Politecnico di Torino

Politecnico di Torino

Corso di Laurea Magistrale

Ingegneria Matematica

A.a. 2025/2026

Inferenza bayesiana nei modelli di copule Archimedee

Una parametrizzazione basata sul coefficiente di Kendall e applicazioni ai mercati finanziari

Relatore: Gianluca Mastrantonio

Candidato: Alessio Attanasi

Sommario

Nella statistica applicata moderna una delle sfide principali è la modellizzazione della dipendenza tra variabili aleatorie. In particolare, nell'analisi finanziaria, le relazioni tra asset risultano spesso non lineari e caratterizzate da fenomeni di dipendenza nelle code. La teoria delle copule, in questo ambito, fornisce strumenti teorici che consentono di separare la struttura di dipendenza dal comportamento marginale delle singole variabili. In questo elaborato l'attenzione è rivolta all'inferenza bayesiana nei modelli di copule Archimedee. Viene presentata una proposta di riparametrizzazione del modello in termini del coefficiente di Kendall τ, in alternativa alla parametrizzazione tradizionale basata sul parametro θ del generatore. Quest'ultima, infatti, presenta diversi punti critici sia interpretativi che computazionali, legati alla diversa natura e al dominio non omogeneo del parametro tra famiglie differenti. Grazie all'utilizzo di τ come parametro di dipendenza è possibile lavorare su uno spazio standard e direttamente interpretabile, migliorando il confronto tra modelli e favorendo una maggiore regolarità delle distribuzioni a posteriori. L'inferenza è implementata tramite l'algoritmo Metropolis–Hastings, seguita da un'analisi della scelta delle distribuzioni a priori, del tuning e della diagnostica di convergenza. L'efficienza computazionale dell'approccio verrà poi confrontata con alcuni metodi frequentisti attraverso uno studio Monte Carlo. Infine l'applicazione empirica ai rendimenti giornalieri degli indici DAX e S&P 500 degli ultimi vent'anni evidenzia la presenza di dipendenza asimmetrica e una marcata intensificazione durante le fasi di crisi, confermando l'utilità della parametrizzazione in τ per l'analisi dei mercati finanziari.

Indice

Elenco delle figure

Introduzione

La modellizzazione della dipendenza tra variabili aleatorie rappresenta una delle sfide centrali della statistica moderna, con implicazioni che spaziano dalla finanza quantitativa alla gestione del rischio, dalla teoria attuariale all'analisi di fenomeni complessi multivariati.

Motivazione e contesto

Tradizionalmente l'analisi della dipendenza si è basata sul coefficiente di correlazione lineare di Pearson, una misura semplice e intuitiva che tuttavia presenta limitazioni significative; tale misura infatti cattura esclusivamente relazioni lineari, è sensibile a valori estremi e non è invariante rispetto a trasformazioni monotone delle variabili [1]. L'inadeguatezza della correlazione lineare emerge in modo particolarmente evidente nell'analisi dei fenomeni finanziari. Durante la crisi finanziaria globale del 2008, ad esempio, si è osservato un fenomeno di contagio tra mercati precedentemente considerati debolmente correlati: asset che in condizioni normali mostravano una dipendenza moderata hanno registrato perdite simultanee di grande entità [2, 3]. Questo comportamento non è adeguatamente catturato da modelli basati sulla sola correlazione lineare e ha evidenziato la necessità di strumenti statistici più flessibili, capaci di descrivere forme di dipendenza non lineare e, soprattutto, di caratterizzare il comportamento congiunto delle variabili in presenza di eventi estremi. La teoria delle copule, formalizzata da Sklar [4] e sviluppata in seguito a partire dagli anni Novanta [5], offre un quadro concettuale rigoroso per affrontare questo problema. Il teorema di Sklar stabilisce che ogni distribuzione multivariata può essere scomposta nella struttura di dipendenza, descritta da una copula, e nelle distribuzioni marginali delle singole variabili. Questa separazione consente di modellare in modo indipendente il comportamento univariato di ciascuna variabile e la loro struttura di associazione, offrendo una flessibilità senza precedenti nella costruzione di modelli multivariati [1]. All'interno della vasta famiglia di copule disponibili, le copule Archimedee occupano un ruolo di particolare rilievo grazie alla loro semplicità algebrica e alla capacità di rappresentare un'ampia varietà di strutture di dipendenza mediante un'unica funzione generatrice unidimensionale

[6]. Famiglie come quella di Clayton, Gumbel e Frank sono diventate strumenti standard nell'analisi finanziaria, poiché permettono di modellare in modo naturale fenomeni di dipendenza asimmetrica e di catturare la tendenza di alcune variabili a muoversi insieme in modo più intenso durante eventi estremi, un fenomeno noto come *tail dependence* [7].

Il problema dell'inferenza

Nonostante la loro eleganza teorica e la loro diffusione applicativa, l'inferenza statistica per le copule Archimedee presenta diverse sfide metodologiche e computazionali. Gli approcci frequentisti più comuni, come la massima verosimiglianza e i metodi basati sui momenti [8, 9], sono ben consolidati ma presentano alcune limitazioni. La stima di massima verosimiglianza, sebbene asintoticamente efficiente, può risultare numericamente instabile in presenza di superfici di verosimiglianza irregolari, multimodali o caratterizzate da regioni piatte [10]. Inoltre, l'inferenza basata su approssimazioni asintotiche può produrre intervalli di confidenza poco accurati quando la dimensione campionaria è limitata, come spesso accade in applicazioni che coinvolgono eventi rari o serie storiche brevi. L'inferenza bayesiana offre un'alternativa naturale, permettendo di incorporare informazione a priori, di quantificare l'incertezza in modo coerente attraverso la distribuzione a posteriori e di propagare tale incertezza a qualsiasi quantità derivata di interesse, come i coefficienti di tail dependence o misure di rischio congiunto [11]. Tuttavia, l'implementazione pratica dell'inferenza bayesiana per copule Archimedee non è priva di difficoltà. La parametrizzazione standard, basata sul parametro naturale θ del generatore della copula, presenta infatti diverse criticità: il dominio e l'interpretazione del parametro varia da una famiglia all'altra, rendendo difficile la specificazione di distribuzioni a priori comparabili e il confronto tra modelli alternativi [12]; inoltre, la funzione di verosimiglianza risulta spesso fortemente asimmetrica, ostacolando la convergenza degli algoritmi Monte Carlo basati su catene di Markov (MCMC).

Contributo della tesi

Questo lavoro affronta il problema dell'inferenza bayesiana per copule Archimedee adottando una parametrizzazione alternativa basata sul coefficiente di Kendall τ, una misura di concordanza robusta, invariante rispetto a trasformazioni monotone e definita su un dominio standardizzato, ovvero $\tau \in (-1,1)$, indipendentemente dalla famiglia di copule considerata [13]. L'idea di utilizzare τ come parametro di dipendenza non è nuova nella letteratura sulle copule: Grazian e Liseo [14] hanno proposto un approccio semiparametrico per l'inferenza bayesiana direttamente su τ come funzionale della copula, senza assumerne una forma parametrica. In

questo lavoro, invece, si considera un modello di copula parametrico e si adotta una riparametrizzazione del parametro θ in termini del coefficiente di Kendall τ, sfruttando la relazione biunivoca esistente per molte famiglie Archimedee. Questa scelta metodologica comporta diversi vantaggi. Dal punto di vista interpretativo, τ fornisce una misura immediata del grado di dipendenza, confrontabile tra famiglie diverse e invariante rispetto a trasformazioni delle marginali. Dal punto di vista computazionale, lavorare su un parametro con dominio limitato può favorire una migliore regolarità, facilitando l'esplorazione dello spazio parametrico mediante algoritmi MCMC e riducendo la sensibilità alle condizioni iniziali [15]. Infine, la parametrizzazione in τ permette una propagazione naturale dell'incertezza verso quantità derivate di interesse, come i coefficienti di tail dependence λ_L e λ_U, che descrivono la dipendenza asintotica nelle code della distribuzione congiunta.

Struttura della tesi

La tesi si articola su tre livelli. Il Capitolo 1 introduce i fondamenti della teoria delle copule e delle misure di dipendenza. Dopo aver discusso i limiti del coefficiente di correlazione lineare, viene presentato il teorema di Sklar e la costruzione generale delle copule. Particolare attenzione è dedicata alle copule Archimedee, di cui vengono analizzate le principali famiglie (Clayton, Gumbel, Frank) e le proprietà caratteristiche, con focus sul comportamento nelle code della distribuzione. Vengono inoltre introdotte le misure di dipendenza basate sui ranghi, il tau di Kendall e il rho di Spearman, evidenziandone le proprietà di invarianza e la relazione con i parametri delle copule Archimedee. Il Capitolo 2 costituisce il nucleo metodologico della tesi. Dopo aver motivato l'adozione dell'approccio bayesiano, viene introdotta la parametrizzazione basata sul coefficiente di Kendall τ e ne vengono discussi i vantaggi rispetto alla parametrizzazione standard. Viene quindi presentato in dettaglio l'algoritmo Metropolis-Hastings adattato a questa parametrizzazione, con particolare attenzione alla scelta delle distribuzioni a priori, al *tuning* dell'algoritmo e alla diagnostica di convergenza. La validazione dell'approccio è condotta attraverso uno studio Monte Carlo che ne caratterizza le proprietà statistiche e l'efficienza computazionale. Il capitolo si conclude con un confronto tra l'approccio bayesiano e i principali metodi frequentisti. Il Capitolo 3 presenta l'analisi empirica della dipendenza tra i rendimenti del DAX e dell'S&P 500. Dopo una descrizione dettagliata del dataset e un'analisi esplorativa preliminare, viene condotta l'inferenza bayesiana per le copule di Clayton, Gumbel e Frank. I risultati vengono confrontati mediante criteri informativi, come il DIC e WAIC, e interpretati alla luce delle caratteristiche dei mercati finanziari. Infine, un'analisi per sotto-periodi permette di evidenziare la dinamica temporale della dipendenza e la sua intensificazione durante le principali crisi finanziarie del periodo considerato.

Capitolo 1

Teoria delle copule e misure di dipendenza

Lo studio della dipendenza tra variabili aleatorie è un tema centrale in moltissimi contesti applicativi, dalla gestione del rischio all'analisi multivariata. Comprendere come diversi fenomeni aleatori si influenzino reciprocamente è infatti essenziale quando l'obiettivo è descrivere il comportamento congiunto di due grandezze, o quando si vuole calcolare la probabilità di eventi estremi che si manifestano simultaneamente. A tal fine, occorre disporre di strumenti appropriati per catturare non soltanto l'intensità, ma anche la tipologia delle relazioni tra le variabili.

1.1 Il coefficiente di correlazione lineare

Tradizionalmente la dipendenza tra due variabili aleatorie X e Y viene misurata attraverso il *coefficiente di correlazione lineare* di Pearson, definito come:

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \tag{1.1}$$

dove μ_X e μ_Y rappresentano i valori attesi delle variabili X e Y, mentre σ_X e σ_Y le rispettive deviazioni standard. Ricordiamo che la covarianza $\mathrm{Cov}(X,Y)$ misura la variazione congiunta delle due variabili rispetto ai loro valori medi, mentre la normalizzazione tramite le deviazioni standard rende il coefficiente adimensionale e confrontabile tra diverse coppie di variabili. Il coefficiente $\rho_{X,Y}$ assume valori nell'intervallo $[-1,1]$ dove valori prossimi a 1 indicano una forte correlazione lineare positiva mentre valori prossimi a -1 una forte correlazione lineare negativa.

1.1.1 Limiti della correlazione lineare

Nonostante si tratti di una misura semplice da calcolare e ampiamente diffusa nell'analisi statistica, il coefficiente di correlazione di Pearson presenta diverse limitazioni che ne riducono l'efficacia in molti contesti applicativi, specialmente quando si analizzano fenomeni con strutture di dipendenza complesse. Ad esempio è importante sottolineare che una correlazione nulla non implica necessariamente indipendenza statistica, poiché possono sussistere relazioni di natura non lineare che il coefficiente di Pearson non è in grado di catturare, come nel seguente esempio.

Esempio 1. *Consideriamo due variabili aleatorie*

$$X \sim \mathcal{N}(0,1) \quad e \quad Y = X^2.$$

Sebbene la distribuzione di Y dipenda dal valore di X, si può dimostrare che il coefficiente di correlazione di Pearson $\rho_{X,Y}$ è nullo. Infatti:

$$\begin{aligned} Cov(X, X^2) &= \mathbb{E}[X \cdot X^2] - \mathbb{E}[X]\mathbb{E}[X^2] \\ &= \mathbb{E}[X^3] - 0 \cdot 1 = 0, \end{aligned}$$

dove

$$\mathbb{E}[X^3] = \int_{-\infty}^{\infty} x^3 f_X(x)\, dx = 0,$$

poiché x^3 è dispari e $f_X(x)$, densità di una normale standard, è simmetrica rispetto all'origine. Ne segue che $\rho_{X,Y} = 0$. Tuttavia, X e Y non sono indipendenti, ad esempio:

$$\mathbb{P}(Y \leq 1 \mid X = 0) = 1, \qquad ma \qquad \mathbb{P}(Y \leq 1) < 1.$$

Questo esempio dimostra dunque che assenza di correlazione lineare non implica indipendenza. Inoltre possiamo osservare che il valore di ρ dipende non solo dalla struttura di dipendenza, ma anche dalle distribuzioni marginali di X e Y. Ciò rende difficile separare l'effetto della dipendenza da quello delle marginali. In generale poi la correlazione lineare non è invariante rispetto a trasformazioni monotone non lineari T, ovvero $\rho_{T(X),T(Y)} \neq \rho_{X,Y}$. Una relazione perfettamente deterministica può quindi avere coefficienti di correlazione diversi a seconda della scala di misurazione scelta; questa proprietà limita la capacità del coefficiente di catturare relazioni stabili quando si considerano trasformazioni delle variabili originali [16], come illustrato nel seguente esempio.

Esempio 2. *Sia* $X \sim \mathcal{N}(0,1)$ *e sia* $Y = 2X + 1$. *In questo caso* Y *è una funzione lineare crescente di* X *e quindi* $\rho_{X,Y} = 1$. *Consideriamo ora la trasformazione monotona crescente* $T(x) = e^x$ *e definiamo*

$$\tilde{X} = T(X) = e^X, \qquad \tilde{Y} = T(Y) = e^{2X+1}.$$

La relazione tra $\tilde{X}$ *e* $\tilde{Y}$ *è ancora deterministica, infatti* $\tilde{Y} = e\,\tilde{X}^2$, *ma non è più lineare. Di conseguenza, il coefficiente di correlazione lineare tra* $\tilde{X}$ *e* $\tilde{Y}$ *non è uguale a uno*

$$\rho_{\tilde{X},\tilde{Y}} \neq 1$$

Inoltre, per distribuzioni con code pesanti o asimmetriche, tipiche nei dati finanziari [7], la correlazione lineare può fornire una rappresentazione inadeguata della struttura di dipendenza, in particolare nelle code della distribuzione. Infine il coefficiente di Pearson nella sua definizione richiede l'esistenza di momenti secondi finiti, condizione non sempre soddisfatta in contesti applicativi, come ad esempio nelle distribuzioni di Cauchy o in altre distribuzioni a code pesanti [1]. Queste limitazioni hanno motivato lo sviluppo di approcci alternativi per la modellizzazione della dipendenza, culminando nella teoria delle copule, che permette di separare la struttura di dipendenza dalle distribuzioni marginali.

1.2 Teorema di Sklar e costruzione di copule

Alla base della teoria delle copule vi è il teorema di Sklar [4], che fornisce una rappresentazione generale della funzione di distribuzione congiunta di un vettore aleatorio in termini delle sue distribuzioni marginali e di una funzione di copula [1].

1.2.1 Definizione di copula

Una funzione $C : [0,1]^d \to [0,1]$ è detta ***copula d-dimensionale*** se soddisfa le seguenti proprietà:

- **Condizioni al bordo.** Il valore della copula si annulla ogni volta che almeno una delle sue componenti è uguale a zero, mentre coincide con la coordinata stessa quando tutte le altre sono pari a uno. Più formalmente, per ogni $\mathbf{u} = (u_1, \dots, u_d) \in [0,1]^d$,

$$C(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_d) = 0, \quad \forall i \in \{1, \dots, d\} \tag{1.2}$$

$$C(1, \dots, 1, u_i, 1, \dots, 1) = u_i, \quad \forall i \in \{1, \dots, d\} \tag{1.3}$$

- **D-crescenza.** Per ogni iperrettangolo

$$[\mathbf{a}, \mathbf{b}] = \prod_{i=1}^{d} [a_i, b_i] \subset [0,1]^d, \qquad \text{con } \mathbf{a} \leq \mathbf{b},$$

il C-*volume* dell'iperrettangolo deve essere non negativo:

$$V_C([\mathbf{a}, \mathbf{b}]) = \sum_{\mathbf{c} \in \{\mathbf{a}, \mathbf{b}\}} (-1)^{N(\mathbf{c})} C(\mathbf{c}) \geq 0,$$

Nella seconda condizione $\{\mathbf{a}, \mathbf{b}\}$ è l'insieme contenente tutti i vertici dell'iperrettangolo ottenuti combinando le coordinate di $\mathbf{a}$ e $\mathbf{b}$, mentre $N(\mathbf{c})$ è il numero di componenti di $\mathbf{c}$ uguali ad a_i. Questo ci assicura che C assegni probabilità non negative a ogni iperrettangolo di $[0,1]^d$, analogamente a quanto accade per una funzione di ripartizione congiunta. Nel caso in cui $d = 2$, l'iperrettangolo diventa

$$[a_1, b_1] \times [a_2, b_2], \qquad a_1 \leq b_1,\ a_2 \leq b_2,$$

cioè semplicemente un rettangolo e la condizione di 2-crescenza si riduce a:

$$C(b_1, b_2) - C(b_1, a_2) - C(a_1, b_2) + C(a_1, a_2) \geq 0. \tag{1.4}$$

ovvero la probabilità assegnata da C a tale rettangolo deve essere non negativa.

1.2.2 Il teorema di Sklar

Il teorema di Sklar non rappresenta soltanto un risultato teorico di grande eleganza, ma costituisce anche il fondamento operativo dell'intera teoria delle copule [1]. La sua importanza risiede soprattutto nelle implicazioni in ambito applicativo, dove garantisce una netta separazione tra il comportamento marginale delle variabili e la loro struttura di dipendenza. La copula C descrive la dipendenza tra le componenti del vettore aleatorio, mentre le distribuzioni marginali $F_1, \dots, F_d$ catturano il comportamento univariato di ciascuna variabile. Questa scomposizione permette di modellare in modo flessibile fenomeni complessi, combinando marginali anche molto diverse tra loro con una struttura di dipendenza scelta in modo appropriato.

Teorema 1 (Teorema di Sklar)**.** *Sia F una funzione di distribuzione congiunta d-dimensionale con distribuzioni marginali $F_1, \dots, F_d$. Allora esiste una copula $C : [0,1]^d \to [0,1]$ tale che per ogni $(x_1, \dots, x_d) \in \overline{\mathbb{R}}^d = [-\infty, +\infty]^d$:*

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \tag{1.5}$$

Se le marginali F_i sono continue, allora C è unica. Viceversa, se C è una copula e $F_1, \dots, F_d$ sono funzioni di distribuzione univariate, allora la funzione F definita dalla (1.5) *è una funzione di distribuzione congiunta con marginali $F_1, \dots, F_d$.*

Dimostrazione. Ci limitiamo al caso bivariato ($d = 2$) assumendo che le marginali F_1 e F_2 siano continue.

- **Esistenza:** Sia F una funzione di distribuzione congiunta con marginali continue F_1 e F_2. Poiché le marginali sono continue, esistono le funzioni quantili, cioè le inverse generalizzate

$$F_i^{-1}(t) = \inf\{x \in \mathbb{R} : F_i(x) \geq t\}, \qquad t \in [0,1], \quad i = 1,2. \tag{1.6}$$

  Definiamo quindi una funzione $C : [0,1]^2 \to [0,1]$ ponendo

$$C(u, v) = F\Big(F_1^{-1}(u),\, F_2^{-1}(v)\Big), \qquad (u, v) \in [0,1]^2. \tag{1.7}$$

  Verifichiamo che C sia una copula.

  a) **Condizioni al bordo:** Per ogni funzione di distribuzione congiunta F valgono le proprietà limite

$$\lim_{x \to -\infty} F(x, y) = 0, \qquad \lim_{y \to -\infty} F(x, y) = 0,$$

$$\lim_{x \to +\infty} F(x, y) = F_2(y), \qquad \lim_{y \to +\infty} F(x, y) = F_1(x).$$

Poiché le marginali sono continue, $\forall v \in [0,1]$ si ha $F_i(F_i^{-1}(v)) = v$, $i = 1,2$. Pertanto:

$$\begin{aligned}
C(0,v) &= \lim_{u\to 0^+} F\Big(F_1^{-1}(u), F_2^{-1}(v)\Big) = 0,\\
C(u,0) &= \lim_{v\to 0^+} F\Big(F_1^{-1}(u), F_2^{-1}(v)\Big) = 0,\\
C(1,v) &= \lim_{u\to 1^-} F\Big(F_1^{-1}(u), F_2^{-1}(v)\Big) = F_2\Big(F_2^{-1}(v)\Big) = v,\\
C(u,1) &= \lim_{v\to 1^-} F\Big(F_1^{-1}(u), F_2^{-1}(v)\Big) = F_1\Big(F_1^{-1}(u)\Big) = u.
\end{aligned}$$

b) **2-crescenza:** Consideriamo un rettangolo $[u_1,u_2]\times[v_1,v_2] \subset [0,1]^2$ con $u_1 \le u_2$ e $v_1 \le v_2$, e poniamo

$$x_1 = F_1^{-1}(u_1), \quad x_2 = F_1^{-1}(u_2), \quad y_1 = F_2^{-1}(v_1), \quad y_2 = F_2^{-1}(v_2).$$

Per la monotonia delle funzioni quantili si ha $x_1 \le x_2$ e $y_1 \le y_2$ e quindi

$$\begin{aligned}
&C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1)\\
&= F(x_2,y_2) - F(x_2,y_1) - F(x_1,y_2) + F(x_1,y_1) \ge 0,
\end{aligned}$$

dove l'ultima disuguaglianza segue dalla 2-crescenza della funzione F.

Pertanto C è una copula. Inoltre, per ogni $(x,y) \in \mathbb{R}^2$, ponendo $u = F_1(x)$ e $v = F_2(y)$ e usando la continuità di F_1, F_2, si ha

$$C(F_1(x),F_2(y)) = C(u,v) = F\Big(F_1^{-1}(u), F_2^{-1}(v)\Big) = F(x,y).$$

- **Unicità:** Supponiamo esistano due copule C_1 e C_2 tali che

$$F(x,y) = C_1\Big(F_1(x),F_2(y)\Big) = C_2\Big(F_1(x),F_2(y)\Big), \qquad \forall (x,y) \in \mathbb{R}^2.$$

Poiché le marginali F_1 e F_2 sono continue, esse sono suriettive su (0,1). Dunque, per ogni $(u,v) \in (0,1)^2$ esistono $x, y \in \mathbb{R}$ tali che $u = F_1(x)$ e $v = F_2(y)$, e

$$C_1(u,v) = F(x,y) = C_2(u,v), \qquad \forall (u,v) \in (0,1)^2.$$

Ricordiamo infine che le copule sono funzioni di distribuzione su $[0,1]^2$, dunque sono uniformemente continue; l'uguaglianza si estende quindi a tutto $[0,1]^2$, garantendo l'unicità.

□

Il teorema di Sklar rende dunque possibile la costruzione di distribuzioni multivariate a partire da specifiche marginali e dalla loro dipendenza [1]. Data una copula C e un insieme arbitrario di distribuzioni marginali $F_1, \ldots, F_d$, è sempre possibile definire una distribuzione congiunta coerente con tali scelte. Questa proprietà è alla base di numerose applicazioni in finanza e in ambito assicurativo, ad esempio nella modellizzazione di portafogli, nella valutazione di rischi congiunti o nell'analisi di eventi estremi [7]. Inoltre un secondo aspetto cruciale riguarda l'invarianza della copula rispetto a trasformazioni monotone crescenti. Se T_1 e T_2 sono funzioni monotone crescenti, allora le coppie di variabili (X, Y) e $(T_1(X), T_2(Y))$ condividono la stessa copula. In termini formali [1]:

$$C_{T_1(X),T_2(Y)}(u, v) = C_{X,Y}(u, v), \qquad \forall (u, v) \in [0,1]^2 \tag{1.8}$$

Questo implica che la copula cattura esclusivamente la struttura di dipendenza tra le variabili, indipendentemente dalla scala o dalla trasformazione applicata alle marginali, a differenza del coefficiente di correlazione di Pearson, che è sensibile a trasformazioni non lineari.

Funzione di densità della copula

Nel caso in cui la copula C sia assolutamente continua, essa ammette una densità, detta densità di copula, definita come

$$c(u_1, \ldots, u_d) = \frac{\partial^d C(u_1, \ldots, u_d)}{\partial u_1 \cdots \partial u_d} \tag{1.9}$$

In questo contesto, la densità congiunta f del vettore aleatorio $(X_1, \ldots, X_d)$ può essere scritta nella forma

$$f(x_1, \ldots, x_d) = c(F_1(x_1), \ldots, F_d(x_d)) \cdot \prod_{i=1}^{d} f_i(x_i) \tag{1.10}$$

dove f_i denota la densità marginale di X_i. Questa rappresentazione rende esplicita, anche a livello di densità, la separazione introdotta dal teorema di Sklar: la funzione c incorpora esclusivamente la struttura di dipendenza, mentre le densità marginali f_i descrivono il comportamento individuale delle singole variabili [1].

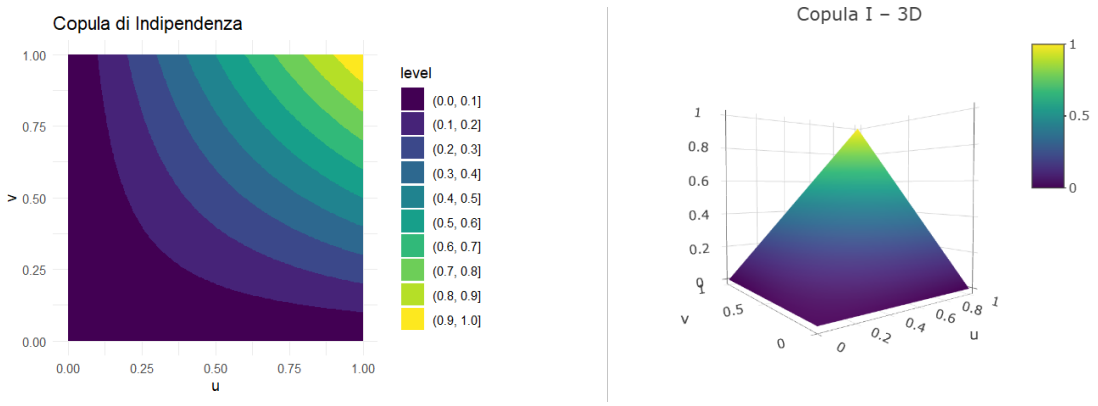


Figura 1.1: Rappresentazioni $2D$ e $3D$ della copula di indipendenza $\Pi(u, v) = uv$.

1.3 Una panoramica delle copule più comuni

Esiste un'ampia varietà di copule, ciascuna con caratteristiche specifiche che la rendono adatta a modellare particolari forme di dipendenza. In questa sezione presentiamo brevemente le famiglie più utilizzate in letteratura, prima di dedicarci alle copule Archimedee nel dettaglio. Le definizioni, le formule e i principali risultati teorici presentati in questa e nelle prossima sezioni seguono la trattazione classica proposta in Nelsen [1]. Le rappresentazioni grafiche delle copule e delle rispettive densità sono invece ottenute utilizzando il codice illustrato in Appendice A.

1.3.1 Copule di indipendenza e comonotonia

Due casi estremi, che rappresentano situazioni limite nella dipendenza tra variabili, sono la copula di indipendenza e le copule di Fréchet-Hoeffding. La **copula di indipendenza**, o copula prodotto, è definita da:

$$\Pi(u_1, \ldots, u_d) = \prod_{i=1}^{d} u_i \tag{1.11}$$

Questa copula descrive la situazione in cui le variabili aleatorie non presentano alcuna dipendenza. Per ottenere la (1.11) supponiamo $X_1, \ldots, X_d$ indipendenti; la loro funzione di distribuzione congiunta soddisferà

$$F(x_1, \ldots, x_d) = \prod_{i=1}^{d} F_i(x_i)$$

e, per il teorema di Sklar, esisterà una copula $\Pi(u_1, \ldots, u_d)$ tale che

$$F(x_1, \ldots, x_d) = \Pi(F_1(x_1), \ldots, F_d(x_d))$$

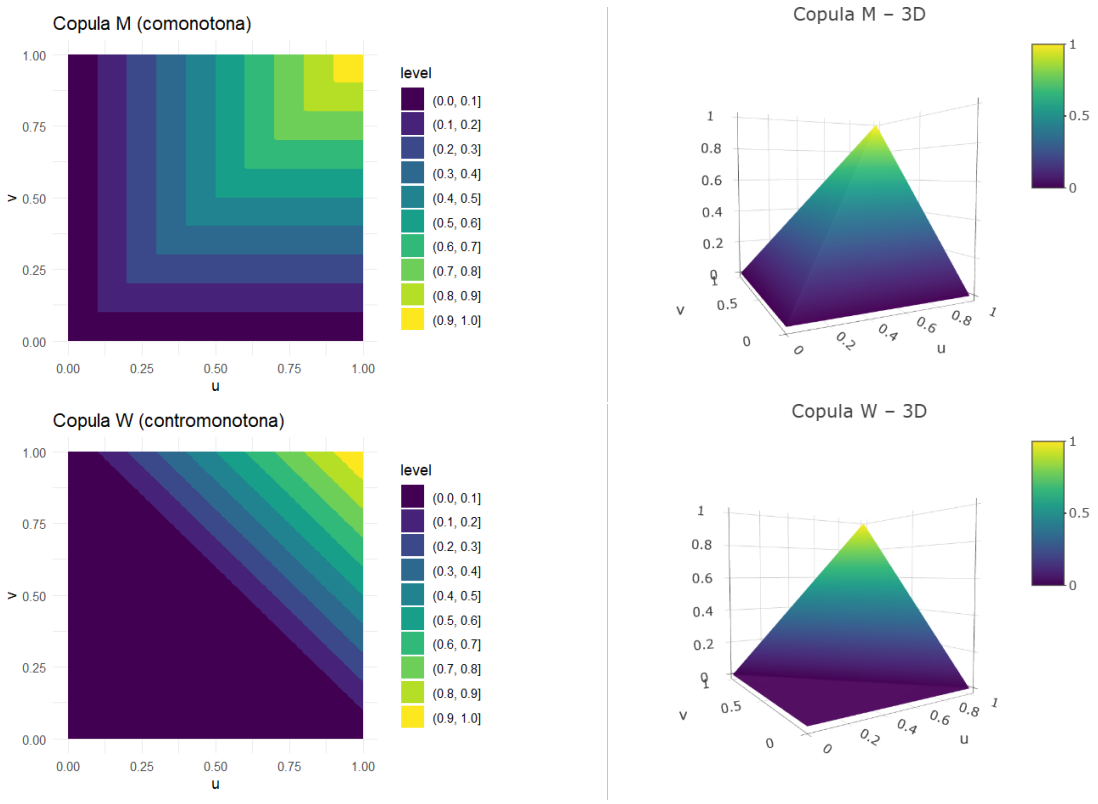


Figura 1.2: Rappresentazioni $2D$ e $3D$ della copula $M(u, v)$ e $W(u, v)$.

Invece le **copule di Fréchet-Hoeffding** superiore e inferiore identificano la massima dipendenza positiva o negativa possibile e sono definite come:

$$\begin{aligned} M(u_1, \ldots, u_d) &= \min(u_1, \ldots, u_d) \\ W(u_1, \ldots, u_d) &= \max\left(\sum_{i=1}^{d} u_i - d + 1,\ 0\right) \end{aligned} \tag{1.12}$$

La copula M rappresenta la dipendenza positiva perfetta, ovvero le variabili si muovono tutte nella stessa direzione. Invece W rappresenta la perfetta dipendenza negativa, ma è formalmente una copula solo nel caso bivariato; per $d \geq 3$ infatti la funzione W non è più d-crescente. Oltre a rappresentare situazioni estreme di dipendenza, le copule di Fréchet-Hoeffding individuano anche i limiti teorici inferiori e superiori per qualsiasi copula, come affermato dal seguente teorema.

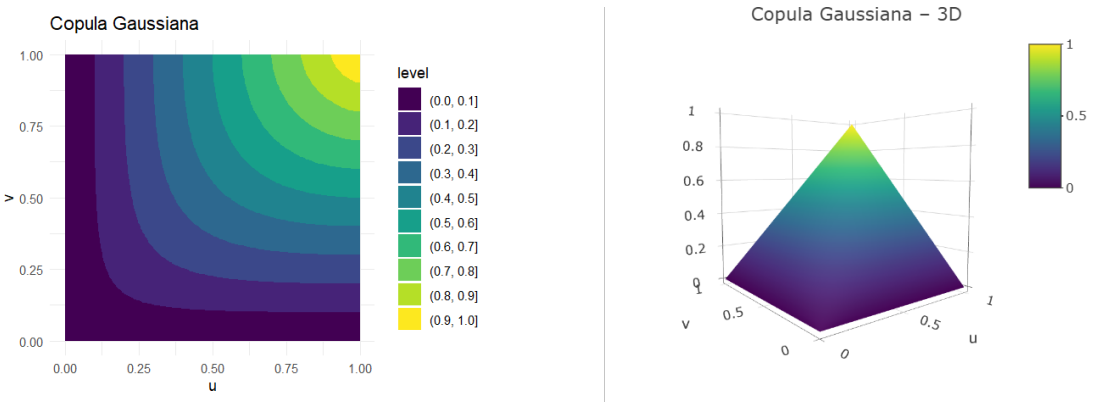


Figura 1.3: Rappresentazioni $2D$ e $3D$ della copula Gaussiana $C_{\mathbf{R}}^{\text{Gauss}}(u, v)$.

Teorema 2. *Per ogni* $(u_1, \ldots, u_d) \in [0,1]^d$ *e per ogni copula* C *vale:*

$$W(u_1, \ldots, u_d) \leq C(u_1, \ldots, u_d) \leq M(u_1, \ldots, u_d) \tag{1.13}$$

dove il limite inferiore è una copula solo nel caso $d = 2$.

Questi vincoli, noti come **limiti di Fréchet-Hoeffding**, delimitano lo spazio di tutte le possibili copule, e quindi di tutte le strutture di dipendenza ammissibili.

1.3.2 Copula Gaussiana

La copula Gaussiana è una delle più utilizzate grazie alla sua semplicità e alla connessione con la distribuzione normale multivariata. Pur essendo molto flessibile, presenta tuttavia una limitazione importante: non presenta dipendenza asintotica nelle code, fenomeno spesso presente nei dati finanziari e assicurativi.

Definizione 1. Sia $\Phi_{\mathbf{R}}$ la funzione di distribuzione di un vettore normale multivariato standard con matrice di correlazione $\mathbf{R}$, e sia Φ la funzione di distribuzione normale standard univariata. La **copula Gaussiana** è definita come:

$$C_{\mathbf{R}}^{\text{Gauss}}(u_1, \ldots, u_d) = \Phi_{\mathbf{R}}(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)) \tag{1.14}$$

La copula Gaussiana presenta simmetria radiale e una struttura di dipendenza simmetrica; tuttavia, essendo i coefficienti di dipendenza nelle code pari a zero, le variabili congiuntamente gaussiane non mostrano tendenza a generare eventi estremi simultanei.

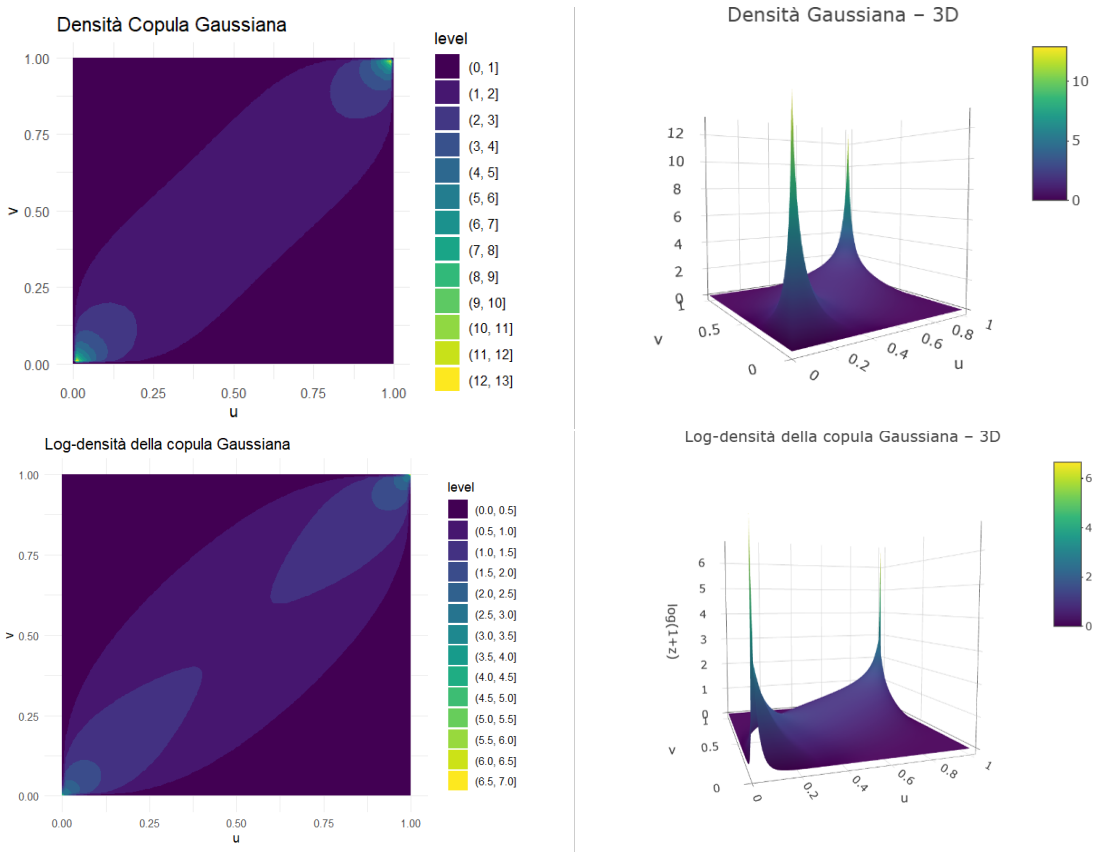


Figura 1.4: Rappresentazioni $2D$ e $3D$ della densità della copula Gaussiana.

Nel caso delle copule ellittiche, la densità di copula può essere espressa in forma compatta come rapporto tra la densità congiunta del vettore sottostante e il prodotto delle densità marginali. In particolare, la densità della copula Gaussiana associata alla matrice di correlazione $\mathbf{R}$ è

$$c_{\mathbf{R}}^{\text{Gauss}}(u_1, \ldots, u_d) = \frac{\varphi_{\mathbf{R}}\Big(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)\Big)}{\prod_{i=1}^{d} \varphi\Big(\Phi^{-1}(u_i)\Big)}$$

dove $\varphi_{\mathbf{R}}$ è la densità normale multivariata standard con correlazione $\mathbf{R}$ e φ è la densità normale standard univariata. Nel caso bivariato, con parametro di correlazione ρ, ponendo $x = \Phi^{-1}(u)$ e $y = \Phi^{-1}(v)$ si ottiene la nota espressione esplicita per la densità della copula Gaussiana:

$$c_{\rho}^{\text{Gauss}}(u, v) = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{\rho^2(x^2+y^2) - 2\rho xy}{2(1-\rho^2)}\right\} \tag{1.15}$$

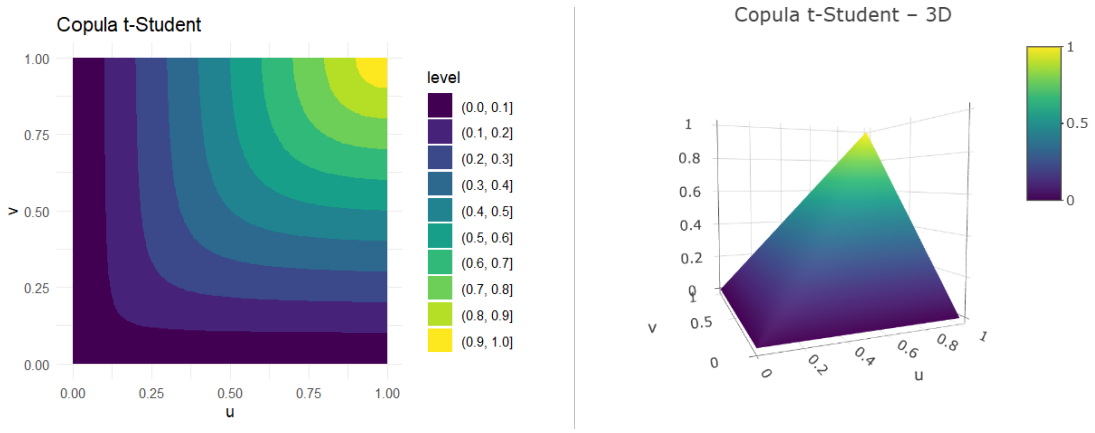


Figura 1.5: Rappresentazioni $2D$ e $3D$ della copula t di Student $C^t_{\nu,\mathbf{R}}(u,v)$.

1.3.3 Copula t di Student

La copula t di Student generalizza la copula Gaussiana ed è particolarmente interessante perché permette di modellare in modo naturale il fenomeno della dipendenza nelle code.

Definizione 2. Sia $t_{\nu,\mathbf{R}}$ la funzione di distribuzione di una t di Student multivariata con ν gradi di libertà e matrice di correlazione $\mathbf{R}$, e sia t_ν la funzione di distribuzione t di Student univariata con ν gradi di libertà. La **copula t di Student** è definita come:

$$C^t_{\nu,\mathbf{R}}(u_1,\dots,u_d) = t_{\nu,\mathbf{R}}(t_\nu^{-1}(u_1),\dots,t_\nu^{-1}(u_d)) \tag{1.16}$$

La copula t presenta dipendenza nelle code superiore e inferiore, con intensità crescente e code più pesanti al diminuire dei gradi di libertà ν. In particolare, i coefficienti di tail dependence sono positivi e simmetrici. Nel limite $\nu \to \infty$ la copula t converge alla copula Gaussiana. La densità della copula t di Student con ν gradi di libertà e matrice di correlazione $\mathbf{R}$ è data da

$$c^t_{\nu,\mathbf{R}}(u_1,\dots,u_d) = \frac{f_{\nu,\mathbf{R}}\Big(t_\nu^{-1}(u_1),\dots,t_\nu^{-1}(u_d)\Big)}{\prod_{i=1}^d f_\nu\Big(t_\nu^{-1}(u_i)\Big)} \tag{1.17}$$

dove $f_{\nu,\mathbf{R}}$ e f_ν denotano rispettivamente la densità t multivariata e la densità t univariata, entrambe con ν gradi di libertà.

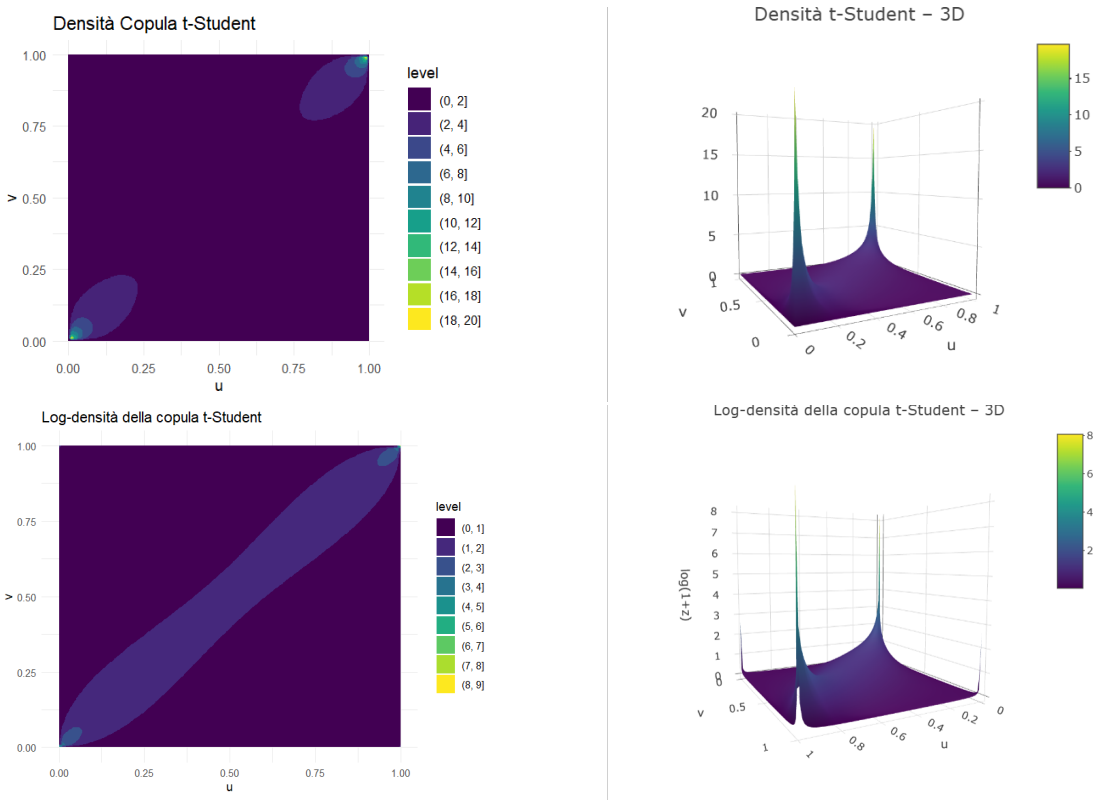


Figura 1.6: Rappresentazioni $2D$ e $3D$ della densità della copula t di Student.

Per la copula t di Student, indicando con $x = t_\nu^{-1}(u)$ e $y = t_\nu^{-1}(v)$, la densità bivariata può essere scritta in forma compatta come

$$c^t_{\nu,\rho}(u,v) = \frac{f_{\nu,\rho}(x,y)}{f_\nu(x)\, f_\nu(y)}, \qquad x = t_\nu^{-1}(u), \quad y = t_\nu^{-1}(v) \tag{1.18}$$

dove $f_{\nu,\rho}$ è la densità t bivariata con ν gradi di libertà e correlazione ρ. Le rappresentazioni tridimensionali delle copule Gaussiana in Figura 1.3 e della copula t di Student in Figura 1.5 appaiono visivamente molto simili a parità di parametro di correlazione. La differenza sostanziale tra le due famiglie emerge tuttavia dall'analisi della densità associata: la copula t di Student mostra una maggiore concentrazione di massa nelle regioni di coda, evidenziando la presenza di tail dependence, che invece è assente nel caso gaussiano. I picchi della densità osservabili in prossimità degli angoli del quadrato unitario non devono essere interpretati come evidenza di dipendenza nelle code. Essi derivano dal fatto che la trasformazione $x = \Phi^{-1}(u)$ e $y = \Phi^{-1}(v)$ diverge ai bordi dell'intervallo unitario. Tuttavia, tale comportamento riguarda esclusivamente la densità locale della copula e non il comportamento asintotico delle probabilità congiunte nelle code.

1.4 Le copule Archimedee nel dettaglio

Le copule Archimedee costituiscono una classe particolarmente importante e flessibile di copule, caratterizzate da una struttura algebrica semplice ma in grado di rappresentare un'ampia varietà di strutture di dipendenza.

1.4.1 Definizione e proprietà fondamentali

Alla base della costruzione Archimedea vi è la funzione ϕ, detta generatore, che determina completamente la struttura di dipendenza della copula. In questa trattazione consideriamo generatori strettamente Archimedei, cioè tali che $\lim_{t\to 0^+} \phi(t) = +\infty$.

Definizione 3. Una funzione $\phi : (0,1] \to [0,\infty)$ è detta ***generatore di copula*** se soddisfa le seguenti proprietà:

- ϕ è continua, convessa e strettamente decrescente su $(0,1]$
- $\phi(1) = 0$
- $\lim_{t\to 0^+} \phi(t) = +\infty$

In queste condizioni, la funzione ϕ è invertibile e ammette un'inversa ordinaria $\phi^{-1} : [0,\infty) \to (0,1)$. Ciò assicura che la copula così costruita sia ben definita e rispetti tutte le proprietà formali richieste.

Definizione 4. Una copula $C : [0,1]^2 \to [0,1]$ è detta ***copula Archimedea*** se può essere rappresentata nella forma:

$$C(u,v) = \phi^{-1}(\phi(u) + \phi(v)) \tag{1.19}$$

dove ϕ è un generatore e ϕ^{-1} denota la inversa ordinaria.

La forma (1.19) mostra chiaramente come la dipendenza tra le variabili sia governata dal modo in cui ϕ combina gli input. Tuttavia è opportuno osservare che la sola convessità di ϕ è sufficiente per garantire che C sia una copula solo in due dimensioni.

Teorema 3. *Sia $\phi : (0,1] \to [0,\infty)$ una funzione continua, strettamente decrescente e convessa con $\phi(1) = 0$. Allora la funzione C definita dalla* (1.19) *è una copula bidimensionale.*

In dimensioni maggiori la situazione è più delicata. Le copule Archimedee possono essere infatti estese al caso d-dimensionale mantenendo la stessa struttura semplice:

$$C(u_1, \ldots, u_d) = \phi^{-1}\left(\sum_{i=1}^{d} \phi(u_i)\right) \tag{1.20}$$

Tuttavia, non tutti i generatori che producono copule valide in dimensione 2 generano copule valide in dimensioni superiori. È necessaria una condizione più restrittiva sulla monotonia di ϕ.

Teorema 4. *Sia ϕ un generatore Archimedeo. La funzione*

$$C(u_1, \ldots, u_d) = \phi^{-1}\left(\sum_{i=1}^{d} \phi(u_i)\right)$$

*è una copula d-dimensionale se e solo se ϕ è d-*monotona *su* $(0,1]$*, cioè ϕ è $(d-2)$ volte derivabile e*

$$(-1)^k \phi^{(k)}(t) \geq 0 \quad \textit{per } k = 0,1,\ldots,d-2,$$

e inoltre $(-1)^{d-2}\phi^{(d-2)}$ è non crescente e convessa su $(0,1]$. Se invece ϕ è completamente monotona*, cioè C^∞ e $(-1)^k\phi^{(k)}(t) \geq 0$ per ogni $k \geq 0$, allora C è una copula Archimedea in ogni dimensione $d \geq 2$.*

Le copule Archimedee godono poi di diverse proprietà particolarmente interessanti. Osserviamo innanzitutto che, per costruzione, ogni copula Archimedea è simmetrica:

$$C(u, v) = C(v, u), \quad \forall (u, v) \in [0,1]^2 \tag{1.21}$$

Inoltre, nelle famiglie di copule più comuni, il generatore ϕ dipende da un parametro θ che controlla il livello di dipendenza. Utilizzeremo la notazione ϕ_θ per esplicitare questa dipendenza. Valori diversi di θ producono strutture di dipendenza molto diverse. Quando poi il generatore ϕ è due volte differenziabile, dunque sufficientemente regolare, è possibile calcolare la densità della copula Archimedea:

$$c(u, v) = \frac{\phi''\Big(\phi^{-1}(\phi(u) + \phi(v))\Big)\, \phi'(u)\, \phi'(v)}{\Big[\phi'\Big(\phi^{-1}(\phi(u) + \phi(v))\Big)\Big]^3} \tag{1.22}$$

Questa espressione, sebbene complessa, ci permetterà nel Capitolo 2 di valutare la verosimiglianza per l'inferenza statistica.

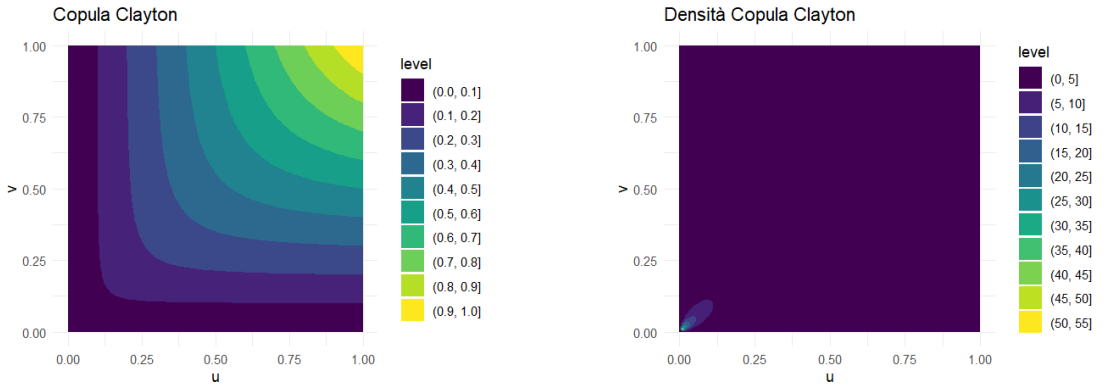


Figura 1.7: Rappresentazioni $2D$ della copula di Clayton $C_\theta^{\text{Clay}}(u,v)$ e $c_\theta^{\text{Clay}}(u,v)$.

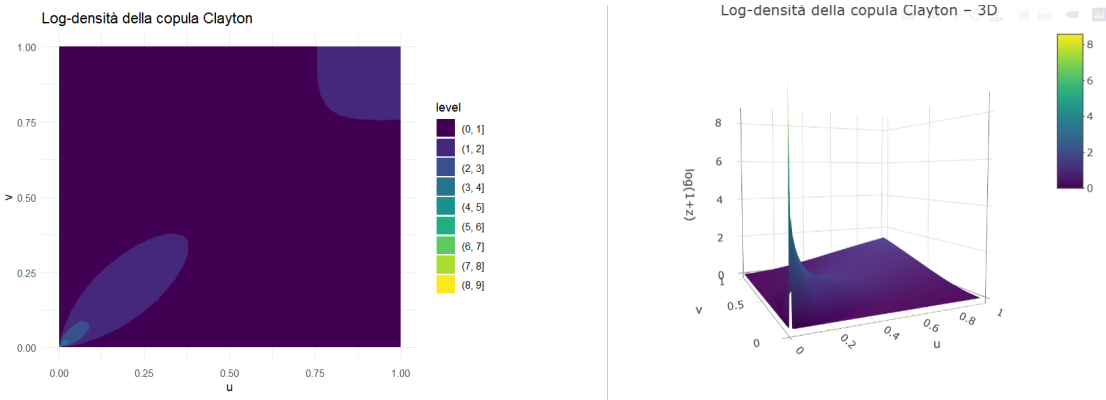


Figura 1.8: Rappresentazioni $2D$ e $3D$ della log-densità della copula di Clayton.

1.4.2 Principali famiglie di copule Archimedee

Presentiamo ora le principali famiglie di copule Archimedee utilizzate in applicazioni, specificando per ciascuna il generatore, il dominio del parametro e le proprietà caratteristiche. Per una trattazione completa delle copule Archimedee e delle loro proprietà si vedano Nelsen [1] e Joe [6].

Copula di Clayton

La copula di Clayton è nota per modellare dipendenza nella coda inferiore, cioè la tendenza delle variabili a realizzare simultaneamente valori molto bassi. Nel caso bidimensionale presenta parametro $\theta \in [-1,\infty) \setminus \{0\}$ ed è definita dal generatore

$$\phi_\theta(t) = \frac{1}{\theta}(t^{-\theta} - 1), \quad t \in (0,1] \tag{1.23}$$

Per estensioni multivariate è necessario $\theta \geq 0$ per avere il generatore d-monotono. La forma esplicita della copula è:

$$C_\theta^{\text{Clay}}(u,v) = \max\left\{(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta},\, 0\right\} \tag{1.24}$$

La forte dipendenza nella coda inferiore, rende la copula di Clayton appropriata per modellare co-movimenti durante crisi di mercato o, più in generale, perdite simultanee, come vedremo nel Capitolo 3. La densità della copula di Clayton è:

$$c_\theta^{\text{Clay}}(u,v) = (1+\theta)\,(uv)^{-\theta-1}\,(u^{-\theta} + v^{-\theta} - 1)^{-2-1/\theta} \tag{1.25}$$

È infine interessante notare come, al variare del parametro θ, otteniamo le copule viste precedentemente.

$$\begin{aligned} \theta \to 0 \quad &\Rightarrow \quad C_\theta^{\text{Clay}} \longrightarrow \Pi, \\ \theta \to \infty \quad &\Rightarrow \quad C_\theta^{\text{Clay}} \longrightarrow M, \\ \theta = -1 \quad &\Rightarrow \quad C_{-1}^{\text{Clay}}(u,v) = \max\{u+v-1,0\} = W(u,v) \end{aligned} \tag{1.26}$$

Copula di Gumbel

La copula di Gumbel è la controparte speculare della Clayton, poiché presenta una forte dipendenza nella coda superiore e indipendenza asintotica nella coda inferiore. Per questo motivo, risulta adatta a modellare eventi estremi congiunti positivi. È definita dal generatore

$$\phi_\theta(t) = (-\ln t)^\theta, \quad t \in (0,1] \tag{1.27}$$

con parametro $\theta \in [1,\infty)$. La sua forma esplicita è

$$C_\theta^{\text{Gumb}}(u,v) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right\} \tag{1.28}$$

Definendo $A := (-\ln u)^\theta + (-\ln v)^\theta$, possiamo scrivere la densità della copula come

$$c_\theta^{\text{Gumb}}(u,v) = C_\theta^{\text{Gumb}}(u,v)\,\frac{(-\ln u)^{\theta-1}(-\ln v)^{\theta-1}}{uv}\,A^{\frac{2}{\theta}-2}\left[1 + (\theta-1)A^{-1/\theta}\right] \tag{1.29}$$

Anche qui, facendo variare il parametro fino agli estremi del suo dominio, otteniamo

$$\begin{aligned} \theta = 1 \quad &\Rightarrow \quad C_\theta^{\text{Gumb}} = \Pi, \\ \theta \to \infty \quad &\Rightarrow \quad C_\theta^{\text{Gumb}} \longrightarrow M \end{aligned} \tag{1.30}$$

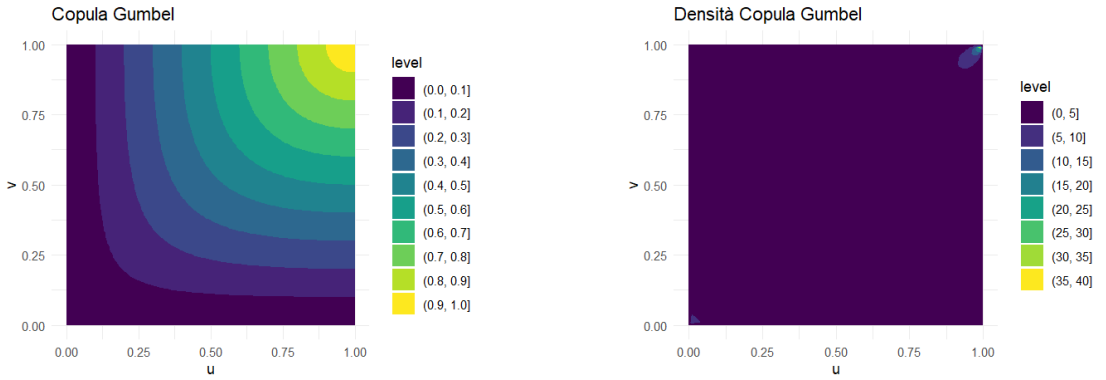


Figura 1.9: Rappresentazioni $2D$ della copula di Gumbel $C_\theta^{\text{Gumb}}(u,v)$ e $c_\theta^{\text{Gumb}}(u,v)$

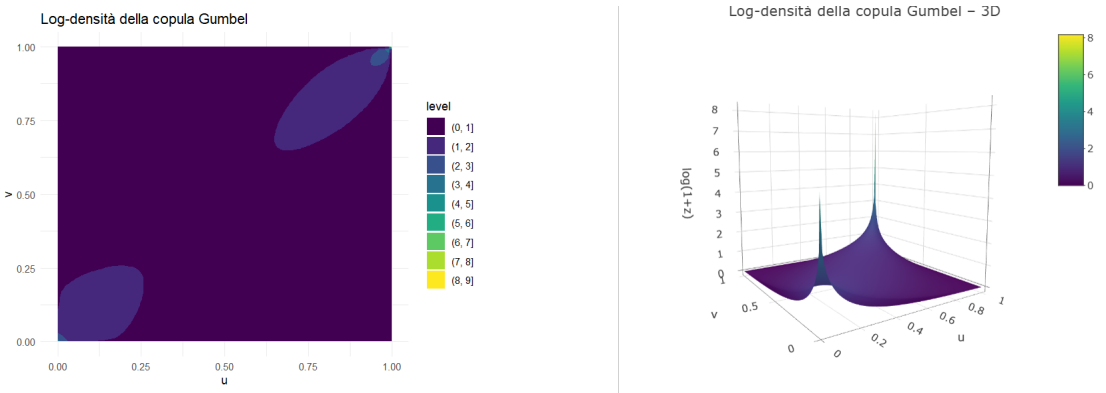


Figura 1.10: Rappresentazioni $2D$ e $3D$ della log-densità della copula di Gumbel.

Copula di Frank

La copula di Frank è molto versatile e adatta quando non si osserva dipendenza estrema nelle code. È definita dal generatore

$$\phi_\theta(t) = -\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right), \quad t \in (0,1] \tag{1.31}$$

con parametro $\theta \in \mathbb{R} \setminus \{0\}$. La sua forma esplicita è

$$C_\theta^{\text{Frank}}(u,v) = -\frac{1}{\theta}\ln\left(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right) \tag{1.32}$$

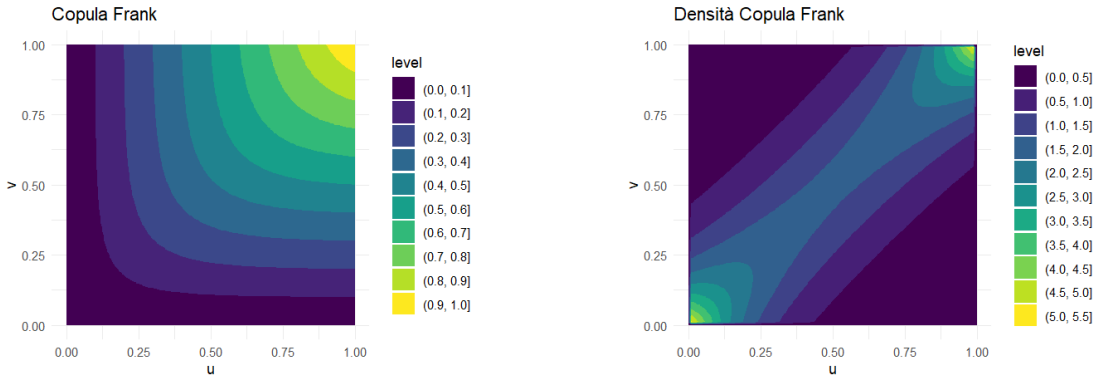


Figura 1.11: Rappresentazioni $2D$ della copula di Frank $C_\theta^{\text{Frank}}(u,v)$ e $c_\theta^{\text{Frank}}(u,v)$.

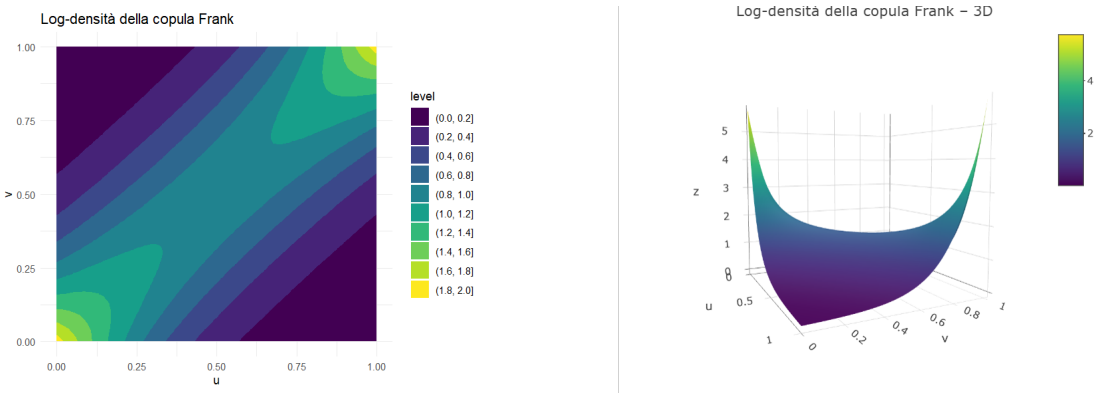


Figura 1.12: Rappresentazioni $2D$ e $3D$ della log-densità della copula di Frank.

Come per le copule precedenti, facendo variare il parametro otteniamo dei casi limite interessanti:

$$\begin{aligned} \theta \to 0 \quad &\Rightarrow \quad C_\theta^{\text{Frank}} \longrightarrow \Pi, \\ \theta \to \infty \quad &\Rightarrow \quad C_\theta^{\text{Frank}} \longrightarrow M, \\ \theta \to -\infty \quad &\Rightarrow \quad C_\theta^{\text{Frank}} \longrightarrow W \end{aligned} \tag{1.33}$$

Anche se la densità associata alla copula di Frank può essere ricavata a partire dal generatore Archimedeo mediante la formula generale per le copule Archimedee (1.22), l'espressione esplicita è complessa e non verrà riportata per esteso. Osservando la figura 1.12 possiamo notare che la copula di Frank, a differenza delle precedenti, non presenta dipendenza di coda, mostrando una struttura simmetrica rispetto alla diagonale, rendendola adatta a modellare situazioni in cui la dipendenza tra le variabili non si intensifica nelle regioni estreme della distribuzione.

1.5 Misure di dipendenza

Quando si studia la relazione tra variabili aleatorie è naturale chiedersi non solo se esiste una dipendenza, ma anche interrogarsi sulla sua intensità e natura. Come anticipato nel paragrafo 1.1, la correlazione di Pearson è la misura di dipendenza più nota, ma presenta un limite importante: non è invariante rispetto a trasformazioni monotone, e quindi non descrive bene relazioni non lineari. Le misure che introdurremo ora, il tau di Kendall e il rho di Spearman, essendo basate sulle copule, mantengono invece la loro interpretazione anche dopo trasformazioni monotone crescenti, risultando quindi più adatte a descrivere forme di dipendenza generali.

1.5.1 Il tau di Kendall

Il tau di Kendall misura la tendenza delle variabili a essere *concordi* o *discordi*. Intuitivamente, due coppie di osservazioni sono concordi se, ordinandole secondo una variabile, l'ordine è lo stesso anche per l'altra variabile.

Definizione 5. Siano (X_1, Y_1) e (X_2, Y_2) due coppie indipendenti di variabili aleatorie con la stessa distribuzione congiunta. Il ***tau di Kendall*** è definito come:

$$\tau = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0] \tag{1.34}$$

ovvero come la differenza tra la probabilità di concordanza e la probabilità di discordanza.

Il coefficiente τ assume valori nell'intervallo $[-1,1]$: il valore -1 corrisponde a dipendenza negativa perfetta, 1 a dipendenza positiva perfetta, mentre 0 indica assenza di dipendenza in termini di concordanza. Uno dei principali vantaggi delle misure di dipendenza fondate sulle copule è la possibilità di esprimerle direttamente in funzione della copula stessa, indipendentemente dalle distribuzioni marginali.

Teorema 5. *Se la coppia di variabili aleatorie (X, Y) ammette copula C, allora il tau di Kendall può essere scritto come*

$$\tau = 4 \int_0^1 \int_0^1 C(u, v)\, dC(u, v) - 1 = 4\mathbb{E}[C(U, V)] - 1 \tag{1.35}$$

dove U e V sono uniformi su $[0,1]$ con copula C e l'integrale è inteso come integrale di Lebesgue-Stieltjes rispetto alla misura indotta da C.

Dimostrazione. Dalla definizione, abbiamo:

$$\begin{aligned}\tau &= \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0] \\ &= 2\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1 \end{aligned} \tag{1.36}$$

Ma $(X_1 - X_2)(Y_1 - Y_2) > 0$ se e solo se ($X_1 < X_2$ e $Y_1 < Y_2$) o ($X_1 > X_2$ e $Y_1 > Y_2$). Quindi per simmetria:

$$\begin{aligned}\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] &= 2\mathbb{P}[X_1 < X_2, Y_1 < Y_2] \\ &= 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x, y)\, dF(x, y)\end{aligned} \tag{1.37}$$

dove F denota la funzione di distribuzione congiunta di (X, Y). Applicando la trasformazione $u = F_1(x)$, $v = F_2(y)$ e usando la rappresentazione di Sklar $F(x, y) = C(F_1(x), F_2(y))$:

$$\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] = 2\int_0^1\int_0^1 C(u, v)\, dC(u, v) \tag{1.38}$$

da cui segue la (1.35). □

Tau di Kendall per copule Archimedee

Per le copule Archimedee, esiste una formula esplicita particolarmente elegante:

Teorema 6. *Sia C una copula Archimedea con generatore ϕ strettamente Archimedeo, derivabile su* (0,1) *e tale che l'integrale seguente sia finito. Allora il tau di Kendall è dato da*

$$\tau = 1 + 4\int_0^1 \frac{\phi(t)}{\phi'(t)}\, dt.$$

Questa formula permette di calcolare τ direttamente dal generatore, senza richiedere integrazioni numeriche complesse della copula stessa. Osserviamo che per le copule Archimedee il generatore ϕ è strettamente decrescente, cioè $\phi'(t) < 0$; il rapporto $\frac{\phi(t)}{\phi'(t)}$, e di conseguenza l'integrale, saranno dunque negativi.

Formule esplicite per copule Archimedee

Per le principali famiglie di copule Archimedee, il tau di Kendall ha espressioni chiuse in funzione del parametro θ. Ciò permette di invertire la relazione per esprimere il parametro della copula in funzione di τ.

Copula	**Kendall τ**	**Inversione di θ**
Clayton	$\tau = \dfrac{\theta}{\theta+2}$	$\theta = \dfrac{2\tau}{1-\tau}$
Gumbel	$\tau = 1 - \dfrac{1}{\theta}$	$\theta = \dfrac{1}{1-\tau}$
Frank	$\tau = 1 - \dfrac{4}{\theta}[1 - D_1(\theta)]$	invertibile numericamente

Tabella 1.1: Relazione tra il parametro di copula θ e il coefficiente di Kendall τ.

Nella Tabella 1.1 $D_1(\theta) = \frac{1}{\theta}\int_0^\theta \frac{t}{e^t-1}\,dt$ è la funzione di Debye di ordine 1. Non esiste una forma chiusa per θ in funzione di τ per la copula di Frank, ma la relazione è monotona e invertibile numericamente. Nel complesso, queste relazioni mostrano come il tau di Kendall rappresenti uno strumento naturale per parametrizzare le copule Archimedee, e ci saranno d'aiuto nei capitoli successivi poiché consentono di esprimere la dipendenza in modo comparabile tra famiglie diverse.

1.5.2 Il rho di Spearman

Il rho di Spearman è una misura di concordanza che quantifica la correlazione monotona tra due variabili casuali attraverso i ranghi delle loro osservazioni.

Definizione 6. Siano X e Y due variabili casuali continue con funzioni di distribuzione marginali F_1 e F_2. Il **rho di Spearman** è definito come il coefficiente di correlazione di Pearson tra le variabili trasformate:

$$\rho_S = \rho_S(X,Y) := \mathrm{Corr}(U,V) = \mathrm{Corr}\Big(F_1(X), F_2(Y)\Big) \tag{1.39}$$

Essendo $U = F_1(X)$ e $V = F_2(Y)$ variabili uniformi su [0,1] con $\mathbb{E}[U] = \mathbb{E}[V] = \frac{1}{2}$ e $\mathrm{Var}(U) = \mathrm{Var}(V) = \frac{1}{12}$, si ottiene

$$\rho_S = \frac{\mathrm{Cov}(U,V)}{\sqrt{\mathrm{Var}(U)\mathrm{Var}(V)}} = 12\,\mathrm{Cov}(U,V) = 12\Big(\mathbb{E}[UV] - \frac{1}{4}\Big) = 12\mathbb{E}[UV] - 3 \tag{1.40}$$

Come per il tau di Kendall, anche il rho di Spearman dipende esclusivamente dalla struttura di dipendenza tra X e Y e non dalle distribuzioni marginali. Pertanto, esso ammette una rappresentazione in termini della copula C associata alla coppia (X,Y). In particolare siano (U,V) una coppia di variabili casuali uniformi su [0,1] con copula C. Allora si può dimostrare che il rho di Spearman può essere espresso come

$$\rho_S = 12\int_0^1\int_0^1 C(u,v)\,du\,dv - 3 \tag{1.41}$$

Formule esplicite per copule Archimedee

Per le copule Archimedee, ricordando che $C(u,v) = \phi^{-1}\Big(\phi(u) + \phi(v)\Big)$, si ottiene

$$\rho_S = 12 \int_0^1 \int_0^1 \phi^{-1}\Big(\phi(u) + \phi(v)\Big)\, du\, dv - 3 \tag{1.42}$$

A differenza del tau di Kendall, questa formula raramente conduce a espressioni chiuse semplici, ma rimane comunque utile dal punto di vista teorico e computazionale. Per la copula di Frank, ad esempio, il rho di Spearman può essere espresso in termini delle funzioni di Debye:

$$\rho_{S,\text{Frank}} = 1 - \frac{12}{\theta}\Big[D_2(\theta) - D_1(\theta)\Big] \tag{1.43}$$

dove D_2 è la funzione di Debye di ordine 2. Pur in assenza di una relazione invertibile in forma chiusa, il legame con il parametro θ è trattabile dal punto di vista numerico.

Relazione tra tau di Kendall e rho di Spearman

Sebbene il tau di Kendall e il rho di Spearman misurino entrambi la dipendenza, essi catturano aspetti leggermente diversi. In generale, la relazione tra le due misure non è lineare e dipende dalla famiglia di copula considerata. Tuttavia, è possibile stabilire vincoli generali del tipo:

$$-1 \leq 3\tau - 2\rho_S \leq 1$$

che forniscono un legame qualitativo tra le due quantità. Nel caso particolare della copula Gaussiana bivariata con coefficiente di correlazione lineare ρ, le relazioni assumono una forma particolarmente semplice ed elegante:

$$\tau = \frac{2}{\pi}\arcsin(\rho), \quad \rho_S = \frac{6}{\pi}\arcsin\left(\frac{\rho}{2}\right) \tag{1.44}$$

Queste identità mostrano come, nel contesto ellittico, tutte le principali misure di dipendenza siano riconducibili a un unico parametro.

1.5.3 Stimatori campionari

Nelle applicazioni pratiche il τ e il ρ_S vengono stimati a partire dai campioni raccolti e dai ranghi osservati.

Definizione 7. Dato un campione $(x_1, y_1), \ldots, (x_n, y_n)$, lo stimatore del tau di Kendall, ovvero il **tau di Kendall campionario** è:

$$\hat{\tau} = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \text{sign}[(x_i - x_j)(y_i - y_j)] \tag{1.45}$$

Analogamente, lo stimatore campionario del rho di Spearman può essere espresso in funzione dei ranghi osservati.

Definizione 8. Siano R_i e S_i rispettivamente i ranghi di x_i e y_i. Il **rho di Spearman** campionario è:

$$\hat{\rho}_S = \frac{\sum (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum (R_i - \bar{R})^2 \, \sum (S_i - \bar{S})^2}} = 1 - \frac{6 \sum_{i=1}^{n} (R_i - S_i)^2}{n(n^2 - 1)}.$$

L'ultima uguaglianza è valida soltanto in assenza di pari ranghi, ossia quando tutte le osservazioni sono distinte. Sotto opportune condizioni di regolarità, sia il tau di Kendall sia il rho di Spearman risultano consistenti e asintoticamente normali. Tali proprietà, insieme alla loro interpretazione intuitiva e all'invarianza rispetto a trasformazioni monotone, spiegano il ruolo centrale che queste misure rivestono nella teoria delle copule e nelle applicazioni statistiche.

1.6 Il comportamento nelle code

Un aspetto centrale nella scelta di una copula per applicazioni pratiche, in particolare in ambito finanziario e nella gestione del rischio, riguarda il modo in cui la dipendenza si manifesta nelle code della distribuzione. In molti contesti, infatti, non è tanto la dipendenza media a essere rilevante, quanto la probabilità che eventi estremi si verifichino simultaneamente. Le misure di *tail dependence* forniscono uno strumento formale per quantificare proprio questo tipo di comportamento.

1.6.1 Coefficienti di dipendenza nelle code

Siano U_1 e U_2 due variabili aleatorie uniformi su [0,1] con copula C.

Definizione 9. Il ***coefficiente di dipendenza nella coda superiore*** è definito come:

$$\lambda_U = \lim_{u \to 1^-} \mathbb{P}[U_2 > u \mid U_1 > u] = \lim_{u \to 1^-} \frac{1 - 2u + C(u,u)}{1-u} \tag{1.46}$$

se il limite esiste in [0,1].

Analogamente, il comportamento congiunto delle variabili in corrispondenza di valori molto piccoli è descritto dal coefficiente di dipendenza nella coda inferiore.

Definizione 10. Il ***coefficiente di dipendenza nella coda inferiore*** è definito come:

$$\lambda_L = \lim_{u \to 0^+} \mathbb{P}[U_2 \leq u \mid U_1 \leq u] = \lim_{u \to 0^+} \frac{C(u,u)}{u} \tag{1.47}$$

se il limite esiste in [0,1].

Valori positivi di λ_U o λ_L indicano la presenza di dipendenza asintotica nella rispettiva coda, mentre il valore nullo segnala indipendenza asintotica. In quest'ultimo caso, ovvero quando $\lambda_U = 0$ o $\lambda_L = 0$, eventi estremi possono verificarsi simultaneamente solo con probabilità trascurabile.

1.6.2 Tail dependence per copule Archimedee

Per una copula Archimedea con generatore ϕ, i coefficienti di tail dependence assumono una forma particolare, come enunciato dal seguente teorema.

Teorema 7. *Sia C una copula Archimedea con generatore ϕ. Allora:*

$$\lambda_L = \lim_{t \to 0^+} \frac{\phi^{-1}(2\phi(t))}{t} \tag{1.48}$$

$$\lambda_U = 2 - \lim_{t \to 1^-} \frac{1 - \phi^{-1}(2\phi(t))}{1-t}, \tag{1.49}$$

se i limiti esistono.

Tabella 1.2: Coefficienti di tail dependence per copule Archimedee

Copula	**Parametro**	λ_L	λ_U
Clayton	$\theta > 0$	$2^{-1/\theta}$	0
Gumbel	$\theta \geq 1$	0	$2 - 2^{1/\theta}$
Frank	$\theta \in \mathbb{R} \setminus \{0\}$	0	0
Gaussiana	$\rho \in (-1,1)$	0	0
t di Student	$\nu > 0, \rho \in (-1,1)$	$\lambda_L > 0$	$\lambda_U > 0$

Queste espressioni mostrano come il comportamento nelle code dipenda direttamente dal generatore della copula, in particolare della sua pseudo-inversa agli estremi del dominio. In Tabella 1.2 riassumiamo i coefficienti di dipendenza nelle code per le principali famiglie di copule considerate in questo elaborato. Dalla tabella emergono differenze sostanziali tra le varie famiglie.

- La copula di Clayton è caratterizzata da dipendenza asintotica nella coda inferiore e risulta quindi adatta a modellare situazioni in cui eventi estremi negativi tendono a manifestarsi congiuntamente, come nel caso di perdite simultanee nei mercati finanziari. La copula di Gumbel, al contrario, concentra la dipendenza nella coda superiore ed è indicata per modellare co-movimenti estremi positivi.

- Le copule di Frank e Gaussiana non presentano dipendenza asintotica in nessuna delle due code: la loro struttura di dipendenza è quindi concentrata principalmente nella regione centrale. Infine, la copula t di Student rappresenta un caso particolarmente interessante, poiché introduce dipendenza simmetrica nelle code, la cui intensità aumenta al diminuire dei gradi di libertà:

$$\lambda_L = \lambda_U = 2\, t_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right),$$

  dove $t_{\nu+1}$ è la funzione di ripartizione univariata della t di Student con $\nu + 1$ gradi di libertà.

Tabella 1.3: Interpretazione del parametro θ

Copula	Dominio	Indipendenza	Comonotonia	Dipendenza
Clayton	$[-1, \infty) \setminus \{0\}$	$\theta \to 0$	$\theta \to \infty$	Positiva per $\theta > 0$
Gumbel	$[1, \infty)$	$\theta = 1$	$\theta \to \infty$	Sempre positiva
Frank	$\mathbb{R} \setminus \{0\}$	$\theta \to 0$	$\theta \to \infty$	Positiva/negativa

Interpretazione del parametro di dipendenza

Per ogni famiglia di copule Archimedee, il parametro θ controlla l'intensità e la forma della dipendenza, come si evince dalla Tabella 1.3. Sebbene il parametro θ governi la dipendenza all'interno di ciascuna famiglia, la sua interpretazione non è uniforme tra copule diverse. Questo rende difficile il confronto diretto tra modelli basati su famiglie differenti. Nel capitolo successivo vedremo come la relazione tra θ e il tau di Kendall permetta di superare questo limite, fornendo una parametrizzazione interpretabile e comparabile, in cui $\tau \in [-1,1]$ misura direttamente il grado di concordanza tra le variabili indipendentemente dalla copula scelta.

1.6.3 Selezione del modello

La scelta della famiglia di copule più appropriata per descrivere dei dati non può basarsi su un singolo criterio, ma deve essere il risultato di una valutazione complessiva che tenga conto sia delle caratteristiche dei dati sia del contesto applicativo [17]. Un primo passo è rappresentato dall'analisi esplorativa, che consente di individuare eventuali pattern di dipendenza attraverso scatterplot, rappresentazioni sui ranghi e un'osservazione preliminare del comportamento nelle code della distribuzione. Accanto all'evidenza empirica, il contesto applicativo gioca un ruolo determinante. In ambito finanziario, ad esempio, la presenza di forti co-movimenti durante fasi di stress di mercato suggerisce l'impiego di copule in grado di catturare dipendenza nelle code, come la copula di Clayton per la coda inferiore o la copula t di Student per una dipendenza simmetrica negli eventi estremi. Per un'interpretazione finanziaria della dipendenza nelle code si rimanda a [7]. La selezione del modello può inoltre essere supportata da strumenti statistici formali, come i test di bontà dell'adattamento che permettono di valutare la coerenza tra il modello di copula ipotizzato e i dati osservati, fornendo un criterio oggettivo per il confronto tra famiglie alternative. Nel contesto bayesiano, criteri come il DIC e il WAIC forniscono un compromesso tra qualità dell'adattamento e complessità del modello [18, 19]. L'adozione di modelli semplici facilita l'interpretazione dei risultati e riduce il rischio di *overfitting*, mantenendo al tempo stesso una buona capacità descrittiva della struttura di dipendenza.

Conclusioni del capitolo

In questo capitolo sono stati introdotti i concetti fondamentali della teoria delle copule, ponendo l'attenzione sul problema della modellizzazione della dipendenza tra variabili aleatorie. A partire dai limiti della correlazione lineare, abbiamo mostrato come il teorema di Sklar consenta di separare in modo rigoroso la struttura di dipendenza dalle distribuzioni marginali, offrendo un quadro concettuale più generale. Un'attenzione particolare è stata dedicata alle copule Archimedee, che grazie alla loro struttura algebrica semplice e all'uso di un generatore unidimensionale rappresentano una classe estremamente versatile. L'analisi delle principali famiglie ha evidenziato come ciascuna copula presenti caratteristiche specifiche, in particolare per quanto riguarda il comportamento nelle code e l'interpretazione del parametro di dipendenza. Sono state inoltre introdotte le principali misure di dipendenza basate sui ranghi, il tau di Kendall e il rho di Spearman, che offrono strumenti robusti, invarianti rispetto alle trasformazioni monotone e strettamente legati alla struttura della copula. Per molte famiglie Archimedee, tali misure ammettono espressioni esplicite, rendendo immediato il collegamento con il parametro di copula θ. Il capitolo successivo sarà dedicato all'inferenza statistica per i modelli di copula, con particolare attenzione all'approccio bayesiano. A differenza dell'approccio semiparametrico di Grazian e Liseo (2017) [14], in questo lavoro adotteremo una riparametrizzazione parametrica in termini del coefficiente di Kendall τ, che consente un'interpretazione più uniforme e vantaggi significativi dal punto di vista computazionale.

Capitolo 2

L'inferenza bayesiana nei modelli di copula

2.1 Motivazioni per l'approccio bayesiano

L'inferenza statistica nei modelli di copula può essere affrontata attraverso diversi approcci, che spaziano dalla massima verosimiglianza ai metodi basati sui momenti, fino all'inferenza bayesiana. In questo capitolo ci concentreremo su quest'ultima, mettendo in evidenza le ragioni che la rendono particolarmente adatta allo studio delle copule Archimedee. L'obiettivo non è soltanto illustrarne i vantaggi teorici, ma anche mostrare come l'approccio bayesiano permetta di superare alcune difficoltà tipiche dei metodi frequentisti.

Limiti degli approcci frequentisti

Gli stimatori frequentisti comunemente impiegati per i modelli di copula presentano diversi limiti, che diventano particolarmente evidenti in presenza di strutture di dipendenza complesse o modelli fortemente non lineari. Un primo esempio è rappresentato dal metodo *IFM (Inference for Margins)* proposto da Joe e Xu (1996) [9]. L'idea alla base di questo approccio consiste nello stimare inizialmente le distribuzioni marginali in modo separato e, successivamente, i parametri della copula, rendendo la procedura computazionalmente efficiente e relativamente semplice da implementare. Tuttavia, questa separazione comporta una perdita di informazione: l'incertezza sulle stime marginali non viene propagata ai parametri della copula, con il rischio di ottenere stime meno precise e intervalli di confidenza poco affidabili [8, 20]. Anche la massima verosimiglianza, sebbene teoricamente efficiente, non è sempre una soluzione praticabile: la funzione di verosimiglianza può diventare difficile da ottimizzare, presentare più massimi locali o zone piatte, e in campioni

finiti gli errori standard asintotici possono risultare fuorvianti [10]. Accanto a questi metodi, esistono approcci semi–parametrici basati su statistiche di rango, come quello fondato sul tau di Kendall, utili in molte applicazioni [8] , ma che forniscono solo stime puntuali, non consentendo di quantificare in modo diretto l'incertezza, limitazione che ha motivato lo sviluppo di approcci bayesiani alternativi. A differenza dell'approccio semiparametrico di Grazian e Liseo (2017)[14], che mira all'inferenza diretta su τ come funzionale della copula senza assumerne una forma parametrica, in questo lavoro si considera un modello di copula parametrico e si adotta una riparametrizzazione del parametro θ in termini del coefficiente di Kendall, sfruttando la relazione biunivoca esistente per molte famiglie Archimedee.

Vantaggi dell'approccio bayesiano

L'inferenza bayesiana permette di affrontare simultaneamente sia la stima sia la quantificazione dell'incertezza. A differenza dei metodi frequentisti, i parametri marginali e quelli della copula vengono stimati congiuntamente, e l'incertezza si propaga poi attraverso tutte le componenti del modello. Un ruolo centrale è svolto dalla distribuzione a posteriori, che permette di descrivere in modo completo la variabilità delle stime, consentendo di costruire intervalli di credibilità spesso più affidabili degli intervalli di confidenza asintotici, soprattutto quando la dimensione campionaria è limitata [10]. Inoltre, la possibilità di introdurre informazione a priori rappresenta un ulteriore vantaggio: in presenza di dati scarsi o quando la teoria suggerisce vincoli o relazioni tra i parametri, le distribuzioni a priori permettono di incorporare tali elementi nella stima. L'impostazione bayesiana mette inoltre a disposizione strumenti naturali per la selezione dei modelli come il DIC e il WAIC [18, 19]. Infine, l'approccio bayesiano risulta particolarmente utile quando si affrontano modelli complessi: strutture gerarchiche, dati mancanti o forme di dipendenza non standard possono essere trattati grazie ai metodi Monte Carlo basati sulle catene di Markov (MCMC) [21].

Sfide computazionali

Naturalmente l'inferenza bayesiana non è priva di difficoltà. L'implementazione pratica richiede la progettazione di algoritmi MCMC che esplorino in modo efficiente lo spazio dei parametri, la scelta di distribuzioni a priori adeguate e di verificare con attenzione la convergenza delle catene. Per modelli complessi o campioni di grandi dimensioni, il costo computazionale può diventare significativo. Negli ultimi anni, tuttavia, lo sviluppo di algoritmi più sofisticati e l'aumento della capacità computazionale hanno reso l'inferenza bayesiana sempre più accessibile. Ciò ha contribuito a diffonderne l'uso anche in contesti, come quello delle copule, che un tempo erano considerati troppo onerosi o difficili da trattare con i metodi MCMC tradizionali [22].

2.2 Riformulazione della parametrizzazione in τ

La parametrizzazione standard delle copule Archimedee utilizza direttamente il parametro θ che caratterizza il generatore ϕ_θ. Questa scelta, sebbene naturale dal punto di vista della costruzione della copula, presenta diverse limitazioni nell'inferenza bayesiana.

Parametrizzazione tramite θ

Consideriamo una copula Archimedea bivariata con generatore $\phi_\theta(t)$ e $\theta \in \Theta \subseteq \mathbb{R}$ parametro di dipendenza. Abbiamo visto nel capitolo precedente che tale copula è definita come:

$$C_\theta(u, v) = \phi_\theta^{-1}(\phi_\theta(u) + \phi_\theta(v)) \tag{2.1}$$

dove ϕ_θ^{-1} indica l'inversa del generatore; supponendo ϕ_θ due volte differenziabile, la densità associata alla copula poteva essere espressa come:

$$c_\theta(u, v) = \frac{\phi_\theta''\Big(\phi_\theta^{-1}(\phi_\theta(u) + \phi_\theta(v))\Big) \cdot \phi_\theta'(u) \cdot \phi_\theta'(v)}{[\phi_\theta'(\phi_\theta^{-1}(\phi_\theta(u) + \phi_\theta(v)))]^3} \tag{2.2}$$

Se ora prendiamo un campione $(u_1, v_1), \ldots, (u_n, v_n)$ dalla copula, possiamo definire la funzione di verosimiglianza per il parametro θ:

$$L(\theta \mid \mathbf{u}, \mathbf{v}) = \prod_{i=1}^{n} c_\theta(u_i, v_i) \tag{2.3}$$

Questa formulazione costituisce il punto di partenza naturale per l'inferenza frequentista e bayesiana [6].

Limitazioni della parametrizzazione in θ

Sebbene l'utilizzo diretto del parametro θ sia la scelta più immediata dal punto di vista costruttivo, questa parametrizzazione presenta diversi inconvenienti quando ci si sposta sul terreno dell'inferenza bayesiana [10]. Alcuni di questi problemi emergono già in fase di modellizzazione, altri diventano particolarmente evidenti nella pratica computazionale con algoritmi MCMC. Una prima difficoltà riguarda l'interpretabilità e la comparabilità tra famiglie diverse. Il parametro θ non possiede infatti un significato uniforme tra le diverse copule Archimedee: ciascuna famiglia è definita su un dominio differente e associa valori numerici simili a livelli di dipendenza molto diversi. Nella copula di Clayton, ad esempio, θ assume valori in $[-1, \infty) \setminus \{0\}$, mentre per la Gumbel è vincolato a $[1, \infty)$ e per la Frank può spaziare su tutto $\mathbb{R} \setminus \{0\}$ [6]. Questa eterogeneità comporta che lo stesso valore numerico di θ descriva livelli di dipendenza molto diversi a seconda della famiglia considerata,

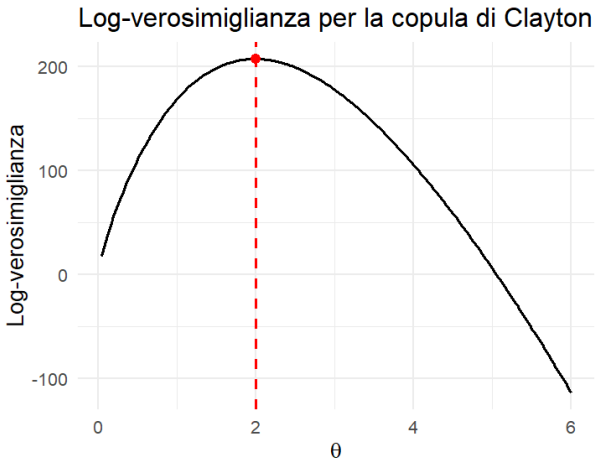


Figura 2.1: Log-verosimiglianza della copula di Clayton con parametro $\theta_{true} = 2$.

rendendo difficile sia la scelta di una priori comune sia il confronto diretto tra modelli diversi. Strettamente connesso a questo aspetto è il comportamento della funzione di verosimiglianza. Per molte copule Archimedee, $L(\theta)$ risulta fortemente asimmetrica e, in tali situazioni, l'uso di priori standard simmetriche, come la distribuzione normale, risulta poco appropriato e può influenzare in modo non trascurabile la forma della distribuzione a posteriori [12, 23]. A ciò si aggiunge la necessità di gestire i vincoli sul dominio, dove diventa necessario introdurre priori troncate o trasformazioni del parametro, aumentando la complessità del modello e rallentando potenzialmente la convergenza delle catene. Alcune di queste difficoltà sono illustrate nel prossimo esempio, dove analizziamo il comportamento della log-verosimiglianza su alcuni dati simulati tramite il codice in Appendice A.

Esempio 3. Consideriamo la copula di Clayton con $\theta_{\text{true}} = 2$, corrispondente a $\tau = 0.5$. Per evidenziare le difficoltà legate alla parametrizzazione diretta in θ, analizziamo l'andamento della log-verosimiglianza della copula di Clayton. Il profilo della log-verosimiglianza, riportato in Figura 2.1, mostra una evidente asimmetria: la funzione decresce rapidamente in prossimità di $\theta = 0$, mentre tende ad appiattirsi per valori più elevati del parametro. Questa forma non parabolica può ridurre l'efficienza di algoritmi MCMC che utilizzano proposte simmetriche, come il random-walk Metropolis [15].

Le caratteristiche emerse dall'esempio appena visto rendono la parametrizzazione in θ poco adatta a un'esplorazione stabile ed efficace dello spazio dei parametri, e motivano l'adozione di parametrizzazioni alternative, come il tau di Kendall.

Motivazione della riparametrizzazione in τ

L'idea di utilizzare τ come parametro di dipendenza si fonda su una serie di proprietà che lo rendono particolarmente adatto all'inferenza sulle copule. Come discusso nel Capitolo 1, il tau di Kendall è una misura di dipendenza che quantifica il grado di concordanza tra due variabili aleatorie ed è definito come [13]:

$$\tau = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0] \tag{2.4}$$

con (X_1, Y_1) e (X_2, Y_2) coppie indipendenti dalla stessa distribuzione bivariata.
In primo luogo, τ assume sempre valori nell'intervallo $[-1,1]$ indipendentemente dalla famiglia considerata, fornendo così un dominio standardizzato. Inoltre, la sua interpretazione è immediata: valori prossimi a zero indicano indipendenza, mentre i limiti ± 1 corrispondono rispettivamente a comonotonia e contromonotonia perfette. Un ulteriore aspetto cruciale è l'invarianza rispetto a trasformazioni monotone crescenti delle variabili, proprietà fondamentale nella teoria delle copule [1]. Infine, essendo basato sui ranghi, il tau di Kendall non dipende dalle distribuzioni marginali, consentendo un confronto diretto del grado di dipendenza anche tra famiglie di copule diverse [1].

Trasformazione $\theta = g(\tau)$

Come mostrato nel Teorema 6, per una copula Archimedea con generatore ϕ_θ il tau di Kendall può essere espresso come funzione del parametro θ:

$$\tau = h(\theta) = 1 + 4 \int_0^1 \frac{\phi_\theta(t)}{\phi'_\theta(t)}\, dt \tag{2.5}$$

dove $h : \Theta \to [-1, 1]$ è una funzione strettamente monotona per le famiglie Archimedee considerate [1]. Di conseguenza, tale relazione ammette un'unica funzione inversa, che consente di esprimere il parametro del generatore direttamente in funzione di τ:

$$\theta = g(\tau) = h^{-1}(\tau) \tag{2.6}$$

Per alcune famiglie di copule Archimedee poi la relazione tra θ e τ ammette un'espressione in forma chiusa particolarmente semplice, come nel caso già visto della Clayton:

$$\tau = \frac{\theta}{\theta + 2} \quad \Rightarrow \quad \theta = g_{\text{Clay}}(\tau) = \frac{2\tau}{1 - \tau} \tag{2.7}$$

con $\tau \in (0, 1)$ per $\theta > 0$ e $\tau \in (-1, 0)$ per $\theta \in [-1, 0)$.

2.2.1 Verosimiglianza nella parametrizzazione τ

Adottando la nuova parametrizzazione, la densità della copula può essere riscritta in funzione di τ, sostituendo il parametro θ con la trasformazione inversa definita nella sezione precedente:

$$c_\tau(u, v) \coloneqq c_{\theta=g(\tau)}(u, v) \tag{2.8}$$

Dato un campione di dimensione n, la verosimiglianza risulta quindi:

$$L(\tau \mid \mathbf{u}, \mathbf{v}) = \prod_{i=1}^{n} c_\tau(u_i, v_i) = \prod_{i=1}^{n} c_{\theta=g(\tau)}(u_i, v_i) \tag{2.9}$$

Questa riformulazione determina proprietà particolarmente favorevoli. In primo luogo la verosimiglianza è definita su un dominio limitato, $\tau \in [-1,1]$, facilitando l'ottimizzazione numerica e l'esplorazione mediante MCMC. Inoltre, per molte famiglie di copule la funzione $L(\tau)$ tende a presentare una forma più regolare e meglio bilanciata rispetto alla corrispondente in θ, rendendo più naturali le scelte di priori simmetriche [12].

2.2.2 Distribuzione a priori e a posteriori su τ

La parametrizzazione in τ consente di specificare distribuzioni a priori direttamente interpretabili in termini di dipendenza. Una scelta naturale, in assenza di informazioni preliminari, è la distribuzione Uniforme, $\tau \sim \text{Unif}(-1,1)$, con densità:

$$\pi(\tau) = \frac{1}{2}, \quad \tau \in (-1, 1) \tag{2.10}$$

Quando invece si dispone di informazione a priori si può utilizzare una distribuzione Beta riscalata su $(-1, 1)$, cioè $\tau \sim 2 \cdot \text{Beta}(a, b) - 1$, con densità:

$$\pi(\tau) = \frac{1}{2}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\left(\frac{\tau+1}{2}\right)^{a-1}\left(\frac{1-\tau}{2}\right)^{b-1} \qquad \tau \in (-1,1) \tag{2.11}$$

che permette di concentrare la massa a priori in regioni di interesse. Ponendo poi $\tilde{\tau} = \frac{(\tau+1)}{2} \in (0, 1)$ si ottiene la riscrittura equivalente:

$$\pi(\tilde{\tau}) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\tilde{\tau}^{a-1}(1-\tilde{\tau})^{b-1} \qquad \tilde{\tau} \in (0,1)$$

Al variare dei parametri a e b si ottengono diverse situazioni:

$$\begin{aligned} a = b = 1 \quad &\Rightarrow \quad \text{priori uniforme,} \\ a = b > 1 \quad &\Rightarrow \quad \text{massa a priori concentrata attorno a } \tau = 0, \\ a > b \quad &\Rightarrow \quad \text{preferenza per dipendenza positiva,} \\ a < b \quad &\Rightarrow \quad \text{preferenza per dipendenza negativa.} \end{aligned} \tag{2.12}$$

L'alternativa alla Beta, qualora siano disponibili informazioni provenienti da studi precedenti o dal contesto applicativo, è una priori troncata:

$$\pi(\tau) \propto \mathbb{I}(\tau \in [\tau_{\min}, \tau_{\max}]) \cdot \pi_0(\tau) \tag{2.13}$$

dove $[\tau_{\min}, \tau_{\max}] \subset (-1, 1)$ rappresenta un intervallo plausibile a priori e $\pi_0(\tau)$ una densità di base. Dopo aver scelto la priori, applicando il teorema di Bayes, si ottiene la distribuzione a posteriori di τ :

$$\pi(\tau \mid \mathbf{u}, \mathbf{v}) \propto L(\tau \mid \mathbf{u}, \mathbf{v}) \cdot \pi(\tau) \tag{2.14}$$

ovvero:

$$\pi(\tau \mid \mathbf{u}, \mathbf{v}) \propto \left[\prod_{i=1}^{n} c_{g(\tau)}(u_i, v_i)\right] \cdot \pi(\tau) \tag{2.15}$$

Questa formulazione consente un'inferenza diretta sulla misura di dipendenza τ, senza necessità di trasformazioni post stima. Gli intervalli di credibilità hanno interpretazione probabilistica immediata e, per campioni sufficientemente grandi, la distribuzione a posteriori si concentra attorno al vero valore del parametro. Osserviamo inoltre che, qualora si desideri recuperare l'inferenza su θ, è sufficiente applicare la trasformazione inversa:

$$\pi(\theta \mid \mathbf{u}, \mathbf{v}) = \pi(\tau = h(\theta) \mid \mathbf{u}, \mathbf{v}) \cdot |h'(\theta)| \tag{2.16}$$

dove $h'(\theta) = \frac{dh(\theta)}{d\theta} = \frac{d\tau}{d\theta}$ è lo Jacobiano della trasformazione.

2.2.3 Vantaggi computazionali

Dal punto di vista numerico, la parametrizzazione in τ può migliorare il comportamento degli algoritmi MCMC [15, 12], poiché il dominio limitato e continuo del parametro semplifica la costruzione delle proposte e favorisce una migliore esplorazione della distribuzione a posteriori. Nel seguito presentiamo un esempio puramente illustrativo, basato sulla valutazione della posteriori su una griglia di valori del parametro, al fine di evidenziare le differenze tra le due parametrizzazioni. L'utilizzo esplicito di metodi MCMC sarà introdotto nelle sezioni successive.

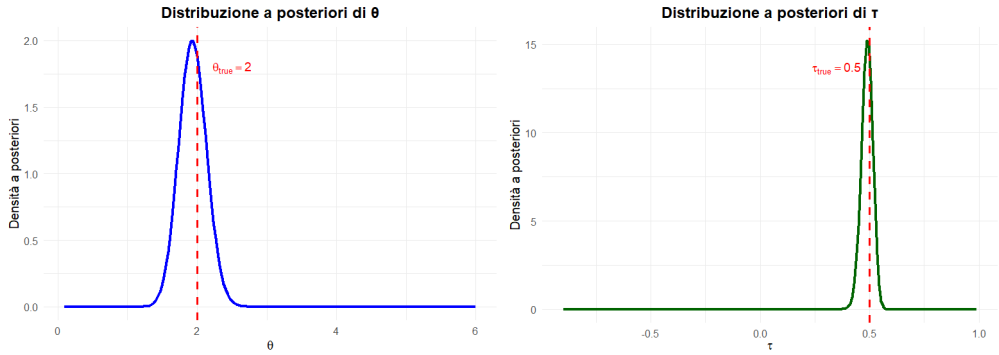


Figura 2.2: Confronto tra le densità a posteriori nelle parametrizzazioni in θ e τ.

Tabella 2.1: Statistiche riassuntive delle distribuzioni a posteriori per le due parametrizzazioni della copula di Clayton.

Statistica	**Parametrizzazione in θ**	**Parametrizzazione in τ**
Valore vero	2.000	0.500
Media a posteriori	1.936	0.488
Deviazione standard	0.200	0.026

Esempio 4. Ricordiamo che per la copula di Clayton la trasformazione esplicita è $\theta = \frac{2\tau}{1-\tau}$. Assumendo una priori uniforme su τ:

$$\pi(\tau) = \frac{1}{2}, \quad \tau \in (-1, 1) \tag{2.17}$$

la distribuzione a posteriori risulta

$$\pi(\tau \mid \mathbf{u}, \mathbf{v}) \propto \prod_{i=1}^{n} c_{\frac{2\tau}{1-\tau}}(u_i, v_i) \tag{2.18}$$

La Figura 2.2 mette a confronto nelle due parametrizzazioni le densità a posteriori, ottenute valutando la funzione di verosimiglianza su una griglia di valori del parametro e successivamente normalizzando i valori ottenuti. Le linee tratteggiate in rosso indicano i valori veri dei parametri $\theta_{\text{true}} = 2$ e $\tau_{\text{true}} = 0.5$. Entrambe le medie a posteriori sono vicine ai valori veri, confermando la consistenza degli stimatori bayesiani. La parametrizzazione in τ presenta una deviazione standard notevolmente più contenuta, indicando una posteriori molto più concentrata attorno al valore vero.

2.3 Algoritmo MCMC: Metropolis-Hastings

L'inferenza bayesiana sulla distribuzione a posteriori $\pi(\tau \mid \mathbf{u}, \mathbf{v})$ non è, in generale, trattabile in forma chiusa. Per questo motivo è necessario ricorrere a metodi Monte Carlo basati sulle catene di Markov (MCMC), che permettono di approssimare la distribuzione a posteriori mediante simulazione. In questa sezione descriviamo l'algoritmo di Metropolis-Hastings adattato alla parametrizzazione in termini del tau di Kendall, soffermandoci sugli aspetti che risultano più rilevanti dal punto di vista applicativo.

2.3.1 Algoritmo di Metropolis-Hastings

L'algoritmo di Metropolis-Hastings (Metropolis et al., 1953 [24]; Hastings, 1970 [25]) è uno dei metodi MCMC più flessibili e diffusi. A partire da un valore iniziale del parametro, genera una sequenza di valori che, sotto opportune condizioni di regolarità, converge in distribuzione alla distribuzione target, in questo caso la posteriori di τ.

Algorithm 1 Metropolis-Hastings per τ

1: **Input:** Valore corrente $\tau^{(t)}$, osservazioni $(\mathbf{u}, \mathbf{v})$
2: Genera un candidato $\tau^* \sim q(\cdot \mid \tau^{(t)})$ dalla densità di proposta
3: Calcola il rapporto di accettazione:

$$\alpha(\tau^{(t)}, \tau^*) = \min\left\{1, \frac{\pi(\tau^* \mid \mathbf{u}, \mathbf{v})}{\pi(\tau^{(t)} \mid \mathbf{u}, \mathbf{v})} \cdot \frac{q(\tau^{(t)} \mid \tau^*)}{q(\tau^* \mid \tau^{(t)})}\right\} \tag{2.19}$$

4: Campiona $u \sim \text{Uniforme}(0, 1)$
5: **If** $u \leq \alpha(\tau^{(t)}, \tau^*)$ **then**
6: Accetta: $\tau^{(t+1)} = \tau^*$
7: **else**
8: Rifiuta: $\tau^{(t+1)} = \tau^{(t)}$
9: **end if**
10: **Return:** $\tau^{(t+1)}$

Sia $\tau^{(t)}$ il valore corrente della catena al passo t. L'algoritmo procede generando un valore candidato τ^* da una distribuzione di proposta $q(\cdot \mid \tau^{(t)})$ e decidendo se accettarlo o meno sulla base di una probabilità di accettazione che dipende dal rapporto tra le densità a posteriori nel punto candidato e nel punto corrente.

Sostituendo le espressioni della posteriori, il rapporto di accettazione diventa:

$$\alpha(\tau^{(t)}, \tau^*) = \min\left\{1,\ \frac{L(\tau^* \mid \mathbf{u}, \mathbf{v}) \cdot \pi(\tau^*)}{L(\tau^{(t)} \mid \mathbf{u}, \mathbf{v}) \cdot \pi(\tau^{(t)})} \cdot \frac{q(\tau^{(t)} \mid \tau^*)}{q(\tau^* \mid \tau^{(t)})}\right\} \tag{2.20}$$

Nel caso in cui la distribuzione di proposta sia simmetrica, il rapporto tra le densità di proposta si semplifica a uno, rendendo il calcolo della probabilità di accettazione particolarmente semplice.

2.3.2 Scelta della distribuzione di proposta

La scelta della distribuzione di proposta $q(\cdot \mid \tau^{(t)})$ influenza in modo determinante l'efficienza dell'algoritmo. Una proposta mal calibrata può infatti produrre catene fortemente autocorrelate o, al contrario, tassi di accettazione molto bassi.

Random Walk Metropolis

Una scelta naturale è la proposta di tipo random walk, in cui il valore candidato viene ottenuto perturbando il valore corrente:

$$\tau^* = \tau^{(t)} + \epsilon, \tag{2.21}$$

dove ε è una variabile aleatoria simmetrica centrata in zero. Le scelte più comuni sono una distribuzione normale $\mathcal{N}(0, \sigma^2)$ o una distribuzione uniforme su un intervallo simmetrico. Essendo simmetriche, tali proposte semplificano il rapporto di accettazione poiché $q(\tau^{(t)} \mid \tau^*) = q(\tau^* \mid \tau^{(t)})$. Inoltre, poiché il parametro τ è vincolato all'intervallo $(-1{,}1)$, è necessario gestire le proposte che cadono al di fuori di tale dominio. In pratica, si può scegliere se rifiutare automaticamente tali proposte ponendo $\alpha = 0$ oppure applicare una trasformazione di riflessione. Nel seguito adotteremo la soluzione più semplice, ovvero il rifiuto automatico.

Proposta basata su distribuzione Beta

Un'alternativa più sofisticata consiste nell'utilizzare una distribuzione Beta riscalata [21], che garantisce automaticamente il rispetto dei vincoli sul dominio di τ. In questo caso il candidato viene generato come

$$\tau^* \sim 2 \cdot \text{Beta}(\alpha(\tau^{(t)}), \beta(\tau^{(t)})) - 1 \tag{2.22}$$

dove i parametri α e β sono scelti in modo che la media della proposta coincida con il valore corrente $\tau^{(t)}$:

$$\mu^* = \frac{\tau^{(t)} + 1}{2} \in (0, 1)$$

Una scelta naturale per i parametri risulta quindi

$$\alpha(\tau^{(t)}) = \mu^*\kappa, \qquad \beta(\tau^{(t)}) = (1-\mu^*)\kappa$$

dove $\kappa > 0$, parametro di tuning, controlla la varianza della proposta: valori più alti di κ producono proposte più concentrate attorno a $\tau^{(t)}$. In questo modo la proposta è automaticamente vincolata in $(-1, 1)$, ma non è più simmetrica, quindi occorre includere il rapporto $\frac{q(\tau^{(t)}|\tau^*)}{q(\tau^*|\tau^{(t)})}$ nel calcolo di α.

Proposta adattiva e tuning dell'algoritmo

Per migliorare l'efficienza iniziale dell'algoritmo, è spesso utile adottare strategie adattive durante le prime iterazioni [26]. In particolare, la varianza della proposta nel random walk può essere aggiornata dinamicamente allo scopo di raggiungere un tasso di accettazione desiderato. Dopo una fase di tuning, l'adattamento viene interrotto per garantire la correttezza teorica dell'algoritmo.

Algorithm 2 Adaptive Random Walk Metropolis

Input: Valore corrente $\tau^{(t)}$, varianza corrente σ_t^2, iterazione corrente t, target di accettazione α_{target}, numero di iterazioni tuning T_{tune}
Esegui un passo Metropolis-Hastings con proposta $\mathcal{N}(\tau^{(t)}, \sigma_t^2)$
Calcola il tasso di accettazione corrente $\hat{\alpha}_t$
If $t \leq T_{\text{tune}}$ **then**
    If $\hat{\alpha}_t > \alpha_{\text{target}}$ **then**
        Aumenta: $\sigma_{t+1}^2 = \sigma_t^2 \cdot 1.05$
    else
        Diminuisci: $\sigma_{t+1}^2 = \sigma_t^2 \cdot 0.95$
    end if
else
    Fissa: $\sigma_{t+1}^2 = \sigma_t^2$
end if

Un'adeguata calibrazione dei parametri dell'algoritmo Metropolis-Hastings è essenziale per garantire un'esplorazione efficiente della distribuzione a posteriori. In pratica, il parametro di scala della proposta viene inizialmente impostato su un valore moderato, ad esempio $\sigma \approx 0.1$, e successivamente adattato durante una fase preliminare di tuning. Durante questa fase iniziale, la varianza della proposta viene aggiornata dinamicamente in base al tasso di accettazione osservato, aumentando o diminuendo progressivamente fino a raggiungere il range desiderato. Al termine del tuning, il parametro di scala viene fissato e l'algoritmo prosegue con

una fase di burn-in e una successiva fase di campionamento vero e proprio. Nel caso unidimensionale considerato in questo lavoro, la letteratura suggerisce come valore ottimale un tasso di accettazione α_{target} compreso approssimativamente tra il 20% e il 40%, con un valore teoricamente ottimale intorno al 23.4% per il random walk Metropolis (Roberts, Gelman e Gilks, 1997) [15]. Ad esempio, nel caso della proposta normale $\mathcal{N}(0, \sigma^2)$, la scelta di σ influenza il tasso di accettazione:

- Se σ è troppo piccolo il tasso di accettazione sarà alto ($\alpha \approx 1$) ma la catena si muoverà lentamente, producendo alta autocorrelazione.

- Se σ è troppo grande, la probabilità di accettazione diminuisce, in quanto le proposte si collocano spesso in regioni di bassa densità a posteriori; la catena mostra quindi scarso movimento.

Diagnostica di convergenza

Una volta eseguita la simulazione MCMC, è fondamentale verificare tramite diversi strumenti diagnostici, come quelli riportati in Tabella 2.2, che la catena abbia raggiunto la distribuzione stazionaria, ovvero che i campioni ottenuti rappresentino correttamente la distribuzione a posteriori di interesse [22].

Tabella 2.2: Strumenti diagnostici per la convergenza MCMC

Strumento	**Descrizione e interpretazione**
Trace plot	Rappresenta l'andamento della sequenza $\{\tau^{(t)}\}$ nel tempo. In presenza di convergenza, la catena mostra comportamento stazionario senza trend evidenti ed esplora ripetutamente la regione di maggiore densità a posteriori.
Autocorrelazione	Misura il grado di dipendenza tra campioni successivi. È definita come $\rho_k = \text{Corr}(\tau^{(t)}, \tau^{(t+k)})$. In condizioni ideali decresce rapidamente al crescere del lag k.
Effective Sample Size	Stima il numero di osservazioni indipendenti. È definita come $\text{ESS} = \frac{T}{1+2\sum_{k=1}^{\infty} \rho_k}$. Valori molto inferiori a T indicano scarsa efficienza dell'algoritmo [11].

2.3.3 Implementazione pratica

Algorithm 3 MCMC per inferenza bayesiana su τ

Input: Dati $(\mathbf{u}, \mathbf{v})$, priori $\pi(\tau)$, T_{tune}, T_{burn}, T_{iter}, famiglia di copula
Inizializzazione:
  Calcola lo stimatore campionario $\hat{\tau}_{\text{Kendall}}$ dai dati
  Imposta $\tau^{(0)} = \hat{\tau}_{\text{Kendall}}$ (o $\tau^{(0)} = 0$ se non disponibile)
  Imposta $\sigma = 0.1$
Fase di Tuning:
for $t = 1$ to T_{tune} **do**
    Genera $\tau^* \sim \mathcal{N}(\tau^{(t-1)}, \sigma^2)$ troncata a $(-1, 1)$
    Calcola $\alpha(\tau^{(t-1)}, \tau^*)$ usando (2.20)
    Accetta/rifiuta secondo MH
    Aggiorna σ per tasso di accettazione target ≈ 0.25
end for
Fissa σ al valore finale del tuning
Fase di Burn-in:
for $t = 1$ to T_{burn} **do**
    Esegui passo MH con σ fisso
    Scarta questi campioni
end for
Fase di Campionamento:
for $t = 1$ to T_{iter} **do**
    Esegui passo MH con σ fisso
    Salva $\tau^{(t)}$ ogni k iterazioni
end for
Output: Campione dalla posteriori $\{\tau^{(j)}\}_{j=1}^{T_{\text{iter}}/k}$

Dal punto di vista operativo, l'algoritmo MCMC viene implementato seguendo una struttura standard, articolata in tre fasi principali: una fase iniziale di tuning, una fase di burn-in e una fase finale di campionamento. Questa suddivisione consente di calibrare preliminarmente i parametri dell'algoritmo, eliminare la dipendenza dalle condizioni iniziali e ottenere infine un campione rappresentativo dalla distribuzione a posteriori. L'algoritmo viene inizializzato a partire dai dati osservati e da una distribuzione a priori specificata sul parametro τ. Come valore iniziale della catena si utilizza, quando disponibile, lo stimatore campionario del tau di Kendall, che fornisce un punto di partenza ragionevole in termini di dipendenza. In alternativa, è possibile inizializzare la catena nel punto $\tau^{(0)} = 0$, corrispondente all'indipendenza. Nella fase di tuning iniziale, la catena viene fatta evolvere per un numero limitato

di iterazioni, durante le quali il parametro di scala della proposta del random walk Metropolis viene adattato al fine di ottenere un tasso di accettazione prossimo al valore ottimale. Questa fase ha lo scopo di calibrare l'algoritmo e non viene utilizzata per l'inferenza. Una volta fissato il valore del parametro di scala, la catena entra nella fase di burn-in, durante la quale le iterazioni iniziali vengono scartate per ridurre l'influenza delle condizioni iniziali. Infine, nella fase di campionamento vera e propria, i valori della catena vengono salvati, eventualmente applicando thinning, e utilizzati per l'analisi a posteriori.

Calcolo efficiente della verosimiglianza

In ciascun passo dell'algoritmo MCMC, per ogni valore candidato τ^* è necessario valutare la funzione di verosimiglianza associata alla copula. Dato un campione di dimensione n, la verosimiglianza è definita come il prodotto delle densità di copula valutate nei punti osservati.

$$L(\tau^* \mid \mathbf{u}, \mathbf{v}) = \prod_{i=1}^{n} c_{g(\tau^*)}(u_i, v_i) \tag{2.23}$$

Dal punto di vista numerico, il calcolo viene effettuato in scala logaritmica, così da evitare problemi di underflow quando la dimensione campionaria è elevata. In particolare, si lavora con la log-verosimiglianza

$$\log L(\tau^* \mid \mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} \log c_{g(\tau^*)}(u_i, v_i) \tag{2.24}$$

Il calcolo procede trasformando innanzitutto il valore candidato τ^* nel corrispondente parametro del generatore $\theta^* = g(\tau^*)$. Successivamente, per ciascuna osservazione (u_i, v_i) viene valutata la log-densità $\log c_{\theta^*}(u_i, v_i)$ della copula parametrizzata in θ^*, e i contributi vengono sommati per ottenere il valore complessivo della log-verosimiglianza

$$\ell(\tau^*) := \sum_{i=1}^{n} \log c_{\theta^*}(u_i, v_i)$$

Per illustrare in modo concreto l'implementazione dell'algoritmo MCMC nella parametrizzazione basata sul tau di Kendall, consideriamo il caso della copula di Clayton assumendo una distribuzione a priori uniforme su τ. Il prossimo esempio metterà in evidenza come la parametrizzazione in τ conduca a un'implementazione particolarmente semplice ed efficiente dell'algoritmo MCMC, pur mantenendo una chiara interpretazione del parametro di dipendenza.

Esempio 5. La relazione tra il parametro del generatore e il coefficiente di Kendall nella cupola di Clayton è esplicita e data da

$$\theta = g_{\text{Clay}}(\tau) = \frac{2\tau}{1-\tau} \tag{2.25}$$

che consente di esprimere direttamente il parametro della copula in funzione del livello di dipendenza. La densità della copula di Clayton ammette una forma chiusa, e la corrispondente log-densità può essere scritta come [1]

$$\log c_\theta(u,v) = \log(1+\theta) - (1+\theta)(\log u + \log v) - \left(2 + \frac{1}{\theta}\right)\log(u^{-\theta} + v^{-\theta} - 1) \tag{2.26}$$

espressione che viene utilizzata per il calcolo efficiente della verosimiglianza lungo la catena MCMC. Assumendo una priori uniforme su τ, la log-posteriori risulta proporzionale alla sola log-verosimiglianza e può essere scritta, a meno di una costante additiva, come

$$\log \pi(\tau \mid \mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} \log c_{g(\tau)}(u_i, v_i) + \text{const} \tag{2.27}$$

A partire dal valore corrente $\tau^{(t)}$, l'algoritmo genera poi un candidato

$$\tau^* \sim \mathcal{N}(\tau^{(t)}, \sigma^2),$$

con supporto troncato all'intervallo $(-1,1)$ per garantire l'ammissibilità del parametro. Il valore proposto viene quindi trasformato nel corrispondente parametro $\theta^* = 2\tau^*/(1-\tau^*)$, analogamente a quanto fatto per il valore corrente $\theta^{(t)} = \frac{2\tau^{(t)}}{1-\tau^{(t)}}$. Per entrambi i valori si calcola la log-verosimiglianza

$$\ell(\tau^*) = \sum_{i=1}^{n} \log c_{\theta^*}(u_i, v_i) \tag{2.28}$$

$$\ell(\tau^{(t)}) = \sum_{i=1}^{n} \log c_{\theta^{(t)}}(u_i, v_i) \tag{2.29}$$

e si costruisce il log-rapporto di accettazione

$$\log \alpha = \min\{0, \ell(\tau^*) - \ell(\tau^{(t)})\} \tag{2.30}$$

In questo caso, grazie alla simmetria della proposta e alla priori uniforme, i termini relativi alla distribuzione a priori e alla densità di proposta si cancellano, semplificando notevolmente il calcolo. Il valore candidato viene infine accettato se $\log u \leq \log \alpha$, con $u \sim$ Uniforme(0,1); in caso contrario, la catena rimane nel punto corrente.

2.4 Confronto tra parametrizzazioni

Prima di discutere l'efficienza computazionale delle due parametrizzazioni, è opportuno chiarire un punto preliminare: lavorare in termini di θ oppure in termini di τ non cambia il modello statistico, a patto di specificare distribuzioni a priori coerenti tra loro [11]. In altre parole, se la priori su θ è ottenuta come immagine della priori su τ tramite la trasformazione biunivoca $\tau = h(\theta)$, allora le due analisi devono produrre la stessa inferenza, a meno di differenze numeriche di approssimazione.

2.4.1 Coerenza delle distribuzioni a priori

Data una priori su τ, la corrispondente priori su θ si ottiene mediante la regola del cambio di variabile [21, 11]:

$$\pi(\theta) = \pi(\tau = h(\theta)) \cdot \left| \frac{d\tau}{d\theta} \right|, \tag{2.31}$$

dove $h(\theta)$ denota la relazione che lega θ al tau di Kendall. Ad esempio, nel caso della copula di Clayton, vale che

$$\tau = h(\theta) = \frac{\theta}{\theta + 2}, \qquad \theta > 0, \tag{2.32}$$

da cui

$$\frac{d\tau}{d\theta} = \frac{d}{d\theta}\left(\frac{\theta}{\theta + 2}\right) = \frac{2}{(\theta + 2)^2} \tag{2.33}$$

Assumendo una priori uniforme su $\tau \in (0,1)$, cioè $\pi(\tau) = 1$, la priori indotta su θ risulta

$$\pi(\theta) = 1 \cdot \frac{2}{(\theta + 2)^2} = \frac{2}{(\theta + 2)^2}, \qquad \theta > 0. \tag{2.34}$$

Tale priori concentra più massa su valori piccoli di θ, cioè su livelli di dipendenza deboli, coerentemente con l'idea di non favorire a priori strutture di dipendenza molto intense in assenza di informazione a priori.

2.4.2 Verifica empirica dell'equivalenza

Per verificare operativamente che le due implementazioni producano la stessa inferenza quando le priori sono rese coerenti dal cambio di variabile, consideriamo un esperimento su dati simulati, riportato in Appendice A. In particolare, fissiamo $\tau_0 = 0.5$ e generiamo un campione di dimensione $n = 200$ dalla copula di Clayton.

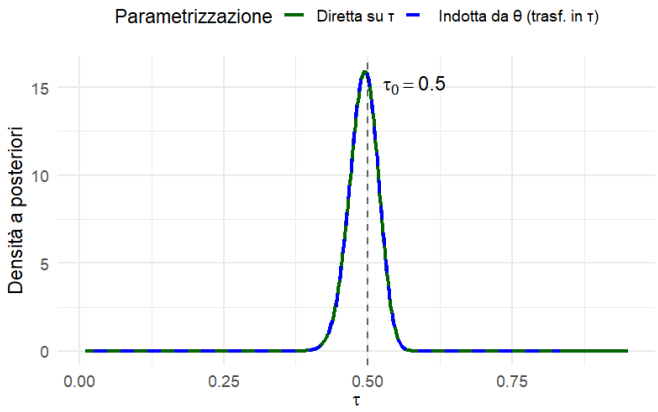


Figura 2.3: Confronto tra le densità a posteriori in scala τ ottenute dalle due parametrizzazioni con priori coerenti.

Tabella 2.3: Risultati della verifica empirica con priori coerenti per la copula di Clayton ($\tau_0 = 0.5$, $n = 200$), con statistiche espresse in scala τ.

Statistica	**Parametrizzazione su τ**	**Parametrizzazione su θ**
Media a posteriori	0.4919	0.4919
Mediana	0.4922	0.4878
Deviazione standard	0.0252	0.0252
IC 95%	[0.4395, 0.5375]	[0.4359, 0.5352]

Su questi dati confrontiamo due procedure. Per entrambe le parametrizzazioni si può lavorare in scala non vincolata. Nel caso ad esempio di $\tau \in (-1{,}1)$ si può usare la trasformazione $z = \operatorname{atanh}(\tau)$, che mappa l'intervallo in $\mathbb{R}$ e permette proposte gaussiane senza gestire vincoli al bordo [22]. Analogamente, per $\theta > 0$ si può adottare $w = \log(\theta)$, così da eliminare il vincolo e rendere più stabile il random walk. Si procede dunque facendo inferenza su τ, campionando in $z = \operatorname{atanh}(\tau)$, e poi facendo inferenza su θ, campionando in $w = \log(\theta)$ e utilizzando la priori indotta. In Tabella 2.3 sono riassunte le principali statistiche della posteriori riportate in scala τ. Dalle analisi emerge che media, deviazione standard e intervalli di credibilità risultano pressoché sovrapposti. Le piccole differenze sono attribuibili principalmente ad approssimazioni numeriche e alla diversa discretizzazione utilizzata per costruire le posteriori su griglia. Nel complesso l'evidenza empirica conferma che, una volta rese coerenti le priori tramite cambio di variabile, le due parametrizzazioni conducono alla stessa inferenza. Le differenze che emergeranno nelle sezioni successive sono quindi da interpretarsi come differenze di natura computazionale e non come divergenze sul contenuto inferenziale del modello.

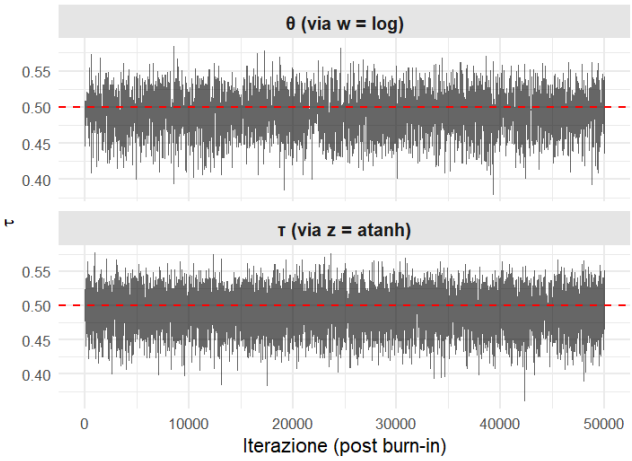


Figura 2.4: Trace plot delle catene MCMC in scala τ per le due parametrizzazioni.

Tabella 2.4: Confronto della convergenza MCMC per Clayton, $\tau_0 = 0.5$, $n = 200$.

Metrica	**Parametrizzazione su τ**	**Parametrizzazione su θ**
ESS	7 217.063	3 994.834
ESS/sec	29.811	9.472
Acceptance rate	0.214	0.126
Tempo (sec)	242.096	421.755

2.4.3 Confronto dell'efficienza computazionale

Verificata la coerenza inferenziale, confrontiamo ora l'efficienza delle due implementazioni MCMC. Il confronto è svolto sullo stesso dataset simulato precedentemente e si basa sugli indicatori riportati in Tabella 2.4. La parametrizzazione basata su τ risulta nettamente più efficiente. A parità di numero di iterazioni, l'ESS è circa raddoppiato e, soprattutto, l'indicatore ESS/sec evidenzia un guadagno netto in termini di costo computazionale. Osserviamo che anche il tasso di accettazione è più favorevole; infine differenze di efficienza emergono anche dallo studio dei grafici diagnostici. Il traceplot relativo alla parametrizzazione su τ in Figura 2.4 mostra un'oscillazione più regolare e un'esplorazione più rapida della regione ad alta densità, mentre la parametrizzazione su θ tende a produrre catene maggiormente autocorrelate, coerentemente con un tasso di accettazione più basso. Infine l'ACF risulta contenuta e decresce rapidamente in entrambi i casi, anche se più lentamente nella parametrizzazione in θ, come si nota dalla Figura 2.5.

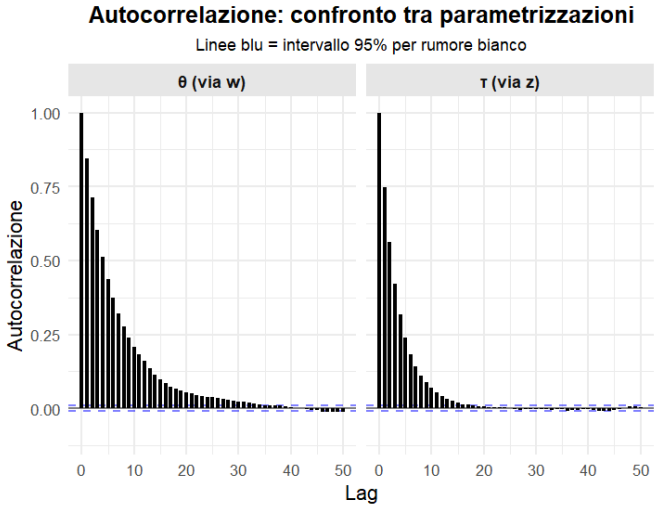


Figura 2.5: Funzione di autocorrelazione (ACF) per le due parametrizzazioni.

Interpretazione e conclusioni

Il vantaggio della parametrizzazione in τ è coerente con quanto atteso: lavorando su questo parametro, che presenta supporto standardizzato e interpretazione immediata, otteniamo una log-posteriori più regolare e più facile da esplorare con proposte gaussiane [23, 12]. La trasformazione $z = \text{atanh}(\tau)$, inoltre, elimina i vincoli di supporto e rende la scala del parametro più compatibile con un random walk di tipo normale. Al contrario, anche adottando $w = \log(\theta)$ per rimuovere il vincolo di positività, la forma della log-posteriori in θ può restare più asimmetrica, con conseguente riduzione del tasso di accettazione e dell'efficienza complessiva. In sintesi, a parità di modello statistico la parametrizzazione in τ sembra offrire un discreto vantaggio computazionale: a fronte di un tempo di esecuzione inferiore, produce un numero maggiore di campioni effettivamente informativi. Questo risultato motiva l'adozione della parametrizzazione in termini di τ per il resto di questo elaborato.

2.5 Confronto tra scelte di priori

Nella sezione precedente è stata verificata l'equivalenza tra diverse parametrizzazioni del modello di copula mediante l'uso di distribuzioni a priori coerenti, ottenute tramite un opportuno cambio di variabile. In questa sezione analizziamo l'impatto di diverse scelte di priori non equivalenti sull'inferenza bayesiana, al fine di valutare la robustezza dei risultati rispetto alle assunzioni a priori. In particolare, consideriamo tre distribuzioni a priori molto utilizzate in letteratura:

1. **Priori A:** $\tau \sim \text{Unif}(-1,1)$. La priori uniforme su τ rappresenta una scelta di riferimento naturale in assenza di informazioni a priori. Essa attribuisce la stessa probabilità a tutti i livelli di dipendenza in scala τ, pur inducendo una distribuzione non uniforme sul parametro del generatore θ.

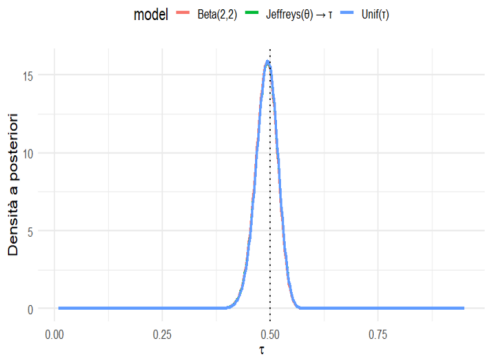
2. **Priori B:** $\pi(\theta) \propto \theta^{-1}$. Una priori di tipo Jeffreys sul parametro naturale della copula, comunemente utilizzata come priori di riferimento per parametri positivi continui [27, 21]. Il suo utilizzo consente di valutare l'effetto di una priori formulata direttamente sulla scala del generatore, in contrapposizione a priori definite sulla misura di dipendenza τ.

3. **Priori C:** $\tau \sim \text{Beta}(2,2)$ riscalata su $(-1,1)$. Essa costituisce un esempio di priori debolmente informativa su τ, simmetrica rispetto allo zero e moderatamente concentrata attorno a valori di dipendenza debole [11]. Tale scelta riflette una situazione realistica in cui si ritiene a priori improbabile una dipendenza estremamente forte, pur senza escluderla.

Le tre distribuzioni a priori analizzate in questa sezione non mirano a esaurire tutte le possibili scelte per la distribuzioni a priori, bensì a rappresentare tre approcci concettualmente distinti e frequentemente impiegati nelle applicazioni pratiche.

2.5.1 Analisi di sensibilità

Per valutare empiricamente l'influenza delle diverse priori, consideriamo un dataset simulato dalla copula di Clayton con $\tau_0 = 0.5$ e numerosità campionaria $n = 200$. Per ciascuna scelta a priori calcoliamo la distribuzione a posteriori di τ utilizzando la stessa verosimiglianza. I risultati mostrano una notevole stabilità dell'inferenza rispetto alla scelta della distribuzione a priori. Le medie a posteriori le deviazioni standard e gli intervalli di credibilità risultano quasi indistinguibili tra le tre scelte. In particolare la priori Beta (2,2), pur essendo informativa e concentrata intorno a valori di dipendenza debole, viene rapidamente superata dall'informazione contenuta nei dati. Anche la priori di Jeffreys sul parametro θ non induce differenze apprezzabili nella distribuzione a posteriori di τ, suggerendo che, per una numerosità



Priori	$\boldsymbol{E}[\boldsymbol{\tau}$ | data$]$	**SD**	**IC 95%**
Unif	0.492	0.025	[0.440, 0.539]
Jeffreys	0.492	0.025	[0.436, 0.536]
Beta(2,2)	0.491	0.025	[0.438, 0.537]

Figura 2.6: Confronto tra le densità a posteriori in scala τ

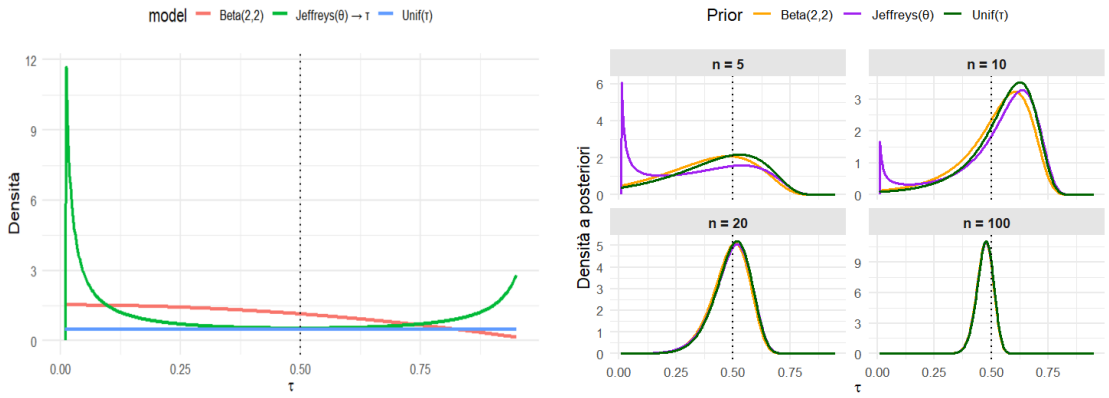


Figura 2.7: Evoluzione delle posteriori al variare della dimensione campionaria n.

campionaria pari a $n = 200$, l'inferenza è dominata dalla verosimiglianza. L'analisi grafica delle posteriori in Figura 2.6 mostra che, nonostante le differenze sostanziali tra le priori, le distribuzioni a posteriori convergono verso una forma molto simile. Questo comportamento è coerente con le proprietà di consistenza dell'inferenza bayesiana: all'aumentare della numerosità campionaria, l'influenza della priori diventa progressivamente trascurabile. La Figura 2.7 mostra l'evoluzione delle distribuzioni a posteriori per le tre priori al crescere di n. Per n piccoli le tre curve presentano differenze apprezzabili sia nella posizione che nella forma. Al crescere di n, le tre curve convergono progressivamente, fino a diventare praticamente indistinguibili per $n = 200$. Nel contesto dell'analisi empirica presentata nel Capitolo 3, in cui si dispone di circa $n \approx 5000$ osservazioni giornaliere, questi risultati indicano che la scelta della priori esercita un'influenza sostanzialmente trascurabile sull'inferenza. Tuttavia, qualora si intendesse analizzare sotto-periodi di breve durata, la sensibilità rispetto alla a priori dovrebbe essere valutata con maggiore attenzione.

2.6 Verifica con simulazioni Monte Carlo

Prima di applicare l'inferenza bayesiana a dati reali, è opportuno verificare il corretto funzionamento dell'algoritmo attraverso uno studio di simulazione Monte Carlo. Questo tipo di analisi permette non solo di controllare la correttezza dell'implementazione, ma anche di valutare l'accuratezza delle stime, la copertura degli intervalli di credibilità e l'efficienza computazionale, al variare del livello di dipendenza e della dimensione campionaria [21].

2.6.1 Schema generale di simulazione

Algorithm 4 Studio di Simulazione Monte Carlo

Input: Famiglia di copula, valore vero τ_0, dimensione campionaria n, numero replicazioni R
for $r = 1$ to R **do**
    Converti $\tau_0 \to \theta_0 = g(\tau_0)$
    Simula dati: $(u_i, v_i) \sim C_{\theta_0}$ per $i = 1, \ldots, n$
    Esegui MCMC per ottenere campioni dalla posteriori: $\{\tau^{(j)}\}_{j=1}^{J}$
    Calcola statistiche riassuntive:
        - Media a posteriori: $\hat{\tau}_r = \frac{1}{J}\sum_{j=1}^{J} \tau^{(j)}$
        - Deviazione standard posteriori: $\hat{\sigma}_{\tau,r}$
        - Intervallo di credibilità 95%: $[\hat{\tau}_{0.025,r}, \hat{\tau}_{0.975,r}]$
        - ESS, tasso di accettazione
    Verifica copertura: $I_r = \mathbb{I}(\tau_0 \in [\hat{\tau}_{0.025,r}, \hat{\tau}_{0.975,r}])$
end for
Calcola metriche aggregate su R replicazioni:
- Bias $= \frac{1}{R}\sum_{r=1}^{R}(\hat{\tau}_r - \tau_0)$
- RMSE $= \sqrt{\frac{1}{R}\sum_{r=1}^{R}(\hat{\tau}_r - \tau_0)^2}$
- Copertura empirica: Cov $= \frac{1}{R}\sum_{r=1}^{R} I_r$
Output: Bias, RMSE, Copertura, ESS medio, tasso accettazione medio

Dopo aver descritto l'algoritmo generale, possiamo entrare nel dettaglio del disegno dello studio, chiarendo come le simulazioni vengono condotte e quali misure vengono raccolte per valutare le prestazioni dell'inferenza.

Parametro	**Valori**
Dimensione campionaria n	$\{50, 100, 200\}$
Livello di dipendenza τ_0	$\{0.3, 0.5, 0.7\}$
Numero di replicazioni R	200 per ciascuna combinazione (n, τ_0)
Algoritmo MCMC	10 000 iterazioni totali
Burn-in	2 000 iterazioni
Campioni utilizzati	8 000

Tabella 2.5: Configurazioni sperimentali dello studio di simulazione.

2.6.2 Disegno dello studio

Lo studio è condotto considerando la copula di Clayton e adottando la parametrizzazione in termini del coefficiente di Kendall τ, con distribuzione a priori uniforme. Per ciascuna combinazione di dimensione campionaria n e valore vero τ_0, si effettuano R replicazioni indipendenti, in ognuna delle quali:

1. si genera un campione di pseudo-osservazioni $(u_i, v_i)_{i=1}^n$ dalla copula di Clayton con parametro $\theta_0 = g(\tau_0) = \frac{2\tau_0}{1-\tau_0}$;

2. sui dati simulati si applica un algoritmo Random Walk Metropolis-Hastings per campionare dalla distribuzione a posteriori $\pi(\tau \mid \mathbf{u}, \mathbf{v})$;

3. a partire dal campione MCMC post burn-in si calcolano la media a posteriori $\hat{\tau}_r$, l'intervallo di credibilità al 95% e gli indicatori di efficienza della catena, come ESS e tasso di accettazione;

4. si registra l'indicatore di copertura $I_r = \mathbb{I}\{\tau_0 \in [\hat{\tau}_{0.025,r}, \hat{\tau}_{0.975,r}]\}$.

Al termine delle R repliche, le statistiche di interesse vengono poi aggregate e riassunte tramite le seguenti quantità:

$$\text{Bias} = \frac{1}{R}\sum_{r=1}^{R}(\hat{\tau}_r - \tau_0), \qquad \text{RMSE} = \sqrt{\frac{1}{R}\sum_{r=1}^{R}(\hat{\tau}_r - \tau_0)^2}, \qquad \text{Coverage} = \frac{1}{R}\sum_{r=1}^{R} I_r.$$

Inoltre, vengono riportati l'ESS medio e il tasso di accettazione medio calcolati sulle repliche, utilizzati come indicatori della qualità della convergenza delle catene. Nel codice, riportato in Appendice A, l'implementazione è stata parallelizzata su 7 core, al fine di ridurre il tempo complessivo di esecuzione dello studio. Le configurazioni sperimentali, inclusi i valori dei parametri, il numero di replicazioni e le dimensioni dei campioni utilizzati sono riportati in Tabella 2.5.

Catene multiple e parallelizzazione

Al fine di migliorare l'affidabilità delle diagnosi di convergenza e sfruttare le risorse computazionali disponibili, è buona pratica eseguire più catene MCMC in parallelo, inizializzate in punti diversi dello spazio dei parametri [22]. In questo contesto, si considerano tipicamente inizializzazioni che rappresentano diversi livelli di dipendenza, come l'indipendenza, una dipendenza positiva o negativa moderata, e lo stimatore campionario di Kendall. Le catene vengono eseguite indipendentemente e, una volta verificata la convergenza tramite strumenti diagnostici standard, i campioni possono essere combinati per ottenere una stima più stabile delle quantità di interesse. Questa strategia consente inoltre di individuare eventuali problemi di multimodalità o di scarsa esplorazione dello spazio parametrico.

Proprietà frequentiste della stima bayesiana

Sebbene l'inferenza sia formulata in un'ottica bayesiana, è utile analizzare anche le sue proprietà frequentiste in esperimenti ripetuti. In particolare, sotto opportune condizioni di regolarità [28], la media a posteriori risulta consistente, nel senso che al crescere della numerosità campionaria converge in probabilità al vero valore τ_0.

Teorema 8. *La media a posteriori $\hat{\tau} = \mathbb{E}[\tau \mid \mathbf{u}, \mathbf{v}]$ converge in probabilità al vero valore τ_0 sotto opportune ipotesi di regolarità per $n \to \infty$:*

$$\hat{\tau} \xrightarrow{P} \tau_0 \tag{2.35}$$

Inoltre, sempre in condizioni di regolarità, l'errore di stima opportunamente riscalato soddisfa una proprietà di normalità asintotica [28].

Teorema 9. *Sotto regolarità, per $n \to \infty$:*

$$\sqrt{n}(\hat{\tau} - \tau_0) \xrightarrow{d} \mathcal{N}(0, V(\tau_0)) \tag{2.36}$$

dove $V(\tau_0)$ è la varianza asintotica.

Dal punto di vista applicativo, ciò implica che il bias della stima tende a diventare trascurabile per campioni di grande dimensione, che l'RMSE decresce con ordine $O(n^{-1/2})$ e che gli intervalli di credibilità al 95% dovrebbero includere il valore vero in circa il 95% delle replicazioni.

Tabella 2.6: Studio Monte Carlo per la copula di Clayton ($R = 200$ replicazioni)

τ_0	n	Bias	RMSE	Coverage	ESS	Acc. Rate
0.3	50	-0.006	0.065	0.945	1125	0.236
0.3	100	-0.008	0.047	0.945	1123	0.229
0.3	200	0.001	0.034	0.940	1089	0.252
0.5	50	-0.010	0.049	0.955	1056	0.226
0.5	100	-0.004	0.034	0.935	1134	0.232
0.5	200	-0.001	0.023	0.960	1006	0.218
0.7	50	-0.004	0.028	0.970	1107	0.214
0.7	100	-0.001	0.023	0.950	1104	0.231
0.7	200	-0.001	0.016	0.935	1028	0.232

2.6.3 Risultati numerici

La Tabella 2.6 riporta i risultati per tutti gli scenari considerati. L'RMSE mostra una riduzione sistematica all'aumentare della numerosità campionaria, coerentemente con quanto atteso dal punto di vista teorico (Teorema 9). Osserviamo inoltre che, a parità di dimensione campionaria, l'RMSE tende a ridursi all'aumentare di τ_0. Questo comportamento è coerente con il fatto che, in presenza di una dipendenza più marcata, la distribuzione a posteriori risulta più concentrata. Relativamente al bias, in tutti gli scenari analizzati, risulta di entità molto contenuta. Questo indica che la media a posteriori fornisce una stima ben centrata attorno al valore vero anche per campioni di dimensione moderata. La copertura empirica degli intervalli di credibilità al 95% varia tra 0.935 e 0.970, risultando complessivamente in buon accordo con il livello nominale. Va tuttavia ricordato che, con $R = 200$ replicazioni, la variabilità Monte Carlo della copertura non è trascurabile, ed è dovuta al numero limitato di repliche. Infine possiamo notare che l'ESS medio si colloca intorno a 10^3 in tutti gli scenari, a fronte di 8 000 iterazioni post burn-in. Il tasso di accettazione varia tra 0.214 e 0.252, un intervallo coerente con quanto atteso per un algoritmo Random Walk Metropolis in dimensione uno, dove valori prossimi a 0.23 sono noti essere vicini all'ottimo teorico in molti contesti [15, 22]. Nel complesso, questi indicatori suggeriscono una buona efficienza delle catene e una scelta accurata della distribuzione di proposta.

Conclusioni

Lo studio Monte Carlo conferma che l'implementazione MCMC per la copula di Clayton nella parametrizzazione in termini di τ produce un'inferenza affidabile. In particolare si osservano bias trascurabile, RMSE decrescente all'aumentare di n, copertura empirica coerente con il livello nominale e indicatori MCMC compatibili con un'esplorazione efficiente della distribuzione a posteriori. Queste evidenze supportano l'utilizzo dell'algoritmo nel capitolo applicativo successivo.

2.7 Un confronto con approcci frequentisti

In questa sezione confrontiamo l'inferenza bayesiana basata sulla parametrizzazione in τ con i principali approcci frequentisti comunemente utilizzati per l'analisi dei modelli di copula. L'obiettivo è valutare differenze e punti di contatto in termini di accuratezza, interpretabilità e costo computazionale.

2.7.1 Metodi frequentisti per copule

Lo stimatore più semplice è basato sul tau di Kendall campionario [13, 1]:

$$\hat{\tau}_n = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \text{sign}[(u_i - u_j)(v_i - v_j)] \tag{2.37}$$

Una volta ottenuto $\hat{\tau}_n$, il parametro della copula viene stimato tramite inversione della trasformazione $\theta = g(\tau)$, ovvero

$$\hat{\theta} = g(\hat{\tau}_n) \tag{2.38}$$

Questo approccio risulta efficiente dal punto di vista computazionale, poiché richiede soltanto il calcolo dei ranghi e non necessita di alcuna ottimizzazione numerica. Tuttavia, lo stimatore ottenuto presenta una varianza generalmente più elevata rispetto allo stimatore di massima verosimiglianza [8, 6]. Inoltre, la costruzione di errori standard e intervalli di confidenza affidabili risulta meno immediata, non permettendo neanche inferenza su altre quantità di interesse.

Massima verosimiglianza (ML)

Lo stimatore di massima verosimiglianza è ottenuto massimizzando la log- verosimiglianza della copula:

$$\hat{\theta}_{\text{ML}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c_\theta(u_i, v_i) \tag{2.39}$$

Dal punto di vista teorico, lo stimatore ML è asintoticamente efficiente e gode di una teoria ben consolidata [6], che consente di ottenere errori standard a partire dalla matrice di informazione di Fisher. Dal punto di vista pratico, tuttavia, la procedura può risultare onerosa: la funzione di verosimiglianza può presentare multimodalità, regioni piatte o problemi di instabilità numerica, soprattutto per copule più complesse o in campioni di piccola dimensione [8, 6]. Inoltre, gli intervalli di confidenza basati su approssimazioni asintotiche possono risultare poco accurati quando n non è sufficientemente grande [6].

Tabella 2.7: Confronto tra i principali metodi di stima per la copula di Clayton

Metodo	Bias(τ)	RMSE(τ)	Copertura 95%	Tempo (sec)
Kendall Campionario	0.001	0.037	–	0.01
ML (in θ)	0.012	0.030	0.878	0.18
Bayes (in τ, Unif)	-0.001	0.024	0.940	56.5

2.7.2 Confronto numerico

Per rendere il confronto più concreto, riportiamo un'analisi basata su simulazioni Monte Carlo per la copula di Clayton con $\tau_0 = 0.5$ e $n = 200$. I risultati sono riportati nella Tabella 2.7, ottenuti su $R = 500$ replicazioni Monte Carlo indipendenti, eseguite in parallelo su 7 core per ridurre il tempo di calcolo complessivo. In tabella il tempo computazionale è riportato come media per singola replica.

Accuratezza delle stime, intervalli di confidenza ed RMSE

Dal punto di vista delle stime puntuali, tutti i metodi mostrano un comportamento complessivamente soddisfacente. Lo stimatore basato sul coefficiente di Kendall e l'approccio bayesiano producono stime praticamente non distorte, con valori di bias prossimi allo zero. La massima verosimiglianza presenta un bias leggermente superiore, che rimane comunque contenuto, ma suggerisce minore stabilità in caso di campioni con dimensione moderata. Per quanto riguarda l'errore quadratico medio lo stimatore di Kendall risulta il meno efficiente, come atteso, poiché non sfrutta completamente l'informazione contenuta nella verosimiglianza, mentre l'approccio bayesiano fornisce il risultato migliore. Un aspetto di particolare rilevanza pratica riguarda la qualità degli intervalli di incertezza. Gli intervalli di confidenza ottenuti tramite la massima verosimiglianza mostrano una sottocopertura rispetto al livello nominale del 95%. In termini concreti, in circa il 12% delle replicazioni l'intervallo di confidenza non contiene il vero valore del parametro. Questo fenomeno è coerente con il fatto che, in campioni moderati, intervalli basati su approssimazioni asintotiche possano risultare meno accurati [6, 11]. Dal punto di vista applicativo, tale sottocopertura è problematica, poiché porta a sottostimare l'incertezza associata alla stimatore. In contesti come la gestione del rischio finanziario, questo può tradursi in decisioni eccessivamente ottimistiche [7]. Al contrario, gli intervalli di credibilità bayesiani mantengono una copertura molto vicina al livello nominale, pari al 94.0%. Questo valore, compatibile con il 95% atteso, indica che la distribuzione a posteriori fornisce una rappresentazione più fedele dell'incertezza reale, rappresentando uno dei principali vantaggi dell'approccio bayesiano in questo contesto.

Costo computazionale e conclusioni

Dal punto di vista computazionale emerge nuovamente il compromesso tra accuratezza e costo di calcolo. Lo stimatore basato sul tau di Kendall risulta estremamente rapido da calcolare, mentre la stima per massima verosimiglianza comporta un costo moderato, legato alla necessità di risolvere un problema di ottimizzazione numerica. L'approccio bayesiano, invece, è sensibilmente più oneroso, poiché richiede la generazione di un elevato numero di campioni per gli algoritmi MCMC. Nel complesso, questi risultati mostrano che, pur a fronte di un maggiore costo computazionale, l'inferenza bayesiana fornisce stime più efficienti e, soprattutto, intervalli di incertezza meglio calibrati rispetto alla massima verosimiglianza per campioni di dimensioni moderate. Questo vantaggio qualitativo risulta rilevante nelle applicazioni in cui una corretta quantificazione dell'incertezza è cruciale e giustifica l'utilizzo di metodi bayesiani nonostante il maggiore costo computazionale.

2.7.3 Propagazione dell'incertezza

Uno dei principali punti di forza dell'inferenza bayesiana risiede nella possibilità di descrivere in modo completo l'incertezza associata alla stima del parametro di dipendenza. L'oggetto centrale dell'analisi non è infatti una singola stima puntuale, ma l'intera distribuzione a posteriori di τ, che riassume in modo naturale l'informazione contenuta nei dati e nella priori. Questo consente di costruire intervalli di credibilità con un'interpretazione probabilistica diretta e di individuare regioni HPD (Highest Posterior Density) [11, 21] che concentrano la massa di probabilità nelle zone più plausibili dello spazio parametrico. Un ulteriore vantaggio è che l'incertezza su τ può essere propagata senza difficoltà a qualsiasi quantità derivata di interesse. Un esempio particolarmente rilevante, come vedremo nel Capitolo 3, riguarda la dipendenza nelle code. Nel caso della copula di Gumbel, la dipendenza nella coda superiore è una funzione esplicita del coefficiente di Kendall τ e può essere espressa come [1]

$$\lambda_U(\tau) = 2 - 2^{1/(1/(1-\tau))} = 2 - 2^{1-\tau} \tag{2.40}$$

Una volta ottenuto un campione MCMC dalla distribuzione a posteriori di τ, la distribuzione a posteriori del coefficiente di tail dependence λ_U si ottiene semplicemente applicando la trasformazione 2.40 a ciascun campione [11]. Lo stesso principio vale in generale per qualsiasi funzione $g(\tau)$ di interesse, la cui distribuzione a posteriori è ottenuta semplicemente trasformando i campioni MCMC, permettendo di svolgere inferenza diretta su quantità che spesso hanno un'interpretazione più immediata del parametro stesso. In questo modo è possibile, ad esempio, valutare la probabilità a posteriori che la dipendenza sia positiva, stimare la probabilità che essa superi una soglia considerata elevata, oppure costruire intervalli di credibilità

per altre misure di associazione, come il coefficiente di Spearman. Analogamente, in applicazioni finanziarie, l'incertezza su τ può essere propagata a misure di rischio congiunte come il Value-at-Risk o l'Expected Shortfall, ottenendo una valutazione più realistica del rischio complessivo [7].

Conclusioni del capitolo

In questo capitolo è stato sviluppato l'approccio bayesiano all'inferenza per copule Archimedee, ponendo l'attenzione sulla parametrizzazione basata sul coefficiente di Kendall τ. Dopo aver messo in luce alcune criticità della parametrizzazione tradizionale in termini del parametro del generatore θ, si è mostrato come la riparametrizzazione in τ consenta di lavorare su un dominio standardizzato e direttamente interpretabile, migliorando al tempo stesso il comportamento numerico delle procedure di stima. L'algoritmo Metropolis-Hastings è stato quindi descritto e adattato a questa nuova parametrizzazione, discutendo in modo dettagliato le scelte operative più rilevanti. Attraverso uno studio di simulazione Monte Carlo, l'approccio è stato validato sia dal punto di vista bayesiano sia sotto il profilo delle sue proprietà frequentiste, evidenziando una buona accuratezza delle stime e degli intervalli di confidenza. Il confronto con i principali metodi frequentisti ha mostrato infine come l'inferenza bayesiana, pur comportando un costo computazionale più elevato, offra una descrizione più completa dell'incertezza e una maggiore flessibilità, soprattutto in campioni di dimensione moderata o in presenza di strutture modellistiche più complesse. Nel capitolo successivo l'attenzione si sposterà sull'analisi empirica della dipendenza tra i rendimenti dell'indice azionario tedesco DAX e dell'indice statunitense S&P 500. L'obiettivo sarà stimare e interpretare il livello di dipendenza attraverso τ, confrontare famiglie diverse di copule Archimedee e valutare, in chiave applicativa, cosa suggeriscono i dati riguardo al comportamento congiunto dei due mercati.

Capitolo 3

Analisi empirica

In questo capitolo i metodi di inferenza bayesiana per copule Archimedee, sviluppati nel Capitolo 2, vengono applicati a dati finanziari reali. L'obiettivo principale è valutare il comportamento della parametrizzazione basata sul coefficiente τ di Kendall quando viene utilizzata su serie storiche reali, caratterizzate da rumore, dipendenza non lineare e possibili fenomeni estremi. In questo senso, l'analisi empirica svolge un duplice ruolo: da un lato permette di validare l'approccio bayesiano dal punto di vista operativo, dall'altro offre l'opportunità di studiare in modo approfondito la struttura di dipendenza tra mercati finanziari.

3.1 Obiettivo dell'analisi

Il caso di studio considerato riguarda la relazione tra l'indice azionario tedesco DAX [29, 30] e l'indice statunitense S&P 500 [31]. Questa scelta è motivata dalla rilevanza economica dei due mercati e dal forte interesse, documentato in letteratura, per la loro interdipendenza [2, 3]. Considerati insieme, essi coprono una quota rilevante del mercato azionario globale e intrattengono forti relazioni commerciali e finanziarie. Questa integrazione suggerisce l'esistenza di una dipendenza significativa tra i due mercati, pur in assenza di una correlazione [3, 2]. Analizzare tale relazione consente quindi di studiare un caso realistico in cui i mercati sono interconnessi ma mantengono caratteristiche strutturali distinte. L'attenzione sarà rivolta anche alla capacità delle diverse famiglie di copule Archimedee di catturare eventuali asimmetrie e co-movimenti estremi. Un ulteriore elemento di interesse è rappresentato dall'evoluzione temporale della dipendenza, che può variare sensibilmente nel corso del tempo e intensificarsi in corrispondenza di fasi di crisi o di elevata volatilità. Nel complesso, questo capitolo intende mostrare come gli strumenti teorici e computazionali introdotti nei capitoli precedenti possano essere utilizzati in modo efficace per analizzare problemi concreti di interesse economico-finanziario.

3.2 Descrizione del dataset

L'analisi empirica si basa sui prezzi di chiusura giornalieri di due indici azionari di riferimento a livello internazionale: il DAX (Deutscher Aktienindex), che rappresenta l'andamento delle principali società quotate alla Borsa di Francoforte, e l'S&P 500 (Standard & Poor's 500), che sintetizza la performance del mercato azionario statunitense attraverso un ampio paniere di titoli quotati al NYSE e al NASDAQ. Il periodo considerato va dal 1 gennaio 2005 al 31 dicembre 2024, coprendo un orizzonte temporale sufficientemente ampio da includere fasi di mercato molto diverse tra loro. All'interno di questa finestra temporale si alternano infatti periodi di relativa stabilità a episodi di forte turbolenza, tra cui la crisi finanziaria globale del 2008–2009, ampiamente documentata nei Global Financial Stability Reports del Fondo Monetario Internazionale [32], la crisi del debito sovrano europeo del periodo 2010-2012, analizzata in dettaglio nei Financial Stability Reviews della Banca Centrale Europea [33], e lo shock sistemico legato alla pandemia da COVID-19, che ha generato una recessione globale sincronizzata nel 2020 [34, 35]; questi rilevanti episodi di stress finanziario degli ultimi decenni rendono il campione particolarmente informativo per lo studio della dipendenza, soprattutto nelle code della distribuzione congiunta. I dati sono stati raccolti tramite la piattaforma Yahoo Finance [36, 37], ampiamente utilizzata nella letteratura empirica per analisi finanziarie di questo tipo.

Pulizia del dataset

Prima di procedere con l'analisi, i dati sono stati sottoposti a una fase preliminare di pulizia e allineamento. Sono state considerate esclusivamente le giornate di contrattazione comuni ai due mercati, escludendo automaticamente weekend, festività nazionali e chiusure straordinarie, così da garantire una perfetta sincronizzazione temporale delle serie. Dopo questo processo di allineamento, il campione finale risulta composto da $n = 4949$ osservazioni giornaliere. È stata inoltre verificata l'assenza di valori mancanti nel dataset finale. Non è stata effettuata alcuna rimozione di osservazioni estreme: i movimenti di prezzo di grande entità, tipicamente osservati durante le fasi di crisi, costituiscono infatti un elemento informativo centrale per l'analisi della dipendenza nelle code e risultano pienamente coerenti con gli obiettivi dello studio.

3.2.1 Statistiche descrittive dei prezzi

La Tabella 3.1 riporta alcune statistiche descrittive calcolate sui livelli dei prezzi di chiusura giornalieri del DAX e dell'S&P 500 nel periodo considerato. Sebbene l'analisi della dipendenza verrà condotta sui rendimenti, un esame preliminare

Tabella 3.1: Statistiche descrittive dei prezzi

Statistica	DAX	S&P 500
Media	10,002.97	2,353.30
Mediana	9,710.70	1,983.53
Deviazione Standard	3,912.15	1,277.11
Minimo	3,666.41	676.53
Massimo	20,426.27	6,090.27
Curtosi	2.21	2.88

dei livelli di prezzo è utile per inquadrare l'evoluzione complessiva delle due serie e per evidenziare eventuali differenze strutturali. Entrambe le serie presentano valori medi superiori alle rispettive mediane, riflettendo la presenza di un marcato trend crescente nel periodo analizzato. Tale comportamento è coerente con la crescita di lungo periodo dei mercati azionari, interrotta solo temporaneamente durante le principali fasi di crisi finanziaria. La deviazione standard risulta elevata per entrambi gli indici, con valori più alti per il DAX, indicando una maggiore variabilità dei livelli di prezzo rispetto all'S&P 500. Questo risultato è coerente con una maggiore esposizione del mercato europeo alle fasi cicliche dell'economia globale [3, 2]. La curtosi leggermente positiva indica invece distribuzioni con una maggiore concentrazione delle osservazioni attorno alla media e code più spesse rispetto alla distribuzione normale. Anche in questo caso, tali caratteristiche riflettono principalmente la presenza di trend e di variazioni di ampia entità nei livelli di prezzo e non forniscono indicazioni dirette sulla struttura di dipendenza tra i mercati, che verrà analizzata in modo appropriato sui rendimenti.

3.2.2 Visualizzazione delle serie temporali

La Figura 3.1 mostra l'evoluzione nel tempo dei prezzi di chiusura del DAX e dell'S&P 500 nel periodo 2005–2024. Già da una prima osservazione emerge un chiaro co-movimento di lungo periodo tra i due indici, indicativo di una dipendenza positiva persistente, pur a fronte di differenze nella dinamica di breve periodo e nell'intensità delle fluttuazioni. Entrambi gli indici registrano un marcato ridimensionamento durante la crisi finanziaria globale del 2008–2009, con perdite di ampiezza comparabile. Questo episodio rappresenta uno dei primi segnali evidenti di una forte interdipendenza nei periodi di stress, quando i benefici della diversificazione tendono a ridursi drasticamente [2]. La fase di recupero successiva è caratterizzata da una crescita sostenuta, ma accompagnata da un'elevata volatilità, in particolare per il mercato europeo. Nel corso della crisi del debito sovrano del 2011–2012, l'andamento dei due indici diverge parzialmente: il DAX risulta più

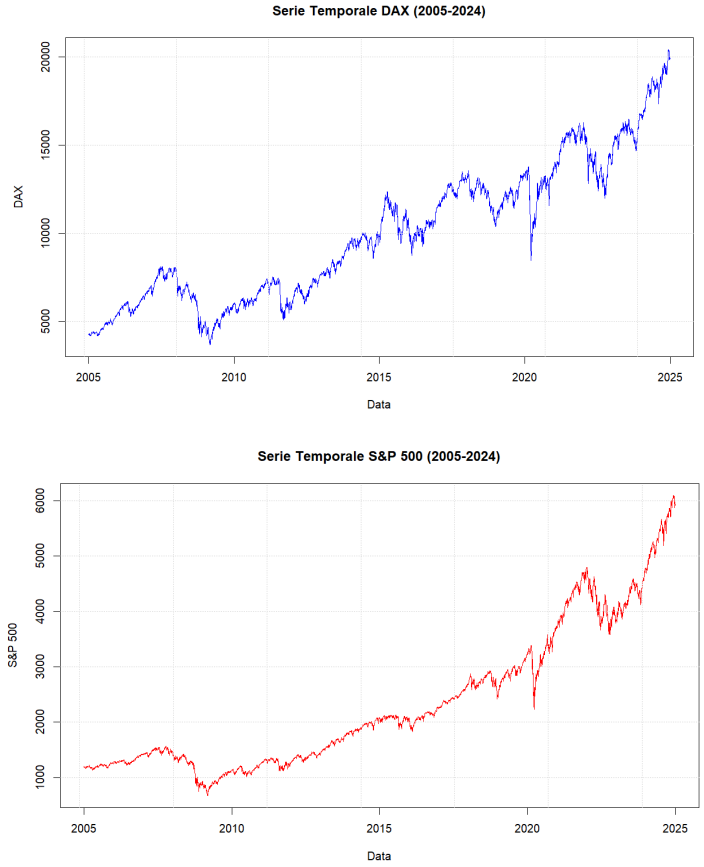


Figura 3.1: Serie temporali dei prezzi di chiusura di DAX e S&P 500 (2005-2024).

penalizzato rispetto all'S&P 500, riflettendo l'impatto asimmetrico della crisi sui mercati dell'area euro. Tra il 2012 e il 2019, invece, entrambi gli indici attraversano una lunga fase di mercato rialzista relativamente stabile, interrotta solo da episodi di correzione di breve durata. Il crollo improvviso osservato nel primo trimestre del 2020, in concomitanza con la pandemia da COVID-19, evidenzia nuovamente un forte co-movimento tra i mercati. Negli anni più recenti, infine, la crescita prosegue in un contesto di maggiore incertezza macroeconomica, con fasi di correzione più frequenti e una volatilità più pronunciata. Nel complesso, queste evidenze grafiche suggeriscono che la relazione tra i due mercati non possa essere adeguatamente descritta dalla sola correlazione lineare. La presenza di co-movimenti accentuati durante le fasi di crisi motiva l'adozione di modelli di copula, che permettono di analizzare in modo più flessibile la struttura di dipendenza e, in particolare, il comportamento congiunto nelle code della distribuzione.

3.3 I rendimenti logaritmici

I prezzi degli asset finanziari sono notoriamente processi non stazionari, caratterizzati da trend di lungo periodo e da una varianza che tende a crescere nel tempo. Per questo motivo, nella modellizzazione della dipendenza tramite copule, si preferisce lavorare non sui livelli di prezzo, ma sui rendimenti. In particolare, in questo elaborato utilizziamo i rendimenti logaritmici, definiti come

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(P_t) - \log(P_{t-1}), \tag{3.1}$$

dove P_t indica il prezzo di chiusura dell'asset al tempo t. I rendimenti, a differenza dei prezzi, risultano generalmente stazionari rendendo più appropriata l'applicazione di modelli statistici [38]. Inoltre, i rendimenti logaritmici godono di una proprietà di additività temporale: il rendimento cumulato su più periodi può essere espresso come somma dei rendimenti giornalieri, semplificando l'analisi su orizzonti temporali più lunghi. Sulla base dei prezzi di chiusura giornalieri, i rendimenti logaritmici vengono calcolati separatamente per i due indici come

$$r_t^{\text{DAX}} = \log(P_t^{\text{DAX}}) - \log(P_{t-1}^{\text{DAX}}) \qquad r_t^{\text{S\&P}} = \log(P_t^{\text{S\&P}}) - \log(P_{t-1}^{\text{S\&P}}) \tag{3.2}$$

La trasformazione comporta la perdita della prima osservazione per ciascuna serie, producendo un campione finale di $n - 1 = 4948$ rendimenti giornalieri per entrambi gli indici. Queste serie costituiranno la base per l'analisi descrittiva e per la successiva modellizzazione della struttura di dipendenza tramite copule.

3.3.1 Statistiche descrittive dei rendimenti

La Tabella 3.2 riporta le principali statistiche descrittive dei rendimenti logaritmici giornalieri del DAX e dell'S&P 500 nel periodo considerato. Entrambi gli indici presentano un rendimento medio giornaliero positivo, coerente con l'esistenza di un premio per il rischio azionario nel lungo periodo. La mediana risulta superiore alla media in entrambi i casi, suggerendo che la distribuzione dei rendimenti è influenzata da episodi estremi negativi che abbassano il valore medio complessivo. Dal punto di vista della volatilità, il DAX mostra una deviazione standard leggermente più elevata rispetto all'S&P 500, indicando una maggiore variabilità giornaliera dei rendimenti del mercato azionario tedesco, coerentemente con l'evidenza empirica secondo cui i mercati europei tendono a essere più volatili rispetto a quello statunitense, soprattutto durante fasi di stress finanziario [39]. Entrambe le distribuzioni presentano asimmetria negativa, più evidente per l'S&P 500, indicando che le grandi perdite tendono a essere più frequenti o più intense rispetto ai grandi guadagni. Il valore molto elevato della curtosi in eccesso segnala la presenza di code pesanti e di una probabilità non trascurabile di eventi estremi, ben superiore a quella prevista da una distribuzione normale.

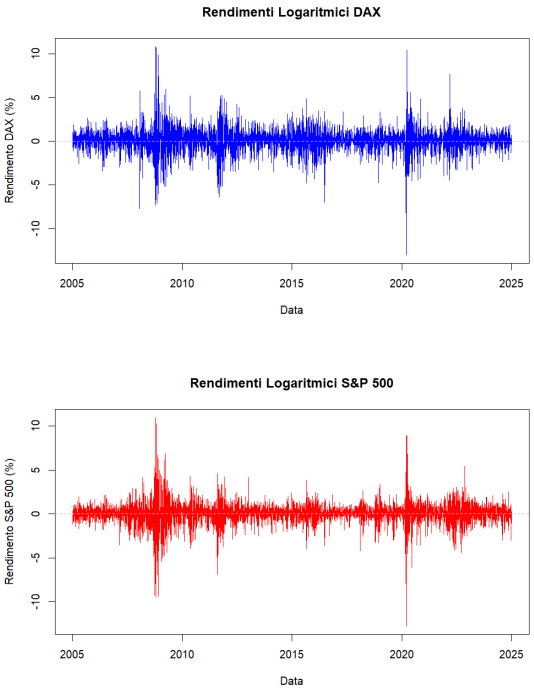


Tabella 3.2: Statistiche descrittive dei rendimenti logaritmici giornalieri

Statistica	DAX	S&P 500
Media (%)	0.0310	0.0322
Mediana (%)	0.0901	0.0710
Dev. Standard (%)	1.3246	1.2179
Minimo (%)	-13.0549	-12.7652
Massimo (%)	10.7975	10.9572
Asimmetria	-0.2434	-0.5248
Curtosi (in eccesso)	11.3247	15.9069

Figura 3.2: Serie temporali dei log-rendimenti giornalieri di DAX e S&P500.

3.3.2 Visualizzazione dei rendimenti

La Figura 3.2 mostra le serie dei rendimenti logaritmici. Dall'osservazione delle serie temporali emergono alcuni pattern ricorrenti tipici dei rendimenti finanziari. In primo luogo, si notano fasi caratterizzate da alta volatilità, come avviene durante la crisi finanziaria del 2008 e nel marzo 2020. Questo andamento indica che la varianza dei rendimenti non è costante nel tempo. Un secondo elemento rilevante riguarda la simultaneità degli eventi estremi. Nei periodi di crisi, entrambi i mercati mostrano movimenti di grande ampiezza nello stesso intervallo temporale, suggerendo una forte dipendenza nelle code della distribuzione congiunta. Tale evidenza è particolarmente importante dal punto di vista della gestione del rischio, poiché indica che le perdite tendono a manifestarsi contemporaneamente proprio nelle fasi più critiche. Nel complesso, le serie mostrano un comportamento eteroschedastico: la volatilità si intensifica nei periodi di stress finanziario e si riduce nelle fasi di relativa stabilità. Queste caratteristiche motivano l'adozione di modelli flessibili, come le copule, in grado di catturare strutture di dipendenza non lineari.

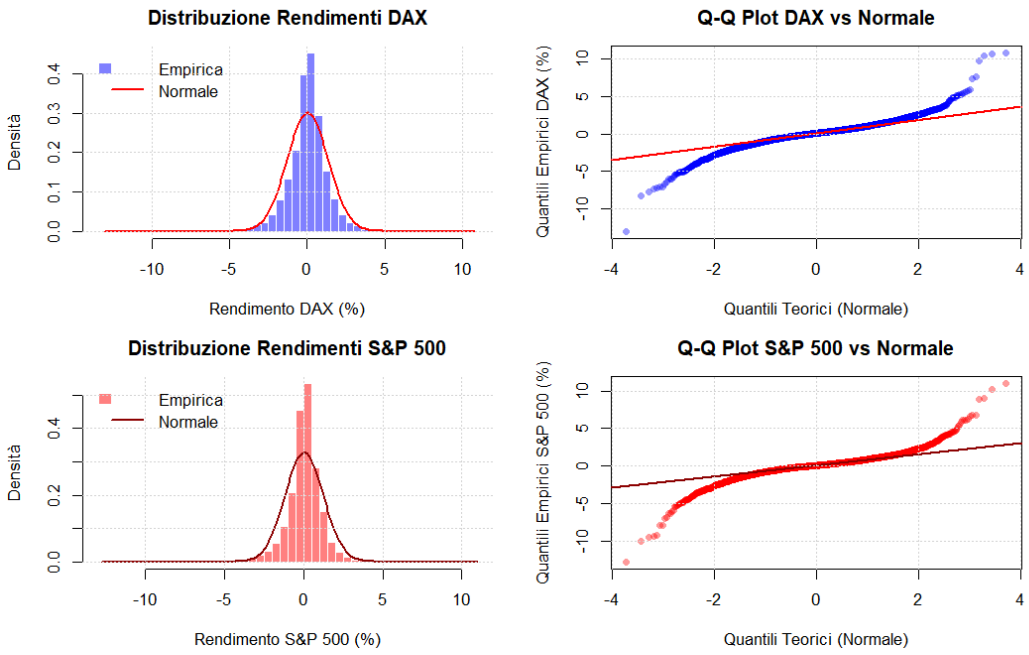


Figura 3.3: Distribuzione dei rendimenti: istogrammi con sovrapposizione della densità normale stimata e Q-Q plot rispetto alla distribuzione normale.

3.3.3 Istogrammi, Q–Q plot e scatterplot

La Figura 3.3 mette a confronto la distribuzione empirica dei rendimenti logaritmici con quella di una distribuzione normale avente stessa media e varianza. Gli istogrammi, accompagnati dalla densità normale sovrapposta, mostrano già visivamente alcune deviazioni rilevanti dall'ipotesi di normalità, in particolare una maggiore concentrazione di osservazioni attorno al centro e la presenza di code più spesse. Queste caratteristiche emergono in modo ancora più chiaro dai Q–Q plot rispetto alla normale. Nella parte centrale della distribuzione, i punti si dispongono in modo allineato lungo la diagonale, suggerendo che il comportamento dei rendimenti moderati sia approssimativamente compatibile con una distribuzione gaussiana. Tuttavia, agli estremi si osservano deviazioni dalla linea a 45°, che indicano la presenza di code più pesanti rispetto alla normale. Tale fenomeno è particolarmente evidente nella coda sinistra, associata a perdite estreme, confermando quanto già emerso dalle statistiche di curtosi elevate. Nel complesso, queste evidenze grafiche rafforzano l'idea che l'ipotesi di normalità sia inadeguata per descrivere la distribuzione dei rendimenti e, soprattutto, che i comportamenti estremi rivestano un ruolo centrale [7]. La Figura 3.4 invece rappresenta lo scatterplot dei rendimenti giornalieri del DAX e dell'S&P 500, fornendo una prima visualizzazione diretta della relazione bivariata tra i due mercati. Già a colpo d'occhio emerge un'associazione

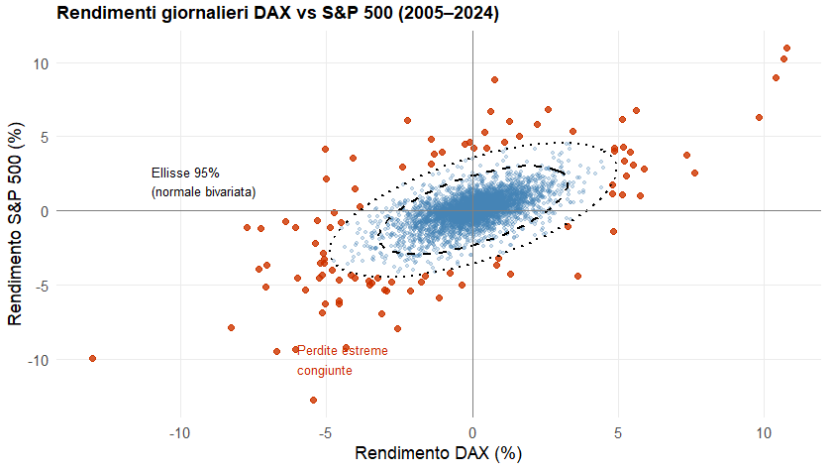


Figura 3.4: Scatterplot dei rendimenti giornalieri DAX vs S&P 500

positiva: nella maggior parte delle osservazioni, rendimenti positivi (negativi) di un indice tendono a essere accompagnati da rendimenti positivi (negativi) dell'altro. Nella regione centrale dello scatterplot la relazione appare approssimativamente lineare, suggerendo che, per movimenti di entità moderata, una dipendenza di tipo quasi lineare possa fornire una descrizione ragionevole. Tuttavia, allontanandosi dal centro, la nube di punti diventa progressivamente più dispersa. La presenza di outlier estremi, soprattutto nelle regioni corrispondenti a forti perdite simultanee, è coerente con quanto osservato nelle statistiche descrittive e nei Q–Q plot e suggerisce l'esistenza di una dipendenza più intensa nelle code della distribuzione congiunta. Inoltre, la forma complessiva della nube non è perfettamente ellittica, indicando che una copula gaussiana, basata esclusivamente sulla correlazione lineare, potrebbe risultare inadeguata per catturare in modo accurato la struttura di dipendenza tra i due mercati. Queste evidenze grafiche rafforzano ulteriormente la motivazione all'utilizzo di copule Archimedee, che consentono di modellare in modo flessibile la dipendenza non lineare e, in particolare, di descrivere in maniera più realistica il comportamento congiunto dei rendimenti in presenza di eventi estremi.

3.3.4 Correlazione preliminare

Prima di procedere alla modellizzazione tramite copule, è utile esaminare alcune misure di correlazione standard tra i rendimenti giornalieri del DAX e dell'S&P 500. La Tabella 3.3 riporta la correlazione lineare di Pearson, il tau di Kendall e il rho di Spearman. Tutte le misure indicano la presenza di una dipendenza positiva di entità medio-alta tra i due mercati. In particolare, la correlazione di Pearson riflette la forte associazione lineare presente soprattutto nella regione centrale

Tabella 3.3: Misure di correlazione tra i rendimenti di DAX e S&P 500

Misura	Valore
Correlazione di Pearson (ρ)	0.6129
Tau di Kendall (τ)	0.3915
Rho di Spearman (ρ_S)	0.5399

della distribuzione. Tuttavia, il divario tra la correlazione lineare e le misure basate sui ranghi suggerisce che la dipendenza non sia completamente catturabile da una struttura puramente lineare. In presenza di code pesanti e di possibili asimmetrie, come evidenziato nelle sezioni precedenti, la correlazione di Pearson può risultare influenzata in modo sproporzionato da osservazioni estreme. Il rho di Spearman, intermedio tra Pearson e Kendall, conferma ulteriormente questo quadro: la dipendenza è significativa ma non perfettamente lineare, e presenta caratteristiche che rendono naturale il ricorso a modelli di copula. In particolare, il valore di τ suggerisce che una parametrizzazione basata su misure di dipendenza a ranghi, come quella adottata nei capitoli precedenti, sia particolarmente adatta per descrivere la relazione tra i rendimenti del DAX e dell'S&P 500.

3.3.5 Costruzione delle pseudo-osservazioni

Per applicare i modelli di copula ai rendimenti dei due indici è necessario trasformare le osservazioni originali in variabili uniformi su [0,1]. Questa trasformazione viene effettuata tramite la trasformazione di probabilità integrale, utilizzando la funzione di distribuzione empirica dei rendimenti. In pratica, a ciascun rendimento viene associato il suo rango all'interno del campione, opportunamente riscalato:

$$\hat{u}_i = \frac{\text{rank}(r_i^{\text{DAX}})}{n+1}, \qquad \hat{v}_i = \frac{\text{rank}(r_i^{\text{S\&P}})}{n+1}. \tag{3.3}$$

La scelta di dividere per $n+1$, anziché per n, è standard nella letteratura sulle copule e consente di evitare valori esattamente pari a 0 o 1 [8, 40]. Tali valori possono infatti causare problemi numerici nel calcolo della densità di alcune famiglie di copule, in particolare in prossimità dei bordi del dominio.

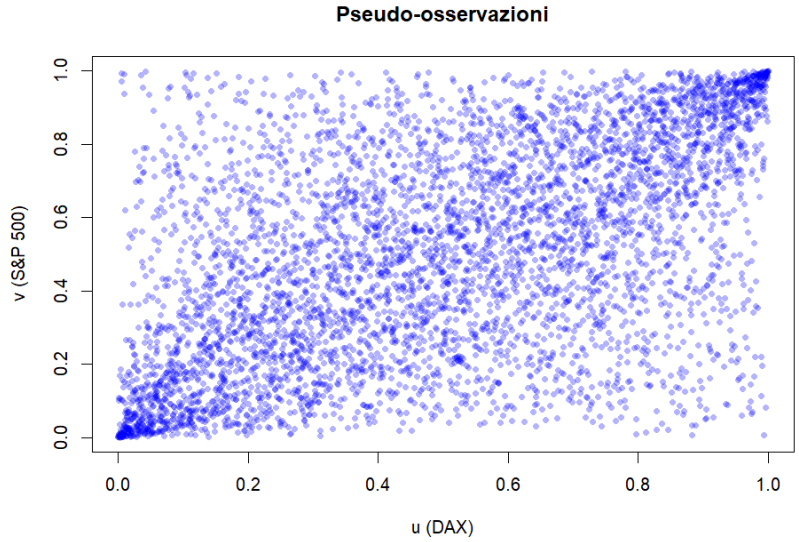


Figura 3.5: Pseudo-osservazioni dei rendimenti DAX e S&P500, uniformi su $[0,1]^2$.

Le pseudo-osservazioni così ottenute assumono valori strettamente compresi tra 0 e 1 e, per costruzione, eliminano l'influenza delle distribuzioni marginali dei rendimenti. In questo modo l'analisi si concentra esclusivamente sulla struttura di dipendenza tra le due serie, che viene preservata attraverso l'ordinamento dei ranghi. Se le marginali sono correttamente trattate, le coppie $(\hat{u}_i, \hat{v}_i)$ possono essere interpretate come osservazioni provenienti direttamente dalla copula sottostante che governa la dipendenza tra i due mercati. La Figura 3.5 riporta la rappresentazione grafica delle pseudo-osservazioni. Sebbene le osservazioni risultino ora distribuite sull'intero quadrato unitario, la presenza di una chiara concentrazione lungo la diagonale principale evidenzia come la dipendenza positiva osservata nei rendimenti originali venga mantenuta anche dopo la trasformazione. Questo passaggio costituisce quindi la base operativa per l'inferenza bayesiana sulla copula, che nei paragrafi successivi verrà condotta utilizzando esclusivamente le pseudo-osservazioni.

3.4 Inferenza bayesiana sulla copula

In questa sezione applichiamo l'approccio di inferenza bayesiana per analizzare la struttura di dipendenza tra i rendimenti giornalieri del DAX e dell'S&P 500. L'attenzione è rivolta al confronto tra diverse famiglie di copule Archimedee, con l'obiettivo di individuare quale di esse sia maggiormente in grado di descrivere le caratteristiche empiriche emerse dai dati. L'inferenza viene condotta direttamente su τ, quantità di immediata interpretazione, evitando così le difficoltà geometriche e numeriche associate alla parametrizzazione tradizionale in termini del parametro naturale θ della copula. Tale scelta consente inoltre una più agevole propagazione dell'incertezza verso misure di interesse, come i coefficienti di tail dependence.

3.4.1 Famiglie di copule considerate e scelta della priori

L'analisi si concentra su tre famiglie di copule Archimedee ampiamente utilizzate in ambito finanziario: Clayton, adatta a modellare perdite simultanee, Gumbel, ideale per descrivere co-movimenti in fasi di espansione e Frank, utile quando non si osservano co-movimenti estremi [1, 6]. La scelta di queste tre famiglie è motivata dal fatto che esse presentano comportamenti molto diversi nelle code della distribuzione congiunta, consentendo di esplorare forme di dipendenza eterogenee. Per tutte le famiglie considerate, adottiamo una prior uniforme sul coefficiente di Kendall. In particolare, poniamo

$$\pi(\tau) = \begin{cases} 1, & \tau \in (0,1) \quad \text{(Clayton, Gumbel)} \\ \frac{1}{2}, & \tau \in (-1,1) \quad \text{(Frank)} \end{cases} \tag{3.4}$$

dove il supporto della prior è ristretto a $\tau \in (0,1)$ per le copule di Clayton e Gumbel, le quali ammettono esclusivamente dipendenza positiva. Questa scelta riflette l'assenza di informazione a priori sulla struttura di dipendenza e lascia che siano i dati a determinare l'intensità e la forma della relazione tra i due mercati.

Copula	**Forma bivariata** $C(u,v)$	**Parametro** $\boldsymbol{\theta}$	λ_L	λ_U
Clayton	$\max\left\{(u^{-\theta}+v^{-\theta}-1)^{-1/\theta},0\right\}$	$\theta > 0$	$2^{-1/\theta}$	0
Gumbel	$\exp\left\{-\left[(-\ln u)^{\theta}+(-\ln v)^{\theta}\right]^{1/\theta}\right\}$	$\theta \geq 1$	0	$2-2^{1/\theta}$
Frank	$-\frac{1}{\theta}\ln\left(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right)$	$\theta \neq 0$	0	0

Tabella 3.4: Confronto tra le principali copule Archimedee

3.4.2 Algoritmo MCMC

L'inferenza bayesiana verrà condotta mediante l'algoritmo Metropolis-Hastings di tipo random walk presentato nel Capitolo 2. Anche qui, per evitare problemi legati ai vincoli sul supporto del parametro, il campionamento viene effettuato nello spazio non vincolato

$$z = \operatorname{atanh}(\tau) \in \mathbb{R}, \tag{3.5}$$

da cui si ottiene $\tau = \tanh(z)$. Ovviamente nella distribuzione a posteriori in z è necessario il termine Jacobiano associato al cambiamento di variabile:

$$\pi(z \mid \mathbf{u}, \mathbf{v}) \propto L(\theta(\tau(z)) \mid \mathbf{u}, \mathbf{v}) \cdot \pi(\tau(z)) \cdot (1 - \tau(z)^2), \tag{3.6}$$

dove L denota la verosimiglianza della copula valutata sulle pseudo-osservazioni e il fattore $(1 - \tau^2)$ rappresenta lo Jacobiano $|\mathrm{d}\tau/\mathrm{d}z|$. In questo modo l'inferenza rimane correttamente formulata pur operando su uno spazio non vincolato. Come si può notare dal codice presente in appendice A, il calcolo della verosimiglianza viene effettuato in scala logaritmica, utilizzando la densità della copula implementata nel pacchetto `copula` in R, mentre la trasformazione tra il parametro τ e il parametro naturale θ della copula è ottenuta tramite l'inversione numerica della relazione tra θ e il tau di Kendall, scelta che garantisce una buona stabilità numerica anche nel caso della copula di Frank, per la quale non vi è una forma chiusa.

Setting dell'algoritmo

Per ciascuna famiglia di copule, la catena è fatta evolvere per un totale di 55 000 iterazioni, di cui le prime 5 000 vengono scartate come burn-in, al fine di ridurre la dipendenza dalle condizioni iniziali. Durante la fase iniziale dell'algoritmo, limitata alle prime 3 000 iterazioni, la varianza della proposta σ viene adattata automaticamente con l'obiettivo di raggiungere un tasso di accettazione prossimo al valore ottimale teorico del 23% per catene unidimensionali. L'adattamento segue la regola moltiplicativa conservativa già vista nel capitolo precedente:

$$\sigma_{t+1} = \begin{cases} 1.02 \cdot \sigma_t, & \text{se } \alpha_t > 0.25 \\ 0.98 \cdot \sigma_t, & \text{altrimenti} \end{cases} \tag{3.7}$$

dove α_t è il tasso di accettazione cumulativo fino all'iterazione t. Al termine del burn-in, la varianza della proposta viene fissata per tutta la fase di campionamento, garantendo che la catena soddisfi le condizioni standard di convergenza.

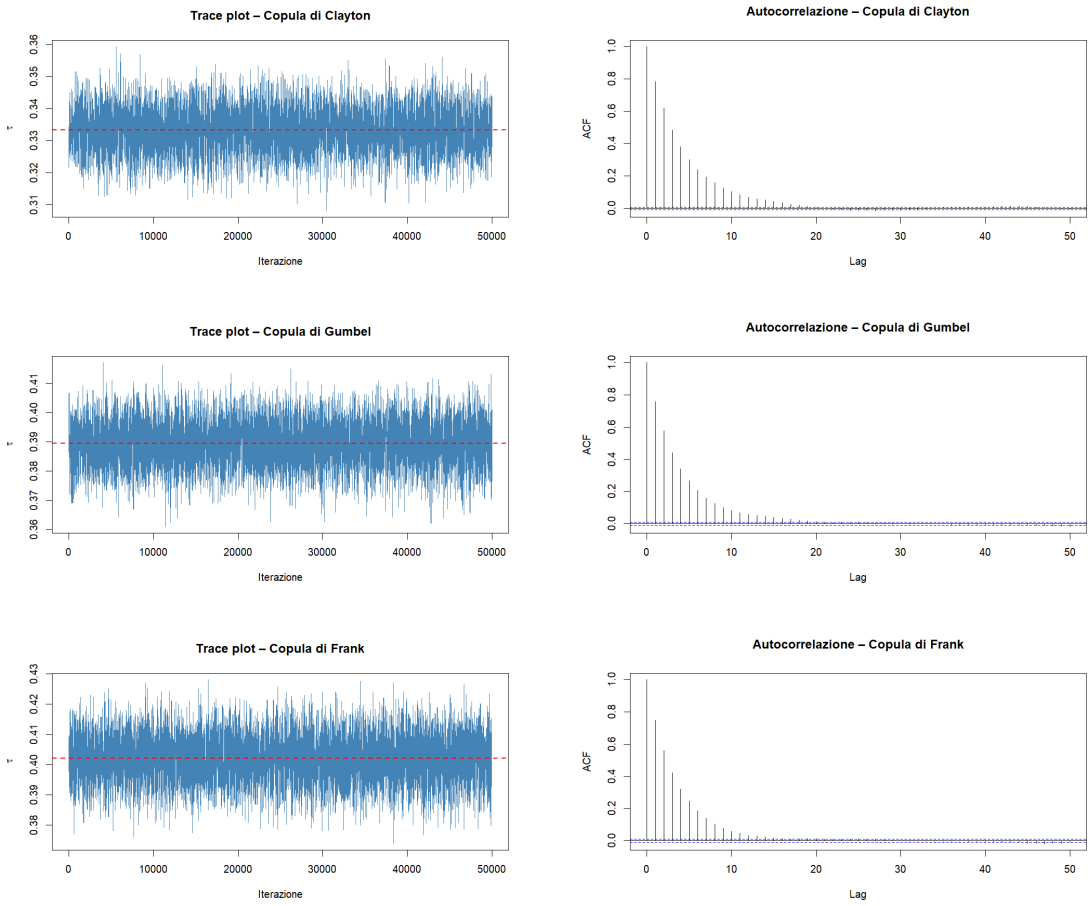


Figura 3.6: Trace plot e ACF per le copule di Clayton, Gumbel e Frank.

Diagnostica MCMC

Prima di discutere i risultati inferenziali, è opportuno verificare che le catene MCMC abbiano effettivamente raggiunto la distribuzione stazionaria e che l'esplorazione dello spazio dei parametri sia stata efficiente. A questo scopo sono state analizzate le principali diagnostiche standard per tutte e tre le famiglie di copule considerate. I trace plot in Figura 3.6 evidenziano, in tutti i casi, un comportamento stazionario della catena, senza trend sistematici né anomalie evidenti. Le oscillazioni attorno alla media a posteriori appaiono regolari e di ampiezza contenuta, indicando che l'algoritmo esplora in modo efficace la regione a maggiore densità di probabilità. La funzione di autocorrelazione decresce rapidamente entro i primi 10-20 lag, indicando una dipendenza limitata tra le iterazioni successive. Infine, osserviamo in Figura 3.7 la densità a posteriori di τ, concentrata in una regione compatibile con una dipendenza positiva di entità moderata, coerente con quanto suggerito dall'analisi

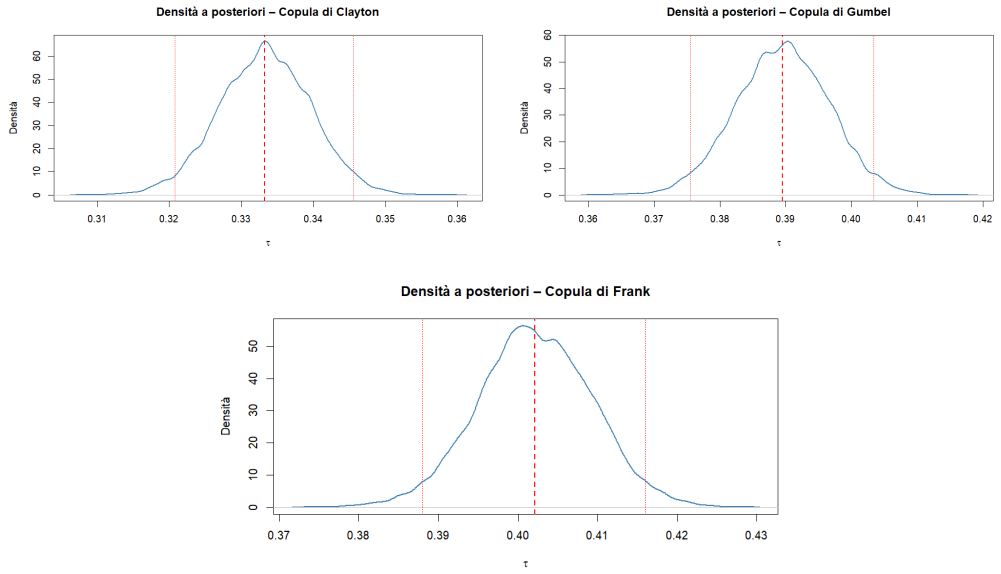


Figura 3.7: Densità a posteriori per le copule di Clayton, Gumbel e Frank

Tabella 3.5: Diagnostica MCMC per le tre famiglie di copule.

Famiglia	Effective Sample Size	Tasso di accettazione	σ finale
Clayton	5 931	0.175	0.050
Gumbel	6 519	0.200	0.050
Frank	7 220	0.208	0.050

descrittiva dei rendimenti. L'assenza di multimodalità rafforza l'affidabilità dell'inferenza. È interessante notare anche come gli intervalli di credibilità associati alle stime di τ per Clayton e Frank, tratteggiati in Figura 3.7, non si sovrappongano. Questa discrepanza suggerisce che la struttura di dipendenza tra i rendimenti del DAX e dell'S&P 500 presenta caratteristiche che vengono catturate in modo diverso dalle tre specificazioni. Un riscontro quantitativo delle buone proprietà di convergenza ed efficienza è fornito dai valori dell'Effective Sample Size e dai tassi di accettazione riportati in Tabella 3.5. In tutti i casi l'ESS risulta elevato, confermando che l'autocorrelazione è contenuta e che la catena fornisce un campione informativo della distribuzione a posteriori. I tassi di accettazione risultano leggermente inferiori al valore teorico ottimale del 23%, ma rientrano pienamente nell'intervallo comunemente considerato soddisfacente per algoritmi di tipo random walk unidimensionali. Nel complesso, le evidenze grafiche e quantitative indicano che le catene hanno raggiunto la distribuzione target e che l'inferenza bayesiana può essere considerata affidabile per tutte le famiglie di copule analizzate.

Tabella 3.6: Inferenza bayesiana per le copule Archimedee (DAX - S&P 500).

Famiglia	$\mathbb{E}[\tau \mid \text{data}]$	IC 95%	$\mathbb{E}[\theta \mid \text{data}]$	$\mathbb{E}[\lambda_L \mid \text{data}]$	$\mathbb{E}[\lambda_U \mid \text{data}]$
Clayton	0.333	[0.321, 0.346]	1.000	0.500	0.000
Gumbel	0.390	[0.376, 0.403]	1.638	0.000	0.473
Frank	0.402	[0.388, 0.416]	4.191	0.000	0.000

3.4.3 Risultati empirici e confronto tra famiglie

La Tabella 3.6 riassume i principali risultati dell'inferenza bayesiana per le tre famiglie di copule Archimedee considerate. Per ciascun modello sono riportate la media a posteriori del coefficiente di Kendall τ, l'intervallo di credibilità al 95%, le stime del parametro naturale θ e dei coefficienti di tail dependence. La stima del coefficiente di Kendall τ varia in modo significativo tra i modelli. La copula di Clayton restituisce il valore più contenuto e un intervallo di credibilità relativamente stretto. Questo valore suggerisce una dipendenza positiva di entità moderata, ma riflette in realtà una struttura fortemente asimmetrica: nel modello di Clayton la dipendenza è concentrata prevalentemente nella coda inferiore. La copula di Gumbel fornisce una stima intermedia di τ, indicando una dipendenza leggermente superiore. Anche in questo caso, tuttavia, la dipendenza non è distribuita in modo uniforme, ma risulta concentrata nella coda superiore della distribuzione. La copula di Frank, infine, produce la stima più elevata del coefficiente di Kendall, coerentemente con la natura simmetrica del modello.

Dipendenza nelle code

Tra le copule considerate, solo la copula di Clayton consente dipendenza asintotica nella coda inferiore. La stima $\hat{\lambda}_L = 0{,}500$ suggerisce che le perdite estreme tendono a verificarsi congiuntamente con una probabilità elevata. In termini pratici, se in uno dei due mercati si realizza un evento fortemente negativo, la probabilità che anche l'altro registri una perdita estrema è intorno al 50%. Le copule di Gumbel e Frank, invece, implicano per costruzione indipendenza asintotica nella coda inferiore, $\lambda_L = 0$, e risultano quindi meno adatte a descrivere co-movimenti estremi negativi. Per quanto riguarda la coda superiore, solo la copula di Gumbel ammette dipendenza asintotica. La stima $\hat{\lambda}_U = 0{,}473$ indica una probabilità non trascurabile di rialzi estremi congiunti. Questo risultato suggerisce che la co-movimentazione tra i mercati può intensificarsi non solo nelle fasi di ribasso, ma anche in presenza di shock fortemente positivi. La dipendenza complessiva, dunque, non è necessariamente unilaterale: può rafforzarsi in entrambe le direzioni, a seconda del regime di mercato. Le copule di Clayton e Frank, al contrario, impongono $\lambda_U = 0$ ed escludono quindi, per costruzione, dipendenza asintotica nella coda superiore.

3.4.4 Selezione del modello

Per individuare quale famiglia di copule descriva in modo più adeguato la struttura di dipendenza tra i rendimenti del DAX e dell'S&P 500 è opportuno affiancare all'analisi delle stime a posteriori un confronto formale tra modelli alternativi. In ambito bayesiano, questo confronto può essere condotto mediante criteri informativi che approssimano la capacità predittiva out-of-sample penalizzando al contempo la complessità effettiva del modello. In questa sezione consideriamo due strumenti standard: il Deviance Information Criterion (DIC) e il Widely Applicable Information Criterion (WAIC).

Deviance Information Criterion

Il DIC, introdotto da Spiegelhalter et al. [18], si basa sulla devianza

$$D(\theta) = -2 \log L(\theta \mid \text{data}),$$

e combina una misura di adattamento medio a posteriori con una penalizzazione per la complessità effettiva. In particolare, indicando con $\overline{D} = \mathbb{E}_{\theta|\text{data}}[D(\theta)]$ la devianza media e con $\bar{\theta}$ la media a posteriori del parametro, il numero effettivo di parametri è definito come

$$p_{\text{DIC}} = \overline{D} - D(\bar{\theta}),$$

e il criterio informativo è dato da

$$\text{DIC} = \overline{D} + p_{\text{DIC}} = 2\overline{D} - D(\bar{\theta}).$$

A parità di scala, valori più bassi del DIC indicano un miglior equilibrio tra qualità dell'adattamento ai dati e complessità del modello.

Widely Applicable Information Criterion

Il WAIC, proposto da Watanabe [19] e discusso in modo sistematico da Gelman et al. [11], è un criterio pienamente bayesiano fondato sulla predictive accuracy pointwise. Sia y_i la i-esima osservazione o, come in questo caso, la i-esima pseudo-osservazione. Definendo la log pointwise predictive density come

$$\text{lppd} = \sum_{i=1}^{n} \log \left(\mathbb{E}_{\theta|\text{data}} \Big[p(y_i \mid \theta) \Big] \right),$$

e la complessità effettiva come

$$p_{\text{WAIC}} = \sum_{i=1}^{n} \text{Var}_{\theta|\text{data}} (\log p(y_i \mid \theta)),$$

il criterio informativo è

$$\mathrm{WAIC} = -2\Big(\mathrm{lppd} - p_{\mathrm{WAIC}}\Big)$$

A differenza del DIC, il WAIC sfrutta l'informazione contenuta nell'intera distribuzione a posteriori, risultando generalmente più stabile quando la distribuzione è asimmetrica o non può essere adeguatamente rappresentata da un singolo valore riassuntivo. Inoltre, essendo calcolato in modo pointwise, può essere interpretato anche come una stima della capacità predittiva del modello.

Calcolo dei criteri informativi

I criteri informativi bayesiani sono stati stimati a partire dai campioni MCMC ottenuti per ciascuna famiglia di copule. Al fine di contenere il costo computazionale senza compromettere l'accuratezza delle stime, entrambi gli indici sono stati calcolati utilizzando un sottocampione casuale di 10 000 iterazioni della catena a posteriori. Nel contesto considerato, ovvero modelli bivariati con un singolo parametro di dipendenza, tali numerosità risultano ampiamente sufficienti a garantire stime numericamente stabili dei criteri informativi. In particolare, l'errore Monte Carlo associato sia alla log pointwise predictive density sia ai termini di penalizzazione (varianze pointwise per il WAIC e numero effettivo di parametri per il DIC) tende a decrescere rapidamente con il numero di campioni disponibili, mentre il guadagno in termini di tempo di calcolo è significativo rispetto all'utilizzo dell'intera catena.

Interpretazione dei risultati

Dalla Tabella 3.7 emerge in modo chiaro che la copula di Gumbel è quella che minimizza sia il DIC sia il WAIC, risultando quindi preferibile secondo entrambi i criteri informativi considerati. Le differenze ampie indicano una superiorità predittiva netta e non riconducibile a fluttuazioni numeriche o a rumore Monte Carlo, fornendo un' evidenza molto forte a favore del modello con valore minimo. I termini p_{DIC} e p_{WAIC} risultano prossimi all'unità per tutte le famiglie considerate, come atteso per modelli con un solo parametro di dipendenza. Le leggere differenze osservate riflettono la diversa modalità con cui i due criteri misurano la complessità effettiva: il DIC si basa sul confronto tra la deviance media e quella valutata nel punto $\bar{\theta}$, mentre il WAIC aggrega contributi pointwise legati alla variabilità della log-verosimiglianza a posteriori. In presenza di osservazioni influenti o di contributi di verosimiglianza eterogenei, tipici dei rendimenti finanziari, il WAIC tende a restituire una penalizzazione leggermente più elevata, come osservato nel caso della copula di Gumbel.

Tabella 3.7: Confronto tra famiglie di copule tramite DIC e WAIC.

Famiglia	DIC	p_{DIC}	WAIC	p_{WAIC}
Clayton	-1817.66	1.01	-1817.33	1.34
Gumbel	**-2107.17**	1.02	**-2106.77**	1.42
Frank	-1840.57	0.95	-1840.28	1.24

Coerenza con l'analisi qualitativa

Il responso dei criteri informativi va interpretato insieme alle evidenze sulle code. La preferenza per la copula di Gumbel suggerisce che, sull'intero campione, la struttura di dipendenza tra i rendimenti del DAX e dell'S&P 500 sia descritta in modo particolarmente efficace da un modello che ammette dipendenza asintotica nella coda superiore. Poiché DIC e WAIC sono criteri orientati alla capacità predittiva media, è naturale che essi privilegino un modello che fornisce un miglior adattamento complessivo, anche a scapito di una descrizione più mirata delle code. Questo risultato non è in contraddizione con quanto emerso dall'analisi delle dipendenze di coda. In particolare, la copula di Clayton, pur non risultando la migliore secondo DIC e WAIC sull'intero periodo, rimane particolarmente adatta a descrivere la dipendenza nelle perdite estreme, che rappresenta l'aspetto più rilevante in applicazioni di risk management e di analisi del contagio finanziario. Gli episodi di crisi, sebbene economicamente rilevanti, rappresentano infatti una porzione limitata delle osservazioni complessive. È pertanto plausibile attendersi che, in analisi condotte su sotto-periodi caratterizzati da forte stress di mercato, la copula di Clayton possa risultare più informativa e più appropriata in contesti di crisi e per la misurazione del rischio congiunto.

3.5 Analisi per sotto-periodi

Un aspetto centrale nell'analisi della dipendenza tra mercati finanziari riguarda la sua stabilità nel tempo. In particolare, è naturale chiedersi se la struttura di dipendenza stimata sull'intero campione rimanga invariata oppure se subisca modifiche rilevanti in corrispondenza di fasi di crisi o di maggiore turbolenza [41]. Per indagare questa dimensione dinamica, il campione viene suddiviso in sotto-periodi omogenei dal punto di vista economico e le copule di Clayton e Gumbel vengono ri-stimate separatamente per ciascuno di essi.

Tabella 3.8: Stima della dipendenza per sotto-periodi (copula di Clayton).

Periodo	n	$\hat{\tau}$	IC 95%	$\hat{\lambda}_L$	IC 95%
Pre-crisi (2005–2007)	744	0.267	[0.233, 0.301]	0.386	[0.320, 0.446]
Crisi GFC (2008–2009)	496	0.376	[0.339, 0.412]	0.562	[0.508, 0.609]
Crisi del debito (2010–2012)	744	0.409	[0.380, 0.436]	0.605	[0.568, 0.639]
Bull market (2013–2019)	1 723	0.310	[0.289, 0.331]	0.462	[0.426, 0.496]
COVID-19 (2020)	248	0.399	[0.345, 0.447]	0.592	[0.518, 0.651]
Post-COVID (2021–2024)	993	0.274	[0.243, 0.302]	0.398	[0.340, 0.450]

Tabella 3.9: Stima della dipendenza per sotto-periodi (copula di Gumbel).

Periodo	n	$\hat{\tau}$	IC 95%	$\hat{\lambda}_U$	IC 95%
Pre-crisi (2005–2007)	744	0.312	[0.271, 0.350]	0.388	[0.343, 0.431]
Crisi GFC (2008–2009)	496	0.440	[0.398, 0.480]	0.525	[0.482, 0.566]
Crisi del debito (2010–2012)	744	0.490	[0.459, 0.520]	0.576	[0.545, 0.605]
Bull market (2013–2019)	1 723	0.364	[0.340, 0.388]	0.446	[0.419, 0.471]
COVID-19 (2020)	248	0.420	[0.359, 0.478]	0.505	[0.440, 0.564]
Post-COVID (2021–2024)	993	0.341	[0.309, 0.370]	0.421	[0.386, 0.453]

3.5.1 Definizione dei sotto-periodi

La suddivisione temporale è guidata da eventi macroeconomici e finanziari di particolare rilievo. In particolare, si considerano sei sotto-periodi distinti: il periodo pre-crisi dal 2005 al 2007, caratterizzato da una crescita relativamente stabile; la crisi finanziaria globale dal 2008 al 2009, dalla bancarotta di Lehman Brothers alle prime fasi di recupero; il periodo di crisi del debito sovrano europeo dal 2010 al 2012; il lungo mercato rialzista dal 2013 al 2019, caratterizzato da volatilità contenuta; lo shock pandemico del 2020; e infine la fase post-COVID dal 2021 al 2024, contraddistinta da una graduale normalizzazione delle condizioni di mercato [32, 33, 34]. Per ciascun sotto-periodo sono state stimate sia la copula di Clayton che la copula di Gumbel, mediante lo stesso approccio bayesiano illustrato nelle sezioni precedenti, basato sulla parametrizzazione in termini del coefficiente di Kendall τ e su campionamento MCMC nello spazio non vincolato.

3.5.2 Confronto Clayton vs Gumbel per sotto-periodi

La Tabelle 3.8 e 3.9 riportano, per ciascun periodo, la stima a posteriori di τ e del coefficiente di dipendenza rispettivamente nella coda inferiore λ_L e superiore λ_U, insieme agli intervalli di credibilità al 95%. A supporto dell'analisi quantitativa, la Figura 3.8 visualizza l'evoluzione temporale del coefficiente di Kendall e dei coefficienti di coda stimati per sotto-periodi. L'andamento evidenzia chiaramente in entrambi i casi un aumento della dipendenza in corrispondenza dei principali

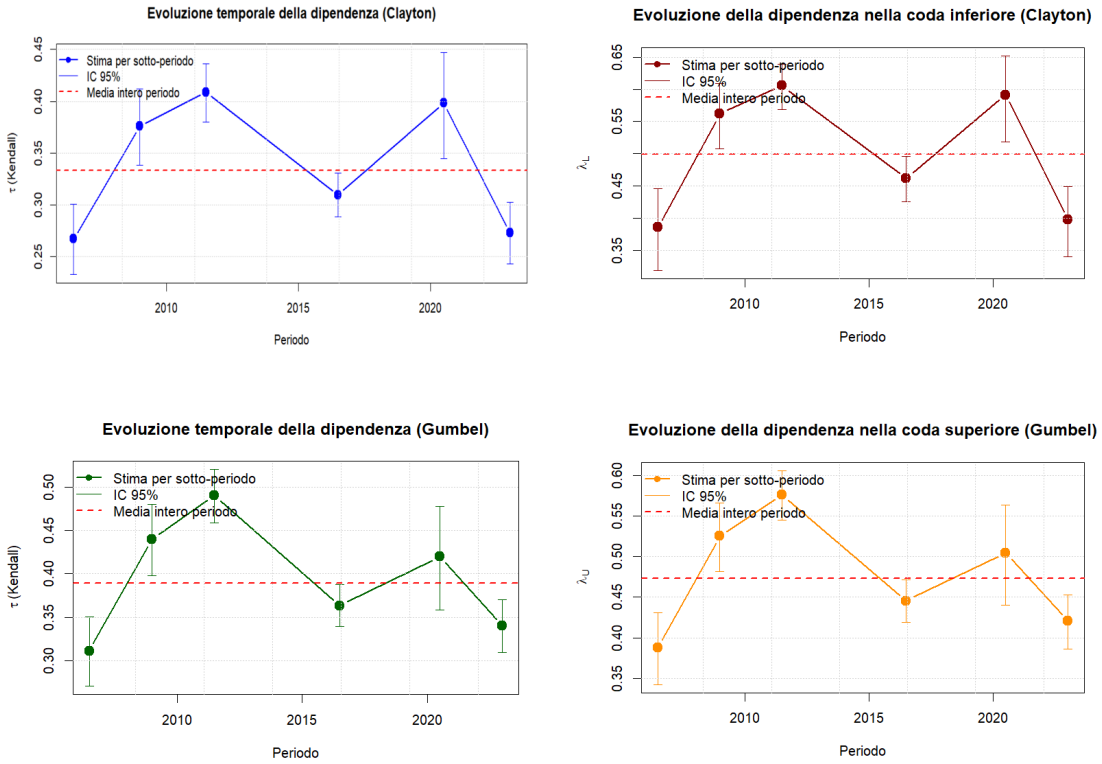


Figura 3.8: Evoluzione della dipendenza nei sotto-periodi per Clayton e Gumbel.

episodi di crisi (2008–2009, 2010–2012, 2020), seguito da una graduale riduzione durante le fasi di normalizzazione. Il confronto tra le due famiglie di copule rivela un'asimmetria importante nella struttura di dipendenza. Sebbene entrambi i coefficienti di tail dependence aumentino durante le fasi di stress, l'intensificazione risulta più pronunciata per la coda inferiore. Questa evidenza conferma che i mercati si sincronizzano di più nelle perdite rispetto ai guadagni. [2, 3]. Al tempo stesso, la presenza di una dipendenza di coda superiore non trascurabile anche durante i periodi di crescita suggerisce che la sincronizzazione nei rialzi estremi è comunque presente, pur essendo meno intensa rispetto a quella osservata nei ribassi. Questo risultato aiuta a comprendere la selezione della copula di Gumbel tramite i criteri informativi DIC e WAIC sull'intero campione: una buona capacità predittiva media deve necessariamente bilanciare le fasi espansive, dominanti in termini di frequenza temporale, e le fasi di crisi, caratterizzate da dipendenza più elevata ma meno frequenti.

Conclusioni

Risultati principali

In questo lavoro è stato sviluppato un approccio di inferenza bayesiana per copule Archimedee basato sulla parametrizzazione in termini del coefficiente di Kendall. L'analisi ha combinato aspetti metodologici, computazionali ed empirici, con particolare attenzione alla modellizzazione della dipendenza tra mercati azionari internazionali. L'adozione del coefficiente τ come parametro di dipendenza si è dimostrata vantaggiosa sotto molteplici profili. Dal punto di vista interpretativo fornisce una misura standardizzata e confrontabile della dipendenza; è inoltre invariante rispetto a trasformazioni monotone e ha un significato immediato indipendentemente dalla famiglia di copule considerata [1]. Sul piano computazionale, la parametrizzazione in τ ha prodotto catene MCMC con efficienza superiore rispetto all'approccio tradizionale, basato sul parametro naturale θ del generatore. Come documentato nella Sezione 2.4.3, la regolarità della distribuzione a posteriori facilita la convergenza degli algoritmi e riduce la sensibilità alle condizioni iniziali [12]. Lo studio Monte Carlo condotto nel Capitolo 2 ha inoltre confermato le proprietà frequentiste della stima bayesiana: bias trascurabile, convergenza dell'RMSE secondo il rate teorico $n^{-1/2}$, e copertura empirica degli intervalli di credibilità coerente con il livello nominale anche per campioni di dimensione moderata [28]. L'analisi comparativa presentata nella Sezione 2.7 ha messo poi in luce differenze rilevanti tra metodi di stima alternativi. Lo stimatore basato sul tau di Kendall campionario, pur essendo computazionalmente immediato, risulta meno efficiente in termini di RMSE rispetto agli approcci basati sulla verosimiglianza. La massima verosimiglianza produce stime puntuali accurate, ma gli intervalli di confidenza costruiti mostrano una sottocopertura per campioni di dimensione moderata [6, 10]. L'approccio bayesiano, al contrario, fornisce intervalli di credibilità ben calibrati e consente una quantificazione più accurata dell'incertezza, particolarmente rilevante in applicazioni di gestione del rischio. La possibilità di propagare in modo naturale l'incertezza parametrica a quantità derivate di interesse, come i coefficienti di tail dependence, rappresenta un ulteriore vantaggio operativo dell'inferenza bayesiana.

L'analisi empirica condotta sui rendimenti giornalieri del DAX e dell'S&P 500 nel periodo 2005–2024 ha evidenziato una struttura di dipendenza tra i due mercati. Sebbene i criteri informativi DIC e WAIC identifichino la copula di Gumbel come il modello con migliore capacità predittiva media sull'intero campione, l'analisi per sotto-periodi rivela che la dipendenza si intensifica in modo differenziato nelle code della distribuzione congiunta. La copula di Clayton, caratterizzata da tail dependence nella coda inferiore, cattura efficacemente la tendenza dei due mercati a muoversi insieme durante episodi di forte ribasso. Il coefficiente λ_L stimato varia da circa 0.40 nei periodi di stabilità fino a 0.60 durante le fasi di crisi, indicando che la probabilità di perdite estreme simultanee aumenta significativamente durante periodi di stress finanziario [7]. Al contrario, la dipendenza nella coda superiore, pur presente, risulta meno intensa e più stabile nel tempo. Questa asimmetria ha, ad esempio, implicazioni dirette per la gestione del rischio di portafoglio; sembrerebbe infatti che i benefici della diversificazione internazionale tendono a ridursi proprio quando sarebbero più necessari, confermando l'evidenza empirica già documentata in letteratura per eventi di contagio finanziario [3, 16].

Limitazioni ed estensioni

È opportuno riconoscere alcune limitazioni riguardo alle conclusioni ottenute che portano poi a estensioni metodologiche e applicative.

Dimensionalità dell'analisi e vine copule

L'intera trattazione si è concentrata sul caso bivariato. Sebbene questa scelta consenta una chiara interpretazione dei risultati, molte applicazioni pratiche richiedono la modellizzazione congiunta di un numero maggiore di variabili. Per superare i limiti imposti dalla dimensionalità bivariata senza rinunciare alla flessibilità delle copule Archimedee, un'estensione naturale consiste nell'adozione di copule vine. Questi modelli, introdotti da Bedford e Cooke [42, 43], permettono di costruire distribuzioni multivariate componendo in modo gerarchico copule bivariate. Le copule vine mantengono la flessibilità nella scelta della famiglia per ciascuna componente bivariata e consentono di modellare strutture di dipendenza eterogenee e asimmetriche in dimensioni elevate [44]. L'inferenza bayesiana e la parametrizzazione in termini di τ potrebbero essere estese a questo contesto, facilitando l'interpretazione e il confronto tra le diverse componenti di dipendenza [45].

Dipendenza dinamica e copule time-varying

L'inferenza bayesiana è stata condotta prima per l'intero campione e poi separatamente per sotto-periodi identificati ex post sulla base di eventi economici noti. Questo approccio, sebbene informativo dal punto di vista descrittivo, presenta due limiti. In primo luogo, la segmentazione temporale è soggetta a un potenziale bias di selezione, poiché le date di interruzione sono scelte sulla base di conoscenze successive. In secondo luogo, trattare ciascun sotto-periodo come indipendente ignora la possibile continuità della dinamica di dipendenza. Come evidenziato dalle Figure 3.9 e 3.11, la dipendenza tra DAX e S&P 500 varia nel tempo, suggerendo che un modello con parametri costanti possa essere inadeguato su orizzonti temporali estesi. Durante le fasi di crisi finanziaria, si osserva un aumento sia del coefficiente di Kendall sia della tail dependence nella coda inferiore, con valori che arrivano a superare del 50% i livelli registrati nei periodi di stabilità. Questa variabilità temporale motiva fortemente l'adozione di modelli con dipendenza dinamica, in grado di catturare l'evoluzione della struttura di dipendenza in modo continuo anziché mediante segmentazioni discrete del campione. Una classe di modelli promettenti in tal senso è rappresentata dalle copule time-varying, nelle quali il parametro di dipendenza evolve nel tempo secondo una dinamica specificata. Ad esempio Patton [46] ha proposto un modello in cui il parametro della copula segue un processo autoregressivo guidato da una trasformazione delle osservazioni passate, garantendo al tempo stesso che il parametro rimanga all'interno del dominio ammissibile. Questo approccio chiaramente presenta sfide computazionali aggiuntive, poiché la dimensione dello spazio parametrico cresce con la serie temporale.

Costruzione di portafoglio e risk management

Sul fronte applicativo, i metodi sviluppati in questa tesi possono essere impiegati in problemi di costruzione ottimale del portafoglio e allocazione del rischio [7]. La modellizzazione accurata della dipendenza nelle code è cruciale per stimare misure di rischio congiunte e per la valutazione dei benefici della diversificazione in scenari di stress. Un'estensione naturale potrebbe essere quella di integrare l'inferenza bayesiana su copule Archimedee con l'ottimizzazione robusta, dove l'incertezza sui parametri di dipendenza viene esplicitamente incorporata nel processo decisionale. Ad esempio, tecniche di risk management potrebbero essere adattate per tener conto della distribuzione a posteriori di τ e dei coefficienti di tail dependence, allocando il rischio in modo da controllare l'esposizione a scenari estremi congiunti con probabilità elevata a posteriori. Analogamente, strategie di hedging dinamico potrebbero beneficiare di modelli time-varying che aggiornano in tempo reale la stima della dipendenza, permettendo aggiustamenti tempestivi della copertura in risposta a variazioni nella struttura di dipendenza nei mercati [16].

Appendice A

Codice R

Per migliorare la leggibilità della tesi e favorire la riproducibilità dei risultati, il codice `R` utilizzato per le analisi, le simulazioni e le visualizzazioni grafiche non è riportato integralmente nel presente documento, ma è reso disponibile in una repository online pubblica. Il codice è organizzato in script distinti, ciascuno corrispondente a una specifica parte della tesi:

- `Capitolo 1 - Visualizzazioni grafiche.R`
  Script utilizzato per la produzione delle rappresentazioni grafiche delle copule e delle densità di copula presentate nel Capitolo 1.
- `Capitolo 2 - Esempi 3 e 4.R`
  Codice relativo agli esempi illustrativi e alle simulazioni preliminari discussi nel Capitolo 2.
- `Capitolo 2 - Inferenza Bayesiana.R`
  Implementazione dell'inferenza bayesiana nei modelli di copula Archimedee, inclusa l'implementazione dell'algoritmo di Metropolis–Hastings e la diagnostica MCMC.
- `Capitolo 3 - Analisi empirica.R`
  Script utilizzato per l'analisi empirica dei rendimenti finanziari e la stima dei modelli di copula sui dati reali.

Il dataset utilizzato nell'analisi empirica è contenuto nel file `Dataset.xlsx`. L'intera repository è disponibile al seguente indirizzo:

`https://github.com/aleatta00/Codice-R-Tesi-Magistrale/`

Bibliografia

[1] Roger B. Nelsen. *An Introduction to Copulas*. 2ª ed. Springer Series in Statistics. Springer, 2006.

[2] Kristin J. Forbes e Roberto Rigobon. «No contagion, only interdependence: Measuring stock market comovements». In: *The Journal of Finance* (2002).

[3] François Longin e Bruno Solnik. «Extreme correlation of international equity markets». In: *The Journal of Finance* (2001).

[4] A. Sklar. «Fonctions de répartition à n dimensions et leurs marges». In: *Publications de l'Institut de Statistique de l'Université de Paris* 8 (1959).

[5] Joe. *Multivariate Models and Dependence Concepts*. Chapman & Hall, 1997.

[6] Joe. *Dependence Modeling with Copulas*. CRC Press, 2014.

[7] Alexander J. McNeil, Rüdiger Frey e Paul Embrechts. *Quantitative Risk Management*. 2ª ed. Princeton University Press, 2015.

[8] Genest, Ghoudi e Rivest. «A semiparametric estimation procedure of dependence parameters in multivariate families of distributions». In: *Biometrika* (1995).

[9] Joe e Xu. *The Estimation Method of Inference Functions for Margins for Multivariate Models*. Rapp. tecn. University of British Columbia, 1996.

[10] Smith. «Bayesian Approaches to Copula Modelling». In: *arXiv* (2011).

[11] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari e Donald B. Rubin. *Bayesian Data Analysis*. 3ª ed. CRC Press, 2013.

[12] Papaspiliopoulos, Roberts e Sköld. «General framework for the parametrization of hierarchical models». In: *Statistical Science* (2007).

[13] G. Kendall. «A new measure of rank correlation». In: *Biometrika* (1938).

[14] Clara Grazian e Brunero Liseo. «Approximate Bayesian inference in semiparametric copula models». In: *Bayesian Analysis* (2017).

[15] Roberts, Gelman e Gilks. «Weak convergence and optimal scaling of random walk Metropolis algorithms». In: *The Annals of Applied Probability* (1997).

[16] Paul Embrechts, Alexander J. McNeil e Daniel Straumann. «Correlation and dependence in risk management: properties and pitfalls». In: *Risk Management: Value at Risk and Beyond*. Cambridge University Press, 2002.

[17] Kole, Koedijk e Verbeek. «Selecting copulas for risk management». In: *Journal of Banking & Finance* (2007).

[18] David J. Spiegelhalter, Nicola G. Best, Bradley P. Carlin e Angelika van der Linde. «Bayesian measures of model complexity and fit». In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* (2002).

[19] Sumio Watanabe. «Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory». In: *Journal of Machine Learning Research* (2010).

[20] Joe. «Asymptotic efficiency of the two-stage estimation method for copula-based models». In: *Journal of Multivariate Analysis* (2005).

[21] Robert e Casella. *Monte Carlo Statistical Methods*. 2ª ed. Springer, 2004.

[22] Steve Brooks, Andrew Gelman, Galin Jones e Xiao-Li Meng. *Handbook of Markov Chain Monte Carlo*. Chapman e Hall/CRC, 2011.

[23] Gareth O. Roberts e Jeffrey S. Rosenthal. «Optimal scaling for various Metropolis-Hastings algorithms». In: *Statistical Science* (2001).

[24] Metropolis, Rosenbluth e Teller. «Equation of State Calculations by Fast Computing Machines». In: *The Journal of Chemical Physics* (1953).

[25] W. Keith Hastings. «Monte Carlo Sampling Methods Using Markov Chains and Their Applications». In: *Biometrika* (1970).

[26] Christophe Andrieu e Johannes Thoms. «A tutorial on adaptive MCMC». In: *Statistics and Computing* (2008).

[27] James O. Berger e Jose M. Bernardo. «On the development of reference priors». In: *Bayesian Statistics* (1992).

[28] E. L. Lehmann e George Casella. *Theory of Point Estimation*. Springer, 1998.

[29] *DAX ®*. Web page.

[30] *DAX (Börse Frankfurt Lexicon)*. Web page.

[31] *S&P 500 ®*. Web page.

[32] International Monetary Fund. *Global Financial Stability Report: Responding to the Financial Crisis and Measuring Systemic Risks*. Web page. 2009.

[33] European Central Bank. *Financial Stability Review*. Web page. 2012.

[34] National Bureau of Economic Research. *Business Cycle Dating Committee Announcement*. Web page. 2020.

[35] International Monetary Fund. *World Economic Outlook: The Great Lockdown*. Web page. 2020.

[36] Yahoo Help. *Download historical data in Yahoo Finance*. Web page.

[37] Yahoo Help. *Exchanges and data providers on Yahoo Finance*. Web page.

[38] Jianqing Fan e Qiwei Yao. *The Elements of Financial Econometrics*. Cambridge University Press, 2017.

[39] Rama Cont. «Empirical properties of asset returns: stylized facts and statistical issues». In: *Quantitative Finance* (2001).

[40] Hidetoshi Tsukahara. «Semiparametric estimation in copula models». In: *The Canadian Journal of Statistics* (2005).

[41] Andrew J. Patton. *Copula-Based Models for Financial Time Series*. Rapp. tecn. University of Oxford, 2008.

[42] Bedford e Cooke. «Probability density decomposition for conditionally dependent random variables modeled by vines». In: *Annals of Mathematics* (2001).

[43] Tim Bedford e Roger M Cooke. «Vines: A new graphical model for dependent random variables». In: *The Annals of Statistics* (2002).

[44] Dorota Kurowicka e Harry Joe. *Dependence Modeling: Vine Copula Handbook*. World Scientific, 2011.

[45] Aas, Czado, Frigessi e Bakken. «Pair-copula constructions of multiple dependence». In: *Insurance: Mathematics and Economics* (2009).

[46] Andrew J Patton. «Modelling asymmetric exchange rate dependence». In: *International Economic Review* (2006).