



Università della Calabria

Dipartimento di Economia, Statistica e Finanza "Giovanni Anania"

Corso di

MODELLI E TECNICHE DI PREVISIONE

Elaborato sull'analisi di una serie storica a carattere finanziario

Studente: Micieli Ottavio

Matricola: 214209

Professore

Perri Pier Francesco

Anno accademico 2020/2021

INDICE

1. L'azienda Nintendo LTD	pag.4
2. Analisi preliminare della serie dei rendimenti	
2.1. Serie storica del titolo azionario Nintendo: prezzi e rendimenti	pag.7
2.2. Analisi sui fatti stilizzati	pag.9
3. Individuazione modello per media condizionata	pag.14
4. Individuazione modello per varianza condizionata	
4.1. Algoritmo iterativo per individuazione modello	pag.18
4.2. Disamina migliori modelli individuati	pag.20
5. Approfondimenti sui modelli individuati	
5.1. Analisi dei possibili candidati	pag.23
5.2. Stima modello eGARCH	pag.25
6. Misure di accuratezza e modello ottimale	
6.1. Accuratezza delle previsioni	pag.28
6.2. Modello ottimale	pag.29
7. Backtesting e previsioni future	
7.1. Test sul VaR dell'adattamento del modello	pag.32
7.2. Previsioni future	pag.35



Le origini dell'azienda

Nintendo Co. LTD è una società giapponese specializzata nella produzione di videogiochi e console per l'intrattenimento videoludico. Tra le maggiori compagnie del Giappone, la Nintendo fu fondata il 23 settembre 1889 da Fusajiro Yamauchi, con lo scopo di produrre e, conseguentemente vendere, le *hanafuda*, carte da gioco usate in Giappone. Tra vari cambi di nome e di presidente, come da tradizione giapponese quasi sempre da padre a figlio o al più genero, l'azienda riscuote un discreto successo nel mercato delle carte da gioco e nel 1953 diventa, inoltre, la prima azienda in Giappone a produrre carte plastificate. La prima importante svolta avvenne nel 1959, quando Nintendo riesce a siglare un accordo con Disney per usare i loro personaggi sulle carte. Questo aiuta ad evitare l'associazione dell'azienda al gioco d'azzardo e ne aumenta enormemente il successo, facendola anche entrare nella borsa di Osaka.



Set di carte Hanafuda

Dal 1963 al 1968 Nintendo occupa un insieme di settori, che in generale non hanno alcun che vedere tra di loro, quali quello dei giocattoli, dei taxi, dei Love Hotel e dei ristoranti. Tuttavia, tutti questi settori fallirono negli anni, tranne quello dei giocattoli, avendo esperienza in quel mercato dovuta alla vendita di carte da gioco.

L'entrata nel mercato videoludico



Nintendo Entertainment System

Nel 1972 Nintendo ha il suo primo contatto col nascente mercato dei videogiochi, rilasciando una *light gun* per il *Magnavox Odyssey*, prima console casalinga mai commercializzata. Nintendo capì così quanto proficua sia l'industria videoludica, e decise quindi di lanciarsi nella produzione della loro primissima console casalinga, il Color TV Game 6 (dove 6 stava proprio nel numero di videogiochi preinstallati al suo interno). Nel frattempo, si occupò, anche, dello sviluppo di giochi arcade e altre console. A questo periodo risalgono, infatti, **Donkey Kong** e il certamente più famoso **Mario Bros**. Dal 1982 Nintendo cominciò a lavorare ai primi prototipi del Famicom, rilasciato in Giappone nell'anno successivo e nel

1985 nel resto del mondo col nome di Nintendo Entertainment System (**NES**). L'azienda continuò a collezionare successi dapprima con il Super Nintendo Entertainment System (**SNES**) nel 1990 e successivamente col Nintendo 64 nel 1995.

L'era del Game Boy e dei suoi successori

Ad affiancarsi all'enorme successo globale del NES si aggiunse nel 1989 il **Game Boy**, che arrivò a vendere la bellezza di 118 milioni di unità in tutto il mondo, segnando l'importante ingresso di Nintendo nel mercato delle console portatili. Il Game Boy nel corso del tempo subì numerosi restyling, in cui quello più rilevante fu nel 1998, quando si immise sul mercato il **Game Boy Color**, chiamato così perché poteva gestire più di cento colori contemporaneamente (mentre il Game Boy prevedeva solo giochi in sfumature di grigio, o al più con un unico colore). Nel 2001 arrivò il vero e proprio successore del Game Boy, ossia il **Game Boy Advance**.

Nel 2001 Nintendo rilasciò, inoltre, una nuova console casalinga, il **Game Cube**, con accoglienza generale di pubblico e critica piuttosto positiva.

Nel 2004 ci fu un ulteriore successo dell'azienda a livello globale, che immise sul mercato la seconda console più venduta di sempre, ma prima considerando solo le console portatili, che sancì un'ennesima rivoluzione del mondo videoludico. Attraverso il **Nintendo DS** i videogiocatori scoprirono la maneggevolezza di una console portatile con due schermi, di cui uno con pannello touch screen; console che riuscì a registrare un record di vendite, pari sino ad oggi a 154.02 milioni di unità. Il successo dell'azienda nipponica venne incrementato anche attraverso l'immissione sul mercato del **Nintendo Wii**, nel 2006. La console ebbe, infatti, un apprezzamento molto positivo da parte del mercato, con un numero di unità vendute pari a 101 milioni, nonché riuscì a portare un'ulteriore innovazione nel mondo videoludico. Il Wii doveva tutto il suo successo al suo sistema di controlli di movimento basati sul Wii Remote, controller usato come sistema di puntamento e in grado di rilevare movimenti nelle 3 dimensioni.

A seguito del Nintendo DS venne lanciato il **Nintendo 3DS**, che prevedeva l'aggiunta dell'effetto 3D stereoscopico, senza l'utilizzo di un qualche particolare tipo di occhiali. Tuttavia, tale innovazione risultò poco sfruttata in quanto scomoda, fastidiosa alla vista e in alcune occasioni anche dannosa, il che comportò una perdita di mercato per l'impresa, che dovette correre ai ripari tagliando il prezzo della console solo dopo pochi mesi dall'uscita.



Game Boy



Game Cube



Nintendo Wii e Wii Remote



Nintendo 3DS

Il fallimento più clamoroso e la sua risalita



Nintendo Switch

Nel 2012 l'azienda sperimentò un vero e proprio fallimento, introducendo nel mercato quello che doveva essere il successore del Nintendo Wii. Tuttavia, il Wii U, venne visto più come un accessorio di quest'ultimo piuttosto che una vera e propria console. Complice un sistema operativo poco reattivo e un gamepad non proprio ergonomico, il Wii U registrò delle bassissime vendite, nonostante la presenza di diversi titoli videoludici realizzati appropriatamente per tale console.

Dalle ceneri del Wii U è sorta, però, **Nintendo Switch**, hardware rilasciato nel 2017 e che portò un'ulteriore innovazione nel mondo videoludico, in quanto si tratta di una console sia casalinga che portatile. Grazie all'introduzione di tale hardware innovativo, l'azienda ha visto un notevole incremento del suo fatturato, nonché proprio del prezzo del titolo azionario associato a tale impresa.

Lockdown e intrattenimento videoludico

Dopo il lancio di Nintendo Switch, l'azienda ha visto il suo patrimonio incrementarsi di anno in anno, con stime nuovamente in rialzo anche per quanto riguarda l'anno fiscale 2020-2021. Il lockdown dovuto alla pandemia in corso ha prodotto un ingente aumento delle vendite, riportando in auge la console e i suoi giochi annessi. Si è registrato un successo oltre le aspettative per la console Nintendo Switch, che con 79.87 milioni di unità in tutto il mondo e in nemmeno quattro anni dalla sua uscita ha superato le performance già alte di Nintendo 3DS, la cui produzione e conseguente vendite si è concluso nell'anno appena terminato. Le vendite di Game Boy Advance, Wii e Game Boy sembrano essere degli obiettivi facilmente alla portata di Switch, che gli esperti indicano essere a metà del suo ciclo vitale, per cui ha ancora davanti a sé almeno 3 di anni di vita.



2.1 Serie storica del titolo azionario Nintendo: prezzi e rendimenti

Da un punto di vista finanziario, Nintendo è quotata sulla borsa di Tokyo dal 1983, con il codice “7974”. In particolare, si è scelto di analizzare un arco temporale di 5 anni dei rendimenti del titolo azionario Nintendo, per una finestra che inizia dal primo giorno del 2016 all’ultimo giorno del 2020. Tale periodo storico coincide pressoché con l’arco di tempo trascorso dall’annuncio di Nintendo Switch, che si sta rilevando essere la nuova pioniera delle vendite dell’azienda, specialmente nell’anno appena concluso, in cui i vari lockdown nazionali ne hanno favorito il crescente interesse da parte dei consumatori. Tali eventi hanno portato ad un forte crescita del mercato videoludico in generale; in particolare la crescita dell’azienda Nintendo può essere analizzata osservando il grafico nel quale si riporta l’andamento dei prezzi rispetto al tempo.

Andamento dei prezzi azioni Nintendo LTD



Come si può osservare all’inizio del periodo di osservazioni il prezzo del titolo azionario 7974, legato all’azienda Nintendo, si trova ad un livello medio di circa poco più di 10000¥. Tuttavia, nel secondo semestre del 2016 si può osservare un’impennata del titolo, che in pochi giorni raggiunge un prezzo superiore ai 30000¥, per poi avere una brusca caduta arrestatasi attorno ai 20000¥. Tale aumento repentino del prezzo delle azioni molto probabilmente è legato al rilascio sul mercato dell’applicazione per dispositivi mobili **Pokémon Go**, gioco per smartphone sviluppato da Niantic in collaborazione con Game Freak, The

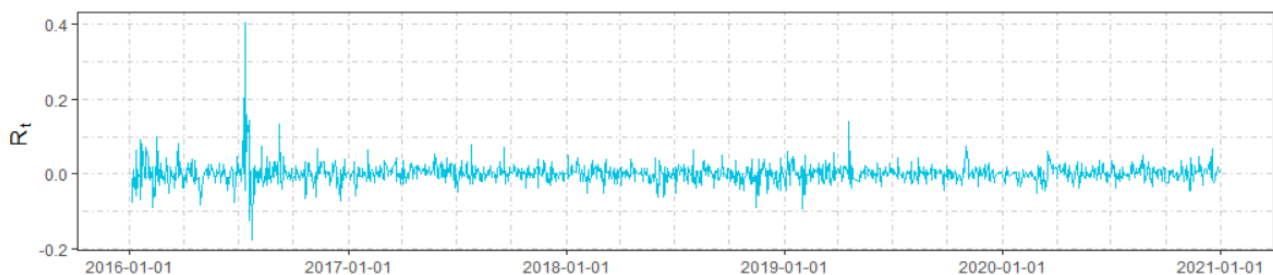


Logo dell’applicazione
Pokémon Go

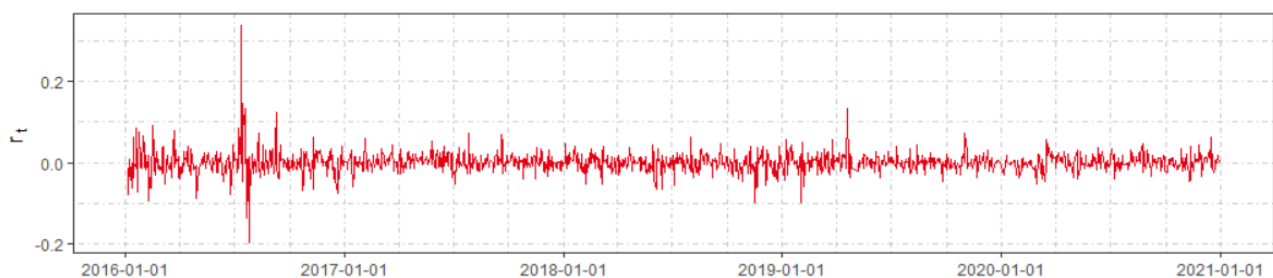
Pokémon Company e Nintendo stesso. Tale applicazione risulta essere ufficialmente rilasciata all'interno degli store virtuali di Google e Apple, per quanto riguarda l'Italia, il 15 luglio 2016. Pokémon Go riscosse subito un notevole successo, attirando verso il mondo dei mostriciattoli tascabili anche coloro i quali non avevano mai intrapreso un'avventura in uno dei giochi della serie principale del brand Pokémon. Tale situazione, a livello finanziario, condusse ad un aumento considerevole del prezzo delle azioni, che però non fu sufficiente a determinare una crescita costante dello stesso. In tale periodo l'impresa stava affrontando gli strascichi del fallimento dovuto a Wii U, anche se si incominciava ad intravedere una nuova opportunità, in termini di investimento, grazie alla presentazione pubblica di Nintendo Switch, avvenuta il 20 ottobre del medesimo anno.

Dal lancio di Nintendo Switch sul mercato mondiale, avvenuto il 3 marzo del 2017, il prezzo delle azioni sperimenta una notevole crescita per tutto l'anno in corso e il primo trimestre dall'anno seguente. I titoli usciti in contemporanea alla console di ottava generazione (ossia Nintendo Switch) hanno fortemente trainato il mercato per il primo anno, ma la mancanza di degni successori in termini videoludici hanno fatto sì che si registrasse una tendenza inversa del tasso di crescita del prezzo del titolo azionario, che è diminuito fino al termine del 2018. Dal 2019 si sperimenta una crescita costante del prezzo del titolo, con la presenza di un picco al ribasso in concomitanza dello scoppio della pandemia globale di COVID-19, avvenuta nel marzo del 2020. Tuttavia, da tale momento, complici i vari lockdown nazionali istituiti per combattere il virus dilagante, il mercato videoludico in generale scopre una nuova fetta di mercato, prima impossibilitata per mancanza di tempo a usufruire di tale tipo di intrattenimento. La crescita del prezzo del titolo azionario è continuata fino a raggiungere il livello di 65830¥ il 12 dicembre 2020.

Andamento dei rendimenti semplici Nintendo LTD



Andamento dei rendimenti logaritmici Nintendo LTD



Tuttavia, per un investitore è più interessante studiare l'andamento dei rendimenti del titolo, piuttosto che del suo prezzo. In maniera particolare, si preferisce analizzare la serie storica del logaritmo dei rendimenti, assieme alla serie storica dei rendimenti semplici.

2.2 Analisi dei fatti stilizzati

Si può facilmente osservare come l'andamento delle due tipologie di rendimenti sia, sostanzialmente, il medesimo, registrando unicamente un cambio di scala. Accanto all'analisi grafica vengono analizzate le proprietà distributive dei log-rendimenti, partendo dall'analisi dei momenti empirici della distribuzione di probabilità legata agli stessi.

Si osservino alcune caratteristiche ricorrenti nell'ambito delle serie storiche di tipo finanziario, partendo dal valore atteso non condizionato. Tale misura risulta essere stimata pari ad un valore molto prossimo allo zero; inoltre, l'intervallo di confidenza per tale parametro risulta contenere proprio il valore 0. Come avviene in molti ambiti finanziari, è un fatto stilizzato che la media non condizionata dei rendimenti sia pari a 0, caratteristica che si verifica anche in tale situazione.

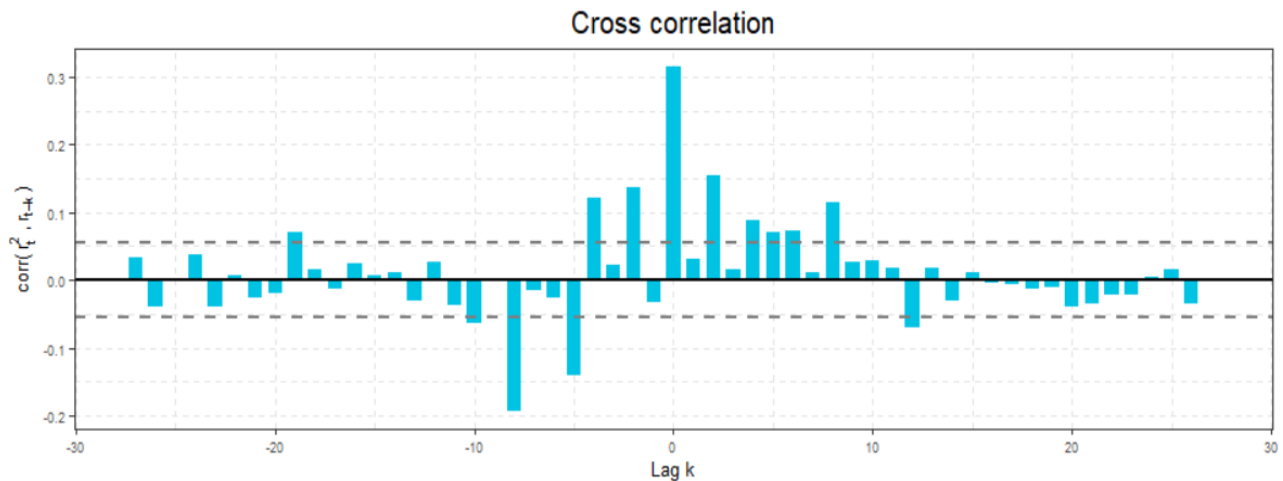
Un altro aspetto da valutare è quello del volatility clustering, caratteristica che si può scorgere osservando la serie storica dei rendimenti. Vi sono delle finestre temporali in cui si riscontra una maggiore volatilità rispetto ad altre in cui le oscillazioni risultano essere più contenute.

Un ulteriore aspetto da valutare è quello della presenza di effetto leva, caratteristica legata alla mancanza di asimmetria all'interno della distribuzione dei rendimenti. Tale mancanza di asimmetria è confermata anche dal test statistico di D'Agostino, in cui si riscontra un p-value inferiore alla soglia di significatività individuata. L'indice di asimmetria della distribuzione dei rendimenti è stimato essere pari a 1.6677, il quale ci suggerisce la presenza di una asimmetria positiva, che va in contrasto con l'usuale definizione di effetto leva. In tale contesto sembrerebbe che le notizie positive che giungono sul mercato generino un aumento di variabilità più di quanto non lo facciano le notizie negative. Tale particolarità può essere legata al contesto economico in cui opera l'impresa che si sta analizzando, ossia quello del mercato videoludico, in generale forte crescita nel periodo considerato e in cui è presente anche una forte concorrenza, il che potrebbe generare una maggiore instabilità proprio nei periodi di crescita. Anche se osservando la serie storica dei rendimenti sembrerebbe che il numero di picchi al ribasso sia maggiore del numero di picchi al rialzo, quest'ultimi risultano avere una magnitudine maggiore rispetto ai primi, che compensa la loro scarsità numerica.

Nobs	1241
NAs	0
Minimum	-0.195017
Maximum	0.339190
1. Quartile	-0.010727
3. Quartile	0.012252
Mean	0.001107
Median	0.001046
Sum	1.374365
SE Mean	0.000737
LCL Mean	-0.000338
UCL Mean	0.002553
Variance	0.000674
Stdev	0.025957
Skewness	1.667722
Kurtosis	28.553356

Una misura grossolana che può essere utilizzata per individuare la presenza di effetto leva all'interna della serie storica dei rendimenti prevede di utilizzare la funzione di autocorrelazione tra il rendimento al quadrato t e i rendimenti ritardati di un numero prefissato di periodo. Di tale funzione di autocorrelazione si fornisce il correlogramma al fine di effettuare un'analisi di tipo grafico.

In generale, tale funzione di autocorrelazione dovrebbe presentare dei valori negativi che si mantengono per un determinato numero di lag, per poi decadere verso lo zero.



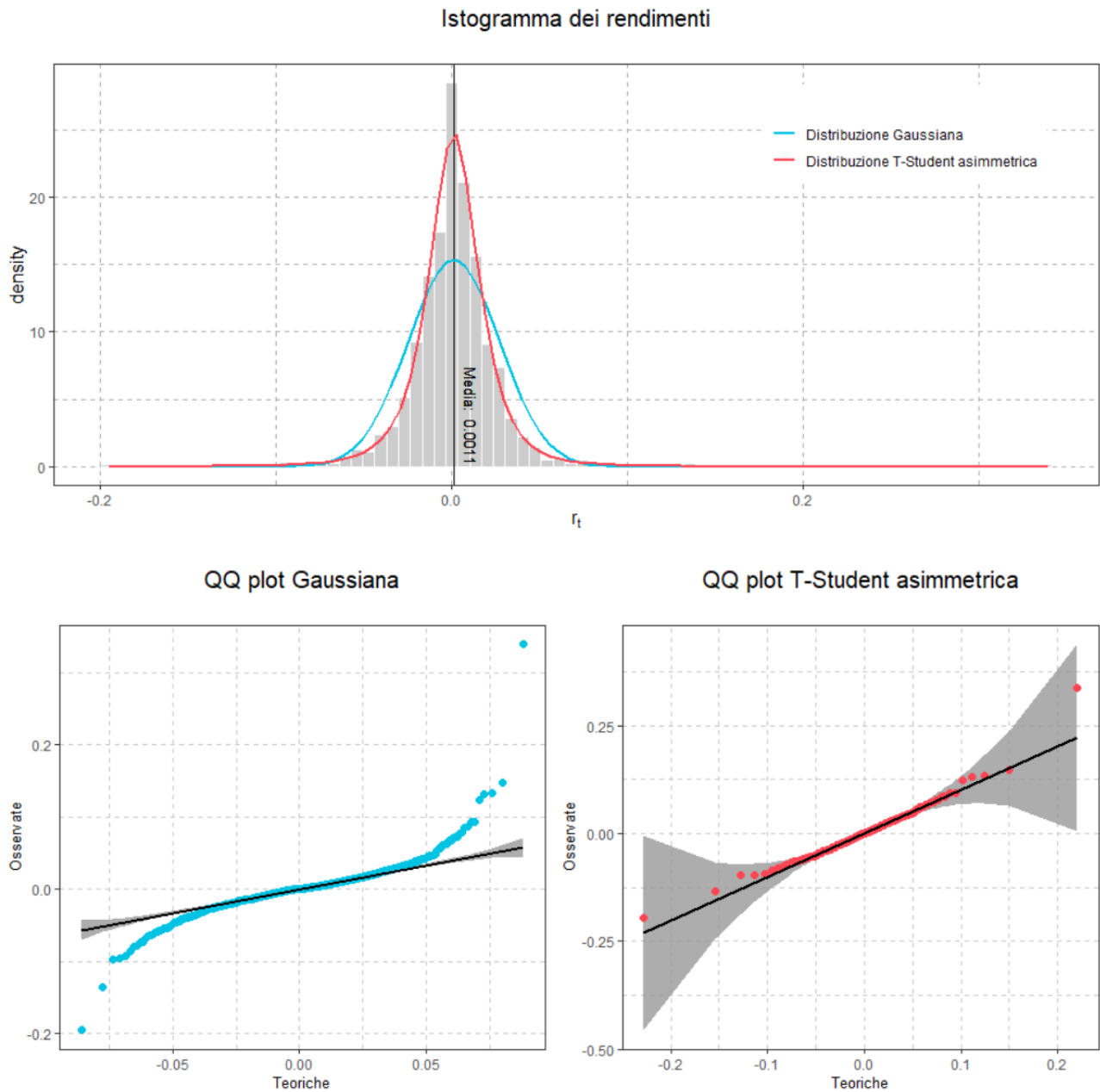
Tuttavia, analizzando tali rendimenti si manifesta una situazione opposta, in cui la funzione di correlazione presenta dei valori positivi per un determinato periodo di lag, per poi annullarsi statisticamente. Questo significa che rendimenti positivi nei periodi precedenti a quello considerato generano un aumento della volatilità mentre rendimenti negativi generano una diminuzione di quest'ultima.

Un altro aspetto da notare è la presenza di una distribuzione leptocurtica, in quanto si è in presenza di picchi positivi e picchi negativi molto grandi rispetto alla media. Tale condizione è riscontrata anche dall'indice di curtosi, il cui eccesso di curtosi è stimato essere pari a 28.55. Il test di Anscombe-Glynn sull'indice di curtosi conferma la presenza di una distribuzione leptocurtica.

Valutando l'indice di asimmetria e quello di curtosi, notiamo un allontanamento da una distribuzione normale, il che ci porta a considerare ulteriori distribuzioni di probabilità. L'allontanamento da una distribuzione normale può essere valutato anche attraverso l'utilizzo di test statistici, quali il test di Kolmogorov-Smirnov, il test di Shapiro-Wilk e il test di Jarque-Bera. Tutti questi test statistici, applicati alla serie dei rendimenti che si sta analizzando, portano a rifiutare l'ipotesi di normalità della distribuzione incognita di tali rendimenti.

Poiché si è stimata la presenza di asimmetria positiva, ma soprattutto di un eccesso di curtosi positivo, quello che si fa è confrontare la distribuzione dei rendimenti sia con una distribuzione Gaussiana, che con una distribuzione T-Student asimmetrica, che generalmente presenta un livello di curtosi maggiore rispetto alla normale. Si riporta l'istogramma di densità di probabilità e il QQ plot relativo alle due distribuzioni considerate. Stimando i parametri che la distribuzione dei rendimenti avrebbe qualora si

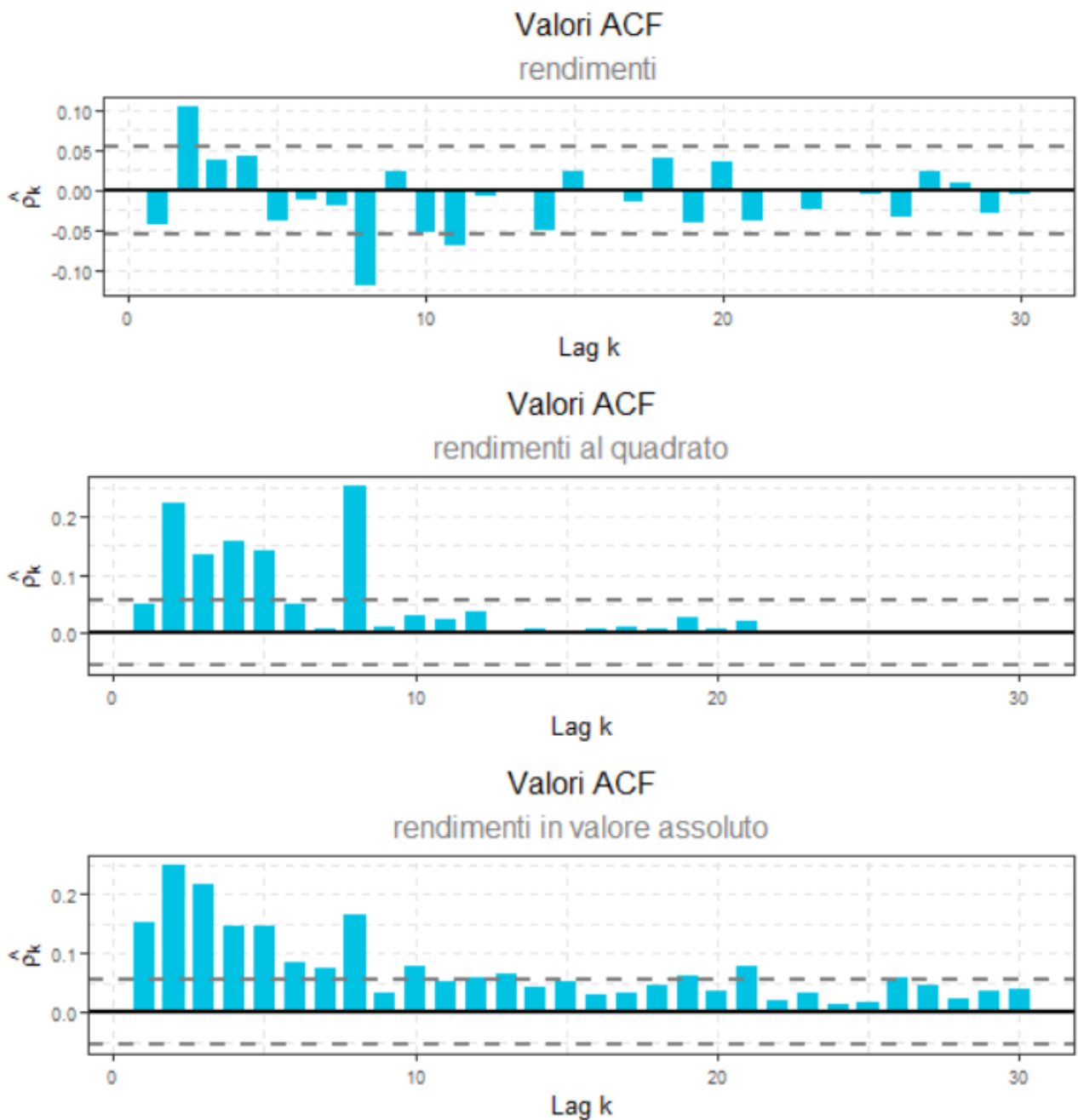
trattasse di una T-Student asimmetrica, notiamo che tale tipo di distribuzione si adatta meglio all'istogramma di densità di probabilità dei rendimenti. Lo stesso aspetto può essere valutato in relazione al QQ plot, in cui tutte le osservazioni risiedono nell'intervallo di confidenza individuato per la distribuzione T-Student asimmetrica.



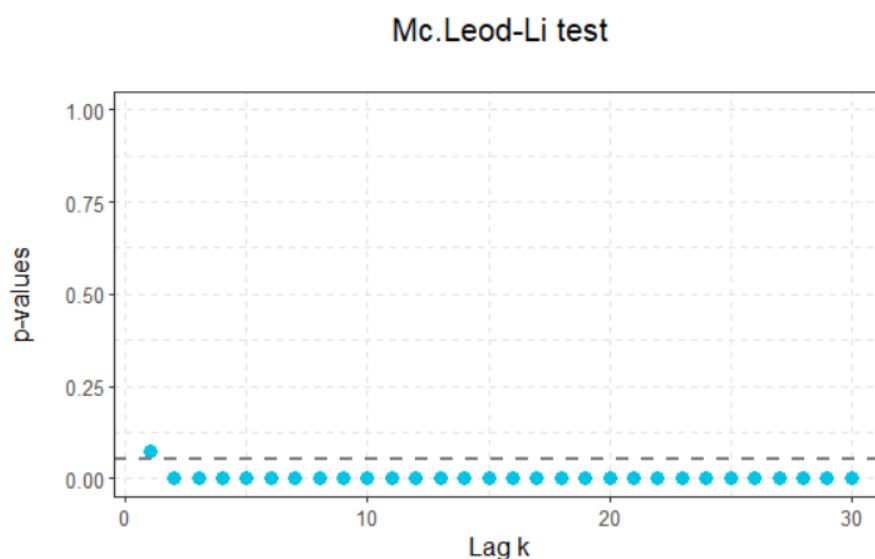
In linea generale, un altro aspetto ricorrente negli studi inerenti alle serie storiche di tipo finanziario, è l'assenza di autocorrelazione seriale tra i rendimenti. Tale condizione si riscontra anche nella serie storica dei rendimenti che si sta analizzando: in particolare si ha un certo livello di correlazione seriale, come anche attestato dal test di Ljung-Box, che tuttavia, dall'analisi del correlogramma, notiamo essere una correlazione sfuria, dovuta alla stima empirica che ne viene fatta.

```
> Box.test(log_rend, lag=30, type="Ljung-Box")  
X-squared = 66.511, df = 30, p-value = 0.0001401
```

L'assenza di correlazione seriale tra i rendimenti non necessariamente indica l'assenza di una qualche altra dipendenza di altro tipo. In particolare, analizzando la correlazione tra i rendimenti al quadrato, oppure tra il valore assoluto dei rendimenti, notiamo che essa per i primi 8 lag circa risulta essere significativamente differente da 0, a ricalcare il fatto che è presente una dipendenza tra i rendimenti, che non è di tipo lineare. Tale presenza di autocorrelazione seriale tra i rendimenti al quadrato e i rendimenti in valore assoluto può essere constatata anche attraverso il test statistico di Ljung-Box, che per entrambi i rendimenti presentano dei valori del p-value nettamente inferiore alla soglia individuata.



L'ultimo aspetto da valutare è la presenza di persistenza all'interna della serie storica. Attraverso l'analisi della funzione di autocorrelazione dei rendimenti al quadrato si può già valutare la presenza di persistenza, in quanto si può notare che tale funzione si mantenga positiva per un periodo di tempo abbastanza lungo e decada lentamente verso lo zero. Tuttavia, tale situazione non si verifica in questo contesto: la funzione di autocorrelazione decade dopo alcuni lag e non si mantiene sempre costantemente positiva. Vi sono anche alcuni test statistici che ci aiutano ad individuare la presenza di persistenza all'interno della serie storica, come il test di McLeod e Li, che verifica la presenza di eteroschedasticità condizionata attraverso l'applicazione del test statistico di Ljung-Box sui rendimenti al quadrato. Tale test statistico fornisce il p-value per ciascuno dei lag indicati, valori che possono essere riportati all'interno di un grafico, di seguito proposto.



Notiamo che, tranne per il primo lag, il p-value risulta essere al di sotto del livello di soglia individuato, pari a 0.05, per tutti i lag a seguire. Tale aspetto è sintomo di presenza di persistenza, di varianza eteroschedastica, all'interno della serie storica, aspetto che può essere ulteriormente testato attraverso il test statistico introdotto da Engle, i cui risultati sono riportati di seguito.

```
> archTest(log_rend, lag=100)
```

Q(m) of squared series (LM test):

Test statistic:	259.4659	p-value:	4.440892e-16
-----------------	----------	----------	--------------

Rank-based Test:

Test statistic:	141.5938	p-value:	0.003971355
-----------------	----------	----------	-------------



Prima di procedere all'individuazione di un modello che possa esprimere l'andamento della volatilità condizionata rispetto al tempo, vi è la necessità di individuare l'equazione della media condizionata. In molte occasioni quello che si fa è porre la media condizionata pari ad una costante, che ancor più spesso tale costante risulta essere pari a 0; tuttavia, si è voluto comunque indagare sul possibile andamento della media condizionata dei rendimenti attraverso l'approccio di Box & Jenkins, che prevede l'uso di modelli indicati con la sigla ARIMA.

Il primo aspetto che si analizza è la presenza di stazionarietà all'interno della serie storica dei rendimenti che si sta analizzando, aspetto che può essere testato attraverso l'uso sia della funzione di autocorrelazione rispetto ai rendimenti, che è stata proposta in precedenza, sia attraverso l'utilizzo di test statistici improntati nell'accertare la presenza di radici unitarie all'interno del polinomio caratteristico derivante dal processo che si presume possa aver generato la serie. L'Augmented Dickey-Fuller test, il test di Phillips-Perron e il KPSS test convergono sul fatto che il processo sia di tipo stazionario, portando a rifiutare l'ipotesi nulla di non stazionarietà qualora l'alternativa sia proprio la stazionarietà dello stesso. Pertanto, non vi è la necessità di ricorrere ad alcuna trasformazione dei dati al fine di rendere stazionaria la serie storica che si andrà ad analizzare.

Un altro aspetto che bisogna analizzare è la presenza di una componente stagionale all'interno della serie storica. Ricordiamo che la serie dei rendimenti riguardanti l'azienda Nintendo presenta un periodo di registrazione pari al giorno. Tuttavia, non per tutti i giorni dell'anno si hanno registrazioni, dovuti ovviamente al fatto che le borse sono chiuse nei fine settimana e nei giorni festivi indetti nella nazione in cui tali borse sono poste, in tal caso il Giappone. Pertanto, il tipo di stagionalità che si ricerca, se presente, è una stagionalità di tipo mensile, quadrimestrale o trimestrale, basando le analisi sulla media aritmetica dei valori all'interno di uno stesso mese. L'esecuzione di test per la presenza di stagionalità porta al rifiuto della presenza della stessa per tutte le frequenze che sono state considerate (rispettivamente 12, 3 e 4). Tale conclusione è rafforzata dal lag plot, che non presenta elementi di sistematicità tali da poter prendere in considerazione la presenza di una componente stagionale all'interno della serie storica dei rendimenti che si sta analizzando. Poiché si stanno analizzando i rendimenti di un'impresa operante nel settore videoludico, si poteva ipotizzare la presenza di una stagionalità di tipo mensile, oppure trimestrale, per la considerazione per cui durante il periodo natalizio la richiesta, e quindi la contestuale vendita di console registra un

aumento. Tuttavia, molto probabilmente l'andamento del settore è caratterizzato più dal periodo di annuncio e uscita dell'hardware di gioco, piuttosto che alle ricorrenze di calendario e, poiché l'immissione sul mercato di nuove console non segue un andamento ben preciso che si ripete di anno in anno, fa sì che non vi sia componente stagionale.

Forti di tali considerazioni circa la stazionarietà della serie storica e dell'assenza di una componente di tipo stagionale all'interno della stessa, si procede all'individuazione di un modello che possa descrivere l'andamento della media condizionata dei rendimenti tra l'insieme di quei modelli indicati con la sigla di ARMA. In prima battuta, il modello migliore viene rintracciato attraverso l'utilizzo della funzione `auto.arima` presente nel pacchetto `forecast` implementato in R. Tale funzione ci fornisce come miglior candidato un modello autoregressivo del secondo ordine.

```
> auto.arima(log_rend)
```

```
Series: log_rend
```

```
ARIMA(2,0,0) with non-zero mean
```

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)
ar1	-0.03935940	0.02822947	-1.3943	0.1632372
ar2	0.10325284	0.028325775	3.6452	0.0002672
intercept	0.00109901	0.00078228	1.4049	0.1600597

```
sigma^2 estimated as 0.0006664 log likelihood=2778.67
```

```
AIC=-5549.34 AICc=-5549.31 BIC=-5528.85
```

Notiamo che tale modello presenta il primo parametro non significativo, così come l'intercetta. La ricerca del miglior modello per descrivere l'andamento della media condizionata viene effettuata anche attraverso l'uso di una funzione iterativa, che stima un modello $ARMA(p, q)$ variando di volta in volta i valori di p e q , analizzando tutte le combinazioni che si originano considerando un valore massimo sia per p che per q pari a 2. La funzione, ad ogni iterazione, registra l'ordine del modello che si sta considerando in quel passo e il valore di AIC e BIC per il modello ARMA che si è stimato; tali valori saranno utilizzati per valutare il miglior modello per descrivere l'andamento della media condizionata dei rendimenti. L'algoritmo, dopo aver valutato 9 differenti modelli, decreta come migliore in base al livello di AIC il modello $ARMA(1,2)$, mentre per l'indice BIC il migliore modello risulta essere un modello a media mobile di ordine 2. Tuttavia, per quest'ultimo modello si può notare come non tutti i coefficienti siano statisticamente significativi, mentre ciò non avviene per il modello $ARMA(1,2)$.

```
> coeftest(Arima(log_rend,order=c(1,0,2)))
```

Z test of coefficients:

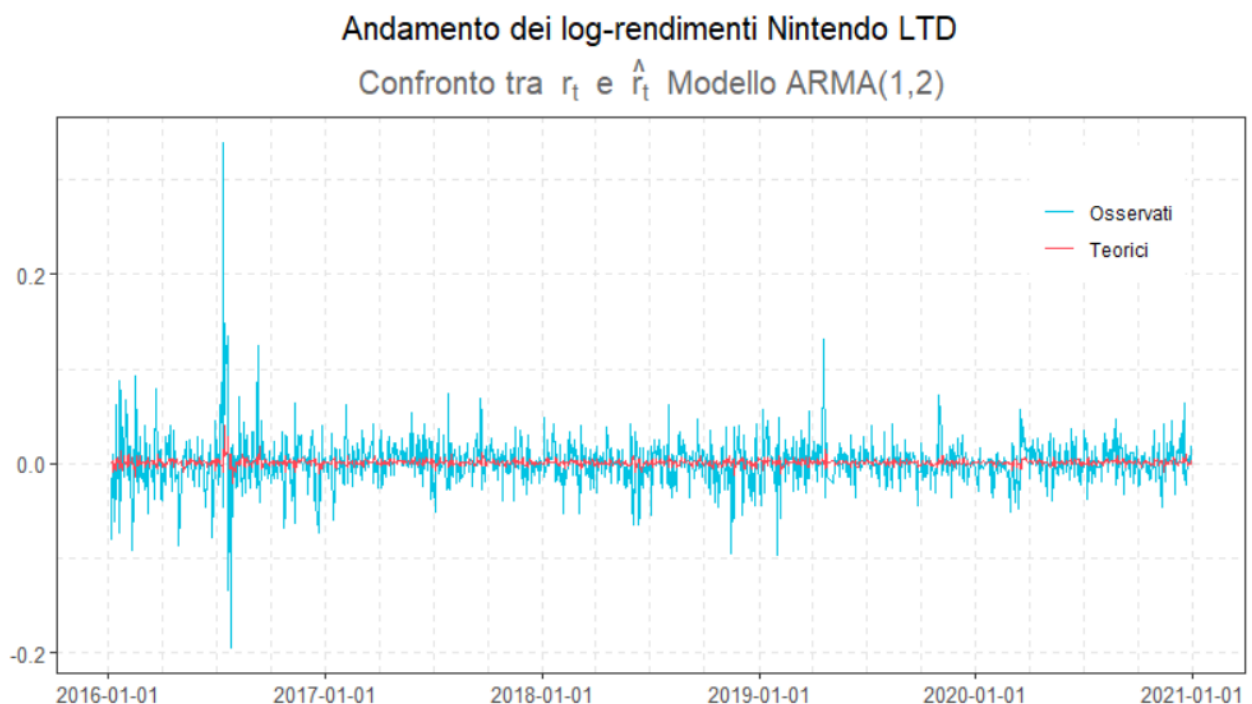
	Estimate	Std.Error	z value	Pr(> z)
ar1	0.42539008	0.15012812	2.8335	0.004604
ma1	-0.47239290	0.14977174	-3.1541	0.001610
ma2	0.12907900	0.02881726	4.4792	7.491e-06
intercept	0.00110336	0.00083535	1.3208	0.186558

```
> coeftest(Arima(log_rend,order=c(0,0,2)))
```

Z test of coefficients:

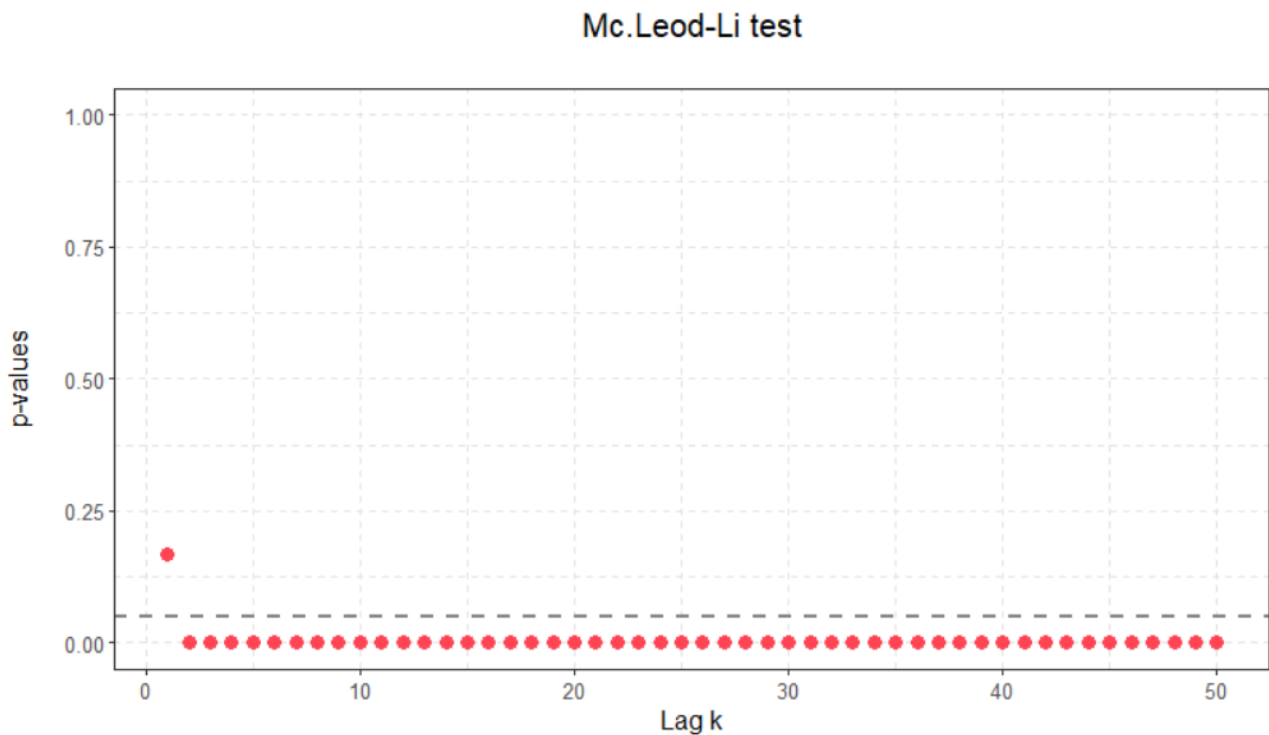
	Estimate	Std.Error	z value	Pr(> z)
ma1	-0.04941630	0.0854331	-1.7313	0.0834028
ma2	0.10276510	0.02757625	3.7266	0.0001941
intercept	0.00110450	0.000777136	1.4319	0.1521782

Dall'analisi che si è condotta si ritiene, pertanto, che il miglior modello, considerando la limitazione circa l'ordine dello stesso, per la media condizionata dei rendimenti sia proprio un modello $ARMA(1,2)$. Anche attraverso l'analisi del grafico in cui si riporta l'andamento dei rendimenti osservati e teorici rispetto al tempo, si può osservare come tale modello riesca a cogliere abbastanza adeguatamente la tendenza di fondo presente.



Tale informazione verrà successivamente sfruttata per l'individuazione del modello migliore che sia in grado di cogliere tutti gli aspetti presenti all'interno della serie storica dei rendimenti circa la loro volatilità condizionata.

Applicando il test di McLeod e Li sui residui del modello relativo alla media condizionata, notiamo che esso suggerisce la presenza di una varianza eteroschedastica, e quindi di persistenza all'interno di essi, come si può osservare dal grafico in cui si riporta il valore del p-value per ogni lag considerato.



Alla medesima conclusione, ossia quella di presenza di effetto ARCH all'interno della serie storica dei residui per quanto riguarda il modello $ARMA(1,2)$, si giunge applicando il test di Ljung-Box sui residui del modello al quadrato, ovvero sul valore assoluto dei medesimi. Entrambi i p-value dei test ci indicano la presenza di una correlazione statisticamente differente da 0 per quanto riguarda il lag massimo considerato.



4.1 Algoritmo iterativo per individuazione del modello

Individuato un possibile modello che possa essere utilizzato come equazione della media condizionata, si continua l'analisi individuando un modello che possa descrivere l'andamento della variabilità condizionata. Tale modello viene ricercato all'interno della classe dei cosiddetti modelli GARCH, introdotti da Bollerslev nel 1986.

La ricerca viene affrontata nuovamente utilizzando un procedimento iterativo, che ad ogni passo stima un nuovo modello sulla base dell'intera serie storica dei rendimenti, ma cambiando di volta in volta non solo l'ordine del modello, ma anche la distribuzione di densità di probabilità a cui, si presume, possa avvicinarsi la distribuzione osservata dei rendimenti. Tale algoritmo iterativo sfrutta, inoltre, un'altra potenzialità del software R, che è quella dell'esecuzione in parallelo delle differenti stime. In generale, l'esecuzione di un comando in R genera un insieme di calcoli ed operazioni di scrittura in maniera che vengono ad essere eseguiti in ordine sequenziale dal processore della macchina su cui si sta lavorando, sfruttando un unico *core* dello stesso e, quindi, non prevedendo la possibilità di sfruttare al pieno la potenza di calcolo della macchina che si ha a disposizione. Tramite l'esecuzione in parallelo si "affida" ad ogni *core* disponibile del processore la stima di un determinato modello pertanto, ad esempio, un computer con 8 *core* potrà stimare in contemporanea 8 differenti modelli, sfruttando al pieno le potenzialità di tale computer. In definitiva si ottiene una notevole diminuzione del tempo di esecuzione dell'algoritmo iterativo che è stato implementato. L'esecuzione in parallelo è resa possibile attraverso l'utilizzo del pacchetto `doParallel` e del comando `parLapply()`. Per ogni modello stimato si registrano un insieme di indicatori che prevedono l'indice AIC, il BIC, lo Shibata e l'Hannan-Quinn, tutti indicatori che valutano il valore della funzione di verosimiglianza stimata e che penalizzano il modello in base al numero di parametri presenti in esso, secondo il concetto di parsimonia.

L'algoritmo iterativo implementato è stato utilizzato in tre differenti occasioni, che differiscono per il fatto che il modello per la media condizionata sia stato fissato o meno.

In primo luogo, si è scelto di individuare il miglior modello per la varianza condizionata della serie storica dei rendimenti fissando il modello per la media condizionata pari al modello $ARMA(1,2)$ stimato nel capitolo precedente. Pertanto, gli unici parametri a variare durante l'esecuzione dell'algoritmo iterativo sono stati l'ordine dei parametri α e dei parametri β relativi alla definizione del modello e la distribuzione di densità di probabilità che veniva ipotizzata per i rendimenti. In particolare, si è scelto di far variare il numero di parametri α da un minimo di 1 ad un massimo di 3 (la funzione di stima

presente nel pacchetto `rugarch` non permette l'esecuzione dell'algoritmo di stima se entrambi i parametri α e β sono fissati pari a 0), mentre il numero di parametri β è libero di variare da un minimo di 0 ad un massimo di 3. Per quanto riguarda le distribuzioni di densità di probabilità che si sono analizzate, esse sono elencate di seguito: distribuzione Gaussiana (indicata con *norm*), distribuzione Gaussiana asimmetrica (indicata con *snorm*), distribuzione T-Student (indicata con *std*), distribuzione T-Student asimmetrica (indicata con *sstd*), distribuzione Gaussiana inversa (indicata con *nig*), distribuzione Generalized error distribution (indicata con *ged*), distribuzione Generalized error distribution asimmetrica (indicata con *sged*) e distribuzione Su di Johnson (indicata con *jsu*). Combinando il numero di parametri con le diverse distribuzioni considerata si arriva a valutare in maniera simultanea 96 differenti modelli, la cui scelta del migliore si basa, in questa prima fase, su quello che presenta i valori meno elevati degli indicatori definiti in precedenza.

Secondo l'indicatore AIC, ma anche secondo l'indice di Shibata il miglior modello per la varianza condizionata risulterebbe essere un modello $GARCH(2,2)$ con distribuzione *sged*; per quanto riguarda l'indice BIC, il migliore modello risulta essere un $GARCH(1,1)$ con distribuzione *ged*; infine, per l'indicatore di Hannan e Quinn il miglior modello risulta essere un $GARCH(2,1)$ con distribuzione *ged*. Tali risultati si ricorda essere frutto del fatto che si è fissato il modello per la media condizionata dei rendimenti.

Successivamente, l'algoritmo iterativo è stato applicato con la volontà di voler stimare quello che viene definito GARCH puro, ossia non considerando un modello specifico per la media condizionata dei rendimenti. Per quanto riguarda il range di variazione del numero di parametri, nonché i tipi di distribuzione di densità di probabilità che sono stati considerati, tali aspetti sono rimasti invariati rispetto l'esecuzione precedente dell'algoritmo, pertanto si avranno nuovamente 96 differenti modelli da valutare.

I risultati che si ottengono dall'esecuzione dell'algoritmo iterativo sono i seguenti: secondo l'indicatore AIC e l'indice di Shibata il miglior modello in grado di descrivere l'andamento della varianza condizionata dei rendimenti risulta essere un $GARCH(2,1)$ con distribuzione *sged*, mentre per quanto riguarda l'indice BIC e l'indicatore di Hannan e Quinn il miglior modello risulterebbe essere un $GARCH(1,1)$ con distribuzione *ged*.

Infine, l'ultima esecuzione dell'algoritmo iterativo per l'individuazione del miglior modello per la varianza condizionata dei rendimenti prevede che non venga fissato alcun modello per la media condizionata, ma che esso venga stimato in contemporanea al modello GARCH per la varianza condizionata. Il numero di parametri p e q varia, per entrambi, da un minimo di 0 ad un massimo di due parametri, così come fatto precedentemente nel momento in cui si ricercava, per l'appunto, un modello per la media condizionata. In questo caso il numero di modelli che vengono confrontati tra di loro sale a 864, ed è proprio in tale fase che si è reso necessario l'adozione di un'esecuzione parallela delle stime dei modelli, al fine di abbattere i tempi di esecuzione dell'algoritmo iterativo. Dall'analisi di tali modelli risulta che per l'indice AIC e per l'indicatore di Shibata il miglior modello risulta essere un modello $GARCH(2,1)$ puro con distribuzione *ged*, mentre secondo l'indicatore BIC e l'indice di Hannan e Quinn si ritiene più adatto un modello $GARCH(1,1)$ sempre puro con distribuzione sempre *ged*.

Riassumendo, i modelli che sono risultati migliori nelle differenti esecuzioni dell'algoritmo iterativo sono i seguenti:

- 1) Modello $ARMA(1,2)$ per la distribuzione della media considerata, unito ad un modello $GARCH(2,2)$ per la distribuzione della varianza condizionata, con distribuzione dei rendimenti pari ad una Generalized Error Distribution.
- 2) Modello $ARMA(1,2)$ per la distribuzione della media considerata, unito ad un modello $GARCH(1,1)$ per la distribuzione della varianza condizionata, con distribuzione dei rendimenti pari ad una Generalized Error Distribution asimmetrica.
- 3) Modello $ARMA(1,2)$ per la distribuzione della media considerata, unito ad un modello $GARCH(2,1)$ per la distribuzione della varianza condizionata, con distribuzione dei rendimenti pari ad una Generalized Error Distribution.
- 4) Modello $GARCH(2,1)$ puro per la distribuzione della varianza condizionata, con distribuzione dei rendimenti pari ad una Generalized Error Distribution.
- 5) Modello $GARCH(1,1)$ puro per la distribuzione della varianza condizionata, con distribuzione dei rendimenti pari ad una Generalized Error Distribution.

4.2 Disamina migliori modelli individuati

Si valutano, in maniera molto sintetica, le informazioni più importanti a tale livello di analisi dei 5 modelli individuati, al fine di decretare il migliore in assoluto.

Modello 1

GARCH model: $sgARCH(2,2)$ Mean model: $ARFIMA(1,0,2)$

Distribution: $sged$

In tale modello il parametro β_2 risulta essere statisticamente non significativo, per cui si potrebbe propendere per l'utilizzo di un modello $GARCH(2,1)$, così come avviene in seguito. Osservando i test relativi alla presenza di effetto ARCH nei residui standardizzati, oppure quelli circa la presenza di autocorrelazione seriale sempre tra i residui standardizzati, per entrambi si è portati all'accettazione dell'ipotesi nulla. I test dei segni sui residui non indicano la presenza di un qualche effetto leva non colto dal modello stimato; inoltre, il test relativo all'aderenza della distribuzione di densità di probabilità ai dati indica che i dati effettivamente è possibile che si distribuiscono secondo una Generalized Error Distribution asimmetrica.

Modello 2

GARCH model: $sgARCH(1,1)$ Mean model: $ARFIMA(1,0,2)$

Distribution: ged

In tale modello tutti i parametri risultano essere significativamente differenti da 0. Non si evidenzia presenza di effetto ARCH latente all'interno della serie dei residui standardizzati, così come non vi è autocorrelazione seriale tra gruppi di quest'ultimi. I test dei segni non evidenziano un qualche effetto leva presente nei rendimenti originali e non

considerato dal modello stimato, così come non si può rifiutare il fatto che la distribuzione dei rendimenti sia approssimabile con una Generalized Error Distribution.

Modello 3

GARCH model: sGARCH(2,1)

Mean model: ARFIMA(1,0,2)

Distribution: ged

I parametri di tale modello risultano essere tutti significativi, a meno del parametro *alpha2*, che risulta significativo se si considera un livello di errore pari al 10%, piuttosto che il 5%. Nuovamente non si riscontano anomalie per quanto riguarda i residui standardizzati di tale modello, e la distribuzione Generalized Error Distribution sembra approssimare in maniera adeguata la distribuzione di densità di probabilità dei rendimenti.

Modello 4

GARCH model: sGARCH(2,1)

Mean model: ARFIMA(0,0,0)

Distribution: sged

In tale modello la media condizionata dei rendimenti viene considerata costante e statisticamente differente da 0; tuttavia, nuovamente, il parametro *alpha2* risulta non essere significativo se si considera un livello di errore del 5%. Non vi è effetto ARCH nei residui standardizzati, così come non vi è correlazione tra quest'ultimi. Ancora una volta il test dei segni ci indica l'assenza di un qualche effetto leva che un modello GARCH per costituzione non può cogliere, mentre si accerta nuovamente l'adattamento dei dati alla distribuzione di densità di probabilità ipotizzata.

Modello 5

GARCH model: sGARCH(1,1)

Mean model: ARFIMA(0,0,0)

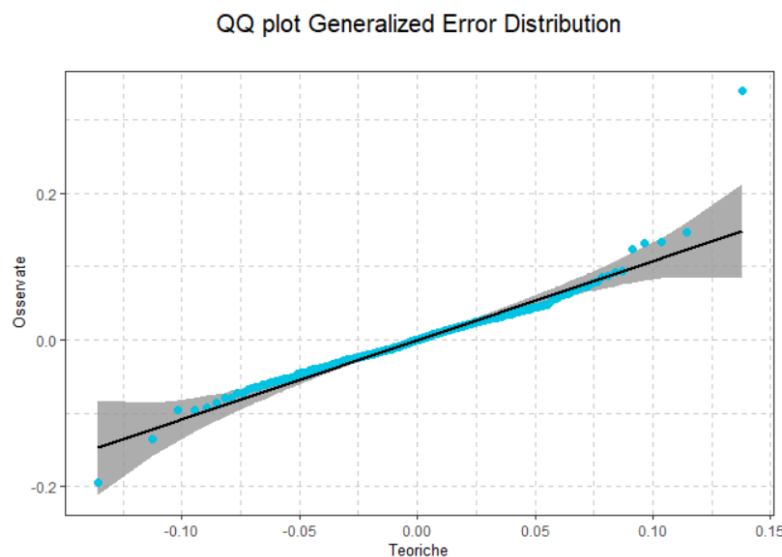
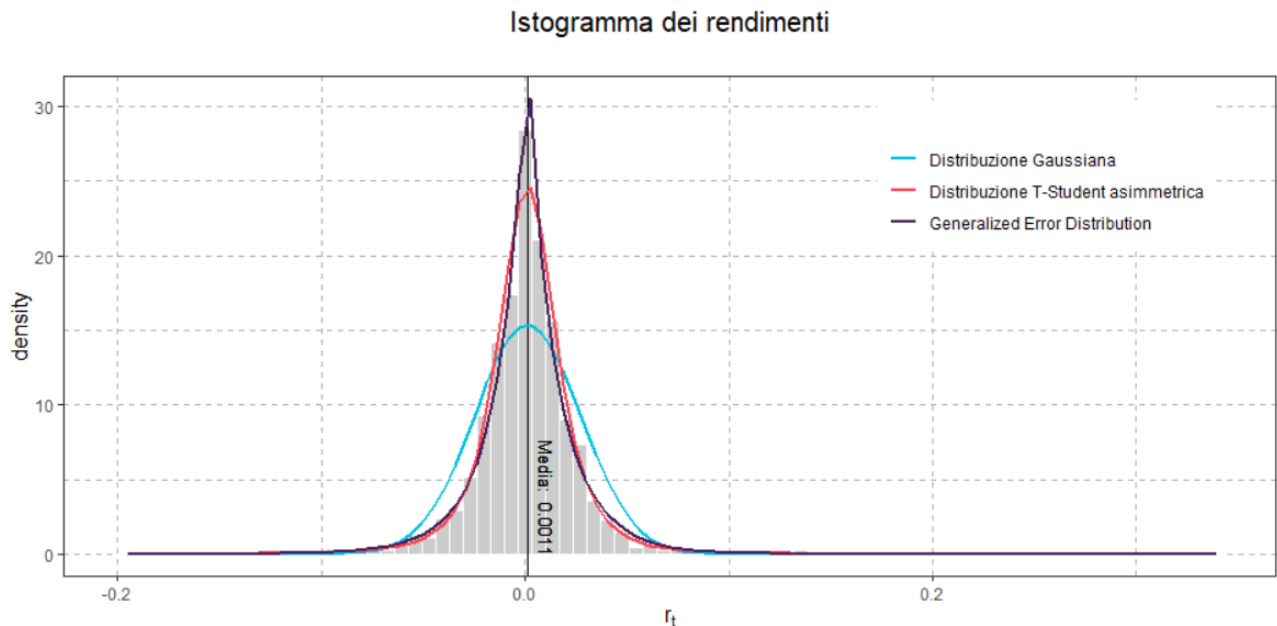
Distribution: ged

In tale modello *GARCH*(1,1) puro i parametri risulta essere tutti significativi, e si rifiuta la presenza di correlazione seriale tra i residui standardizzati, come anche si rifiuta l'assenza di effetto ARCH negli stessi, non colto dal modello stimato. I test dei segni, per una soglia di 0.05, non indicano presenza di effetto leva nei residui, mentre si accerta l'adattamento della distribuzione Generalized Error Distribution ai dati osservati.

In definitiva, si ha che i modelli che risultano essere i più adeguati a descrivere l'andamento dei rendimenti dell'azienda Nintendo LTD risultano essere il modello 2 e il modello 5. La differenza tra quest'ultimi risiede nel fatto che in uno si considera un modello *ARMA*(1,2) per descrivere l'andamento della media condizionata dei rendimenti, mentre nell'alto tale parametro viene posto pari ad una costante.

Prima di considerare quale tra i due modelli sia il migliore in termini di previsioni, si analizza l'adattamento ai dati della distribuzione di densità di probabilità Generalized Error Distribution.

Osservando l'istogramma di frequenze proposto di seguito, si può notare come tale distribuzione si adatti in maniera migliore ai dati osservati rispetto alla distribuzione T-Student asimmetrica, e ovviamente rispetto anche alla distribuzione Gaussiana. La distribuzione GED è in grado di cogliere in maniera più ottimale l'eccesso di curtosi presente nella serie dei rendimenti.

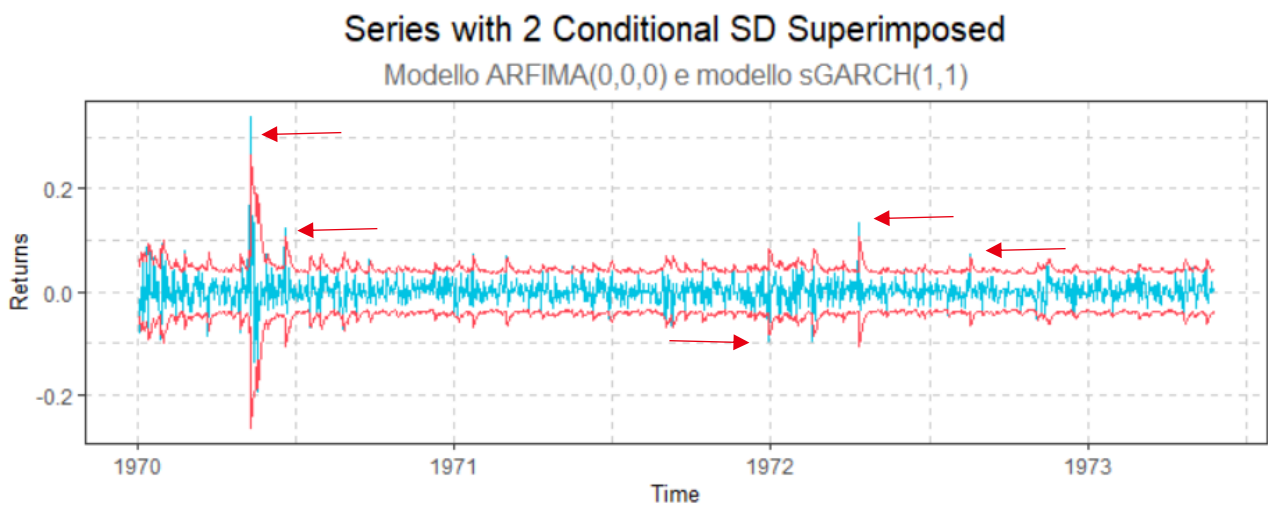
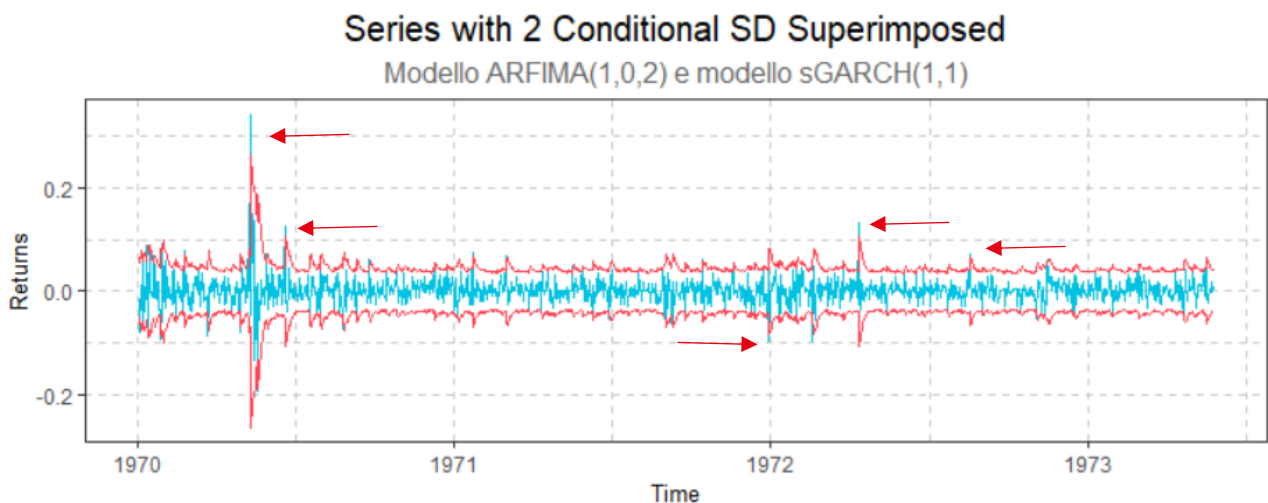


Osservando il QQ plot, si può osservare come i quantili teorici corrispondano in maniera quasi impeccabile ai quantili osservati, con giusto qualche deviazione sulle code.



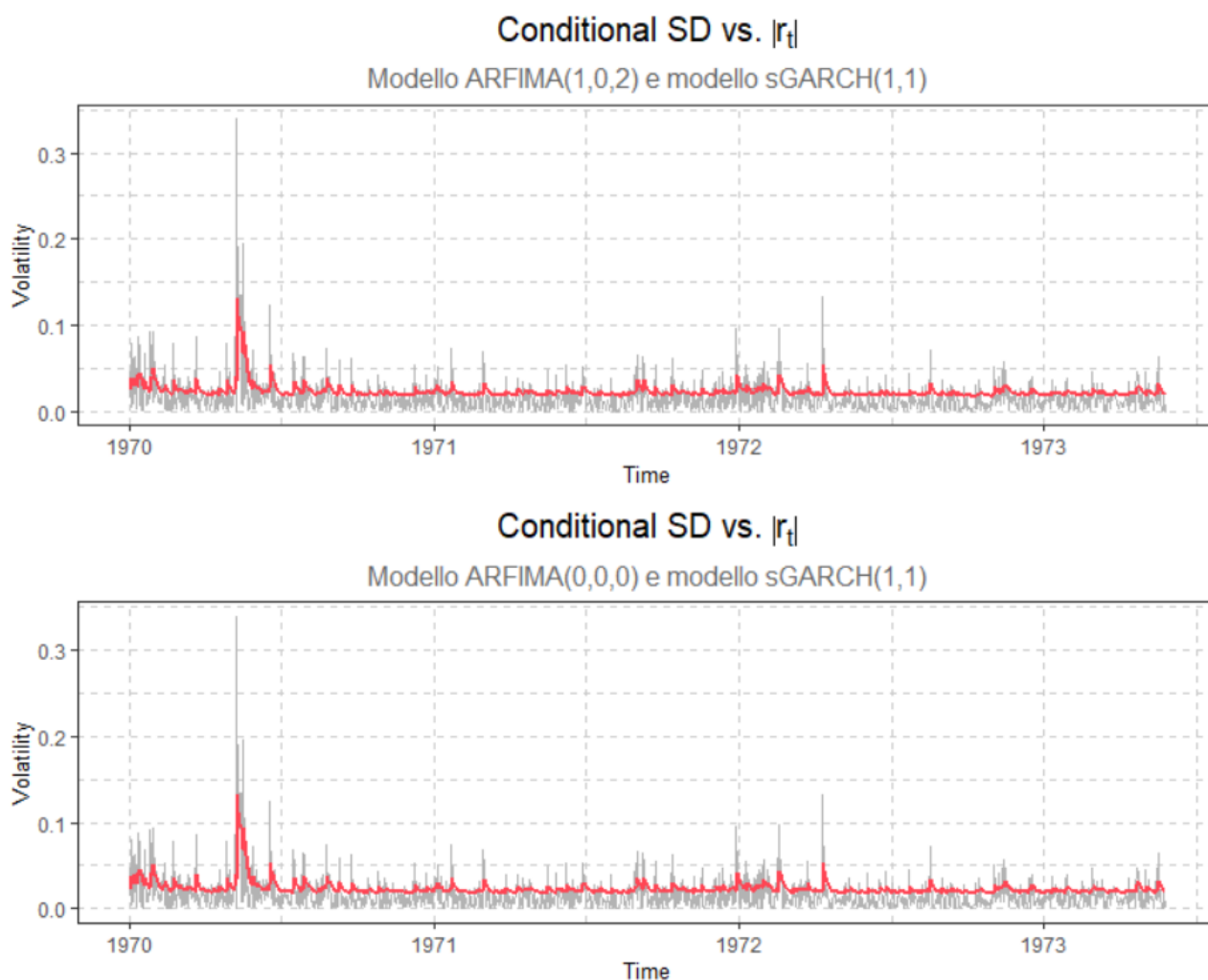
5.1 Analisi dei possibili candidati

Si analizzano i due modelli individuati in termini di aderenza ai dati osservati e di corretta specificazione degli stessi, al fine di individuare quale sia il migliore. Si cerca, anche, di individuare se i modelli rintracciati riescono a descrivere tutte le caratteristiche presenti all'interno della serie storica dei rendimenti dell'azienda Nintendo LTD. Al fine di confrontare i due modelli, ci si avvale in prima battuta di strumenti grafici.



Confrontiamo i grafici in cui si riporta l'andamento dei rendimenti rispetto al tempo e il valore della deviazione standard condizionata stimato, moltiplicato per 2 e per -2 . Tale grafico ci aiuta a comprendere se i modelli stimati siano in grado di catturare i valori estremi presenti all'interno della serie dei rendimenti, ossia tutti quei casi in cui vi sono dei picchi positivi oppure dei picchi negativi. Ovviamente, il confronto effettuato sulla base di tale grafico risulta avere poco senso, in quanto il modello stimato per la volatilità condizionata risulta essere il medesimo per entrambe le specificazioni. In altre parole, i valori stimati per la varianza condizionata, e di conseguenza per la deviazione standard condizionata, sono i medesimi per entrambi i modelli.

L'osservazione che bisogna fare riguarda, invece, il fatto che un modello $GARCH(1,1)$, che sia puro o meno, non riesca a cogliere alcuni picchi, specialmente a rialzo, che sono presenti all'interno della serie storica dei rendimenti. Pertanto, tale modello sembrerebbe non essere del tutto adatto a modellare la varianza condizionata dei rendimenti, anche se allo stesso tempo, vista la sua semplicità, risulta essere abbastanza soddisfacente.



Tramite l'ausilio di quest'alta coppia di grafici, notiamo che il modello stimato segua comunque l'andamento della volatilità, individuata tramite la variabile proxy dei rendimenti in valore assoluto. Tuttavia, non coglie completamente quei casi in cui sono presenti degli eccessi di volatilità.

Da un punto di vista di specificazione del modello, attraverso il test di Hansen, implementato nel pacchetto `rugarch` sotto la funzione `GMMTest()`, possiamo notare come la distribuzione Generalized Error Distribution si adatti in maniera ottimale ai dati osservati e, pertanto entrambi i modelli risultano essere correttamente specificati.

Detto ciò, gli unici elementi, a netto delle misure di accuratezza circa le previsioni, che rimangono al fine di individuare il miglior modello per la varianza condizionata relativi ai rendimenti dell'azienda Nintendo LTD sono gli indicatori quali l'AIC, il BIC, lo Shibata e l'Hannan-Quinn. Secondo tutti tali criteri, il miglior modello risulterebbe essere il *GARCH(1,1)* puro, per il semplice motivo che contiene un numero di parametri inferiore rispetto all'altro modello individuato.

5.2 Stima modello eGARCH

Tuttavia, poiché abbiamo notato la presenza di picchi non completamente catturati dai modelli stimati, si ricerca un modello che riesca a tenere in considerazione valori della volatilità condizionata estremi rispetto a tutti gli altri presenti all'interno della serie. Si è, pertanto, scelto di individuare il miglior modello all'interno della classe dei cosiddetti modelli eGARCH, ossia modelli GARCH esponenziali, la cui caratteristica è modellare il logaritmo della varianza condizionata piuttosto che la varianza condizionata stessa. Si è fatto nuovamente ricorso alla funzione iterativa per la ricerca del miglior modello vista in precedenza, con l'unica differenza di considerare un insieme di funzioni di densità di probabilità minore, escludendo le distribuzioni Gaussiane. Da un insieme di 324 differenti modelli, è risultato il migliore secondo l'indice BIC un modello *eGARCH(1,1)* puro, mentre per tutti gli altri indici il migliore risulta un modello *eGARCH(3,3)* sempre puro. Per quanto riguarda il modello *eGARCH(3,3)* si ha che i parametri di tale modello risultano essere tutti statisticamente significativi. Bisogna tenere in considerazione il fatto che all'interno del pacchetto `rugarch` la formulazione proposta per un modello eGARCH prevede che siano i coefficienti α a determinare l'effetto del segno dei rendimenti.

$$\ln(\sigma_t^2) = \left(\omega + \sum_{j=1}^m \zeta_j v_{jt} \right) + \sum_{j=1}^q \left(\alpha_j z_{t-j} + \gamma_j (|z_{t-j}| - E|z_{t-j}|) \right) + \sum_{j=1}^p \beta_j \ln(\sigma_{t-1}^2)$$

In presenza di un qualche effetto leva con la stessa formulazione che si osserva all'interno dei fatti stilizzati inerenti alle serie storiche di tipo finanziario, ci si aspetta che i parametri α abbiano segno negativo, in accordo alla considerazione per cui shock negativi aumentino la volatilità condizionata più di quando non lo facciano gli shock positivi. Tuttavia, dalla stima risulta che i primi due parametri α presentano segno positivo, a rimarcare la caratteristica di tale serie storica dei rendimenti vista all'inizio di tale discussione. Nell'ambito del mercato videoludico sono gli shock positivi a far aumentare la volatilità dei rendimenti in maniera maggiore di quanto non lo facciano gli shock negativi. Vi è un ulteriore aspetto da tenere in considerazione, che riguarda il fatto che i residui standardizzati derivanti da tale modello presentino per un lag pari a 1 una

autocorrelazione non statisticamente pari a 0. Questo risulta essere un sintomo che il modello non è del tutto adatto a cogliere gli aspetti presenti all'interno della serie storica dei rendimenti. Non si ha, comunque, la presenza di effetto ARCH all'interno dei residui, nonché anche i test dei segni risultano essere tutti non significativi, per cui non si può rifiutare l'assenza di effetto leva all'interno della serie di tali residui.

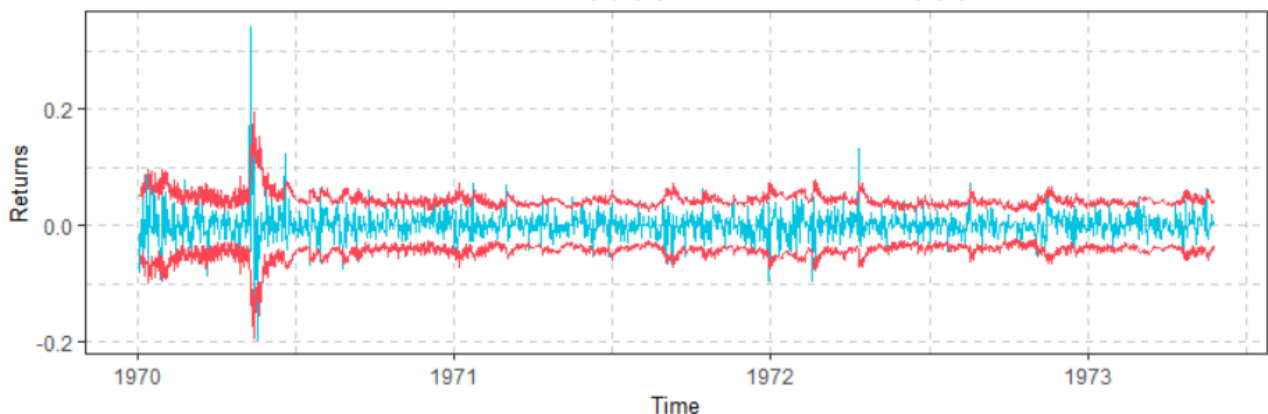
Per quanto riguarda il modello $eGARCH(1,1)$ si può immediatamente notare come il parametro α_1 , che si ricorda essere quello che identifica l'effetto del segno dei rendimenti, risulti essere non significativo. Questo ci indica, a differenza del modello precedente, assenza di effetto leva all'interno della serie storica dei rendimenti. I vari test sui residui forniscono comunque risultati ottimali in relazione ad una corretta specificazione del modello. Tuttavia, il test di Pearson sulla bontà di adattamento per alcuni gruppi di rendimenti risulta essere statisticamente non significativo, pertanto tale modello sembrerebbe non essere molto adatto e la mancanza di effetto leva ci porterebbe nuovamente a considerare un modello $GARCH(1,1)$ puro.

Per quanto riguarda i criteri di scelta del modello più adatto, l'Akaike, lo Shibata e l'Hannan-Quinn concordano sul fatto che il miglior modello risulterebbe essere il modello $eGARCH(3,3)$, mentre l'indicatore Bayesiano indica come miglior modello il $GARCH(1,1)$ puro.

Per quanto riguarda i test di corretta specificazione del modello, sia per l' $eGARCH(3,3)$, che per l' $eGARCH(1,1)$ si hanno esiti positivi.

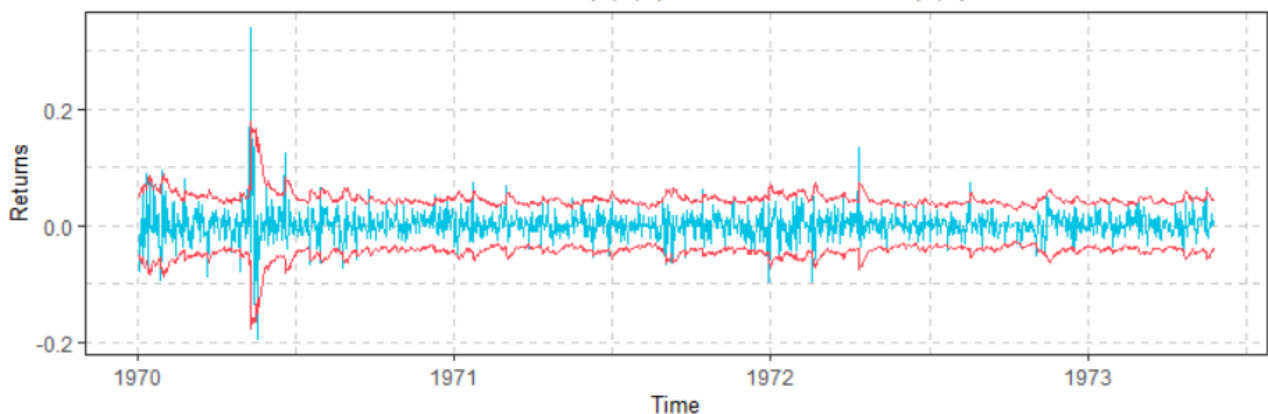
Series with 2 Conditional SD Superimposed

Modello ARFIMA(0,0,0) e modello $eGARCH(3,3)$

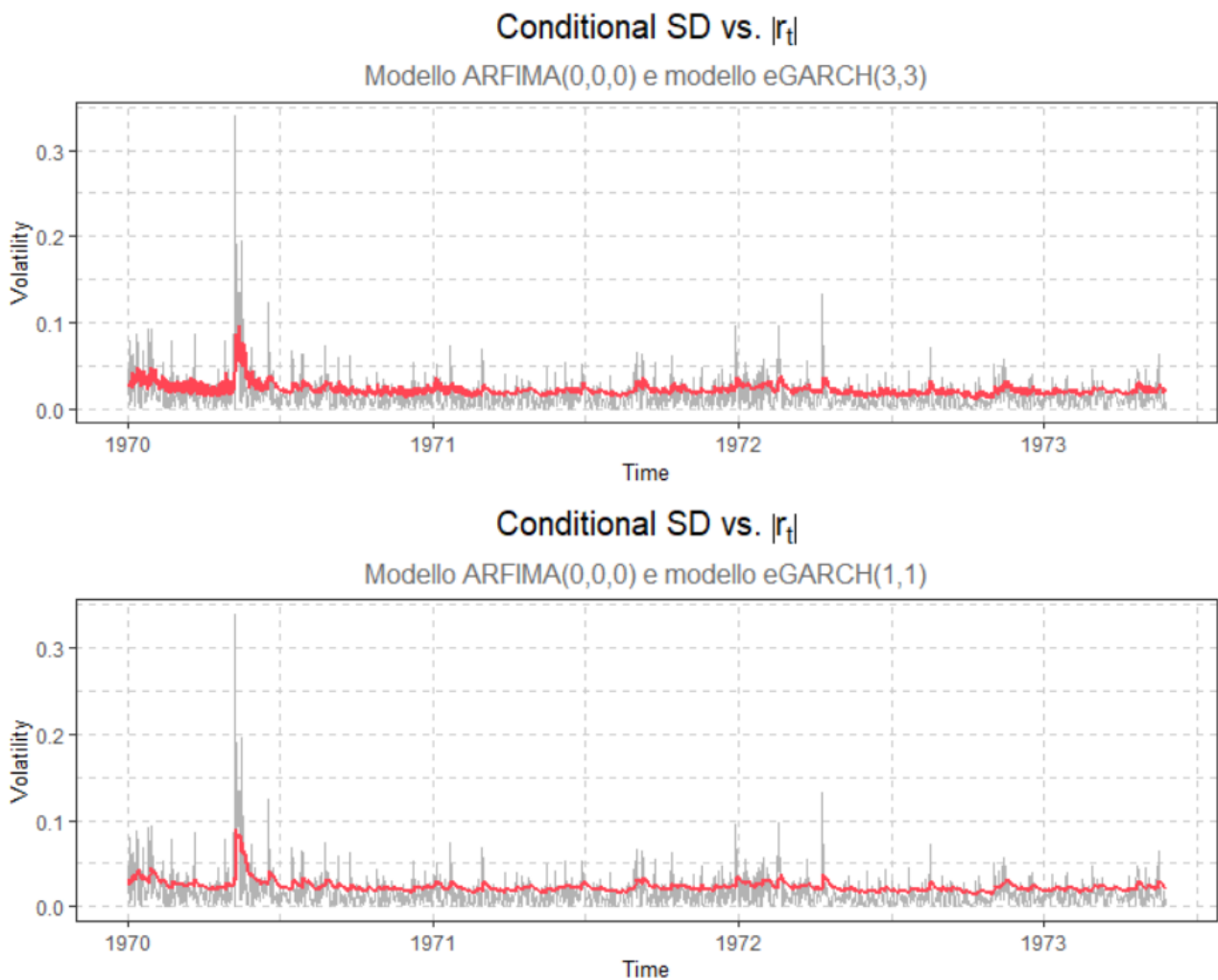


Series with 2 Conditional SD Superimposed

Modello ARFIMA(0,0,0) e modello $eGARCH(1,1)$



Osservando i grafici proposti notiamo che i modelli eGARCH non riescano comunque a descrivere i picchi, sia in positivo che in negativo, presenti all'interno della serie storica dei dati. Il modello *eGARCH*(3,3), tuttavia, sembra essere maggiormente aderente ai dati osservati rispetto agli altri modelli analizzati, il che potrebbe essere un sintomo di *overfitting* del modello.



Stesse considerazione possono essere fatte sulla base di quest'altra coppia di grafici. I modelli GARCH riuscivano a cogliere meglio i picchi presenti rispetto ai modelli eGARCH, che risultano pertanto essere meno adeguati.



6.1 Accuratezza delle previsioni

La valutazione su quale sia il modello migliore da utilizzare anche in vista di previsioni future che si possono ottenere, avviene sulla base dell'accuratezza delle previsioni.

In maniera pratica, quello che si è fatto è stimare nuovamente i parametri dei modelli visti in precedenza, escludendo dal set di dati il mese di dicembre 2020, che verrà utilizzato come partizione per testare l'accuratezza delle previsioni. Successivamente, utilizzando la funzione `ugarchforecast()` si sono previsti i valori per la media condizionata e per la varianza condizionata che la serie storica teoricamente presenta per il mese di dicembre. La previsione sfrutta il meccanismo del *rolling*, in cui ogni previsione avverrà considerando sempre l'intera storia passata della serie. Sfruttando questo meccanismo si hanno delle stime più precise, in quanto l'insieme di informazioni che si hanno a disposizione oggi viene sfruttato per stimare quale sarà il valore della media condizionata e della varianza condizionata di domani. Inoltre, applicando la funzione `fpm()` alle previsioni così ottenute, si riescono ad ottenere le misure dell'Mean Square Error e del Mean Absolute Error; in particolare applicando la radice quadrata all'MSE si ottiene il RMSE, che risulta migliore in quanto si tratta di una misura di errore adimensionale. Si confrontano i risultati che si ottengono sulle previsioni derivanti dai quattro modelli considerati (i risultati sono ordinati partendo dal modello che presenta il minore valore di RMSE).

MODELLO		RMSE	MAE
ARFIMA(1,2)	sGARCH(1,1)	0.01931877	0.01319215
ARFIMA(0,0)	sGARCH(1,1)	0.01959198	0.01291612
ARFIMA(0,0)	eGARCH(1,1)	0.01960097	0.01291393
ARFIMA(0,0)	eGARCH(3,3)	0.01960233	0.01291360

Notiamo che il miglior modello in base all'accuratezza misurata tramite il Root Mean Square Error risulta essere il modello $GARCH(1,1)$, unito ad un modello $ARMA(1,2)$ per quanto riguarda la media condizionata dei rendimenti; invece, per l'indicatore MAE il modello migliore risulta essere il modello $eGARCH(3,3)$ puro. Vi è, tuttavia, il sospetto che quest'ultimo modello abbia un effetto di *overfitting* sui dati. Si testa quindi, nuovamente, l'accuratezza di tali modelli aumentando l'arco temporale di previsione, che passa da un numero di 22 osservazioni, ad un numero di 220 osservazioni. Per le

previsioni viene nuovamente sfruttata la tecnica *rolling*, considerando sempre un passo avanti. I risultati sono di seguito riportati:

MODELLO	RMSE	MAE
ARFIMA(1,2) sGARCH(1,1)	0.01825652	0.01412633
ARFIMA(0,0) sGARCH(1,1)	0.01825725	0.01412674
ARFIMA(0,0) eGARCH(1,1)	0.01825725	0.01412674
ARFIMA(0,0) eGARCH(3,3)	0.01825755	0.01412709

In tale situazione entrambi gli indicatori di errore suggeriscono come miglior modello un modello *GARCH*(1,1) per la varianza condizionata della serie storica dei rendimenti, e un modello *ARMA*(1,2) per la media condizionata della medesima serie.

6.2 Modello ottimale

Visti gli indici di errore e gli indicatori che tengono conto della parsimonia del modello individuato, il miglior modello tra quelli analizzati in grado di cogliere le caratteristiche presenti all'interno della serie storica dei rendimenti del titolo azionario dell'azienda Nintendo LTD, risulta essere il modello *ARMA*(1,2) per la media condizionata unito al modello per la varianza condizionata *GARCH*(1,1). Di seguito si riportano le stime sui parametri e i test effettuati sui residui.

```

*-----*
*          GARCH Model Fit          *
*-----*

```

Conditional Variance Dynamics

```

-----
GARCH Model:      sGARCH(1,1)
Mean Model:       ARFIMA(1,0,2)
Distribution:      ged

```

Optimal Parameters

```

-----
      Estimate  Std. Error   t value   Pr(>|t|)
mu      0.001066    0.000147    7.2445    0.000000
ar1     0.461445    0.023961   19.2579    0.000000
ma1    -0.492555    0.023737  -20.7502    0.000000
ma2     0.004272    0.002094    2.0398    0.041372
omega   0.000092    0.000043    2.1509    0.031482

```

alpha1	0.142392	0.056405	2.5244	0.011588
beta1	0.692020	0.119119	5.8095	0.000000
shape	0.982422	0.050345	19.5139	0.000000

Robust Standard Errors:

	Estimate	Std. Error	t value	Pr(> t)
mu	0.001066	0.000058	18.3811	0.000000
ar1	0.461445	0.011355	40.6386	0.000000
ma1	-0.492555	0.011844	-41.5876	0.000000
ma2	0.004272	0.000237	18.0596	0.000000
omega	0.000092	0.000059	1.5600	0.118752
alpha1	0.142392	0.076061	1.8721	0.061196
beta1	0.692020	0.165698	4.1764	0.000030
shape	0.982422	0.068112	14.4237	0.000000

LogLikelihood : 3042.961

Information Criteria

Akaike	-4.8912	Bayes	-4.8581
Shibata	-4.8912	Hannan-Quinn	-4.8787

Weighted Ljung-Box Test on Standardized Residuals

	statistic	p-value
Lag[1]	0.5522	0.4574
Lag[2*(p+q)+(p+q)-1][8]	4.8940	0.2512
Lag[4*(p+q)+(p+q)-1][14]	7.4132	0.4795
d.o.f=3		
H0 : No serial correlation		

Weighted Ljung-Box Test on Standardized Squared Residuals

	statistic	p-value
Lag[1]	0.3368	0.5617
Lag[2*(p+q)+(p+q)-1][5]	4.5204	0.1959
Lag[4*(p+q)+(p+q)-1][9]	5.7098	0.3329

d.o.f=2

Weighted ARCH LM Tests

	Statistic	Shape	Scale	P-Value
ARCH Lag[3]	0.1354	0.500	2.000	0.7129
ARCH Lag[5]	0.3297	1.440	1.667	0.9323
ARCH Lag[7]	0.7188	2.315	1.543	0.9545

Sign Bias Test

	t-value	prob sig
Sign Bias	0.4679	0.6399
Negative Sign Bias	0.3081	0.7581
Positive Sign Bias	0.7777	0.4369
Joint Effect	0.8574	0.8357

Adjusted Pearson Goodness-of-Fit Test:

group	statistic	p-value(g-1)
1 20	21.93	0.2876
2 30	24.34	0.7119
3 40	45.67	0.2145
4 50	44.70	0.6481

Tenendo conto della stima dei parametri, le equazioni che definiscono il modello possono essere così esplicitate

$$r_t = \mu_t + \varepsilon_t$$

$\varepsilon_t \sim WN$ distribuzione GED con media 0 e varianza σ_t^2

$$\mu_t = 0.001066 + 0.461445 r_{t-1} - 0.492555 \varepsilon_{t-1} + 0.00422 \varepsilon_{t-2}$$

$$\sigma_t^2 = 0.000092 + 0.142392 r_{t-1}^2 + 0.69202 \sigma_{t-1}^2$$



7.1 Test sul Var dell'adattamento del modello

Non ci rimane, pertanto, di utilizzare il modello individuato precedentemente al fine di prevedere la volatilità futura dei rendimenti non prima, però, di aver testato l'accuratezza delle previsioni attraverso l'utilizzo del metodo del *backtesting*.

Si tratta di test retrospettivi, basati sul confronto tra i dati osservati circa i rendimenti della serie e quanto viene stimato attraverso l'utilizzo dei modelli individuati. La logica sottostante al backtesting di un modello per la varianza condizionata dei rendimenti è relativamente semplice: se il modello è qualitativamente adeguato, da confronto periodico della stima giornaliera del VaR con le perdite giornaliere dell'attività di negoziazione relative al giorno successivo dovrebbe emergere che le perdite effettive risultino superiori al VaR con una frequenza coerente con quella definita dal livello di confidenza. La maggior parte delle tecniche di backtesting si basa su test d'ipotesi. Nel caso in cui l'ipotesi nulla venga rigettata, le previsioni del VaR non presentano le caratteristiche richieste e, pertanto, tali previsioni sono da considerarsi inadeguate.

In termini pratici, quello che si fa è stimare nuovamente il modello prescelto non considerando un determinato arco temporale, su cui successivamente verranno effettuati i test. In questo caso si è scelto di non considerare l'ultimo anno di osservazione. Successivamente, tramite la funzione `ugarchroll()` si è stimato il valore del VaR sempre considerando unicamente il giorno seguente e sfruttando le informazioni contenute nella prima partizione della serie. La funzione permette l'utilizzo della tecnica *sliding window* per cui ad ogni iterazione la partizione su cui viene stimato il modello "scivola" di un passo avanti, modello che verrà utilizzato per predire il valore futuro della varianza condizionata, ma anche della media condizionata, per un unico lag. Inoltre, ad ogni iterazione vengono ristimati i parametri del modello individuato, iterazioni che continuano fino a che non si ha un numero di previsioni pari al numero di osservazioni non incluse nella partizione cosiddetta di training.

Sfruttando la stima del VaR sia puntuale, che in riferimento al livello di confidenza indicato, si possono effettuare alcuni test sfruttando il report che ci viene fornito.

Il primo test previsto è il test di Kupiec, che è basato sull'esame della frequenza con cui le perdite superano il VaR. In pratica, ciò che si intende sottoporre a verifica empirica è se la frequenza delle eccezioni empiricamente rilevate è significativamente diversa da quella "teorica" desiderata, ossia dal livello di *alpha* indicato durante le previsioni del VaR. Tuttavia, in generale il test di Kupiec richiede un campione composto da un elevato

numero di dati (si parla di 10 anni di dati giornalieri) per poter generare risultati affidabili. In particolare, la potenza statistica di questo test è piuttosto bassa.

Il test di Christoffersen, invece, valuta anche se le eccezioni presenti siano serialmente indipendenti. In altri termini, si accerta che la probabilità che si verifichi un'eccezione in un determinato giorno sia indipendente dal fatto che il giorno precedente si sia o meno verificata un'altra eccezione. In generale, la metodologia di backtesting proposta da Christoffersen risulta più completa ed efficiente di quella analizzata in precedenza. I risultati che si ottengono testando il modello per la media condizionata $ARMA(1,2)$ unito al modello per la varianza condizionata $GARCH(1,1)$ sono i seguenti

VaR Backtest Report

```
=====
Model:                sGARCH-ged
Backtest Length:      242
=====
alpha:                1%
Expected Exceed:      2.4
Actual VaR Exceed:    1
Actual %:              0.4%
```

Unconditional Coverage (Kupiec)

```
Null-Hypothesis:      Correct Exceedances
LR.uc Statistic:      1.081
LR.uc Critical:       3.841
LR.uc p-value:        0.299
Reject Null:          NO
```

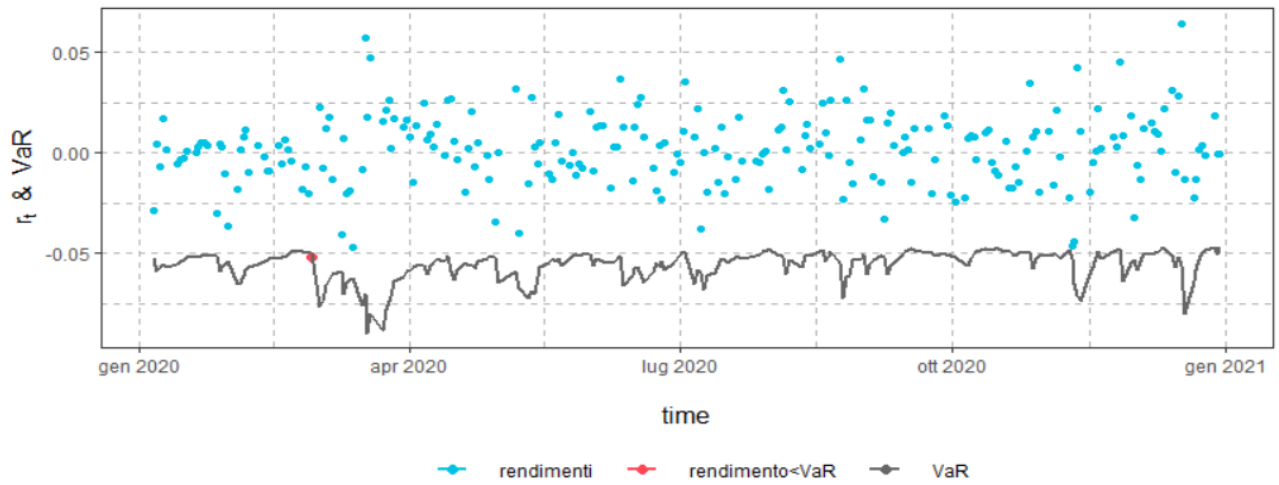
Conditional Coverage (Christoffersen)

```
Null-Hypothesis:      Correct Exceedances and Independence of Failures
LR.cc Statistic:      1.089
LR.cc Critical:       5.991
LR.cc p-value:        0.58
Reject Null:          NO
```

Notiamo che entrambi i test non consentono il rifiuto dell'ipotesi nulla, per cui le eccezioni presenti, che come si può anche osservare dal grafico riportato di seguito si manifesta una sola eccezione, sono coerenti con il livello di significatività prescelto e, inoltre, non dipendono le une dalle altre.

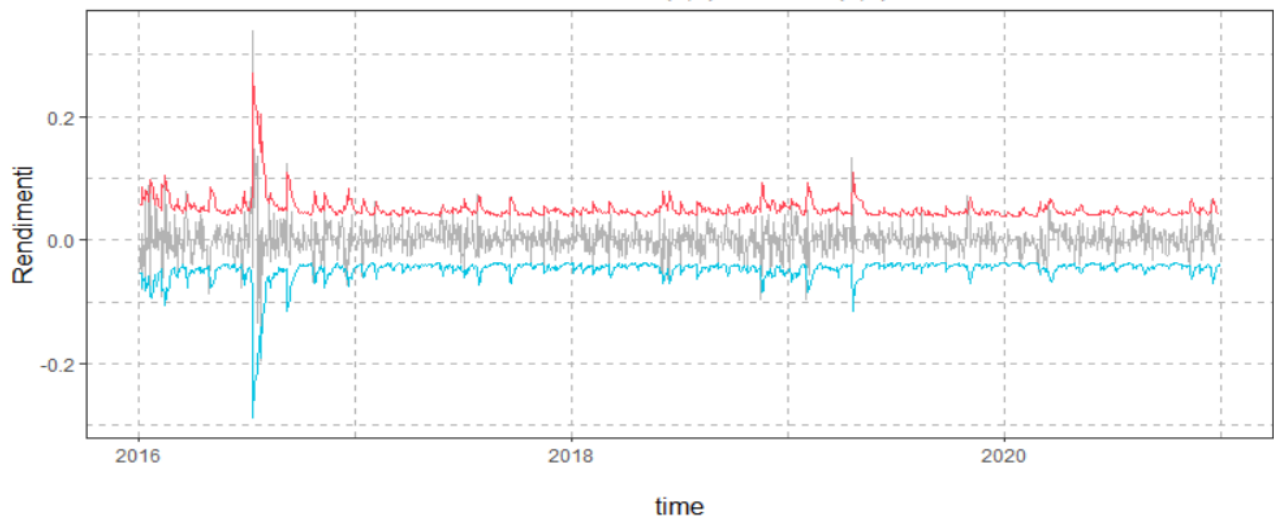
Rendimenti vs previsioni VaR

Modello ARFIMA(1,2) e GARCH(1,1)



Rendimenti con VaR all'95% di confidenza

Modello ARFIMA(1,2) e GARCH(1,1)



Rendimenti con VaR all'99% di confidenza

Modello ARFIMA(1,2) e GARCH(1,1)

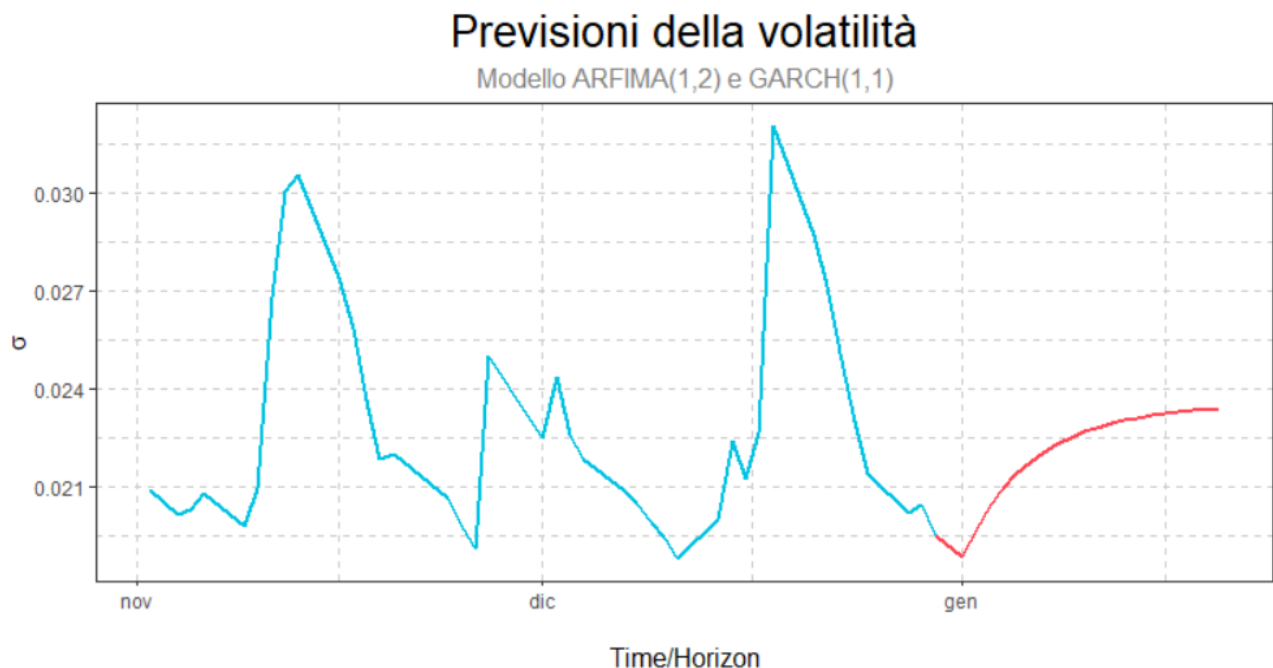


L'intervallo di confidenza del VaR per un livello di significatività del 99% contiene interamente la serie storica dei rendimenti, mentre ciò non avviene se si restringe l'intervallo di confidenza al 95%.

Per un livello di confidenza del 99%, un agente finanziario che investe nelle azioni dell'azienda Nintendo LTD stima una perdita che risulta essere superiore alla reale perdita che invece ha subito. Anche nei momenti di maggior volatilità, il rendimento negativo che si è registrato non supera mai quello che si è preventivato. Di contro, si ha che i rendimenti positivi che si sono previsti non si realizzano mai, rimanendo ad un livello sempre inferiore. Per un intervallo di confidenza del 95% vi sono alcune situazioni che un agente finanziario possa aver perso, o guadagnato più di quanto non avesse preventivato, ma si tratta comunque di scostamenti non tanto significativi.

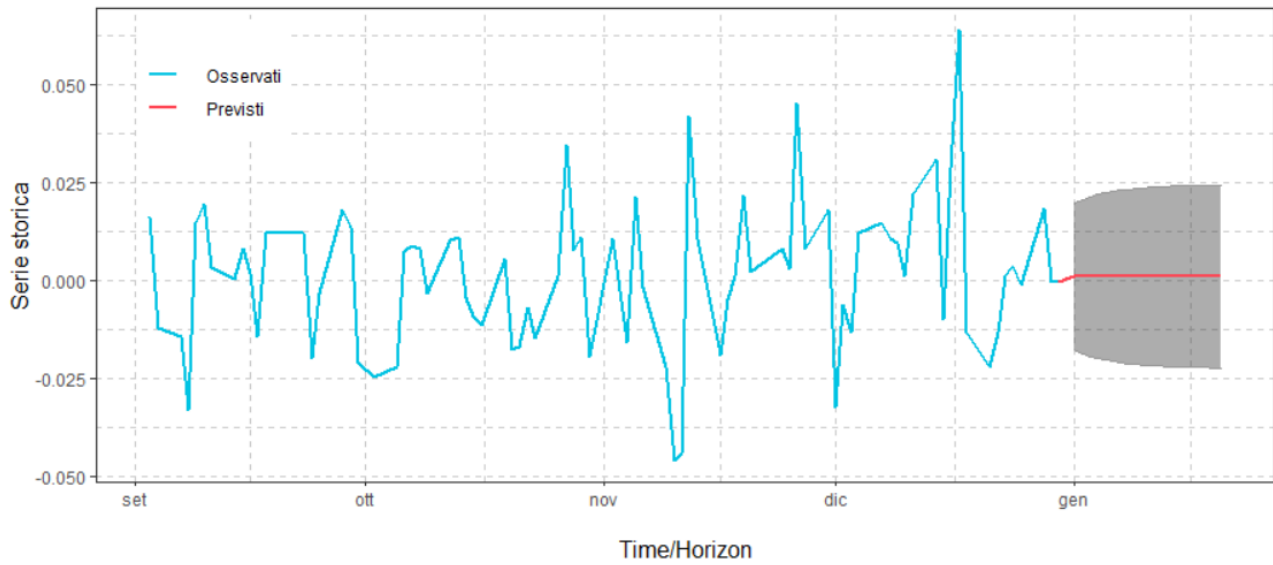
7.2 Previsioni future

Terminiamo l'analisi con le previsioni della volatilità condizionata, nonché della media condizionata, per un arco di tempo pari a 20 giorni successivi al termine del periodo di osservazione.



Previsioni dei rendimenti

Modello ARFIMA(1,2) e GARCH(1,1)



```

*-----*
*      GARCH Model Forecast      *
*-----*

```

```

Model: sGARCH      Horizon: 20
Roll Steps: 0      Out of Sample: 0

```

0-roll forecast [T0=2020-12-30]:

	Series	Sigma		Series	Sigma
T+1	0.001043	0.01887	T+11	0.001066	0.02284
T+2	0.001049	0.01972	T+12	0.001066	0.02296
T+3	0.001058	0.02040	T+13	0.001066	0.02306
T+4	0.001062	0.02095	T+14	0.001066	0.02314
T+5	0.001064	0.02140	T+15	0.001066	0.02320
T+6	0.001065	0.02177	T+16	0.001066	0.02326
T+7	0.001066	0.02207	T+17	0.001066	0.02331
T+8	0.001066	0.02232	T+18	0.001066	0.02335
T+9	0.001066	0.02253	T+19	0.001066	0.02338
T+10	0.001066	0.02270	T+20	0.001066	0.02341