



SAPIENZA  
UNIVERSITÀ DI ROMA

## Copule per la modellizzazione di dipendenze non monotone: applicazioni in finanza e demografia

Scuola di Dottorato in Scienze Statistiche - XXXVI Ciclo  
Dottorato di Ricerca in Scienze Attuariali

**Manfred Marvin Marchione**

Matricola 1647551

Relatore

Prof. Fabio Baione

Anno Accademico 2022-2023

---

**Copule per la modellizzazione di dipendenze non monotone: applicazioni in  
finanza e demografia**

Tesi di Dottorato. Sapienza Università di Roma

© 2023 Manfred Marvin Marchione. Tutti i diritti riservati

Questa tesi è stata composta con  $\text{\LaTeX}$  e la classe Sapthesis.

Email dell'autore: [manfredmarvin.marchione@uniroma1.it](mailto:manfredmarvin.marchione@uniroma1.it)

## Sommario

L'oggetto del presente elaborato è lo studio dell'uso delle copule per la modellizzazione di dipendenze non monotone. In particolare, la tesi propone un metodo generale per la costruzione di copule che possano descrivere dipendenze di forma arbitraria. La metodologia proposta si basa sull'utilizzo di funzioni che preservano la distribuzione uniforme. A differenza della letteratura esistente sull'argomento, il metodo qui proposto consente di costruire tali funzioni in forma parametrica non limitando la scelta alle sole funzioni lineari a tratti, permettendo quindi di ottenere modelli maggiormente flessibili che meglio si adattano a descrivere la complessità di fenomeni reali. L'applicazione dei risultati ottenuti sarà illustrata esaminando due fenomeni in cui è tipicamente presente una dipendenza di tipo non monotono. Il primo riguarda la relazione tra rendimento e volume degli scambi nei mercati finanziari. Il secondo, in ambito demografico, riguarda l'impatto delle temperature sulla mortalità, di cui proponiamo in questa tesi lo studio attraverso l'uso di copule di dipendenza non monotona.

# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
<b>2</b>	<b>Copule: definizione e proprietà</b>	<b>3</b>
2.1	Risultati preliminari . . . . .	3
2.2	Copule e teorema di Sklar . . . . .	5
2.3	Confini di Fréchet-Hoeffding . . . . .	11
<b>3</b>	<b>Misure di dipendenza</b>	<b>14</b>
3.1	Tau di Kendall . . . . .	16
3.2	Rho di Spearman . . . . .	17
3.3	Relazione tra tau di Kendall e rho di Spearman . . . . .	19
3.4	Dipendenza di coda . . . . .	23
<b>4</b>	<b>Alcuni esempi di copule</b>	<b>24</b>
4.1	Copule Archimedee . . . . .	24
4.1.1	Copula Frank . . . . .	27
4.1.2	Copula Clayton . . . . .	28
4.1.3	Copula Gumbel . . . . .	29
4.2	Copule ellittiche . . . . .	30
4.2.1	Copula Gaussiana . . . . .	32
4.2.2	Copula $t$ . . . . .	33
4.3	Un metodo generale per la costruzione di copule . . . . .	34
<b>5</b>	<b>Dipendenze non monotone</b>	<b>37</b>
5.1	Costruzione di copule per dipendenze non monotone . . . . .	40
5.2	Stima dei parametri e selezione del modello . . . . .	43
<b>6</b>	<b>La relazione rendimento-volume in finanza</b>	<b>45</b>
6.1	Dati e risultati . . . . .	46
<b>7</b>	<b>Effetto delle temperature sulla mortalità</b>	<b>51</b>
7.1	Dati e risultati . . . . .	52
<b>8</b>	<b>Simulazione Monte Carlo da copule</b>	<b>62</b>
8.1	Metodi di carattere generale . . . . .	62
8.2	Simulazione da copule archimedee . . . . .	63
8.3	Simulazione da copule ellittiche . . . . .	66

<b>Indice</b>	<b>v</b>
8.4 Simulazione da copule non monotone . . . . .	70
<b>Conclusioni</b>	<b>71</b>
<b>Bibliografia</b>	<b>72</b>
<b>A Codici R: stima dei parametri</b>	<b>78</b>
<b>B Codici R: simulazioni Monte Carlo</b>	<b>91</b>

# Capitolo 1

## Introduzione

Mentre l'introduzione delle copule in ambito statistico risale al lavoro di Sklar [64], secondo Genest et al. [28] la produzione scientifica sull'argomento ha visto una considerevole crescita solamente a partire dal 1999. Un forte impulso a tale crescita è stato dato in particolare dai libri di Joe [37] e Nelsen [53] che hanno fortemente contribuito alla diffusione della teoria delle copule nella comunità scientifica. Il potenziale delle copule nelle applicazioni finanziarie e attuariali è stato evidenziato da Frees e Valdez [21] e da Embrechts [18] e le copule sono oggi uno strumento di normale utilizzo nella modellizzazione delle dipendenze tra variabili aleatorie.

La maggior parte delle copule note nella letteratura si adatta tuttavia alla descrizione di fenomeni che manifestano una dipendenza di tipo monotono. Sorprendentemente, l'utilizzo delle copule per la descrizione di dipendenze non monotone è stato trattato di rado e la letteratura sull'argomento appare limitata. Un'analisi matematica del problema è stata proposta da Scarsini e Venetoulis [58]. Gli autori hanno osservato che strutture di dipendenza arbitrarie possono essere ottenute applicando trasformazioni che conservano la distribuzione uniforme alle marginali di un vettore avente componenti uniformi. In particolare, se due variabili uniformemente distribuite esibiscono massima dipendenza positiva, l'applicazione di una funzione che preservi la distribuzione uniforme porta a una distribuzione bivariata il cui supporto coincide con il grafico della funzione utilizzata. Nello stesso lavoro, gli autori propongono diversi esempi in cui le funzioni preservanti la distribuzione uniforme sono lineari a tratti. Lo scopo del presente elaborato è quello di sviluppare l'argomento e, in particolare, ottenere un metodo generale per la costruzione di funzioni preservanti la distribuzione uniforme. Ciò consente, di conseguenza, di costruire copule in grado di descrivere dipendenze di tipo arbitrario. Tra i vantaggi del metodo proposto vi è la possibilità di costruire tali copule in forma parametrica, consentendo quindi di effettuare la stima dei parametri nelle applicazioni pratiche.

La metodologia di stima dei parametri e di selezione del modello sarà illustrata applicando le metodologie descritte allo studio della relazione tra rendimento e volume nei mercati finanziari. La nascita della letteratura scientifica sulla dipendenza non monotona tra rendimento e volume è generalmente attribuita al lavoro di Osborne [56], in seguito al quale sono stati proposti diversi modelli matematici (si veda per esempio Clark [11] e Epps [19]) e studi empirici (tra cui Crouch [13], Tauchen e

Pitts [66], Lee e Rui [45] e Gündüz e Hatemi-J [30]) nel tentativo di investigare il fenomeno. Una dettagliata revisione dei primi lavori sull'argomento è reperibile nell'articolo di Karpoff [39]. Un'analisi della relazione rendimento-volume mediante l'utilizzo delle copule è stata effettuata da Neto [55] che ha proposto a tale scopo una copula di dipendenza non monotona. Come si vedrà, la copula utilizzata dall'autore può essere facilmente ottenuta attraverso un metodo generale qui proposto per la costruzione di copule che descrivono dipendenze di forma arbitraria.

Un secondo originale esempio di dipendenza non monotona, ampiamente trattato nella letteratura demografica ed epidemiologica, è costituito dall'impatto che le temperature hanno sulla mortalità. Infatti, sia le basse temperature che le ondate di calore hanno un impatto sulla salute che conduce tipicamente a un incremento della mortalità. Tale fenomeno è stato identificato, per esempio, da autori come Kovats et al. [42] e Linares e Díaz [47]. La letteratura più recente sul tema, tra cui figurano i lavori di Masselot et al. [49] e Gasparrini et al. [25], propone l'utilizzo di modelli di tipo regressivo non lineare per descrivere la relazione non monotona tra temperature e mortalità. Utilizzando un dataset relativo alle temperature rilevate in diversi comuni italiani, proporremo in questa tesi l'utilizzo delle copule per la modellizzazione della dipendenza non monotona tra temperatura e mortalità. In particolare, illustreremo quali informazioni è possibile ricavare da modelli basati sull'uso delle copule e quali sono invece i limiti metodologici dell'analisi proposta.

La tesi si articola nel modo seguente. Nel capitolo 2 sono descritte le idee fondamentali sulle copule, a partire dalla definizione, le proprietà elementari e il loro ruolo nella descrizione delle dipendenze in probabilità. Il capitolo 3 è invece dedicato allo studio di misure di dipendenza alternative proposte nella letteratura. Di queste si studierà il legame con le copule e se ne discuteranno brevemente i limiti derivanti dal fatto che sono misure inadatte a descrivere dipendenze non monotone. Il capitolo 4 è dedicato allo studio di alcune delle famiglie di copule più note nella letteratura. Si passerà poi, nel capitolo 5, allo studio delle dipendenze non monotone. In particolare, si dimostrerà il risultato principale dell'elaborato, che fornisce di fatto un metodo di costruzione per copule in grado di descrivere strutture di dipendenza arbitrarie. Estendendo e generalizzando il lavoro di Neto [55] sulla relazione rendimento-volume, illustreremo la metodologia di stima dei parametri per copule di dipendenza non monotona nel capitolo 6, per poi applicare i risultati ottenuti allo studio della relazione tra temperatura e mortalità nel capitolo 7. Concludiamo con un capitolo dedicato alla simulazione Monte Carlo di campioni generati da una copula data, illustrando in particolare i metodi da utilizzare nel caso di dipendenze non monotone.

## Capitolo 2

# Copule: definizione e proprietà

Come anticipato nell'introduzione, le copule rappresentano uno strumento fondamentale nello studio delle distribuzioni multivariate. La loro utilità risiede nel fatto che consentono di separare lo studio della struttura di dipendenza tra due variabili aleatorie dallo studio delle distribuzioni marginali. In virtù del teorema di Sklar [64], è infatti possibile depurare una distribuzione multivariata dall'effetto che le marginali hanno sulla forma distribuzionale, studiando quindi solo la componente che descrive la dipendenza. Da un punto di vista probabilistico, una copula può essere definita come una funzione di ripartizione multivariata avente marginali uniformi. Lo scopo del presente capitolo è quello di formalizzare le idee appena descritte, definendo le copule e studiandone le proprietà fondamentali fino ad arrivare al teorema di Sklar.

### 2.1 Risultati preliminari

Prima di presentare le definizioni fondamentali per lo studio delle copule, introduciamo alcuni elementi notazionali che saranno utilizzati nel presente elaborato. Innanzitutto, sarà fatto ampio utilizzo della cosiddetta retta reale estesa  $\overline{\mathbb{R}}$ , definita come

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}.$$

Inoltre, data una generica funzione  $H$ , identificheremo il suo dominio mediante il simbolo  $\text{Dom}(H)$  e la sua immagine con  $\text{Im}(H)$ .

**Definizione 2.1.1.** Siano  $S_1$  ed  $S_2$  due sottinsiemi non vuoti di  $\overline{\mathbb{R}}$  e sia  $B = [x_1, x_2] \times [y_1, y_2]$ , con  $x_1 \leq x_2$  e  $y_1 \leq y_2$ , un rettangolo i cui vertici si trovano in  $S_1 \times S_2$ . Data una funzione  $H : S_1 \times S_2 \rightarrow \mathbb{R}$ , definiamo l' $H$ -volume di  $B$  come

$$V_H(B) = H(x_2, y_2) - H(x_1, y_2) - H(x_2, y_1) + H(x_1, y_1).$$

■

**Definizione 2.1.2.** Diciamo che la funzione  $H$  è 2-crescente se

$$V_H(B) \geq 0 \tag{2.1}$$

per ogni rettangolo  $B$  i cui vertici appartengono a  $\text{Dom}(H)$ .

■



Si osservi che la proprietà di 2-crescenza è tipica delle funzioni di ripartizione. Assumiamo per esempio che  $H$  sia la funzione di ripartizione della variabile aleatoria bivariata  $(X, Y)$  con valori in  $\mathbb{R}^2$ , ovvero

$$H(x, y) = \mathbb{P}(X \leq x, Y \leq y), \quad x, y, \in \mathbb{R}.$$

Si verifica immediatamente che, per il generico rettangolo  $B = [x_1, x_2] \times [y_1, y_2]$ , si ha

$$V_H(B) = \mathbb{P}(x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2) \geq 0.$$

Tuttavia, è altresì evidente come la proprietà di 2-crescenza della funzione  $H$  non sia sufficiente a concludere che  $H$  è una funzione di ripartizione. Infatti, perché  $H$  possa essere una funzione di ripartizione, deve essere non decrescente in ciascuno dei due argomenti. Tale proprietà non è però implicata dalla 2-crescenza. Affinché sia verificata anche la non decrescenza rispetto a entrambi gli argomenti di  $H$ , è necessario imporre un'ulteriore ipotesi che è descritta nel risultato seguente.

**Lemma 2.1.1.** *Siano  $S_1$  ed  $S_2$  due sottinsiemi non vuoti di  $\overline{\mathbb{R}}$  e  $H$  una funzione 2-crescente con dominio  $\text{Dom}(H) = S_1 \times S_2$ . Se  $S_1$  e  $S_2$  hanno due elementi minimi, che indichiamo rispettivamente con  $a_1$  e  $a_2$ , tali che*

$$H(a_1, y) = H(x, a_2) = 0 \quad \forall (x, y) \in S_1 \times S_2 \quad (2.2)$$

*allora  $H$  è non decrescente rispetto ad entrambi i suoi argomenti.*

*Dimostrazione.* Scrivendo per esteso la disuguaglianza (2.1) per  $x_1 = a_1$  si ottiene  $H(x_2, y_2) \geq H(x_2, y_1)$ . Dato che tale disuguaglianza vale per ogni  $x_2$  quando  $y_2 \geq y_1$ , si può immediatamente concludere che  $H$  è non decrescente rispetto a  $y$  per  $x$  fissato. La non decrescenza rispetto a  $x$  si dimostra in maniera analoga ponendo  $y_1 = a_2$  nella (2.1).  $\square$

Concludiamo la sezione con un risultato che riguarda il legame di una funzione di due variabili con le sue marginali.

**Lemma 2.1.2.** *Siano  $S_1$  ed  $S_2$  due sottinsiemi non vuoti di  $\overline{\mathbb{R}}$  e sia  $H : S_1 \times S_2 \rightarrow \mathbb{R}$  una funzione 2-crescente che soddisfa l'ipotesi (2.2). Siano  $(x_1, y_1)$  e  $(x_2, y_2)$  due punti di  $S_1 \times S_2$ . Se  $+\infty \in S_1 \cap S_2$  allora le marginali*

$$F(x) = H(x, \infty), \quad G(y) = H(\infty, y)$$

*soddisfano la disuguaglianza*

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|.$$

*Dimostrazione.* Assumiamo, senza perdita di generalità, che  $x_1 \leq x_2$  e  $y_1 \leq y_2$ . Dato che  $H$  è 2-crescente possiamo scrivere

$$H(x_2, y_1) - H(x_1, y_1) \leq H(x_2, y_2) - H(x_1, y_2). \quad (2.3)$$

Ponendo  $y_2 = \infty$  la formula (2.3) implica che

$$|H(x_2, y_1) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)|. \quad (2.4)$$

In maniera analoga, si può mostrare che

$$|H(x_2, y_2) - H(x_2, y_1)| \leq |G(y_2) - G(y_1)|. \quad (2.5)$$

Applicando la disuguaglianza triangolare e le disuguaglianze (2.4) e (2.5) si ottiene quindi

$$\begin{aligned} |H(x_2, y_2) - H(x_1, y_1)| &\leq |H(x_2, y_2) - H(x_2, y_1)| + |H(x_2, y_1) - H(x_1, y_1)| \\ &\leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|. \end{aligned}$$

□

## 2.2 Copule e teorema di Sklar

Nella sezione precedente abbiamo mostrato i principali risultati che saranno necessari nello studio delle copule. Possiamo a questo punto introdurre le copule iniziando dal concetto di subcopula.

**Definizione 2.2.1.** Una funzione di due variabili  $C'$  si dice subcopula se gode delle seguenti proprietà:

- (i)  $\text{Dom}(C') = S_1 \times S_2$  dove  $S_1$  e  $S_2$  sono sottinsiemi dell'intervallo  $[0, 1]$  entrambi contenenti i punti 0 e 1
- (ii)  $C'$  è 2-crescente e vale la relazione

$$C'(u, 0) = C'(0, v) = 0 \quad \forall (u, v) \in S_1 \times S_2$$

- (iii) per ogni  $u \in S_1$ ,  $v \in S_2$  si ha che

$$C'(u, 1) = u \quad \text{e} \quad C'(1, v) = v.$$

■

Si noti che il punto (ii) della definizione 2.2.1 non è altro che la proprietà (2.2) per  $a_1 = a_2 = 0$ . Pertanto, dalla (ii) segue che una subcopula è una funzione non decrescente di entrambi i suoi argomenti in virtù del Lemma 2.1.1.

La definizione di subcopula generalizza quella di copula. Quest'ultima è infatti una subcopula avente per dominio il quadrato  $[0, 1]^2$ . La definizione per esteso è riportata di seguito.

**Definizione 2.2.2.** Una funzione di due variabili  $C : [0, 1]^2 \rightarrow [0, 1]$  si dice copula se gode delle seguenti proprietà:

- (i)  $C$  è 2-crescente e vale la relazione

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

- (ii) per ogni  $(u, v) \in [0, 1]^2$  si ha che

$$C(u, 1) = u \quad \text{e} \quad C(1, v) = v.$$



È opportuno sottolineare che una copula non è altro che una funzione di ripartizione avente marginali uniformi, come si può facilmente verificare a partire dalla definizione 2.2.2.

Abbiamo a questo punto a disposizione tutti gli elementi necessari a dimostrare il teorema di Sklar, che costituisce la base del fondamentale ruolo ricoperto dalle copule in ambito statistico per lo studio delle distribuzioni multivariate. Il teorema sarà dimostrato sfruttando alcuni lemmi preliminari di seguito riportati. Per i risultati di questa sezione facciamo riferimento, ancora una volta, al testo di Nelsen [54].

**Lemma 2.2.1.** *Sia  $H$  una funzione di ripartizione bivariata con marginali  $F$  e  $G$ . Allora esiste ed è unica la subcopula  $C'$  tale che*

1.  $\text{Dom}(C') = \text{Im}(F) \times \text{Im}(G)$
2.  $H(x, y) = C'(F(x), G(y)) \quad \forall x, y, \in \overline{\mathbb{R}}$ .

*Dimostrazione.* Comunque vengano scelti  $x, y \in \overline{\mathbb{R}}$ , è possibile associare alla coppia  $(F(x), G(y))$  il valore  $H(x, y)$ . Inoltre, si può verificare che tale associazione è una funzione

$$(F(x), G(y)) \mapsto H(x, y). \quad (2.6)$$

Ciò segue dal fatto che  $H$  soddisfa le ipotesi del Lemma 2.1.2 e che quindi, scelti due punti  $(x_1, y_1)$  e  $(x_2, y_2)$ , se  $F(x_1) = F(x_2)$  e  $G(y_1) = G(y_2)$  si deve avere  $H(x_1, y_1) = H(x_2, y_2)$ . Pertanto, abbiamo dimostrato l'esistenza della funzione  $C'$  tale che

$$H(x, y) = C'(F(x), G(y))$$

e il cui dominio è, per costruzione,  $\text{Dom}(C') = \text{Im}(F) \times \text{Im}(G)$ . Per concludere la dimostrazione è sufficiente mostrare che  $C'$  è una subcopula. Ciò si verifica facilmente sfruttando il fatto che  $H$  è una funzione di ripartizione.  $\square$

**Lemma 2.2.2.** *Sia  $C'$  una subcopula. Allora esiste una copula  $C$  tale che  $C(u, v) = C'(u, v)$  per ogni  $(u, v) \in \text{Dom}(C')$ .*

*Dimostrazione.* Assumiamo che  $\text{Dom}(C') = S_1 \times S_2$ . Verifichiamo innanzitutto che, sfruttando l'uniforme continuità di  $C'$  nel suo dominio, si può estendere la subcopula  $C'$  a una subcopula  $C''$  avente dominio  $\text{Dom}(C'') = \overline{S_1} \times \overline{S_2}$ , dove  $\overline{S_1}$  rappresenta la chiusura di  $S_1$  e  $\overline{S_2}$  la chiusura di  $S_2$ . Infatti, dato un punto  $(u, v) \in (\overline{S_1} \times \overline{S_2}) \setminus (S_1 \times S_2)$ , esiste una successione di punti  $(u_n, v_n) \in S_1 \times S_2$ ,  $n \in \mathbb{N}$ , che converge a  $(u, v)$ . La successione è quindi di Cauchy e di conseguenza, per l'uniforme continuità di  $C'$ , deve essere di Cauchy anche la successione  $C'(u_n, v_n)$ ,  $n \in \mathbb{N}$ . Pertanto esiste il limite  $\lim_{n \rightarrow +\infty} C'(u_n, v_n)$  da cui segue che  $C'$  può essere estesa alla funzione  $C''$  cercata. Inoltre, si verifica facilmente che  $C''$  è una subcopula.

Il passo successivo è quello di estendere la subcopula  $C''$  a una copula  $C$  con dominio  $\text{Dom}(C) = [0, 1]^2$ . L'estensione viene effettuata mediante un'opportuna interpolazione. Si consideri un generico punto  $(a, b) \in [0, 1]^2$ . Siano  $a_1$  e  $a_2$  rispettivamente il massimo e il minimo elemento di  $\overline{S_1}$  tali che valga la relazione  $a_1 \leq a \leq a_2$ . Analogamente, siano  $b_1$  e  $b_2$  rispettivamente il massimo e il minimo elemento di  $\overline{S_2}$

tali che valga la relazione  $b_1 \leq b \leq b_2$ . Chiaramente, se  $a \in \overline{S_1}$  si ha che  $a_1 = a_2 = a$  e se  $b \in \overline{S_2}$  si ha che  $b_1 = b_2 = b$ . Ponendo ora

$$\lambda_1 = \begin{cases} \frac{a-a_1}{a_2-a_1} & \text{se } a_1 < a_2 \\ 1 & \text{se } a_1 = a_2 \end{cases} \quad \mu_1 = \begin{cases} \frac{b-b_1}{b_2-b_1} & \text{se } b_1 < b_2 \\ 1 & \text{se } b_1 = b_2 \end{cases}$$

si definisca

$$\begin{aligned} C(a, b) = & (1 - \lambda_1)(1 - \mu_1)C''(a_1, b_1) + (1 - \lambda_1)\mu_1C''(a_1, b_2) \\ & + \lambda_1(1 - \mu_1)C''(a_2, b_1) + \lambda_1\mu_1C''(a_2, b_2). \end{aligned} \quad (2.7)$$

Risulta immediato verificare che  $\text{Dom}(C) = [0, 1]^2$  e che, per ogni  $(a, b) \in \text{Dom}(C'')$ , si ha  $C(a, b) = C''(a, b)$ . Occorre quindi dimostrare che  $C$  è una copula. In particolare, per  $a \leq c$  e  $b \leq d$ , mostriamo che  $V_C([a, c] \times [b, d]) \geq 0$ . Sfruttando la definizione (2.7) e scrivendo in forma esplicita tutti i termini, si ha che

$$\begin{aligned} V_C([a, c] \times [b, d]) = & C(c, d) - C(a, d) - C(c, b) + C(a, b) \\ = & (1 - \lambda_1)\mu_2V_C([a_1, a_2] \times [d_1, d_2]) + \mu_2V_C([a_2, c_1] \times [d_1, d_2]) \\ & + \lambda_2\mu_2V_C([c_1, c_2] \times [d_1, d_2]) + (1 - \lambda_1)V_C([a_1, a_2] \times [b_2, d_1]) \\ & + V_C([a_2, c_1] \times [b_2, d_1]) + \lambda_2V_C([c_1, c_2] \times [b_2, d_1]) \\ & + (1 - \lambda_1)(1 - \mu_1)V_C([a_1, a_2] \times [b_1, b_2]) + (1 - \mu_1)V_C([a_2, c_1] \times [b_1, b_2]) \\ & + \lambda_2(1 - \mu_1)V_C([c_1, c_2] \times [b_1, b_2]) \geq 0 \end{aligned}$$

dove  $c_1, d_1, c_2, d_2, \lambda_2, \mu_2$  svolgono per  $c$  e  $d$  lo stesso ruolo svolto da  $a_1, b_1, a_2, b_2, \lambda_1, \mu_1$  per  $a$  e  $b$ . Le rimanenti proprietà che rendono la funzione  $C$  una copula si verificano direttamente attraverso la definizione (2.7).  $\square$

Il risultato appena dimostrato richiede alcuni commenti. In particolare, si è visto che ogni subcopula può essere estesa a una copula mediante un'opportuna interpolazione. Va tuttavia sottolineato il fatto che la scelta del metodo di interpolazione è arbitraria. Seguendo la linea proposta da Nelsen [54], abbiamo dimostrato il Lemma 2.2.2 effettuando un'interpolazione bilineare. Se si utilizzasse un metodo differente sarebbe di conseguenza differente la copula  $C$  risultante. Si può quindi concludere che, in generale, la copula che estende una data subcopula non è unica. Alla luce di tali considerazioni si può enunciare il risultato seguente.

**Teorema 2.2.3** (Teorema di Sklar). *Sia  $H$  una funzione di ripartizione bivariata con marginali  $F$  e  $G$ . Allora esiste una copula  $C$  tale che*

$$H(x, y) = C(F(x), G(y)) \quad (2.8)$$

per ogni  $x, y$  in  $\overline{\mathbb{R}}$ . Se  $F$  e  $G$  sono continue allora la copula  $C$  è unica, altrimenti è univocamente determinata sull'insieme  $\text{Im}(F) \times \text{Im}(G)$ . Inoltre, se  $C$  è una copula e  $F$  e  $G$  sono due funzioni di ripartizione, allora la funzione  $H$  definita dalla formula (2.8) è una funzione di ripartizione.

*Dimostrazione.* L'esistenza della copula  $C$  che soddisfa la relazione (2.8) segue immediatamente dai Lemmi 2.2.1 e 2.2.2. Inoltre, se  $F$  e  $G$  sono continue, la subcopula relativa al Lemma 2.2.1 è una copula da cui segue l'unicità di  $C$ . La verifica dell'ultima parte dell'enunciato è banale.  $\square$

La relazione (2.8) consente di decomporre una funzione di ripartizione multivariata distinguendo le componenti che descrivono le distribuzioni marginali dalla componente di copula che descrive la dipendenza tra le variabili aleatorie. L'importanza del teorema di Sklar risiede nel fatto che, in virtù della rappresentazione (2.8), le distribuzioni marginali e la componente di copula possono essere studiate disgiuntamente. Nella pratica, nell'ambito della modellizzazione delle dipendenze tra variabili aleatorie, si è generalmente interessati allo studio della sola componente di copula. Mentre la forma di una distribuzione multivariata è il risultato della combinazione delle distribuzioni marginali con la copula, il teorema di Sklar consente di isolare quest'ultima trascurando l'effetto delle marginali sulla forma complessiva della distribuzione. Allo scopo di isolare la componente di copula di una distribuzione, vengono generalmente utilizzati opportuni corollari del teorema di Sklar, per introdurre i quali è innanzitutto necessario definire il concetto di pseudo-inversa di una funzione di ripartizione.

**Definizione 2.2.3.** Sia  $F : \overline{\mathbb{R}} \rightarrow [0, 1]$  la funzione di ripartizione di una variabile aleatoria. La pseudo-inversa di  $F$  è la funzione  $F^{(-1)} : [0, 1] \rightarrow \overline{\mathbb{R}}$  definita come

$$F^{(-1)}(t) = \begin{cases} -\infty & \text{se } t = 0 \\ \inf\{x \in \mathbb{R} : F(x) \geq t\} & \text{se } t \in (0, 1) \\ +\infty & \text{se } t = 1. \end{cases}$$

■

La seguente proposizione illustra una proprietà fondamentale della pseudo-inversa di una funzione di ripartizione continua.

**Proposizione 2.2.4.** Sia  $F$  la funzione di ripartizione di una variabile aleatoria continua e sia  $F^{(-1)}$  la sua pseudo-inversa. Per ogni  $t \in [0, 1]$ , si ha che

$$F(F^{(-1)}(t)) = t.$$

*Dimostrazione.* Assumiamo che  $t \in (0, 1)$  dato che la dimostrazione sarebbe altrimenti banale. Si verifica facilmente che

$$\inf\{x \in \mathbb{R} : F(x) \geq t\} \leq \sup\{x \in \mathbb{R} : F(x) \leq t\}. \quad (2.9)$$

Inoltre, per ogni  $t \in (0, 1)$ , l'insieme  $\{x \in \mathbb{R} : F(x) \geq t\}$  è inferiormente limitato ed è chiuso perché  $F$  è continua, per cui deve contenere il suo estremo inferiore. Analogamente, l'insieme  $\{x \in \mathbb{R} : F(x) \leq t\}$  è chiuso e superiormente limitato e deve quindi contenere il suo estremo superiore. Possiamo pertanto riscrivere la disuguaglianza (2.9) nella forma

$$\min\{x \in \mathbb{R} : F(x) \geq t\} = F^{(-1)}(t) \leq \max\{x \in \mathbb{R} : F(x) \leq t\}. \quad (2.10)$$

Tenendo conto del fatto che  $F$  è non decrescente e applicandola a tutti i membri della disuguaglianza (2.10), si verifica che

$$t \leq F(\min\{x \in \mathbb{R} : F(x) \geq t\}) = F(F^{(-1)}(t)) \leq F(\max\{x \in \mathbb{R} : F(x) \leq t\}) \leq t.$$

da cui segue immediatamente la tesi. □

Dimostriamo di seguito un noto corollario del teorema di Sklar.

**Corollario 2.2.5.** *Sia  $H$  una funzione di ripartizione bivariata con marginali continue  $F$  e  $G$  e sia  $C$  la copula tale che*

$$H(x, y) = C(F(x), G(y))$$

per ogni  $x, y \in \overline{\mathbb{R}}$ . Allora si ha che

$$C(u, v) = H(F^{(-1)}(u), G^{(-1)}(v)), \quad \forall u, v \in [0, 1].$$

*Dimostrazione.* Si verifica immediatamente che

$$H(F^{(-1)}(u), G^{(-1)}(v)) = C(F(F^{(-1)}(u)), G(G^{(-1)}(v))) = C(u, v)$$

dove è stato utilizzato il risultato della proposizione 2.2.4 nell'ultimo passaggio.  $\square$

Il risultato appena mostrato, oltre a descrivere come si possa ricavare esplicitamente la copula sottostante a una data distribuzione multivariata, fornisce di fatto un metodo per la costruzione di copule. Infatti, è sufficiente costruire una distribuzione multivariata con marginali qualunque per poi applicare il corollario 2.2.5 e ricavare la relativa copula. Come vedremo, con questo metodo si ricavano, ad esempio, le copule relative alle distribuzioni ellittiche.

Illustriamo ora un'ulteriore proprietà delle pseudo-inverse di funzioni di ripartizione. Si può dimostrare che, se una data funzione di ripartizione  $F$  è invertibile, la sua pseudo-inversa coincide con l'inversa. In generale, la funzione di ripartizione  $F$  non è necessariamente invertibile perchè può essere costante su alcuni intervalli. Inoltre, si verifica facilmente che, in corrispondenza di tali intervalli, la pseudo-inversa  $F^{(-1)}$  presenta delle discontinuità di salto. Per un dato  $x \in \mathbb{R}$ , non è quindi vero in generale che  $F^{(-1)}(F(x)) = x$ . Vale tuttavia il seguente risultato.

**Teorema 2.2.6.** *Sia  $X$  una variabile aleatoria continua con funzione di ripartizione  $F$ . Indicando con  $F^{(-1)}$  la pseudo-inversa di  $F$ , si ha che*

$$\mathbb{P}\left(F^{(-1)}(F(X)) = X\right) = 1.$$

*Dimostrazione.* Osserviamo innanzitutto che la funzione  $F^{(-1)}$  è strettamente monotona crescente ed è pertanto invertibile. In virtù della proposizione 2.2.4, l'inversa di  $F^{(-1)}$  è  $F$ . Ciò significa che, restringendo la funzione  $F$  alla sola immagine di  $F^{(-1)}$ , possiamo scrivere che

$$F^{(-1)}(F(x)) = x, \quad \forall x \in \text{Im}\left(F^{(-1)}\right). \quad (2.11)$$

Indichiamo con  $D$  l'insieme dei punti di discontinuità della funzione  $F^{(-1)}$ . Vogliamo ora dimostrare che, per  $x \in \mathbb{R}$ ,

$$x \notin \text{Im}\left(F^{(-1)}\right) \implies F(x) \in D. \quad (2.12)$$

Infatti, se  $x \notin \text{Im}(F^{(-1)})$  allora è vero che

$$\nexists t : \min\{z \in \mathbb{R} : F(z) \geq t\} = x$$

o equivalentemente

$$\min\{z \in \mathbb{R} : F(z) \geq t\} \neq x, \quad \forall t \in (0, 1)$$

che implica a sua volta che

$$\left(F(x) < t\right) \vee \left(F(x) > t\right) \vee \left(F(x) = t \wedge \exists \varepsilon > 0 : F(x - \varepsilon) \geq t\right), \quad \forall t \in (0, 1). \quad (2.13)$$

Si consideri quindi un punto  $x \in \mathbb{R}$  tale che  $x \notin \text{Im}(F^{(-1)})$  e si ponga  $t = F(x)$ . In virtù della relazione (2.13), deve essere vero che per qualche  $\varepsilon > 0$  si ha  $F(x - \varepsilon) \geq t$ . Pertanto, dato che  $F$  è non decrescente, deve essere costante sull'intervallo  $[x - \varepsilon, x]$ . Si può verificare di conseguenza che  $F(x)$  è un punto di discontinuità per la funzione  $F^{(-1)}$ . Abbiamo quindi mostrato che vale l'implicazione (2.12). Da questa segue ovviamente che

$$\mathbb{P}\left(X \notin \text{Im}\left(F^{(-1)}\right)\right) \leq \mathbb{P}(F(X) \in D). \quad (2.14)$$

Inoltre, dato che  $F^{(-1)}$  è monotona, l'insieme  $D$  dei suoi punti di discontinuità deve essere al più numerabile (si veda il teorema 4.30 di Rudin [57]) e deve quindi avere misura di Lebesgue nulla. Dato che  $X$  è una variabile aleatoria continua, la sua distribuzione è assolutamente continua rispetto alla misura di Lebesgue. Di conseguenza, si può dimostrare che  $\mathbb{P}(F(X) \in D) = 0$  che implica, in virtù della (2.14), che

$$\mathbb{P}\left(X \in \text{Im}\left(F^{(-1)}\right)\right) = 1.$$

Ciò permette di concludere che

$$\begin{aligned} \mathbb{P}\left(F^{(-1)}(F(X)) = X\right) &= \mathbb{P}\left(F^{(-1)}(F(X)) = X, X \in \text{Im}\left(F^{(-1)}\right)\right) \\ &= \mathbb{P}\left(X = X, X \in \text{Im}\left(F^{(-1)}\right)\right) = 1 \end{aligned}$$

dove abbiamo utilizzato, nel passaggio dal secondo al terzo membro dell'uguaglianza, la relazione (2.11). Il teorema è quindi dimostrato.  $\square$

Nel seguito della trattazione, diverrà evidente come le idee utilizzate nella dimostrazione del teorema 2.2.6 siano rilevanti ai fini dello studio delle dipendenze non monotone qui presentato. Per il momento, ci limitiamo a dimostrare il seguente corollario del teorema di Sklar e del teorema 2.2.6.

**Corollario 2.2.7.** *Siano  $X$  e  $Y$  due variabili aleatorie continue con distribuzione congiunta  $H$  e marginali rispettivamente  $F$  e  $G$ . Sia  $C$  la copula tale che*

$$H(x, y) = C(F(x), G(y)).$$

Allora la distribuzione congiunta del vettore  $(F(X), G(Y))$  è

$$\mathbb{P}(F(X) \leq u, G(Y) \leq v) = C(u, v), \quad u, v \in [0, 1].$$

*Dimostrazione.* Se  $F$  e  $G$  sono invertibili, la dimostrazione si effettua facilmente osservando che

$$\begin{aligned} \mathbb{P}(F(X) \leq u, G(Y) \leq v) &= \mathbb{P}(X \leq F^{-1}(u), Y \leq G^{-1}(v)) = H(F^{-1}(u), G^{-1}(v)) \\ &= C(F(F^{-1}(u)), G(G^{-1}(v))) = C(u, v). \end{aligned}$$

Se invece le funzioni di ripartizione non sono invertibili, per il teorema 2.2.6 le variabili  $X$  e  $Y$  assumono valori, con probabilità 1, su un insieme in cui le inverse esistono. La dimostrazione è quindi identica.  $\square$

Alla luce del corollario 2.2.7, è evidente come, data una variabile aleatoria continua  $X$  con funzione di ripartizione  $F$ , la variabile aleatoria  $F(X)$  abbia distribuzione uniforme. La trasformazione che consiste nell'applicare ad una variabile aleatoria continua la sua stessa funzione di ripartizione, riducendo quindi la sua distribuzione ad uniforme, è nota nella letteratura come *trasformazione integrale di probabilità*. Il corollario 2.2.7 stabilisce che la trasformazione integrale di probabilità, applicata alle componenti di un vettore aleatorio, trasforma le marginali in uniformi ma preserva la struttura di dipendenza, nel senso che mantiene invariata la copula che lega le variabili.

## 2.3 Confini di Fréchet-Hoeffding

Concludiamo questo capitolo studiando alcune proprietà elementari delle copule bivariate che, seppur apparentemente semplici, aprono la via ad alcuni interrogativi che sono alla base del presente elaborato. Iniziamo con il seguente noto risultato.

**Teorema 2.3.1.** *Sia  $C'$  una subcopula. Allora per ogni  $(u, v) \in \text{Dom}(C')$  si ha che*

$$\max(u + v - 1; 0) \leq C'(u, v) \leq \min(u, v).$$

*Dimostrazione.* Dato che  $C'$  è non decrescente in entrambi i suoi argomenti, deve valere che  $C'(u, v) \leq C'(u, 1) = u$  e  $C'(u, v) \leq C'(1, v) = v$  da cui segue che  $C'(u, v) \leq \min(u, v)$  per qualunque punto  $(u, v)$  nel dominio della subcopula. Inoltre, dato che  $V_{C'}([u, 1] \times [v, 1]) \geq 0$ , si deve avere che  $C'(u, v) \geq u + v - 1$ . Per concludere la dimostrazione è sufficiente osservare che  $C'(u, v) \geq 0$ .  $\square$

Si osservi che, dato che ogni copula è una subcopula, il teorema 2.3.1 vale anche per le copule. Il risultato appena mostrato stabilisce che per qualunque copula valgono un confine superiore

$$M(u, v) = \min(u, v) \tag{2.15}$$

e un confine inferiore

$$W(u, v) = \max(u + v - 1; 0) \tag{2.16}$$

noti rispettivamente come *confine superiore di Fréchet-Hoeffding* e *confine inferiore di Fréchet-Hoeffding*. Si verifica facilmente che entrambi i confini di Fréchet-Hoeffding  $M$  e  $W$  sono delle copule. Tuttavia, ciò che è particolarmente rilevante è il tipo di dipendenza che tali copule descrivono. Infatti, le copule  $M$  e  $W$  rappresentano rispettivamente il caso di perfetta dipendenza positiva e perfetta dipendenza negativa



tra due variabili aleatorie. Per chiarire questo punto, si consideri la coppia di variabili aleatorie  $(U, V)$ , aventi distribuzione uniforme sull'intervallo  $[0, 1]$ , definite come segue:

$$\begin{cases} U \sim \text{Unif}[0, 1] \\ V = U. \end{cases}$$

È evidente che le due variabili così definite sono perfettamente dipendenti in quanto legate da una funzione deterministica. Inoltre, tale dipendenza perfetta è *positiva* perché le due variabili sono legate da una funzione monotona crescente. Dato il valore di una delle due variabili aleatorie, il valore dell'altra è univocamente determinato. La funzione di ripartizione congiunta della coppia  $(U, V)$  è una copula, come si può evincere dal fatto che le marginali sono uniformi. In particolare, per  $u, v \in [0, 1]$ , tale copula è

$$\begin{aligned} C(u, v) &= \mathbb{P}(U \leq u, V \leq v) = \mathbb{P}(U \leq u, U \leq v) \\ &= \mathbb{P}(U \leq \min(u, v)) = \min(u, v) \end{aligned}$$

che coincide proprio con il confine superiore di Fréchet-Hoeffding. Analogamente, ipotizziamo che la coppia  $(U, V)$  sia definita come

$$\begin{cases} U \sim \text{Unif}(0, 1) \\ V = 1 - U. \end{cases}$$

In tal caso, la dipendenza perfetta è *negativa* perché le due variabili aleatorie sono una funzione decrescente dell'altra. La copula che le lega è, per  $u, v \in [0, 1]$ ,

$$\begin{aligned} C(u, v) &= \mathbb{P}(U \leq u, V \leq v) = \mathbb{P}(U \leq u, 1 - U \leq v) \\ &= \mathbb{P}(1 - v \leq U \leq u) = \max(u + v - 1, 0) \end{aligned}$$

che coincide con il confine inferiore di Fréchet-Hoeffding.

Alla luce di quanto abbiamo appena visto, è chiaro che, se due variabili uniformi esibiscono perfetta dipendenza positiva, la copula che le lega coincide con il confine superiore di Fréchet-Hoeffding. Per i nostri scopi, sono a questo proposito necessarie delle osservazioni. È infatti cruciale osservare che l'unica forma di perfetta dipendenza positiva tra due variabili uniformi è la dipendenza lineare che si ottiene ponendo  $U = V$ . Infatti, si ipotizzi che, per  $U$  e  $V$  uniformi, esista una funzione monotona crescente  $g$  tale che  $V = g(U)$ . Studiando la distribuzione di  $V$ , si verifica immediatamente che

$$\mathbb{P}(V \leq v) = \mathbb{P}(U \leq g^{-1}(v)) = g^{-1}(v), \quad v \in (0, 1).$$

È quindi del tutto evidente che, a meno che  $g$  non coincida con la funzione identità, viene contraddetta l'ipotesi che  $V$  abbia distribuzione uniforme. Possiamo pertanto concludere che la perfetta dipendenza lineare sia l'unica forma di perfetta dipendenza monotona tra due variabili uniformi. Simili conclusioni possono essere tratte nel caso di perfetta dipendenza negativa. Ne consegue che i confini di Fréchet-Hoeffding sono le uniche copule che possono descrivere la perfetta dipendenza monotona tra variabili uniformi. Il quesito fondamentale è se questo sia vero anche per distribuzioni più

generali, quando la perfetta dipendenza monotona non è più vincolata ad essere di tipo lineare e può assumere le forme più varie. Una risposta è immediatamente fornita dal seguente noto risultato (si veda, per esempio, Fréchet [22]).

**Teorema 2.3.2.** *Si considerino due variabili aleatorie  $X_1$  e  $X_2$  tali che  $X_2 = \phi(X_1)$  dove  $\phi$  è una funzione strettamente monotona crescente. Allora la distribuzione congiunta di  $(X_1, X_2)$  è*

$$F(x_1, x_2) = \min(F_{X_1}(x_1), F_{X_2}(x_2)).$$

Se  $\phi$  è strettamente monotona decrescente si ha invece

$$F(x_1, x_2) = \max(F_{X_1}(x_1) + F_{X_2}(x_2) - 1, 0).$$

*Dimostrazione.* Se  $\phi$  è monotona crescente si ha che

$$\begin{aligned} F(x_1, x_2) &= \mathbb{P}(X_1 \leq x_1, X_1 \leq \phi^{-1}(x_2)) = \mathbb{P}(X_1 \leq \min(x_1, \phi^{-1}(x_2))) \\ &= F_{X_1}(\min(x_1, \phi^{-1}(x_2))) = \min(F_{X_1}(x_1), F_{X_1}(\phi^{-1}(x_2))) \\ &= \min(F_{X_1}(x_1), F_{X_2}(x_2)) \end{aligned}$$

Per  $\phi$  monotona decrescente si può invece scrivere, per  $x_1 \geq \phi^{-1}(x_2)$ ,

$$\begin{aligned} F(x_1, x_2) &= \mathbb{P}(X_1 \leq x_1, X_1 \geq \phi^{-1}(x_2)) = \mathbb{P}(X_1 \leq x_1) - \mathbb{P}(X_1 \leq x_1, X_1 < \phi^{-1}(x_2)) \\ &= F_{X_1}(x_1) - \mathbb{P}(X_1 < \phi^{-1}(x_2)) = F_{X_1}(x_1) - \mathbb{P}(X_2 > x_2) \\ &= F_{X_1}(x_1) + F_{X_2}(x_2) - 1. \end{aligned}$$

Per  $x_1 < \phi^{-1}(x_2)$  si ottiene  $F(x_1, x_2) = 0$  che completa la dimostrazione.  $\square$

Il risultato appena mostrato conferma che, nel caso di dipendenze monotone, le copule sottostanti coincidono sempre con i confini di Fréchet-Hoeffding indipendentemente dalla forma della funzione deterministica che lega le marginali della distribuzione. A qualunque dipendenza perfetta di tipo monotono corrisponde quindi una copula che concentra l'intera massa di probabilità su una delle diagonali del quadrato unitario. In altre parole, le funzioni  $f(u) = u$  e  $f(u) = 1 - u$  sono le uniche due possibili curve monotone su cui una copula può concentrare la sua intera massa di probabilità.

Tutte le considerazioni fin qui effettuate sono basate sull'ipotesi che le variabili in esame siano tra loro legate da una perfetta dipendenza monotona. Viene dunque spontaneo chiedersi cosa accada se si indebolisce tale ipotesi ammettendo dipendenze perfette che non siano di tipo monotono. Proprio questo quesito è alla base del presente elaborato. Nel caso di dipendenze non monotone, il numero di possibili curve che descrivono una dipendenza perfetta è infinito. Come si vedrà, si possono infatti avere curve dalle forme più varie, purchè queste corrispondano a funzioni che preservano la distribuzione uniforme.

## Capitolo 3

# Misure di dipendenza

Le metodologie proposte in letteratura per la descrizione delle dipendenze tra variabili aleatorie sono molteplici. In linea generale, una proprietà desiderabile per una data misura di dipendenza è che questa sia invariante rispetto a trasformazioni monotone (Hoeffding [32, 33]). In virtù del teorema 2.3.2, possiamo concludere che le copule sono uno strumento in grado di cogliere il fatto che una trasformazione monotona non altera la dipendenza stocastica tra variabili aleatorie. Uno dei principali limiti delle copule sta nel fatto che il loro utilizzo è strettamente legato a un'ipotesi di tipo parametrico. In questo capitolo, descriviamo alcuni indici non parametrici per la misurazione della dipendenza tra due variabili. È opportuno sottolineare che le misure che esamineremo sono misure di concordanza. Informalmente, due variabili si dicono concordanti quando una tende ad assumere valori elevati in corrispondenza di valori elevati dell'altra, mentre tende ad assumere valori bassi in corrispondenza di valori bassi assunti dall'altra. Più precisamente, si considerino due realizzazioni  $(x_i, y_i)$  e  $(x_j, y_j)$  di una variabile aleatoria bidimensionale  $(X, Y)$ . Si dice che  $(x_i, y_i)$  e  $(x_j, y_j)$  sono concordanti se  $x_i < x_j$  e  $y_i < y_j$  oppure se  $x_i > x_j$  e  $y_i > y_j$ . In caso contrario si dice che le realizzazioni  $(x_i, y_i)$  e  $(x_j, y_j)$  sono discordanti. Equivalentemente,  $(x_i, y_i)$  e  $(x_j, y_j)$  sono concordanti se  $(x_i - x_j)(y_i - y_j) > 0$  e sono discordanti se  $(x_i - x_j)(y_i - y_j) < 0$ . È quindi evidente che tenderanno ad essere concordanti due variabili legate da una dipendenza di tipo monotono. Come già discusso, questa tipologia di dipendenza non esaurisce però l'insieme delle possibili dipendenze osservabili. Pertanto, come approfondiremo in seguito, le misure di concordanza si rivelano adatte solo in una ristretta minoranza dei casi. Concluderemo infine il capitolo introducendo il concetto di dipendenza di coda.

Nel presentare i risultati dei prossimi paragrafi, sarà fatto ampio uso di integrali che saranno da intendersi nel senso di Lebesgue-Stieltjes. Per giustificare l'uso dell'integrale di Lebesgue-Stieltjes, si osservi che ad una data copula  $C$  corrisponde una premisura che associa ad ogni rettangolo  $B \subseteq [0, 1]^2$  il suo  $C$ -volume  $V_C(B)$ . Applicando il teorema di estensione di Carathéodory e indicando con  $\mathcal{B}([0, 1]^2)$  la  $\sigma$ -algebra di Borel su  $[0, 1]^2$ , tale premisura può essere estesa ad una misura sullo spazio misurabile  $([0, 1]^2, \mathcal{B}([0, 1]^2))$ . Indicando per semplicità con il simbolo  $C$  sia la copula che la misura da essa generata, è quindi possibile definire uno spazio di misura  $([0, 1]^2, \mathcal{B}([0, 1]^2), C)$  e costruire su di esso l'integrale di Lebesgue-Stieltjes

con le usuali tecniche (si veda Athreya e Lahiri [4] maggiori dettagli). Pertanto, dati un insieme  $A \in \mathcal{B}([0, 1]^2)$  e una funzione  $g : [0, 1]^2 \rightarrow \mathbb{R}$  misurabile secondo Lebesgue, indichiamo l'integrale su  $A$  della funzione  $g$  rispetto alla misura  $C$  con la notazione

$$\int \int_A g(u, v) dC(u, v).$$

Nello studio della concordanza tra variabili aleatorie, ci avvaliamo innanzitutto del seguente risultato.

**Teorema 3.0.1.** *Siano  $(X_1, Y_1)$  e  $(X_2, Y_2)$  due vettori indipendenti di variabili aleatorie continue. Siano  $H_1$  e  $H_2$  le funzioni di ripartizione dei due vettori e si assuma che queste abbiano le stesse marginali  $F$  (per  $X_1$  e  $X_2$ ) e  $G$  (per  $Y_1$  e  $Y_2$ ). Siano  $C_1$  e  $C_2$  le copule di  $(X_1, Y_1)$  e  $(X_2, Y_2)$ , ovvero le funzioni tali che  $H_1(x, y) = C_1(F(x), G(y))$  e  $H_2(x, y) = C_2(F(x), G(y))$ .*

*Definiamo la funzione di concordanza*

$$Q = \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) < 0\right). \quad (3.1)$$

Allora vale la formula

$$Q = Q(C_1, C_2) = 4 \int \int_{[0,1]^2} C_2(u, v) dC_1(u, v) - 1. \quad (3.2)$$

*Dimostrazione.* Si ha che

$$\begin{aligned} Q &= \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) < 0\right) \\ &= 2 \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - 1 \\ &= 2 \mathbb{P}(X_1 > X_2, Y_1 > Y_2) + 2 \mathbb{P}(X_1 < X_2, Y_1 < Y_2) - 1. \end{aligned} \quad (3.3)$$

Osserviamo ora che

$$\begin{aligned} \mathbb{P}(X_1 > X_2, Y_1 > Y_2) &= \mathbb{P}(X_2 < X_1, Y_2 < Y_1) \\ &= \int \int_{\mathbb{R}^2} \mathbb{P}(X_2 < x, Y_2 < y) dC_1(F(x), G(y)) \\ &= \int \int_{\mathbb{R}^2} C_2(F(x), G(y)) dC_1(F(x), G(y)) \\ &= \int \int_{[0,1]^2} C_2(u, v) dC_1(u, v) \end{aligned} \quad (3.4)$$

dove nell'ultimo passaggio si è posto  $u = F(x)$  e  $v = G(y)$ . Analogamente, possiamo scrivere

$$\begin{aligned}
& \mathbb{P}(X_2 > X_1, Y_2 > Y_1) \\
&= \int \int_{\mathbb{R}^2} \mathbb{P}(X_2 > x, Y_2 > y) dC_1(F(x), G(y)) \\
&= \int \int_{\mathbb{R}^2} [1 - F(x) - G(y) + C_2(F(x), G(y))] dC_1(F(x), G(y)) \\
&= \int \int_{[0,1]^2} [1 - u - v + C_2(u, v)] dC_1(u, v) \\
&= 1 - \frac{1}{2} - \frac{1}{2} + \int \int_{[0,1]^2} C_2(u, v) dC_1(u, v) \\
&= \int \int_{[0,1]^2} C_2(u, v) dC_1(u, v) \tag{3.5}
\end{aligned}$$

dove abbiamo posto  $u = F(x)$  e  $v = G(y)$  e abbiamo tenuto in considerazione il fatto che  $C_1$  è la funzione di ripartizione di una coppia  $(U, V)$  di variabili uniformi tali che  $\mathbb{E}[U] = \mathbb{E}[V] = \frac{1}{2}$ . Sostituendo le formule (3.4) e (3.5) nella (3.3) si ottiene la formula (3.2). □

Prima di passare allo studio delle misure di concordanza nei paragrafi che seguono, è opportuno dedicare alcuni commenti al risultato appena dimostrato. Il teorema 3.0.1 stabilisce che, date due coppie di variabili aleatorie  $(X_1, Y_1)$  e  $(X_2, Y_2)$  aventi distribuzioni bivariate differenti con marginali comuni, la misura di concordanza  $Q$  definita dalla formula (3.1) è determinata dalle sole copule delle due coppie di variabili aleatorie e non dipende quindi dalle distribuzioni marginali, come si evince dalla formula (3.2). Viene tuttavia spontaneo chiedersi la ragione per cui sia necessario, nell'enunciato del teorema, assumere che le variabili aleatorie coinvolte siano continue. La risposta è data dal teorema di Sklar, che stabilisce che la copula che descrive il legame tra due variabili aleatorie è unica solo se le variabili aleatorie sono continue. Se quindi la misura di concordanza (3.1) è ben definita anche nel caso di variabili aleatorie che non siano continue, lo stesso non si può dire della (3.2). Infatti, nel caso di variabili che non siano continue, sorge un'ambiguità nella definizione della formula (3.2) legata al fatto che, nella rappresentazione (2.8) fornita dal teorema di Sklar per le funzioni di ripartizione, sono possibili diverse scelte per le copule  $C_1$  e  $C_2$  che condurrebbero a diversi valori della misura di concordanza (3.2). Allo scopo di evitare tale ambiguità, l'ipotesi che le variabili aleatorie considerate siano continue sarà adottata in diversi teoremi del presente capitolo.

### 3.1 Tau di Kendall

Consideriamo una coppia  $(X, Y)$  avente funzione di ripartizione congiunta  $H$ . Siano  $(X_1, Y_1)$  e  $(X_2, Y_2)$  due coppie indipendenti aventi entrambe funzione di ripartizione  $H$ . La misura di associazione tra  $X$  e  $Y$  nota come tau di Kendall può essere espressa come (Kruskal [44])

$$\tau_{X,Y} = \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) < 0\right). \tag{3.6}$$

Si può verificare immediatamente che  $-1 \leq \tau_{X,Y} \leq 1$ . Inoltre, se  $X$  e  $Y$  sono indipendenti, si deve avere che  $\tau_{X,Y} = 0$  perché  $X_1$  e  $X_2$  (o equivalentemente  $Y_1$  e  $Y_2$ ) sono intercambiabili. È infine evidente come il valore di  $\tau_{X,Y}$  non cambi se una tra  $X$  e  $Y$  viene trasformata mediante una funzione monotona.

La versione campionaria del tau di Kendall, storicamente precedente alla versione probabilistica (3.6) e introdotta da Kendall [40], si ottiene come segue. Si consideri il campione di  $n$  realizzazioni  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  della variabile  $(X, Y)$ . Ovviamente, è possibile selezionare  $\binom{n}{2}$  coppie di realizzazioni. Tra queste vi sarà un numero  $c$  di coppie concordanti e un numero  $d$  di coppie discordanti. La tau di Kendall campionaria si calcola come

$$t = \frac{c - d}{c + d} = \frac{c - d}{\binom{n}{2}}.$$

La misura di concordanza così costruita è non parametrica. Tuttavia, se si considera la variante teorica (3.6), si può verificare che il tau di Kendall è esprimibile in termini della copula che descrive la dipendenza tra le variabili  $X$  e  $Y$  se queste ultime sono continue. Vale infatti il seguente risultato.

**Teorema 3.1.1.** *Sia  $(X, Y)$  una coppia di variabili aleatorie continue la cui copula è  $C$ . Allora il tau di Kendall tra  $X$  e  $Y$  è dato da*

$$\tau_{X,Y} = 4 \int \int_{[0,1]^2} C(u, v) dC(u, v) - 1. \quad (3.7)$$

*Dimostrazione.* Indicando con  $Q$  la funzione di concordanza definita nel teorema 3.0.1 e tenendo in considerazione la definizione (3.6) della tau di Kendall, si verifica immediatamente che

$$\tau_{X,Y} = Q(C, C).$$

L'enunciato segue quindi immediatamente dalla formula (3.2).  $\square$

Si osservi che, nel caso in cui  $X$  e  $Y$  siano due variabili continue tra loro indipendenti, sostituendo la copula di indipendenza  $\Pi(u, v) = uv$  nella formula (3.2) si ricava ancora una volta che  $\tau_{X,Y} = 0$ . Inoltre, avendo espresso  $\tau_{X,Y}$  in termini della sola copula, è evidente come tale misura di concordanza sia invariante rispetto a trasformazioni monotone in virtù del teorema 2.3.2.

## 3.2 Rho di Spearman

Descriviamo ora una seconda misura di concordanza nota come rho di Spearman. Consideriamo la coppia di variabili aleatorie  $(X, Y)$  e siano  $(X_1, Y_1)$ ,  $(X_2, Y_2)$  e  $(X_3, Y_3)$  tre coppie tra loro indipendenti e tutte distribuite come  $(X, Y)$ . La versione probabilistica del rho di Spearman, ancora una volta introdotta da Kruskal [44], può essere espressa nella forma

$$\rho_{X,Y} = 3 \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_3) > 0\right) - 3 \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_3) < 0\right). \quad (3.8)$$

Come per il tau di Kendall, anche per il rho di Spearman la versione campionaria è storicamente precedente a quella probabilistica ed è stata introdotta da Spearman

[65]. Per costruirla, si consideri il campione di  $n$  coppie  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  e si indichino con  $r_j$  e  $s_j$  rispettivamente il rango di  $x_j$  e  $y_j$ , ovvero

$$r_j = \#\{k : x_j \geq x_k, k = 1, \dots, n\}, \quad s_j = \#\{k : y_j \geq y_k, k = 1, \dots, n\}.$$

Il rho di Spearman campionario si ottiene come

$$\rho = \frac{\sum_{j=1}^n (r_j - \bar{r})(s_j - \bar{s})}{\sqrt{\sum_{j=1}^n (r_j - \bar{r})^2} \sqrt{\sum_{j=1}^n (s_j - \bar{s})^2}} \quad (3.9)$$

dove  $\bar{s} = \frac{1}{n} \sum_{j=1}^n s_j$  e  $\bar{r} = \frac{1}{n} \sum_{j=1}^n r_j$ . In altre parole, il rho di Spearman non è altro che un coefficiente di correlazione lineare di Pearson calcolato sui ranghi delle variabili anzichè sui valori stessi delle variabili. Per chiarire il legame tra le formule (3.8) e (3.9), riportiamo il seguente risultato dal testo di Nelsen [54].

**Teorema 3.2.1.** *Siano  $X$  e  $Y$  due variabili aleatorie continue con copula  $C$ . Si ha che*

$$\begin{aligned} \rho_{X,Y} &= 3 Q(C, \Pi) = 12 \int \int_{[0,1]^2} uv \, dC(u, v) - 3 \\ &= 12 \int \int_{[0,1]^2} C(u, v) \, du \, dv - 3. \end{aligned} \quad (3.10)$$

*Dimostrazione.* L'enunciato si dimostra riformulando la (3.8) alla luce del teorema 3.0.1 e considerando che le variabili  $X_2$  e  $Y_3$  sono legate dalla copula di indipendenza  $\Pi$ . Per concludere si sfrutta la simmetria della funzione di concordanza da cui segue che  $Q(C, \Pi) = Q(\Pi, C)$ .  $\square$

Passiamo a questo punto all'interpretazione del teorema appena dimostrato. Supponiamo che  $U$  e  $V$  siano due variabili aleatorie uniformemente distribuite nell'intervallo  $(0, 1)$  con copula  $C$ . Dalla formula (3.10) segue che il rho di Spearman che lega  $U$  e  $V$  è

$$\begin{aligned} \rho &= 12 \int \int_{[0,1]^2} C(u, v) \, du \, dv - 3 = 12 \mathbb{E}[UV] - 3 \\ &= \frac{\mathbb{E}[UV] - 1/4}{1/12} = \frac{\mathbb{E}[UV] - \mathbb{E}[U] \mathbb{E}[V]}{\sqrt{\text{Var}[U] \text{Var}[V]}}. \end{aligned} \quad (3.11)$$

La formula (3.11) illustra che il rho di Spearman è il coefficiente di correlazione lineare che lega le distribuzioni uniformi corrispondenti con le marginali della copula. Seguono quindi alcune proprietà elementari, tra cui per esempio che  $-1 \leq \rho_{X,Y} \leq 1$  oppure che  $\rho_{X,Y} = 0$  se  $X$  e  $Y$  sono indipendenti. Diventa altresì evidente la ragione per cui la formula (3.9) può essere considerata la variante campionaria della (3.8). Sottolineiamo che l'interpretazione del rho di Spearman come un coefficiente di correlazione lineare tra variabili uniformi, è esemplificativa della ragione per cui tale misura di concordanza non è adatta a descrivere dipendenze non monotone. Queste, infatti, non rientrano ovviamente tra le dipendenze lineari.

Concludiamo questo paragrafo facendo presente che, nella pratica, si preferisce in

genere utilizzare uno stimatore campionario del rho di Spearman diverso da quello indicato nella formula (3.9). In particolare, si usa la formula

$$\rho = 1 - \frac{6 \sum_{j=1}^n (r_j - s_j)^2}{n(n^2 - 1)}.$$

Per una giustificazione all'utilizzo di quest'ultima formula si rimanda alla lettura dell'articolo di Kruskal [44].

### 3.3 Relazione tra tau di Kendall e rho di Spearman

Nei paragrafi precedenti, abbiamo illustrato due diverse misure di concordanza. Naturalmente, essendo tali misure costruite in modo diverso, non ci si aspetta che coincidano per una data distribuzione bivariata. In generale non esiste nemmeno una relazione biunivoca tra il tau di Kendall e il rho di Spearman. In altre parole, nota una delle due misure, non si può sapere in maniera esatta il valore dell'altra. Esistono però alcune combinazioni di  $\tau$  e di  $\rho$  che non è possibile ottenere in nessun caso. Intuitivamente ciò significa che, per quanto diversi possano essere i due coefficienti, questi non potranno essere troppo in contraddizione tra loro. Lo scopo del presente capitolo è quello di presentare i risultati che permettono di definire una regione ammissibile di coppie  $(\tau, \rho)$  che è possibile ottenere per qualche distribuzione. Cominciamo con il seguente risultato ottenuto per la prima volta da Daniels [14].

**Teorema 3.3.1.** *Siano  $X$  e  $Y$  due variabili aleatorie continue e siano  $\tau$  e  $\rho$  rispettivamente il tau di Kendall e il rho di Spearman che le legano. Allora si ha*

$$-1 \leq 3\tau - 2\rho \leq 1.$$

*Dimostrazione.* Siano  $(X_1, Y_1)$ ,  $(X_2, Y_2)$  e  $(X_3, Y_3)$  tre coppie tra loro indipendenti e tutte distribuite come  $(X, Y)$ . Si ha che

$$\tau = 2 \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - 1$$

$$\rho = 6 \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_3) > 0\right) - 3.$$

Dato che i vettori  $(X_1, Y_1)$ ,  $(X_2, Y_2)$  e  $(X_3, Y_3)$  sono intercambiabili ai fini della definizione di  $\tau$  e  $\rho$ , possiamo considerare tutte le diverse possibili combinazioni dei pedici e scrivere

$$\tau = \frac{2}{3} \left\{ \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) + \mathbb{P}\left((X_2 - X_3)(Y_2 - Y_3) > 0\right) + \mathbb{P}\left((X_3 - X_1)(Y_3 - Y_1) > 0\right) \right\} - 1$$

e

$$\begin{aligned} \rho = & \left\{ \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_3) > 0\right) + \mathbb{P}\left((X_1 - X_3)(Y_1 - Y_2) > 0\right) \right. \\ & + \mathbb{P}\left((X_2 - X_1)(Y_2 - Y_3) > 0\right) + \mathbb{P}\left((X_2 - X_3)(Y_2 - Y_1) > 0\right) \\ & \left. + \mathbb{P}\left((X_3 - X_1)(Y_3 - Y_2) > 0\right) + \mathbb{P}\left((X_3 - X_2)(Y_3 - Y_1) > 0\right) \right\} - 3. \end{aligned}$$



Le espressioni così costruite sono invarianti per permutazioni degli indici. Pertanto, applicando il teorema delle probabilità totali e condizionando rispetto a tutti i possibili ordinamenti delle variabili  $X_1$ ,  $X_2$  e  $X_3$ , si può dimostrare che i valori di  $\tau$  e  $\rho$  non cambiano se le probabilità vengono condizionate ad uno specifico ordinamento. Possiamo assumere senza perdita di generalità che  $X_1 < X_2 < X_3$  da cui segue che

$$\tau = \frac{2}{3} \left\{ \mathbb{P}(Y_1 < Y_2) + \mathbb{P}(Y_2 < Y_3) + \mathbb{P}(Y_1 < Y_3) \right\} - 1$$

e

$$\begin{aligned} \rho &= \left\{ \mathbb{P}(Y_1 < Y_3) + \mathbb{P}(Y_1 < Y_2) + \mathbb{P}(Y_2 > Y_3) \right. \\ &\quad \left. + \mathbb{P}(Y_2 < Y_1) + \mathbb{P}(Y_3 > Y_2) + \mathbb{P}(Y_3 > Y_1) \right\} - 3 \\ &= 2 \mathbb{P}(Y_1 < Y_3) - 1. \end{aligned}$$

Indichiamo ora con  $p_{ijk}$  la probabilità che  $Y_i < Y_j < Y_k$ , mantenendo sempre come evento condizionante l'evento  $X_1 < X_2 < X_3$ . Si osservi che esistono sei diversi  $p_{ijk}$  e che la somma dei sei è pari a 1. Le espressioni per  $\tau$  e  $\rho$  possono essere riformulate come

$$\begin{aligned} \tau &= \frac{2}{3} \left\{ (p_{123} + p_{132} + p_{312}) + (p_{123} + p_{213} + p_{231}) + (p_{123} + p_{132} + p_{213}) \right\} - 1 \\ &= p_{123} + \frac{1}{3} (p_{132} + p_{213}) - \frac{1}{3} (p_{231} + p_{312}) - p_{321} \end{aligned}$$

e

$$\begin{aligned} \rho &= 2(p_{123} + p_{132} + p_{213}) - 1 \\ &= p_{123} + p_{132} + p_{213} - p_{231} - p_{312} - p_{321}. \end{aligned} \tag{3.12}$$

Segue quindi che

$$3\tau - 2\rho = (p_{123} + p_{231} + p_{312}) - (p_{132} + p_{321} + p_{213})$$

che conclude la dimostrazione.  $\square$

Il risultato che segue è stato dimostrato per la prima volta da Durbin e Stuart [17]

**Teorema 3.3.2.** *Siano  $X$ ,  $Y$ ,  $\tau$  e  $\rho$  definite come nel teorema 3.3.1. Allora si ha che*

$$\frac{1 + \rho}{2} \geq \left( \frac{1 + \tau}{2} \right)^2 \tag{3.13}$$

e

$$\frac{1 - \rho}{2} \geq \left( \frac{1 - \tau}{2} \right)^2. \tag{3.14}$$

*Dimostrazione.* Siano  $(X_1, Y_1)$ ,  $(X_2, Y_2)$  e  $(X_3, Y_3)$  tre coppie tra loro indipendenti e tutte distribuite come  $(X, Y)$  e aventi funzione di ripartizione  $H$ . Indichiamo con

$p$  la probabilità che le coppie  $(X_2, Y_2)$  e  $(X_3, Y_3)$  siano entrambe concordanti con  $(X_1, Y_1)$ . Si ha che

$$\begin{aligned}
p &= \mathbb{P}\left((X_2 - X_1)(Y_2 - Y_1) > 0, (X_3 - X_1)(Y_3 - Y_1) > 0\right) \\
&= \mathbb{E}\left[\mathbb{P}\left((X_2 - X_1)(Y_2 - Y_1) > 0, (X_3 - X_1)(Y_3 - Y_1) > 0 \mid X_1, Y_1\right)\right] \\
&= \int \int_{\mathbb{R}^2} \mathbb{P}\left((X_2 - x)(Y_2 - y) > 0, (X_3 - x)(Y_3 - y) > 0\right) dH(x, y) \\
&= \int \int_{\mathbb{R}^2} \mathbb{P}\left((X_2 - x)(Y_2 - y) > 0\right) \mathbb{P}\left((X_3 - x)(Y_3 - y) > 0\right) dH(x, y) \\
&= \int \int_{\mathbb{R}^2} \mathbb{P}\left((X_2 - x)(Y_2 - y) > 0\right)^2 dH(x, y) \\
&\geq \left(\int \int_{\mathbb{R}^2} \mathbb{P}\left((X_2 - x)(Y_2 - y) > 0\right) dH(x, y)\right)^2 \\
&= \left(\mathbb{P}\left((X_2 - X_1)(Y_2 - Y_1) > 0\right)\right)^2 = \left(\frac{1 + \tau}{2}\right)^2
\end{aligned}$$

dove si è fatto uso della disuguaglianza di Jensen. Osserviamo inoltre che, permutando opportunamente i vettori  $(X_1, Y_1)$ ,  $(X_2, Y_2)$  e  $(X_3, Y_3)$ , possiamo scrivere

$$\begin{aligned}
p &= \frac{1}{3} \left\{ \mathbb{P}\left((X_2 - X_1)(Y_2 - Y_1) > 0, (X_3 - X_1)(Y_3 - Y_1) > 0\right) \right. \\
&\quad + \mathbb{P}\left((X_3 - X_2)(Y_3 - Y_2) > 0, (X_1 - X_2)(Y_1 - Y_2) > 0\right) \\
&\quad \left. + \mathbb{P}\left((X_1 - X_3)(Y_1 - Y_3) > 0, (X_2 - X_3)(Y_2 - Y_3) > 0\right) \right\}.
\end{aligned}$$

Procediamo a questo punto come nel teorema 3.3.1. Utilizzando opportunamente le probabilità condizionate, assumiamo senza perdita di generalità che  $X_1 < X_2 < X_3$  e indichiamo con  $p_{ijk}$  la probabilità che  $Y_i < Y_j < Y_k$  condizionatamente all'evento  $X_1 < X_2 < X_3$ . Segue che

$$\begin{aligned}
p &= \frac{1}{3} \left\{ (p_{123} + p_{132}) + p_{123} + (p_{123} + p_{213}) \right\} \\
&= p_{123} + \frac{1}{3} p_{132} + \frac{1}{3} p_{213}.
\end{aligned}$$

Sfruttando ora la formula (3.12) si ottiene

$$\frac{1 + \rho}{2} = p_{123} + p_{132} + p_{213} \geq p \geq \left(\frac{1 + \tau}{2}\right)^2$$

che completa la dimostrazione della formula (3.13).

Per dimostrare la (3.14), assumiamo ora che  $p$  la probabilità che le coppie  $(X_2, Y_2)$  e  $(X_3, Y_3)$  siano entrambe discordanti con  $(X_1, Y_1)$ . Seguendo passaggi simili a quelli effettuati fino ad ora, si dimostra che

$$p \geq \left(\mathbb{P}\left((X_2 - X_1)(Y_2 - Y_1) < 0\right)\right)^2 = \left(\frac{1 - \tau}{2}\right)^2.$$

Inoltre, si ha che

$$\begin{aligned} p &= \frac{1}{3} \{ (p_{231} + p_{321}) + p_{321} + (p_{312} + p_{321}) \} \\ &= p_{321} + \frac{1}{3} p_{231} + \frac{1}{3} p_{312}. \end{aligned}$$

Segue che

$$\frac{1-\rho}{2} = 1 - (p_{123} + p_{132} + p_{213}) = p_{321} + p_{231} + p_{312} \geq p \geq \left( \frac{1-\tau}{2} \right)^2$$

che completa la dimostrazione della (3.14).  $\square$

Abbiamo a questo punto tutti gli elementi necessari a definire la cosiddetta regione  $\tau - \rho$ , ovvero l'insieme delle possibili coppie  $(\tau, \rho)$ .

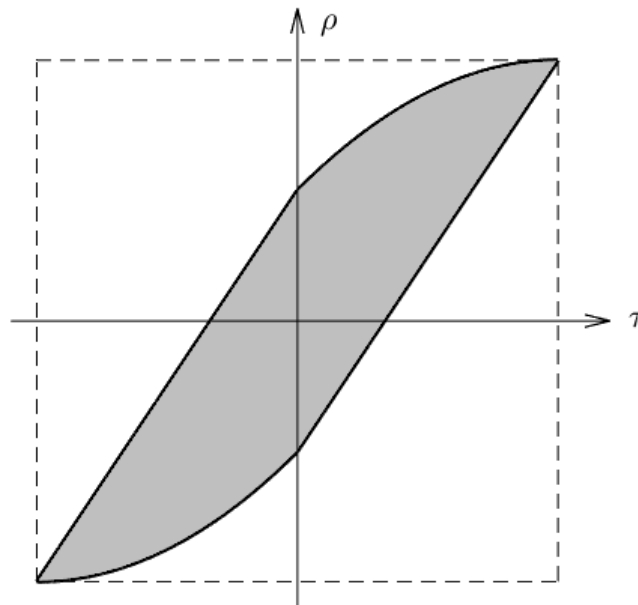
**Corollario 3.3.3.** *Siano  $X, Y, \tau$  e  $\rho$  definite come nel teorema 3.3.1. Si ha che*

$$\frac{3\tau - 1}{2} \leq \rho \leq \frac{1 + 2\tau - \tau^2}{2}, \quad \tau \geq 0$$

e

$$\frac{\tau^2 + 2\tau - 1}{2} \leq \rho \leq \frac{1 + 3\tau}{2}, \quad \tau \leq 0.$$

*Dimostrazione.* Il risultato si dimostra immediatamente a partire dai Teoremi 3.3.1 e 3.3.2.  $\square$



**Figura 3.1.** In grigio è evidenziata la regione definita dalle coppie  $(\tau, \rho)$  che possono essere ottenute per qualche copula. La regione all'interno della linea tratteggiata è il quadrato  $[-1, 1]^2$ .

### 3.4 Dipendenza di coda

In questo paragrafo definiamo brevemente i parametri di dipendenza di coda, il cui scopo è quello di fornire indicazioni sulla dipendenza tra due variabili aleatorie condizionatamente al fatto che una delle due assume valori estremi. In particolare, consideriamo due variabili aleatorie continue  $X$  e  $Y$  che hanno rispettivamente funzione di ripartizione  $F$  e  $G$  con copula  $C$ . Il relativo parametro di dipendenza di coda superiore  $\lambda_U$  è definito come

$$\lambda_U = \lim_{t \rightarrow 1^-} \mathbb{P} \left[ Y > G^{(-1)}(t) | X > F^{(-1)}(t) \right]$$

ammesso che il limite esista. Il parametro di dipendenza di coda inferiore  $\lambda_L$  è definito come

$$\lambda_L = \lim_{t \rightarrow 1^-} \mathbb{P} \left[ Y \leq G^{(-1)}(t) | X \leq F^{(-1)}(t) \right]$$

se il limite esiste.

La caratteristica fondamentale dei parametri  $\lambda_L$  e  $\lambda_U$  è che questi non dipendono dalle distribuzioni marginali ma solo dalla copula, come dimostra il seguente risultato.

**Teorema 3.4.1.** *Siano  $X$  e  $Y$  due variabili aleatorie continue con copula  $C$ . Siano  $\lambda_L$  e  $\lambda_U$  i parametri di dipendenza di coda inferiore e superiore di  $X$  e  $Y$ . Si ha che*

$$\lambda_U = 2 - \lim_{t \rightarrow 1^-} \frac{1 - C(t, t)}{1 - t}$$

e

$$\lambda_L = \lim_{t \rightarrow 0^+} \frac{C(t, t)}{t}.$$

*Dimostrazione.* Indichiamo con  $F$  e  $G$  le funzioni di ripartizione rispettivamente di  $X$  e  $Y$ . Possiamo quindi scrivere

$$\begin{aligned} \lambda_U &= \lim_{t \rightarrow 1^-} \mathbb{P} \left[ Y > G^{(-1)}(t) | X > F^{(-1)}(t) \right] \\ &= \lim_{t \rightarrow 1^-} \mathbb{P} [G(Y) > t | F(X) > t] \\ &= \lim_{t \rightarrow 1^-} \frac{\mathbb{P} [G(Y) > t, F(X) > t]}{\mathbb{P} [F(X) > t]} \\ &= \lim_{t \rightarrow 1^-} \frac{1 - \mathbb{P} [F(X) \leq t] - \mathbb{P} [G(Y) \leq t] + \mathbb{P} [G(Y) \leq t, F(X) \leq t]}{1 - t} \\ &= \lim_{t \rightarrow 1^-} \frac{1 - 2t + C(t, t)}{1 - t} = 2 - \lim_{t \rightarrow 1^-} \frac{1 - C(t, t)}{1 - t}. \end{aligned}$$

Analogamente

$$\begin{aligned} \lambda_L &= \lim_{t \rightarrow 0^+} \mathbb{P} \left[ Y \leq G^{(-1)}(t) | X \leq F^{(-1)}(t) \right] \\ &= \lim_{t \rightarrow 0^+} \frac{\mathbb{P} \left[ Y \leq G^{(-1)}(t), X \leq F^{(-1)}(t) \right]}{\mathbb{P} [F(X) \leq t]} = \lim_{t \rightarrow 0^+} \frac{C(t, t)}{t}. \end{aligned}$$

□

## Capitolo 4

# Alcuni esempi di copule

Il presente capitolo è dedicato allo studio di alcune note famiglie di copule e delle loro proprietà. Ci soffermeremo in particolare sulle copule Archimedee, di cui descriveremo il metodo di costruzione, e su quelle ellittiche che emergono in modo naturale dallo studio delle distribuzioni ellittiche.

### 4.1 Copule Archimedee

Le copule Archimedee costituiscono una famiglia di copule ampiamente studiate in letteratura per via della loro trattabilità analitica. Si tratta di copule frequentemente utilizzate nella pratica attuariale e finanziaria, anche per via della varietà delle strutture di dipendenza che possono descrivere. Restringendo ancora una volta l'analisi al caso di copule bivariate, la definizione di copula Archimedea è strettamente legata al seguente risultato.

**Teorema 4.1.1.** *Sia  $\varphi : [0, 1] \rightarrow [0, +\infty]$  una funzione continua e strettamente decrescente tale che  $\varphi(1) = 0$ . Indicando con  $\varphi^{[-1]}$  la pseudo-inversa di  $\varphi$ , ovvero*

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi(0) \\ 0, & \varphi(0) \leq t \leq +\infty \end{cases}$$

*allora si ha che la funzione*

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)), \quad u, v \in [0, 1] \quad (4.1)$$

*è una copula se e solo se  $\varphi$  è convessa.*

*Dimostrazione.* Osserviamo preliminarmente che, per  $t \in [0, 1]$ , si ha  $\varphi^{[-1]}(\varphi(t)) = t$  e  $\varphi(\varphi^{[-1]}(t)) = \min(t, \varphi(0))$ . Si verifica facilmente che la funzione  $C$  soddisfa le proprietà delle copule sulla frontiera del quadrato unitario. Ci resta quindi da mostrare che  $\varphi$  è convessa se e solo se  $C$  è 2-crescente. Dividiamo la dimostrazione in due parti. Nella prima parte mostriamo che  $C$  è 2-crescente se e solo se, per  $u_1 \leq u_2$ , vale la disuguaglianza

$$C(u_2, v) - C(u_1, v) \leq u_2 - u_1. \quad (4.2)$$

La seconda parte della dimostrazione consiste nel verificare che la condizione (4.2) equivale alla convessità di  $\varphi$ .

Si verifica immediatamente che la formula (4.2) può essere riscritta nella forma  $V_C([u_1, u_2] \times [v, 1]) \geq 0$  che è certamente vera se  $C$  è 2-crescente. Per verificare l'implicazione contraria, osserviamo che, per  $v_1 \leq v_2$ , vale la relazione  $C(0, v_2) = 0 \leq v_1 \leq v_2 = C(1, v_2)$ . Pertanto, dato che  $C$  è continua, per il teorema dei valori intermedi deve esistere un punto  $t \in [0, 1]$  tale che  $C(t, v_2) = v_1$  o, equivalentemente,  $\varphi(v_2) + \varphi(t) = \varphi(v_1)$ . Segue che, se la condizione (4.2) è soddisfatta, possiamo scrivere

$$\begin{aligned} C(u_2, v_1) - C(u_1, v_1) &= \varphi^{[-1]}(\varphi(u_2) + \varphi(v_1)) - \varphi^{[-1]}(\varphi(u_1) + \varphi(v_1)) \\ &= \varphi^{[-1]}(\varphi(u_2) + \varphi(v_2) + \varphi(t)) - \varphi^{[-1]}(\varphi(u_1) + \varphi(v_2) + \varphi(t)) \\ &= C(C(u_2, v_2), t) - C(C(u_1, v_2), t) \leq C(u_2, v_2) - C(u_1, v_2) \end{aligned} \quad (4.3)$$

da cui si conclude che la relazione (4.2) è equivalente alla 2-crescenza di  $C$ . Mostriamo ora che la (4.2) equivale alla convessità di  $\varphi$ . Cominciamo con il riformulare la (4.2) come

$$u_1 + \varphi^{[-1]}(\varphi(u_2) + \varphi(v)) \leq u_2 + \varphi^{[-1]}(\varphi(u_1) + \varphi(v))$$

che, posti  $a = \varphi(u_1)$ ,  $b = \varphi(u_2)$  e  $c = \varphi(v)$ , può essere a sua volta riscritta come

$$\varphi^{[-1]}(a) + \varphi^{[-1]}(b + c) \leq \varphi^{[-1]}(b) + \varphi^{[-1]}(a + c) \quad (4.4)$$

dove  $a \geq b$  e  $c \geq 0$ . Comunque siano scelti due punti  $s, t \in [0, \infty]$  con  $s < t$ , ponendo  $a = \frac{s+t}{2}$ ,  $b = s$  e  $c = \frac{t-s}{2}$  si può scrivere la formula (4.4) nella forma

$$\varphi^{[-1]} \left( \frac{s+t}{2} \right) \leq \frac{\varphi^{[-1]}(s) + \varphi^{[-1]}(t)}{2}. \quad (4.5)$$

La formula (4.5), congiuntamente alla continuità di  $\varphi^{[-1]}$ , implica la convessità di  $\varphi^{[-1]}$  e di conseguenza la convessità di  $\varphi$ . Ipotizziamo ora che  $\varphi^{[-1]}$  sia convessa. Fissati tre punti  $a, b, c \in [0, 1]$  tali che  $a \geq b$  e  $c \geq 0$ , si ponga  $\gamma = \frac{a-b}{a-b+c}$  e si osservi che  $a = (1-\gamma)b + \gamma(a+c)$  e  $b+c = \gamma b + (1-\gamma)(a+c)$ . Per la convessità di  $\varphi^{[-1]}$  si può scrivere

$$\varphi^{[-1]}(a) \leq (1-\gamma)\varphi^{[-1]}(b) + \gamma\varphi^{[-1]}(a+c)$$

e

$$\varphi^{[-1]}(b+c) \leq \gamma\varphi^{[-1]}(b) + (1-\gamma)\varphi^{[-1]}(a+c).$$

Sommando le ultime due disuguaglianze si ottiene proprio la (4.4).  $\square$

Il teorema 4.1.1 fornisce un semplice metodo per la costruzione di copule. Infatti, scelta opportunamente una funzione  $\varphi$  che soddisfi le ipotesi del teorema, si ottiene immediatamente una copula mediante la costruzione descritta dalla formula (4.1). Le copule così costruite si dicono Archimedee.

**Definizione 4.1.1.** Una copula che ammette la rappresentazione (4.1) è detta Archimedea. La funzione  $\varphi$  è detta generatore della copula.  $\blacksquare$

Nello studio delle copule Archimedee, il generatore riveste un ruolo fondamentale e consente di studiare diverse proprietà della copula. Con riferimento alle misure di dipendenza studiate nel capitolo precedente, vediamo di seguito come queste possono essere espresse in termini del generatore nel caso di copule Archimedee. Per farlo, ci avvaliamo innanzitutto del risultato seguente.

**Teorema 4.1.2.** *Siano  $U$  e  $V$  due variabili aleatorie uniformi in  $(0, 1)$  con copula Archimedea  $C$  generata da  $\varphi$ . La funzione di ripartizione della variabile aleatoria  $C(U, V)$  è data da*

$$K_C(t) = t - \frac{\varphi(t)}{\varphi'(t^+)}$$

dove  $\varphi'(t^+)$  rappresenta la derivata destra di  $\varphi$  nel punto  $t$ .

*Dimostrazione.* Cominciamo con l'osservare che  $K_C(t)$  coincide con la  $C$ -misura dell'insieme

$$\{(u, v) \in [0, 1]^2 \mid C(u, v) \leq t\}$$

o, equivalentemente,

$$\{(u, v) \in [0, 1]^2 \mid \varphi(u) + \varphi(v) \geq \varphi(t)\}.$$

Si noti che, se  $u \leq t$ , allora è certamente vero che  $\varphi(u) + \varphi(v) \geq \varphi(t)$ . Se invece  $u > t$ , allora si ha che  $\varphi(u) + \varphi(v) \geq \varphi(t)$  se e solo se  $v \leq \varphi^{[-1]}(\varphi(t) - \varphi(u))$ . Di conseguenza, ponendo

$$S = \{(u, v) \in [0, 1]^2 \mid v \leq \varphi^{[-1]}(\varphi(t) - \varphi(u)), u > t\},$$

possiamo scrivere

$$K_C(t) = V_C([0, t] \times [0, 1]) + V_C(S) = t + V_C(S). \quad (4.6)$$

Partizioniamo ora l'intervallo  $[t, 1]$  nei punti  $t_k = \varphi^{[-1]} \left( \frac{(n-k)\varphi(t)}{n} \right)$  per  $k = 0, \dots, n$ . Definiamo  $S_n$  come

$$S_n = \cup_{k=1}^n R_k$$

dove

$$R_k = [t_{k-1}, t_k] \times [0, \varphi^{[-1]}(\varphi(t) - \varphi(t_{k-1}))].$$

Si può verificare che

$$V_C(S) = \lim_{n \rightarrow +\infty} V_C(S_n).$$

Inoltre, si ha che

$$\begin{aligned} V_C(S_n) &= \sum_{k=1}^n V_C(R_k) = \sum_{k=1}^n \left[ C(t_k, \varphi^{[-1]}(\varphi(t) - \varphi(t_{k-1}))) - C(t_{k-1}, \varphi^{[-1]}(\varphi(t) - \varphi(t_{k-1}))) \right] \\ &= - \sum_{k=1}^n \left[ t - \varphi^{[-1]}(\varphi(t_k) - \varphi(t_{k-1}) + \varphi(t)) \right] \\ &= - \varphi(t) \frac{\varphi^{[-1]}(\varphi(t)) - \varphi^{[-1]} \left( \varphi(t) - \frac{\varphi(t)}{n} \right)}{\frac{\varphi(t)}{n}}. \end{aligned}$$

Calcolando il limite per  $n \rightarrow +\infty$  si ottiene

$$V_C(S) = -\frac{\varphi(t)}{\varphi'(t^+)}$$

che sostituita nella formula (4.6) completa la dimostrazione.  $\square$

Vediamo ora come si esprime il tau di Kendall di una copula Archimedeana in termini del generatore.

**Teorema 4.1.3.** *Siano  $X$  e  $Y$  due variabili aleatorie con copula  $C$  generata da  $\varphi$ . Il tau di Kendall per  $X$  e  $Y$  è dato da*

$$\tau_C = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt.$$

*Dimostrazione.* Abbiamo visto nel capitolo precedente che il tau di Kendall non dipende dalle marginali ma solo dalla copula. Pertanto, consideriamo per semplicità il caso di due variabili uniformi in  $(0, 1)$  con copula  $C$ . Grazie alla formula (3.7), possiamo scrivere che

$$\tau_C = 4 \mathbb{E}[C(U, V)] - 1 = 4 \int_0^1 K_C(t) dt - 1.$$

L'enunciato si dimostra utilizzando l'espressione esplicita di  $K_C(t)$  ottenuta nel teorema 4.1.2 e svolgendo l'integrale. Si osservi, in particolare, che all'interno dell'integrale l'espressione  $\varphi'(t^+)$  può essere sostituita con  $\varphi'(t)$  perchè una funzione convessa è derivabile quasi ovunque.  $\square$

Un risultato simile vale anche per i parametri di dipendenza di coda.

**Teorema 4.1.4.** *Sia  $C$  una copula Archimedeana con generatore  $\varphi$ . Allora*

$$\lambda_U = 2 - \lim_{t \rightarrow 1^-} \frac{1 - \varphi^{[-1]}(2\varphi(t))}{1 - t} = 2 - \lim_{x \rightarrow 0^+} \frac{1 - \varphi^{[-1]}(2x)}{1 - \varphi^{[-1]}(x)}$$

e

$$\lambda_L = \lim_{t \rightarrow 0^+} \frac{\varphi^{[-1]}(2\varphi(t))}{t} = \lim_{x \rightarrow +\infty} \frac{\varphi^{[-1]}(2x)}{\varphi^{[-1]}(x)}.$$

*Dimostrazione.* Il risultato è un'immediata conseguenza del teorema 3.4.1 e della definizione di copula Archimedeana.  $\square$

Mostriamo di seguito alcuni esempi di copule Archimedee uniparametriche.

### 4.1.1 Copula Frank

La copula corrispondente al generatore

$$\varphi(t) = -\log \left( \frac{e^{-\theta t} - 1}{e^{-\theta} - 1} \right), \quad \theta \in \mathbb{R}$$



è nota come copula Frank e assume la forma

$$C_{\theta}(u, v) = -\frac{1}{\theta} \log \left[ 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right].$$

La corrispondente densità di copula è

$$c_{\theta}(u, v) = \frac{\theta (1 - e^{-\theta}) e^{-\theta(u+v)}}{(e^{-\theta} + e^{-\theta(u+v)} - e^{-\theta u} - e^{-\theta v})^2}.$$

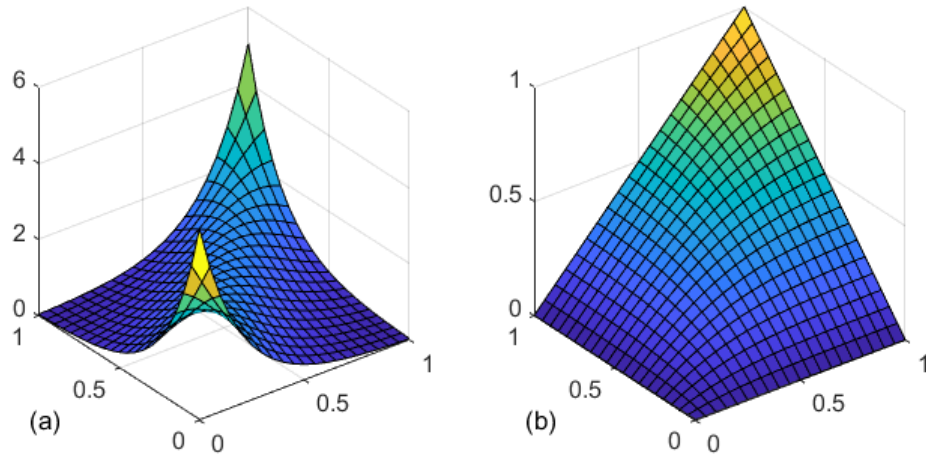
Il tau di Kendall per la copula Frank è

$$\tau = 1 - \frac{4}{\theta} [1 - D_1(\theta)]$$

dove

$$D_n(\theta) = \frac{n}{\theta^n} \int_0^{\theta} \frac{t^n}{e^t - 1} dt$$

è la funzione di Debye. La copula Frank non possiede dipendenza di coda, ovvero si ha che  $\lambda_U = \lambda_L = 0$ .



**Figura 4.1.** (a) densità della copula Frank. (b) copula Frank. Entrambi i grafici sono stati ottenuti per  $\theta = 5$ .

#### 4.1.2 Copula Clayton

Si consideri il generatore

$$\varphi(t) = \frac{1}{\theta} (t^{-\theta} - 1), \quad \theta \in [-1, 0) \cup (0, \infty).$$

La corrispondente copula, nota come copula di Clayton, può essere scritta come

$$C_{\theta}(u, v) = \left[ \max(u^{-\theta} + v^{-\theta} - 1; 0) \right]^{-\frac{1}{\theta}}.$$

È evidente che la densità di probabilità della copula Clayton non è ben definita sull'intero quadrato unitario. Per questa ragione, nella pratica, si preferisce restringere l'insieme dei possibili valori di  $\theta$  ai reali positivi, escludendo quindi i valori negativi. Infatti, se  $\theta \in (0, \infty)$ , la copula si riduce alla forma

$$C_\theta(u, v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}}$$

che ammette densità di probabilità

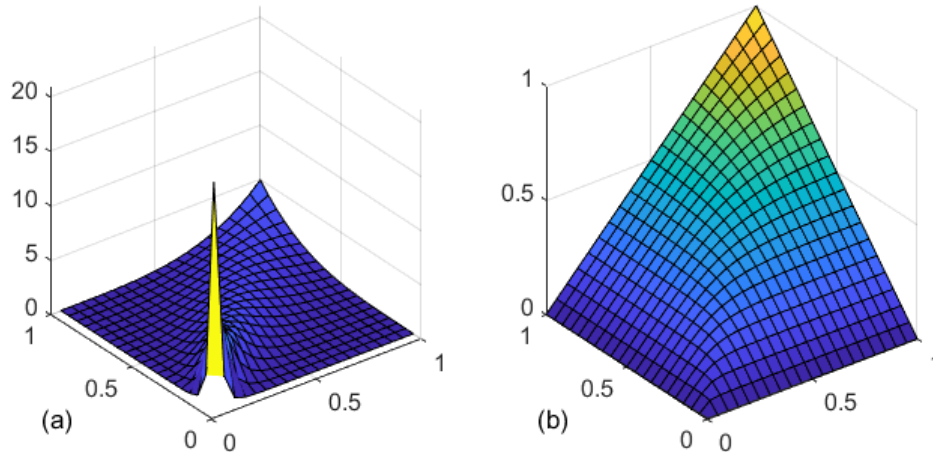
$$c_\theta(u, v) = \frac{(1 + \theta)}{(uv)^{1+\theta}} \left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1+2\theta}{\theta}}.$$

Il tau di Kendall per la copula Clayton è

$$\tau = \frac{\theta}{2 + \theta}.$$

La copula Clayton non possiede dipendenza di coda superiore mentre possiede dipendenza di coda inferiore. In particolare si ha

$$\lambda_U = 0, \quad \lambda_L = 2^{-\frac{1}{\theta}}.$$



**Figura 4.2.** (a) densità della copula Clayton. (b) copula Clayton. I grafici sono stati ottenuti per  $\theta = 4$ .

### 4.1.3 Copula Gumbel

La copula Gumbel si ottiene dal generatore

$$\varphi(t) = (-\log t)^\theta, \quad \theta \in (1, \infty)$$

e assume quindi la forma

$$C_\theta(u, v) = \exp\left(-\left[(-\log u)^\theta + (-\log v)^\theta\right]^{\frac{1}{\theta}}\right)$$

con densità

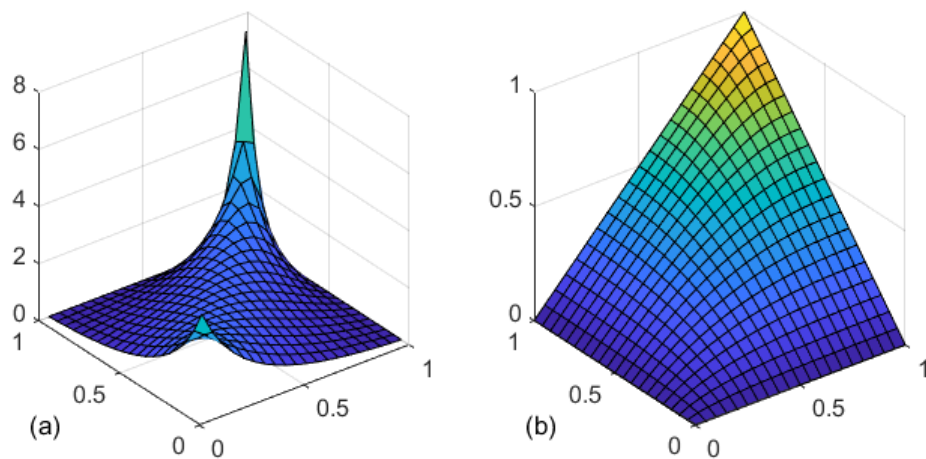
$$c_{\theta}(u, v) = \frac{(\log u \log v)^{\theta-1}}{uv} \exp\left(-\left[(-\log u)^{\theta} + (-\log v)^{\theta}\right]^{\frac{1}{\theta}}\right) \cdot \left\{ \left[(-\log u)^{\theta} + (-\log v)^{\theta}\right]^{2\left(\frac{1}{\theta}-1\right)} - (1-\theta) \left[(-\log u)^{\theta} + (-\log v)^{\theta}\right]^{\frac{1}{\theta}-2} \right\}.$$

Il tau di Kendall per la copula Gumbel è

$$\tau = \frac{\theta}{2 + \theta}.$$

La copula Gumbel non possiede dipendenza di coda inferiore mentre possiede dipendenza di coda superiore. I parametri di dipendenza di coda sono

$$\lambda_U = 2 - 2^{\frac{1}{\theta}}, \quad \lambda_L = 0.$$



**Figura 4.3.** (a) densità della copula Gumbel. (b) copula Gumbel. I grafici sono stati ottenuti per  $\theta = 2$ .

## 4.2 Copule ellittiche

Si definiscono ellittiche le copule che legano le marginali delle distribuzioni ellittiche multivariate. Tali copule ereditano quindi dalle distribuzioni ellittiche alcune interessanti proprietà. In particolare, vedremo che le distribuzioni ellittiche, e di conseguenza le copule ellittiche, ammettono una rappresentazione stocastica che consente, per esempio, di effettuare simulazioni Monte Carlo. Nel presente capitolo introduciamo l'argomento presentando la definizione e le proprietà fondamentali delle distribuzioni ellittiche. A questo scopo facciamo riferimento al libro di Anderson [3]. Iniziamo esaminando il sottinsieme delle distribuzioni ellittiche costituito dalle cosiddette distribuzioni sferiche.

**Definizione 4.2.1.** La variabile aleatoria  $d$ -dimensionale  $\mathbf{X} = (X_1, \dots, X_d)^T$  ha distribuzione sferica se la sua funzione caratteristica è esprimibile, per  $\mathbf{t} = (t_1, \dots, t_d)^T$ , nella forma

$$\mathbb{E} \left[ e^{it^T \mathbf{X}} \right] = \psi \left( \mathbf{t}^T \mathbf{t} \right) \quad (4.7)$$

per qualche funzione scalare  $\psi$ . ■

Al fine di comprendere il significato della definizione di distribuzione sferica, supponiamo di applicare al vettore  $\mathbf{X}$  una trasformazione lineare associata a una certa matrice ortogonale  $\mathbf{O}$ . Ricordiamo che  $\mathbf{O}$  si dice ortogonale se  $\mathbf{O}^T \mathbf{O} = \mathbf{I}$ . Una matrice ortogonale descrive una trasformazione rigida che non altera il prodotto scalare, come si può facilmente verificare. Se poniamo  $\mathbf{Y} = \mathbf{O}\mathbf{X}$ , la funzione caratteristica di  $\mathbf{Y}$  sarà data da

$$\mathbb{E} \left[ e^{it^T \mathbf{Y}} \right] = \psi \left( (\mathbf{O}^T \mathbf{t})^T (\mathbf{O}^T \mathbf{t}) \right) = \psi \left( \mathbf{t}^T \mathbf{t} \right) = \mathbb{E} \left[ e^{it^T \mathbf{X}} \right].$$

Ciò significa che una distribuzione sferica è invariante rispetto a trasformazioni ortogonali. Come è evidente, la definizione 4.2.1 non fa alcun riferimento alla densità di probabilità della variabile con distribuzione sferica. L'esistenza di una densità di probabilità non è infatti garantita in generale. Tuttavia, nel caso in cui questa esista, il suo valore deve dipendere solo dalla norma dell'argomento  $\mathbf{y}$  e non dalla sua direzione. In caso contrario sarebbe contraddetta l'invarianza rispetto a trasformazioni ortogonali. Pertanto, una distribuzione sferica ha densità di probabilità della forma

$$g(\mathbf{y}^T \mathbf{y})$$

ammesso che questa esista.

Possiamo a questo punto definire le distribuzioni ellittiche. Un vettore  $\mathbf{X}$  ha distribuzione ellittica se ammette la rappresentazione

$$\mathbf{X} = \boldsymbol{\mu} + \mathbf{C}\mathbf{Y}$$

dove  $\boldsymbol{\mu} \in \mathbb{R}^d$ ,  $\mathbf{Y}$  è un vettore avente distribuzione sferica e  $\mathbf{C}$  è una matrice tale che il prodotto  $\boldsymbol{\Sigma} = \mathbf{C}^T \mathbf{C}$  è una matrice definita positiva. Equivalentemente, è possibile dare la seguente definizione.

**Definizione 4.2.2.** La variabile aleatoria  $d$ -dimensionale  $\mathbf{X} = (X_1, \dots, X_d)^T$  ha distribuzione ellittica se la sua funzione caratteristica è esprimibile, per  $\mathbf{t} = (t_1, \dots, t_d)^T$ , nella forma

$$\mathbb{E} \left[ e^{it^T (\mathbf{X} - \boldsymbol{\mu})} \right] = \psi \left( \mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t} \right) \quad (4.8)$$

per un'opportuna scelta di un parametro di posizione  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)^T$ , una matrice definita positiva  $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$  e una funzione scalare  $\psi$ . ■

Nei casi in cui una distribuzione ellittica ammette densità di probabilità, questa è esprimibile nella forma

$$\frac{1}{\sqrt{|\boldsymbol{\Sigma}|}} g \left( (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right).$$

Nel seguito presentiamo le due copule ellittiche più note, ovvero la copula Gaussiana e la copula  $t$ , nelle loro versioni multivariate.

### 4.2.1 Copula Gaussiana

La copula Gaussiana emerge in maniera naturale dallo studio della distribuzione normale multivariata. Si consideri un vettore aleatorio  $d$ -dimensionale  $\mathbf{X} = (X_1, X_2, \dots, X_d)^\top$  avente distribuzione normale multivariata, ovvero avente densità di probabilità

$$f(x_1, x_2, \dots, x_d; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}}$$

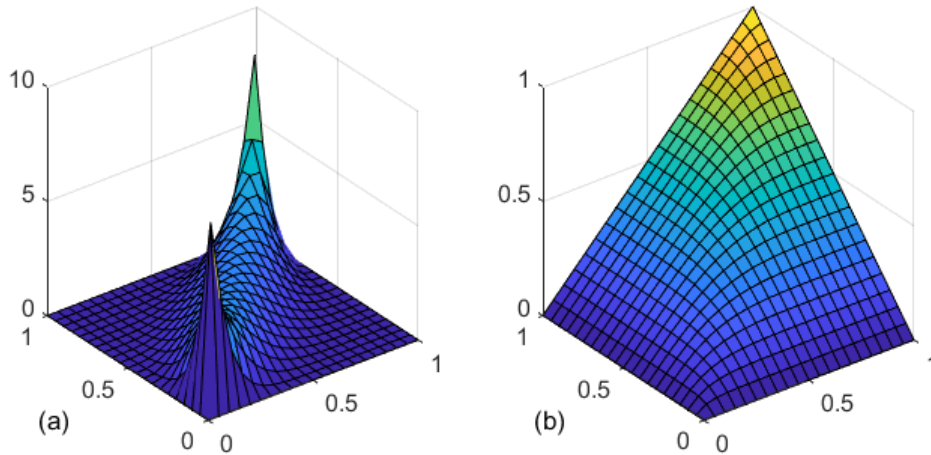
dove  $\mathbf{x} = (x_1, x_2, \dots, x_d)^\top$  e  $\boldsymbol{\mu}$  e  $\boldsymbol{\Sigma}$  sono rispettivamente il vettore delle medie e la matrice di varianze e covarianze di  $\mathbf{X}$ . Consideriamo per semplicità il caso in cui la normale multivariata abbia marginali standard. In altre parole, assumiamo che  $\boldsymbol{\mu} = \mathbf{0}$  e che la matrice  $\boldsymbol{\Sigma} = \mathbf{R}$  sia una matrice di correlazioni. Indicando con  $\Phi_{\mathbf{R}}$  la funzione di ripartizione multivariata di  $\mathbf{X}$  e con  $\Phi$  la funzione di ripartizione univariata di una normale standard, si può ovviamente scrivere

$$\Phi_{\mathbf{R}}(x_1, x_2, \dots, x_d) = \Phi_{\mathbf{R}}\left(\Phi^{-1}(\Phi(x_1)), \Phi^{-1}(\Phi(x_2)), \dots, \Phi^{-1}(\Phi(x_d))\right). \quad (4.9)$$

Applicando il teorema di Sklar, e in particolare confrontando le formule (2.8) e (4.9), si può verificare che le marginali della normale multivariata sono legate dalla copula

$$C_{\mathbf{R}}(u_1, u_2, \dots, u_d) = \Phi_{\mathbf{R}}\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_d)\right).$$

Lo stesso risultato può essere ottenuto, naturalmente, applicando il corollario 2.2.5 alla distribuzione normale multivariata.



**Figura 4.4.** (a) densità della copula Gaussiana. (b) copula Gaussiana. I grafici sono stati ottenuti per  $\rho = 0.9$ .

La densità della copula è data da

$$c_{\mathbf{R}}(u_1, u_2, \dots, u_d) = \frac{\exp\left(-\frac{1}{2} \Phi^{-1}(\mathbf{u})^\top (\mathbf{R}^{-1} - \mathbf{I}_d) \Phi^{-1}(\mathbf{u})\right)}{\sqrt{|\mathbf{R}|}}$$

dove  $\Phi^{-1}(\mathbf{u}) = (\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_d))^T$  e  $\mathbf{I}_d$  è la matrice identità  $d \times d$ . Si osservi che, nel caso in cui la matrice delle correlazioni coincida con  $\mathbf{I}_d$ , la copula Gaussiana coincide con la copula di indipendenza. In altre parole, quando le marginali sono incorrelate allora sono anche indipendenti.

### 4.2.2 Copula $t$

Consideriamo ora la distribuzione  $t$  di Student multivariata. Si consideri quindi un vettore aleatorio  $d$ -dimensionale  $\mathbf{X} = (X_1, X_2, \dots, X_d)^T$  avente densità di probabilità

$$f_\nu(x_1, x_2, \dots, x_d; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu\pi)^d |\boldsymbol{\Sigma}|}} \left[1 + \frac{1}{\nu}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right]^{-\frac{\nu+d}{2}} \quad (4.10)$$

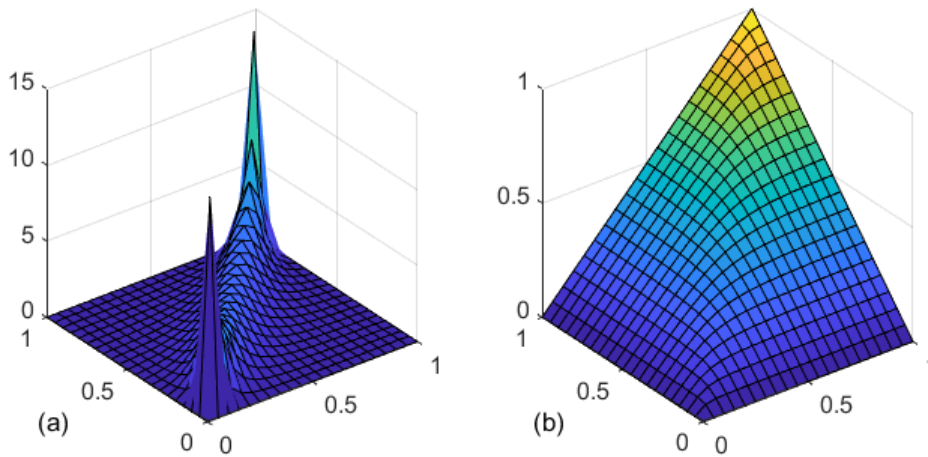
dove  $\boldsymbol{\mu}$  e  $\boldsymbol{\Sigma}$  sono rispettivamente il vettore dei parametri di posizione e una matrice definita positiva,  $\nu > 0$  è il parametro che indica i gradi di libertà della distribuzione e  $\mathbf{x} = (x_1, x_2, \dots, x_d)^T$ . Ricordiamo che, essendo  $\boldsymbol{\Sigma}$  una matrice definita positiva, è di fatto una potenziale matrice di varianze e covarianze. Tuttavia, questa non coincide con la matrice di varianze e covarianze di  $\mathbf{X}$ , che è invece

$$\mathbb{V}[\mathbf{X}] = \frac{\nu}{\nu - 2} \boldsymbol{\Sigma}, \quad \nu > 2.$$

Per  $\nu \leq 2$  la matrice di varianze e covarianze non è definita. La media è invece definita solo per  $\nu > 1$  e si ha

$$\mathbb{E}[\mathbf{X}] = \boldsymbol{\mu}, \quad \nu > 1.$$

Per maggiori dettagli su come dimostrare i risultati appena riportati si rimanda al libro di Johnson [38].



**Figura 4.5.** (a) densità della copula  $t$ . (b) copula  $t$ . I grafici sono stati ottenuti per  $\rho = 0.9$  e  $\nu = 2$ .

Come per la copula Gaussiana, consideriamo per semplicità il caso in cui  $\boldsymbol{\mu} = \mathbf{0}$  e la matrice  $\boldsymbol{\Sigma} = \mathbf{R}$  è una matrice di correlazioni. Indicando con  $t_{\nu, \mathbf{R}}$  la corrispondente

funzione di ripartizione multivariata e con  $t_\nu$  la funzione di ripartizione di una  $t$  di Student univariata standard, si ottiene immediatamente che le marginali della  $t$  di Student multivariata sono legate dalla copula

$$C_{\nu, \mathbf{R}}(u_1, u_2, \dots, u_d) = \mathbf{t}_{\nu, \mathbf{R}} \left( t_\nu^{-1}(u_1), t_\nu^{-1}(u_2), \dots, t_\nu^{-1}(u_d) \right).$$

### 4.3 Un metodo generale per la costruzione di copule

Nei paragrafi precedenti abbiamo descritto alcune delle copule più note in letteratura. Viene naturale, a questo punto, domandarsi quali metodi siano stati utilizzati per ottenere tali copule e come si possa fare per costruirne di nuove. Un metodo generale per la costruzione di copule deriva direttamente dal teorema di Sklar. Sfruttando quest'ultimo, è infatti possibile costruire una copula semplicemente costruendo una distribuzione multivariata per poi separare la componente di copula dalle distribuzioni marginali. Molte delle copule note in letteratura possono essere ottenute proprio con questo metodo. Mostriamo di seguito alcuni esempi a scopo illustrativo.

Cominciamo con il mostrare come si può costruire la copula Gaussiana bivariata. Si consideri la coppia  $(X, Y)$  di variabili aleatorie definite come

$$\begin{cases} X = Z_1 \\ Y = \rho Z_1 + \sqrt{1 - \rho^2} Z_2 \end{cases} \quad |\rho| \leq 1 \quad (4.11)$$

dove  $Z_1$  e  $Z_2$  sono due variabili aleatorie indipendenti e aventi distribuzione normale standard. La distribuzione del vettore  $(X, Y)$  è ovviamente una normale bivariata con coefficiente di correlazione  $\rho$ . Indicando con  $\Phi_\rho$  la sua funzione di ripartizione congiunta e con  $\Phi$  la funzione di ripartizione di una normale standard univariata, si ha che

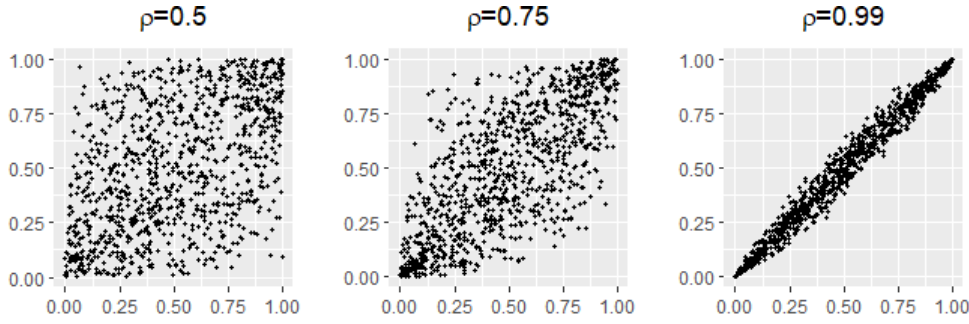
$$\Phi_\rho(x, y) = \Phi_\rho \left( \Phi^{-1}(\Phi(x)), \Phi^{-1}(\Phi(y)) \right)$$

da cui si ottiene, effettuando un semplice confronto con la formula (2.8), che la copula sottostante alla distribuzione di  $(X, Y)$  è la copula Gaussiana bivariata

$$C(u, v) = \Phi_\rho \left( \Phi^{-1}(u), \Phi^{-1}(v) \right).$$

A partire da due variabili indipendenti  $Z_1$  e  $Z_2$ , abbiamo quindi costruito delle variabili  $X$  e  $Y$  tra loro dipendenti. Il teorema di Sklar consente a questo punto di ricavare la copula dalla distribuzione multivariata ottenuta. Il metodo descritto è noto come metodo di inversione.

La figura 4.6 mostra dei campioni simulati da una distribuzione uniforme bivariata con copula Gaussiana. Tale rappresentazione grafica illustra un fatto di cui si è discusso nel capitolo 2. Si vede infatti che, quando la correlazione si avvicina a  $\rho = 1$ , i punti tendono a distribuirsi lungo la curva di equazione  $f(u) = u, u \in [0, 1]$ . Ciò è ovviamente dovuto al fatto che, per  $\rho = 1$ , la copula Gaussiana descrive, per costruzione, la massima dipendenza positiva. Ovviamente, se la correlazione fosse pari a  $\rho = -1$ , i punti si disporrebbero lungo la curva di equazione  $f(u) = 1 - u, u \in [0, 1]$ . In ogni caso, in presenza di massima dipendenza i punti del campione bivariato saranno concentrati su una delle diagonali del quadrato  $C$  come è immediatamente evidente dalla definizione (4.11). Simili conclusioni possono essere raggiunte per una



**Figura 4.6.** Simulazioni Monte Carlo da una copula Gaussiana per diversi valori di  $\rho$ . Per ciascuno dei grafici il numero di punti simulati è pari a 1000.

moltitudine di copule note in letteratura.

Concludiamo il capitolo proponendo un ulteriore esempio di applicazione del metodo di inversione. Siano  $Z_1$  e  $Z_2$  due variabili aleatorie indipendenti aventi distribuzione normale standard. Poniamo

$$\begin{cases} X = \rho Z_1 + \sqrt{1 - \rho^2} Z_2 \\ Y = Z_1^2 \end{cases} \quad (4.12)$$

con  $-1 \leq \rho \leq 1$ . Si verifica immediatamente che  $X$  ha distribuzione normale standard mentre  $Y$  ha distribuzione chi-quadro con un grado di libertà. A differenza del precedente esempio, è evidente in questo caso che le variabili  $X$  e  $Y$  non sono legate da una dipendenza di tipo monotono. In corrispondenza di valori estremi di  $Z_1$ , la variabile  $X$  tenderà ad assumere valori estremi concordanti (o discordanti se  $\rho < 0$ ) con  $Z_1$  mentre la variabile  $Y$  tenderà ad assumere valori estremi sulla coda destra. La dipendenza è infatti di tipo quadratico e diventa una dipendenza perfetta nel caso in cui  $|\rho| = 1$ . Per studiare la distribuzione congiunta di  $X$  e  $Y$ , osserviamo che

$$\begin{aligned} \mathbb{P}(X \leq x, Y \leq y) &= \mathbb{P}(\rho Z_1 + \sqrt{1 - \rho^2} Z_2 \leq x, Z_1^2 \leq y) \\ &= \mathbb{P}(\rho Z_1 + \sqrt{1 - \rho^2} Z_2 \leq x, -\sqrt{y} \leq Z_1 \leq \sqrt{y}) \\ &= \mathbb{P}(\rho Z_1 + \sqrt{1 - \rho^2} Z_2 \leq x, Z_1 \leq \sqrt{y}) - \mathbb{P}(\rho Z_1 + \sqrt{1 - \rho^2} Z_2 \leq x, Z_1 \leq -\sqrt{y}) \\ &= \Phi_\rho(x, \sqrt{y}) - \Phi_\rho(x, -\sqrt{y}). \end{aligned}$$

Indicando con  $\chi$  la funzione di ripartizione di una variabile chi-quadro con un grado di libertà, possiamo quindi scrivere la copula che lega  $X$  e  $Y$  nella forma

$$C_\rho(u, v) = \Phi_\rho\left(\Phi^{-1}(u), \sqrt{\chi^{-1}(v)}\right) - \Phi_\rho\left(\Phi^{-1}(u), -\sqrt{\chi^{-1}(v)}\right). \quad (4.13)$$

Siamo ora interessati a comprendere il comportamento della copula (4.13) nel caso



limite in cui  $|\rho| = 1$ . In particolare, per  $\rho = 1$ , segue dalla definizione (4.12) che

$$\begin{aligned} \mathbb{P}(X \leq x, Y \leq y) &= \mathbb{P}(Z_1 \leq x, -\sqrt{y} \leq Z_1 \leq \sqrt{y}) = \mathbb{P}(-\sqrt{y} \leq Z_1 \leq \min(\sqrt{y}, x)) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\min(-\sqrt{y}, x)}^{\min(\sqrt{y}, x)} e^{-\frac{z^2}{2}} dz = \Phi(\min(\sqrt{y}, x)) - \Phi(\min(-\sqrt{y}, x)) \end{aligned}$$

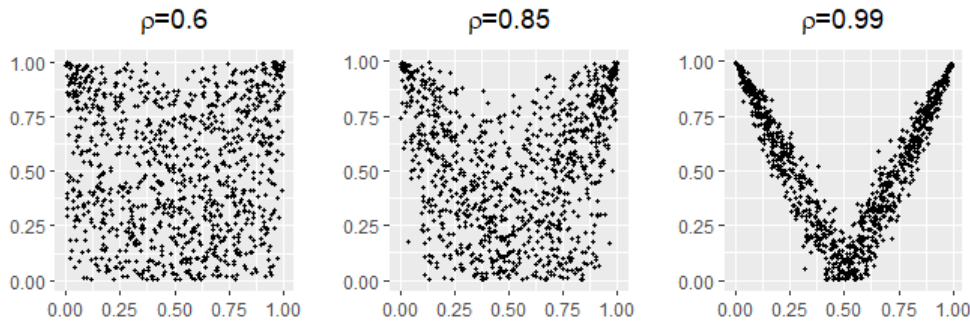
da cui, utilizzando il fatto che  $\Phi(-\sqrt{\chi^{-1}(v)}) = \Phi(\sqrt{\chi^{-1}(v)}) - v$ , si ottiene

$$C_1(u, v) = \min\left(\Phi\left(\sqrt{\chi^{-1}(v)}\right), u\right) - \min\left(\Phi\left(\sqrt{\chi^{-1}(v)}\right) - v, u\right).$$

Considerando che  $\Phi\left(\sqrt{\chi^{-1}(v)}\right) = \frac{1+v}{2}$  otteniamo infine

$$C_1(u, v) = M\left(u, \frac{1+v}{2}\right) - M\left(u, \frac{1-v}{2}\right)$$

dove  $M$  rappresenta il confine superiore di Fréchet-Hoeffding.



**Figura 4.7.** Simulazioni Monte Carlo della copula non monotona definita nell'equazione (4.13) per diversi valori di  $\rho$ . Per ciascuno dei grafici il numero di punti simulati è pari a 1000.

La figura 4.7 illustra il risultato della simulazione Monte Carlo di alcuni campioni della distribuzione (4.13) per diversi valori di  $\rho$ . È immediatamente evidente come, al crescere di  $\rho$ , la distribuzione tenda a concentrarsi lungo il grafico di una funzione non monotona. Come avremo modo di dimostrare nel prossimo capitolo, la funzione in questione corrisponde a  $g(u) = |2u - 1|$ .

Utilizzando il metodo di inversione, abbiamo quindi ottenuto un procedimento per la costruzione di una copula che descrive una dipendenza non monotona. È infatti sufficiente costruire una generica distribuzione multivariata, con marginali arbitrarie, che esibisca una dipendenza non monotona. Mediante il teorema di Sklar si può di conseguenza isolare la copula che descrive tale dipendenza.

L'obiettivo del prossimo capitolo è quello di formalizzare il procedimento appena illustrato, definendo un metodo generale per la costruzione di copule che possano descrivere dipendenze di tipo arbitrario.

## Capitolo 5

# Dipendenze non monotone

Riprendiamo ora alcune idee esposte nel paragrafo 2.3. In particolare, abbiamo visto che l'unica forma di perfetta dipendenza monotona tra due variabili uniformi è la dipendenza lineare. Se quindi  $U$  e  $V$  sono due variabili aleatorie uniformemente distribuite sull'intervallo  $[0, 1]$  e  $g : [0, 1] \rightarrow [0, 1]$  è una funzione strettamente monotona, si può avere la relazione  $V = g(U)$  solo nel caso in cui  $g(u) = u$  o  $g(u) = 1 - u$ . In tal caso, i punti di coordinate  $(U, V)$  saranno quindi concentrati sulla curva coincidente con il grafico di  $g$ . Lo scopo del presente capitolo è quello di generalizzare le idee appena descritte al caso in cui  $g$  non sia una funzione monotona. Nel seguito della trattazione, per indicare che due variabili aleatorie  $X$  e  $Y$  sono identiche in distribuzione, adotteremo la notazione

$$X \stackrel{i.d.}{=} Y.$$

Introduciamo lo studio delle dipendenze non monotone presentando la seguente definizione.

**Definizione 5.0.1.** Diciamo che una funzione  $g : [0, 1] \rightarrow [0, 1]$  preserva la distribuzione uniforme se, per una variabile aleatoria  $U$  uniformemente distribuita sull'intervallo  $[0, 1]$ , si ha che

$$g(U) \stackrel{i.d.}{=} U.$$

■

Osserviamo ora che, se  $g$  preserva la distribuzione uniforme e  $V = g(U)$ , la coppia  $(U, g(U))$  avrà ovviamente una distribuzione con marginali uniformi e la sua funzione di ripartizione sarà quindi una copula. Inoltre, la dipendenza è per costruzione perfetta ed è quindi immediato verificare che la massa di probabilità sarà concentrata sulla curva che coincide con il grafico di  $g$ . Naturalmente, ciò è vero indipendentemente dal fatto che  $g$  sia una funzione monotona. Di conseguenza, a qualunque funzione che preservi la distribuzione uniforme corrisponde una curva su cui è concentrata la massa di probabilità di un'opportuna copula che descrive perfetta dipendenza. Proprio queste osservazioni sono alla base del lavoro di Scarsini e Venetoulis [58]. Seguendo le idee degli stessi autori, è quindi possibile costruire copule che descrivono dipendenze di tipo arbitrario nel modo seguente.

Si consideri una coppia  $(U, V)$  di variabili aleatorie uniformemente distribuite su  $[0, 1]$  con copula  $C$  e sia  $g$  una funzione che preserva la distribuzione uniforme. La coppia

$(U, g(V))$  avrà a sua volta una distribuzione avente marginali uniformi. Pertanto, la sua funzione di ripartizione sarà una copula che indichiamo con  $Q$ . L'applicazione della funzione  $g$  a una delle componenti di  $(U, V)$  non modifica le distribuzioni marginali e fa sì che la distribuzione risultante sia quindi ancora una copula. Tuttavia, trasformando una delle due componenti del vettore iniziale si altera la struttura di dipendenza. Scegliendo quindi una copula iniziale  $C$  e applicando in maniera opportuna una funzione  $g$  che preservi la distribuzione uniforme, è in linea teorica possibile costruire distribuzioni con strutture di dipendenza arbitrarie. In particolare, assumiamo per semplicità che la copula di  $(U, V)$  descriva una dipendenza monotona e che sia dipendente da un parametro  $\theta$ . Indichiamo quindi tale copula con  $C_\theta$ . Assumiamo inoltre che  $C_\theta$  includa come caso particolare, o come caso limite, il confine superiore di Fréchet-Hoeffding, ovvero assumiamo che per qualche  $\theta_0$  si abbia che

$$\lim_{\theta \rightarrow \theta_0} C_\theta(u, v) = M(u, v) \quad \forall u, v \in [0, 1].$$

Ciò significa che più  $\theta$  sarà vicino a  $\theta_0$  più i punti di coordinate  $(U, V)$  tenderanno a concentrarsi vicino alla diagonale del quadrato unitario. Se indichiamo a questo punto con  $Q_\theta$  la copula che lega  $U$  e  $g(V)$ , è evidente che al tendere di  $\theta$  a  $\theta_0$  i punti di coordinate  $(U, g(V))$  tenderanno a concentrarsi in prossimità del grafico di  $g$ . Infatti, il grafico di  $g$  rappresenta per  $Q_\theta$  la curva lungo la quale i punti saranno concentrati nel caso di perfetta dipendenza, ovvero nel caso in cui  $\theta = \theta_0$ . Nell'articolo di Scarsini e Venetoulas [58] sono proposti diversi esempi di funzioni non monotone che preservano la distribuzione uniforme. Un esempio è dato dalla funzione

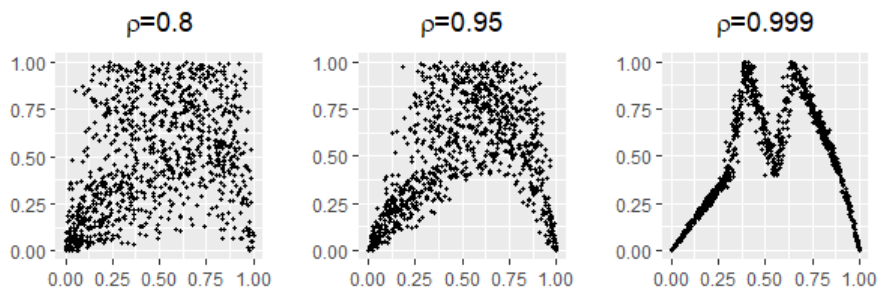
$$g(v) = \begin{cases} \frac{(1 - \zeta + \beta)}{\beta} v & \text{se } 0 \leq v \leq \beta \\ (1 - \zeta + \beta) + \frac{(\zeta - \beta)}{(\gamma - \beta)}(v - \beta) & \text{se } \beta < v \leq \gamma \\ (1 - \zeta + \beta) + \frac{(\zeta - \beta)}{(\delta - \gamma)}(\delta - v) & \text{se } \gamma < v \leq \delta \\ (1 - \zeta + \beta) + \frac{(\zeta - \beta)}{(\epsilon - \delta)}(v - \delta) & \text{se } \delta < v \leq \epsilon \\ (1 - \zeta + \beta) + \frac{(\zeta - \beta)}{(\zeta - \epsilon)}(\zeta - v) & \text{se } \epsilon < v \leq \zeta \\ \frac{(1 - \zeta + \beta)}{(1 - \zeta)}(1 - v) & \text{se } \zeta < v \leq 1 \end{cases} \quad (5.1)$$

dove  $0 \leq \beta \leq \gamma \leq \delta \leq \epsilon \leq \zeta \leq 1$ . Come osservato dagli autori, la funzione (5.1) contiene diversi casi particolari. Se infatti è una funzione lineare a tratti e continua quando i parametri sono tutti diversi tra loro, diventa per esempio discontinua nel caso particolare in cui  $\beta = \gamma = \delta = \epsilon$  e  $\zeta = 1$ . Le figure 5.1 e 5.2 riportano i risultati di alcune simulazioni Monte Carlo di copule costruite mediante l'utilizzo della funzione (5.1), sia nel caso in cui questa è continua sia nel caso in cui è presente una discontinuità.

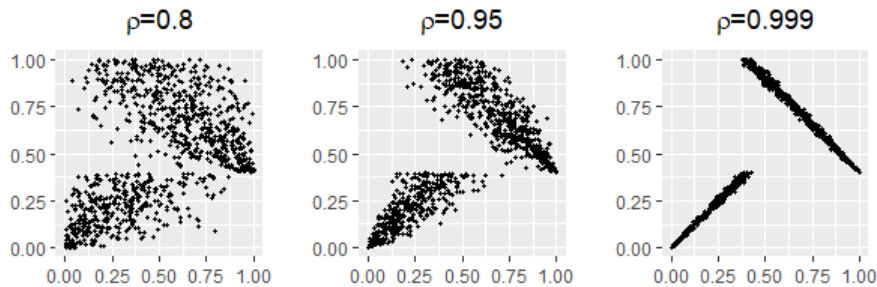
Come ulteriore esempio, consideriamo la funzione

$$g(v) = \begin{cases} 2v & \text{se } 0 \leq v \leq \frac{1}{2} \\ 2v - 1 & \text{se } \frac{1}{2} < v \leq 1. \end{cases} \quad (5.2)$$

Si verifica facilmente che la funzione (5.2) preserva la distribuzione uniforme. Un'e-semplificazione di una copula costruita mediante la trasformazione (5.2) è data nella figura 5.3.

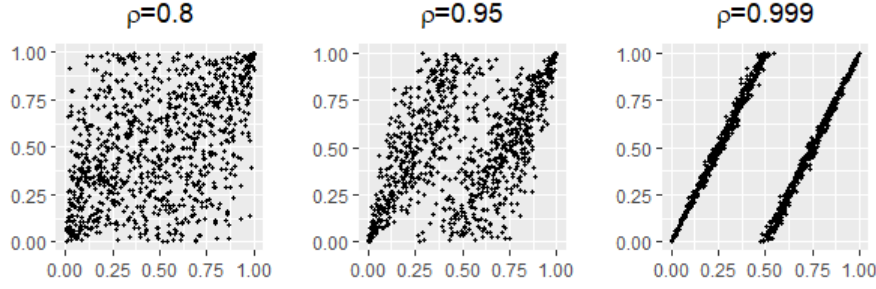


**Figura 5.1.** Simulazioni Monte Carlo della coppia  $(U, g(V))$ , dove la variabile  $(U, V)$  ha marginali uniformi e copula Gaussiana con correlazione  $\rho$  e  $g$  è la funzione (5.1). I parametri della funzione sono  $\beta = 0.3$ ,  $\gamma = 0.4$ ,  $\delta = 0.55$ ,  $\epsilon = 0.65$ ,  $\zeta = 0.9$ . La simulazione è stata effettuata per diversi valori di  $\rho$  e per ciascuno dei grafici il numero di punti simulati è pari a 1000.



**Figura 5.2.** Simulazioni Monte Carlo della coppia  $(U, g(V))$ , dove la variabile  $(U, V)$  ha marginali uniformi e copula Gaussiana con correlazione  $\rho$  e  $g$  è la funzione (5.1) con parametri  $\beta = \gamma = \delta = \epsilon = 0.4$  e  $\zeta = 1$ . La simulazione è stata effettuata per diversi valori di  $\rho$  e per ciascuno dei grafici il numero di punti simulati è pari a 1000.

È opportuno osservare che tutti gli esempi proposti fino a questo momento riguardano funzioni lineari a tratti. Lo studio di un metodo per la costruzione di funzioni preservanti la distribuzione uniforme che abbiano una forma più generale è un problema di interesse pratico oltre che teorico. Infatti, l'uso di funzioni più generali permetterebbe di costruire copule più flessibili e di conseguenza più idonee a descrivere fenomeni complessi.



**Figura 5.3.** Simulazioni Monte Carlo della coppia  $(U, g(V))$ , dove la variabile  $(U, V)$  ha marginali uniformi e copula Gaussiana con correlazione  $\rho$  e  $g$  è la funzione (5.2). La simulazione è stata effettuata per diversi valori di  $\rho$  e per ciascuno dei grafici il numero di punti simulati è pari a 1000.

## 5.1 Costruzione di copule per dipendenze non monotone

Presentiamo a questo punto il principale risultato del presente elaborato, che fornisce indirettamente un metodo per la costruzione di funzioni preservanti la distribuzione uniforme.

**Teorema 5.1.1.** *Sia  $C$  una copula e  $f : [0, 1] \rightarrow [0, 1]$  una funzione che gode delle seguenti proprietà:*

1.  $f$  è monotona crescente con  $f(1) = 1$
2.  $f(v) \geq v \quad \forall v \in [0, 1]$
3.  $f$  è lipschitziana e la sua costante di Lipschitz è pari a 1, ovvero

$$\left| \frac{f(v_2) - f(v_1)}{v_2 - v_1} \right| \leq 1 \quad \forall v_1, v_2 \in [0, 1].$$

Allora la funzione

$$Q(u, v) = C(u, f(v)) - C(u, f(v) - v) \quad (5.3)$$

è una copula.

*Dimostrazione.* Utilizzando la formula (5.3) e la definizione 2.2.2, si verifica immediatamente che  $Q(u, 0) = Q(0, v) = 0$  e  $Q(1, v) = v$ . Inoltre, dato che  $f(1) = 1$ , si ha che  $Q(u, 1) = u$ . Per completare la dimostrazione è quindi sufficiente verificare che

$$V_Q([u_1, u_2] \times [v_1, v_2]) \geq 0$$

per  $u_1 \leq u_2$  e  $v_1 \leq v_2$ . Osserviamo innanzitutto che

$$V_Q([u_1, u_2] \times [v_1, v_2]) = V_C([u_1, u_2] \times [f(v_2) - v_2, f(v_2)]) - V_C([u_1, u_2] \times [f(v_1) - v_1, f(v_1)]). \quad (5.4)$$

Il fatto che  $f$  ha costante di Lipschitz pari a 1 unitamente alla monotonia di  $f$  implica la disuguaglianza  $f(v_2) - v_2 \leq f(v_1) - v_1$ , da cui segue che

$$[f(v_1) - v_1, f(v_1)] \subseteq [f(v_2) - v_2, f(v_2)].$$

È quindi possibile riscrivere la formula (5.4) nella forma

$$\begin{aligned} V_Q([u_1, u_2] \times [v_1, v_2]) \\ = V_C([u_1, u_2] \times [f(v_2) - v_2, f(v_1) - v_1]) + V_C([u_1, u_2] \times [f(v_1), f(v_2)]) \geq 0 \end{aligned}$$

che completa la dimostrazione.  $\square$

Il teorema 5.1.1 fornisce un metodo per la costruzione di una copula  $Q$  a partire da una data copula  $C$  e una funzione  $f$  che soddisfa opportune proprietà. La caratteristica fondamentale del metodo proposto è che permette di ottenere copule che descrivono dipendenze non monotone. Per illustrare questo punto, si consideri una coppia  $(U, V)$  di variabili uniformi su  $[0, 1]$  avente copula  $C$ . Per una data funzione  $f$  che soddisfi le ipotesi del teorema 5.1.1, si definisca la copula

$$Q(u, v) = C(u, f(v)) - C(u, f(v) - v).$$

Indichiamo ora con  $g^+(x)$  e  $g^-(x)$  le pseudo-inverse rispettivamente di  $f(v)$  e  $f(v) - v$ , ovvero

$$g^+(x) = \inf\{v \in [0, 1] : f(v) \geq x\}$$

e

$$g^-(x) = \sup\{v \in [0, 1] : f(v) - v \geq x\}.$$

Ponendo

$$g(x) = \begin{cases} g^+(x) & x > f(0) \\ g^-(x) & x \leq f(0) \end{cases} \quad (5.5)$$

si può scrivere

$$\begin{aligned} Q(u, v) &= C(u, f(v)) - C(u, f(v) - v) = \mathbb{P}(U \leq u, f(v) - v < V \leq f(v)) \\ &= \mathbb{P}(U \leq u, f(v) - v < V \leq f(v) \mid V > f(0)) \cdot \mathbb{P}(V > f(0)) \\ &\quad + \mathbb{P}(U \leq u, f(v) - v < V \leq f(v) \mid V \leq f(0)) \cdot \mathbb{P}(V \leq f(0)) \\ &= \mathbb{P}(U \leq u, V \leq f(v) \mid V > f(0)) \cdot \mathbb{P}(V > f(0)) \\ &\quad + \mathbb{P}(U \leq u, f(v) - v < V \mid V \leq f(0)) \cdot \mathbb{P}(V \leq f(0)) \end{aligned} \quad (5.6)$$

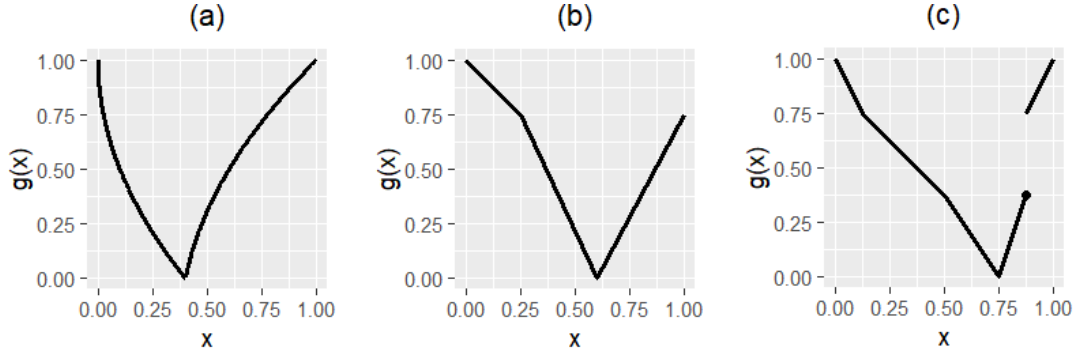
$$\begin{aligned} &= \mathbb{P}(U \leq u, g^+(V) \leq v \mid V > f(0)) \cdot \mathbb{P}(V > f(0)) \\ &\quad + \mathbb{P}(U \leq u, g^-(V) < v \mid V \leq f(0)) \cdot \mathbb{P}(V \leq f(0)) \end{aligned} \quad (5.7)$$

$$= \mathbb{P}(U \leq u, g(V) \leq v). \quad (5.8)$$

Si osservi che, nel passaggio dalla formula (5.6) alla (5.7), le funzioni  $g^+$  e  $g^-$  sono state trattate come inverse rispettivamente delle funzioni  $v \mapsto f(v)$  e  $v \mapsto f(v) - v$ . Se queste ultime non sono invertibili, è tuttavia vero che, con probabilità 1, le variabili aleatorie  $g^+(V)$  e  $g^-(V)$  assumono valori su insiemi dove le inverse di  $v \mapsto f(v)$  e  $v \mapsto f(v) - v$  esistono, come si può dimostrare con argomenti simili a quelli utilizzati per il teorema 2.2.6.

La formula (5.8) richiede ora alcuni commenti. Innanzitutto, il teorema 5.1.1 garantisce che  $Q$  sia una copula e ciò implica, in virtù della (5.8), che la variabile  $g(V)$  ha distribuzione uniforme sull'intervallo  $[0, 1]$ . In altre parole, la funzione  $g$  è per costruzione una funzione che preserva la distribuzione uniforme. Sono inoltre

necessarie delle osservazioni sulla forma della funzione  $g$ . Si può verificare facilmente che nei casi banali in cui  $f(0) = 1$  o  $f(0) = 0$  la funzione  $g$  è lineare. Se si escludono questi due casi banali, è possibile osservare che il comportamento di  $g$  ricorda quello di una parabola. Essa assume infatti valore nullo in corrispondenza del punto  $f(0)$ , a destra e a sinistra del quale esibisce due rami verso l'alto che possono eventualmente presentare delle discontinuità di salto (si veda, per esempio, la figura 5.4).



**Figura 5.4.** Alcuni esempi di funzioni che preservano la distribuzione uniforme ottenute mediante il teorema 5.1.1 e la formula (5.5) per diverse scelte di  $f$ :

- (a)  $f(v) = \frac{1}{5}(2v^2 + v + 2)$   
 (b)  $f(v) = \min\left(1, \frac{8}{15}v + \frac{3}{5}\right)$   
 (c)  $f(v) = \min\left(\max\left(\frac{7}{8}, \frac{1+v}{2}\right), \frac{v}{3} + \frac{3}{4}\right)$ .

Il comportamento della funzione è quindi non monotono. Ciò implica che, se la copula  $C$  descrive la distribuzione di due variabili stocasticamente crescenti (o decrescenti) l'una rispetto all'altra, la copula  $Q$  descrive una dipendenza non monotona. Tuttavia, la struttura di  $g$  sembra a prima vista non essere sufficientemente generale da poter cogliere strutture di dipendenza più complesse. In realtà, dipendenze più complicate possono essere descritte applicando in maniera iterativa il teorema 5.1.1 o, equivalentemente, componendo tra loro funzioni preservanti la distribuzione uniforme aventi una struttura semplice come quella della formula (5.5). Si considerino, a scopo esemplificativo, le funzioni

$$f_1(v) = 1, \quad f_2(v) = \max\left(v, \frac{\zeta + \gamma - \beta - \epsilon}{\zeta - \beta}v + (\epsilon - \gamma)\right),$$

$$f_3(v) = \begin{cases} \frac{(\epsilon - \delta)}{(\epsilon - \gamma)}v + \delta & \text{se } 0 \leq v \leq \epsilon - \gamma \\ \frac{(\zeta - \epsilon)}{(\zeta + \gamma - \epsilon - \beta)}v - \frac{(\epsilon - \gamma)(\zeta - \epsilon)}{(\zeta + \gamma - \epsilon - \beta)} + \epsilon & \text{se } \epsilon - \gamma < v \leq \zeta - \beta \\ \frac{(1 - \zeta)}{(1 - \zeta + \beta)}v + \frac{\beta}{(1 - \zeta + \beta)} & \text{se } \zeta - \beta < v \leq 1. \end{cases}$$

che soddisfano, per  $0 \leq \beta \leq \gamma \leq \delta \leq \epsilon \leq \zeta \leq 1$ , le ipotesi del teorema 5.1.1. Definendo come nella formula (5.5) le corrispondenti funzioni preservanti la distribuzione uniforme

$$g_j(x) = \begin{cases} g_j^+(x) & x > f_j(0) \\ g_j^-(x) & x \leq f_j(0) \end{cases} \quad j = 1, 2, 3$$

dove  $g_j^+(x)$  e  $g_j^-(x)$  rappresentano le pseudo-inverse rispettivamente di  $f_j(v)$  e  $f_j(v) - v$ , si può verificare che la funzione

$$g = g_1 \circ g_2 \circ g_3$$

coincide con la funzione (5.1) esaminata da Scarsini e Venetoulas [58]. Analogamente, considerando le funzioni

$$f_1(v) = \max(1, v), \quad f_2(v) = \frac{1+v}{2}$$

e definendo le relative funzioni  $g_j$ ,  $j = 1, 2$  come nella (5.5), si ha che la funzione

$$g = g_1 \circ g_2$$

coincide con la funzione (5.2).

Possiamo quindi concludere che il teorema 5.1.1 fornisce uno strumento flessibile per la costruzione di copule che descrivono dipendenze di forma arbitraria. È naturalmente possibile costruire tali copule in forma parametrica semplicemente costruendo in forma parametrica le funzioni che preservano la distribuzione uniforme.

## 5.2 Stima dei parametri e selezione del modello

Si consideri una copula  $C_\theta : [0, 1]^2 \rightarrow [0, 1]$  che dipende da un vettore di parametri  $\theta$ . Sia inoltre  $f_\alpha : [0, 1] \rightarrow [0, 1]$  una funzione dipendente dal vettore di parametri  $\alpha$ . Se  $f_\alpha$  soddisfa le ipotesi del teorema 5.1.1, è possibile definire una copula  $Q_{\alpha, \theta}$  come

$$Q_{\alpha, \theta}(u, v) = C_\theta(u, f_\alpha(v)) - C_\theta(u, f_\alpha(v) - v) \quad u, v \in [0, 1] \quad (5.9)$$

Assumendo che  $C_\theta$  abbia densità di probabilità  $c_\theta$ , e sotto l'ipotesi che  $f_\alpha$  sia derivabile, è possibile esprimere la densità di  $Q_{\alpha, \theta}$  nella forma

$$q_{\alpha, \theta}(u, v) = [c_\theta(u, f_\alpha(v)) - c_\theta(u, f_\alpha(v) - v)] \frac{df_\alpha(v)}{dv} + c_\theta(u, f_\alpha(v) - v). \quad (5.10)$$

Grazie alla rappresentazione (5.10), è possibile effettuare una stima di massima verosimiglianza dei parametri  $\alpha$  e  $\theta$  della copula (5.9). Si consideri il campione  $(u_1, v_1), (u_2, v_2), \dots, (u_n, v_n)$  di  $n$  osservazioni provenienti dalla distribuzione bivariata  $Q_{\alpha, \theta}$  dove i parametri  $\alpha, \theta$  sono ignoti. Massimizzando la funzione di log-verosimiglianza

$$\ell(\alpha, \theta; \mathbf{u}, \mathbf{v}) = \sum_{k=1}^n \log q_{\alpha, \theta}(u_k, v_k)$$

rispetto ai parametri  $\alpha, \theta$ , si ottiene la stima di massima verosimiglianza

$$(\hat{\alpha}, \hat{\theta}) = \arg \max_{\alpha, \theta} \ell(\alpha, \theta; \mathbf{u}, \mathbf{v}).$$

Naturalmente, nelle applicazioni pratiche non è generalmente nota la distribuzione da cui provengono i dati osservati. Per questa ragione, è opportuno mettere a confronto diverse ipotesi modellistiche ed effettuare, per ciascuno dei modelli presi in



considerazione, una diversa stima dei parametri. Diviene quindi necessario adottare un criterio di selezione del modello. Vengono incontro a tale esigenza i due indicatori noti come *Akaike information criterion* (AIC) e *Bayesian information criterion* (BIC), introdotti rispettivamente da Akaike [1] e Schwarz [61] e definiti come

$$\text{AIC} = 2k - 2\hat{\ell}$$

e

$$\text{BIC} = \log(n) k - 2\hat{\ell}$$

dove  $\hat{\ell} = \ell(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\theta}}; \mathbf{u}, \mathbf{v})$  e  $k$  è il numero di parametri del modello. Il criterio di selezione del modello basato sugli indicatori AIC e BIC prevede che il modello con punteggi AIC e BIC più basso sia preferibile rispetto agli altri. Nella pratica, viene quindi effettuata la stima dei parametri per diverse copule  $Q_{\boldsymbol{\alpha}, \boldsymbol{\theta}}$  aventi la forma (5.9) e, per ciascuna delle copule stimate, viene calcolato il corrispondente punteggio AIC o BIC. Il modello di copula da selezionare sarà quello con il punteggio più basso. Il criterio così definito tende a privilegiare i modelli che presentano una verosimiglianza più elevata, e quindi una maggior capacità di adattarsi ai dati, penalizzando però i modelli che presentano un maggior numero di parametri allo scopo di favorire la scelta di un modello più parsimonioso. Come è evidente, l'indicatore BIC prevede, rispetto all'AIC, una penalizzazione maggiore dei modelli che hanno un numero elevato di parametri.

## Capitolo 6

# La relazione rendimento-volume in finanza

Abbiamo illustrato, nei capitoli precedenti, i principali aspetti teorici riguardanti lo studio delle dipendenze non monotone. Nei mercati finanziari, un fenomeno in cui si osserva tipicamente una dipendenza non monotona è la relazione tra rendimento e volume degli scambi di un titolo negoziato. La letteratura sulla relazione rendimento-volume ha avuto origine con il lavoro di Osborne [56], in cui la dinamica del prezzo di un titolo fu modellizzata come un processo diffusivo la cui varianza dipende dal numero di transazioni. Come osservato da Karpoff [39], Osborne ipotizzò che le transazioni fossero uniformemente distribuite nel tempo e riuscì quindi ad esprimere la distribuzione del prezzo del titolo in termini dell'ampiezza degli intervalli temporali. Non ottenne quindi un risultato che descrivesse in maniera esplicita il legame tra volume e rendimenti. In un articolo successivo, Clark [11] propose una variante del modello di Osborne in cui la velocità delle transazioni era modellizzata subordinando il processo del prezzo con un subordinatore stabile di Lévy. Trattando il volume dei rendimenti come una misura del numero di scambi, l'autore giunse alla conclusione che il volume degli scambi e il valore assoluto dei rendimenti sono positivamente correlati. Inoltre, l'autore supportò tale conclusione con uno studio empirico su volumi e rendimenti dei futures sul cotone. Diversi studi empirici basati sull'uso di modelli di regressione sono stati effettuati, per esempio, da Crouch [13], Tauchen e Pitts [66] e Lee e Rui [45]. Gündüz e Hatemi-J [30] hanno proposto uno studio sulle relazioni causali tra prezzi e volumi nei mercati finanziari di Ungheria, Repubblica Ceca, Russia, Polonia e Turchia, con risultati che variano da paese a paese. L'uso delle copule per la descrizione della relazione tra rendimento e volume è stato proposto per la prima volta da Neto [55]. L'autore ha esaminato la relazione tra rendimenti e volumi relativamente a titoli azionari emessi da diverse società francesi. In particolare, Neto ha stimato, per ognuna delle società prese in considerazione, il parametro  $\rho$  della copula

$$C_\rho(u, v) = \Phi_\rho\left(\Phi^{-1}(u), \sqrt{\chi^{-1}(v)}\right) - \Phi_\rho\left(\Phi^{-1}(u), -\sqrt{\chi^{-1}(v)}\right)$$

che, come abbiamo visto nel paragrafo 4.3, si ottiene costruendo una dipendenza di tipo quadratico una una variabile normale standard e una con distribuzione

chi-quadro. In questo capitolo generalizzeremo il lavoro di Neto prendendo in considerazione copule più generali costruite mediante l'utilizzo del teorema 5.1.1.

## 6.1 Dati e risultati

Consideriamo la serie storica delle osservazioni giornaliere dei prezzi di chiusura e dei volumi degli scambi per le quote del fondo QQQ, un fondo d'investimento gestito da Invesco Ltd. che replica l'andamento dell'indice Nasdaq 100. In particolare, consideriamo  $n + 1$  osservazioni, con  $n = 624$ , di dati relativi al periodo compreso tra il 1° febbraio 2020 il 24 giugno 2022. Indichiamo con  $p_k$ ,  $k = 0, \dots, n$ , la  $k$ -esima osservazione del prezzo di chiusura giornaliero e calcoliamo il corrispondente rendimento giornaliero  $r_k$  come

$$r_k = \frac{p_k - p_{k-1}}{p_{k-1}}, \quad k = 1, \dots, n.$$

Indichiamo quindi con  $\mathbf{r} = (r_1, \dots, r_n)$  i rendimenti giornalieri e con  $\mathbf{w} = (w_1, \dots, w_n)$  i corrispondenti volumi giornalieri degli scambi. Allo scopo di studiare la dipendenza tra rendimenti e volumi, applichiamo innanzitutto la trasformazione integrale di probabilità (si veda, per esempio, Joe [36], pag. 38) che consente di trasformare le distribuzioni di  $\mathbf{r}$  e  $\mathbf{w}$  in distribuzioni uniformi. Ciò permette di conseguenza di studiare la dipendenza mediante la metodologia descritta nel paragrafo 5.2. Utilizziamo la versione empirica della trasformazione integrale di probabilità introdotta nel paragrafo 2.2 e calcoliamo quindi le pseudo-osservazioni

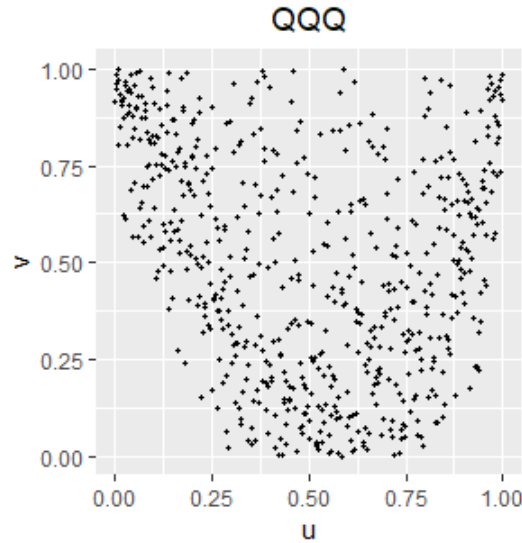
$$u_k = \frac{1}{n+1} \sum_{j=1}^n \mathbb{1}(r_j \leq r_k) = \frac{n}{n+1} \widehat{F}_r(r_k), \quad k = 1, \dots, n$$

$$v_k = \frac{1}{n+1} \sum_{j=1}^n \mathbb{1}(w_j \leq w_k) = \frac{n}{n+1} \widehat{F}_w(w_k), \quad k = 1, \dots, n$$

dove  $\widehat{F}_r(\cdot)$  e  $\widehat{F}_w(\cdot)$  sono rispettivamente la funzione di ripartizione empirica dei rendimenti e quella dei volumi.

La figura 6.1 illustra il grafico dei punti aventi per coordinate le pseudo-osservazioni dei rendimenti giornalieri  $\mathbf{u} = (u_1, \dots, u_n)$  e quelle dei volumi degli scambi  $\mathbf{v} = (v_1, \dots, v_n)$ . Se nel caso di dipendenze monotone ci si aspetta che i punti tendano ad avvicinarsi lungo una delle diagonali del quadrato unitario, in questo caso i punti tendono a disporsi lungo una curva che ha un comportamento di tipo parabolico. In particolare, si osserva che il volume tende ad assumere valori elevati sia in corrispondenza di rendimenti bassi che in corrispondenza di rendimenti alti. Questo comportamento è coerente con la maggior parte della letteratura esistente sulla relazione rendimento-volume. Ovviamente, le classiche copule che descrivono dipendenze monotone non sono adatte ad esaminare il fenomeno in analisi.

Allo scopo di descrivere la dipendenza non monotona che emerge dal dataset considerato, utilizziamo il teorema 5.1.1 per costruire copule adeguate. Effettueremo quindi la stima dei parametri per copule aventi la forma descritta dall'equazione (5.9) per diverse scelte di  $C_\theta$  and  $f_\alpha$ . In particolare, confronteremo le copule Frank,



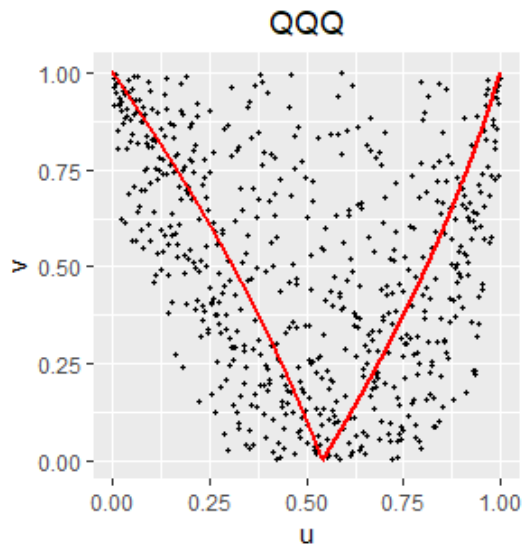
**Figura 6.1.** Grafico dei punti corrispondenti alle pseudo-osservazioni dei rendimenti giornalieri (alle ascisse) e dei volumi giornalieri degli scambi (alle ordinate) per il fondo QQQ. Il grafico suggerisce la presenza di una dipendenza di tipo non monotono.

Clayton, Gumbel, Gaussiana e t. Per ciascuna di queste copule proponiamo tre possibili scelte della funzione  $f$  che determina la forma della dipendenza, ovvero

$$\begin{aligned}
 f_1(v; c) &= \frac{v + c}{1 + c}, & c &\geq 0 \\
 f_2(v; a, c) &= av^2 + (1 - a - c)v + c, & c &\geq 0, 0 \leq a + c \leq 1, 0 \leq 1 + a - c \leq 1 \\
 f_3(v; a, c) &= \left(\frac{v + c}{1 + c}\right)^a, & a &\geq 0, c \geq 0, a \leq 1 + c, \left(\frac{c}{1 + c}\right)^{a-1} \leq \frac{1 + c}{a}
 \end{aligned}$$

dove i vincoli sono imposti in maniera tale da soddisfare le ipotesi del teorema 5.1.1. Effettueremo quindi la stima dei parametri per quindici copule diverse. È opportuno sottolineare che i vincoli imposti sulle funzioni  $f_1$ ,  $f_2$  e  $f_3$  hanno nature diverse tra loro. Infatti, mentre il vincolo su  $f_1$  può essere imposto con una semplice riparametrizzazione, i vincoli per  $f_2$  e  $f_3$  sono rispettivamente di tipo lineare e non lineare. Segue che in ciascun caso è necessario scegliere un algoritmo di ottimizzazione appropriato. I risultati della procedura di stima dei parametri sono riportati nella tabella 6.1, mentre i corrispondenti indici AIC e BIC sono riportati rispettivamente nelle tabelle 6.2 e 6.3. Da queste ultime si evince immediatamente come la funzione quadratica  $f_2$ , combinata con la copula Frank, sia la scelta migliore in termini di AIC e BIC per la descrizione della forma della dipendenza tra rendimenti e volumi per la serie storica considerata. Si osservi a questo punto che la funzione lineare  $f_1$  non è altro che un caso particolare di  $f_2$  che si ottiene ponendo  $a = 0$ . Ciò significa che il criterio di selezione del modello basato sugli indici AIC e BIC giustifica l'introduzione di un parametro aggiuntivo per migliorare la capacità del modello di adattarsi ai dati. Infatti, l'aumento dei valori dell'AIC e del BIC dovuto alla penalizzazione del parametro aggiuntivo è più che compensato dal miglioramento in termini di

verosimiglianza. Nella figura 6.2 mostriamo la funzione preservante la distribuzione uniforme, costruita a partire dalla funzione quadratica  $f_2$ , che meglio approssima, secondo i risultati della stima dei parametri, la dipendenza tra volume e rendimento per il dataset esaminato.



**Figura 6.2.** Grafico dei punti corrispondenti alle pseudo-osservazioni dei rendimenti giornalieri e dei volumi giornalieri degli scambi. La curva in rosso coincide con la funzione preservante la distribuzione uniforme che è implicita nel modello stimato combinando la copula Frank e la funzione quadratica  $f_2$ .

Concludiamo l'analisi osservando che la scelta di mettere a confronto diverse copule di tipo non monotono, escludendo del tutto le copule di tipo monotono dall'analisi, è stata basata esclusivamente sull'osservazione del grafico nella figura 6.1. Per completezza, è opportuno stimare i parametri di copule di tipo monotono per effettuare un confronto e verificare in maniera più rigorosa che le copule che rappresentano dipendenze non monotone sono le più adatte a descrivere i dati considerati. Ciò confermerebbe, coerentemente con la letteratura esistente, la presenza di una dipendenza non monotona tra rendimento e volume. A questo scopo, utilizziamo il pacchetto R `VineCopula` e, in particolare, la funzione `BiCopSelect`. Quest'ultima effettua la stima dei parametri per diverse copule ellittiche, Archimedee e alcune loro generalizzazioni tenendo in considerazione anche le possibili rotazioni delle distribuzioni. In totale, vengono messi a confronto trentasette modelli, per un elenco dei quali si rimanda alla lettura della documentazione di Schepsmeier et al. [59]. La funzione `BiCopSelect` fornisce in output il modello che meglio si adatta ai dati secondo i criteri impostati dall'utente, che nel nostro caso sono l'AIC e il BIC. La migliore tra tutte le copule confrontate, sia in termini di AIC che di BIC, risulta essere la copula di Tawn di tipo 1 ruotata di  $90^\circ$ . La copula di Tawn (si veda Tawn [67] e Bernard e Czado [7]) è una copula a tre parametri  $\theta, \psi_1$ , e  $\psi_2$ , con  $\psi_1, \psi_2 \in [0, 1]$  e  $\theta \in [1, +\infty)$ , definita come

$$C_{\theta, \psi_1, \psi_2}(u, v) = (uv)^{A_{\theta, \psi_1, \psi_2}\left(\frac{\log(u)}{\log(uv)}\right)} \quad (6.1)$$

	$f_1$	$f_2$	$f_3$
Clayton	c=1.179097 $\theta=1.721766$	a=0.003635 c=0.543055 $\theta=1.720576$	a=1.028725 c=1.234383 $\theta=1.720647$
Gumbel	c=2.241228 $\theta=2.265978$	a=-0.226821 c=0.593773 $\theta=2.288596$	a=0.277944 c=0.160345 $\theta=2.294540$
Frank	c=1.460967 $\theta=9.034415$	a=-0.159780 c=0.541587 $\theta=9.085100$	a=0.452935 c=0.337502 $\theta=9.068456$
Gaussian	c=1.815000 $\theta=0.782685$	a=-0.186255 c=0.563700 $\theta=0.784902$	a=0.339790 c=0.181386 $\theta=0.784803$
t	c=1.787246 $\theta=0.775160$ $\nu=5.884596$	a=0.204121 c=0.441193 $\theta=-0.775673$ $\nu=5.257317$	a=0.337119 c=0.183122 $\theta=0.775672$ $\nu=5.045899$

**Tabella 6.1.** Risultati della stima di massima verosimiglianza per le copule ottenute con diverse combinazioni di  $C_\theta$  e  $f_\alpha$ . Nel caso della copula t, il parametro aggiuntivo  $\nu$  rappresenta il numero di gradi di libertà.

	$f_1$	$f_2$	$f_3$
Clayton	-171.1221	-169.1243	-168.6280
Gumbel	-218.6425	-236.0971	-233.3283
Frank	-234.7418	<b>-243.0672</b>	-241.2312
Gaussian	-214.0477	-220.5011	-219.6703
t	-229.4568	-238.4797	-236.7481

**Tabella 6.2.** Valore dell'indice AIC per ciascuno dei modelli stimati.

	$f_1$	$f_2$	$f_3$
Clayton	-162.2498	-155.8158	-155.3196
Gumbel	-209.7702	-222.7886	-220.0199
Frank	-225.8695	<b>-229.7588</b>	-227.9227
Gaussian	-205.1754	-207.1927	-206.3618
t	-216.1483	-220.7351	-219.0035

**Tabella 6.3.** Valore dell'indice BIC per ciascuno dei modelli stimati.

dove

$$A_{\theta, \psi_1, \psi_2}(t) = (\psi_2 - \psi_1)t + (1 - \psi_2) + \left( [\psi_2(1 - t)]^\theta + (\psi_1 t)^\theta \right)^{\frac{1}{\theta}}.$$

Si osservi che per  $\psi_1 = \psi_2 = 1$  si ottiene la copula Gumbel. Quando si parla di copula di Tawn di tipo 1 ci si riferisce alla copula biparametrica che si ottiene ponendo  $\psi_2 = 1$  nella (6.1). Inoltre, a seguito della rotazione di  $90^\circ$ , la copula risultante può essere espressa nella forma

$$C_{\theta, \psi_1}^{Rot}(u, v) = v - C_{\theta, \psi_1, 1}(1 - u, v). \quad (6.2)$$

I parametri stimati dalla funzione `BiCopSelect` per la copula (6.2) sono

$$\theta = 2.5905362, \quad \psi_1 = 0.2815409$$

mentre gli indici AIC e BIC sono

$$\text{AIC} = -155.1409, \quad \text{BIC} = -146.2686.$$

Ribadiamo che la copula selezionata è la migliore, in termini di AIC e BIC, tra trentasette copule di tipo monotono messe a confronto. Tale copula risulta avere valori degli indici AIC e BIC più elevati di quelli di tutte le copule di tipo non monotono che abbiamo stimato in precedenza. Questo risultato supporta la conclusione che ci sia una dipendenza di tipo non monotono tra volume e rendimento.

## Capitolo 7

# Effetto delle temperature sulla mortalità

Diversi studi in ambito demografico ed epidemiologico hanno esaminato l'impatto che temperature elevate hanno sulla salute umana e, di conseguenza, sulla mortalità. La letteratura sul tema, per una revisione della quale facciamo riferimento al lavoro di Basu e Samet [6], è concorde sull'esistenza di un'associazione tra temperature elevate e mortalità, seppur con definizioni non sempre uniformi dei criteri secondo cui la causa di un decesso possa essere attribuita ad una temperatura elevata (si veda, per esempio, Donoghue et al. [15] e Shen et al. [63]). Alcuni studi, come quelli di Semenza et al. [62] e Tertre et al. [68], sono nati a seguito di specifici eventi meteorologici caratterizzati da ondate di calore anomale che hanno avuto un notevole impatto sulle popolazioni colpite. L'analisi dell'effetto delle temperature sulla mortalità diventa ancor più rilevante se contestualizzata nello studio dei cambiamenti climatici, a causa dei quali si prevede che eventi meteorologici estremi saranno sempre più frequenti (Imada et al. [35], Dunn et al. [16], Clarke et al. [12]). Per questa ragione, Gasparrini et al. [26] hanno evidenziato come lo studio della relazione tra temperatura e mortalità possa essere utilizzato per effettuare proiezioni sulla futura sovrarmortalità legata all'aumento delle temperature medie.

Studi analoghi sono stati proposti per l'analisi dell'impatto delle basse temperature sulla mortalità umana. Anche nel caso di periodi freddi, alcuni autori hanno rilevato un incremento della mortalità dovuto principalmente all'aumento dell'incidenza di malattie cardiovascolari (Analitis et al. [2], Barnett et al. [5]). Se la mortalità tende quindi ad aumentare sia nel caso di temperature basse sia nel caso di temperature alte, non sorprende il fatto che diversi autori, tra cui Kovats et al. [42] e Linares e Díaz [47], abbiano identificato una relazione di tipo non monotono tra temperatura e mortalità. Gli stessi autori hanno anche tentato di individuare delle soglie di temperatura a partire dalle quali inizia a manifestarsi l'aumento di mortalità.

Masselot et al. [49] hanno esaminato la relazione tra temperature e mortalità in diverse città europee mediante un approccio regressivo in più fasi. Dopo aver descritto la relazione temperatura-mortalità mediante un'opportuna curva dose-risposta non lineare (Gasparrini et al. [24]) per 232 città, hanno utilizzato un modello di regressione per spiegare i parametri delle curve dose-risposta in funzione di fattori geografici, ambientali e socioeconomici legati a ciascuna città. Prendendo in considerazione



l'intero intervallo delle temperature osservate, e non limitando quindi l'analisi alle sole temperature basse o elevate, gli autori hanno ottenuto curve dose-risposta non monotone che mostrano una mortalità più elevata in corrispondenza di picchi di temperatura sia verso l'alto che verso il basso, giungendo alla conclusione che il numero di decessi causati dal freddo è notevolmente superiore a quello dei decessi causati dalle ondate di calore. Utilizzando un approccio simile, Gasparrini et al. [25] hanno esaminato il fenomeno in diverse località in Europa, America meridionale e settentrionale, Asia e Australia, giungendo anche in questo caso alla conclusione che la dipendenza tra temperatura e mortalità è di tipo non monotono e che il maggior numero di decessi associati a temperature anomale sia da attribuire al freddo. Gli autori hanno inoltre stimato i quantili di minima mortalità, mostrando che questi vanno dal sessantesimo fino al novantesimo percentile variando da paese a paese. In particolare, quantili di minima mortalità vicino al sessantesimo percentile si osservano tipicamente in zone tropicali, mentre percentili vicini al novantesimo sono osservabili nelle aree temperate. Tale conclusione è del tutto ragionevole se si considera il fatto che percentili dello stesso livello possono corrispondere, in aree climatiche tra loro differenti, a temperature molto diverse. È tuttavia necessario tenere in considerazione che la temperatura di minima mortalità può comunque variare da paese a paese per via della naturale capacità di adattamento dell'uomo al clima in cui vive (Ye et al. [70], Heutel et al. [31]).

Lo scopo di questo capitolo è quello di proporre l'uso delle copule per la descrizione e la modellizzazione della dipendenza tra temperatura e mortalità. Inoltre, esaminando i dati relativi al caso italiano, si illustrerà come l'approccio basato sulle copule consente di ricavare informazioni sul quantile di minima mortalità conducendo a conclusioni che sono in linea con la letteratura esistente.

## 7.1 Dati e risultati

Al fine di esaminare la relazione tra temperature e mortalità in Italia utilizziamo due diverse fonti di dati. Il dataset relativo alle temperature proviene da un'istituzione privata ed è costruito sulla base delle rilevazioni delle temperature massime, minime e medie giornaliere fornite a livello comunale da diverse stazioni di rilevamento dal 1/1/2010 al 31/10/2022. I dati sono riferiti a 1021 comuni appartenenti a quindici differenti province: Agrigento, Bologna, Brescia, Catania, Cuneo, Foggia, Forlì-Cesena, Latina, Matera, Perugia, Piacenza, Siena, Sondrio, Trapani e Verona. Dato che per alcuni comuni le rilevazioni sono effettuate da più stazioni, il dataset riporta, per ciascun comune, il minimo, la media aritmetica e il massimo dei valori rilevati dalle diverse stazioni. Pertanto, il dataset fornisce per ciascun comune i seguenti valori:

- il minimo delle temperature minime giornaliere rilevate dalle diverse stazioni
- la media aritmetica delle temperature minime giornaliere rilevate dalle diverse stazioni
- il massimo delle temperature minime giornaliere rilevate dalle diverse stazioni
- il minimo delle temperature medie giornaliere rilevate dalle diverse stazioni

- la media aritmetica delle temperature medie giornaliere rilevate dalle diverse stazioni
- il massimo delle temperature medie giornaliere rilevate dalle diverse stazioni
- il minimo delle temperature massime giornaliere rilevate dalle diverse stazioni
- la media aritmetica delle temperature massime giornaliere rilevate dalle diverse stazioni
- il massimo delle temperature massime giornaliere rilevate dalle diverse stazioni.

Lo scopo dell'analisi è quello di indagare l'effetto della temperatura sulla mortalità tenendo in considerazione sia il caso di temperature basse che quello di temperature elevate. Per questa ragione, è sufficiente considerare dei valori di temperatura che siano rappresentativi dell'intera giornata e indicativi di quanto una data giornata sia stata complessivamente più fredda o più calda delle altre. Ai fini dell'analisi utilizzeremo quindi la media aritmetica delle temperature medie giornaliere.

I dati relativi alla mortalità sono invece pubblici e provengono dal sito dell'ISTAT. In particolare, utilizziamo i dati relativi al numero di decessi giornalieri che l'ISTAT fornisce per ciascun comune. Osserviamo tuttavia che, per via della scarsa popolosità di molti comuni di piccole dimensioni, il numero di decessi registrati è spesso esiguo e presenta una distribuzione di probabilità fortemente concentrata su pochi valori bassi con un'elevata frequenza di giorni in cui il numero di decessi registrati è pari a zero. Una simile distribuzione porta, dopo l'applicazione della trasformata integrale di probabilità, a delle pseudo-osservazioni che sono concentrate su pochi punti dell'intervallo unitario rendendo quindi difficile cogliere in maniera efficace eventuali fenomeni di dipendenza. Per questa ragione, si è scelto di aggregare i decessi a livello provinciale sommando i decessi registrati a livello comunale. Abbiamo in questo modo ottenuto la serie storica del numero di decessi giornalieri per le quindici province sopra menzionate. Naturalmente, va tenuto conto del fatto che tale scelta porta necessariamente ad una perdita di informazione dovuta al fatto che comuni appartenenti alla stessa provincia possono presentare, dal punto di vista topografico e climatico, delle differenze tra loro.

Un'ulteriore aggregazione dei dati viene effettuata rispetto alla dimensione temporale. La necessità di tale aggregazione nasce dal fatto che l'impatto delle temperature sulla mortalità ha una durata che varia da pochi giorni a diverse settimane, tipicamente con durate maggiori nel caso di temperature basse (Chen et al. [10], Gasparrini et al. [25]). Inoltre, l'effetto si manifesta generalmente con un ritardo di alcuni giorni rispetto al momento in cui si è registrato il picco di temperatura (Gasparrini et al. [25]). Per le ragioni appena descritte, l'utilizzo di dati giornalieri non consentirebbe di cogliere la dipendenza tra temperature e mortalità. Ai fini dell'analisi abbiamo scelto di aggregare i dati a livello mensile. Per ciascuna provincia abbiamo quindi calcolato il numero di decessi mensili sommando il numero di decessi giornalieri. Per quanto riguarda le temperature, la temperatura media mensile è stata ottenuta come media aritmetica delle temperature medie giornaliere. Anche in questo caso, l'aggregazione comporta naturalmente una perdita di informazione. Ciò è dovuto al fatto che un dato valore di temperatura media mensile non fornisce necessariamente

un'indicazione precisa sulla presenza o meno di un picco di temperatura nel mese, impedendo quindi di cogliere in maniera efficace l'impatto dell'eventuale picco sul numero di decessi. Va tuttavia osservato che, secondo alcuni autori, l'impatto sulla mortalità di picchi di temperature anomale è minore di quello di temperature più moderate ma non ottimali che si protraggono nel tempo (Gasparrini et al. [25]).

Per concludere, è necessario tenere in considerazione il fatto che il numero assoluto di decessi è correlato alla popolosità di ciascuna provincia. Per questo motivo dividiamo il numero di decessi provinciali per il numero annuo di cittadini residenti in ciascuna provincia, ottenendo in questo modo dei tassi di mortalità mensili provinciali. I dati relativi alla popolazione residente sono reperibili sul sito dell'ISTAT.

L'analisi è stata effettuata separatamente per nord, centro e sud Italia per la popolazione di sesso femminile, di sesso maschile e totale. Nelle figure 7.1, 7.2 e 7.3 sono illustrati i punti corrispondenti alle pseudo-osservazioni delle temperature medie mensili e dei corrispondenti tassi mensili di mortalità per i diversi sottoinsiemi della popolazione presi in considerazione. La relazione non monotona tra temperatura e mortalità è particolarmente evidente al sud Italia per entrambi i generi, mentre tende a diventare meno evidente procedendo verso il nord Italia. Per le diverse aree geografiche e per entrambi i generi, procediamo a stimare i parametri di modelli che descrivono sia dipendenze monotone che dipendenze non monotone. La selezione del modello sarà poi effettuata mediante l'uso del BIC per favorire la scelta di modelli più parsimoniosi dal punto di vista del numero di parametri.

La stima dei parametri per i modelli di dipendenza monotona viene effettuata in R mediante la funzione `BiCopSelect`. Le tabelle 7.1, 7.2 e 7.3 riportano i risultati relativi alle copule di dipendenza monotona migliori in termini di BIC. Sottolineiamo che le copule ottenute sono le più performanti anche in termini di AIC. Per completezza abbiamo quindi riportato nelle tabelle anche i valori dell'indicatore AIC. Per tutti i sottoinsiemi della popolazione presi in esame, la copula di dipendenza monotona che meglio si adatta ai dati è la copula BB8 ruotata di  $90^\circ$ . La copula BB8 (si veda Joe [37], pag. 153) è la copula Archimedeana biparametrica, con parametri  $\theta \geq 1$  e  $\delta \in (0, 1]$ , avente per generatore la funzione

$$\varphi(t) = -\log \left( \frac{1 - (1 - \delta t)^\theta}{1 - (1 - \delta)^\theta} \right)$$

ed assume pertanto la forma

$$C_{\theta, \delta}(u, v) = \frac{1}{\delta} - \frac{1}{\delta} \left[ 1 - \frac{(1 - (1 - \delta u)^\theta)(1 - (1 - \delta v)^\theta)}{1 - (1 - \delta)^\theta} \right]^{\frac{1}{\theta}}.$$

A seguito della rotazione di  $90^\circ$  la copula risultante è quindi

$$C_{\theta, \delta}^{Rot}(u, v) = v - C_{\theta, \delta}(1 - u, v).$$

Per quanto riguarda le copule di dipendenza non monotona abbiamo messo a confronto le copule Frank, Clayton, Gumbel, Gaussiana e t combinate con le tre possibili

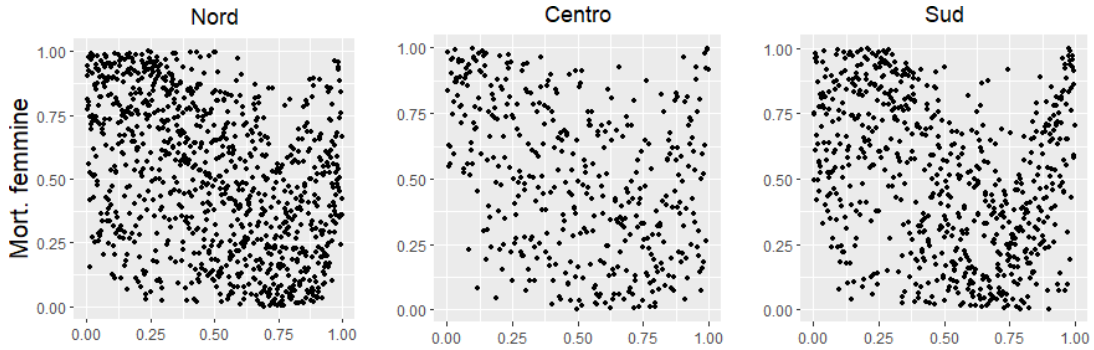
scelte della funzione  $f$ , che determina la forma della dipendenza, proposte nel capitolo 6. Le copule migliori in termini di BIC per ciascun sottinsieme della popolazione sono riportate nelle tabelle 7.4, 7.5 e 7.6. In particolare, per il nord e il sud Italia la copula migliore tra quelle confrontate risulta essere la Clayton combinata con la funzione quadratica  $f_2$  sia per la popolazione divisa per generi sia per la popolazione totale. Tale copula è più performante delle altre anche rispetto al criterio AIC. Nel caso del centro Italia, la copula Clayton combinata con la funzione  $f_2$  è la migliore rispetto al criterio AIC, mentre in termini di BIC la copula più performante è la Gaussiana combinata con la funzione lineare  $f_1$ . Abbiamo quindi selezionato quest'ultima per favorire la scelta di un modello con un numero minore di parametri.

In tutti i casi, si osserva che alle copule di dipendenza non monotona corrispondono valori di indici AIC e BIC minori rispetto a quelli ottenuti per le copule di dipendenza non monotona, con differenze particolarmente marcate per il sud Italia. Possiamo concludere che i modelli di dipendenza non monotona sono più adatti a descrivere i dati considerati, confermando quindi l'esistenza di una relazione non monotona tra temperatura e mortalità.

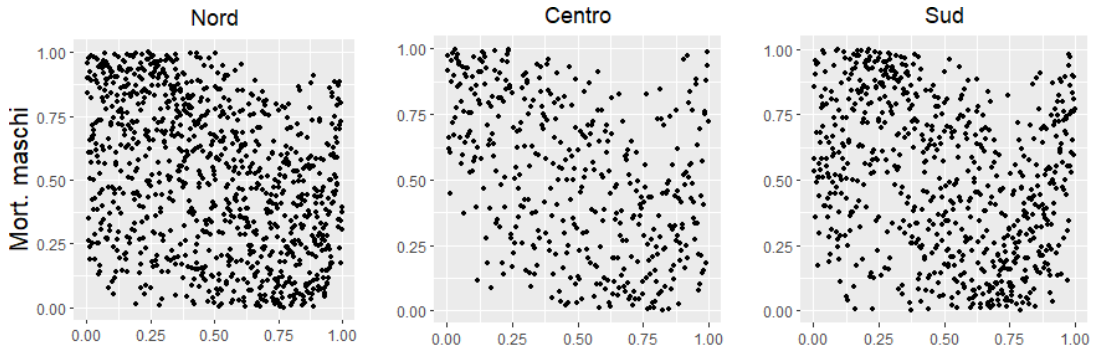
Utilizzando i metodi descritti nel capitolo 5, possiamo infine costruire le funzioni preservanti la distribuzione uniforme che descrivono la forma della dipendenza tra temperatura e mortalità per ciascun sottinsieme della popolazione considerato. I grafici di tali funzioni sono illustrati nelle figure 7.4, 7.5 e 7.6. Al fine di interpretare le funzioni preservanti la distribuzione uniforme ottenute, osserviamo che queste possiedono dei punti di minimo a sinistra dei quali sono funzioni decrescenti e a destra dei quali sono crescenti. Il punto di minimo rappresenta quindi il punto prima del quale la mortalità è funzione decrescente della temperatura e a partire dal quale la mortalità inizia a crescere all'aumentare della temperatura. Pertanto, la stima del punto di minimo della funzione preservante la distribuzione uniforme fornisce un'approssimazione naturale del livello del quantile corrispondente alla temperatura di minima mortalità. Osserviamo che i quantili di minima mortalità sono approssimativamente compresi tra il sessantottesimo e l'ottantaquattresimo, come mostrato nella tabella 7.7. Tale risultato è coerente con quanto osservato da Gasparrini et al. [25] e supporta la conclusione di altri autori come Fu et al. [23] e Zhao et al. [72] secondo cui la maggior parte dei decessi attribuibili a temperature anomale è causata dal freddo. Infatti, se il quantile di minima mortalità si trova a destra della mediana significa che è più probabile osservare un aumento di mortalità causato dal calo delle temperature rispetto a un aumento di mortalità causato da temperature elevate. La mortalità diventa funzione crescente della temperatura solo quando quest'ultima assume valori che si trovano vicini alla coda destra della distribuzione.

Utilizzando le stime dei quantili di minima mortalità e le distribuzioni marginali della temperatura, illustrate nella figura 7.7 per le diverse aree della penisola, è possibile ottenere le stime delle temperature di minima mortalità i cui valori sono riportati nella tabella 7.8. Per i diversi sottinsiemi della popolazione, le temperature di minima mortalità stimate sono comprese tra i 20 e i 23 gradi centigradi circa. Come è ragionevole aspettarsi, le temperature di minima mortalità per la popolazione totale

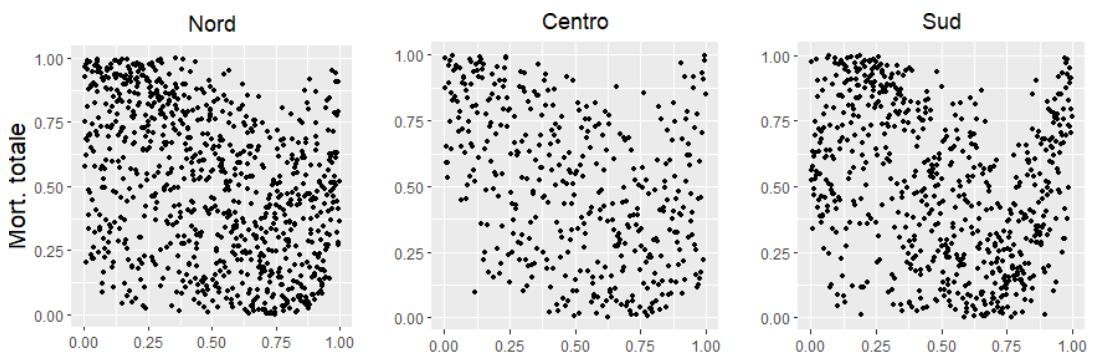
sono comprese tra quelle della popolazione femminile e quelle della popolazione maschile. Non si evidenzia tuttavia un legame chiaro tra l'area geografica e la temperatura di minima mortalità. Alcuni autori (Tobías et al. [69]) hanno osservato che la temperatura di minima mortalità di una data area geografica è correlata alla temperatura media dell'area stessa, suggerendo quindi che vi sia una capacità di adattamento della popolazione alle condizioni meteorologiche e climatiche in cui vive. È plausibile che tale effetto non sia colto nell'analisi qui effettuata per ragioni legate alla qualità dei dati utilizzati, per via del numero esiguo di province considerate e per via delle aggregazioni dei dati sopra discusse con la conseguente perdita di informatività del dataset. Altri autori hanno invece sostenuto che la temperatura media non sia il fattore più adatto a spiegare la temperatura di minima mortalità. Secondo Yin et al. [71] è più opportuno l'utilizzo della temperatura più frequente come variabile esplicativa, mentre Honda et al. [34] hanno proposto l'utilizzo dei quantili di livello compreso tra l'80% e l'85% della temperatura. In ogni caso, le difficoltà relative alla stima della temperatura di minima mortalità sono state evidenziate da diversi autori (Krummenauer et al. [43], Gasparrini et al. [27]). I fattori che concorrono a determinare la capacità di adattamento di una popolazione alle condizioni climatiche sono infatti molteplici e tra questi figurano, per esempio, anche fattori di natura socioeconomica.



**Figura 7.1.** Pseudo-osservazioni di temperatura media mensile (ascissa) e tasso di mortalità mensile (ordinata) per la popolazione di sesso femminile nelle province del nord, centro e sud Italia.



**Figura 7.2.** Pseudo-osservazioni di temperatura media mensile (ascissa) e tasso di mortalità mensile (ordinata) per la popolazione di sesso maschile nelle province del nord, centro e sud Italia.



**Figura 7.3.** Pseudo-osservazioni di temperatura media mensile (ascissa) e tasso di mortalità mensile (ordinata) per la popolazione totale nelle province del nord, centro e sud Italia.

## Popolazione femminile

	Copula (param.)	AIC	BIC
Nord	BB8, rot. 90° $\theta = 3.089545$ $\delta = 0.6836381$	-184.8051	-175.0016
Centro	BB8, rot. 90° $\theta = 1.970302$ $\delta = 0.9095634$	-78.12939	-70.02051
Sud	BB8, rot. 90° $\theta = 2.302423$ $\delta = 0.7351900$	-74.14097	-65.01044

**Tabella 7.1.** La tabella riporta le copule che, tra quelle che descrivono dipendenze monotone, rappresentano al meglio la dipendenza tra temperatura e mortalità per la popolazione femminile al nord, centro e sud Italia.

## Popolazione maschile

	Copula (param.)	AIC	BIC
Nord	BB8, rot. 90° $\theta = 3.104590$ $\delta = 0.6187415$	-134.7286	-124.9251
Centro	BB8, rot. 90° $\theta = 2.380482$ $\delta = 0.8517227$	-95.52117	-87.41229
Sud	BB8, rot. 90° $\theta = 2.728345$ $\delta = 0.6231565$	-68.08302	-58.95249

**Tabella 7.2.** La tabella riporta le copule che, tra quelle che descrivono dipendenze monotone, rappresentano al meglio la dipendenza tra temperatura e mortalità per la popolazione maschile al nord, centro e sud Italia.

## Popolazione totale

	Copula (param.)	AIC	BIC
Nord	BB8, rot. 90° $\theta = 3.076116$ $\delta = 0.6776163$	-178.1819	-168.3785
Centro	BB8, rot. 90° $\theta = 2.180438$ $\delta = 0.8927646$	-95.40733	-87.29845
Sud	BB8, rot. 90° $\theta = 2.553350$ $\delta = 0.6944035$	-82.72518	-73.59465

**Tabella 7.3.** La tabella riporta le copule che, tra quelle che descrivono dipendenze monotone, rappresentano al meglio la dipendenza tra temperatura e mortalità per la popolazione totale al nord, centro e sud Italia.

## Popolazione femminile

	Copula (param.)	Funzione $f_j$ (param.)	AIC	BIC
Nord	Gumbel $\theta = 1.698619832$	$f_2$ , quadratica $a = -0.1576250431$ $c = 0.8124867397$	-208.4702	-193.7650
Centro	Gaussiana $\theta = 0.6892493999$	$f_1$ , lineare $c = 3.5934901292$	-111.61815	-103.50927
Sud	Gumbel $\theta = 1.96543959$	$f_2$ , quadratica $a = -0.16361188708$ $c = 0.68091626332$	-174.03934	-160.34354

**Tabella 7.4.** La tabella riporta le copule che, tra quelle che descrivono dipendenze non monotone, descrivono al meglio la dipendenza tra temperatura e mortalità per la popolazione femminile al nord, centro e sud Italia.

## Popolazione maschile

	Copula (param.)	Funzione $f_j$ (param.)	AIC	BIC
Nord	Gumbel $\theta = 1.5736646950$	$f_2$ , quadratica $a = -0.2114375866$ $c = 0.7885623670$	-159.0600	-144.35478
Centro	Gaussiana $\theta = 0.66871918938$	$f_1$ , lineare $c = 5.3166748601$	-116.48645	-108.37757
Sud	Gumbel $\theta = 1.7741835665$	$f_2$ , quadratica $a = -0.23844642977$ $c = 0.69839684334$	-133.85664	-120.16085

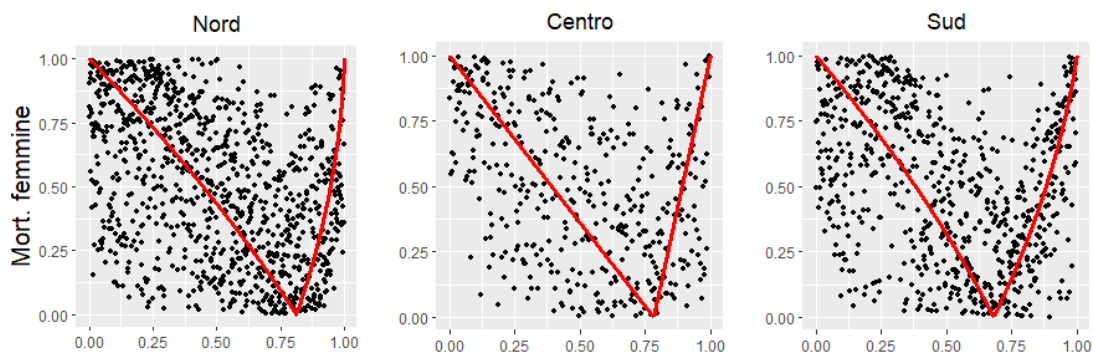
**Tabella 7.5.** La tabella riporta le copule che, tra quelle che descrivono dipendenze non monotone, descrivono al meglio la dipendenza tra temperatura e mortalità per la popolazione maschile al nord, centro e sud Italia.

## Popolazione totale

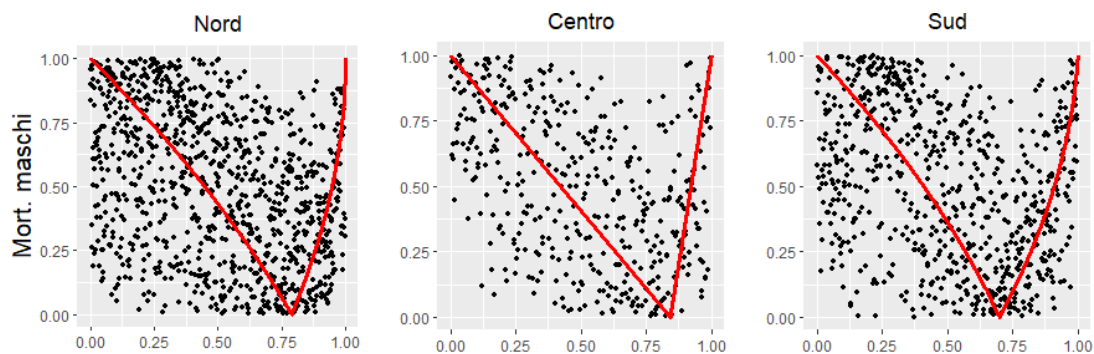
	Copula (param.)	Funzione $f_j$ (param.)	AIC	BIC
Nord	Gumbel $\theta = 1.67282988017$	$f_2$ , quadratica $a = -0.1807137950$ $c = 0.8020956364$	-201.9772	-187.2720
Centro	Gaussiana $\theta = 0.69605580295$	$f_1$ , lineare $c = 4.4019044059$	-126.44639	-118.33751
Sud	Gumbel $\theta = 1.9749529059$	$f_2$ , quadratica $a = -0.2206170614$ $c = 0.6820940681$	-185.85491	-172.15911

**Tabella 7.6.** La tabella riporta le copule che, tra quelle che descrivono dipendenze non monotone, descrivono al meglio la dipendenza tra temperatura e mortalità per la popolazione totale al nord, centro e sud Italia.

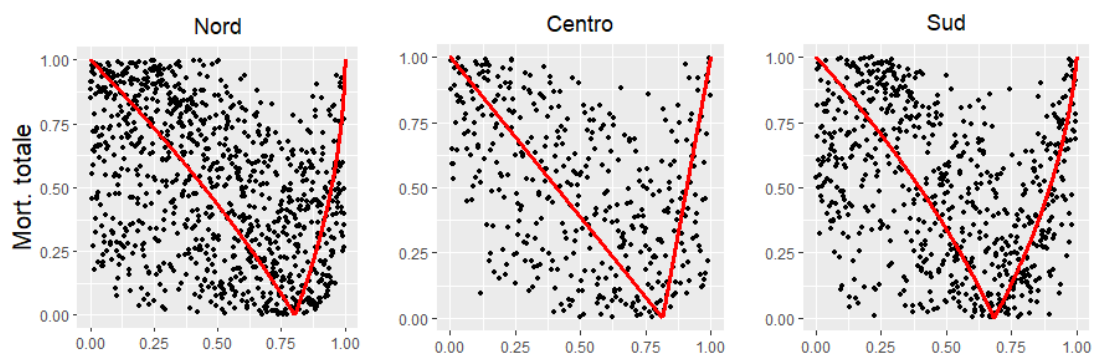




**Figura 7.4.** Pseudo-osservazioni di temperatura media mensile (ascissa) e tasso di mortalità mensile (ordinata) per la popolazione di sesso femminile nelle province del nord, centro e sud Italia.



**Figura 7.5.** Pseudo-osservazioni di temperatura media mensile (ascissa) e tasso di mortalità mensile (ordinata) per la popolazione di sesso maschile nelle province del nord, centro e sud Italia.



**Figura 7.6.** Pseudo-osservazioni di temperatura media mensile (ascissa) e tasso di mortalità mensile (ordinata) per la popolazione totale nelle province del nord, centro e sud Italia.

### Quantili di minima mortalità

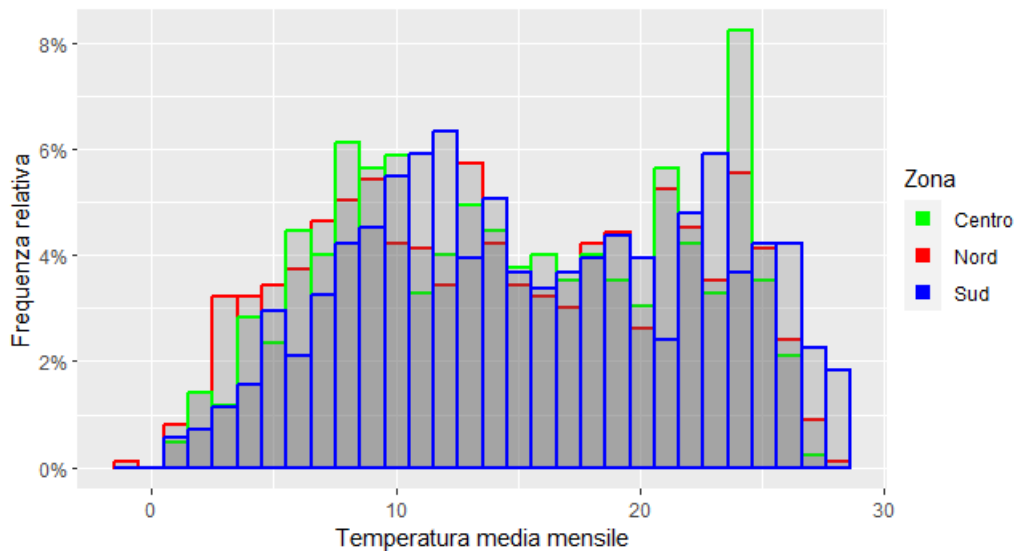
	Femmine	Maschi	Totale
Nord	81.25%	78.86%	80.21%
Centro	78.23%	84.17%	81.49%
Sud	68.09%	69.84%	68.21%

**Tabella 7.7.** Nella tabella sono riportati i livelli dei quantili di minima mortalità per ciascuno dei modelli stimati.

### Temperature di minima mortalità

	Femmine	Maschi	Totale
Nord	22.17	21.56	21.97
Centro	21.57	23.25	22.18
Sud	19.95	20.32	19.96

**Tabella 7.8.** Temperature, in gradi centigradi, di minima mortalità corrispondenti ai quantili della tabella .



**Figura 7.7.** Distribuzione delle temperature medie mensili delle province del nord, centro e sud Italia.

## Capitolo 8

# Simulazione Monte Carlo da copule

Il presente capitolo è dedicato alla descrizione dei principali algoritmi per la simulazione di copule. Dopo aver illustrato alcuni metodi generali, tratteremo alcuni algoritmi specifici per la simulazione di copule Archimedee ed ellittiche. Il capitolo sarà concluso con un paragrafo dedicato ai metodi simulativi, e ai relativi limiti, nel caso di dipendenze non monotone.

### 8.1 Metodi di carattere generale

Il primo metodo di simulazione che descriviamo è un metodo generale che risulta essere di facile implementazione nel caso di copule bivariate. Il metodo è basato essenzialmente sul fatto che le derivate parziali di una copula sono interpretabili, da un punto di vista probabilistico, come distribuzioni condizionate (si veda, ad esempio, il capitolo 1 di Mai e Scherer[48]). Consideriamo una copula bivariata  $C$  e assumiamo che questa abbia una densità di probabilità. Siano inoltre  $U$  e  $V$  due variabili aleatorie uniformemente distribuite su  $[0, 1]$  e aventi distribuzione congiunta  $C$ . Si verifica facilmente che

$$\begin{aligned} \frac{\partial}{\partial u} C(u, v) &= \lim_{h \rightarrow 0} \frac{C(u+h, v) - C(u, v)}{h} = \lim_{h \rightarrow 0} \frac{\mathbb{P}(u < U \leq u+h, V \leq v)}{h} \\ &= \mathbb{P}(V \leq v \mid U = u) \end{aligned} \quad (8.1)$$

e, analogamente,

$$\frac{\partial}{\partial v} C(u, v) = \mathbb{P}(U \leq u \mid V = v). \quad (8.2)$$

Per simulare la coppia  $(U, V)$  è quindi sufficiente simulare la variabile  $U$  per poi simulare  $V$  mediante inversione della funzione di ripartizione condizionata (8.1), o viceversa simulare  $V$  per poi utilizzare l'inversa della funzione (8.2) rispetto a  $u$ . In entrambi i casi, la simulazione di un campione bivariato viene ridotta alla simulazione di due campioni univariati.

Naturalmente, anche nel caso in cui la copula  $C$  ammetta le derivate parziali, non è detto in generale che le distribuzioni condizionate (8.1) e (8.2) siano invertibili in forma chiusa. È il caso, per esempio, della copula Gaussiana. Proprio la copula

Gaussiana costituisce però un caso in cui è possibile applicare un algoritmo di simulazione alternativo. Come discusso nei capitoli precedenti, la copula Gaussiana emerge in modo naturale nello studio della distribuzione normale multivariata, per cui sono disponibili diversi algoritmi simulativi. Nel caso della copula Gaussiana, la scelta più semplice è quindi quella di simulare i dati direttamente dalla distribuzione normale, riportando poi le marginali a uniformi mediante l'utilizzo della trasformazione integrale di probabilità. Il campione simulato risultante avrà, ovviamente, marginali uniformi e copula Gaussiana. Il metodo appena descritto è applicabile in tutti i casi in cui la funzione di ripartizione condizionata non è facilmente invertibile ma è noto un metodo per simulare da una distribuzione multivariata la cui dipendenza è descritta dalla copula cercata. L'effetto delle marginali sulla forma distribuzionale può essere eliminato semplicemente mediante l'utilizzo della trasformazione integrale di probabilità. Con questo metodo abbiamo effettuato, per esempio, le simulazioni dalla copula (4.13) mostrate nella figura 4.7 del paragrafo 4.3. Lo stesso metodo può essere utilizzato per simulare dati provenienti da una copula  $t$ . Infatti, la simulazione di un campione proveniente da una  $t$  di Student multivariata può essere effettuata grazie al fatto che, se  $\mathbf{X}$  è una variabile  $t$  di Student  $d$ -variata con densità (4.10), allora si ha che

$$\mathbf{X} \stackrel{i.d.}{=} \boldsymbol{\mu} + \frac{\sqrt{\nu}}{\sqrt{Y}} \mathbf{Z}$$

dove  $\mathbf{Z} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  e  $Y$  è una variabile aleatoria indipendente da  $\mathbf{Z}$  e avente distribuzione chi-quadro con  $\nu$  gradi di libertà (Lin [46]).

## 8.2 Simulazione da copule archimedee

In questo paragrafo descriviamo un algoritmo di simulazione specifico per alcune copule Archimedee. A questo scopo è opportuno estendere al caso multivariato le definizioni e alcuni risultati su tale famiglia di copule. Una copula Archimedea multivariata può essere definita come segue.

**Definizione 8.2.1.** Una copula  $C : [0, 1]^d \rightarrow [0, 1]$  si dice Archimedea se ammette la rappresentazione

$$C(u_1, u_2, \dots, u_d) = \varphi^{-1} \left( \sum_{k=1}^d \varphi(u_k) \right)$$

dove  $\varphi : [0, 1] \rightarrow [0, \infty)$  è una funzione strettamente decrescente tale che  $\varphi(1) = 0$  e  $\lim_{t \rightarrow 0^+} \varphi(t) = +\infty$ . La funzione  $\varphi$  è detta generatore della copula. ■

Si osservi innanzitutto che la definizione proposta, tratta dall'articolo di McNeil e Nešlehová [51], è più stringente della definizione 4.1.1 che abbiamo utilizzato per le copule bivariate. Infatti, la definizione 8.2.1 ammette come generatori della copula solo le funzioni invertibili. In questo paragrafo adottiamo tale limitazione aggiuntiva sulla scelta del generatore perché è funzionale alla definizione dell'algoritmo di simulazione che sarà di seguito illustrato.

Come visto nel paragrafo 4.1, il teorema 4.1.1 stabilisce che nel caso bivariato la funzione  $\varphi$  debba essere convessa perché  $C$  possa essere una copula. Nel caso

multivariato, McNeil e Nešlehová [51] hanno mostrato un risultato più generale che illustreremo, per semplicità, senza dimostrazione. L'enunciazione di tale risultato richiede una definizione preliminare per la quale introduciamo alcuni elementi notazionali. In particolare, per una generica funzione  $f$ , indicheremo con  $f^{(k)}(x)$  la sua derivata di ordine  $k$  nel punto  $x$ . Inoltre, la parte interna di un dato intervallo  $A$  sarà indicata con il simbolo  $\text{Int}(A)$ .

**Definizione 8.2.2.** Per  $d \geq 2$ , una funzione  $f$  si dice  $d$ -monotona su un intervallo  $A$  se è derivabile  $d - 2$  volte in  $\text{Int}(A)$ , se per ogni  $x \in \text{Int}(A)$  si ha che

$$(-1)^k f^{(k)}(x) \geq 0 \quad \text{per } k = 0, 1, \dots, d - 2$$

e se  $(-1)^{d-2} f^{(d-2)}(x)$  è una funzione non crescente e convessa su  $\text{Int}(A)$ . ■

Enunciamo quindi il risultato di McNeil e Nešlehová che generalizza il teorema 4.1.1 al caso multivariato.

**Teorema 8.2.1.** Sia  $\varphi$  un generatore Archimedeo nel senso della definizione 8.2.1. La funzione  $C$  definita come

$$C(u_1, u_2, \dots, u_d) = \varphi^{-1} \left( \sum_{k=1}^d \varphi(u_k) \right)$$

è una copula se e solo se la funzione  $\varphi^{-1}$  è  $d$ -monotona su  $[0, \infty)$ .

Il teorema 8.2.1 stabilisce le condizioni sotto cui un generatore  $\varphi$  definisce una copula  $d$ -dimensionale per un certo  $d$  fissato. Esiste tuttavia una categoria particolare di generatori che permette di costruire copule Archimedee  $d$ -dimensionali per ogni valore di  $d \geq 2$ . Tale famiglia di copule, studiata da Kimberling [41] in un articolo precedente a quello di McNeil e Nešlehová [51], è strettamente legata al concetto di funzione completamente monotona.

**Definizione 8.2.3.** Una funzione  $f$  si dice completamente monotona su un intervallo  $A$  se in ogni punto  $x \in \text{Int}(A)$  è derivabile infinite volte e si ha che

$$(-1)^k f^{(k)}(x) \geq 0 \quad \forall k \in \mathbb{N}_0.$$

■

Enunciamo di seguito, senza dimostrazione, il risultato di Kimberling [41].

**Teorema 8.2.2.** Sia  $\varphi$  un generatore Archimedeo nel senso della definizione 8.2.1. La funzione  $C$  definita come

$$C(u_1, u_2, \dots, u_d) = \varphi^{-1} \left( \sum_{k=1}^d \varphi(u_k) \right)$$

è una copula per ogni  $d \geq 2$  se e solo se la funzione  $\varphi^{-1}$  è completamente monotona.

I generatori Archimedei la cui inversa è completamente monotona possono quindi essere utilizzati per costruire copule Archimedee di qualunque dimensione. Tuttavia, vi è un'altra proprietà che rende tale famiglia di generatori particolarmente interessante. È infatti ben noto il risultato di Bernstein [8] secondo cui ogni funzione completamente monotona è esprimibile come trasformata di Laplace di un'opportuna misura. In virtù di tale risultato, se  $\varphi^{-1}$  è completamente monotona, allora esiste una funzione di ripartizione  $F$  tale che

$$\varphi^{-1}(s) = \int_0^\infty e^{-sx} dF(x), \quad s \in [0, \infty). \quad (8.3)$$

Per una dimostrazione della formula (8.3) rimandiamo al libro di Schilling et al. [60]. Per i nostri scopi è invece rilevante il seguente risultato (si veda, per esempio, McNeil [50]).

**Teorema 8.2.3.** *Sia  $\varphi$  un generatore Archimedeo avente inversa completamente monotona. Sia  $M$  una variabile aleatoria avente funzione generatrice dei momenti*

$$\mathbb{E} \left[ e^{-sM} \right] = \varphi^{-1}(s)$$

*e siano  $X_1, X_2, \dots, X_d$  delle variabili aleatorie tra loro indipendenti, e tutte indipendenti da  $M$ , aventi distribuzione uniforme sull'intervallo unitario. Allora la variabile  $d$ -dimensionale  $(U_1, U_2, \dots, U_d)$  definita come*

$$(U_1, U_2, \dots, U_d) = \left( \varphi^{-1} \left( \frac{-\log X_1}{M} \right), \varphi^{-1} \left( \frac{-\log X_2}{M} \right), \dots, \varphi^{-1} \left( \frac{-\log X_d}{M} \right) \right)$$

*ha distribuzione*

$$\mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d) = \varphi^{-1} \left( \sum_{k=1}^d \varphi(u_k) \right).$$

*Dimostrazione.* Si verifica che

$$\begin{aligned} & \mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d) \\ &= \mathbb{P} \left( \varphi^{-1} \left( \frac{-\log X_1}{M} \right) \leq u_1, \dots, \varphi^{-1} \left( \frac{-\log X_d}{M} \right) \leq u_d \right) \\ &= \mathbb{P} \left( X_1 \leq e^{-\varphi(u_1)M}, \dots, X_d \leq e^{-\varphi(u_d)M} \right) \\ &= \mathbb{E} \left[ \mathbb{P} \left( X_1 \leq e^{-\varphi(u_1)M}, \dots, X_d \leq e^{-\varphi(u_d)M} \mid M \right) \right] \\ &= \mathbb{E} \left[ \prod_{k=1}^d \mathbb{P} \left( X_k \leq e^{-\varphi(u_k)M} \mid M \right) \right] \\ &= \mathbb{E} \left[ \prod_{k=1}^d e^{-\varphi(u_k)M} \right] = \mathbb{E} \left[ e^{-\sum_{k=1}^d \varphi(u_k)M} \right] \\ &= \varphi^{-1} \left( \sum_{k=1}^d \varphi(u_k) \right). \end{aligned}$$

□

Il teorema 8.2.3 fornisce un algoritmo per la simulazione delle copule Archimedee il cui generatore  $\varphi$  abbia inversa completamente monotona. L'algoritmo è utilizzabile a condizione che si sappia effettuare la simulazione Monte Carlo della variabile  $M$  la cui funzione generatrice dei momenti coincide con la funzione  $\varphi^{-1}$ . Alcuni esempi di copule per cui è possibile utilizzare l'algoritmo appena descritto sono i seguenti (McNeil e Nešlehová [51]):

- la copula Frank, per cui la variabile  $M$  è discreta e ha distribuzione di probabilità

$$\mathbb{P}(M = k) = \frac{(e^{-\theta})^k}{k\theta}, \quad k = 1, 2, \dots$$

dove  $\theta$  rappresenta il parametro della copula;

- la copula Clayton, il cui generatore ha inversa

$$\varphi^{-1}(s) = (1 + s\theta)^{-\frac{1}{\theta}}$$

che coincide con la trasformata di Laplace della densità di una variabile aleatoria gamma;

- la copula Gumbel, il cui generatore ha inversa

$$\varphi^{-1}(s) = e^{-s^{\frac{1}{\theta}}}$$

che corrisponde alla funzione generatrice dei momenti di una variabile aleatoria con distribuzione stabile (si veda Zolotarev [73]).

### 8.3 Simulazione da copule ellittiche

Un algoritmo specifico per la simulazione di alcune copule ellittiche si può ricavare dallo studio di una rappresentazione stocastica per le distribuzioni ellittiche. In questo paragrafo ricaveremo tale rappresentazione stocastica proponendo una dimostrazione ispirata all'articolo di Cambanis et al. [9] e al libro di Johnson [38]. Come nel paragrafo 4.2, consideriamo una variabile aleatoria  $\mathbf{Y}$  con valori in  $\mathbb{R}^d$  avente distribuzione sferica e funzione caratteristica

$$\mathbb{E} \left[ e^{it^T \mathbf{Y}} \right] = \psi \left( \|\mathbf{t}\|^2 \right).$$

Sia ora  $\mathbf{U}_d$  una variabile aleatoria avente distribuzione uniforme sulla superficie della sfera unitaria in  $\mathbb{R}^d$ . È evidente che, con probabilità 1, si avrà la relazione

$$\|\mathbf{t}\|^2 = \left\| \|\mathbf{t}\| \mathbf{U}_d \right\|^2$$

da cui segue che

$$\psi \left( \|\mathbf{t}\|^2 \right) = \psi \left( \left\| \|\mathbf{t}\| \mathbf{U}_d \right\|^2 \right) = \mathbb{E} \left[ \psi \left( \left\| \|\mathbf{t}\| \mathbf{U}_d \right\|^2 \right) \right]. \quad (8.4)$$

Osserviamo ora che formalmente, se  $\mathbf{Y}$  e  $\mathbf{U}_d$  sono indipendenti, vale con probabilità 1 la relazione

$$\psi\left(\|\mathbf{t}\| \|\mathbf{U}_d\|^2\right) = \mathbb{E}\left[e^{i\|\mathbf{t}\| \mathbf{U}_d^T \mathbf{Y}} \mid \mathbf{U}_d\right]$$

da cui segue che la formula (8.4) può essere scritta come

$$\psi\left(\|\mathbf{t}\|^2\right) = \mathbb{E}\left[e^{i\|\mathbf{t}\| \mathbf{U}_d^T \mathbf{Y}}\right] = \mathbb{E}\left[e^{i\|\mathbf{t}\| \mathbf{Y}^T \mathbf{U}_d}\right].$$

È cruciale a questo punto osservare che anche  $\mathbf{U}_d$  ha una distribuzione sferica, per cui la sua funzione caratteristica dipende solo dalla norma dell'argomento. Pertanto possiamo scrivere

$$\begin{aligned} \psi\left(\|\mathbf{t}\|^2\right) &= \mathbb{E}\left[\mathbb{E}\left[e^{i\|\mathbf{t}\| \mathbf{Y}^T \mathbf{U}_d} \mid \mathbf{Y}\right]\right] = \mathbb{E}\left[\mathbb{E}\left[e^{i\|\mathbf{Y}\| \mathbf{t}^T \mathbf{U}_d} \mid \mathbf{Y}\right]\right] \\ &= \mathbb{E}\left[e^{i\|\mathbf{Y}\| \mathbf{t}^T \mathbf{U}_d}\right] \end{aligned}$$

che equivale a scrivere che

$$\mathbb{E}\left[e^{i\mathbf{t}^T \mathbf{Y}}\right] = \mathbb{E}\left[e^{i\mathbf{t}^T \|\mathbf{Y}\| \mathbf{U}_d}\right].$$

Segue immediatamente che

$$\mathbf{Y} \stackrel{i.d.}{=} R \mathbf{U}_d \tag{8.5}$$

dove  $R$  è una variabile aleatoria indipendente da  $\mathbf{U}_d$  e avente distribuzione identica a quella di  $\|\mathbf{Y}\|$ . Ricordiamo ora che una variabile  $\mathbf{X}$  con distribuzione ellittica ammette la rappresentazione

$$\mathbf{X} = \boldsymbol{\mu} + \mathbf{C}\mathbf{Y} \tag{8.6}$$

dove  $\boldsymbol{\mu} \in \mathbb{R}^d$ ,  $\mathbf{Y}$  è un vettore avente distribuzione sferica e  $\mathbf{C}$  è una matrice tale che il prodotto  $\boldsymbol{\Sigma} = \mathbf{C}^T \mathbf{C}$  è una matrice definita positiva. Utilizzando le relazioni (8.5) e (8.6) si ottiene, che per la variabile aleatoria  $\mathbf{X}$  con distribuzione ellittica e funzione caratteristica

$$\mathbb{E}\left[e^{i\mathbf{t}^T (\mathbf{X} - \boldsymbol{\mu})}\right] = \psi\left(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\right)$$

vale la rappresentazione stocastica

$$\mathbf{X} = \boldsymbol{\mu} + R \mathbf{C} \mathbf{U}_d \tag{8.7}$$

dove  $R$  è una variabile aleatoria indipendente da  $\mathbf{U}_d$  e avente distribuzione identica a quella di  $\sqrt{(\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu})}$ .

Come osservato da Johnson [38], la rappresentazione stocastica (8.7) permette, nel caso in cui siano noti i momenti della variabile  $R$ , di studiare in maniera immediata l'esistenza dei momenti della distribuzione ellittica. Inoltre, la stessa rappresentazione stocastica (8.7) permette di effettuare simulazioni Monte Carlo dalla distribuzione di  $\mathbf{X}$  sotto la condizione che si sia in grado di simulare dalle distribuzioni di  $R$  e  $\mathbf{U}_d$ . Per una dettagliata trattazione degli algoritmi per la simulazione di  $\mathbf{U}_d$  rimandiamo alla lettura del capitolo 7 di Johnson [38]. Un algoritmo particolarmente pratico, per la sua efficienza e per il fatto che è valido in ogni dimensione, è quello proposto da



Muller [52]. L'autore ha dimostrato che se  $\mathbf{Z} = (Z_1, \dots, Z_d)^T$  è un vettore di variabili aleatorie normali standard tra loro indipendenti, allora si ha che

$$\mathbf{U}_d \stackrel{i.d.}{=} \frac{\mathbf{Z}}{\|\mathbf{Z}\|}.$$

Per quanto riguarda la simulazione della variabile  $R$ , è di aiuto il seguente risultato.

**Teorema 8.3.1.** *Sia  $\mathbf{X}$  una variabile aleatoria  $d$ -variata con distribuzione ellittica e densità di probabilità*

$$\frac{1}{\sqrt{|\boldsymbol{\Sigma}|}} g\left((\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right).$$

Allora la variabile  $Z = (\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu})$  ha densità di probabilità

$$h(z) = \frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)} z^{\frac{d}{2}-1} g(z), \quad z > 0.$$

*Dimostrazione.* Osserviamo innanzitutto che

$$\begin{aligned} \mathbb{P}(Z \in A) &= \int_A \mathbb{P}\left((\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \in d\tau\right) \\ &= \int_{\{\mathbf{y}: \mathbf{y}^T \mathbf{y} \in A\}} g(\mathbf{y}^T \mathbf{y}) dy_1 \dots dy_n. \end{aligned}$$

Passando alle coordinate ipersferiche

$$\begin{cases} y_1 = r \cos(\theta_1) \\ y_2 = r \sin(\theta_1) \cos(\theta_2) \\ y_3 = r \sin(\theta_1) \sin(\theta_2) \cos(\theta_3) \\ \vdots \\ y_{d-1} = r \sin(\theta_1) \sin(\theta_2) \dots \cos(\theta_{d-1}) \\ y_d = r \sin(\theta_1) \sin(\theta_2) \dots \sin(\theta_{d-1}) \end{cases}$$

dove  $\theta_1, \dots, \theta_{d-2} \in [0, \pi]$  e  $\theta_{d-1} \in [0, 2\pi]$ , e tenendo conto che il valore assoluto del determinante dello Jacobiano è

$$|J_d| = r^{d-1} \sin^{d-2}(\theta_1) \sin^{d-3}(\theta_2) \dots \sin(\theta_{d-2})$$

possiamo scrivere

$$\begin{aligned} \mathbb{P}(Z \in A) &= \int_{\{r: r^2 \in A\}} r^{d-1} g(r^2) dr \\ &\quad \cdot \int_0^{2\pi} \int_0^\pi \dots \int_0^\pi \sin^{d-2}(\theta_1) \sin^{d-3}(\theta_2) \dots \sin(\theta_{d-2}) d\theta_1 \dots d\theta_{d-2} d\theta_1 \\ &= S_d(1) \int_{\{r: r^2 \in A\}} r^{d-1} g(r^2) dr \end{aligned}$$

dove  $S_d(1)$  rappresenta la misura  $(d-1)$ -dimensionale della superficie di un'ipersfera  $d$ -dimensionale di raggio 1. Tenendo ora presente che

$$S_d(1) = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}$$

ed effettuando il cambiamento di variabili  $r = \sqrt{z}$  si ottiene

$$\mathbb{P}(Z \in A) = \frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)} \int_A z^{\frac{d}{2}-1} g(z) dz.$$

che conclude la dimostrazione.  $\square$

Il teorema 8.3.1 fornisce la distribuzione di probabilità di  $R^2$ , dove  $R$  è definita come nella rappresentazione stocastica (8.7). È noto che, nel caso in cui  $\mathbf{X}$  abbia distribuzione normale  $d$ -variata, la variabile  $R^2$  ha distribuzione chi-quadro con  $d$  gradi di libertà. Altri esempi di distribuzioni ellittiche per cui è possibile ricavare la distribuzione di  $R^2$  sono reperibili nel capitolo 6 di Johnson [38]. A scopo esemplificativo consideriamo qui il caso della distribuzione  $t$  di Student multivariata. In virtù del teorema 8.3.1 e tenendo in considerazione la forma della densità della  $t$  di Student multivariata (4.10), si verifica facilmente che la variabile  $R^2$  ha densità di probabilità

$$h(z; \nu) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\nu^{\frac{d}{2}} \Gamma\left(\frac{\nu}{2}\right) \Gamma\left(\frac{d}{2}\right)} z^{\frac{d}{2}-1} \left(1 + \frac{z}{\nu}\right)^{-\frac{\nu+d}{2}}, \quad z > 0 \quad (8.8)$$

e la corrispondente funzione di ripartizione è

$$\begin{aligned} H(z; \nu) &= \frac{1}{\nu^{\frac{d}{2}} B\left(\frac{\nu}{2}, \frac{d}{2}\right)} \int_0^z s^{\frac{d}{2}-1} \left(\frac{\nu}{s+\nu}\right)^{\frac{\nu+d}{2}} ds \\ &= \frac{1}{B\left(\frac{\nu}{2}, \frac{d}{2}\right)} \int_{\frac{\nu}{\nu+z}}^1 (1-y)^{\frac{d}{2}-1} y^{\frac{\nu}{2}-1} dy \\ &= \frac{B\left(\frac{\nu}{2}, \frac{d}{2}\right) - B\left(\frac{\nu}{\nu+z}; \frac{\nu}{2}, \frac{d}{2}\right)}{B\left(\frac{\nu}{2}, \frac{d}{2}\right)} \end{aligned} \quad (8.9)$$

dove abbiamo effettuato la sostituzione  $y = \frac{\nu}{s+\nu}$  e abbiamo indicato con  $B$  la funzione beta. Nel caso della  $t$  di Student è quindi possibile simulare la variabile  $R^2$  utilizzando la densità di probabilità (8.8) congiuntamente a un algoritmo di tipo Accetta-Rifiuta (si veda, per esempio, Glasserman [29], pag. 58) oppure invertendo numericamente la funzione di ripartizione (8.8). Un'implementazione dell'inversa della funzione beta incompleta è fornita in R dal pacchetto `zipfR` (Evert e Baroni [20]). Si osservi che nel caso bivariato la densità (8.8) si riduce a quella di una Pareto generalizzata che semplifica notevolmente la simulazione Monte Carlo.

## 8.4 Simulazione da copule non monotone

Concludiamo il capitolo esaminando alcuni aspetti legati agli algoritmi di simulazione nel caso di copule non monotone. Essendo queste ultime delle copule a tutti gli effetti, valgono naturalmente, ove applicabili, gli algoritmi generali descritti nel paragrafo 8.1. Si considerino una copula  $C$  e una funzione  $f$  che soddisfi le ipotesi del teorema 5.1.1. Definiamo quindi la copula

$$Q(u, v) = C(u, f(v)) - C(u, f(v) - v) \quad (8.10)$$

e osserviamo che le sue derivate parziali sono

$$\frac{\partial}{\partial u} Q(u, v) = \frac{\partial C}{\partial u}(u, f(v)) - \frac{\partial C}{\partial u}(u, f(v) - v) \quad (8.11)$$

e

$$\frac{\partial}{\partial v} Q(u, v) = \frac{\partial C}{\partial v}(u, f(v)) \cdot f'(v) - \frac{\partial C}{\partial v}(u, f(v) - v) \cdot [f'(v) - 1]. \quad (8.12)$$

Come già anticipato nel paragrafo 8.1, il principale difetto del metodo simulativo basato sulle distribuzioni condizionate è che questo richiede il calcolo dell'inversa di una tra le funzioni (8.11) o (8.12). I casi in cui è possibile esprimere l'inversa in forma chiusa sono tuttavia rari. Tra questi rientra, ad esempio, il caso in cui  $C$  è la copula Frank. Nella maggior parte dei casi l'inversione delle formule (8.11) o (8.12) può essere effettuata solo per via numerica, rendendo quindi la simulazione Monte Carlo notevolmente più pesante dal punto di vista computazionale.

Un altro possibile metodo di simulazione proviene invece dall'interpretazione della copula  $Q$  fornita dalla formula (5.8). Infatti, la copula  $Q$  corrisponde alla distribuzione della coppia

$$(U, g(V))$$

dove la coppia  $(U, V)$  è distribuita secondo la copula  $C$  e  $g$  è definita come nella formula (5.5), ovvero

$$g(x) = \begin{cases} g^+(x) & x > f(0) \\ g^-(x) & x \leq f(0) \end{cases}$$

avendo indicato con  $g^+(x)$  e  $g^-(x)$  rispettivamente le pseudo-inverse di  $f(v)$  e  $f(v) - v$ . Se è quindi disponibile un algoritmo per simulare dati dalla copula  $C$ , è sufficiente simulare la coppia  $(U, V)$  e poi trasformare la variabile  $V$  mediante la funzione  $g$ . Il metodo in questione è stato impiegato per effettuare le simulazioni mostrate nelle figure 5.1, 5.2 e 5.3. Facciamo presente tuttavia che anche il calcolo della funzione  $g$  non è sempre banale. Infatti, anche nel caso in cui sia possibile trovare in forma esplicita l'inversa di  $f(v)$ , non è necessariamente detto che lo stesso sia vero anche per la funzione  $f(v) - v$ . Prendiamo ad esempio le tre funzioni  $f_1$ ,  $f_2$  ed  $f_3$  proposte nel paragrafo 6.1 per la descrizione della relazione rendimento-volume nella serie storica del fondo QQQ. In particolare, si verifica immediatamente che, mentre la funzione  $f_3(v; a, c)$  è invertibile in forma chiusa, lo stesso non si può dire della funzione  $f_3(v; a, c) - v$  che richiede quindi che la sua inversa  $g_3^-(x)$  sia calcolata con metodi numerici. Si osservi che il calcolo della funzione  $g_3(x)$  che preserva la

distribuzione uniforme associata a  $f_3(v; a, c)$ , definita come

$$g_3(x) = \begin{cases} g_3^+(x) & x > f_3(0; a, c) \\ g_3^-(x) & x \leq f_3(0; a, c) \end{cases}$$

con l'usuale significato dei simboli, deve essere calcolata numericamente solo quando l'argomento è più piccolo di  $f_3(0; a, c)$ . Dato un campione di  $n$  osservazioni  $v_1, v_2, \dots, v_n$  provenienti da una distribuzione uniforme, la funzione  $g_3(x)$  dovrà quindi essere calcolata numericamente per un numero di osservazioni pari circa a  $n \cdot f_3(0; a, c)$  mentre potrà essere calcolata in forma chiusa per le restanti osservazioni.

## Conclusioni

I risultati ottenuti in questa tesi forniscono un metodo generale per la costruzione di copule che descrivono dipendenze di forma arbitraria. Come illustrato da Scarsini e Venetoulas [58], la definizione di copule di dipendenza non monotona è strettamente legata alla costruzione di funzioni che possiedono la proprietà di preservare la distribuzione uniforme. I risultati qui presentati permettono di costruire in forma parametrica funzioni che godono di tale proprietà non limitando però la scelta, a differenza della letteratura attualmente esistente sull'argomento, alle sole funzioni lineari a tratti. Ciò conferisce ai modelli di copula risultanti una flessibilità che li rende adatti a descrivere fenomeni reali che presentano strutture di dipendenza complesse. La semplicità di implementazione e la flessibilità della metodologia proposta sono state illustrate prendendo in considerazione due fenomeni reali in cui emerge una dipendenza di tipo non monotono. Il primo fenomeno esaminato è la relazione tra rendimento e volume degli scambi nei mercati finanziari, studiato per la prima volta mediante l'uso di copule di dipendenza non monotona da Neto [55]. Rispetto alla copula originariamente utilizzata da Neto, le copule qui definite sono ottenute a partire da funzioni preservanti la distribuzione uniforme più generali, dipendenti da parametri che regolano la concavità o la convessità dei rami della funzione stessa. I criteri di selezione del modello qui adottati giustificano l'introduzione di tali parametri aggiuntivi, supportando quindi la conclusione che l'utilizzo di modelli parametrici più flessibili sia utile ai fini di una descrizione più accurata dei dati analizzati.

Abbiamo in seguito applicato gli stessi modelli di copula allo studio della relazione tra temperatura e mortalità, proponendo per la prima volta l'uso di copule di dipendenza non monotona per l'analisi del fenomeno. Anche in questo caso, i criteri impiegati per la selezione del modello confermano l'adeguatezza dei modelli parametrici proposti e la loro capacità di adattarsi ai dati considerati. Si è osservato come i modelli adottati conducono a conclusioni coerenti con la letteratura relativamente al maggior impatto sulla mortalità delle basse temperature rispetto alle ondate di calore. Infine, abbiamo illustrato come i modelli di copula possono essere utilizzati per la stima del quantile della temperatura corrispondente al minimo livello di mortalità. Si è tuttavia osservato che le sole differenze climatiche tra le aree geografiche non sembrano essere sufficienti a spiegare le temperature di minima mortalità, per lo studio delle quali è verosimilmente più adeguato un approccio di tipo multivariato.

# Bibliografia

- [1] Akaike, H. “Information Theory and an Extension of the Maximum Likelihood Principle”. In: *Petrov, B.N. and Csaki, F., Eds., International Symposium on Information Theory*. 1973, pp. 267–281.
- [2] Analitis, A. et al. “Effects of cold weather on mortality: results from 15 European cities within the PHEWE project.” In: *American journal of epidemiology* 168 12 (2008), pp. 1397–408.
- [3] Anderson, T. W. *An Introduction to Multivariate Statistical Analysis*. Third. Stanford: John Wiley & Sons, 2003.
- [4] Athreya, K. B. e Lahiri, S. N. *Measure theory and probability theory*. Vol. 19. Springer, 2006.
- [5] Barnett, A. G. et al. “Cold periods and coronary events: an analysis of populations worldwide”. In: *Journal of Epidemiology & Community Health* 59.7 (2005), pp. 551–557.
- [6] Basu, R. e Samet, J. M. “Relation between Elevated Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence”. In: *Epidemiologic Reviews* 24.2 (dic. 2002), pp. 190–202.
- [7] Bernard, C. e Czado, C. “Conditional quantiles and tail dependence”. In: *Journal of Multivariate Analysis* 138 (2015), pp. 104–126.
- [8] Bernstein, S. “Sur les fonctions absolument monotones”. In: *Acta Mathematica* 52.none (1929), pp. 1–66.
- [9] Cambanis, S., Huang, S. e Simons, G. “On the theory of elliptically contoured distributions”. In: *Journal of Multivariate Analysis* 11.3 (1981), pp. 368–385.
- [10] Chen, R. et al. “Association between ambient temperature and mortality risk and burden: time series study in 272 main Chinese cities”. In: *BMJ* 363 (2018).
- [11] Clark, P. K. “A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices”. In: *Econometrica* 41.1 (1973), pp. 135–155.
- [12] Clarke, B. et al. “Extreme weather impacts of climate change: an attribution perspective”. In: *Environmental Research: Climate* 1.1 (2022), p. 012001.
- [13] Crouch, R. L. “A Nonlinear Test of the Random-Walk Hypothesis”. In: *The American Economic Review* 60.1 (1970), pp. 199–202.
- [14] Daniels, H. E. “Rank Correlation and Population Models”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 12.2 (1950), pp. 171–191.

- [15] Donoghue, E. et al. “Criteria for the diagnosis of heat-related deaths: National Association of Medical Examiners. Position paper. National Association of Medical Examiners Ad Hoc Committee on the Definition of Heat-Related Fatalities”. In: *The American journal of forensic medicine and pathology* 18.1 (1997), 11–14.
- [16] Dunn, R. J. H. et al. “Development of an Updated Global Land In Situ-Based Data Set of Temperature and Precipitation Extremes: HadEX3”. In: *Journal of Geophysical Research: Atmospheres* 125.16 (2020), e2019JD032263.
- [17] Durbin, J. e Stuart, A. “Inversions and Rank Correlation Coefficients”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 13.2 (1951), pp. 303–309. (Visitato il 10/08/2023).
- [18] Embrechts, P. “Correlation: pitfalls and alternatives”. In: *Risk Magazine* (1999), pp. 69–71.
- [19] Epps, T. W. “Security Price Changes and Transaction Volumes: Theory and Evidence”. In: *The American Economic Review* 65.4 (1975), pp. 586–597.
- [20] Evert, S. e Baroni, M. “*zipfR*: Word Frequency Distributions in R”. In: *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, Posters and Demonstrations Sessions*. (R package version 0.6-70 of 2020-10-10). Prague, 2007, pp. 29–32.
- [21] Frees, E. W. e Valdez, E. A. “Understanding Relationships Using Copulas”. In: *North American Actuarial Journal* 2.1 (1998), pp. 1–25.
- [22] Fréchet, M. “Sur les tableaux dont les marges et des bornes sont données”. In: *Revue de l’Institut International de Statistique / Review of the International Statistical Institute* 28.1/2 (1960), pp. 10–32.
- [23] Fu, S. H. et al. “Mortality attributable to hot and cold ambient temperatures in India: a nationally representative case-crossover study”. In: *PLoS medicine* 15.7 (2018), e1002619.
- [24] Gasparrini, A., Armstrong, B. e Kenward, M. G. “Distributed lag non-linear models”. In: *Statistics in Medicine* 29.21 (2010), pp. 2224–2234.
- [25] Gasparrini, A. et al. “Mortality risk attributable to high and low ambient temperature: a multicountry observational study”. In: *The Lancet* 386.9991 (2015), pp. 369–375.
- [26] Gasparrini, A. et al. “Projections of temperature-related excess mortality under climate change scenarios”. In: *The Lancet Planetary Health* 1.9 (2017), e360–e367.
- [27] Gasparrini, A. et al. “Small-area assessment of temperature-related mortality risks in England and Wales: a case time series analysis”. In: *The Lancet Planetary Health* 6.7 (2022), e557–e564.
- [28] Genest, C., Gendron, M. e Bourdeau-Brien, M. “The Advent of Copulas in Finance”. In: *The European Journal of Finance* 15.7-8 (2009), pp. 609–618.
- [29] Glasserman, P. *Monte Carlo Methods in Financial Engineering*. New York: Springer, 2004.

- [30] Gündüz, L. e Hatemi-J, A. “Stock Price and Volume Relation in Emerging Markets”. In: *Emerging Markets Finance and Trade* 41.1 (2005), pp. 29–44.
- [31] Heutel, G., Miller, N. H. e Molitor, D. “Adaptation and the Mortality Effects of Temperature across U.S. Climate Regions”. In: *The Review of Economics and Statistics* 103.4 (set. 2021), pp. 740–753.
- [32] Hoeffding, W. “Scale-Invariant Correlation Measures for Discontinuous Distributions”. In: *The Collected Works of Wassily Hoeffding*. A cura di N. I. Fisher e P. K. Sen. New York, NY: Springer New York, 1994, pp. 109–133.
- [33] Hoeffding, W. “Scale-Invariant Correlation Theory”. In: *The Collected Works of Wassily Hoeffding*. A cura di N. I. Fisher e P. K. Sen. New York, NY: Springer New York, 1994, pp. 57–107.
- [34] Honda, Y. et al. “Determination of optimum daily maximum temperature using climate data”. In: *Environmental health and preventive medicine* 12 (2007), pp. 209–216.
- [35] Imada, Y. et al. “Climate Change Increased the Likelihood of the 2016 Heat Extremes in Asia”. In: *Bulletin of the American Meteorological Society* 99.1 (2018), S97–S101.
- [36] Joe, H. *Dependence modeling with copulas*. Vancouver: CRC Press, 2014.
- [37] Joe, H. *Multivariate models and dependence concepts*. London: Chapman e Hall, 1997.
- [38] Johnson, M. E. *Multivariate statistical simulation*. New York: John Wiley e Sons, 1987.
- [39] Karpoff, J. M. “The Relation Between Price Changes and Trading Volume: A Survey”. In: *The Journal of Financial and Quantitative Analysis* 22.1 (1987), pp. 109–126.
- [40] Kendall, M. G. “A New Measure of Rank Correlation”. In: *Biometrika* 30.1/2 (1938), pp. 81–93.
- [41] Kimberling, C. “A probabilistic interpretation of complete monotonicity”. In: *aequationes mathematicae* 10 (1974), pp. 152–164.
- [42] Kovats, R. S., Hajat, S e Wilkinson, P. “Contrasting patterns of mortality and hospital admissions during hot weather and heat waves in Greater London, UK”. In: *Occupational and Environmental Medicine* 61.11 (2004), pp. 893–898.
- [43] Krummenauer, L. et al. “Global drivers of minimum mortality temperatures in cities”. In: *Science of The Total Environment* 695 (2019), p. 133560.
- [44] Kruskal, W. H. “Ordinal Measures of Association”. In: *Journal of the American Statistical Association* 53.284 (1958), pp. 814–861.
- [45] Lee, B.-S. e Rui, O. M. “The dynamic relationship between stock returns and trading volume: Domestic and cross-country evidence”. In: *Journal of Banking & Finance* 26.1 (2002), pp. 51–78.
- [46] Lin, P.-E. “Some characterizations of the multivariate t distribution”. In: *Journal of Multivariate Analysis* 2.3 (1972), pp. 339–344.



- [47] Linares, C. e Díaz, J. “Impact of high temperatures on hospital admissions: comparative analysis with previous studies about mortality (Madrid)”. In: *European Journal of Public Health* 18.3 (nov. 2007), pp. 317–322.
- [48] Mai, J.-F. e Scherer, M. *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*. Imperial College Press, 2012.
- [49] Masselot, P. et al. “Excess mortality attributed to heat and cold: a health impact assessment study in 854 cities in Europe”. In: *The Lancet Planetary Health* 7.4 (2023), e271–e281.
- [50] McNeil, A. J. “Sampling nested Archimedean copulas”. In: *Journal of Statistical Computation and Simulation* 78.6 (2008), pp. 567–581.
- [51] McNeil, A. J. e Nešlehová, J. “Multivariate Archimedean copulas, d-monotone functions and  $\ell_1$ -norm symmetric distributions”. In: *The Annals of Statistics* 37.5B (2009), pp. 3059–3097.
- [52] Muller, M. E. “A Note on a Method for Generating Points Uniformly on N-Dimensional Spheres”. In: *Commun. ACM* 2.4 (1959), 19–20.
- [53] Nelsen, R. B. *An introduction to copulas*. New York: Springer, 1999.
- [54] Nelsen, R. B. *An introduction to copulas*. Second edition. New York: Springer, 2006.
- [55] Neto, D. “Dépendance non-monotone: Une application à la relation rendement-volume”. In: *Annals of Economics and Statistics* 82 (2006), pp. 187–216.
- [56] Osborne, M. F. M. “Brownian Motion in the Stock Market”. In: *Operations Research* 7.2 (1959), pp. 145–173.
- [57] Rudin, W. *Principles of mathematical analysis*. Third edition. New York: McGraw-hill, 1976.
- [58] Scarsini, M. e Venetoulas, A. “Bivariate Distributions With Nonmonotone Dependence Structure”. In: *Journal of the American Statistical Association* 88.421 (1993), pp. 338–344.
- [59] Schepsmeier, U. et al. “Package ‘vinecopula’”. In: *R package version 2.5* (2015).
- [60] Schilling, R. L., Song, R. e Vondracek, Z. *Bernstein Functions*. Berlin, Boston: De Gruyter, 2012.
- [61] Schwarz, G. “Estimating the Dimension of a Model”. In: *The Annals of Statistics* 6.2 (1978), pp. 461–464.
- [62] Semenza, J. C. et al. “Heat-Related Deaths during the July 1995 Heat Wave in Chicago”. In: *New England Journal of Medicine* 335.2 (1996), pp. 84–90.
- [63] Shen, T et al. “Toward a broader definition of heat-related death: comparison of mortality estimates from medical examiners’ classification with those from total death differentials during the July 1995 heat wave in Chicago, Illinois”. In: *The American journal of forensic medicine and pathology* 19.2 (1998), 113–118.
- [64] Sklar, A. “Fonctions de Répartition à n Dimensions et Leurs Marges”. In: *Publications de l’Institut Statistique de l’Université de Paris* 8 (1959), pp. 229–231.

- [65] Spearman, C. “The Proof and Measurement of Association between Two Things”. In: *The American Journal of Psychology* 100.3/4 (1987), pp. 441–471.
- [66] Tauchen, G. E. e Pitts, M. “The Price Variability-Volume Relationship on Speculative Markets”. In: *Econometrica* 51.2 (1983), pp. 485–505.
- [67] Tawn, J. A. “Bivariate Extreme Value Theory: Models and Estimation”. In: *Biometrika* 75.3 (1988), pp. 397–415.
- [68] Tertre, A. L. et al. “Impact of the 2003 Heatwave on All-Cause Mortality in 9 French Cities”. In: *Epidemiology* 17.1 (2006), pp. 75–79.
- [69] Tobías, A. et al. “Geographical variations of the minimum mortality temperature at a global scale: a multicountry study”. In: *Environmental epidemiology* 5.5 (2021).
- [70] Ye, X. et al. “Ambient temperature and morbidity: a review of epidemiological evidence.” In: *Environmental Health Perspectives* 120.1 (2012), pp. 19–28.
- [71] Yin, Q. et al. “Mapping the increased minimum mortality temperatures in the context of global climate change”. In: *Nature communications* 10.1 (2019), p. 4640.
- [72] Zhao, Q. et al. “Global, regional, and national burden of mortality associated with non-optimal ambient temperatures from 2000 to 2019: a three-stage modelling study”. In: *Lancet Planetary Health* 5.7 (2021), e415–e425.
- [73] Zolotarev, V. M. *One-dimensional stable distributions*. Providence: American Mathematical Society, 1986.

## Appendice A

# Codici R: stima dei parametri

```

# IMPORTAZIONE E PREPARAZIONE DATI
Data <- read.csv("QQQ.csv")

Close=Data$Close
n=length(Close)
Ret=(Close[2:n]-Close[1:(n-1)])/Close[1:(n-1)]
Vol=Data$Volume[2:n]

library(copula)
u=pobs(Ret)
v=pobs(Vol)
n=length(u)

hist(u)
hist(v)
plot(u,v,pch=19,cex=0.5)

# DEFINIZIONE DELLE DENSITA' DI COPULA
Frank_Density=function(u,v,theta){
  if(abs(theta)<1e-12){out=rep(1,length(u))}else{
    out=-theta*exp(-theta*(u+v))*(exp(-theta)-1)/
      (exp(-theta)-exp(-theta*u)-
        exp(-theta*v)+exp(-theta*(u+v)))^2}
  return(out)}

Clayton_Density=function(u,v,theta){
  if(theta<1e-12){out=rep(1,length(u))}else{
    out=(1+theta)*(u*v)^(-theta-1)*(u^(-theta)+v^(-theta)-1)^
      (-(2*theta+1)/theta)}
  return(out)}

Gumbel_Density=function(u,v,theta){
  temp=(-log(u))^theta+(-log(v))^theta
  Cop=exp(-temp^(1/theta))
  out=Cop*(log(u)*log(v))^(theta-1)/(u*v)*
    (temp^(2/theta-2)-(1-theta)*temp^(1/theta-2))
}

```



```

colnames(Params_Quadratic_Output)=c("a","c","theta")

ML_Power_Output=data.frame(matrix(NA, nrow = 0, ncol = 3))
colnames(ML_Power_Output)=c("ML","AIC","BIC")

Params_Power_Output=data.frame(matrix(NA, nrow = 0, ncol = 3))
colnames(Params_Power_Output)=c("a","c","theta")

# FUNZIONE AUSILIARIA PER DISEGNARE GRAFICI:
# (INVERSIONE NUMERICA DI f_pow(v)-v MEDIANTE L'ALGORITMO DI
# HALLEY)
aux_inv=function(x,a,c){
  x0=rep(0.5,length(x))
  error=max(abs(f_pow(x0,a,c)-x0-x))
  while(error>1e-12){
    x_next=x0-2*(f_pow(x0,a,c)-x0-x)*(df_pow(x0,a,c)-1)/
      (2*(df_pow(x0,a,c)-1)^2-
        (f_pow(x0,a,c)-x0-x)*ddf_pow(x0,a,c))
    error=max(abs(f_pow(x_next,a,c)-x_next-x))
    x0=x_next}
  return(x0)}

#####
# CASO 1: f LINEARE
#####

# COPULA FRANK
L=function(c,theta){
  -log(Frank_Density(u,f_lin(v,c),theta)/(c+1)+
    c*Frank_Density(u,f_lin(v,c)-v,theta)/(c+1))}
# Stima dei parametri
fobj=function(vec){sum(L(exp(vec[1]),vec[2]))}
out=optim(par=c(0,1), fn=fobj,method="BFGS")
out$convergence
params=c(exp(out$par[1]),out$par[2])
ML=-out$value
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
c=params[1]
theta=params[2]
plot(u,v,pch=19,cex=0.5)
lines(c(c/(1+c),1),c(0,1),col="red",lwd=3)
lines(c(0,c/(1+c)),c(1,0),col="red",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Linear_Output=rbind(ML_Linear_Output,dframe)
row.names(ML_Linear_Output)[1]="Frank"
dframe <- data.frame(c, theta)
Params_Linear_Output=rbind(Params_Linear_Output,dframe)
row.names(Params_Linear_Output)[1]="Frank"

```

```

# COPULA CLAYTON
L=function(c,theta){
  -log(Clayton_Density(u,f_lin(v,c),theta)/(c+1)+
        c*Clayton_Density(u,f_lin(v,c)-v,theta)/(c+1))}
# Stima dei parametri
fobj=function(vec){sum(L(exp(vec[1]),exp(vec[2])))}
out=optim(par=c(0,0), fn=fobj)
out$convergence
params=c(exp(out$par[1]),exp(out$par[2]))
ML=-out$value
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
c=params[1]
theta=params[2]
plot(u,v,pch=19,cex=0.5)
lines(c(c/(1+c),1),c(0,1),col="red",lwd=3)
lines(c(0,c/(1+c)),c(1,0),col="red",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Linear_Output=rbind(ML_Linear_Output, dframe)
row.names(ML_Linear_Output)[2]="Clayton"
dframe <- data.frame(c, theta)
Params_Linear_Output=rbind(Params_Linear_Output, dframe)
row.names(Params_Linear_Output)[2]="Clayton"

# COPULA GUMBEL
L=function(c,theta){
  -log(Gumbel_Density(u,f_lin(v,c),theta)/(c+1)+
        c*Gumbel_Density(u,f_lin(v,c)-v,theta)/(c+1))}
# Stima dei parametri
fobj=function(vec){sum(L(exp(vec[1]),1+exp(vec[2])))}
out=optim(par=c(1,0), fn=fobj,method="BFGS")
out$convergence
params=c(exp(out$par[1]),1+exp(out$par[2]))
ML=-out$value
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
c=params[1]
theta=params[2]
plot(u,v,pch=19,cex=0.5)
lines(c(c/(1+c),1),c(0,1),col="red",lwd=3)
lines(c(0,c/(1+c)),c(1,0),col="red",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Linear_Output=rbind(ML_Linear_Output, dframe)
row.names(ML_Linear_Output)[3]="Gumbel"
dframe <- data.frame(c, theta)
Params_Linear_Output=rbind(Params_Linear_Output, dframe)
row.names(Params_Linear_Output)[3]="Gumbel"

```

```

# COPULA GAUSSIANA
L=function(c,theta){
  -log(Gauss_Density(u,f_lin(v,c),theta)/(c+1)+
        c*Gauss_Density(u,f_lin(v,c)-v,theta)/(c+1))}
# Stima dei parametri
fobj=function(vec){sum(L(exp(vec[1]),0.99*tanh(vec[2])))}
out=optim(par=c(1,0), fn=fobj,method="BFGS")
out$convergence
params=c(exp(out$par[1]),0.99*tanh(out$par[2]))
ML=-out$value
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
c=params[1]
theta=params[2]
plot(u,v,pch=19,cex=0.5)
lines(c(c/(1+c),1),c(0,1),col="red",lwd=3)
lines(c(0,c/(1+c)),c(1,0),col="red",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Linear_Output=rbind(ML_Linear_Output,dframe)
row.names(ML_Linear_Output)[4]="Gaussian"
dframe <- data.frame(c, theta)
Params_Linear_Output=rbind(Params_Linear_Output,dframe)
row.names(Params_Linear_Output)[4]="Gaussian"

# COPULA t
L=function(c,theta,df){
  -log(t_Density(u,f_lin(v,c),theta,df)/(c+1)+
        c*t_Density(u,f_lin(v,c)-v,theta,df)/(c+1))}
# Stima dei parametri
fobj=function(vec){
  sum(L(exp(vec[1]),tanh(vec[2]),exp(vec[3])))}
out=optim(par=c(1,1,1), fn=fobj,method="BFGS")
out$convergence
params=c(exp(out$par[1]),tanh(out$par[2]),exp(out$par[3]))
ML=-out$value
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
c=params[1]
theta=params[2]
dof_lin=params[3]
plot(u,v,pch=19,cex=0.5)
lines(c(c/(1+c),1),c(0,1),col="red",lwd=3)
lines(c(0,c/(1+c)),c(1,0),col="red",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Linear_Output=rbind(ML_Linear_Output,dframe)
row.names(ML_Linear_Output)[5]="t"
dframe <- data.frame(c, theta)
Params_Linear_Output=rbind(Params_Linear_Output,dframe)
row.names(Params_Linear_Output)[5]="t"

```

```
#####
# CASO 2: f QUADRATICA
#####

# COPULA FRANK
L=function(a,c,theta){
  -log(Frank_Density(u,f_quad(v,a,c),theta)*df_quad(v,a,c)+
      Frank_Density(
        u,f_quad(v,a,c)-v,theta)*(1-df_quad(v,a,c)))}
# Stima dei parametri
fobj=function(vec){sum(L(vec[1],vec[2],vec[3]))}
ui=matrix(c(0,1,0,0,-1,0,1,1,0,-1,-1,0,-1,1,0,1,-1,0),
          ncol=3,byrow=T)
ci=c(0,-1,0,-1,0,-1)
out=constrOptim(theta=c(0,0.5,5),
                f=fobj,ui=ui,ci=ci,method="Nelder-Mead")
out$convergence
params=out$par
ML=-out$value
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
a=params[1]
c=params[2]
theta=params[3]
plot(u,v,pch=19,cex=0.5)
x=seq(round(c,4),1,by=0.0001)
y=(-(1-a-c)+sqrt((1-a-c)^2-4*a*(c-x)))/2/a
points(x,y,col="red",type="l",lwd=3)
x=seq(0,round(c,4),by=0.0001)
y=(a+c-sqrt((a+c)^2-4*a*(c-x)))/2/a
points(x,y,col="red",type="l",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Quadratic_Output=rbind(ML_Quadratic_Output,dframe)
row.names(ML_Quadratic_Output)[1]="Frank"
dframe <- data.frame(a, c, theta)
Params_Quadratic_Output=rbind(Params_Quadratic_Output,dframe)
row.names(Params_Quadratic_Output)[1]="Frank"

# COPULA CLAYTON
L=function(a,c,theta){
  -log(Clayton_Density(u,f_quad(v,a,c),theta)*df_quad(v,a,c)+
      Clayton_Density(
        u,f_quad(v,a,c)-v,theta)*(1-df_quad(v,a,c)))}
# Stima dei parametri
fobj=function(vec){sum(L(vec[1],vec[2],vec[3]))}
ui=matrix(c(0,1,0,0,-1,0,1,1,0,-1,-1,0,-1,1,0,1,-1,0),
          ncol=3,byrow=T)
ci=c(0,-1,0,-1,0,-1)
ui=rbind(ui,c(0,0,1))
```



```

ci=c(ci,0)
out=constrOptim(theta=c(0,0.5,1),
                f=fobj,ui=ui,ci=ci,method="Nelder-Mead")
out$convergence
params=out$par
ML=-out$value
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
a=params[1]
c=params[2]
theta=params[3]
plot(u,v,pch=19,cex=0.5)
x=seq(round(c,4),1,by=0.0001)
y=(-(1-a-c)+sqrt((1-a-c)^2-4*a*(c-x)))/2/a
points(x,y,col="red",type="l",lwd=3)
x=seq(0,round(c,4),by=0.0001)
y=(a+c-sqrt((a+c)^2-4*a*(c-x)))/2/a
points(x,y,col="red",type="l",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Quadratic_Output=rbind(ML_Quadratic_Output,dframe)
row.names(ML_Quadratic_Output)[2]="Clayton"
dframe <- data.frame(a, c, theta)
Params_Quadratic_Output=rbind(Params_Quadratic_Output,dframe)
row.names(Params_Quadratic_Output)[2]="Clayton"

# COPULA GUMBEL
L=function(a,c,theta){
  -log(Gumbel_Density(u,f_quad(v,a,c),theta)*df_quad(v,a,c)+
       Gumbel_Density(
         u,f_quad(v,a,c)-v,theta)*(1-df_quad(v,a,c)))}
# Stima dei parametri
fobj=function(vec){sum(L(vec[1],vec[2],vec[3]))}
ui=matrix(c(0,1,0,0,-1,0,1,1,0,-1,-1,0,-1,1,0,1,-1,0),
          ncol=3,byrow=T)
ci=c(0,-1,0,-1,0,-1)
ui=rbind(ui,c(0,0,1))
ci=c(ci,1)
out=constrOptim(theta=c(0,0.5,5),
                f=fobj,ui=ui,ci=ci,method="Nelder-Mead")
out$convergence
params=out$par
ML=-out$value
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
a=params[1]
c=params[2]
theta=params[3]
plot(u,v,pch=19,cex=0.5)
x=seq(round(c,4),1,by=0.0001)
y=(-(1-a-c)+sqrt((1-a-c)^2-4*a*(c-x)))/2/a

```

```

points(x,y,col="red",type="l",lwd=3)
x=seq(0,round(c,4),by=0.0001)
y=(a+c-sqrt((a+c)^2-4*a*(c-x)))/2/a
points(x,y,col="red",type="l",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Quadratic_Output=rbind(ML_Quadratic_Output,dframe)
row.names(ML_Quadratic_Output)[3]="Gumbel"
dframe <- data.frame(a, c, theta)
Params_Quadratic_Output=rbind(Params_Quadratic_Output,dframe)
row.names(Params_Quadratic_Output)[3]="Gumbel"

# COPULA GAUSSIANA
L=function(a,c,theta){
  -log(Gauss_Density(u,f_quad(v,a,c),theta)*df_quad(v,a,c)+
      Gauss_Density(
        u,f_quad(v,a,c)-v,theta)*(1-df_quad(v,a,c)))}
# Stima dei parametri
fobj=function(vec){sum(L(vec[1],vec[2],vec[3]))}
ui=matrix(c(0,1,0,0,-1,0,1,1,0,-1,-1,0,-1,1,0,1,-1,0),
          ncol=3,byrow=T)
ci=c(0,-1,0,-1,0,-1)
ui=rbind(ui,c(0,0,1))
ci=c(ci,-1)
ui=rbind(ui,c(0,0,-1))
ci=c(ci,-1)
out=constrOptim(theta=c(0,0.5,0),
                f=fobj,ui=ui,ci=ci,method="Nelder-Mead")
out$convergence
params=out$par
ML=-out$value
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
a=params[1]
c=params[2]
theta=params[3]
plot(u,v,pch=19,cex=0.5)
x=seq(round(c,4),1,by=0.0001)
y=(-(1-a-c)+sqrt((1-a-c)^2-4*a*(c-x)))/2/a
points(x,y,col="red",type="l",lwd=3)
x=seq(0,round(c,4),by=0.0001)
y=(a+c-sqrt((a+c)^2-4*a*(c-x)))/2/a
points(x,y,col="red",type="l",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Quadratic_Output=rbind(ML_Quadratic_Output,dframe)
row.names(ML_Quadratic_Output)[4]="Gaussian"
dframe <- data.frame(a, c, theta)
Params_Quadratic_Output=rbind(Params_Quadratic_Output,dframe)
row.names(Params_Quadratic_Output)[4]="Gaussian"

# COPULA t

```

```

L=function(a,c,theta,dof){
  -log(t_Density(u,f_quad(v,a,c),theta,dof)*df_quad(v,a,c)+
      t_Density(
        u,f_quad(v,a,c)-v,theta,dof)*(1-df_quad(v,a,c)))}

fobj=function(vec){sum(L(vec[1],vec[2],vec[3],vec[4]))}
ui=matrix(
  c(0,1,0,0,0,-1,0,0,1,1,0,0,-1,-1,0,0,-1,1,0,0,1,-1,0,0),
  ncol=4,byrow=T)
ci=c(0,-1,0,-1,0,-1)
ui=rbind(ui,c(0,0,1,0))
ci=c(ci,-1)
ui=rbind(ui,c(0,0,-1,0))
ci=c(ci,-1)
ui=rbind(ui,c(0,0,0,1))
ci=c(ci,0)
out=constrOptim(theta=c(0,0.5,0,1),
                f=fobj,ui=ui,ci=ci,method="Nelder-Mead")
out$convergence
params=out$par
ML=-out$value
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
a=params[1]
c=params[2]
theta=params[3]
dof_quad=params[4]
plot(u,v,pch=19,cex=0.5)
x=seq(round(c,4),1,by=0.0001)
y=(-(1-a-c)+sqrt((1-a-c)^2-4*a*(c-x)))/2/a
points(x,y,col="red",type="l",lwd=3)
x=seq(0,round(c,4),by=0.0001)
y=(a+c-sqrt((a+c)^2-4*a*(c-x)))/2/a
points(x,y,col="red",type="l",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Quadratic_Output=rbind(ML_Quadratic_Output,dframe)
row.names(ML_Quadratic_Output)[5]="t"
dframe <- data.frame(a, c, theta)
Params_Quadratic_Output=rbind(Params_Quadratic_Output,dframe)
row.names(Params_Quadratic_Output)[5]="t"

#####
# CASO 3: f POTENZA
#####
library(nloptr)

# Funzione per la definizione del vincolo non lineare
eval_g_ineq=function(vec){
  a=vec[1]
  c=vec[2]

```

```

return(c(a-1-c,a/(1+c)*(c/(1+c))^(a-1)-1))}

# COPULA FRANK
L=function(a,c,theta){
  -log(Frank_Density(u,f_pow(v,a,c),theta)*df_pow(v,a,c)+
      Frank_Density(
        u,f_pow(v,a,c)-v,theta)*(1-df_pow(v,a,c)))}
# Stima dei parametri
fobj=function(vec){sum(L(vec[1],vec[2],vec[3]))}
lower=c(0,0,0)
out=nloptr(c(0.25,0.25,5),fobj,
  lb = lower,
  eval_g_ineq = eval_g_ineq,
  opts = list("algorithm"="NLOPT_LN_COBYLA",
              "xtol_rel"=1.0e-8,
              "xtol_abs"=1.0e-10,
              "maxeval"= 5000))

out$status
params=out$solution
ML=-out$objective
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
a=params[1]
c=params[2]
theta=params[3]
plot(u,v,pch=19,cex=0.5)
x=seq(round((c/(1+c))^a,4),1,by=0.0001)
y=x^(1/a)*(1+c)-c
points(x,y,col="red",type="l",lwd=3)
x=seq(0,round((c/(1+c))^a,4)-0.0001,by=0.0001)
y=aux_inv(x,a,c)
points(x,y,col="red",type="l",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Power_Output=rbind(ML_Power_Output,dframe)
row.names(ML_Power_Output)[1]="Frank"
dframe <- data.frame(a, c, theta)
Params_Power_Output=rbind(Params_Power_Output,dframe)
row.names(Params_Power_Output)[1]="Frank"

# COPULA CLAYTON
L=function(a,c,theta){
  -log(Clayton_Density(u,f_pow(v,a,c),theta)*df_pow(v,a,c)+
      Clayton_Density(
        u,f_pow(v,a,c)-v,theta)*(1-df_pow(v,a,c)))}
# Stima dei parametri
fobj=function(vec){sum(L(vec[1],vec[2],vec[3]))}
lower=c(0,0,0)
out=nloptr(c(0.25,0.25,5),fobj,
  lb = lower, eval_g_ineq = eval_g_ineq,
  opts = list("algorithm"="NLOPT_LN_COBYLA",
              "xtol_rel"=1.0e-8,

```

```

                                "xtol_abs"=1.0e-10,
                                "maxeval"= 5000))
out$status
params=out$solution
ML=-out$objective
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
a=params[1]
c=params[2]
theta=params[3]
plot(u,v,pch=19,cex=0.5)
x=seq(round((c/(1+c))^a,4),1,by=0.0001)
y=x^(1/a)*(1+c)-c
points(x,y,col="red",type="l",lwd=3)
x=seq(0,round((c/(1+c))^a,4)-0.0001,by=0.0001)
y=aux_inv(x,a,c)
points(x,y,col="red",type="l",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Power_Output=rbind(ML_Power_Output,dframe)
row.names(ML_Power_Output)[2]="Clayton"
dframe <- data.frame(a, c, theta)
Params_Power_Output=rbind(Params_Power_Output,dframe)
row.names(Params_Power_Output)[2]="Clayton"

# COPULA GUMBEL
L=function(a,c,theta){
  -log(Gumbel_Density(u,f_pow(v,a,c),theta)*df_pow(v,a,c)+
      Gumbel_Density(
        u,f_pow(v,a,c)-v,theta)*(1-df_pow(v,a,c)))}
# Stima dei parametri
fobj=function(vec){sum(L(vec[1],vec[2],vec[3]))}
lower=c(0,0,1)
out=nloptr(c(0.25,0.25,2),fobj,
          lb = lower, eval_g_ineq = eval_g_ineq,
          opts = list("algorithm"="NLOPT_LN_COBYLA",
                      "xtol_rel"=1.0e-8,
                      "xtol_abs"=1.0e-10,
                      "maxeval"= 5000))

out$status
params=out$solution
ML=-out$objective
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
a=params[1]
c=params[2]
theta=params[3]
plot(u,v,pch=19,cex=0.5)
x=seq(round((c/(1+c))^a,4),1,by=0.0001)
y=x^(1/a)*(1+c)-c
points(x,y,col="red",type="l",lwd=3)

```

```

x=seq(0,round((c/(1+c))^a-0.0001,4),by=0.0001)
y=aux_inv(x,a,c)
points(x,y,col="red",type="l",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Power_Output=rbind(ML_Power_Output,dframe)
row.names(ML_Power_Output)[3]="Gumbel"
dframe <- data.frame(a, c, theta)
Params_Power_Output=rbind(Params_Power_Output,dframe)
row.names(Params_Power_Output)[3]="Gumbel"

# COPULA GAUSSIANA
L=function(a,c,theta){
  -log(Gauss_Density(u,f_pow(v,a,c),theta)*df_pow(v,a,c)+
      Gauss_Density(
        u,f_pow(v,a,c)-v,theta)*(1-df_pow(v,a,c)))}
# Stima dei parametri
fobj=function(vec){sum(L(vec[1],vec[2],vec[3]))}
lower=c(0,0,-1)
upper=c(Inf,Inf,1)
out=nloptr(c(0.25,0.25,0),fobj,
          lb = lower, ub = upper, eval_g_ineq = eval_g_ineq,
          opts = list("algorithm"="NLOPT_LN_COBYLA",
                     "xtol_rel"=1.0e-10,
                     "xtol_abs"=1.0e-12,
                     "maxeval"= 10000))

out$status
params=out$solution
ML=-out$objective
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
a=params[1]
c=params[2]
theta=params[3]
plot(u,v,pch=19,cex=0.5)
x=seq(round((c/(1+c))^a,4),1,by=0.0001)
y=x^(1/a)*(1+c)-c
points(x,y,col="red",type="l",lwd=3)
x=seq(0,round((c/(1+c))^a,4)-0.0001,by=0.0001)
y=aux_inv(x,a,c)
points(x,y,col="red",type="l",lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Power_Output=rbind(ML_Power_Output,dframe)
row.names(ML_Power_Output)[4]="Gaussian"
dframe <- data.frame(a, c, theta)
Params_Power_Output=rbind(Params_Power_Output,dframe)
row.names(Params_Power_Output)[4]="Gaussian"

# COPULA t
L=function(a,c,theta,dof){
  -log(t_Density(u,f_pow(v,a,c),theta,dof)*df_pow(v,a,c)+

```

```

        t_Density(
          u, f_pow(v, a, c) - v, theta, dof) * (1 - df_pow(v, a, c)))}
# Stima dei parametri
fobj=function(vec){sum(L(vec[1], vec[2], vec[3], vec[4]))}
lower=c(0,0,-1,0)
upper=c(Inf, Inf, 1, Inf)
out=nloptr(c(0.25,0.25,0,1), fobj,
          lb = lower, ub = upper, eval_g_ineq = eval_g_ineq,
          opts = list("algorithm"="NLOPT_LN_COBYLA",
                      "xtol_rel"=1.0e-8,
                      "xtol_abs"=1.0e-10,
                      "maxeval"= 5000))

out$status
params=out$solution
ML=-out$objective
k=length(params)
AIC=2*k-2*ML
BIC=log(n)*k-2*ML
a=params[1]
c=params[2]
theta=params[3]
dof_pow=params[4]
plot(u, v, pch=19, cex=0.5)
x=seq(round((c/(1+c))^a, 4), 1, by=0.0001)
y=x^(1/a)*(1+c)-c
points(x, y, col="red", type="l", lwd=3)
x=seq(0, round((c/(1+c))^a, 4)-0.0001, by=0.0001)
y=aux_inv(x, a, c)
points(x, y, col="red", type="l", lwd=3)
dframe <- data.frame(ML, AIC, BIC)
ML_Power_Output=rbind(ML_Power_Output, dframe)
row.names(ML_Power_Output)[5]="t"
dframe <- data.frame(a, c, theta)
Params_Power_Output=rbind(Params_Power_Output, dframe)
row.names(Params_Power_Output)[5]="t"

# RISULTATI
ML_Linear_Output
Params_Linear_Output
dof_lin

ML_Quadratic_Output
Params_Quadratic_Output
dof_quad

ML_Power_Output
Params_Power_Output
dof_pow

```

## Appendice B

# Codici R: simulazioni Monte Carlo

```
#####
#  SIMULAZIONE COPULA FRANK MEDIANTE IL METODO
#  DELLA DISTRIBUZIONE CONDIZIONATA
#####
nsim=10000
theta=10

v=runif(nsim)
unif=runif(nsim)
u=-log(1+unif*(exp(-theta)-1)/(exp(-theta*v)-
                                     unif*(exp(-theta*v)-1)))/theta

hist(u)
hist(v)
plot(u, v, pch=19, cex=0.5)

#####
#  SIMULAZIONE COPULA GAUSSIANA MEDIANTE
#  TRASFORMAZIONE INTEGRALE DI PROBABILITA'
#####
nsim=10000
rho=0.9

mu=c(0,0)
Sigma=matrix(c(1,rho,rho,1),ncol=2)

library(MASS)
Z=mvrnorm(n=nsim, mu, Sigma)
u=pnorm(Z[,1])
v=pnorm(Z[,2])

hist(u)
hist(v)
```



```

plot(u, v, pch=19, cex=0.5)

#####
#  SIMULAZIONE COPULA NON MONOTONA DEL CAPITOLO 4
#    (FORMULA (4.13)) MEDIANTE TRASFORMAZIONE
#    INTEGRALE DI PROBABILITA'
#####
nsim=10000
rho=0.9

mu=c(0,0)
Sigma=matrix(c(1,rho,rho,1),ncol=2)

Z=mvrnorm(n=nsim, mu, Sigma)
u=pnorm(Z[,1])
v=pchisq(Z[,2]^2, df=1)

hist(u)
hist(v)
plot(u, v, pch=19, cex=0.5)

#####
#  SIMULAZIONE COPULA t MEDIANTE
#    TRASFORMAZIONE INTEGRALE DI PROBABILITA'
#####
nsim=10000
rho=0.9
df=3

Sigma=matrix(c(1, rho, rho, 1),ncol=2)

Z=mvrnorm(n=nsim, mu, Sigma)
Y=rchisq(n=nsim, df=df)
X=sqrt(df)/sqrt(Y)*Z
u=pt(X[,1], df=df)
v=pt(X[,2], df=df)

hist(u)
hist(v)
plot(u, v, pch=19, cex=0.5)

#####
#  SIMULAZIONE COPULA CLAYTON MEDIANTE
#    UTILIZZO DEL GENERATORE ARCHIMEDEO
#####
nsim=10000
theta=10

```

```

d=2

phi_inv=function(s,theta){
  out=(1+theta*s)^(-1/theta)
  return(out)}

X=matrix(runif(nsim*d), ncol=d)
M=rgamma(n=nsim, shape=1/theta, scale=theta)

Cop=phi_inv(-log(X)/M,theta)

u_clayton=Cop[,1]
v_clayton=Cop[,2]
hist(u_clayton)
hist(v_clayton)
plot(u_clayton, v_clayton, pch=19, cex=0.5)

#####
# SIMULAZIONE COPULA t MEDIANTE RAPPRESENTAZIONE
# STOCASTICA DELLA DISTR. ELLITTICA E
# TRASFORMAZIONE INTEGRALE DI PROBABILITA'
#####
nsim=10000
d=2
rho=0.9
Sigma=matrix(c(1,rho,rho,1),ncol=2)
Chol=chol(Sigma)
df=3

library(zipfR)
Z=mvrnorm(n=nsim,mu=rep(0,d),Sigma=diag(d))
Ud=Z/apply(Z,1,function(vec){sqrt(sum(vec^2))})
temp=Ibeta.inv(runif(n=nsim)*beta(df/2,d/2),df/2,d/2)
R=sqrt(df*(1-temp)/temp)
X=(R*Ud)%*%Chol

u=pt(X[,1], df=df)
v=pt(X[,2], df=df)

hist(u)
hist(v)
plot(u, v, pch=19, cex=0.5)

#####
# SIMULAZIONE COPULA NON MONOTONA Q CON:
# - C copula Frank
# - f(v)=((v+c)/(1+c))^a
# METODO: DISTRIBUZIONE CONDIZIONATA

```

```
#####
a=0.75
c=0.22
theta=20

f_pow=function(v,a,c){((v+c)/(1+c))^a}
df_pow=function(v,a,c){a/(1+c)*((v+c)/(1+c))^(a-1)}
ddf_pow=function(v,a,c){a*(a-1)/(1+c)^2*((v+c)/(1+c))^(a-2)}

f=function(v){f_pow(v,a,c)}
df=function(v){df_pow(v,a,c)}

QFrankSimul=function(v,unif,theta){
  t1=exp(-theta*f(v))
  t2=exp(-theta*(f(v)-v))
  k1=df(v)
  k2=1-df(v)
  a=unif*(t1-1)*(t2-1)-t1*k1*(t2-1)-t2*k2*(t1-1)
  b=(exp(-theta)-1)*(unif*(t1+t2-2)-(k1*t1+k2*t2))
  c=unif*(exp(-theta)-1)^2
  sol=(-b+sqrt(b^2-4*a*c))/(2*a)
  print(sol)
  out=-log(1+sol)/theta
  return(out)}

nsim=10000
temp=runif(2*nsim)

v=temp[1:nsim]
u=QFrankSimul(v,temp[(nsim+1):(2*nsim)],theta)

hist(u)
hist(v)
plot(u, v, pch=19, cex=0.5)

#####
# SIMULAZIONE COPULA NON MONOTONA Q CON:
# - C copula Clayton
# - f(v)=((v+c)/(1+c))^a
# METODO: TRASFORMAZIONE PRESERVANTE LA
# DISTR. UNIFORME
#####
a=0.75
c=0.22

g_minus=function(x,a,c){
  x0=rep(0.5,length(x))
  error=max(abs(f_pow(x0,a,c)-x0-x))
  while(error>1e-12){
    x_next=x0-2*(f_pow(x0,a,c)-x0-x)*(df_pow(x0,a,c)-1)/
      (2*(df_pow(x0,a,c)-1)^2-
```

```
        (f_pow(x0,a,c)-x0-x)*ddf_pow(x0,a,c))
    error=max(abs(f_pow(x_next,a,c)-x_next-x))
    x0=x_next}
return(x0)}

g_plus=function(x,a,c){return(x^(1/a)*(1+c)-c)}

g=function(x,a,c){
  threshold=(c/(1+c))^a
  greater=which(x>=threshold)
  lower=which(x<threshold)
  x[greater]=g_plus(x[greater],a,c)
  x[lower]=g_minus(x[lower],a,c)
  return(x)}

u=u_clayton
v=g(v_clayton, a, c)
hist(u)
hist(v)
plot(u, v, pch=19, cex=0.5)
```