

Sotto la condizione di stazionarietà, la varianza non condizionale del processo $\{Y_t\}$ è pari a:

$$\text{Var}(Y_t) = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j}. \quad (4.27)$$

La dimostrazione del Teorema si può trovare in [5, Teorema 1].

Vediamo ora il caso elementare di processi GARCH(1,1) stazionari. Infatti, proprio come per i processi ARCH(1), utilizzeremo i processi GARCH(1,1) per l'analisi dei dati nel Capitolo 5.

Esempio 3. GARCH(1,1) Stazionari

La varianza condizionale di questa classe di processi è definita da:

$$\text{Var}(Y_t|\mathcal{F}_{t-1}) = \sigma_t^2 = \alpha_0 + \alpha_1 Y_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

dove $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$ e $\alpha_1 + \beta_1 < 1$, affinché sia garantita la condizione di stazionarietà.

Dal Teorema 4.2 segue che la varianza non condizionale risulta pari a

$$\text{Var}(Y_t) = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)}, \quad (4.28)$$

dove $(\alpha_1 + \beta_1)$ è il fattore di persistenza della volatilità.

Si verifica poi facilmente che i momenti condizionali e non condizionali di primo ordine sono nulli, infatti:

$$E[Y_t] = E[E[Y_t|\mathcal{F}_{t-1}]] = 0.$$

Per la simmetria della distribuzione, i momenti di terzo ordine sono nulli:

$$E[Y_t^3] = E[E[Y_t^3|\mathcal{F}_{t-1}]] = 0.$$

Questo risultato viene poi generalizzato da Bollerslev [5]: un processo GARCH per il quale esiste il momento di ordine $2m$, $m \in \mathbb{N}$, ha momento centrale di ordine $2m - 1$ nullo.

Vediamo ora come calcolare il momento di quarto ordine per i processi GARCH(1,1):

$$\begin{aligned} E[Y_t^4] &= E[E[Y_t^4|\mathcal{F}_{t-1}]] = \\ &= E[E[\sigma_t^4 e_t^4|\mathcal{F}_{t-1}]] = \\ &= E[\sigma_t^4 E[e_t^4]] = \\ &= E[3\sigma_t^4] = \\ &= 3E[(\alpha_0 + \alpha_1 Y_{t-1}^2 + \beta_1 \sigma_{t-1}^2)^2] = \\ &= 3\alpha_0 + 3\alpha_1^2 E[Y_{t-1}^4] + 3\beta_1^2 E[\sigma_{t-1}^4] + 6\alpha_0\alpha_1 E[Y_{t-1}^2] + 6\alpha_0\beta_1 E[\sigma_{t-1}^2] + 6\alpha_1\beta_1 E[Y_{t-1}^2\sigma_{t-1}^2]. \end{aligned}$$

Considerando poi che

$$\begin{aligned} E[Y_{t-1}^2\sigma_{t-1}^2] &= E[E[Y_{t-1}^2\sigma_{t-1}^2|\mathcal{F}_{t-2}]] = \\ &= E[\sigma_{t-1}^2 E[Y_{t-1}^2|\mathcal{F}_{t-2}]] = \\ &= E[\sigma_{t-1}^2\sigma_{t-1}^2] = \\ &= E[\sigma_{t-1}^4] \end{aligned}$$

e

$$E[\sigma_t^4] = \frac{E[Y_t^4]}{3},$$

il momento di quarto ordine risulta pari a:

$$E[Y_t^4] = \frac{3\alpha_0^2(1 + \alpha_1 + \beta_1)}{(1 - \alpha_1 - \beta_1)(1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_1^2)}. \quad (4.29)$$

Di conseguenza, tenendo conto delle condizioni di esistenza dei momenti di quarto e secondo ordine, l'indice di curtosi risulta:

$$K = \frac{E[Y_t^4]}{(E[Y_t^2])^2} = \frac{3(1 - \alpha_1 - \beta_1)(1 + \alpha_1 + \beta_1)}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2} > 3. \quad (4.30)$$

Verifichiamo ora che il processo $\{v_t\}$ introdotto per poter rappresentare i quadrati di un processo GARCH come un processo ARMA, sia effettivamente un white noise.

Preso $(Y_t)_{t \in \mathbb{Z}}$ un processo GARCH(1,1), il processo ARMA(1,1) per i quadrati delle osservazioni risulta pari a:

$$Y_t^2 = \alpha_0 + (\alpha_1 + \beta_1)Y_{t-1}^2 - \beta_1 v_{t-1} + v_t. \quad (4.31)$$

$E[v_t] = 0$ e $\text{Cov}(v_t, v_{t-h}) = 0$ si verificano analogamente all'*Esempio 2*. Sfruttando poi i calcoli precedenti per il momento di quarto ordine, si trova infine la formula per la varianza, che risulta costante.

$$\begin{aligned} \text{Var}(v_t) &= E[v_t^2] = E[(Y_t^2 - \sigma_t^2)^2] = E[Y_t^4] + E[\sigma_t^4] - 2E[Y_t^2 \sigma_t^2] = \\ &= \frac{2\alpha_0^2(1 + \alpha_1 + \beta_1)}{(1 - \alpha_1 - \beta_1)(1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_1^2)}. \end{aligned}$$

Sfruttando ora la possibilità di rappresentare i quadrati delle osservazioni di un processo GARCH, come un processo ARMA, vediamo come calcolare le funzioni di autocovarianza e autocorrelazione. Come già osservato in precedenza queste sono molto utili per studiare il comportamento di una serie storica e capire se questo si può rappresentare con un processo GARCH.

Ricordiamo ora la formula per la funzione di autocovarianza di un processo ARMA(1,1) trovata nel capitolo precedente:

$$\gamma(h) = \begin{cases} \phi^{h-1} \sigma^2 \left[\frac{(\theta + \phi)(1 + \theta\phi)}{1 - \phi^2} \right] & h \neq 0 \\ \sigma^2 \left[\frac{1 + 2\theta\phi + \theta^2}{1 - \phi^2} \right] & h = 0. \end{cases} \quad (4.32)$$

In questo caso i parametri sono:

- $\phi = \alpha_1 + \beta_1$;
- $\theta = -\beta_1$;
- $\sigma^2 = \text{Var}(v_t) = E[v_t^2] = E[Y_t^4] + E[\sigma_t^4] - 2E[Y_t^2\sigma_t^2] = \frac{2}{3}E[Y_t^4] = \frac{2\alpha_0^2(1+\alpha_1+\beta_1)}{(1-\alpha_1-\beta_1)(1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2)}$.

Andando a sostituire in (4.32), si trova:

$$\gamma(h) = \begin{cases} (\alpha_1 + \beta_1)^{h-1} \frac{2\alpha_0^2\alpha_1(1-\beta_1\alpha_1-\beta_1^2)}{(1-\alpha_1-\beta_1)^2(1-3\alpha_1^2-\beta_1^2-2\alpha_1\beta_1)} & h \neq 0 \\ \frac{2\alpha_0^2(1-2\beta_1\alpha_1-\beta_1^2)}{(1-\alpha_1-\beta_1)^2(1-3\alpha_1^2-\beta_1^2-2\alpha_1\beta_1)} & h = 0. \end{cases} \quad (4.33)$$

La funzione di autocorrelazione per un processo GARCH(1,1) risulta quindi pari a:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = (\alpha_1 + \beta_1)^{h-1} \frac{\alpha_1(1 - \beta_1\alpha_1 - \beta_1^2)}{1 - 2\beta_1\alpha_1 - \beta_1^2}. \quad (4.34)$$

4.3 Processi EGARCH

Sebbene i modelli GARCH siano in grado di riprodurre parecchie caratteristiche delle serie storiche finanziarie, come la leptocurtosi e la volatility clustering, non sono in grado di spiegare l'effetto leverage. Questa lacuna è dovuta alla simmetria della classe di modelli GARCH, per cui shock positivi o negativi hanno lo stesso effetto sulla volatilità. Empiricamente però questo spesso non accade; infatti, come già anticipato, l'incremento della volatilità è usualmente maggiore a seguito di variazioni negative dei prezzi rispetto a variazioni positive.

Per tener conto di questo effetto sono state proposte diverse estensioni dei modelli GARCH, una in particolare è quella dei processi EGARCH, introdotta da Daniel B. Nelson nel 1991 [6].

Definizione 4.3. Un processo $(Y_t)_{t \in \mathbb{Z}}$ si dice **autoregressivo a eteroschedasticità condizionata generalizzato esponenziale di ordine p e q** o in breve **EGARCH(p,q)**, con $p, q \in \mathbb{N}^*$ se soddisfa la seguente equazione:

$$Y_t = \sigma_t e_t \quad (4.35)$$

dove

$$\log(\sigma_t^2) = \alpha_0 + \sum_{i=1}^q \alpha_i g(e_{t-i}) + \sum_{j=1}^p \beta_j \log(\sigma_{t-j}^2) \quad (4.36)$$

e

$$g(e_t) = w e_t + \lambda(|e_t| - E[|e_t|]) \quad (4.37)$$

con $\alpha_0, \alpha_i, \beta_j, w, \lambda \in \mathbb{R}$ ed $\{e_t\} \sim \text{IID}(0, 1)$ e simmetrico.

Questa classe di processi ha numerosi vantaggi rispetto a quelle descritte in precedenza. Innanzitutto il modello è più flessibile, in quanto non sono necessarie condizioni sui parametri che garantiscano la positività della varianza condizionale. Infatti, il modello richiede un'espressione del logaritmo della varianza condizionale e non della varianza condizionale stessa.

Per garantire la stabilità dei processi EGARCH(p,q), è necessario imporre $\sum_{i=1}^p \beta_i < 1$.

Inoltre, riscrivendo $g(e_t)$ come segue,

$$g(e_t) = \begin{cases} (w + \lambda)e_t - \lambda E[|e_t|] & e_t \geq 0 \\ (w - \lambda)e_t - \lambda E[|e_t|] & e_t < 0 \end{cases} \quad (4.38)$$

si osserva che si tratta di una funzione lineare a tratti con pendenza $(w + \lambda)$ su $(0, \infty)$ e $(w - \lambda)$ su $(-\infty, 0)$. Questa asimmetria di g fa sì che $\log(\sigma_t^2)$ risponda in modo diverso a variazioni positive o negative della stessa ampiezza. Se $w = 0$ la funzione è simmetrica.

Affinché il modello sia in grado di cogliere l'effetto leverage, ovvero il fatto che shock negativi abbiano un effetto maggiore sulla volatilità rispetto a shock positivi, è naturale aspettarsi che il parametro w sia negativo.

Sebbene questo modello abbia portato un notevole miglioramento rispetto a quelli precedenti, soprattutto poiché riesce a spiegare l'effetto leverage, ha un limite: è stato strutturato per analizzare il logaritmo naturale della varianza condizionale e non della varianza condizionale stessa. Questo potrebbe essere un problema nelle previsioni di valori futuri della volatilità dato che $E[\ln(\sigma_t^2)] \neq \ln E(\sigma_t^2)$.

Capitolo 5

Studio della serie storica dell'EU ETS

In questo capitolo vediamo in un primo momento il dettaglio dell'evoluzione del prezzo delle quote di carbonio negli anni più recenti, dalla fine della Fase 3, fino ad oggi. Utilizziamo i rapporti GSE sulle aste di quote europee di emissione, per capire quali sono state le oscillazioni del prezzo delle EUA (European Union Allowances) e le possibili cause che ne hanno determinato il trend (Figura 5.1).

In seguito, viene analizzata numericamente la serie storica in Fase 4, da inizio 2021 fino al 28 agosto 2023. Utilizziamo il software Matlab, più nello specifico l'applicazione *Econometric Modeler*, grazie alla quale vedremo come scegliere un modello opportuno per i dati, come stimare i parametri ed infine come valutare la bontà del modello tramite l'analisi dei residui.

5.1 Analisi qualitativa

Dalla sua introduzione, come già anticipato nel Capitolo 1, il prezzo delle quote di emissione ha subito notevoli cambiamenti (Figura 1.1). Dalla seconda metà della terza fase però, tale prezzo ha cominciato a crescere in modo significativo. Infatti, a partire dal 2018, questo si è quasi triplicato rispetto all'anno precedente, determinando così un aumento sostanziale dei proventi generati dalla vendita (+156%), nonostante i volumi di quote disponibili fossero diminuiti.

Il 2019 invece è stato caratterizzato dall'entrata in funzione della Riserva della Stabilità, che ha comportato una consistente diminuzione (-35%) dei volumi di EUA all'asta rispetto al 2018. La riduzione del numero di quote disponibili ha quindi causato l'incremento del prezzo medio, che è passato da 15,4 € a 24,6 €.

Nel 2020 c'è stato un aumento considerevole delle quote disponibili (da 588 milioni di EUA nel 2019 a 778 nel 2020), che ha portato i proventi generati a 19 miliardi di euro,

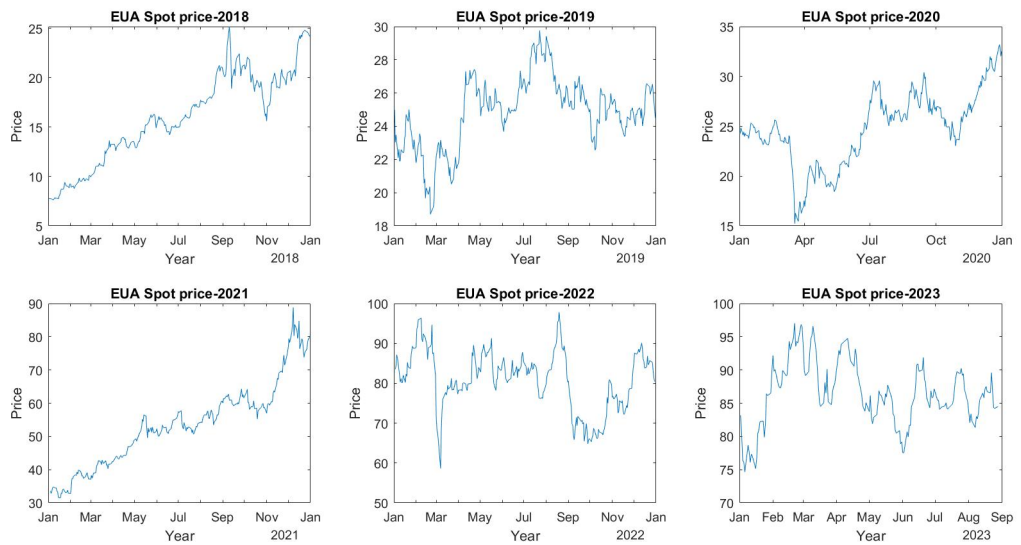


Figura 5.1: Andamento prezzi dell'EU ETS negli anni 2018-2023

il 31% in più rispetto al 2019, nonostante il prezzo medio di EUA fosse diminuito. Infatti la pandemia ha provocato una repentina diminuzione dei prezzi in concomitanza con il primo lockdown fino ad un valore minimo di 14,60 €. Da fine maggio, il trend delle quotazioni è stato nuovamente rialzista, ad eccezione del mese di ottobre. Le principali cause di questo andamento, sono state l'annuncio dell'impossibilità di godere del borrowing per le quote di quarta fase, il momento positivo sui mercati e la crescita stagionale del prezzo del gas e delle principali commodities energetiche.

Nel 2021, con l'inizio della quarta fase, sono state collocate sul mercato 583 milioni di EUA, il 13% in meno rispetto all'anno precedente. Nonostante ciò, i ricavi generati sono aumentati di 14 miliardi di euro, grazie alla variazione positiva di prezzo (+116%). Anche in questo anno è stato fondamentale il meccanismo della Riserva Stabilizzatrice del Mercato, che ha prelevato 378 milioni di quote dal mercato. Infatti le emissioni dei settori ETS, anche a causa dell'avvento della pandemia, sono state inferiori alle attese, determinando così un surplus di quote. Nel 2021 l'andamento del prezzo della quota è stato caratterizzato da una tendenza al rialzo, dovuto in parte alla discussione del regolamento di UE n° 447/2021, in cui sono stati aggiornati in modo stringente rispetto alla terza fase i benchmark per il periodo 2021-2025.

Il trend di crescita è proseguito anche all'inizio del 2022, in cui sono state messe sul mercato 482 milioni di quote (17% in meno rispetto al 2021), ad un prezzo medio ponderato di 79,7 €. A febbraio è stato raggiunto un massimo storico pari a 96 €, ma il trend crescente è stato interrotto a causa dell'inizio del conflitto tra Russia ed Ucraina; la forte svendita di quote sul mercato secondario, ha spinto i prezzi verso il basso (-30% all'inizio di marzo). In seguito il trend è tornato al rialzo nei mesi estivi per la ridotta offerta di

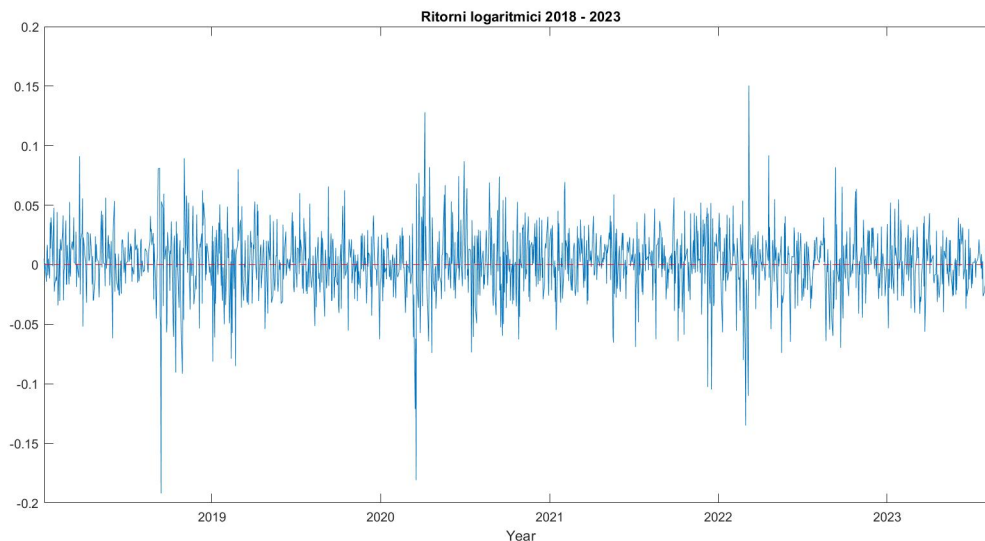


Figura 5.2: Rendimenti giornalieri dei prezzi delle quote EUA negli anni 2018 - 2023

volumi all'asta e per le temperature che hanno ridotto le previsioni di fabbisogno di gas.

Nel primo trimestre del 2023 il prezzo delle quote ha seguito un trend rialzista, raggiungendo un nuovo massimo storico il 28 febbraio 2023, pari a 100 €. Nei mesi seguenti però il prezzo di EUA è rimasto pressoché stabile, con una media pari a 86€. L'impatto del disincentivo alle emissioni GHG attraverso l'ETS è pensabile che spinga i settori maggiormente emissivi ad accelerare per raggiungere la completa decarbonizzazione dei loro impianti e, dove ciò non sarà possibile in tempi brevi, ad affidarsi e finanziare nuove tecnologie CCUS.

Focalizzando invece l'attenzione sull'andamento dei ritorni logaritmici (Figura 5.2), si notano cambiamenti nel livello della volatilità e il fenomeno del *volatility clustering*, con particolari picchi a fine 2018, inizio 2020 e 2022.

Se l'importante variazione di volatilità del prezzo della quota nel 2020 è dovuta alla diffusione del COVID e nel 2022 alla guerra Russia-Ucraina, a fine 2018 non è stata identificata una causa scatenante. Infatti, come riportato nel Rapporto GSE sulle aste del 2018, l'aumento sostanziale della volatilità è dovuto alle elevate oscillazioni del prezzo registrate tra il 6 e il 14 settembre, dove la variazione massima si è verificata il 13 settembre, con un ritorno logaritmico di quasi il 20%.

5.2 Identificazione dei modelli

In questa sezione mostriamo le principali proprietà della serie dei prezzi in Fase 4 (in notazione $P = P_1, \dots, P_n$) e alcuni modelli opportuni.

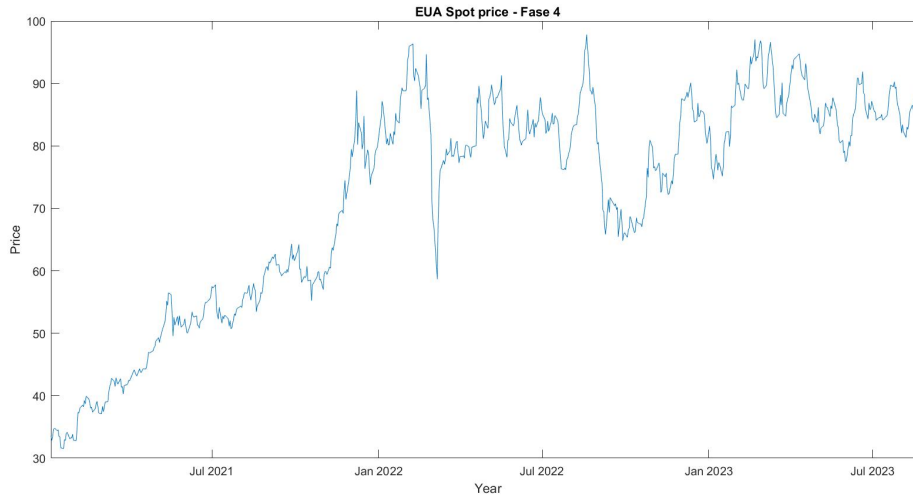


Figura 5.3: Serie storica dei prezzi di spot dell'EU ETS in Fase 4

In particolare consideriamo le osservazioni da inizio 2021 fino al 28 agosto 2023 (in notazione $\mathcal{T} = t_1, \dots, t_n$). Il data set considerato riporta i prezzi delle quote dell'EU ETS nei giorni feriali, quindi dal lunedì al venerdì con esclusione delle festività, per un totale di $n = 680$ osservazioni.

È necessario fare subito un'osservazione: come si evince dalla Figura 5.3, la serie storica dei prezzi non è stazionaria. Trattandosi di una serie finanziaria, come suggerito nel Capitolo 4, consideriamo la successione dei ritorni logaritmici definita da:

$$Y_{t_k} = \log \frac{P_{t_k}}{P_{t_{k-1}}}, \quad \text{per ogni } k = 2, \dots, n. \quad (5.1)$$

5.2.1 Considerazioni iniziali

Vediamo ora le principali caratteristiche distributive della serie dei ritorni logaritmici in Fase 4 (Figura 5.4), che risulta essere effettivamente stazionaria. Valutiamo innanzitutto se la distribuzione della serie è Normale e se si evidenziano particolari autocorrelazioni.

Per quanto riguarda la distribuzione della serie, si può osservare un comportamento non Normale in molteplici modi. Innanzitutto, dalla Tabella 5.1, si evince che l'indice di curtosi è pari a 6.3809, molto superiore a 3, valore di una distribuzione Normale. Inoltre, l'indice di asimmetria è leggermente negativo, dunque con una coda più lunga a sinistra, ed evidenzia una limitata asimmetria della distribuzione, non presente in una distribuzione Normale.

Queste caratteristiche si possono osservare anche tramite l'istogramma e il grafico Q-Q Plot (Figura 5.5) e con entrambi è possibile confrontare la distribuzione campionaria

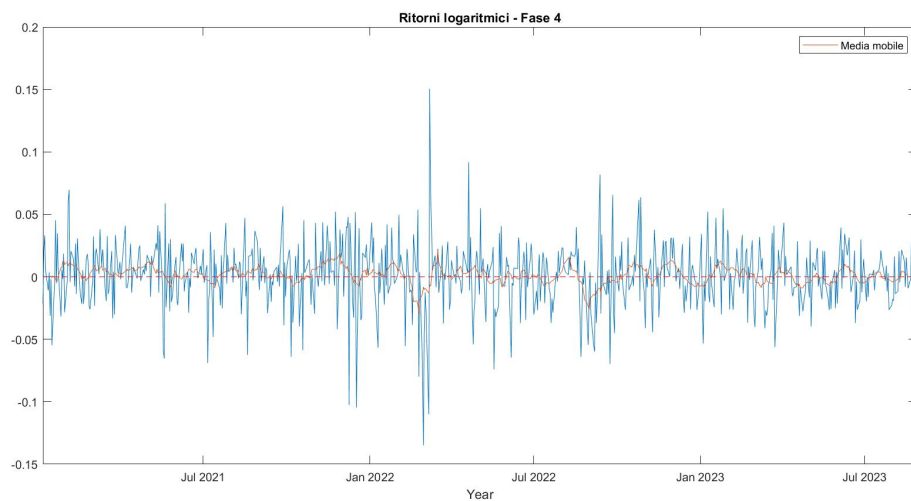


Figura 5.4: Rendimenti giornalieri dei prezzi delle quote EUA in Fase 4. La linea arancione indica la media mobile ottenuta con 15 osservazioni, mentre la linea rossa tratteggiata è la media della serie dei rendimenti.

Media	0.0013608
Mediana	0.0028784
Deviazione Standard	0.026526
Valore Massimo	0.15056
Valore Minimo	-0.13498
Varianza	0.00070361
Coefficiente di variazione	19.4933
Curtosi	6.3809
Indice di Asimmetria	-0.30944
1° Quartile	-0.012164
3° Quartile	0.01716

Tabella 5.1: Caratteristiche della distribuzione dei rendimenti in Fase 4

con quella Normale. In particolare nel Q-Q Plot, qualora la variabile osservata presenti effettivamente una distribuzione Normale, i punti del grafico si disporranno vicino alla bisettrice del I quadrante.

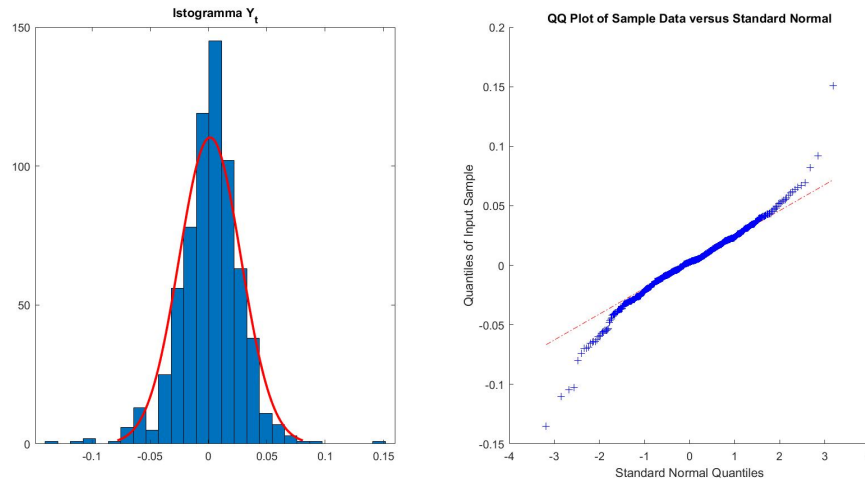


Figura 5.5: Istogramma dei ritorni logaritmici e Q-Q plot

La distribuzione non Normale della serie osservata si può valutare più rigorosamente utilizzando il test di Jarque-Bera (Figura 5.6). L'ipotesi nulla di questo test è infatti la distribuzione Normale dei dati, e, in questo caso, il valore piccolo del P-Value porta a rifiutare tale ipotesi.

	P-Value	Test Statistic	Critical Value
Jarque-Bera TEST	0.001	334.22	5.8939

Figura 5.6: Test Jarque-Bera sui ritorni logaritmici della serie in esame

Un'altra considerazione iniziale da fare sulla serie storica, riguarda l'autocorrelazione e l'autocorrelazione parziale tra i dati. Come è mostrato in Figura 5.7, sembra esserci una leggera correlazione al lag 5, indice di una possibile stagionalità settimanale.

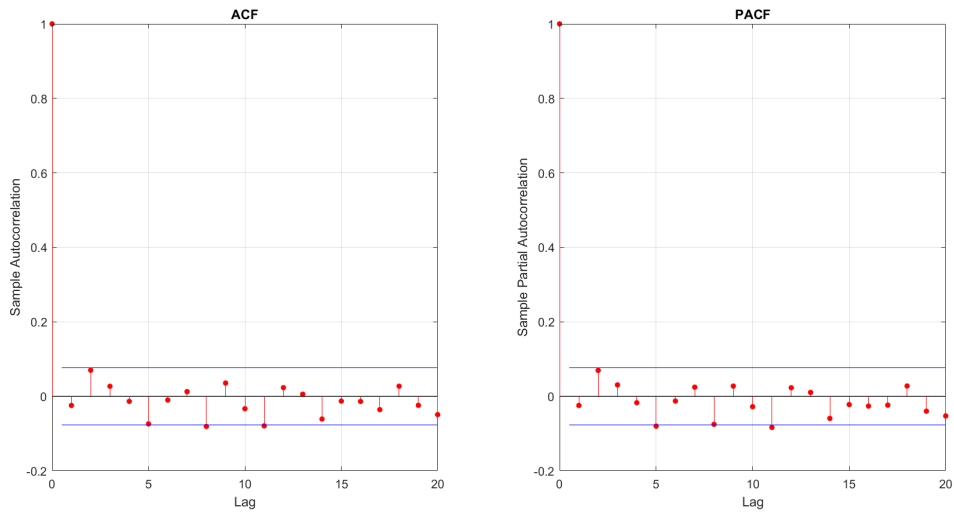


Figura 5.7: Funzioni di autocorrelazione e autocorrelazione parziale campionarie

Tuttavia questa correlazione non è ritenuta significativa dal test di Ljung-Box. Questo test infatti permette di valutare la presenza o meno di autocorrelazione tra le osservazioni. L'ipotesi nulla indica l'assenza di autocorrelazione tra i dati e, nel caso in esame, questa ipotesi viene accettata, come mostrato in Figura 5.8 dal valore del P-Value superiore a 0.05.

	P-Value	Test Statistic	Critical Value
Ljung-Box TEST	0.18165	25.537	31.41

Figura 5.8: Test Box-Ljung sui ritorni logaritmici della serie in esame

Vediamo adesso diversi approcci che sono stati utilizzati per identificare vari modelli, verificarli e scegliere il più opportuno.

5.2.2 Metodo 1

Una volta accertato che la distribuzione campionaria è non Normale e che le osservazioni sono scorrelate, proseguiamo l'analisi andando a valutare le funzioni di autocorrelazione e autocorrelazione parziale dei quadrati dei dati.

Come mostrato in Figura 5.9, si osserva che queste sono significativamente diverse da zero, il che suggerisce che un modello ARCH o GARCH potrebbe essere appropriato per

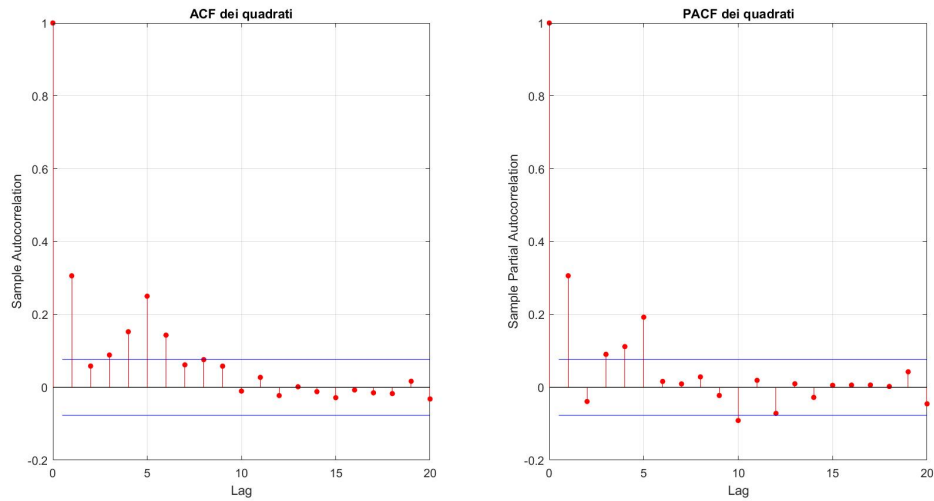


Figura 5.9: ACF e PACF dei quadrati delle osservazioni

la serie storica. Infatti queste classi di modelli permettono di dare una rappresentazione che lega i quadrati di un'osservazione al tempo t ai quadrati ai tempi precedenti.

Utilizziamo quindi il **test ARCH**, un test che permette di testare i cosiddetti *effetti ARCH*, utilizzato per valutare la presenza di eteroschedasticità condizionale nella serie storica. Questo test fa parte della famiglia di test dei moltiplicatori di Lagrange; per questo motivo spesso viene chiamato LM-ARCH. L'obiettivo è quello di testare l'ipotesi nulla

$$H_0 : \alpha_1 = \dots = \alpha_q = 0,$$

ovvero l'omoschedasticità nella serie in esame. Qualora l'ipotesi nulla venisse accettata, la serie avrà un comportamento omoschedastico; qualora invece venisse rifiutata, si è in presenza di varianza condizionale non costante nel tempo.

Come si evince dall'immagine 5.10, l'ipotesi nulla che valuta l'assenza di eteroschedasticità condizionale viene rifiutata.

	P-Value	Test Statistic	Critical Value
ARCH TEST	1.5543e-15	63.55	3.8415

Figura 5.10: Test ARCH

A questo punto possiamo specificare i modelli ARCH, GARCH ed EGARCH per descrivere la varianza condizionale. Sono stati considerati due casi sulla possibile distribuzione di $\{e_t\}$, ovvero la distribuzione Normale e t-Student. Si sono ottenuti risultati migliori

ipotizzando una distribuzione t , con i gradi di libertà stimati dal modello; questo è il caso che viene presentato nel seguito della discussione.

Per quanto riguarda la stima parametrica, Matlab utilizza la funzione di verosimiglianza. Per questa classe di processi infatti, sfruttando il fatto che la distribuzione condizionale di Y_t è gaussiana, $Y_t|\mathcal{F}_{t-1} \sim N(0, \sigma_t^2)$, risulta semplice scrivere la funzione di verosimiglianza e massimizzarla per ottenere gli stimatori dei parametri.

Prendendo ad esempio la sequenza di osservazioni y_1, \dots, y_n , di un processo GARCH(p,q), la funzione di verosimiglianza condizionale di y_2, \dots, y_n dato $Y_1 = y_1$ è pari a:

$$\mathcal{L} = \prod_{t=2}^n \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left\{-\frac{y_t^2}{2\sigma_t^2}\right\}.$$

Si può poi costruire la funzione di log-verosimiglianza, più semplice da utilizzare numericamente per massimizzarla.

$$\log \mathcal{L} = -\frac{n-1}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \log \sigma_t^2 - \frac{1}{2} \sum_{t=2}^n \left(\frac{y_t^2}{2\sigma_t^2}\right)$$

Vediamo ora nel dettaglio quali modelli sono stati stimati.

Partendo dalla classe dei processi ARCH, è necessario inizialmente determinare un ordine q opportuno. Osservando la funzione di autocorrelazione parziale dei quadrati, si notano correlazioni significative ai lag 1,3,4,5,10. Per non utilizzare un ordine troppo elevato e rischiare di fare overfitting, consideriamo un modello ARCH(5).

Stimiamo quindi il valore dei parametri. Dalla Figura 5.11 si evince che i parametri stimati non sono tutti significativi. Infatti, l'indice *TStatistic*, dato dal rapporto tra *Value* e *Standard Error*, indica se il valore dei parametri stimati si discosta significativamente da 0. Nel caso dei parametri ARCH 2,3,4,5, questo valore è piuttosto basso, e la non significatività viene confermata anche dal *P-Value*, maggiore di 0.05.

GARCH(0,5) Conditional Variance Model (t Distribution):

	Value	StandardError	TStatistic	PValue
Constant	0.00045929	6.1876e-05	7.4227	1.1478e-13
ARCH{1}	0.1378	0.052234	2.6382	0.0083348
ARCH{2}	0.069875	0.057454	1.2162	0.22391
ARCH{3}	0.046324	0.049537	0.93513	0.34972
ARCH{4}	0.050382	0.039747	1.2676	0.20494
ARCH{5}	0.015866	0.037491	0.4232	0.67215
DoF	6.7873	2.0735	3.2733	0.0010632

Figura 5.11: ARCH(5)

Consideriamo quindi altri due modelli più semplici: ARCH(1) e ARCH(5), con $\alpha_1, \alpha_2, \alpha_3, \alpha_4 = 0$, per tener conto della periodicità settimanale. In entrambi i modelli l'unico parametro stimato è significativo (Figure 5.12, 5.13).

GARCH(0,1) Conditional Variance Model (t Distribution):

	Value	StandardError	TStatistic	PValue
Constant	0.00056613	6.0664e-05	9.3322	1.0367e-20
ARCH{1}	0.16674	0.052469	3.1778	0.0014837
DoF	6.3305	1.7499	3.6175	0.00029741

Figura 5.12: ARCH(1)

GARCH(0,5) Conditional Variance Model (t Distribution):

	Value	StandardError	TStatistic	PValue
Constant	0.00063861	6.9649e-05	9.1689	4.7801e-20
ARCH{5}	0.084662	0.041851	2.0229	0.04308
DoF	5.3505	1.2096	4.4234	9.7164e-06

Figura 5.13: ARCH(5) con $\alpha_1, \alpha_2, \alpha_3, \alpha_4 = 0$

Il secondo modello che consideriamo è il GARCH(1,1). Come nel caso precedente sono stati stimati i parametri, i cui valori sono tutti significativi (Figura 5.14).

GARCH(1,1) Conditional Variance Model (t Distribution):

	Value	StandardError	TStatistic	PValue
Constant	0.00010588	5.0485e-05	2.0972	0.035974
GARCH{1}	0.7274	0.10002	7.2723	3.5342e-13
ARCH{1}	0.11682	0.040128	2.9111	0.0036012
DoF	6.9125	2.0893	3.3085	0.00093811

Figura 5.14: GARCH(1,1)

Infine, vediamo il modello EGARCH(1,1) (Figura 5.15). È bene osservare che il parametro di Leverage risulta negativo, come ci si aspettava dalla teoria. Infatti questo coglie

EGARCH(1,1) Conditional Variance Model (t Distribution):

	Value	StandardError	TStatistic	PValue
Constant	-1.0665	0.54974	-1.94	0.052377
GARCH{1}	0.85427	0.075049	11.383	5.0865e-30
ARCH{1}	0.23331	0.071465	3.2646	0.0010961
Leverage{1}	-0.030157	0.038998	-0.7733	0.43934
DoF	6.4402	1.7782	3.6218	0.00029261

Figura 5.15: EGARCH(1,1)

l'asimmetria, secondo cui rendimenti negativi hanno maggior effetto sull'aumento della volatilità.

Vengono poi calcolati gli indici di AIC e BIC dei quattro modelli con i parametri significativi, in modo da confrontarli e scegliere il migliore. Questi criteri permettono infatti di fare un compromesso tra bontà del fitting e complessità del modello. In particolare, nel caso di modelli GARCH (e quindi ARCH, ponendo $p = 0$), vengono calcolati utilizzando le seguenti espressioni.

$$AIC := -2 \ln \mathcal{L} + 2(p + q + 1)$$

$$BIC := -2 \ln \mathcal{L} + (p + q + 1) \ln(n)$$

Nel caso di modelli EGARCH invece, il numero dei parametri viene aumentato di 1, per tenere conto anche della leverage. Va osservato inoltre che nel caso in esame viene stimato anche ν , ovvero il numero di gradi di libertà della distribuzione t .

Come è mostrato in Figura 5.16, per entrambi i criteri è preferibile il modello GARCH(1,1).

	AIC	BIC
ARCH (1)	-3076.3	-3062.7
GARCH (1, 1)	-3084	-3065.9
EGARCH (1, 1)	-3079	-3056.3
ARCH (5) con alpha 1,2,3,4=0	-3062.8	-3049.3

Figura 5.16: Criteri di AIC e BIC

Vediamo ora le stime della volatilità dei modelli (Figura 5.17). Tutti e quattro mostrano il periodo di maggior volatilità a inizio 2022, in corrispondenza dell'inizio della guerra Russia-Ucraina. Seppur di minor importanza, anche a fine 2021 c'è stata notevole variazione della volatilità giornaliera a seguito di oscillazioni sul prezzo.

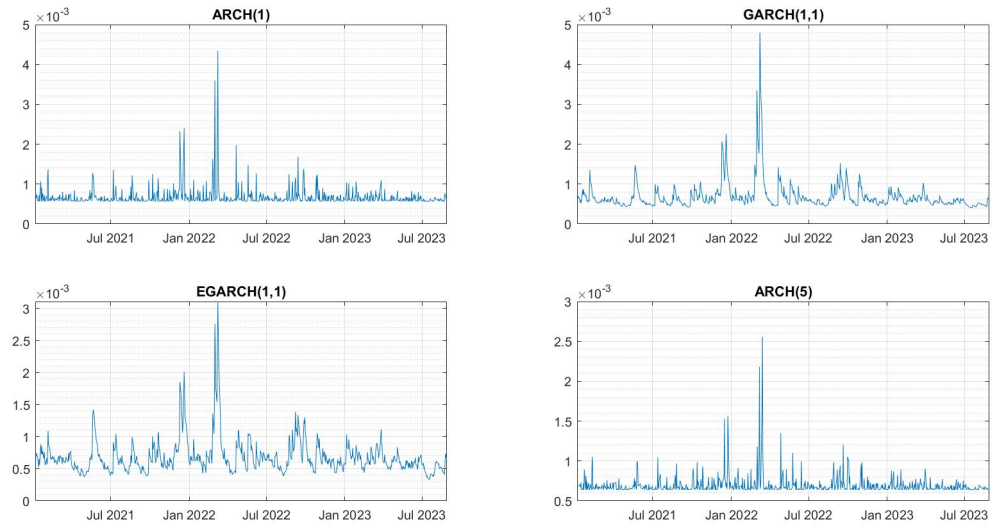


Figura 5.17: Volatilità condizionale stimata dai modelli ARCH(1), ARCH(5) con $\alpha_1, \alpha_2, \alpha_3, \alpha_4 = 0$, GARCH(1,1) e EGARCH(1,1)

Lo step successivo è quello di verificare la bontà dei modelli stimati. Vediamo quindi come fare l'analisi dei residui standardizzati. Innanzitutto, i residui restituiti da Matlab vengo calcolati così:

$$\hat{e}_t = \frac{Y_t}{\hat{\sigma}_t}. \quad (5.2)$$

Affinché il modello sia adeguato ai dati, ci si aspetta che i residui mostrino le stesse proprietà e abbiano la stessa distribuzione del processo $\{e_t\}$. Per la serie in esame ci si aspetta quindi che $\{\hat{e}_t\} \sim \text{IID}(0, 1)$ e che tale processo sia distribuito come una $t - Student$ con ν gradi di libertà.

Per ciascun dei quattro modelli, presentiamo le caratteristiche distributive dei residui (Figura 5.22). Inoltre, oltre al grafico dei residui standardizzati (in alto a sinistra), mostriamo la funzione di autocorrelazione campionaria (in alto a destra), il Q-Q Plot (in basso a sinistra) e il confronto tra la distribuzione campionaria e la distribuzione t_ν con ν il valore risultante dalla stima (in basso a destra).

I risultati ottenuti, mostrano che i residui di tutti e quattro i modelli hanno effettivamente media 0 e varianza 1. Inoltre, osservando il grafico delle funzioni di autocorrelazione, simile a quello di un white noise, non ci sono particolare correlazioni tra i residui. Questo fatto viene confermato anche dal test di Ljung-Box, in cui i valori del P-Value superiori a 0.05 permettono di accettare l'ipotesi nulla di assenza di autocorrelazione (Figura 5.23).

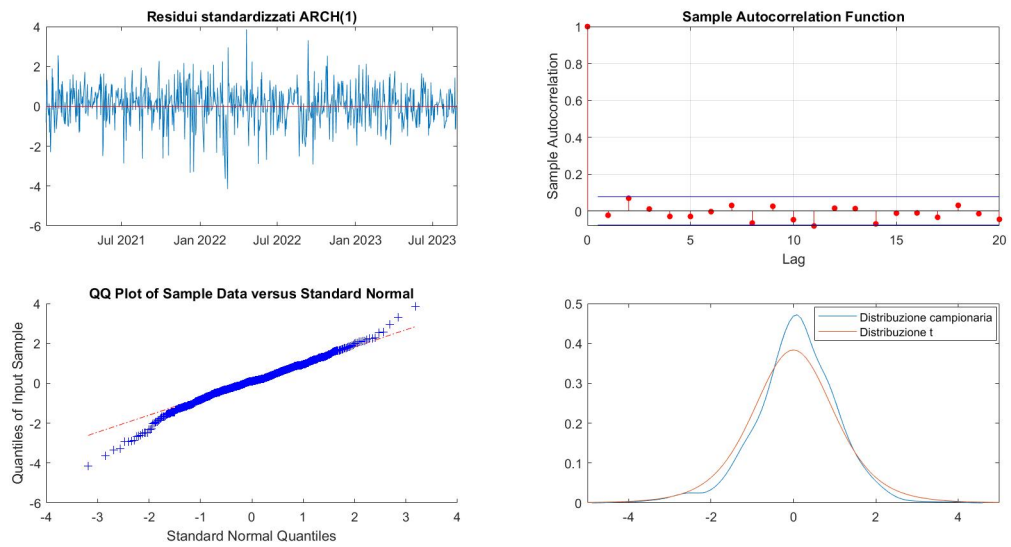


Figura 5.18: Residui standardizzati ARCH(1)

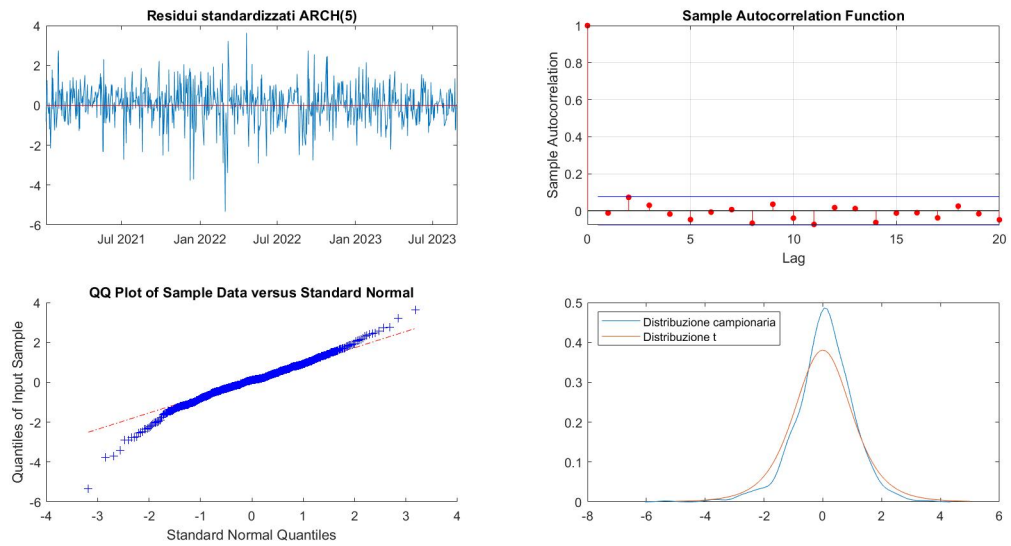


Figura 5.19: Residui standardizzati ARCH(5) con $\alpha_1, \alpha_2, \alpha_3, \alpha_4 = 0$

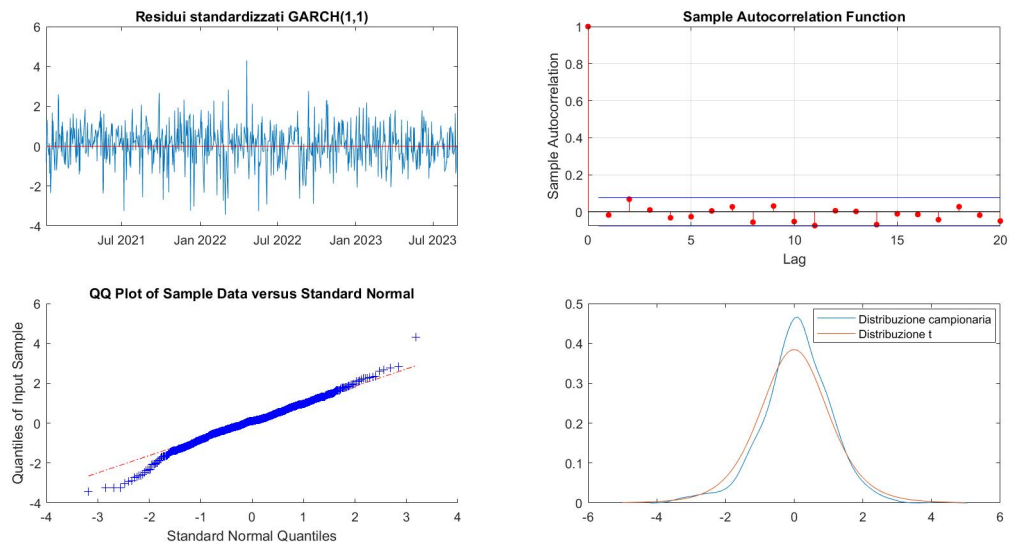


Figura 5.20: Residui standardizzati GARCH(1,1)

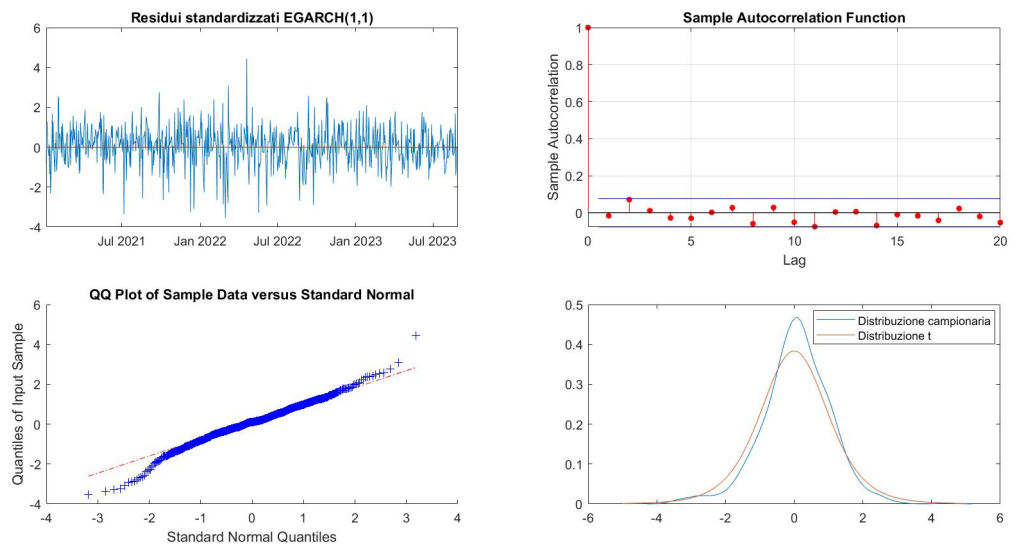


Figura 5.21: Residui standardizzati EGARCH(1,1)

RESIDUI ARCH(1)	RESIDUI ARCH(5)
Media: 0.059434	Media: 0.052077
Mediana: 0.10093	Mediana: 0.11026
Deviazione Standard: 0.99218	Deviazione Standard: 0.99082
Valore Massimo: 3.8544	Valore Massimo: 3.6313
Valore Minimo: -4.1518	Valore Minimo: -5.3317
Varianza: 0.98443	Varianza: 0.98172
Coefficiente di Variazione: 16.6937	Coefficiente di Variazione: 19.0261
Curtosi: 4.4041	Curtosi: 5.2328
Indice di Asimmetria: -0.37014	Indice di Asimmetria: -0.48672
1° Quartile: -0.46344	1° Quartile: -0.45496
3° Quartile: 0.69234	3° Quartile: 0.64885
RESIDUI GARCH(1,1)	RESIDUI EGARCH(1,1)
Media: 0.065071	Media: 0.066156
Mediana: 0.10337	Mediana: 0.10255
Deviazione Standard: 0.99263	Deviazione Standard: 0.99215
Valore Massimo: 4.3005	Valore Massimo: 4.4356
Valore Minimo: -3.4469	Valore Minimo: -3.5436
Varianza: 0.98531	Varianza: 0.98437
Coefficiente di Variazione: 15.2546	Coefficiente di Variazione: 14.9971
Curtosi: 4.1676	Curtosi: 4.3669
Indice di Asimmetria: -0.26449	Indice di Asimmetria: -0.27113
1° Quartile: -0.47463	1° Quartile: -0.46341
3° Quartile: 0.69721	3° Quartile: 0.69252

Figura 5.22: Statistiche descrittive dei residui standardizzati nei 4 casi

	P-Value Ljung-Box
ARCH (1)	0.3336
GARCH (1,1)	0.3836
EGARCH (1,1)	0.37025
ARCH (5) con alpha 1,2,3,4=0	0.36607

Figura 5.23: Test di Ljung-Box sui residui standardizzati dei quattro modelli

Tuttavia, osservando il Q-Q Plot si osserva una particolare pesantezza e asimmetria delle code, il che si evince anche dai valori degli indici di curtosi e di asimmetria in Figura 5.22. Inoltre, dal confronto tra la distribuzione campionaria e quella teorica di una t-Student si osserva una discrepanza.

Dalla verifica dei residui, si può quindi concludere che i modelli proposti non sono dei buoni candidati per la serie storica.

5.2.3 Metodo 2

Il secondo tentativo è consistito nel cercare di capire (o nell'indagare) se la correlazione al lag 5 individuata nei grafici di autocorrelazione potesse essere la causa dei problemi riscontrati nell'approccio precedente e fosse quindi importante da tenere in considerazione.

Sono state provate diverse soluzioni. La prima proposta è stata provare ad identificare un modello ARMA-GARCH, nello specifico il modello AR(5)-GARCH(1,1). È stato stimato inizialmente modello ARMA, ai cui residui, dato che presentavano effetti ARCH, è stato applicato un modello GARCH. Tuttavia non si sono riscontrati miglioramenti.

Un altro tentativo è stato valutare se il giorno della settimana fosse effettivamente il responsabile del comportamento anomalo. In particolare, si sono fatte due osservazioni. La prima è stata valutare la presenza di un *day of the week effect* (o *weekend effect*), secondo cui il rendimento medio del venerdì è anormalmente alto, mentre il rendimento medio del lunedì è anormalmente basso. Questo effetto è anomalo, in quanto ci si aspetta che i rendimenti del lunedì siano più alti (il triplo) rispetto agli altri giorni, a causa del gap di tre giorni tra venerdì e lunedì. Tuttavia la serie storica in esame non mostra questo comportamento.

La seconda osservazione è stata provare ad utilizzare un modello GARCH modificato seguendo [16], dove sono state introdotte delle dummy variables d_k , $k = 1, 2, 3, 4, 5$ per indicare il giorno della settimana:

$$Y_t = \sigma_t e_t$$
$$\sigma_t^2 = \sum_{k=1}^5 d_k + \sum_{i=1}^q \alpha_i Y_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2.$$

Anche questo tentativo non ha portato miglioramenti significativi.

5.2.4 Metodo 3

Un ulteriore tentativo è stato fatto a partire dall'osservazione della media mobile fatta sulla serie dei ritorni logaritmici presente in Figura 5.4. Per la media mobile sono state utilizzate 15 osservazioni e si evidenziano diversi cluster di media.

È stata quindi sottratta questa media mobile alla serie originaria dei rendimenti, ottenendo così una nuova serie, a cui sono state tolte solamente le prime e le ultime 7 osservazioni (Figura 5.24).

Così come è stato fatto per la serie dei ritorni logaritmici precedente, facciamo alcune considerazioni iniziali. La nuova serie continua ad essere non Normale, il che si può evincere dall'istogramma, dal Q-Q Plot, dal test di Jarque-Bera e dall'indice di curtosi (Figure 5.25, 5.26 e Tabella 5.2).

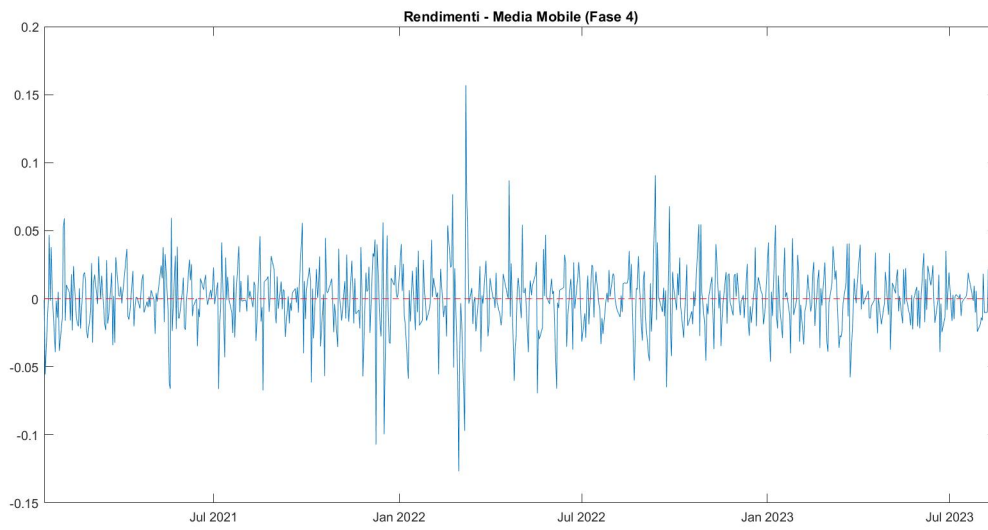


Figura 5.24: Nuova serie storica, ottenuta dalla differenza della serie precedente con la media mobile

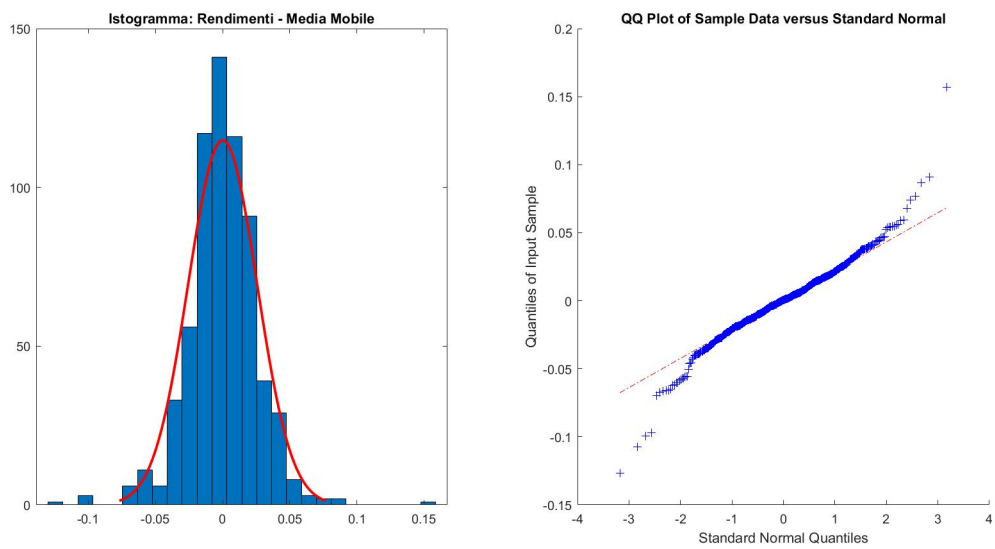


Figura 5.25: Istogramma della nuova serie dei ritorni logaritmici e Q-Q Plot

	<u>P-Value</u>	<u>Test Statistic</u>	<u>Critical Value</u>
Jarque-Bera TEST	0.001	420.1	5.8918

Figura 5.26: Test Jarque-Bera sui ritorni logaritmici della nuova serie in esame

Media	$3.2809 \cdot 10^{-5}$
Mediana	0.00053936
Deviazione Standard	0.025647
Valore Massimo	0.15686
Valore Minimo	-0.12679
Varianza	0.00065776
Coefficiente di variazione	781.7064
Curtosi	6.893
Indice di Asimmetria	-0.03849
1° Quartile	-0.013992
3° Quartile	0.014902

Tabella 5.2: Caratteristiche della distribuzione dei rendimenti in Fase 4

Per quanto riguarda invece la correlazione tra le osservazioni, i grafici delle funzioni di autocorrelazione e autocorrelazione parziale (Figura 5.28), in accordo con il test di Ljung-Box (Figura 5.27), mostrano la presenza di autocorrelazione tra i dati.

	<u>P-Value</u>	<u>Test Statistic</u>	<u>Critical Value</u>
Ljung-Box TEST	7.5391e-06	59.84	31.41

Figura 5.27: Test Ljung-Box sui ritorni logaritmici della nuova serie in esame

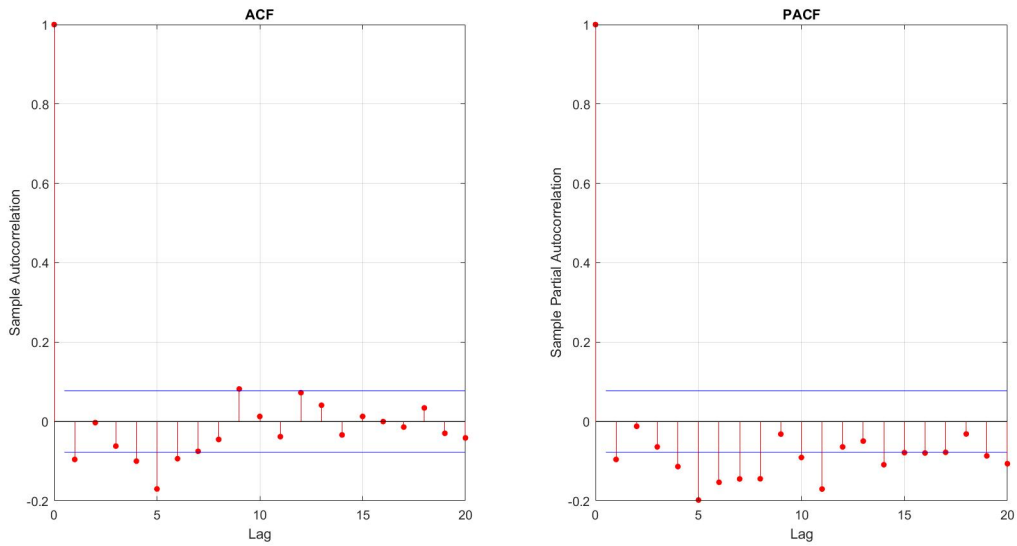


Figura 5.28: Funzioni di autocorrelazione e autocorrelazione parziale campionarie

Per cercare di rimuovere questa autocorrelazione provo ad utilizzare un modello ARMA(p,q).

Tra i diversi ordini p e q, il modello migliore che effettivamente rimuove l'autocorrelazione è ARMA(5,1) con distribuzione del processo white noise t-Student. Dato che i coefficienti ϕ_2 e ϕ_3 sono risultati non significativi, ho optato per un modello ARMA(5,1) imponendo $\phi_2, \phi_3 = 0$ (Figura 5.29).

ARIMA(5,0,1) Model (t Distribution):

	Value	StandardError	TStatistic	PValue
Constant	5.2317e-06	7.917e-06	0.66083	0.50872
AR{1}	0.64939	0.023265	27.913	1.8744e-171
AR{4}	-0.10144	0.030029	-3.378	0.00073013
AR{5}	-0.11847	0.032057	-3.6955	0.00021945
MA{1}	-1	0.0072358	-138.2	0
DoF	6.1553	1.2738	4.8321	1.3511e-06
Variance	0.00049371	3.9217e-05	12.589	2.4214e-36

Figura 5.29: ARMA(5,1) con $\phi_2, \phi_3 = 0$

Andando poi ad analizzare i residui di questo modello, si osserva dal test di Ljung-Box (Figura 5.31) che l'ipotesi nulla viene accettata, ed è quindi verificata l'assenza

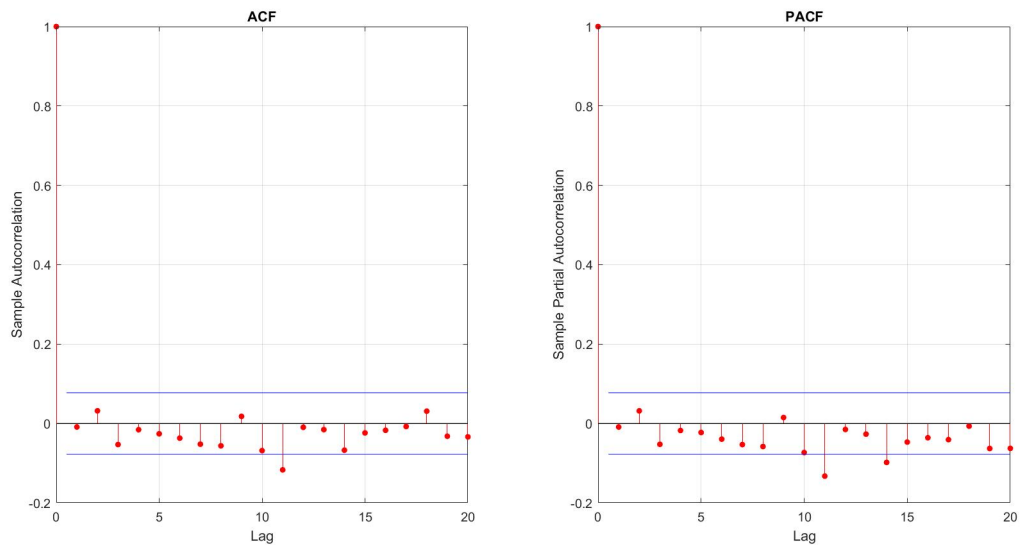


Figura 5.30: Funzioni di autocorrelazione e autocorrelazione parziale dei residui ARMA(5,1) con $\phi_2, \phi_3 = 0$

	P-Value	Test Statistic	Critical Value
Ljung-Box TEST	0.14124	26.791	31.41

Figura 5.31: Test Ljung-Box sui residui ARMA(5,1) con $\phi_2, \phi_3 = 0$

di autocorrelazione tra i residui. Sebbene dal grafico in Figura 5.30 sembra ci sia una correlazione al lag 11, il test non la ritiene significativa.

A questo punto dell'analisi valutiamo la presenza di effetti ARCH sui residui del modello ARMA(5,1). Per far ciò utilizziamo sia il Test ARCH che i grafici delle funzioni di autocorrelazione e autocorrelazione parziale dei quadrati dei residui.

Effettivamente dal test ARCH si evince la presenza di effetti ARCH, di conseguenza specifico i modelli ARCH, GARCH ed EGARCH, per i quali il processo $\{e_t\}$ è distribuito come una t-Student.

Per quanto riguarda l'ordine q del processo ARCH, guardando il grafico della funzione di autocorrelazione parziale dei quadrati delle osservazioni, si notano correlazioni significative ai lag 1,4,8,12. Per evitare di utilizzare un grado polinomiale eccessivamente elevato scegliamo un modello ARCH(4).

Dalla Figura 5.34 si osserva che il coefficiente α_3 non è significativo. Imponendo quindi

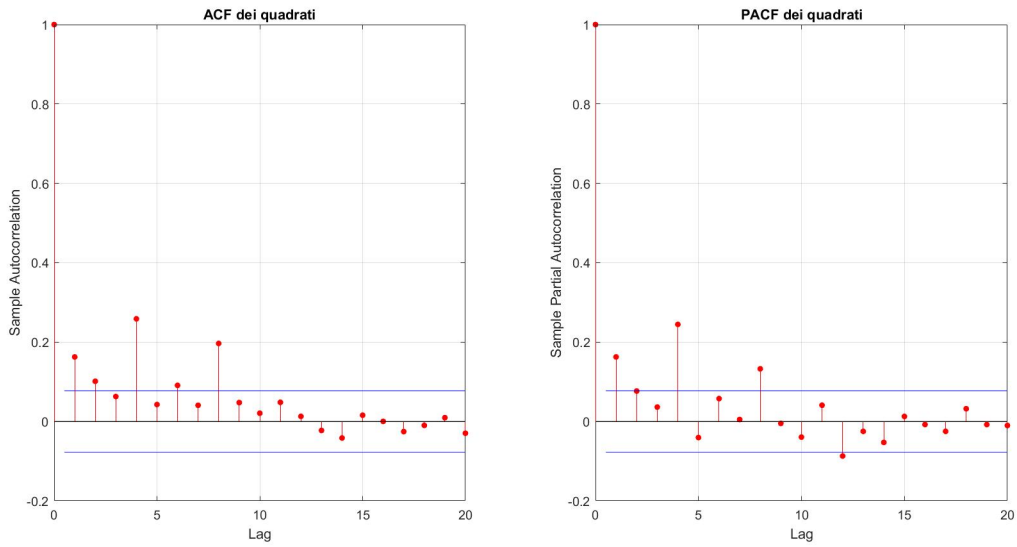


Figura 5.32: Funzioni di autocorrelazione e autocorrelazione parziale dei quadrati dei residui ARMA(5,1) con $\phi_2, \phi_3 = 0$

	P-Value	Test Statistic	Critical Value
ARCH TEST	2.7539e-05	17.581	3.8415

Figura 5.33: Test ARCH sui residui ARMA(5,1) con $\phi_2, \phi_3 = 0$

GARCH(0,4) Conditional Variance Model (t Distribution):

	Value	StandardError	TStatistic	PValue
Constant	0.00025219	3.7454e-05	6.7333	1.6588e-11
ARCH{1}	0.11	0.05271	2.0868	0.036905
ARCH{2}	0.17373	0.074576	2.3295	0.019831
ARCH{3}	0.091854	0.051505	1.7834	0.074519
ARCH{4}	0.11352	0.046858	2.4226	0.015411
DoF	11.59	5.8262	1.9893	0.046669

Figura 5.34: ARCH(4)

GARCH(0,4) Conditional Variance Model (t Distribution):

	Value	StandardError	TStatistic	PValue
Constant	0.00028487	3.474e-05	8.2	2.4031e-16
ARCH{1}	0.12573	0.057119	2.2012	0.027719
ARCH{2}	0.17922	0.075324	2.3793	0.017346
ARCH{4}	0.11198	0.047211	2.372	0.017694
DoF	10.099	4.1334	2.4433	0.014552

Figura 5.35: ARCH(4), con $\alpha_3 = 0$

GARCH(1,1) Conditional Variance Model (t Distribution):

	Value	StandardError	TStatistic	PValue
Constant	5.8789e-05	2.4262e-05	2.4231	0.015389
GARCH{1}	0.72892	0.082086	8.8799	6.6924e-19
ARCH{1}	0.15086	0.046849	3.2202	0.0012809
DoF	10.718	4.1337	2.5928	0.0095191

Figura 5.36: GARCH(1,1)

$\alpha_3 = 0$, stimiamo il nuovo modello ARCH(4) (Figura 5.35).

Proviamo poi ad implementare anche il modello GARCH(1,1) (Figura 5.36) ed EGARCH(1,1) (Figura 5.37). In quest'ultimo modello, seppur tutti i coefficienti sono significativamente diversi da zero, si osserva che il valore di λ è positivo. Questo non è coerente con quanto ci si aspetta dal punto di vista teorico. Infatti, affinché il modello sia in grado di cogliere l'effetto leverage, il coefficiente relativo dovrebbe essere negativo.

Consideriamo quindi due candidati per la serie dei residui del modello ARMA(5,1), ovvero l'ARCH(4) con $\alpha_3 = 0$ e il GARCH(1,1). Guardando i risultati dei criteri di AIC e BIC (Figura 5.38), il miglior candidato risulta essere il GARCH(1,1).

In Figura 5.39 viene mostrata la volatilità stimata dai modelli ARCH(4) e GARCH(1,1). Entrambi i grafici mostrano diversi cluster di volatilità, coerentemente con quanto atteso.

Vediamo ora l'analisi dei residui, che aveva portato a scartare i modelli stimati nell'Approccio 1.

In questo caso, c'è stato un notevole miglioramento, soprattutto per quanto riguarda la simmetria dei residui. Questo si evince sia dal grafico Q-Q Plot che dall'indice di asimmetria. Osservando le statistiche descrittive riportate in Figura 5.42, i residui hanno

EGARCH(1,1) Conditional Variance Model (t Distribution):

	Value	StandardError	TStatistic	PValue
Constant	-0.66787	0.29771	-2.2433	0.024876
GARCH{1}	0.91425	0.038575	23.701	3.5624e-124
ARCH{1}	0.23365	0.066414	3.5181	0.00043466
Leverage{1}	0.096602	0.03687	2.6201	0.0087913
DoF	9.839	3.0969	3.177	0.0014881

Figura 5.37: EGARCH(1,1)

	AIC	BIC
ARCH(4) con $\alpha_3 = 0$	-3240	-3217.5
GARCH(1,1)	-3243	-3225

Figura 5.38: Criteri di AIC e BIC

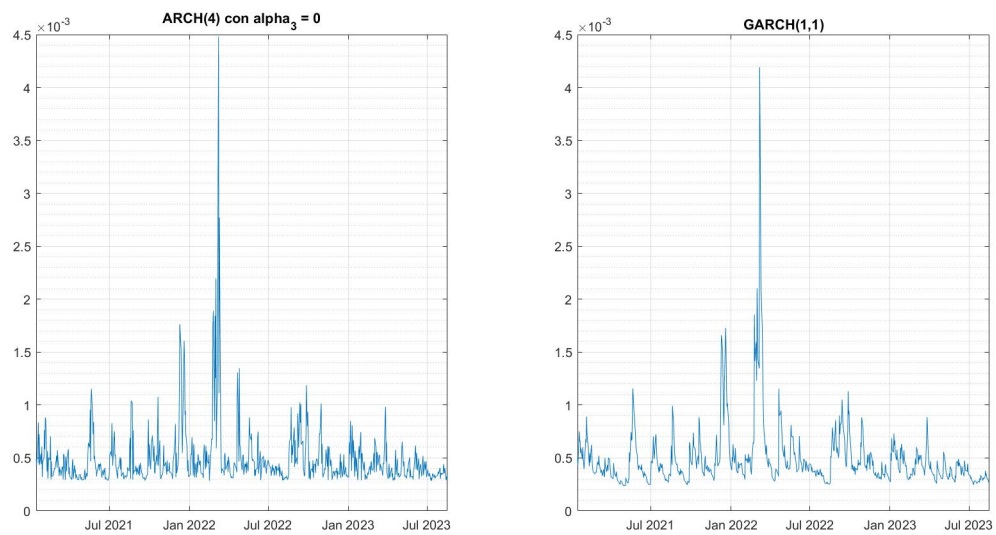


Figura 5.39: Volatilità condizionale stimata dai modelli ARCH(4) con $\alpha_3 = 0$ e GARCH(1,1)

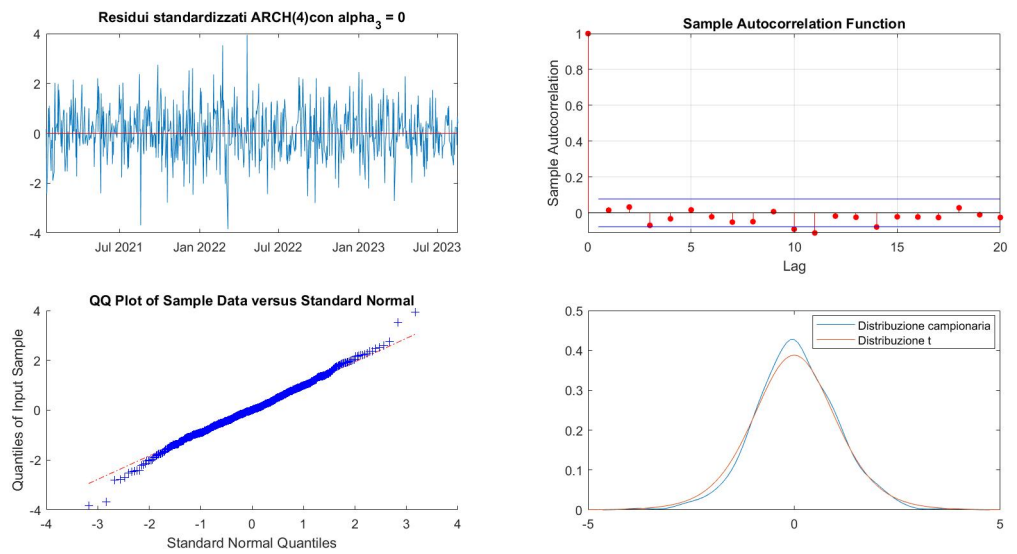


Figura 5.40: Residui standardizzati ARCH(4) con $\alpha_3 = 0$

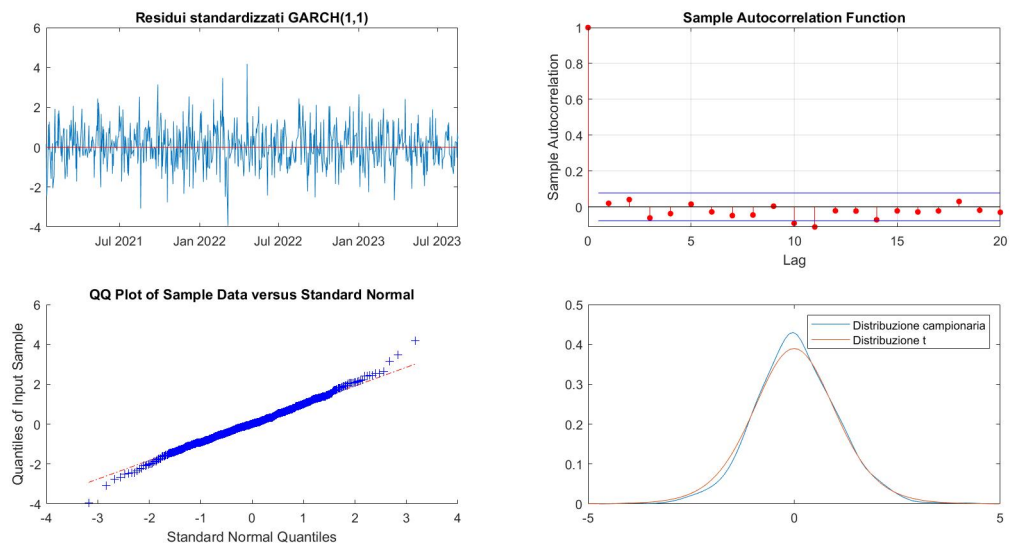


Figura 5.41: Residui standardizzati GARCH(1,1)

RESIDUI ARCH(4)	RESIDUI GARCH(1,1)
Media: 0.037146	Media: 0.043511
Mediana: 0.01106	Mediana: 0.011989
Deviazione Standard: 0.99922	Deviazione Standard: 1.0011
Valore Massimo: 3.9409	Valore Massimo: 4.1753
Valore Minimo: -3.8448	Valore Minimo: -3.9394
Varianza: 0.99844	Varianza: 1.0023
Coefficiente di Variazione: 26.9001	Coefficiente di Variazione: 23.0091
Curtosi: 3.8183	Curtosi: 3.7858
Indice di Asimmetria: -0.010211	Indice di Asimmetria: 0.081822
1° Quartile: -0.58076	1° Quartile: -0.58132
3° Quartile: 0.69172	3° Quartile: 0.67949

Figura 5.42: Statistiche descrittive dei residui standardizzati

	P-Value Ljung-Box
ARCH(4) con alpha_3 = 0	0.069789
GARCH(1,1)	0.077417

Figura 5.43: Test di Ljung-Box sui residui standardizzati

la media pari a 0 e varianza 1, come ci si aspettava. Inoltre, come si osserva dai plot in basso a destra di Figura 5.40 e 5.41, la distribuzione campionaria è simile a quella attesa di una t-Student.

In entrambi i casi il grafico della funzione di autocorrelazione sembra però mostrare alcune correlazioni significative. Tuttavia il test di Ljung-Box con valore del P-Value superiore a 0.05, porta ad accettare l'ipotesi nulla di assenza di autocorrelazione, confermando l'adeguatezza di questi modelli.

Per completezza raffiguriamo graficamente la previsione nei 20 giorni successivi della volatilità, ottenuta dal modello migliore, ovvero ARMA(5,1) - GARCH(1,1). In Figura 5.44 viene indicata in rosso la previsione, mentre in nero la stima della volatilità, che, per rendere il grafico più visibile, non parte dall'inizio di Fase 4, ma circa 400 osservazioni dopo.

5.2.5 Conclusioni

Riassumiamo quindi i tentativi di analisi della serie storica in Fase 4. Le considerazioni iniziali sulla serie dei ritorni logaritmici hanno evidenziato una distribuzione non Normale e l'assenza di correlazioni significative.

Nel primo approccio poi sono stati stimati diversi modelli e quelli ritenuti significativi sono stati ARCH(1), ARCH(5) con $\alpha_1, \alpha_2, \alpha_3, \alpha_4 = 0$, GARCH(1,1) ed EGARCH(1,1).

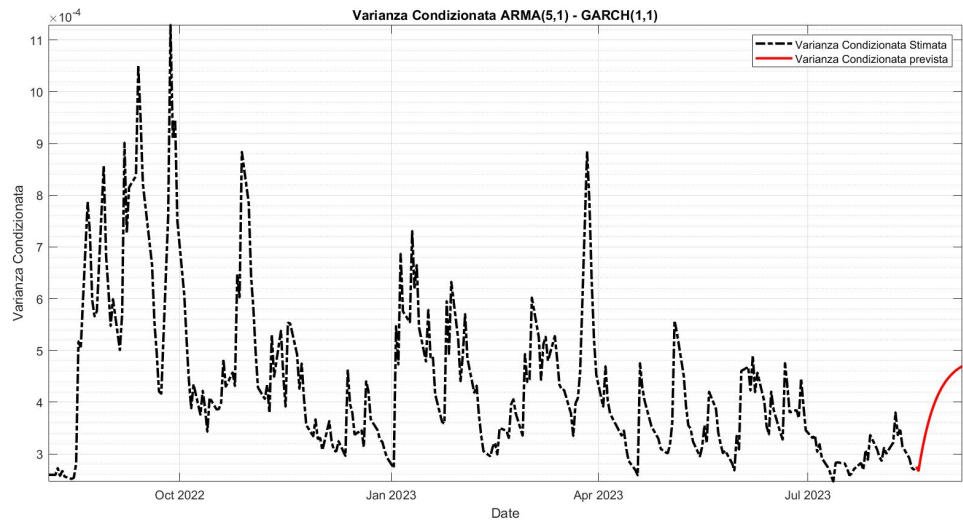


Figura 5.44: Previsione della volatilità

In ciascun modello il processo $\{e_t\}$ era distribuito come una $t - Student$. Nonostante la stima della volatilità ottenuta da ciascun modello fosse coerente con quanto atteso, l'analisi dei residui ha portato a scartare questi modelli. Innanzitutto i cluster di volatilità presenti nella serie dei ritorni logaritmici erano ancora visibili nel grafico dei residui standardizzati, indice che il modello non è riuscito ad assorbirli. Inoltre, la distribuzione campionaria dei residui standardizzati, si discosta notevolmente da quella teorica di una $t - Student$, e mostra un'accentuata pesantezza e asimmetria delle code.

Sebbene questo primo approccio non sia il migliore per descrivere la serie dei prezzi in Fase 4, è risultato ottimale per l'analisi della serie in Fase 3, proposta da Villar Rubio in [18]. Inoltre, i risultati riportati in [18] sono molto simili con quanto trovato in questo elaborato, a differenza del grafico Q-Q Plot dei residui. Per i modelli stimati in Fase 3 infatti, questo mostrava buone proprietà di simmetria, il che ha portato alla scelta di questa classe di modelli come buoni candidati.

La seconda classe di approcci valutava la correlazione settimanale delle osservazioni e la presenza del *day of the week effect*. Ancora una volta si è riscontrato che i modelli osservati non erano opportuni per la descrizione della serie storica in esame.

Infine, il terzo tentativo si è concentrato sull'analisi della serie storica dei ritorni logaritmici a cui è stata tolta la componente di trend tramite una media mobile calcolata con 15 osservazioni. Qui è stato stimato un modello ARMA - GARCH congiunto, in particolare l'ARMA(5,1) - GARCH(1,1) è risultato il migliore. Infatti la volatilità stimata mostra il comportamento atteso e la verifica del modello tramite l'analisi dei residui ha mostrato dei buoni risultati. Questi infatti non mostrano correlazioni significative, hanno buone

proprietà di simmetria e la loro distribuzione campionaria è simile a quella teorica attesa di una $t - Student$.

I risultati ottenuti con quest'ultimo approccio sono molto soddisfacenti, considerata la natura dei dati osservati. Tuttavia, in caso si ambisca ad ottenere modelli più accurati, si potrebbero ad esempio analizzare le serie temporali di osservazioni congiunte, andando a considerare anche variabili correlate alle quote di carbonio, quali il prezzo del petrolio e del gas. Queste informazioni non erano disponibili per l'arco temporale della Fase 4, il che ha fatto propendere per il modello senza tali variabili. Alcuni dettagli in merito relativi alla Fase 3 sono presenti in [17].

Acronimi

ACF funzione di autocorrelazione parziale
ACF funzione di autocorrelazione
ACVF funzione di autocovarianza
CBAM Carbon Border Adjustment Mechanism
CCUS Carbon Capture Utilization & Storage
CDM Clean Development Mechanism
CER Certified Emissions Reduction
CH₄ metano
CO₂ anidride carbonica
COP Conferenze delle Parti
EEX European Energy Exchange
ERU Emissions Reduction Unit
ETS Emissions Trading System
EU ETS European Union Emissions Trading System
EUA European Union Allowances
EUA A European Union Allowances Aviation
GHG Greenhouse Gases
IET International Emissions Trading
IPCC Intergovernmental Panel on Climate Change
JI Joint Implementation
MSR Riserva Stabilizzatrice del Mercato
N₂O protossido di azoto

NDC Nationally Determined Contribution

RBCF Result – Based Climate Finance

UNFCCC Convenzione quadro delle Nazioni Unite sui cambiamenti climatici

Bibliografia

- [1] Peter J. Brockwell, Richard A. Davis, *Introduction to Time Series and Forecasting*, 1996, terza edizione, Springer;
- [2] Peter J. Brockwell, Richard A. Davis, *Time Series: Theory and Methods*, 1991, seconda edizione, Springer;
- [3] Ruey S. Tsay, *Analysis of Financial Time Series*, 2010, terza edizione, Wiley;
- [4] Engle Robert F. *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation*, *Econometrica: Journal of the econometric society*, 1982, 987-1007;
- [5] Bollerslev Tim, *Generalized autoregressive conditional heteroskedasticity*, *Journal of Econometrics*, 1986, 307-327;
- [6] Nelson, Daniel B. *Conditional heteroskedasticity in asset returns: a new approach.*, *Econometrica* 59, 1991, 347-370;
- [7] Rama Cont, *Empirical properties of asset returns: stylized facts and statistical issues.*, *Quantitative finance* 1.2, 2001, 223;
- [8] ISPRA, Istituto Superiore per la Protezione e la Ricerca Ambientale <https://www.isprambiente.gov.it/>;
- [9] COP 26, <https://ukcop26.org/it>;
- [10] COP 27, <https://cop27.eg>;
- [11] A Green Deal Industrial Plan for the Net-Zero Age, Brussels 2023;
- [12] Eni, MAV&F (2021), An overview on Carbon Pricing and ETS: The financial mechanisms to lower CO₂ emissions;
- [13] Eni for 2021 – Neutralità carbonica al 2050 <https://www.eni.com/assets/documents/ita/sostenibilita/2021/eni-for-2021-neutralita-carbonica-2050-ita.pdf>;
- [14] U.S. Bureau of Transportation Statistics, Natural Gas Consumption [NATURAL-GAS], retrieved from FRED, Federal Reserve Bank of St. Louis;

- [15] SENDECO₂, Prezzi CO₂, <https://www.sendeco2.com/it/prezzi-co2>;
- [16] Choudhry Taufiq, *Day of the week effect in emerging Asian stock markets: evidence from the GARCH model*, Applied Financial Economics, 10:3, 2000, 235-242;
- [17] Dhamija, Ajay K., Surendra S. Yadav, and P. K. Jain., *Forecasting volatility of carbon under EU ETS: a multi-phase study*, Environmental Economics and Policy Studies, 19, 2017, 299-335;
- [18] Villar-Rubio, E., Huete-Morales, MD. & Galán-Valdivieso, F., *Using EGARCH models to predict volatility in unconsolidated financial markets: the case of European carbon allowances*, J Environ Stud Sci 13, 2023, 500–509, <https://doi.org/10.1007/s13412-023-00838-5>.