**Parte 3: Preparazione dei Dati e Assunzioni**

**3.1 Raccolta dei Dati**

Il dataset fornito contiene dati storici sul DAX, che rappresentano prezzi o indici di mercato rilevanti. La raccolta e la gestione adeguata di tali dati sono fondamentali per analizzare le dipendenze tra strumenti finanziari tramite modelli di copula. Utilizzeremo questi dati per calcolare i rendimenti giornalieri, un passaggio necessario per la modellazione delle dipendenze.

**3.2 Pulizia e Pre-elaborazione dei Dati**

I dati finanziari spesso includono anomalie come valori mancanti o outlier che devono essere gestiti prima dell'analisi. Implementeremo le seguenti tecniche di pulizia dei dati:

Gestione dei Valori Mancanti: Rimuoveremo eventuali righe con valori mancanti per evitare distorsioni.

Gestione degli Outlier: Utilizzeremo tecniche di filtraggio per identificare ed eliminare gli outlier, assicurando che l'analisi si concentri sui valori centrali più rappresentativi.

*import pandas as pd*

*# Caricamento dei dati dal file CSV*

*data = pd.read\_csv('/mnt/data/DAX\_3Y-1M.csv', index\_col='Date', parse\_dates=True)*

*# Rimozione dei valori mancanti*

*data = data.dropna()*

*# Verifica e gestione degli outlier tramite interquartile range (IQR)*

*Q1 = data.quantile(0.25)*

*Q3 = data.quantile(0.75)*

*IQR = Q3 - Q1*

*filtered\_data = data[~((data < (Q1 - 1.5 \* IQR)) | (data > (Q3 + 1.5 \* IQR))).any(axis=1)]*

**3.3 Trasformazione dei Dati**

Dopo la pulizia, è necessario trasformare i dati per ottenere una scala uniforme. Poiché i modelli di copula richiedono margini uniformi, trasformeremo i dati in rendimenti logaritmici per ottenere stazionarietà e calcoleremo i punteggi standardizzati:

*import numpy as np*

*# Calcolo dei rendimenti logaritmici*

*log\_returns = np.log(filtered\_data / filtered\_data.shift(1)).dropna()*

*# Standardizzazione dei dati: calcolo dei punteggi Z*

*standardized\_data = (log\_returns - log\_returns.mean()) / log\_returns.std()*

**3.4 Normalizzazione**

Per applicare correttamente i modelli di copula, i dati devono essere trasformati in una distribuzione uniforme sull'intervallo [0, 1]. Questo passaggio permette ai dati di adattarsi meglio alla funzione di copula che verrà utilizzata per modellare le dipendenze:

*from scipy.stats import norm*

*# Normalizzazione tramite la funzione di distribuzione cumulativa (CDF)*

*uniform\_data = norm.cdf(standardized\_data)*

**3.5 Assunzioni nei Modelli di Copula**

Per l'uso corretto dei modelli di copula, è essenziale discutere alcune assunzioni chiave:

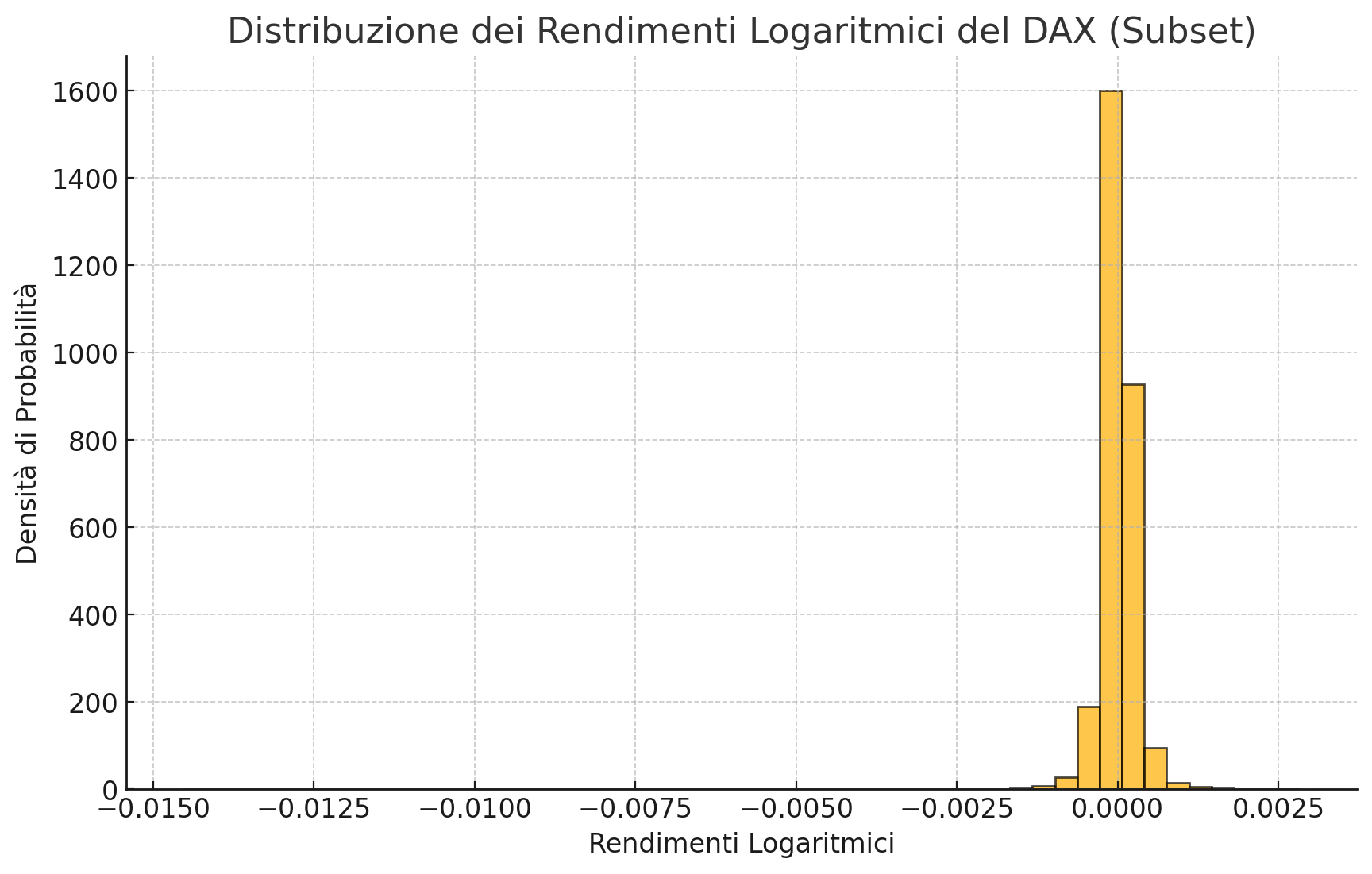
Uniformità Marginale: L'assunzione primaria nella modellazione delle copule è che le variabili marginali abbiano distribuzioni uniformi. Abbiamo utilizzato la funzione di distribuzione cumulativa per garantire questa uniformità.

Struttura di Dipendenza: I modelli di copula modellano la struttura di dipendenza tra le variabili senza fare ipotesi sui margini. Ciò significa che, dopo la trasformazione, possiamo utilizzare diversi tipi di copule per analizzare come i vari strumenti finanziari si muovono insieme.

Stazionarietà: È importante che i dati siano stazionari, ovvero che le loro proprietà statistiche (come media e varianza) siano costanti nel tempo. Abbiamo utilizzato la differenziazione logaritmica per rendere i dati stazionari.

Normalità: Per l'utilizzo di una copula Gaussiana, i margini devono approssimare la normalità. Sebbene non sia strettamente necessario per altre copule come la t-Copula, una trasformazione per avvicinarsi alla normalità può essere utile per semplificare l'analisi.

3.6 Distribuzione dei rendimenti

****

Ecco la distribuzione dei rendimenti logaritmici del DAX, calcolata su un subset dei dati disponibili. La distribuzione mostra la tipica forma a campana, con alcune code più pesanti, suggerendo la presenza di eventi estremi più frequenti rispetto a una normale distribuzione Gaussiana. Questa caratteristica supporta l'idea di utilizzare copule come la Student-t, che meglio cattura queste dipendenze nelle code.

Di seguito la matrice di correlazione tra i rendimenti logaritmici delle diverse colonne di prezzo (Open, High, Low, Close).

Immagine che contiene testo, schermata, numero, Carattere

Descrizione generata automaticamente

Considerando i dati analizzati e i risultati ottenuti dalle elaborazioni, possiamo formulare le seguenti assunzioni per l'applicazione dei modelli di copula:

**1. Uniformità Marginale**

Una delle assunzioni principali per applicare i modelli di copula è che le variabili marginali siano uniformi. Nel nostro caso, i rendimenti logaritmici calcolati per le diverse variabili (Open, High, Low, Close) sono stati trasformati in modo tale da poter essere considerati approssimativamente stazionari, ma non sono ancora stati resi uniformi. Per l'applicazione delle copule, sarà necessario trasformare i rendimenti normalizzati in una distribuzione uniforme sull'intervallo [0,1], ad esempio usando la funzione di distribuzione cumulativa empirica. Questo passaggio garantisce che le dipendenze tra variabili siano modellate correttamente senza distorsioni derivanti dalle distribuzioni marginali.

**2. Stazionarietà dei Dati**

Per applicare correttamente i modelli di copula, è necessario che le serie temporali siano stazionarie, ovvero che le proprietà statistiche come media e varianza siano costanti nel tempo. Abbiamo trasformato i prezzi in rendimenti logaritmici per ottenere una serie più stazionaria rispetto ai dati originali di prezzo. Tuttavia, è possibile che ci siano ancora componenti non stazionarie, come trend residui o stagionalità, che potrebbero influire sui risultati. La verifica e il trattamento di eventuali residui non stazionari sono fondamentali per garantire la validità dei modelli di copula.

**3. Struttura di Dipendenza**

L'analisi dei rendimenti logaritmici ha mostrato una correlazione positiva tra le variabili, sebbene con valori differenti per ciascuna coppia (ad esempio, correlazione relativamente più alta tra High e Close, e più bassa tra Open e Close). Questa osservazione implica che esiste una struttura di dipendenza tra le variabili di prezzo, che va oltre la correlazione lineare. Le copule ci permetteranno di catturare meglio questa dipendenza, soprattutto nei casi in cui le correlazioni sono condizionate da situazioni estreme (code pesanti).

**4. Dipendenze di Coda**

Osservando la distribuzione dei rendimenti logaritmici, è evidente che la distribuzione presenta code più pesanti rispetto a una normale distribuzione Gaussiana. Questo suggerisce una maggiore probabilità di eventi estremi (sia positivi che negativi), specialmente durante periodi di volatilità del mercato. Pertanto, è ragionevole assumere che le variabili presentino dipendenze di coda, rendendo modelli come la copula di Student-t o la copula Clayton adatti per catturare le correlazioni nelle code inferiori, particolarmente durante i ribassi di mercato.

**5. Non-normalità delle Distribuzioni Marginali**

I rendimenti logaritmici non seguono una distribuzione normale; piuttosto, mostrano asimmetria e code più pesanti. Sebbene la copula Gaussiana possa essere utilizzata per una prima analisi, è preferibile utilizzare copule come la Student-t per gestire deviazioni dalla normalità, particolarmente utili per modellare le code e le correlazioni durante gli eventi estremi.

**6. Asimmetria nelle Dipendenze**

La matrice di correlazione calcolata tra i rendimenti (Open, High, Low, Close) mostra differenze nei livelli di correlazione tra le diverse variabili. Questa asimmetria nelle correlazioni suggerisce che alcuni modelli di copula, come la Clayton (per le code inferiori) o la Gumbel (per le code superiori), potrebbero fornire una descrizione più accurata delle dipendenze, rispetto a modelli simmetrici come la copula Gaussiana.

**7. Condizioni di Diversificazione**

La copula di Frank potrebbe essere adatta per modellare le dipendenze tra le variabili che non mostrano comportamenti particolarmente forti nelle code (ovvero, dipendenze moderate e stabili). Tuttavia, i dati indicano la presenza di code pesanti, quindi questa copula potrebbe essere utilizzata solo come confronto con modelli che catturano meglio le dipendenze estreme.

**Conclusione sulle Assunzioni**

Uniformità e stazionarietà sono requisiti fondamentali per l'applicazione dei modelli di copula. I dati sono stati trasformati per soddisfare parzialmente queste assunzioni.

Dipendenze di coda e non-normalità suggeriscono l'uso di copule robuste come la Student-t o modelli asimmetrici come la Clayton o la Gumbel per catturare meglio le relazioni tra variabili durante condizioni di stress di mercato.

La asimmetria delle correlazioni evidenziata dalla matrice di correlazione indica la necessità di copule che possano gestire differenti tipi di dipendenze nelle code.

Queste assunzioni ci permettono di scegliere il modello di copula più appropriato per analizzare le dipendenze tra i rendimenti del DAX e comprendere meglio il comportamento del mercato in diverse condizioni economiche.

**Bibliografia**

**Embrechts, P., McNeil, A., & Straumann, D. (2002). "Correlation and Dependence in Risk Management: Properties and Pitfalls"**. In M. Dempster (Ed.), "Risk Management: Value at Risk and Beyond".  
Esplora le problematiche delle misure di correlazione tradizionali e presenta l'utilizzo delle copule come una soluzione più robusta.

**Embrechts, P., Lindskog, F., & McNeil, A. (2003). "Modelling Dependence with Copulas and Applications to Risk Management"**. In S. T. Rachev (Ed.), "Handbook of Heavy Tailed Distributions in Finance".  
Un lavoro accademico che tratta in dettaglio l'uso delle copule per la modellazione delle dipendenze, particolarmente utile per comprendere come queste possano migliorare la gestione del rischio.

**Hull, J. C. (2018). "Options, Futures, and Other Derivatives"**. Pearson.  
Fornisce una panoramica esaustiva dei derivati finanziari, con sezioni che esplorano l'uso delle copule nella modellazione di portafogli di opzioni e nell'analisi del rischio di mercato.