



**UNIVERSITA' DEGLI STUDI
DI PADOVA**

FACOLTA' DI SCIENZE STATISTICHE

CORSO DI LAUREA SPECIALIASTICA
IN SCIENZE STATISTICHE, ECONOMICHE, FINANZIARIE
E AZIENDALI

TESI DI LAUREA

**IL METODO DELLE COPULE:
inferenza, calcolo del VaR ed
implementazione in R**

RELATORE: CH.MO PROF. NUNZIO CAPPUCCIO

LAUREANDO: ALESSANDRO ZILIO

ANNO ACCADEMICO 2004-2005

*Alla memoria di mia madre,
a mio padre e mio fratello
che mi hanno sempre dato fiducia.
Ai miei amici,
ai compagni d'Università e
a quanti mi sono stati vicini
in questi mesi difficili,
ma indimenticabili.*

INDICE

<i>INTRODUZIONE</i>	9
<i>1. COPULE MULTIVARIATE</i>	11
1.1 DEFINIZIONE E PROPRIETA' BASE.....	11
1.2 LIMITI DI FRECHET E ORDINE DI CONCORDANZA:.....	15
1.3 IL TEOREMA DI SKLAR E L'INTERPRETAZIONE PROBABILISTICA DI BASE.....	16
1.4 DISTRIBUZIONE CONDIZIONALE VIA COPULA.....	19
1.5 COPULA DI SOPRAVVIVENZA E FUNZIONE DI SOPRAVVIVENZA CONGIUNTA.....	20
1.6 DENSITA' E RAPPRESENTAZIONE CANONICA DI UNA COPULA MULTIDIMENSIONALE.....	23
<i>2. MISURE DI DIPENDENZA E FAMIGLIE DI COPULE</i>	25
2.1 LE COPULE COME FUNZIONI DI DIPENDENZA:.....	25
2.1.1 <i>Indipendenza</i>	25
2.1.2 <i>Comonotonicità</i>	26
2.2 MISURE DI ASSOCIAZIONE.....	27
2.2.1 <i>CONCORDANZA</i>	28
2.2.2 <i>τ DI KENDALL</i>	30
2.2.3 <i>ρ DI SPEARMAN</i>	32
2.2.4 <i>CORRELAZIONE LINEARE</i>	34
2.2.5 <i>DIPENDENZA DI CODA</i>	37
2.2.6 <i>DIPENDENZA POSITIVA DI QUADRANTE</i>	39

2.3 FAMIGLIE PARAMETRICHE DI COPULE N-DIMENSIONALI.....	41
2.3.1 COPULA GAUSSIANA MULTIVARIATA (M.G.C).....	41
2.3.2 COPULA T-STUDENT MULTIVARIATA (MTC).....	42
2.3.3 COPULA DI DISPERSIONE MULTIVARIATA (MDC).....	43
2.3.4 COPULE DI ARCHIMEDE.....	44
2.3.5 PLACKETT COPULA.....	47
3. INFERENZA STATISTICA PER LE COPULE.....	49
3.1 METODO DI MASSIMA VEROSIMIGLIANZA ESATTA.....	49
3.2 METODO IFM.....	50
3.3 IL METODO IFM APPLICATO AD UN CASO PRATICO.....	52
3.3.1 I DATI.....	52
3.3.2 STIMA DEI PARAMETRI DELLE MARGINALI.....	55
3.3.3 STIMA DEI PARAMETRI DELLA COPULA.....	57
3.3.4 STIMA DELLA MATRICE DI VARIANZA E COVARIANZA.....	57
4. APPLICAZIONE DELLE COPULE AL CALCOLO DEL VALUE AT RISK (VAR).....	63
4.1 DEFINIZIONE DI VAR.....	63
4.2 CALCOLO DEL VaR CON IL METODO DELLE COPULE.....	65
4.3 ESEMPIO PRATICO DI CALCOLO DEL VAR CON LE COPULE.....	66
5. IMPLEMENTAZIONE IN R: LA FUNZIONE COPULA.....	77
6. CONCLUSIONI.....	89
APPENDICE.....	94
BIBLIOGRAFIA.....	116

INTRODUZIONE

Il metodo delle copule è uno strumento relativamente recente che permette di manipolare in maniera flessibile i fattori di rischio ed altre variabili rilevanti studiate in finanza. Sebbene gli strumenti di questa metodologia siano adottati dalla statistica, essi sono stati fatti propri da molti addetti ai lavori nel campo della finanza per far fronte ai problemi legati :

- alla volatilità;
- alla non-normalità della distribuzione dei rendimenti degli asset finanziari (in particolare per il problema delle “code-pesanti”);
- alla dipendenza tra gli asset, spesso diversa da quella normale (spiegata solamente dalla correlazione lineare).

In particolare, quando ci si trova a lavorare in ambiti multidimensionali, dove sono presenti più fattori di rischio, l’assunzione di una distribuzione normale multivariata per essi porta a conclusioni nelle analisi svolte che sono distanti da quelle riscontrate empiricamente.

La ricerca di nuove distribuzioni multivariate che meglio si adattino al reale comportamento dei fattori di rischio, ha portato la metodologia delle copule ad essere uno strumento utile per la risoluzione di queste problematiche. La caratteristica che meglio definisce questo nuovo strumento di analisi finanziaria è il fatto di poter scindere le questioni legate all’analisi univariata delle marginali, da quelle che si riferiscono alla struttura di dipendenza.

La presente tesi ha lo scopo di presentare questa nuova metodologia e di fornirne un’applicazione all’analisi finanziaria. Nel concreto, gli obiettivi che ci si propone di ottenere sono:

- Raccogliere le proprietà matematiche delle funzioni copula ed evidenziare le basi su cui si fonda la loro teoria.
- Dare una definizione di copula dal punto di vista probabilistico, e di procedere quindi ad affrontare i problemi di natura inferenziale sui parametri che caratterizzano la funzione copula.
- Definire le più importanti famiglie di copule

- Applicare la metodologia delle copule al rischio di mercato e precisamente al calcolo del Value at Risk, per vedere come la non-normalità della distribuzione multivariata cambi la “consistenza” di questa misura.
- Implementare in un linguaggio informatico i precedenti due punti

1. COPULE MULTIVARIATE

In questo primo capitolo vengono presentate le definizioni e le proprietà matematiche che stanno alla base del concetto di copula e che consentono alla stessa di essere una funzione di distribuzione.

Da questa caratterizzazione si arriva poi ad enunciare il **teorema di Sklar** (il risultato maggiore della teoria delle copule) che consente di scrivere la copula come funzione delle marginali e viceversa.

1.1 DEFINIZIONE E PROPRIETA' BASE

Per dare una definizione di copula, abbiamo bisogno di 2 proprietà che consentono alla copula di rispettare le proprietà di funzione di distribuzione. Le proprietà di **crescita** e **consistenza**.

Definizione 1.1 Sia la funzione $G : \mathfrak{R}^n \rightarrow \mathfrak{R}$ con dominio $\text{Dom } G = A_1 \times \dots \times A_n$, dove gli insiemi non vuoti A_i hanno come valore più piccolo il valore a_i . La funzione G è detta essere **consistente** se e solo se si annulla in corrispondenza di un valore $v_k \in \text{Dom } G$, tale che $v_k = a_k$. In simboli

$$G(v) = G(v_1, \dots, v_{k-1}, a_k, v_{k+1}, \dots, v_n) = 0.$$

In pratica, una funzione è consistente, se si annulla quando uno solo degli insiemi A_i assume il suo valore più piccolo nel suo dominio.

Per mostrare la proprietà di n-crescita si deve partire dal considerare un generico n-intervallo A . Esso è definito come il prodotto cartesiano di n intervalli chiusi:

$$A = [u_{11}, u_{12}] \times \dots \times [u_{n1}, u_{n2}]$$

con $u_{i1} \leq u_{i2}$, $i = 1, 2, \dots, n$.

Denotiamo inoltre con w un generico vertice di A e con $\text{ver}(A)$ l'insieme di tutti i vertici di A : allora, $w \in \text{ver}(A)$ se e solo se la sua i -esima componente w_i , con $i=1, \dots, n$, è uguale a u_{i1} o u_{i2} .

Si consideri il prodotto

$$\prod_{i=1}^n \text{sgn}(2w_i - u_{i1} - u_{i2})$$

Poiché ciascun fattore del prodotto è -1 se $w_i = u_{i1} < u_{i2}$, 0 se $w_i = u_{i1} = u_{i2}$, e $+1$ se $w_i = u_{i1} > u_{i2}$, abbiamo che:

$$\prod_{i=1}^n \text{sgn}(2w_i - u_{i1} - u_{i2}) = \begin{cases} -1 & \text{se } u_{i1} \neq u_{i2}, \forall i, \# [i : w_i = u_{i2}] = 2m + 1 \\ 0 & \text{se } \exists i : u_{i1} = u_{i2} \\ +1 & \text{se } u_{i1} \neq u_{i2}, \forall i, \# [i : w_i = u_{i1}] = 2m \end{cases} \quad m \in \mathbb{N}$$

Se $\text{ver}(A) \subset \text{Dom } G$, definiamo il G -volume di A come la somma

$$\sum_{w \in \text{ver}(A)} G(w) \prod_{i=1}^n \text{sgn}(2w_i - u_{i1} - u_{i2}) \quad (1.1)$$

La (1.1) misura il volume dell' n -scatola A ed è stata introdotta per spiegare il concetto di funzione **n -crescente**.

Definizione 1.2 La funzione $G: A_1 \times \dots \times A_n \rightarrow \mathfrak{R}$ è n -crescente se il G -volume di A è non negativo per ogni n -scatola A per la quale $\text{ver}(A) \subset \text{Dom } G$:

$$\sum_{w \in \text{ver}(A)} G(w) \prod_{i=1}^n \text{sgn}(2w_i - u_{i1} - u_{i2}) \geq 0$$

Adesso che abbiamo definito le proprietà di consistenza e di funzione n -crescente, possiamo presentare alcune definizioni che valgono per le funzioni con queste proprietà.

Teorema 1.1 La funzione $G: A_1 \times \dots \times A_n \mapsto \mathfrak{R}$ consistente ed n -crescente, è non decrescente in ogni suo argomento.

Dimostrazione: Dobbiamo dimostrare che se $(u_1, \dots, u_{i-1}, x, u_{i+1}, \dots, u_n) \in \text{Dom } G$ e $x \leq y$, allora

$$G(u_1, \dots, u_{i-1}, x, u_{i+1}, \dots, u_n) \leq G(u_1, \dots, u_{i-1}, y, u_{i+1}, \dots, u_n) \quad (1.2)$$

Per questo scopo consideriamo la n -scatola :

$$A = [a_1, u_1] \times \dots \times [a_{i-1}, u_{i-1}] \times [x, y] \times [a_{i+1}, u_{i+1}] \dots \times [a_n, u_n].$$

Dalla definizione di consistenza, il G -volume di A è

$$G(u_1, \dots, u_{i-1}, y, u_{i+1}, \dots, u_n) - G(u_1, \dots, u_{i-1}, x, u_{i+1}, \dots, u_n)$$

Inoltre , poiché G è n -crescente , il G -volume è non negativo e ne segue quindi la (1.2).

Definiamo ora il concetto di marginali k -dimensionale e 1-dimensionale.

Definizione 1.3 Se tutti i sottoinsiemi A_i sono non vuoti, le marginali k -dimensionali della funzione $G : A_1 \times \dots \times A_n \rightarrow \mathfrak{R}$, sono le funzioni $C_{i_1 \dots i_k}(u_{i_1}, \dots, u_{i_k}) : A_{i_1} \times \dots \times A_{i_k} \rightarrow \mathfrak{R}$, con $1 \leq k \leq n$, $k \in \mathbb{N}$, definite da

$$G_{i_1 \dots i_k}(u_{i_1}, \dots, u_{i_k}) = G(\overline{a_1}, \dots, u_{i_1}, \dots, u_{i_2}, \dots, u_{i_k}, \dots, \overline{a_n}) , \quad (1.3)$$

dove $\overline{a_i}$ è l'elemento più grande del generico sottoinsieme A_i , ed i_1, \dots, i_k è ciascuna selezione di k indici tra gli originali n .

In particolare , abbiamo che:

Definizione 1.4 Se ciascun A_i è non vuoto e sia $\overline{a_i}$ il suo elemento più grande, allora la i -esima marginale 1-dimensionale della funzione $G: A_1 \times \dots \times A_n \rightarrow \mathfrak{R}$, è la funzione $G_i(u) : A_i \rightarrow \mathfrak{R}$ definita da

$$G_i(u) = G(\overline{a_1}, \dots, \overline{a_{i-1}}, u, \overline{a_{i+1}}, \dots, \overline{a_n}).$$

Una funzione consistente e n -crescente con marginali 1-dimensionali soddisfa i seguenti lemmi, che sono usati nella dimostrazione del teorema di Sklar.

Lemma 1.1 : Una funzione $G : A_1 \times \dots \times A_n \rightarrow \mathfrak{R}$ consistente , n -crescente, con marginali 1-dimensionale , è tale che

$$G(u^y) - G(u^x) \leq G_i(y) - G_i(x)$$

per ogni $1 \leq i \leq n$, $x < y$ e con

$$u^y = (u_1, \dots, u_{i-1}, y, u_{i+1}, \dots, u_n) \in \text{Dom } G$$

$$u^x = (u_1, \dots, u_{i-1}, x, u_{i+1}, \dots, u_n) \in \text{Dom } G.$$

Lemma 1.2 : Per ogni funzione G del precedente lemma,

$$|G(u) - G(z)| \leq \sum_{i=1}^n |G_i(u_i) - G_i(\dot{u}_i)|$$

per ogni $u = (u_1, \dots, u_n)$ e $\dot{u} = (\dot{u}_1, \dots, \dot{u}_n) \in \text{Dom } G$.

Una volta assunta questa terminologia si arriva a dare una definizione di copula e subcopula n-dimensionale:

Definizione 1.5 Una subcopula n-dimensionale è una funzione $C: A_1 \times \dots \times A_n \rightarrow \mathfrak{R}$, con $A_i \subset [0,1]$ per ogni i , tale che

- i. C è consistente
- ii. Le sue marginali 1-dimensionale sono le funzione identità in $[0,1]$:
 $C_i(u)=u, i=1, \dots, n$
- iii. C è n-crescente.

Nel caso in cui $A_i=[0,1]$, per ogni i , una subcopula n-dimensionale diventa una copula C .

In particolare,

Teorema 1.2 Per $n>2$, $1<k<n$, le marginali k-dimensionali di C sono copule k-dimensionali.

Dimostrazione: Si deve mostrare che, quando $A_i=[0,1]$ per ogni i , per la funzione

$$C_{i_1 \dots i_k}(u_{i_1}, \dots, u_{i_k}) : I^k \rightarrow \mathfrak{R}$$

definita come (1.3), valgono le 3 caratteristiche della definizione 1.5

La (i) vale se per almeno un indice i_j tale che $u_{i_j}=a_i$, abbiamo

$$C_{i_1 \dots i_k}(u_{i_1}, \dots, u_{i_k}) = C(\overline{a_1}, \dots, u_{i_1}, \dots, u_{i_j-1}, a_{i_j}, u_{i_j+1}, \dots, u_{i_k}, \overline{a_n})=0,$$

Dato che il vettore $C(\overline{a_1}, \dots, u_{i_1}, \dots, u_{i_j-1}, a_{i_j}, u_{i_j+1}, \dots, u_{i_k}, \overline{a_n})$ ha la caratteristica di v nella definizione 1.1, la (i) è dimostrata

La (ii) deriva dal fatto che le marginali 1-dimensionali di C , sono anche marginali 1-dimensionali per le loro marginali k-dimensionali.

La (iii) è una conseguenza del fatto che il $C_{i_1 \dots i_k}$ -volume di ciascuna k-scatola

$$[u_{i_1}, u_{i_2}] \times \dots \times [u_{i_k}, u_{i_{k+1}}]$$

è la C-colonna della n-scatola

$$[\overline{a_1}, \overline{a_1}] \times [u_{i_1}, u_{i_2}] \times \dots \times [u_{i_k}, u_{i_{k+1}}] \times \dots \times [\overline{a_n}, \overline{a_n}]$$

che è non negativa perché n-crescente.

Dal teorema 1.2 segue che la copula C è una funzione di distribuzione in $[0,1]^n$, con marginali uniformemente distribuite in $[0,1]$. Come conseguenza di questo, la copula C

- È non decrescente in ogni suo argomento (teorema 1.1)
- è uniformemente crescente (conseguenza del lemma 1.2) :

$$|C(u)-C(\dot{u})| \leq \sum_{i=1}^n |u_i - \dot{u}_i|,$$

per ogni $u, \dot{u} \in [0,1]^n$.

- ammette derivate parziali miste con $1 \leq k \leq n$ e

$$0 \leq \frac{\partial^k C(u)}{\partial u_1 \dots \partial u_k} \leq 1$$

1.2 LIMITI DI FRECHET E ORDINE DI CONCORDANZA:

Introduciamo nei seguenti teoremi i **limiti di Frechet**, che consentono di porre dei limiti superiori ed inferiori alle copule e permettono di ordinarle secondo l'**ordine di concordanza**.

Teorema 1.3 ogni copula soddisfa la disuguaglianza:

$$\max(u_1 + \dots + u_{n-1} - 1, 0) \leq C(u) \leq \min(u_1, \dots, u_n)$$

per ogni $u \in I^n$.

Il limite superiore è denotato come C^+ e soddisfa la definizione di copula, mentre il limite inferiore non la rispetta per $n > 2$.

Teorema 1.4 Per $n > 2$, per ogni $u \in I^n$ esiste una copula C_u tale che

$$C_u(u) = \max(u_1 + \dots + u_{n-1} - 1, 0)$$

Si veda Nelsen (1999) per la dimostrazione

Introduciamo ora la nozione di **funzione di sopravvivenza** per vettori n-dimensionali di variabili uniformi.

Definizione 1.6 La funzione di sopravvivenza congiunta per i vettori (U_1, \dots, U_n) di variabili casuali uniformi con copula C , denotato come \mathcal{C} , rappresenta, quando è valutato in (u_1, \dots, u_n) , la probabilità congiunta che (U_1, \dots, U_n) sia più grande di u_1, \dots, u_n :

$$\mathcal{C}(u_1, \dots, u_n) = \Pr(U_1 > u_1, \dots, U_n > u_n).$$

Definizione 1.7 Diamo ora la definizione di **ordine di concordanza** (Nelsen, 1999).

La copula C_1 è più piccola della copula C_2 se e solo se

$$C_1(u) \leq C_2(u)$$

$$\mathcal{C}_1(u) \leq \mathcal{C}_2(u)$$

per ogni $u \in I^n$, dove con I denotiamo il dominio $[0,1]$

1.3 IL TEOREMA DI SKLAR E L'INTERPRETAZIONE PROBABILISTICA DI BASE

Il **teorema di Sklar** è il prodotto più rilevante di tutta la teoria alla base del metodo delle copule, poiché fornisce il fondamento per l'uso della copula in tutte le sue possibili applicazioni.

Il teorema garantisce che non solo ogni copula è una funzione di distribuzione congiunta, se i suoi argomenti sono funzioni di distribuzione marginali, ma lo è anche l'opposto. Ogni funzione di distribuzione congiunta può essere estesa a copula ed inoltre, se le marginali sono continue, l'estensione è unica.

Teorema 1.5 (SKLAR, 1959) Siano date le funzioni di distribuzione marginali $F_1(x_1), \dots, F_n(x_n)$. Quindi per ogni $x = (x_1, \dots, x_n) \in \mathfrak{R}^n$:

- i. Se C è una copula il cui dominio contiene $\text{Ran } F_1 \times \dots \times \text{Ran } F_n$,

$$C(F_1(x_1), \dots, F_n(x_n))$$

è una funzione di distribuzione congiunta con marginali $F_1(x_1), \dots, F_n(x_n)$.

- ii. Al contrario, se H è una funzione di distribuzione congiunta con marginali

$$F_1(x_1), \dots, F_n(x_n)$$

esiste un'unica copula C , con dominio $\text{Ran } F_1 \times \dots \times \text{Ran } F_n$, tale che

$$H(x) = C(F_1(x_1), \dots, F_n(x_n)) \tag{1.4}$$

Se $F_1(x_1), \dots, F_n(x_n)$ sono continue, la copula è unica; altrimenti la copula C è unicamente determinata in $\text{Ran } F_1 \times \dots \times \text{Ran } F_n$.

Per una dimostrazione completa si veda Sklar (1996)

Per definire il corollario al teorema di Sklar abbiamo bisogno di definire F^{-1} come la funzione inversa generalizzata di una data funzione di distribuzione univariata F . $F^{-1}(t) = \inf \{x \in \mathfrak{R} \mid F(x) \geq t\}$, per ogni t in $[0,1]$.

Corollario 1.1 Sia H una funzione di distribuzione n -dimensionale con marginali $F_1(x_1), \dots, F_n(x_n)$ continue e copula C , allora per ogni $u \in I^n$ vale che

$$C(u_1, \dots, u_n) = H(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)),$$

In altre parole una funzione di distribuzione multivariata H , con marginali $F_1(x_1), \dots, F_n(x_n)$ continue, si lega alla corrispondente copula C se e solo se gli argomenti di H sono le funzioni inverse delle funzioni di distribuzione marginali univariate.

Il teorema di Sklar assicura che la probabilità congiunta cumulativa può essere scritta come funzione delle marginali cumulative e viceversa

$$H(x) = C(F_1(x_1), \dots, F_n(x_n)).$$

Questa possibilità (cioè di scrivere la probabilità congiunta cumulata in funzione delle marginali) viene chiamata **interpretazione probabilistica di base delle copule** ed permette alle copule di avere due importanti proprietà che saranno discusse di seguito:

- 1) Sia $X = (X_1, \dots, X_n)$ un vettore casuale e siano $F(X_1), \dots, F(X_n)$ le rispettive funzioni di distribuzione univariate. È noto che le variabili casuali del vettore casuale X sono indipendenti se e solo se $F(X) = F(X_1) \cdot \dots \cdot F(X_n)$

Definiamo **la copula prodotto** come

$$C^{\perp}(u) = u_1 \dots u_n.$$

con $u_1 = F(X_1), \dots, u_n = F(X_n)$

Il teorema di Sklar implica che:

Teorema 1.6. I vettori casuali in X sono indipendenti se hanno il prodotto della copula dentro la $\prod_{i=1}^n \text{Ran} F_i$.

2) Un'altra proprietà delle copule è di essere **invarianti** rispetto a trasformazioni strettamente monotone.

Teorema 1.7 (Schweizer e Wolff, 1976) Siano X_1, \dots, X_n variabili casuali continue con funzioni di distribuzione $F(X_1), \dots, F(X_n)$ e con copula C . Se $\alpha_1, \dots, \alpha_n$ sono trasformazioni crescenti in $\alpha_i: \text{Ran} F_i \rightarrow \mathfrak{R}^n$, le variabili casuali $\alpha_1(X_1), \dots, \alpha_n(X_n)$ con marginali

$$G_1 = F_1(\alpha_1^{-1}), \dots, G_n = F_n(\alpha_n^{-1})$$

e distribuzione congiunta G , hanno pure la copula C :

$$C_{\alpha_1(X_1), \dots, \alpha_n(X_n)}(u) = C_{X_1, \dots, X_n}(u)$$

per ogni $u = (u_1, \dots, u_n) \in I^n$ o equivalentemente

$$G(u) = C(G_1(u_1), \dots, G_n(u_n)).$$

Dimostrazione: Sia C la copula di $(X_1, \dots, X_n)^T$ e C_α quella di $(\alpha_1(X_1), \dots, \alpha_n(X_n))^T$. Poiché α_k è strettamente crescente per ogni k , $G_k(x) = \Pr\{\alpha_k(X_k) \leq x\} = \Pr\{X_k \leq \alpha_k^{-1}(x)\} = F_k(\alpha_k^{-1}(x))$ per ogni x in \mathbb{R} , quindi

$$\begin{aligned} C_\alpha(G_1(x_1), \dots, G_n(x_n)) &= \Pr\{\alpha_1(X_1) \leq x_1, \dots, \alpha_n(X_n) \leq x_n\} \\ &= \Pr\{X_1 \leq \alpha_1^{-1}(x_1), \dots, X_n \leq \alpha_n^{-1}(x_n)\} \\ &= C(F_1(\alpha_1^{-1}(x_1)), \dots, F_n(\alpha_n^{-1}(x_n))) \\ &= C(G_1(x_1), \dots, G_n(x_n)). \end{aligned}$$

Poiché X_1, \dots, X_n sono continue, $\text{Ran} G_1 = \dots = \text{Ran} G_n = [0, 1]$, Ne deriva che $C_\alpha = C$ in $[0, 1]^n$.

Le copule sono quindi invarianti rispetto a trasformazioni strettamente monotone.

Il teorema di Sklar consente di separare nelle funzioni di distribuzione multivariate le marginali univariate e la struttura di dipendenza multivariata, e la struttura di dipendenza può essere rappresentata dalla copula.

1.4 DISTRIBUZIONE CONDIZIONALE VIA COPULA

Il concetto di **distribuzione condizionale via copula** sarà utile quando si applicherà il metodo delle copule al calcolo del Value at Risk (vedi capitolo 4)

Siano $C_k(u_1, \dots, u_k) = C(u_1, \dots, u_k, 1, \dots, 1)$, con $k=2, \dots, n-1$, le marginali k -dimensionali di $C = C_n(u_1, \dots, u_n) = C(u_1, \dots, u_n)$.

C , inoltre, sia la funzione di distribuzione congiunta delle variabili casuali U_1, \dots, U_n . Allora la **distribuzione condizionale** di U_k dati i valori di U_1, \dots, U_{k-1} è data da:

$$\begin{aligned} C_k(u_k | u_1, \dots, u_{k-1}) &= \Pr\{U_k \leq u_k | U_1 = u_1, \dots, U_{k-1} = u_{k-1}\} \\ &= \frac{\frac{\partial^{k-1} C_k(u_1, \dots, u_k)}{\partial u_1 \dots \partial u_{k-1}}}{\frac{\partial^{k-1} C_{k-1}(u_1, \dots, u_{k-1})}{\partial u_1 \dots \partial u_{k-1}}} \end{aligned}$$

dato che il numeratore e il denominatore esistono (come conseguenza della caratterizzazione della copula come funzione di distribuzione congiunta) e che il denominatore è diverso da zero.

Questa proprietà è utile nel caso si vogliano generare variabili casuali (u_1, \dots, u_n) dalla copula C ed è usata in esempi di simulazione della dipendenza di una copula.

Osservazione 1.1: Conseguenze della proprietà precedente, per $n=2$, sono:

$$\Pr(X \leq x, Y > y) = F_1(x) - C(F_1(x), F_2(y))$$

$$\Pr(X > x, Y \leq y) = F_2(y) - C(F_1(x), F_2(y))$$

$$\Pr(X \leq x | Y \leq y) = C(F_1(x), F_2(y)) / F_2(y)$$

$$\Pr(X \leq x, Y > y) = F_1(x) - C(F_1(x), F_2(y)) / (1 - F_2(y))$$

$$\Pr(X \leq x, Y = y) = C_{|2}(F_1(x), F_2(y)) = \frac{\partial C(v, z)}{\partial z} \Big|_{v = F_1(x), z = F_2(y)}$$

1.5 COPULA DI SOPRAVVIVENZA E FUNZIONE DI SOPRAVVIVENZA CONGIUNTA

In questo paragrafo introduciamo la nozione di **copula di sopravvivenza** e la confrontiamo con **la funzione di sopravvivenza congiunta** che abbiamo definito precedentemente (definizione 1.6). Successivamente le applicheremo la prima alla valutazione delle funzioni di distribuzione delle funzioni **massimo** e **minimo** per $n > 2$.

Consideriamo la probabilità

$$\bar{F}(\mathbf{x}) = \Pr(X_1 > x_1, \dots, X_n > x_n) \quad (1.5)$$

e la denotiamo come probabilità di sopravvivenza congiunta o funzione di sopravvivenza di n componenti X_i . Le funzioni di sopravvivenza (probabilità di sopravvivenza marginali) sono invece definite come segue:

$$\bar{F}_i(X_i) = \Pr(X_i > x_i).$$

La copula che rappresenta la probabilità di sopravvivenza congiunta degli n componenti è definita **copula di sopravvivenza**.

Definizione 1.8 Il teorema di Sklar garantisce l'esistenza di una copula \bar{C} , unica su $\text{Ran } \bar{F}_1 \times \dots \times \text{Ran } \bar{F}_n$, tale che

$$\bar{F}(\mathbf{x}) = \bar{C}(\bar{F}_1(X_1), \dots, \bar{F}_n(X_n)). \quad (1.6)$$

E' consuetudine distinguere la copula sopravvivenza dalla funzione di sopravvivenza di n variabili uniformi. In questo caso $\bar{F}(\mathbf{x})$ è:

$$\bar{F}(\mathbf{x}) = \mathcal{C}(F_1(x_1), \dots, F_n(x_n)).$$

Ricordando che le trasformate integrali di probabilità sono uniformemente distribuite e denotando con U_i l'ennesima trasformata:

$$U_i = F_i(X_i),$$

si può scrivere $F_i(x_i) = u_i$ e $1 - \bar{F}_i(x_i) = u_i$.

Con questa notazione è facile trovare la relazione tra copula di sopravvivenza e la funzione di sopravvivenza per variabili uniformi. Poiché è possibile esprimere $\bar{F}(\mathbf{x})$ rispettivamente come

$$\bar{C}(\bar{F}_1(X_1), \dots, \bar{F}_n(X_n))$$

e

$$C(F_1(x_1), \dots, F_n(x_n))$$

abbiamo che

$$\bar{C}(\bar{F}_1(x_1), \dots, \bar{F}_n(x_n)) = C(F_1(x_1), \dots, F_n(x_n)) = C(1 - \bar{F}_1(x_1), \dots, 1 - \bar{F}_n(x_n)).$$

Definita questa relazione, siamo ora interessati a trovare la relazione tra la copula di sopravvivenza \bar{C} e la corrispondente copula C

Teorema 1.8 La copula di sopravvivenza \bar{C} può essere scritta in termini della corrispondente copula C come segue:

$$\bar{C}(u_1, \dots, u_n) = C(1 - u_1, \dots, 1 - u_n) = \sum_{i=0}^n \left[(-1)^i \sum_{w(u) \in Z(n-i, n, 1)} C(1 - w) \right]$$

dove $Z(n-i, n, 1)$ è l'insieme dei $\binom{n}{i}$ possibili vettori con $n-i$ componenti uguali a 1, i uguali ad u_i e

$$\mathbf{1-w} = (1-w_1, \dots, 1-w_n)$$

Simmetricamente, la copula C può essere scritta in termini della corrispondente copula di sopravvivenza \bar{C} come segue:

$$C(u_1, \dots, u_n) = \sum_{i=0}^n \left[(-1)^i \sum_{w(u) \in Z(n-i, n, 1)} \bar{C}(1 - w) \right]$$

Per concludere la discussione sulla copula sopravvivenza, presentiamo il seguente teorema, che permette di ottenere la copula sopravvivenza per alcune trasformazioni delle componenti di un vettore casuale.

Teorema 1.9 Siano X_1, \dots, X_n n variabili aleatorie continue con funzioni di distribuzione F_1, \dots, F_n e copula C . Consideriamo n funzioni di distribuzione continue G_1, \dots, G_n e denotiamo con T_j le variabili casuali $T_j = G_j^{-1}(1 - F_j(X_j))$. Di conseguenza, le marginali e la copula del vettore casuale (T_1, \dots, T_n) sono rispettivamente G_1, \dots, G_n e \bar{C} , ovvero la copula di sopravvivenza di C .

Dimostrazione :

$$\begin{aligned} \Pr(T_j \leq t_j) &= \Pr(G_j^{-1}(1 - F_j(X_j)) \leq t_j) \\ &= 1 - \Pr(1 - F_j(X_j) \geq G_j(t_j)) \\ &= 1 - \Pr(F_j(X_j) \leq 1 - G_j(t_j)) \\ &= 1 - F_j(F_j^{-1}(1 - G_j(t_j))) = G_j(t_j). \end{aligned}$$

Il fatto che la copula del vettore casuale (T_1, \dots, T_n) è \bar{C} dipende dal teorema di Schweizer e Wolff, perché G_j^{-1} e $1 - F_j$ sono rispettivamente funzioni crescenti e decrescenti.

Vediamo ora come le funzioni di distribuzione delle funzioni massimo e minimo di n variabili casuali, definite rispettivamente nel modo seguente:

$$m = \min(X_1, \dots, X_n)$$

e

$$M = \max(X_1, \dots, X_n)$$

siano esprimibili nei termini della loro copula o della loro copula di sopravvivenza

Per la funzione *massimo* abbiamo

$$F_M(a) = \Pr(M \leq a) = \Pr(X_1 \leq a, \dots, X_n \leq a) = F(a, \dots, a) = C(F_1(a), \dots, F_n(a)).$$

Per la funzione *minimo* :

$$\begin{aligned} F_m(a) &= \Pr(m \leq a) = 1 - \Pr(m > a) \\ &= 1 - \Pr(X_1 > a, \dots, X_n > a) \\ &= 1 - \mathcal{C}(F_1(a), \dots, F_n(a)) \\ &= 1 - \bar{C}(\bar{F}_1(a), \dots, \bar{F}_n(a)). \end{aligned}$$

1.6 DENSITA' E RAPPRESENTAZIONE CANONICA DI UNA COPULA MULTIDIMENSIONALE

Questo paragrafo introduce la nozione di **densità** e di **rappresentazione canonica** di una copula, insieme a quelle di componenti assolutamente continue e singolari.

Definizione 1.9 La densità $c(u_1, \dots, u_n)$ associata ad una copula $C(u_1, \dots, u_n)$ è:

$$c(u_1, \dots, u_n) = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n}$$

Essa esiste quasi ovunque in I^n .

La densità può essere usata per definire le componenti assolutamente continue e singolare di C , denotate come A_C e S_C , come segue:

$$A_C(u_1, \dots, u_n) = \int_0^{u_1} \int_0^{u_2} \dots \int_0^{u_n} \frac{\partial^n C(s_1, \dots, s_n)}{\partial s_1 \dots \partial s_n} ds_1 \dots ds_n$$

$$S_C(u_1, \dots, u_n) = C(u_1, \dots, u_n) - A_C(u_1, \dots, u_n)$$

Una copula per cui $C = A_C$ su I^n è chiamata **assolutamente continua**, mentre è chiamata **singolare** se $C = S_C$ su I^n . una copula è detta avere sia una **componente singolare** che una **assolutamente continua** se non appartiene né all'una, né all'altra classe.

Infine per concludere la discussione sul concetto di densità per una copula, diciamo che, per variabili casuali continue, la densità della copula è collegata alla densità della distribuzione F , denominata f , dalla **rappresentazione canonica**:

$$f(u_1, \dots, u_n) = c(F_1(x_1), \dots, F_n(x_n)) \cdot \prod_{j=1}^n f_j(x_j),$$

con

$$c(F_1(x_1), \dots, F_n(x_n)) = \frac{\partial^n (C(F_1(x_1), \dots, F_n(x_n)))}{\partial F_1(x_1) \dots \partial F_n(x_n)}, \quad (1.7)$$

mentre f_j sono le densità delle marginali

$$f_j(x_j) = \frac{dF_j(x_j)}{dx_j}$$

2. MISURE DI DIPENDENZA E FAMIGLIE DI COPULE

Le copule rappresentano una maniera “naturale” per studiare e misurare la dipendenza tra variabili casuali, dato che esse stesse sono funzioni di dipendenza. Inoltre, separando il comportamento marginale da quello di dipendenza, sono invarianti rispetto a trasformazioni monotone delle variabili casuali continue che le caratterizzano.

In questo capitolo saranno studiati alcuni concetti legati a quello di dipendenza e verranno presentate alcune misure ad essi associati.

Il primo paragrafo sarà, tuttavia, dedicato all'esposizione di alcune definizioni che verranno riprese più avanti.

2.1 LE COPULE COME FUNZIONI DI DIPENDENZA:

Le definizioni presentate di seguito riguardano le nozioni d'indipendenza e di comonotonicità. I teoremi e le definizioni esposte trattano un contesto bivariato.

2.1.1 Indipendenza

Siano X e Y due variabili casuali continue in I^2 . Poiché X e Y sono indipendenti se e solo se $F(x,y) = F_1(x)F_2(y)$, dal teorema di Sklar scaturisce che X e Y sono indipendenti se e solo se possiedono la copula prodotto C^\perp .

Teorema 2.1 Siano X e Y due variabili casuali continue in I^2 , tali da ammettere almeno una tra C^+ e C^- . Allora esistono due funzioni monotone $\alpha, \beta: \mathfrak{R} \rightarrow \mathfrak{R}$ e una variabile casuale Z tale che

$$(X,Y) = (\alpha(Z), \beta(Z)),$$

con α crescente e β decrescente nel caso di C^- ed entrambe crescenti nel caso di C^+ . Vale anche il contrario di questo risultato. Per la dimostrazione, si veda Embrechts (1999).

2.1.2 Comonotonicità

Due variabili casuali sono **comonotone** o **contromonotone**, e quindi perfettamente dipendenti, se e solo se le loro copule corrispondono rispettivamente ai limiti superiore ed inferiore di Frechét (teorema 1.3).

Definizione 2.1 L'insieme $A \subset \mathfrak{R}^2$ è detto essere comonotono se e solo se, per ogni $(x_1, y_1), (x_2, y_2)$ in A, vale

$$\begin{cases} x_1 \leq y_1 \\ x_2 \leq y_2 \end{cases} \text{ oppure } \begin{cases} x_1 \geq y_1 \\ x_2 \geq y_2 \end{cases}$$

Definizione 2.2 Un vettore casuale (X, Y) è comonotono o perfettamente dipendente positivo se e solo se esiste un insieme comonotono $A \subset \mathfrak{R}^2$ tale che

$$\Pr((X, Y) \in A) = 1$$

In altre parole, una coppia di variabili casuali comonotone X e Y è tale che a realizzazioni di X più grandi corrispondano realizzazioni di Y anch'esse più grandi.

Il seguente teorema è un'ulteriore caratterizzazione della proprietà comonotona

Teorema 2.2 Un vettore casuale (X, Y) , con funzioni di distribuzione F_1, F_2 e distribuzione congiunta $F(x, y)$ è comonotono se e solo se soddisfa i seguenti argomenti:

1) (X, Y) ha supporto comonotono

2) per ogni $(x, y) \in \mathfrak{R}^2$

$$F(x, y) = \min(F_1(x), F_2(y))$$

3) $C(v, z) = C^+(v, z)$

4) (X, Y) è distribuito come $(F_1^{-1}(U), F_2^{-1}(U))$, dove U è la variabile casuale uniforme standardizzata

5) (X, Y) è distribuito come $(F_1^{-1}(F_2(y)), F_2^{-1}(F_1(x)))$,

Dal teorema 2.2 deriva il seguente corollario:

Corollario 2.1 Se $F_1=F_2$, allora X e Y sono comonotone se e solo se sono uguali

Per introdurre la definizione di contromonotonicità o di dipendenza perfettamente negativa, si può procedere simmetricamente alla definizione data per comonotonicità

Definizione 2.3 L'insieme $A \subset \mathfrak{R}^2$ è detto essere contromonotono se e solo se , per ogni $(x_1,y_1), (x_2,y_2)$ in A, vale

$$\begin{cases} x_1 \leq y_1 \\ x_2 \geq y_2 \end{cases} \text{ oppure } \begin{cases} x_1 \geq y_1 \\ x_2 \leq y_2 \end{cases}$$

Definizione 2.4 Un vettore casuale (X,Y) è contromonotono o perfettamente dipendente negativo se e solo se esiste un insieme comonotono $A \subset \mathfrak{R}^2$ tale che

$$\Pr((X,Y) \in A) = 1$$

Per simmetria rispetto al concetto di comonotonicità vale il seguente teorema

Teorema 2.3 Un vettore casuale (X,Y) , con funzioni di distribuzione F_1,F_2 e distribuzione congiunta $F(x,y)$ è contromonotono se e solo se soddisfa le seguenti asserzioni:

- 1) (X,Y) ha supporto contromonotono
- 2) per ogni $(x,y) \in \mathfrak{R}^2$

$$F(x,y) = \max(F_1(x)+F_2(y)-1, 0)$$
- 3) $C(v,z) = C^-(v,z)$
- 4) (X,Y) è distribuito come $(F_1^{-1}(U), 1-F_2^{-1}(U))$, dove U è la variabile casuale uniforme standardizzata
- 5) (X,Y) è distribuito come $(F_1^{-1}(1-F_2(y)), F_2^{-1}(1-F_1(x)))$.

2.2 MISURE DI ASSOCIAZIONE

Questo paragrafo tratta le relazioni tra funzioni copula e le **misure di associazione** fra coppie di variabili casuali.

Generalmente, le variabili casuali X e Y sono dette essere **associate** quando non sono indipendenti, ovvero quando la funzione di ripartizione congiunta bivariata $F(x,y) \neq F_1(x)F_2(x)$.

Nel seguito saranno presentati alcuni concetti di associazione:

- Concordanza (diversa da dipendenza), correlazione lineare, dipendenza delle code, dipendenza positiva di quadrante

e alcune misure associate ad esse

- la T di Kendall, la ρ di Spearman, il coefficiente di correlazione lineare e gli indici di dipendenza delle code.

Tutte queste misure sono legate alle copule, poiché, nel mettere insieme una distribuzione congiunta con le sue marginali, la copula “cattura certi aspetti della relazione tra le variabili e di conseguenza i concetti di dipendenza sono proprietà della copula” (Nelsen, 1991).

2.2.1 CONCORDANZA

Il concetto di **concordanza** ha il compito di mettere in mostra il fatto che la probabilità di avere grandi (piccoli) valori delle due variabili è alta, mentre la probabilità di avere valori grandi per una variabile e bassa per l'altra (o viceversa) è piccola.

D'ora in poi assumeremo che X e Y siano variabili casuali continue.

Una misura di concordanza tra due variabili casuali può essere denotata con $M_{X,Y}$ o M_C ed ha le seguenti proprietà assiomatiche:

Definizione 2.5 $M_{X,Y}$ è una misura di concordanza tra due variabili casuali X e Y se e solo se:

- 1) È definita per ogni paio di variabili casuali (completezza)
- 2) È una misura relativa: $M_{X,Y} \in [-1,1]$
- 3) È simmetrica: $M_{X,Y} = M_{Y,X}$
- 4) Se X e Y sono indipendenti, allora $M_{X,Y} = 0$
- 5) $M_{-X,Y} = M_{Y,-X} = -M_{X,Y}$
- 6) Converge quando lo fa anche la copula: se $\{(X_n, Y_n)\}$ è una sequenza di variabili casuali continue con copula C_n , e

$$\lim_{n \rightarrow +\infty} C_n(v, z) = C(v, z) \text{ per ogni } (v, z) \in I^2$$

allora

$$\lim_{n \rightarrow +\infty} M_{X_n, Y_n} = M_{X, Y}$$

7) Rispetta l'ordine di concordanza : se $C_1 \prec C_2$, allora $M_{C_1} \leq M_{C_2}$

Questa definizione implica invarianza rispetto a trasformazioni crescenti e l'esistenza di limiti per M in corrispondenza di comonotonicità e contromonotonicità. I seguenti teoremi ne sono la conferma:

Teorema 2.4 Se α_i , $i=1,2$, sono funzioni crescenti su $\text{Ran } F_i$, allora $M_{X, Y} = M_{\alpha_1(X), \alpha_2(Y)}$

Teorema 2.5 Se X e Y sono comonotone , $M_{X, Y}=1$; se sono contromonotone , $M_{X, Y}=-1$

Scarsini (1984) ha provato che vale la seguente rappresentazione, dalla quale si possono ottenere diverse misure di associazione

Teorema 2.6 Data una funzione f limitata , debolmente monotona e dispari, con $\text{Dom } f = [-\frac{1}{2}, \frac{1}{2}]$, allora

$$k \int \int_{I^2} f(v - \frac{1}{2}) f(z - \frac{1}{2}) dC(v, z) \tag{2.1}$$

Dove $k^{-1} = \int_I f^2(u - \frac{1}{2}) du$ è una misura di concordanza

Dalla specificazione di f , si possono ottenere diverse misure di concordanza. Ad esempio, per $f(u) = u$ si ottiene la **ρ di Spearman** (che definiamo dopo); per $f(u) = \text{sgn}(u)$ otteniamo la β di Blomqvist, definita come

$$q = 4C(\frac{1}{2}, \frac{1}{2}) - 1 \tag{2.2}$$

Tuttavia altre misure di concordanza non possono essere ottenute dalla rappresentazione precedente: è il caso della **τ di Kendall** o del coefficiente γ di Gini, definito come

$$\gamma = 2 \int \int_{I^2} (|v+z-1| - |v-z|) dC(v, z) \quad (2.3)$$

Prima di passare in rassegna alcune delle più importanti misure di concordanza, forniamo altre due caratteristiche proprie di queste misure:

- Affinché una misura di concordanza sia pari a 0, l'indipendenza è una condizione sufficiente, ma non necessaria
- Esse Sono diverse dalle misure di dipendenza, poiché le seconde assumono il valore più basso quando le due variabili casuali sono indipendenti, non quando sono contromonotone

Considereremo ora due misure di concordanza (τ di Kendall e ρ di Spearman) che sono valide alternative al coefficiente di correlazione lineare per misurare la dipendenza in distribuzioni non ellittiche, per le quali il coefficiente di correlazione lineare è inappropriato e fuorviante.

2.2.2 τ DI KENDALL

Questa misura, riscoperta da Kendall (1938) . è stata introdotta da Fechner intorno al 1900.

Definizione 2.6 La **tau di Kendall** per le variabili casuali X e Y con copula C e denotata come τ o τ_C è:

$$\tau = 4 \int \int_{I^2} C(v, z) dC(v, z) - 1 \quad (2.4)$$

Si può dimostrare che essa misura la differenza tra la probabilità di concordanza e quella di discordanza tra 2 vettori casuali indipendenti, (X_1, Y_1) e (X_2, Y_2) , ognuno con la stessa funzione di distribuzione congiunta e copula C . I vettori sono detti essere concordanti se $X_1 > X_2$ mentre $Y_1 > Y_2$, e $X_1 < X_2$ mentre $Y_1 < Y_2$; discordanti nel caso opposto. Da questo segue che:

Teorema 2.7 Dati (X_1, Y_1) e (X_2, Y_2) i.i.d. con copula C

$$\tau = \Pr((X_1 - X_2)(Y_1 - Y_2) > 0) - \Pr((X_1 - X_2)(Y_1 - Y_2) < 0) \quad (2.5)$$

Per una dimostrazione del fatto che la tau di Kendall soddisfa gli assiomi i e vii per una misura di concordanza , si veda Scarsini (1984).

Questa misura di concordanza, inoltre, è compresa tra -1 e 1 ($-1 \leq \tau \leq 1$) e se consideriamo variabili casuali continue , vale che:

$$\tau = -1 \text{ se e solo se } C = C^-$$

$$\tau = 1 \text{ se e solo se } C = C^+$$

Vediamo ora il legame tra la tau di Kendall e la densità di una copula.

Quando la copula è assolutamente continua , il differenziale

$$dC = \frac{\partial^2 C}{\partial v \partial z} dv dz$$

può essere sostituito nella definizione di τ (equazione 2.4) .

Se invece C ha una componente assolutamente continua e una singolare, o solo singolare, vale il seguente teorema:

Teorema 2.8 La τ di Kendall può essere calcolata come

$$\tau = 1 - 4 \int \int_{I^2} \frac{\partial C(v, z)}{\partial v} \frac{\partial C(v, z)}{\partial z} dv dz \quad (2.6)$$

L'equivalenza fra la 2.4 e la 2.6 segue dal seguente lemma (Nelsen,1991)

Lemma 2.1 Se C è una copula

$$\int \int_{I^2} C(v, z) dC(v, z) + \int \int_{I^2} \frac{\partial C(v, z)}{\partial v} \frac{\partial C(v, z)}{\partial z} dv dz = \frac{1}{2}$$

Usando la 2.6 è semplice dimostrare che

Teorema 2.9 La tau di Kendall di una copula e quella della sua associata copula di sopravvivenza coincidono.

$$\tau_C = \tau_{\bar{C}}$$

Per stimare τ da un campione casuale di n coppie (X_i, Y_i) , $i = 1, \dots, n$, si definisce il seguente indicatore

$$A_{ij} \equiv \text{sgn}(X_i - X_j)(Y_i - Y_j)$$

dal quale si può notare che

$$E(A_{ij}) = (+1) \Pr((X_i - X_j)(Y_i - Y_j) > 0) + (-1) \Pr((X_i - X_j)(Y_i - Y_j) < 0) = \tau$$

Segue che uno stimatore non distorto e consistente per il coefficiente di Kendall è il cosiddetto τ campionario di Kendall:

$$\frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j>i} A_{ij} \quad (2.7)$$

2.2.3 ρ DI SPEARMAN

Definizione 2.7 La **rho di Spearman** per le variabili casuali X e Y con copula C ,denotata come ρ_S o ρ_{SC} , è

$$\rho_S = 12 \int \int_{I^2} C(v, z) dv dz - 3 = 12 \int \int_{I^2} v z dC(v, z) - 3 \quad (2.8)$$

Anche questa misura sfrutta le probabilità di concordanza e discordanza. Essa parte dal considerare tre coppie di vettori casuali i.i.d. $[(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)]$, con copula C.

La rho di Spearman è un multiplo della differenza tra la probabilità di concordanza e discordanza tra i vettori (X_1, Y_1) , (X_2, Y_3) , con X_2 e Y_3 variabili casuali tra loro indipendenti. Perciò , le probabilità di concordanza e discordanza sono misurate in un caso d'indipendenza.

Abbiamo, quindi, che:

Teorema 2.10 Dati $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$ i.i.d. con copula C , allora

$$\rho_S = 3 \Pr((X_1 - X_2)(Y_1 - Y_3) > 0) - \Pr((X_1 - X_2)(Y_1 - Y_3) < 0) \quad (2.9)$$

Sostituendo nella 2.8, si può scrivere

$$\rho_S = 12 \int \int_{I^2} [C(v, z) - v z] dv dz$$

La seguente osservazione permette di scrivere l'espressione per la rho di Spearman in una maniera più semplice.

Osservazione 2.1 Poiché le trasformate integrali $U_1=F_1(X)$, $U_2=F_2(x)$ sono uniformi standardizzate, con funzione di distribuzione congiunta C , l'integrale precedente può essere visto come $E[U_1U_2]$. Di conseguenza

$$\rho_s = 12E[U_1U_2] - 3 = \frac{E[U_1U_2] - \frac{1}{4}}{1/12}$$

Inoltre, dato che $\frac{1}{2}$ e $\frac{1}{12}$ sono la media e la varianza delle uniformi standard, ne segue che

$$\rho_s = \frac{\text{cov}(F_1(X), F_2(Y))}{\sqrt{\text{var}(F_1(X)) \text{var}(F_2(Y))}} \quad (2.10)$$

La rho di Spearman è, dunque, la correlazione di rango, cioè la correlazione delle trasformazioni integrali, di X e Y .

Inoltre ρ_s soddisfa la definizione di misura di concordanza e raggiunge i suoi limiti se e solo se X e Y sono rispettivamente variabili casuali contromonotone e comonotone:

$$\rho_s = -1 \text{ se e solo se } C=C^-$$

$$\rho_s = 1 \text{ se e solo se } C=C^+$$

Usando la 2.9 per definire la rho di Spearman,

Teorema 2.11 E' facile dimostrare, che la ρ_s di una copula e quella della copula sopravvivenza coincidono

$$\rho_{sc} = \rho_{s\bar{c}}$$

Per stimare ρ_s partendo da un campione casuale di n coppie (X_i, Y_i) , con $i=1, \dots, n$ ricordando che ρ_s è la correlazione di rango, secondo la 2.10, si procede considerando i ranghi delle variate campionarie:

$$R_i = \text{rango}(X_i), \quad S_i = \text{rango}(Y_i)$$

dove il rango deve essere fatto in ordine ascendente.

Così facendo, si ottiene il seguente risultato per ρ_s , che prende il nome di ρ_s campionario.

$$\frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2 \sum_{i=1}^n (S_i - \bar{S})^2}}$$

Tenendo in considerazione il fatto che i ranghi di n dati sono i primi n numeri interi, si semplifica l'espressione precedente e si ottiene

$$12 \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{n(n^2 - 1)}$$

o

$$1 - 6 \frac{\sum_{i=1}^n (R_i - S_i)^2}{n(n^2 - 1)} \quad (2.11)$$

Questo stimatore campionario è non distorto rispetto a quello della popolazione.

Presenteremo ora il concetto di correlazione lineare, che è la misura di dipendenza più usata nella pratica.

Tuttavia, poiché la correlazione lineare non è una misura di dipendenza basata sulla copula, spesso porta a risultati fuorvianti e di conseguenza non è considerata come una misura di dipendenza canonica.

2.2.4 CORRELAZIONE LINEARE

Per variabili casuali con momenti secondi finiti, la concordanza può essere spiegata anche dalla **covarianza**. Poiché la covarianza non è una misura normalizzata normalizzata, né relativa, è stato introdotto il **coefficiente di correlazione lineare**.

Definizione 2.8 Per variabili casuali non degeneri X e Y con momenti secondi finiti, il coefficiente di correlazione lineare ρ_{XY} è

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \text{var}(Y)}}$$

Teorema 2.12 Il coefficiente di correlazione lineare soddisfa gli assiomi da (i) a (v) e (vii) della definizione di misura di concordanza

Dimostrazione: (i) è sempre verificato, con l'esclusione delle variabili degeneri.

L'assioma (ii) segue dal fatto che $|\text{cov}(X, Y)| \leq \sqrt{\text{var}(X) \text{var}(Y)}$, mentre l'assioma (iii) dipende dalla simmetria della covarianza, $\text{cov}(X, Y) = \text{cov}(Y, X)$.

La proprietà (iv) segue dal fatto che l'indipendenza tra X e Y implica $\text{cov}(X; Y) = 0$

La (v) è una conseguenza del fatto che se $Y = aX + b$, con $a \in \mathfrak{R} \setminus \{0\}$, $b \in \mathfrak{R}$, allora $|\rho_{XY}| = 1$ e viceversa.

Per quanto riguarda la proprietà (vii), si deve ricorrere all'espressione di Hoeffding (1940) per la covarianza:

$$\text{Cov}(X, Y) = \iint_D (F(x, y) - F_1(x)F_2(y)) dx dy \quad (2.12)$$

con $D = \text{Dom } F_1 \times \text{Dom } F_2$.

Da questa espressione e dalla disuguaglianza di Fréchet segue che, se $C_1 \prec C_2$, e denotiamo con ρ_1 e ρ_2 le corrispondenti correlazioni lineari, allora da

$$\begin{aligned} & \iint_D (C_1(F_1(x)F_2(y)) - F_1(x)F_2(y)) dx dy \\ & \leq \iint_D (C_2(F_1(x)F_2(y)) - F_1(x)F_2(y)) dx dy \end{aligned}$$

si ottiene che $\rho_1 \leq \rho_2$.

Ciononostante, il coefficiente di correlazione non soddisfa l'assioma (vi) e quindi non è una misura di concordanza, tuttavia esso verifica le proprietà seguenti, che violano i teoremi 2.4 e 2.5.

Proprietà 2.1 ρ_{XY} è invariante sotto trasformazioni lineari crescenti, ma non sotto trasformazioni crescenti non lineari.

Dimostrazione: Per dimostrare il primo fatto, si consideri che

$$\rho_{aX+b, cY+d} = \text{sgn}(ac) \rho_{XY} \quad \text{per } a, c \in \mathfrak{R} \setminus \{0\}, b, d \in \mathfrak{R}$$

Trasformazioni lineari crescenti producono $\text{sgn}(ac) = 1$ e quindi $\rho_{aX+b, cY+d} = \rho_{XY}$.

Proprietà 2.2 ρ_{XY} è limitata

$$\rho_l \leq \rho_{XY} \leq \rho_u$$

dove i limiti ρ_l e ρ_u sono definiti come

$$\rho_l = \frac{\iint_D (C^-(F_1(x), F_2(y)) - F_1(x)F_2(y)) dx dy}{\sqrt{\int_{\text{Dom}F_1} (x - EX)^2 dF_1(x) \int_{\text{Dom}F_2} (y - EY)^2 dF_2(y)}} \quad (2.13)$$

$$\rho_u = \frac{\iint_D (C^+(F_1(x), F_2(y)) - F_1(x)F_2(y)) dx dy}{\sqrt{\int_{DomF_1} (x - EX)^2 dF_1(x) \int_{DomF_2} (y - EY)^2 dF_2(y)}} \quad (2.14)$$

Questi limiti sono raggiunti quando rispettivamente X e Y sono contromonotoni e comonotoni.

Dimostrazione: I limiti per ρ_{XY} possono essere ottenuti dall'espressione di Hoeffding (1940) per la covarianza (3.12), insieme con la disuguaglianza di Fréchet:

$$\begin{aligned} \iint_D (C^-(F_1(x), F_2(y)) - F_1(x)F_2(y)) dx dy &\leq \text{cov}(X, Y) \\ \text{cov}(X, Y) &\leq \iint_D (C^+(F_1(x), F_2(y)) - F_1(x)F_2(y)) dx dy \end{aligned}$$

Dividendo per la radice quadrata delle varianze, si ottengono rispettivamente i limiti C^- e C^+ .

Proprietà 2.3 ρ_{XY} per variabili casuali comonotone può essere diverso da 1 (-1)

Dimostrazione: si consideri un esempio di comonotonicità

Proprietà 2.4 $\rho_{XY}=0$ non implica indipendenza tra X e Y, a meno che una non sia Gaussiana.

Dimostrazione: Se X e Y sono Gaussiane, $\rho_{XY}=0$ implica $F(x,y)=F_1(x)F_2(y)$. Questo può essere controllato sostituendo $\rho_{XY}=0$ nell'espressione della funzione di distribuzione congiunta della Gaussiana: ad esempio con $\rho_{XY}=0$, la distribuzione congiunta di due variabili casuali normali standardizzate diventa:

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y \frac{1}{2\pi} \exp\left(\frac{-s^2 - t^2}{2}\right) ds dt$$

Poiché le marginali sono $F_i = \phi$, si dimostra che $F(x,y)=F_1(x)F_2(y)$.

Per altre funzioni di distribuzione questa proprietà non è dimostrabile.

Proprietà 2.5 Se $\rho_{XY}=0$, non significa che una variabile casuale non possa essere funzione dell'altra.

Dimostrazione: Per la dimostrazione, si consideri la seguente copula definita da Nelsen (1999)

$$\begin{cases} v & 0 \leq v \leq z/2 \leq 1/2 \\ z/2 & 0 \leq z/2 \leq v \leq 1 - z/2 \\ v + z - 1 & 1/2 \leq 1 - z/2 \leq v \leq 1 \end{cases}$$

Data questa copula, $\text{cov}(U_1, U_2) = 0$, ma

$$\Pr(U_2 = 1 - |2U_1 - 1|) = 1$$

Le variabili casuali U_1 e U_2 sono incorrelate, ma una è funzione dell'altra.

Come abbiamo visto il coefficiente di correlazione lineare non soddisfa la definizione di misura di concordanza: esso porta infatti a risultati fuorvianti per distribuzioni congiunte non ellittiche e talvolta anche per quelle ellittiche.

2.2.5 DIPENDENZA DI CODA

La **dipendenza di coda** bivariata guarda alla concordanza dei valori in "coda" alle variabili X e Y .

Geometricamente parlando si concentra l'attenzione sulle code, inferiore e superiore, della funzione di distribuzione congiunta.

E' un concetto rilevante per lo studio della dipendenza tra valori estremi.

La dipendenza di coda tra due variabili casuali è una proprietà della copula ed è perciò invariante per trasformazioni crescenti delle variabili casuali

Una volta definita la funzione di sopravvivenza congiunta per variabili uniformi, \mathcal{C} , abbiamo la seguente definizione:

Definizione 2.9 Sia il

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

finito. \mathcal{C} si dice avere **dipendenza di coda superiore** se e solo se $\mu_U \in (0, 1]$, e non avere dipendenza di coda superiore se e solo se $\mu_U = 0$. Analogamente, sia

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

finito. C è detto avere **dipendenza di coda inferiore** se e solo se $\mu_L \in (0,1]$, e non avere dipendenza di coda inferiore se e solo se $\mu_L=0$.

Per capire la corrispondenza tra questa definizione e l'intuizione di base, si ricordi che

$$\mathcal{C}(v, v) = \Pr(U_1 > v, U_2 > v)$$

Cosicché il rapporto $\mathcal{C}(v, v)/(1-v)$ è la seguente probabilità condizionale:

$$\frac{\mathcal{C}(v, v)}{1-v} = \Pr(U_1 > v | U_2 > v) = \Pr(U_2 > v | U_1 > v)$$

Perciò

$$\lambda_U = \lim_{v \rightarrow 1^-} \Pr(U_1 > v | U_2 > v) = \lim_{v \rightarrow 1^-} \Pr(U_2 > v | U_1 > v)$$

Il valore μ_U rappresenta il limite della probabilità condizionale che la funzione di distribuzione di X ecceda la soglia v , data dalla corrispondente funzione di Y , quando v tende a 1. Analogamente vale per μ_L .

Introduciamo ora i coefficienti di dipendenza di coda per la copula di sopravvivenza \bar{C} e vediamo come essi si relazionano ai coefficienti di dipendenza di coda della copula C calcolati prima:

$$\lim_{v \rightarrow 1^-} \frac{1 - 2v + \bar{C}(v, v)}{1 - v} = \bar{\lambda}_U$$

$$\lim_{v \rightarrow 0^+} \frac{\bar{C}(v, v)}{v} = \bar{\lambda}_L$$

Se questi limiti sono finiti, allora vale la proprietà seguente:

Teorema 2.13 Se \bar{C} è la copula di sopravvivenza associata a C , allora

$$\bar{\lambda}_U = \lambda_L$$

$$\bar{\lambda}_L = \lambda_U$$

dimostrazione:

$$\lim_{v \rightarrow 1^-} \frac{1 - 2v + \bar{C}(v, v)}{1 - v} = \lim_{v \rightarrow 1^-} \frac{C(1-v, 1-v)}{1 - v}$$

implica $\bar{\lambda}_U = \lambda_L$. Simmetricamente

$$\lim_{v \rightarrow 0^+} \frac{\bar{C}(v, v)}{v} = \lim_{v \rightarrow 0^+} \frac{2v - 1 + C(1 - v, 1 - v)}{v}$$

$$\lim_{v \rightarrow 1^-} \frac{1 - 2w + C(w, w)}{1 - w}$$

da $\bar{\lambda}_L = \lambda_U$.

2.2.6 DIPENDENZA POSITIVA DI QUADRANTE

Introduciamo ora un altro concetto di dipendenza: la **dipendenza positiva di quadrante** (PDQ), che può essere espressa in termini di copule nella maniera seguente.

Definizione 2.10 Le variabili casuali X e Y hanno dipendenza positiva di quadrante se e solo se

$$C(v, z) \geq vz$$

Per ogni $(v, z) \in I^2$.

In alternativa, usando l'ordine di concordanza tra copule, X e Y sono PDQ se e solo se la loro copula è più grande di quella prodotto.

$$C \succ C^\perp$$

In termini di funzioni di distribuzione, la PDQ può essere formalizzata come

$$F(x, y) \geq F_1(x)F_2(y) \quad \text{per ogni } (x, y) \in \mathfrak{R}^2$$

La probabilità congiunta in ogni punto non deve essere più piccola del prodotto delle 2 funzioni di distribuzione.

La PDQ implica la non negatività della tau di Kendall, della rho di Spearman e del coefficiente di correlazione lineare, dato che le variabili casuali indipendenti, per cui $C = C^\perp$, rendono questi coefficienti pari a zero e i coefficienti stessi rispettano l'ordine di concordanza.

Applicando la regola di Bayes, la disuguaglianza della PDQ può essere riscritta come:

$$\Pr(X \leq x | Y \leq y) \geq \Pr(X \leq x) \quad (2.15)$$

Perciò la condizione di PDQ di Lehmann (1966) può essere rinforzata ponendo la probabilità condizionale come funzione non crescente di y . Questo implica che,

ponendo ad esempio (X_t, Y_t) il vettore casuale bivariato dei rendimenti di due titoli X e Y al tempo t , la probabilità che il rendimento X_t prenda un valore basso non cresce col valore preso dall'altro rendimento Y_t . Questo corrisponde ad una particolare monotonicità nelle code.

Analogamente, si può dire che una variabile casuale X ha una **coda sinistra decrescente** in Y , denotata come $LTD(X|Y)$, se

$\Pr(X \leq x | Y \leq y)$ è una funzione non decrescente di y per ogni x .

Questo è equivalente alla condizione che, per ogni v in $[0,1]$, $C(v,z)/z$ è una funzione non decrescente in z , o:

$$\frac{\partial C(v, z)}{\partial z} \leq \frac{C(v, z)}{z} \text{ per quasi tutti gli } z. \quad (2.16)$$

Per $n > 2$ come generalizzazione del quadrante abbiamo l'ortante.

Definizione 2.11 (dipendenza positiva dell'ortante). Sia $X=(X_1, \dots, X_n)$ un vettore casuale n -dimensionale.

- 1) X è detto avere la più bassa dipendenza positiva dell'ortante se per ogni (x_1, \dots, x_n) in \mathfrak{R}^n ,

$$\Pr(X_1 \leq x_1, \dots, X_n \leq x_n) \geq \prod_{i=1}^n \Pr(X_i \leq x_i).$$

Ne segue che $C \succ C^\perp$.

- 2) X è detto avere la più alta dipendenza positiva dell'ortante se per ogni (x_1, \dots, x_n) in \mathfrak{R}^n ,

$$\Pr(X_1 > x_1, \dots, X_n > x_n) \geq \prod_{i=1}^n \Pr(X_i > x_i).$$

ne segue che $C \succ C^\perp$

- 3) X è detta avere dipendenza positiva dell'ortante se per ogni (x_1, \dots, x_n) in \mathfrak{R}^n , possiede le due proprietà precedenti.

Si possono definire anche le **dipendenze negative d'ortante** invertendo il senso delle precedenti disuguaglianze.

2.3 FAMIGLIE PARAMETRICHE DI COPULE n-DIMENSIONALI

In questa sezione presenteremo alcune famiglie o classi di copule. Ogni famiglia di copule è caratterizzato dall'aver un parametro o un vettore di parametri.

Per ogni famiglia, daremo la definizione e l'espressione della densità.

Le classi di copule che tratteremo si distinguono principalmente in **copule ellittiche** e **non ellittiche** e per le seconde in particolare vedremo di definire le cosiddette **copule archimediane**

La classe delle copule ellittiche forniscono molte distribuzioni multivariate, fra queste definiremo la copula Gaussiana e la copula T-student. Le copule ellittiche, in quanto simmetriche, hanno il difetto di non modellare bene i valori estremi delle distribuzioni multivariate.

Le copule archimediane, invece (che verranno definite meglio nella sezione 2.3.4) trattano meglio il problema dell'asimmetria delle code che si presenta spesso nelle applicazioni pratiche. Questa famiglia di copule hanno inoltre la caratteristica di dipendere da un parametro solamente. In seguito tratteremo, per questa classe di copule, la Frank copula e la Gumbel copula.

Da ultimo vedremo una famiglia di copule (Plackett) caratterizzata da proprietà desiderabili.

2.3.1 COPULA GAUSSIANA MULTIVARIATA (M.G.C)

Diamo la definizione di **copula Gaussiana**:

Definizione 2.12 Sia R una matrice simmetrica e definita positiva, con $\text{diag}(R) = (1, \dots, 1)^T$. Sia Φ_R la distribuzione normale multivariata standardizzata, con matrice di correlazione R .

La M.G.C. è definita come

$$C_R^{\text{Ga}}(\mathbf{u}) = \Phi_R(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))$$

dove Φ^{-1} è l'inversa della funzione di distribuzione normale univariata Φ .

Proprietà 2.6 La copula gaussiana genera la funzione la distribuzione congiunta normale, se le marginali sono normali standardizzate.

Usando la rappresentazione canonica , la densità della copula è facilmente determinabile:

$$\frac{1}{(2\pi)^{n/2} |R|^{1/2}} \cdot \exp\left(-\frac{1}{2} x^T R^{-1} x\right) = c_R^{Ga}(\phi(x_1), \dots, \phi(x_n)) \times \prod_{j=1}^n \left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} x_j^2\right) \right)$$

dove $|R|$ è il determinante della matrice di correlazione.

Di conseguenza:

$$c_R^{Ga}(\phi(x_1), \dots, \phi(x_n)) = \frac{\frac{1}{(2\pi)^{n/2} |R|^{1/2}} \cdot \exp\left(-\frac{1}{2} x^T R^{-1} x\right)}{\prod_{j=1}^n \left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} x_j^2\right) \right)}$$

Siano $u_j = \Phi(x_j)$, tale che $x_j = \Phi^{-1}(u_j)$. La densità, allora, può essere riscritta come segue

$$c_R^{Ga}(u_1, \dots, u_n) = \frac{1}{|R|^{1/2}} \cdot \exp\left(-\frac{1}{2} \zeta^T (R^{-1} - I) \zeta\right)$$

dove $\zeta = (\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))^T$

2.3.2 COPULA T-STUDENT MULTIVARIATA (MTC)

Un'altra appartenente alla classe delle copule ellittiche è la **copula t-student**.

Definizione 2.13 Sia R definita come sopra, e sia $t_{R,v}$ la distribuzione t-student standardizzata multivariata con matrice di correlazione R e v gradi di libertà:

$$t_{R,v}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} \frac{\Gamma\left(\frac{v+n}{2}\right) |R|^{-\frac{1}{2}}}{\Gamma\left(\frac{v}{2}\right) (v\pi)^{\frac{n}{2}}} \left(1 + \frac{1}{v} x^T R^{-1} x\right)^{-\frac{v+n}{2}} dx_1 \dots dx_n$$

La MTC è quindi definita come segue:

$$T_{R,v}(u_1, \dots, u_n) = t_{R,v}(t_v^{-1}(u_1), \dots, t_v^{-1}(u_n))$$

$$= \int_{-\infty}^{t_v^{-1}(u_1)} \int_{-\infty}^{t_v^{-1}(u_2)} \dots \int_{-\infty}^{t_v^{-1}(u_n)} \frac{\Gamma\left(\frac{v+n}{2}\right) |R|^{-\frac{1}{2}}}{\Gamma\left(\frac{v}{2}\right) (v\pi)^{\frac{n}{2}}} \left(1 + \frac{1}{v} x^T R^{-1} x\right)^{-\frac{v+n}{2}} dx_1 \dots dx_n$$

dove t_v^{-1} è l'inversa della funzione di distribuzione di una t-student con v gradi di libertà.

Ne scaturisce che la densità della copula per la MTC è:

$$c_{R,v}(u_1, \dots, u_n) = |R|^{-\frac{1}{2}} \frac{\Gamma\left(\frac{v+n}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \left(\frac{\Gamma\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v+1}{2}\right)}\right)^n \frac{\left(1 + \frac{1}{v} \zeta^T R^{-1} \zeta\right)^{-\frac{v+n}{2}}}{\prod_{j=1}^n \left(1 + \frac{\zeta_j^2}{v}\right)^{\frac{v+1}{2}}}$$

con $\zeta_j = t_v^{-1}(u_j)$

2.3.3 COPULA DI DISPERSIONE MULTIVARIATA (MDC)

Di seguito presentiamo una famiglia di copule con le caratteristiche della classe di copule ellittiche

Definizione 2.14 Sia $\mu = (\mu_1, \dots, \mu_n)$ un parametro di posizione, $\sigma^2 = (\sigma_1^2, \dots, \sigma_n^2)$ un parametro di dispersione e R una matrice di correlazione. Diciamo che $X \sim \text{MDC}(\mu, \sigma^2, R)$ se

$$f(y; \mu, \sigma^2, R) = \frac{1}{|R|^{1/2}} \cdot \exp\left(-\frac{1}{2} \zeta^T (R^{-1} - I) \zeta\right) \prod_{j=1}^n f_j(y_j; \mu_j, \sigma_j^2)$$

dove

$$\zeta_j = \Phi^{-1}(F_j(y_j; \mu_j, \sigma_j^2)) \quad \text{per } j=1, \dots, n \text{ e}$$

$$f_j(y_j, \mu_j, \sigma_j^2) = \frac{\partial F_j(y_j; \mu_j, \sigma_j^2)}{\partial y_j}$$

per ogni insieme di funzioni di distribuzione $F_j(y_j; \mu_j, \sigma_j^2)$.

Vediamo ora un esempio di applicazione della copula di dispersione multivariata:

Esempio 2.1 Costruiamo una MDC assumendo delle marginali con distribuzione Weibull.

In questo caso abbiamo

$$f(x) = \frac{\alpha x^{\alpha-1}}{\beta} \exp\left(-\frac{x^\alpha}{\beta}\right)$$

tramite la quale otteniamo la densità della MDC

$$f(x_1, \dots, x_n) = \frac{1}{|R|^{1/2}} \cdot \exp\left(-\frac{1}{2} \zeta^T (R^{-1} - I) \zeta\right) \prod_{j=1}^n \frac{\alpha_j x_j^{\alpha_j-1}}{\beta_j} \exp\left(-\frac{x_j^{\alpha_j}}{\beta_j}\right)$$

dove

$$\zeta_j = \Phi^{-1}\left(1 - \exp\left(-\frac{x_j^{\alpha_j}}{\beta_j}\right)\right).$$

2.3.4 COPULE DI ARCHIMEDE

Le copule **archimediane** hanno il vantaggio dal punto di vista computazionale di avere pochi parametri (di solito 1 o 2) per rappresentare la struttura di dipendenza, ma dall'altro lato ciò provoca una poca flessibilità della copula nell'adattarsi ai dati, soprattutto nei casi di grandi dimensioni.

La loro trattazione in questa sezione riguarderà prima alcune definizioni e proprietà e successivamente la presentazione di 2 particolari famiglie (**Gumbel**, **Franck**)

La famiglia delle copule di Archimede, per essere costruita ha bisogno di una **funzione generatrice** o **generatore** ϕ .

Questa funzione viene formulata dalle seguenti definizioni:

La funzione ϕ è definita da $I \rightarrow \mathfrak{R}^+$, essa è continua e decrescente.

Nelsen (1999) ha dimostrato che tale funzione deve essere convessa e tale che $\phi(1)=0$. Il generatore si definisce stretto quando $\phi(0)=+\infty$

In simboli

$$\phi(u): [0,1] \mapsto [0,+\infty]$$

Definiamo ora la funzione pseudo-inversa della funzione generatrice:

Definizione 2.15 La **pseudo-inversa** di ϕ è definita come segue:

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

Questa funzione pseudo-inversa è tale che, composta col generatore, dà l'identità, come fanno le inverse ordinarie con dominio e rango \mathfrak{R} .

$$\phi^{[-1]}(\phi(v)) = v \text{ per ogni } v \in I.$$

Teorema 2.14 (Kimberling(1974)) Sia data ϕ , la funzione generatrice, e la sua pseudo-inversa. La funzione $C: [0,1]^n \rightarrow [0,1]$ definita da

$$C(u_1, \dots, u_n) = \phi^{-1}(\phi(u_1) + \dots + \phi(u_n))$$

è una copula se e solo se ϕ^{-1} è completamente monotona in $[0, +\infty]$.

Come conseguenza del teorema precedente, possiamo dare la seguente definizione:

Definizione 2.16 Sia ϕ un generatore stretto, con ϕ^{-1} completamente monotona in $[0, +\infty]$.

Segue che la copula archimediana n-variata è la funzione

$$C(u_1, \dots, u_n) = \phi^{-1}(\phi(u_1) + \dots + \phi(u_n)).$$

Un'importante fonte di generatori per le n-copule archimediane consiste nelle inverse delle **trasformazioni di Laplace** di funzioni di distribuzione. La prova è data dal teorema 2.15.

Prima però diamo la definizione di trasformazione di Laplace,

Definizione 2.17 La trasformazione di Laplace di una variabile casuale γ , con funzione di distribuzione F_γ , è definita come:

$$\tau(s) = E_\gamma(e^{-s\gamma}) = \int_0^{+\infty} e^{-st} dF_\gamma(t) \quad (2.17)$$

Teorema 2.15 (Feller 1971) Una funzione ϕ in $[0, +\infty]$ è la trasformazione di Laplace di una funzione di distribuzione, se e solo se ϕ è completamente monotona e $\phi(0)=1$.

Le copule archimediane sono simmetriche; nel senso che , ad esempio per n=2;

$$C^A(v, z) = C^A(z, v) \quad \text{per ogni } (v, z) \in I^2$$

Inoltre, esse possiedono la proprietà associativa. Sempre come esempio, per n=3

$$C^A(C^A(v, z), u) = C^A(v, C^A(z, u)) \quad \text{per ogni } (v, z, u) \in I^3$$

Che implica entrambi i lati dell'uguaglianza precedente pari a $\phi^{-1}(\phi(v) + \phi(z) + \phi(u))$.

Definizione 2.18 Definiamo ora la formula generale della densità di una copula di Archimede nel caso bivariato, utile per i nostri calcoli successivi:

$$C^A(v, z) = \frac{-\phi''(C(v, z))\phi'(v)\phi'(z)}{(\phi'(C(v, z)))^3} \quad (2.18)$$

Allo stesso risultato, per la densità di una copula archimediana, si arriva usando la 1.7, con n=2

E' relativamente semplice a partire dalla definizione 2.17 generare copule di Archimede multivariate. Di seguito ne definiremo alcune:

copula di Gumbel

Il generatore è dato da $\phi(u) = (-\ln(u))^\alpha$, perciò $\phi^{-1}(t) = \exp(-t^{\frac{1}{\alpha}})$; essa è completamente monotona se $\alpha > 1$.

La n-copula di Gumbel è quindi

$$C(u_1, \dots, u_n) = \exp \left\{ - \left[\sum_{i=1}^n (-\ln u_i)^\alpha \right]^{\frac{1}{\alpha}} \right\} \quad \text{con } \alpha > 1.$$

Per n=2 la copula di Gumbel diventa:

$$C(u_1, u_2) = \exp \left\{ - \left[(-\ln u_1)^\alpha + (-\ln u_2)^\alpha \right]^{\frac{1}{\alpha}} \right\} \quad (2.19)$$

con u_1 e u_2 funzioni di ripartizione di due variabili casuali continue e pertanto $\in [0, 1]$.

La corrispondente densità bivariata, calcolata per mezzo della 2.18 è:

$$c(u_1, u_2) = \frac{\left(\exp\left\{-\left[-\log(u_1)\right]^\alpha\right\} \cdot (\alpha-1) \cdot \log(u_1) \cdot \log(u_2) \cdot \left\{\left[-\log(u_2)\right]^\alpha + \left[-\log(u_1)\right]^\alpha + \alpha - 1\right\} \right)}{\left(\left[\frac{1}{(u_1 \cdot u_2)} \right] \cdot \left[2 - \left(\frac{1}{\alpha} \right) \right] \cdot \left\{ \left[-\log(u_1)\right]^\alpha + \left[-\log(u_2)\right]^\alpha \right\} \right)} \quad (2.20)$$

copula di Frank

Il generatore è dato da $\varphi(u) = \ln\left(\frac{\exp(-\alpha u) - 1}{\exp(-\alpha) - 1}\right)$,

perciò

$$\varphi^{-1}(t) = -\frac{1}{\alpha} \ln(1 + e^t (e^{-\alpha} - 1));$$

essa è completamente monotona se $\alpha > 0$.

La n-copula di Frank è data quindi da

$$C(u_1, \dots, u_n) = -\frac{1}{\alpha} \ln \left\{ 1 + \frac{\prod_{i=1}^n (e^{-\alpha u_i} - 1)}{(e^{-\alpha} - 1)^{n-1}} \right\} \text{ con } \alpha > 0 \text{ quando } n \geq 3.$$

Per $n=2$ la copula di Frank diventa:

$$C(u_1, u_2) = -\frac{1}{\alpha} \ln \left(1 + \frac{(\exp(-\alpha u_1) - 1)(\exp(-\alpha u_2) - 1)}{(\exp(-\alpha) - 1)} \right) \quad (2.21)$$

con u_1 e u_2 funzioni di ripartizione di due variabili casuali continue e pertanto $\in [0,1]$.

La corrispondente densità, calcolata tramite la 2.18 è:

$$c(u_1, u_2) = \frac{(-\alpha \cdot [1 - \exp(-\alpha)])}{(2 \cdot \exp(-\alpha \cdot u_1 - \alpha \cdot u_2) \cdot \{1 - \exp(-\alpha) - [1 - \exp(-\alpha \cdot u_1)] \cdot [1 - \exp(-\alpha \cdot u_2)]\})} \quad (2.22)$$

2.3.5 PLACKETT COPULA

Presentiamo ora una famiglia di copule non archimediane (in quanto non derivate da alcuna funzione generatrice), ma molto utile per le sue proprietà: la famiglia delle **copule di Plackett** (1965).

Nel caso bivariato la copula di Plackett è definita nella maniera seguente: siano X e Y , due variabili casuali continue e $u=F(x)$, $v=F(y) \in [0,1]$ le rispettive funzioni di distribuzione. Allora

$$C(u, v) = \begin{cases} \frac{1}{2(\theta - 1)} \left[1 + (\theta - 1)(u + v) - \sqrt{[1 + (\theta - 1)(u + v)]^2 - 4uv\theta(\theta - 1)} \right] & \text{se } \theta \neq 1 \\ uv & \text{se } \theta = 1 \end{cases} \quad (2.23)$$

Con $\theta > 0$.

Questo tipo di copula è stata costruita in modo da avere correlazione positiva e quindi dipendenza positiva per $\theta > 1$ e ,viceversa, correlazione e dipendenza negativa per $\theta < 1$.

La conseguente funzione di densità ricavata mediante l'espressione 2.18 è:

$$c(u, v) = \theta \cdot [1 + (\theta - 1) + (u + v - 2 \cdot u \cdot v)] \cdot \left\{ [1 + (\theta - 1) \cdot (u + v)]^2 - 4 \cdot \theta \cdot (1 - \theta) \cdot u \cdot v \right\}^{\frac{3}{2}} \quad (2.24)$$

La famiglia di Plackett è simmetrica e positivamente ordinata; inoltre è assolutamente continua ed ammette i limiti di Frechét come casi estremi.

3. INFERENZA STATISTICA PER LE COPULE

Da un punto di vista statistico, come abbiamo visto dal teorema di Sklar , una funzione copula è un'espressione di un modello multivariato. E come per molti modelli multivariati le teorie inferenziali classiche non sono applicabili. L'unico metodo utilizzabile è quello della massima verosimiglianza asintotica.

Le metodologie inferenziali proposte per le copule sono lo stimatore di massima verosimiglianza (esatto) propriamente inteso e una sua versione , detta IFM method (inference for the margins).

Entrambi i metodi di stima richiedono un'ottimizzazione numerica della funzione obiettivo, in quanto un modello multivariato (qual'è una copula) richiede il calcolo di derivate miste.

Come prima cosa presentiamo i due metodi di stima e li mettiamo a confronto. Successivamente applichiamo il metodo IFM ad un contesto di serie storiche

3.1 METODO DI MASSIMA VEROSIMIGLIANZA ESATTA

Per introdurre i due metodi basati sulla massima verosimiglianza alle copule, riprendiamo la rappresentazione canonica per la densità di una copula

$$f(x_1, \dots, x_n) = c(F_1(x_1), \dots, F_n(x_n)) \cdot \prod_{j=1}^n f_j(x_j) \quad (3.1)$$

dove

$$c(F_1(x_1), \dots, F_n(x_n)) = \frac{\partial^n (C(F_1(x_1), \dots, F_n(x_n)))}{\partial F_1(x_1) \dots \partial F_n(x_n)} \quad (3.2)$$

è la n-sima derivata parziale mista della copula C, c è la densità della copula e f è la funzione di densità univariata.

Questa rappresentazione per la densità multivariata ci permette di scomporre in due parti il modello statistico per le copule:

1. identificazione delle distribuzioni marginali
2. definizione di un'appropriata funzione copula

Sia $\mathbf{x} = \{x_{1t}, \dots, x_{nt}\}_{t=1}^T$ la matrice dei dati campionari. L'espressione per la funzione di log-verosimiglianza diventa:

$$l(\theta) = \sum_{t=1}^T \ln c(F_1(x_{1t}), \dots, F_n(x_{nt})) + \sum_{t=1}^T \sum_{j=1}^n \ln f_j(x_{jt}) \quad (3.3)$$

dove θ è l'insieme di tutti i parametri del modello (sia delle marginali, che della copula).

Massimizzando l'espressione precedente rispetto all'insieme dei parametri , otteniamo lo stimatore di massima verosimiglianza:

$$\hat{\theta}_{MLE} = \max_{\theta \in \Theta} l(\theta)$$

Assumiamo che valgano le condizioni di regolarità della teoria della massima verosimiglianza asintotica per i modelli multivariati e per le marginali (Serfling,1980, Shao,1999) .

Sotto queste condizioni di regolarità lo stimatore di massima verosimiglianza esiste ed è consistente e asintoticamente efficiente. Inoltre è verificata la proprietà di normalità asintotica:

$$\sqrt{T}(\hat{\theta}_{MLE} - \theta_0) \rightarrow N(0, \mathfrak{I}^{-1}(\theta_0)) \quad (3.4)$$

con \mathfrak{I} l'usuale matrice d'informazione di Fisher e θ_0 il valore esatto del parametro.

La matrice di covarianza di $\hat{\theta}$ può essere stimata tramite l'inversa della matrice hessiana della funzione di verosimiglianza.

3.2 METODO IFM

Il metodo di massima verosimiglianza può essere computazionalmente difficile , specie nei casi di grandi dimensioni, poiché è necessario stimare congiuntamente i parametri delle marginali e quelli della struttura di dipendenza rappresentata dalla copula.

Joe e Xu (1996) hanno proposto di stimare quest'insieme di parametri in due passi:

1. il primo passo consiste nel stimare i parametri delle marginali , a partire dalle distribuzioni marginali univariate

$$\hat{\theta}_1 = \text{ArgMax}_{\theta_1} \sum_{t=1}^T \sum_{j=1}^n \ln f_j(x_{jt}; \theta_1) \quad (3.5)$$

2. in un secondo tempo, dato $\hat{\theta}_1$, si procede alla stima del parametro della copula θ_2 :

$$\hat{\theta}_2 = \text{ArgMax}_{\theta_2} \sum_{t=1}^T \ln c(F_1(x_{1t}), \dots, F_n(x_{nt}); \theta_2, \hat{\theta}_1) \quad (3.6)$$

Questo metodo è chiamato inferenza per le marginali o IFM. Lo stimatore IFM è definito come il vettore :

$$\theta_{IFM} = (\hat{\theta}_1, \hat{\theta}_2)'$$

Chiamiamo l'intera funzione di log-verosimiglianza , l_j la log-verosimiglianza per la j -esima marginale e l_c la log-verosimiglianza per la copula stessa.

Lo stimatore IFM è la soluzione di:

$$\left(\frac{\partial l_1}{\partial \theta_{11}}, \dots, \frac{\partial l_n}{\partial \theta_{1n}}, \frac{\partial l_c}{\partial \theta_2} \right) = 0'$$

mentre lo stimatore MLE deriva dalla soluzione di:

$$\left(\frac{\partial l}{\partial \theta_{11}}, \dots, \frac{\partial l}{\partial \theta_{1n}}, \frac{\partial l}{\partial \theta_2} \right) = 0'$$

L'equivalenza tra i due stimatori di solito non si mantiene. Tuttavia le stime IFM sono un buon punto di partenza per ottenere le stime MLE (massima verosimiglianza esatta).

Per quanto riguarda l'efficienza asintotica dello stimatore IFM, si deve considerare la matrice di covarianza asintotica.

Joe (1997) ha provato che sotto condizioni regolari, lo stimatore IFM verifica la proprietà di normalità asintotica che vale

$$\sqrt{T}(\hat{\theta}_{IFM} - \theta_0) \rightarrow N(0, \Sigma^{-1}(\theta_0)) \quad (3.7)$$

Con $\Sigma(\theta_0)$ matrice d'informazione di Godambe, definita come

$$\Sigma(\theta_0) = D V^{-1} (D)' \quad (3.8)$$

Con

$$D = E \left[\frac{\partial s(\theta)}{\partial \theta} \right] \text{ e } V = E[s(\theta)s(\theta)']$$

e con

$$s(\theta) = \left(\frac{\partial l_1}{\partial \theta_{11}}, \dots, \frac{\partial l_n}{\partial \theta_{1n}}, \frac{\partial l_c}{\partial \theta_2} \right)'$$

Abbiamo così diviso la log-verosimiglianza in due parti : l_1, \dots, l_n per ogni marginale e l_c per la copula.

$\Sigma(\theta_0)$ assume la forma dello stimatore di White per la matrice di varianza e covarianza dei parametri stimati col metodo dei minimi quadrati ordinari in caso di eteroschedasticità.

La stima di questa matrice di covarianza richiede il calcolo di molte derivate.

Joe (1997) ha definito il metodo IFM più efficiente del metodo di massima verosimiglianza esatta (MLE).

Di seguito faremo sempre riferimento al metodo IFM per l'inferenza sui parametri e per il calcolo degli standard error asintotici.

3.3 IL METODO IFM APPLICATO AD UN CASO PRATICO

Vediamo nel seguito come il metodo IFM per l'inferenza sui parametri delle copule si applica a dati di tipo finanziario ed in particolare alle serie storiche di indici azionari.

Per semplicità di calcolo si è scelto di lavorare con dati bivariati, cosa che ha permesso di valutare con relativa facilità le matrici di varianze e covarianze dei parametri delle copule. La matrice dei dati campionari pertanto è la seguente

$\mathfrak{X} = \{x_{1t}, x_{2t}\}_{t=1}^T$, dove 1 e 2 rappresentano le serie storiche prese in considerazione, mentre t si riferisce alla numerosità di entrambe le serie (nel nostro caso $T=550$).

Si sono stimati i parametri di 5 copule (Gaussiana, student-t, Plackett, Gumbel e Franck) col metodo IFM, e successivamente per ogni parametro è stata calcolata la matrice d'informazione di Godambe, la quale rappresenta la matrice di varianze asintotiche del parametro secondo la distribuzione asintotica descritta nella 3.7.

3.3.1 I DATI

I dati utilizzati riguardano i valori settimanali di due indici azionari (il Dow-Jones americano e il DAX30 tedesco), presi in considerazione a partire da ottobre 1994, fino ad aprile 2005, per un totale di 550 osservazioni.

Successivamente, a partire dai valori di entrambi gli indici, sono stati calcolati i rendimenti logaritmici.

Poiché siamo interessati alla distribuzione marginale delle serie dei rendimenti, "depurate" della dipendenza temporale, si è proceduto a filtrare le due serie, ipotizzando per esse un andamento descritto da un processo GARCH(1,1).

Notiamo dai grafici dell'ACF e del PACF per i residui e per i residui al quadrato delle 2 serie dei rendimenti che non ci sono "ritardi" significativamente diversi da 0 e che quindi il modello Garch(1,1) elimina l'effetto di dipendenza temporale.

Di seguito vengono riportati, per entrambi gli indici azionari, i grafici delle serie:

- dei valori degli indici,
- dei rendimenti degli indici:
- delle funzioni di autocorrelazione e autocorrelazione parziale per i residui del Garch(1,1)

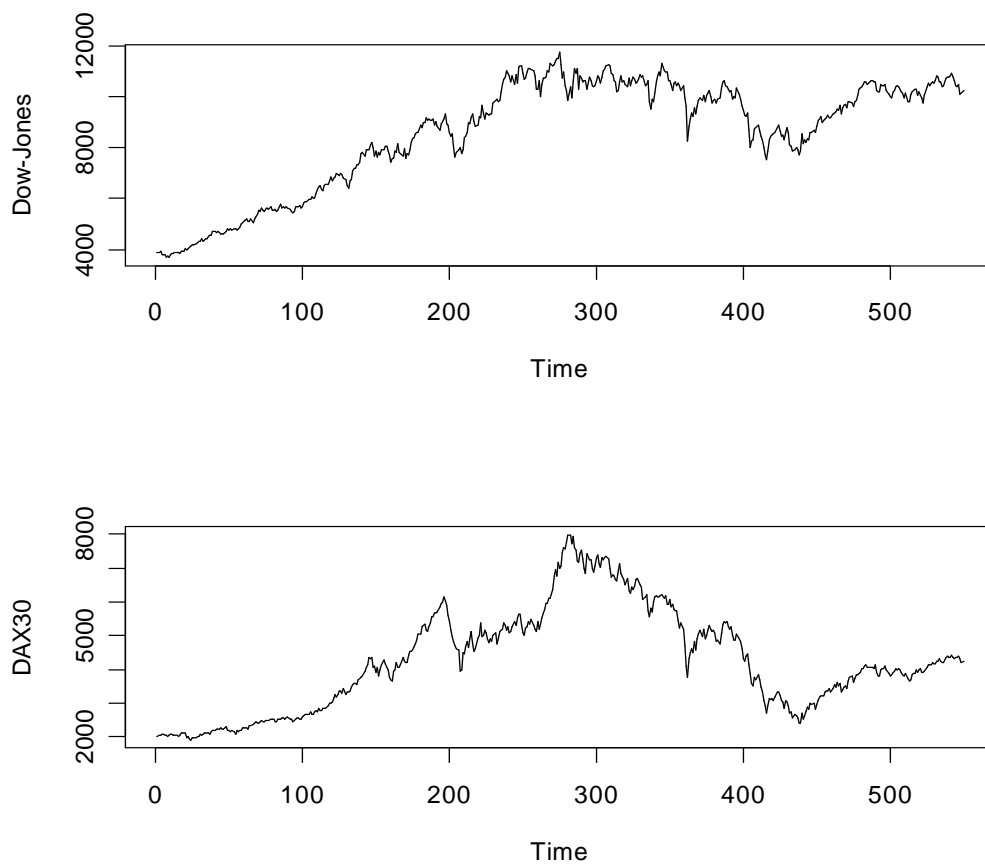


Figura 1 Serie storica dei valori degli indici azionari Dow-Jones e DAX30

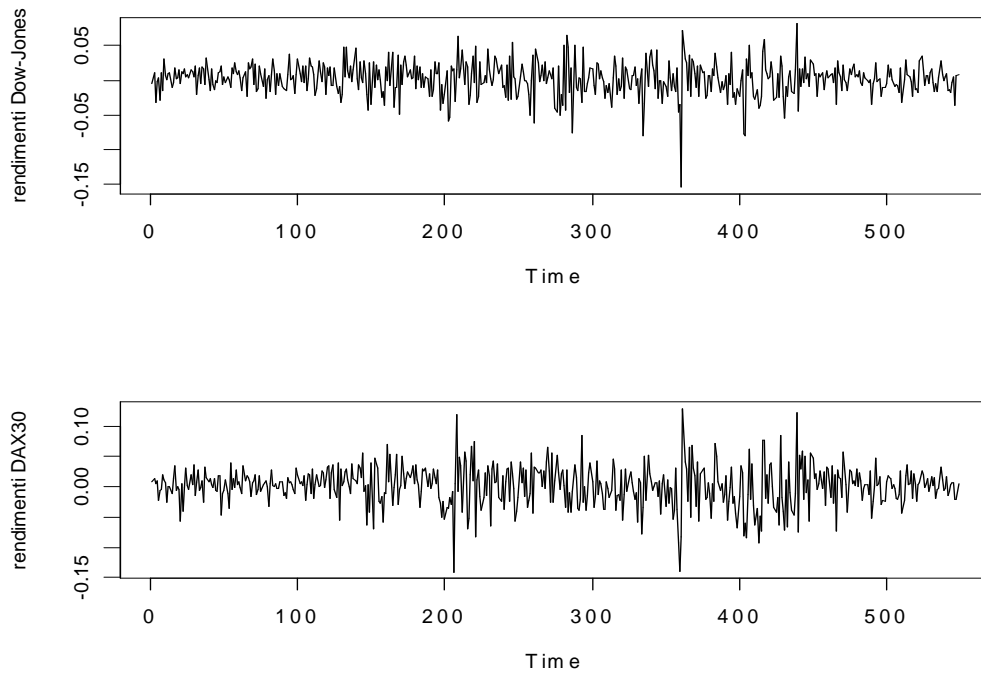


Figura 2 Serie storica dei rendimenti del Dow-Jones e del DAX30

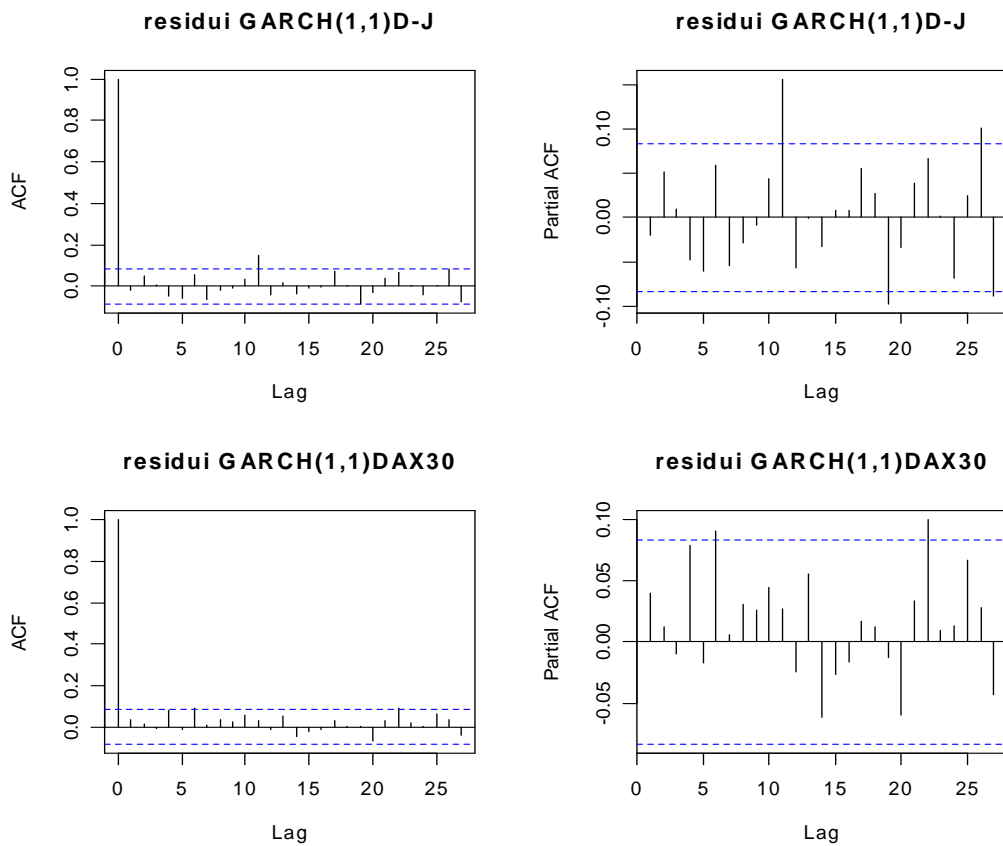


Figura 3 ACF e PACF dei residui standardizzati del Garch(1,1) per i rendimenti del Dow-Jones e del DAX30

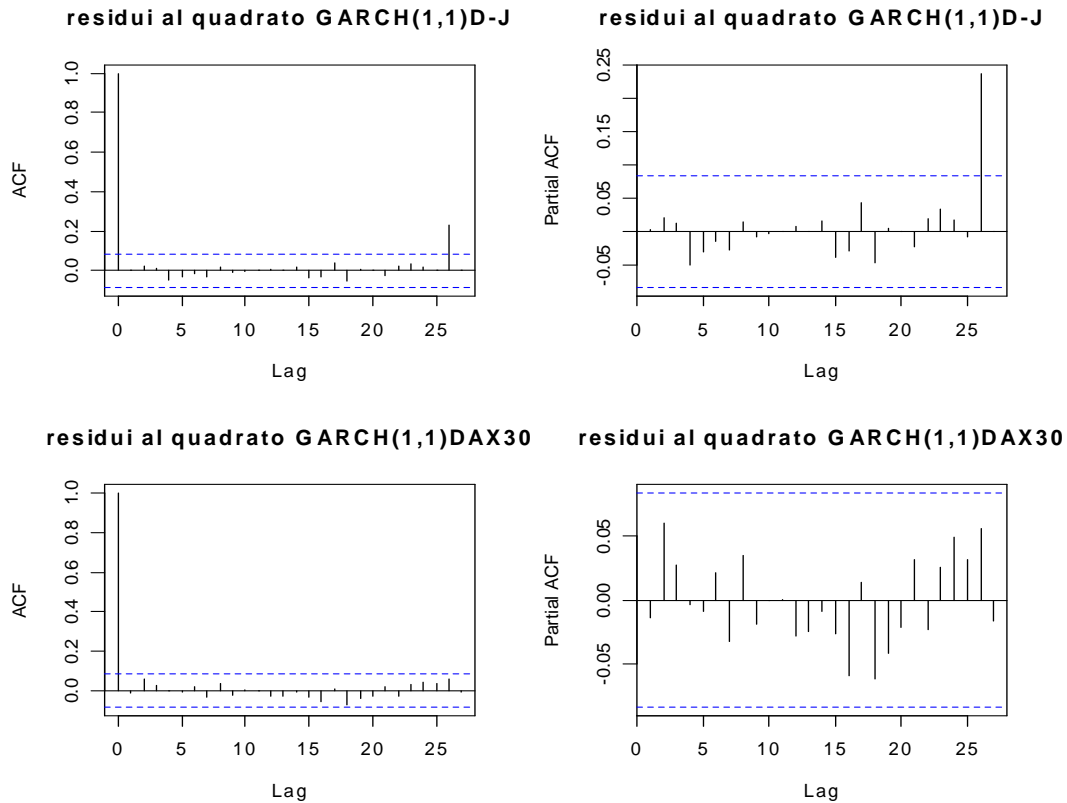


Figura 4 ACF e PACF dei residui standardizzati al quadrato del Garch(1,1) per i rendimenti del Dow-Jones e del DAX30

3.3.2 STIMA DEI PARAMETRI DELLE MARGINALI

Il metodo IFM ci permette, come abbiamo visto, di stimare i parametri delle marginali e delle copule in 2 momenti diversi.

Vediamo ora di stimare la distribuzione marginale dei residui standardizzati del Garch(1,1) per entrambe le serie dei rendimenti.

Ipotizziamo per le 2 marginali una distribuzione t-student con v gradi di libertà, con funzione di densità:

$$f(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)\Gamma\left(\frac{v}{2}\right)}{\sqrt{v\pi}} \cdot \left(1 + \frac{x^2}{v}\right)^{-\frac{(v+1)}{2}} \quad (3.9)$$

dove v è il parametro d'interesse.

Questa distribuzione ha la proprietà di tener conto della curtosi, presente nei residui di molte serie storiche

Col tendere di v verso l'infinito la funzione di densità della t-student si avvicina a quella di una normale di media 0 e varianza unitaria. Pertanto prima di passare

alla stima del parametro v , si è verificata l'ipotesi H_0 di normalità della distribuzione marginale, applicando, ai dati dei residui standardizzati del Garch(1,1), il test di Jarque-Bera.

La statistica test è $n \cdot \frac{asimmetria}{6} + n \cdot \frac{(curtosi - 3)^2}{24}$, che sotto H_0 si distribuisce come un χ^2 .

Tabella 1 Test di Jarque-Bera per i rendimenti dei due indici azionari

Indici Azionari	Dow-Jones	DAX30
Curtosi	5.06	3.89
Asimmetria	0.33	0.25
Statistica Jarque-Bera	127.58	40.89
p-value Jarque-Bera	0.0000	0.0000

Per entrambi gli indici si rifiuta l'ipotesi di normalità, pertanto possiamo applicare la 3.5 per il calcolo del parametro v della distribuzione t-student ipotizzata per le marginali. La tabella seguente riporta i parametri stimati per v .

Tabella 2 stima col metodo IFM dei gradi di libertà delle marginali t-student

STIMA IFM	Dow-Jones	DAX30
\hat{v}	18.94	34.07

I parametri, per avere senso, dovrebbero essere accompagnati dai relativi standard error asintotici, che permettono di fare verifiche d'ipotesi e test di significatività su di essi, approfittando della distribuzione normale asintotica dei parametri.

Nel metodo della massima verosimiglianza le varianze asintotiche dei parametri stimati corrispondono agli elementi sulla diagonale dell'inversa della matrice d'informazione (cambiata di segno). Il metodo IFM sebbene basato sulla verosimiglianza non fornisce stime efficienti per le varianze dei parametri semplicemente invertendo la matrice d'informazione. Pertanto, bisogna ricorrere ad uno stimatore di esse più complesso che sarà presentato successivamente. Inoltre saranno confrontati i risultati di questo metodo con quelli ottenuti dalla semplice inversione della matrice d'informazione.

3.3.3 STIMA DEI PARAMETRI DELLA COPULA

Prima di stimare i parametri delle copule per mezzo della 3.6, si devono trasformare le serie dei residui standardizzati del Garch(1,1) a seconda della distribuzione di probabilità che si è ipotizzata per essi.

In sostanza i valori dei residui standardizzati sono i quantili x_i della funzione di distribuzione trovata e in base ad essa vengono trasformati nelle rispettive probabilità y_i .

$$y_i = F(x_i)$$

Questa operazione deriva direttamente dalla definizione di copula. Infatti la copula è uno strumento mediante il quale si dà una struttura di dipendenza alle funzioni di distribuzione di variabili casuali.

Sono state considerate 5 copule: la gaussiana, la student-t, due copule di Archimede (Franck e Gumbel) e la copula di Plackett.

Di queste, 4 sono caratterizzate dall'aver solo un parametro da stimare, mentre la copula student ne ha 2.

La tabella che segue riporta le stime per i parametri. Come per i parametri delle marginali, anche per le copule gli standard error calcolati dall'inversione della matrice d'informazione non sono efficienti. Rinviamo al paragrafo successivo il calcolo delle devianze asintotiche mediante la matrice d'informazione di Godambe.

Tabella 3 stima col metodo IFM dei parametri delle copule

COPULA	$\hat{\rho}$	$\hat{\nu}$	$\hat{\alpha}$
Gaussian	0.31		
Student	0.32	8.07	
Franck			1.99
Gumbel			1.26
Plackett			2.79

3.3.4 STIMA DELLA MATRICE DI VARIANZA E COVARIANZA

Come abbiamo già detto la stima dei parametri col metodo IFM, impropriamente detto "verosimiglianza a due passi", non porta a risultati efficienti per il calcolo

degli standard error asintotici, se si considera l'inversa cambiata di segno della matrice d'informazione di Fisher.

Per arrivare a risultati che siano efficienti si utilizza l'espressione 3.8, dove V e D, definite sopra, non sono altro che, rispettivamente, la matrice attesa dei prodotti delle funzioni punteggio (derivate prime rispetto ai parametri delle funzioni di verosimiglianza delle marginali e della copula) e la matrice attesa delle derivate seconde e delle derivate seconde miste delle funzioni di verosimiglianza.

Nel nostro caso pratico consideriamo due distribuzioni marginali, che implicano 2 funzioni di verosimiglianza e quindi 2 parametri (uno per marginale) da stimare. Inoltre ogni copula comporta un'ulteriore funzione di verosimiglianza e 1 o 2 (nel caso della copula t-student) parametri in più da stimare.

Pertanto nel caso della copula student (viene presa come esempio poiché, avendo un parametro in più delle altre, è la più complessa) le matrici V e D diventano :

$$V = E \left(\begin{array}{c} \left(\begin{array}{c} \frac{\partial l_1}{\partial \theta_1} \\ \frac{\partial l_2}{\partial \theta_2} \\ \frac{\partial l_c}{\partial \theta_{11}} \\ \frac{\partial l_c}{\partial \theta_{12}} \end{array} \right) \left(\begin{array}{cccc} \frac{\partial l_1}{\partial \theta_1} & \frac{\partial l_2}{\partial \theta_2} & \frac{\partial l_c}{\partial \theta_{11}} & \frac{\partial l_c}{\partial \theta_{12}} \end{array} \right) \end{array} \right) \quad (3.9)$$

e

$$D = E \left(\begin{array}{cccc} \frac{\partial l_1}{\partial \theta_1 \partial \theta_1} & \frac{\partial l_1}{\partial \theta_1 \partial \theta_2} & \frac{\partial l_1}{\partial \theta_1 \partial \theta_{11}} & \frac{\partial l_1}{\partial \theta_1 \partial \theta_{12}} \\ \frac{\partial l_2}{\partial \theta_2 \partial \theta_1} & \frac{\partial l_2}{\partial \theta_2 \partial \theta_2} & \frac{\partial l_2}{\partial \theta_2 \partial \theta_{11}} & \frac{\partial l_2}{\partial \theta_2 \partial \theta_{12}} \\ \frac{\partial l_c}{\partial \theta_{11} \partial \theta_1} & \frac{\partial l_c}{\partial \theta_{11} \partial \theta_2} & \frac{\partial l_c}{\partial \theta_{11} \partial \theta_{11}} & \frac{\partial l_c}{\partial \theta_{11} \partial \theta_{12}} \\ \frac{\partial l_c}{\partial \theta_{12} \partial \theta_1} & \frac{\partial l_c}{\partial \theta_{12} \partial \theta_2} & \frac{\partial l_c}{\partial \theta_{12} \partial \theta_{11}} & \frac{\partial l_c}{\partial \theta_{12} \partial \theta_{12}} \end{array} \right) \quad (3.10)$$

dove con l_1, l_2, l_c denotiamo le funzioni di log-verosimiglianza, rispettivamente, della prima e della seconda marginale e della copula. Inoltre θ_1 è il parametro della prima marginale, θ_2 della seconda e θ_{11}, θ_{12} sono 2 i parametri della copula. Poiché l_1 ed l_2 dipendono solo dai propri parametri θ_1 e θ_2 , le derivate di esse rispetto a qualsiasi altra variabile è 0. l_c , invece, dipende da tutti i parametri.

Per questo motivo la matrice D può essere riscritta come segue:

$$D = E \begin{pmatrix} \frac{\partial l_1}{\partial \theta_1 \partial \theta_1} & 0 & 0 & 0 \\ 0 & \frac{\partial l_2}{\partial \theta_2 \partial \theta_2} & 0 & 0 \\ \frac{\partial l_c}{\partial \theta_1 \partial \theta_1} & \frac{\partial l_c}{\partial \theta_1 \partial \theta_2} & \frac{\partial l_c}{\partial \theta_1 \partial \theta_{11}} & \frac{\partial l_c}{\partial \theta_1 \partial \theta_{12}} \\ \frac{\partial l_c}{\partial \theta_2 \partial \theta_1} & \frac{\partial l_c}{\partial \theta_2 \partial \theta_2} & \frac{\partial l_c}{\partial \theta_2 \partial \theta_{11}} & \frac{\partial l_c}{\partial \theta_2 \partial \theta_{12}} \end{pmatrix} \quad (3.11)$$

I valori attesi valgono a livello di “popolazione” e possono essere sostituiti da stime a livello campionario.

Avremmo perciò:

$$\hat{V} = \frac{1}{T} \sum_{t=1}^T \begin{pmatrix} \left(\frac{\partial l_{1t}}{\partial \theta_1} \right) \\ \frac{\partial l_{2t}}{\partial \theta_2} \\ \frac{\partial l_{ct}}{\partial \theta_{11}} \\ \frac{\partial l_{ct}}{\partial \theta_{12}} \end{pmatrix} \begin{pmatrix} \frac{\partial l_{1t}}{\partial \theta_1} & \frac{\partial l_{2t}}{\partial \theta_2} & \frac{\partial l_{ct}}{\partial \theta_{11}} & \frac{\partial l_{ct}}{\partial \theta_{12}} \end{pmatrix} \quad (3.12)$$

e

$$\hat{D} = \frac{1}{T} \sum_{t=1}^T \begin{pmatrix} \frac{\partial l_{1t}}{\partial \theta_1 \partial \theta_1} & 0 & 0 & 0 \\ 0 & \frac{\partial l_{2t}}{\partial \theta_2 \partial \theta_2} & 0 & 0 \\ \frac{\partial l_{ct}}{\partial \theta_{11} \partial \theta_1} & \frac{\partial l_{ct}}{\partial \theta_{11} \partial \theta_2} & \frac{\partial l_{ct}}{\partial \theta_{11} \partial \theta_{11}} & \frac{\partial l_{ct}}{\partial \theta_{11} \partial \theta_{12}} \\ \frac{\partial l_{ct}}{\partial \theta_{12} \partial \theta_1} & \frac{\partial l_{ct}}{\partial \theta_{12} \partial \theta_2} & \frac{\partial l_{ct}}{\partial \theta_{12} \partial \theta_{11}} & \frac{\partial l_{ct}}{\partial \theta_{12} \partial \theta_{12}} \end{pmatrix} \quad (3.13)$$

dove il pedice t indica la log-verosimiglianza di ogni singola osservazione. Sappiamo infatti che la log-verosimiglianza è definita a partire dalla somma di quelle di tutte le osservazioni

Sostituendo le 3.12 e 3.13 nella 3.8, troviamo che

$$\hat{\Sigma}(\theta_0) = \hat{D} \hat{V}^{-1} (\hat{D})' \quad (3.14)$$

Che rappresenta lo stimatore della matrice di Godambe e di conseguenza della matrice di varianze e covarianze asintotiche dei parametri.

Possiamo, a questo punto, calcolare in maniera efficiente gli standard error dei parametri e confrontarli con quelli ottenuti semplicemente invertendo “l’hessiano” delle funzioni di verosimiglianza (quelle delle 2 marginali e quella della copula). I risultati sono raccolti nella tabelle seguenti:

Tabella 4 parametri stimati ed errori standard ottenuti dalla matrice di Godambe

COPULA	1° marginale	2° marginale	ρ	v	α
Gaussian	18.94 (10.87)	34.07 (31.57)	0.31 (0.046)		
Student	18.94 (10.18)	34.07 (32.06)	0.32 (0.045)	8.07 (4.16)	
Frank	18.94 (10.76)	34.07 (32.60)			1.99 (0.310)
Gumbel	18.94 (10.78)	34.07 (32.62)			1.26 (0.054)
Plackett	18.94 (10.70)	34.07 (32.20)			2.79 (0.424)

Tabella 5 parametri stimati ed errori standard ottenuti dalla matrice d'informazione

COPULA	1° marginale	2° marginale	ρ	v	A
Gaussian	18.94 (7.77)	34.07 (24.87)	0.31 (0.040)		
Student	18.94 (7.77)	34.07 (24.87)	0.32 (0.044)	8.07 (3.48)	
Frank	18.94 (7.77)	34.07 (24.87)			1.99 (0.287)
Gumbel	18.94 (7.77)	34.07 (24.87)			1.26 (0.045)
Plackett	18.94 (7.77)	34.07 (24.87)			2.79 (0.378)

Come si nota gli standard error calcolati nella tabella 1 a partire dalla matrice di Godambe sono più grandi rispetto al caso in cui sono calcolati con la semplice inversione della matrice d'informazione. Questo fatto sta ad indicare come il primo metodo tenga in maggiore considerazione l'eteroschedasticità.

Sfruttando la distribuzione asintotica dei parametri, si sono calcolati i valori dei t-test dei parametri ed i rispettivi p-value mediante l'utilizzo degli standard error

stimati dalla matrice di Godambe. I p-value sono calcolati secondo la distribuzione normale standardizzata.

Tabella 6 Test di significatività dei parametri e rispettivi p-value

COPULA	1° marginale	2° marginale	ρ	ν	A
Gaussian	1.74 (0.041)	1.08 (0.140)	6.74 (0.000)		
Student	1.86 (0.031)	1.06 (0.145)	7.11 (0.000)	1.94 (0.026)	
Frank	1.76 (0.039)	1.05 (0.147)			6.42 (0.000)
Gumbel	1.76 (0.039)	1.04 (0.149)			23.33 (0.000)
Plackett	1.77 (0.038)	1.06 (0.145)			6.59 (0.000)

Si noti come ad un livello di significatività del 5%, i parametri che non risultano significativi sono quelli riferiti ai gradi di libertà della distribuzione t-student delle marginali.

4. APPLICAZIONE DELLE COPULE AL CALCOLO DEL Value at Risk (VaR)

Vediamo ora una delle possibili applicazioni del metodo delle copule in finanza: la misurazione del rischio di mercato, o più precisamente il calcolo del Value at Risk

Scopo di questo studio è vedere se l'uso di copule diverse porta a risultati significativamente differenti nel calcolo del VaR.

In sostanza vogliamo capire come varia questa misura di performance al variare delle strutture di dipendenza e delle assunzioni su come si distribuiscono le marginali (ricordiamo che una copula si può scindere ed analizzare in due parti: le marginali e la struttura di dipendenza tra di esse).

Inoltre, si desidera studiare anche come varia il VaR al mutare del peso dato alle singole componenti del portafoglio.

Gli esiti della misura del VaR per copule diverse viene infine confrontato col risultato che si avrebbe calcolando il Value at Risk nella maniera usuale, cioè assumendo una distribuzione normale multivariata per le osservazioni.

4.1 DEFINIZIONE DI VaR

Come prima cosa diamo una definizione di Value at Risk.

Il VaR è una stima della massima perdita potenziale attesa su di un portafoglio di strumenti finanziari, in un arco temporale definito (holding period), con un certo grado di probabilità, a seguito del verificarsi di condizioni di mercato sfavorevoli.

Attraverso il VaR si definisce un indicatore sintetico che consente il confronto tra diverse tipologie di strumenti presenti all'interno di un portafoglio, sulla base di un'unica unità di misura che è la perdita potenziale. Il VaR esprime quale sarà la perdita massima che si potrà avere sulla posizione aperta al rischio nelle successive 24 ore o nei giorni seguenti, a fronte di cambiamenti nelle variabili di riferimento ovvero una perdita che può essere generata da un movimento avverso di un prezzo, un cambio, un tasso, in altre parole di un "fattore di rischio".

Il VaR indica quindi qual è l'ammontare di capitale adeguato a coprire le perdite inattese del portafoglio.

La metodologia VaR è in grado di calcolare il rischio complessivo del portafoglio tenendo conto delle dipendenze esistenti tra i diversi fattori di rischio e quindi dell'effetto diversificazione.

E' qui che le copule hanno un ruolo chiave, in quanto essendo ognuna costruita con una struttura di dipendenza diversa dalle altre, di conseguenza determinano i legami tra i fattori di rischio all'interno del portafoglio, portando a risultati a volte molto discordanti.

Il valore del VaR complessivo di un portafoglio è inferiore alla somma dei singoli VaR.

In termini matematici il Value at Risk può essere rappresentato come segue; sia "X" una variabile casuale che rappresenta la distribuzione dei profitti e delle perdite di un portafoglio di strumenti finanziari, il VaR a livello di confidenza $(1-\alpha)$ può essere definito come:

$$VaR(X) = -\inf \{x \mid P(X \leq x) > \alpha\}$$

La metodologia VaR considera la distribuzione di frequenza dei rendimenti che un portafoglio di strumenti finanziari può possedere in un determinato orizzonte temporale e focalizza l'attenzione solamente su un percentile, l' α -esimo, di tale distribuzione.

Tradizionalmente il Value at Risk viene calcolato assumendo una distribuzione normale multivariata dei fattori di rischio, cosicché la dipendenza tra di essi sia spiegabile semplicemente dalla correlazione.

Ad esempio, consideriamo un portafoglio di due indici azionari e indichiamo con X e Y le variabili casuali dei rispettivi rendimenti. Siano μ_X e μ_Y le medie e σ_X^2 , σ_Y^2 le varianze dei rendimenti dei due titoli.

Il portafoglio composto dai due titoli $R_p = w_X X + w_Y Y$, dove con w_X e w_Y indichiamo i pesi dei due titoli nel portafoglio, ha una distribuzione normale bivariata $N_2(\mu_p, \sigma_p^2)$.

$$\mu_p = w_X \mu_X + w_Y \mu_Y$$

$$\sigma_p^2 = w_X^2 \sigma_X^2 + w_Y^2 \sigma_Y^2 + \rho_{XY} \sigma_X \sigma_Y w_X w_Y.$$

$$w_X + w_Y = 1$$

La formula per il calcolo del VaR di questo portafoglio ad un livello di confidenza α del 5% diventa:

$$\begin{aligned} \text{VaR}_{95\%}(\mathbf{R}_P) &= -\mu_P - q_{0.05} \sigma_P \\ &= -\mu_P - q_{0.05} (w_X^2 \sigma_X^2 + w_Y^2 \sigma_Y^2 + \rho_{XY} \sigma_X \sigma_Y w_X w_Y) \\ &= -(w_X^2 \text{VaR}_{95\%}^2(X) + w_Y^2 \text{VaR}_{95\%}^2(Y) + w_X w_Y \rho_{12} \text{VaR}_{95\%}(X) \text{VaR}_{95\%}(Y))^{1/2} \end{aligned}$$

4.2 CALCOLO DEL VaR CON IL METODO DELLE COPULE

Abbiamo visto come usualmente viene misurato il VaR per un portafoglio di due titoli. Cerchiamo ora una maniera per applicare il metodo delle copule al calcolo di questo indicatore.

Dato un livello di confidenza θ , il VaR è la soglia sotto la quale i rendimenti cadono solamente con probabilità θ .

Se indichiamo con Z il rendimento del portafoglio di un dato orizzonte di tempo T , il VaR è il valore per cui:

$$\Pr(Z \leq \text{VaR}_Z) = \theta.$$

Equivalentemente, usando la funzione di distribuzione di Z , $\xi(z)$, il Value at Risk può essere definito come la soluzione z^* dell'equazione $\xi(z^*) = \theta$.

La funzione di distribuzione ξ può essere scritta in termini di copule nella maniera seguente:

consideriamo un portafoglio di due titoli e siano X e Y i rendimenti in un uguale orizzonte di tempo T . Sia $\beta \in (0,1)$ il peso di allocazione. Il rendimento del portafoglio è $Z = \beta X + (1 - \beta)Y$, con funzione di distribuzione

$$\begin{aligned} \xi(z) &= \Pr(Z \leq z) = \Pr(\beta X + (1 - \beta)Y \leq z) \\ &= \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{\frac{1}{\beta}z - \frac{1-\beta}{\beta}y} c(F_1(x), F_2(y)) f_1(x) dx \right\} f_2(y) dy \quad (4.1) \\ &= \int_{-\infty}^{+\infty} \Pr\left(X \leq \frac{1}{\beta}z - \frac{1-\beta}{\beta}y, Y = y\right) f_2(y) dy \\ &= \int_{-\infty}^{+\infty} C_{12}\left(F_1\left(\frac{1}{\beta}z - \frac{1-\beta}{\beta}y\right), F_2(y)\right) f_2(y) dy \end{aligned}$$

dove è stata usata la probabilità condizionale via copula (vedi paragrafo 1.4)

Nel calcolo del VaR col metodo delle copule non entra la serie storica dei rendimenti azionari, i quali entrano nella funzione del VaR solo tramite la loro distribuzione.

4.3 ESEMPIO PRATICO DI CALCOLO DEL VaR CON LE COPULE

Mettendo in pratica l'espressione precedente, si è calcolato il Value at Risk di 4 portafogli, composti ognuno da coppie di indici azionari.

In particolare si è misurato il VaR al livello di significatività dell'1% e del 5% per i seguenti abbinamenti di indici: Mibtel-CaC40, Mibtel-DAX30, Mibtel-Dow-Jones e Mibtel-FTSE100.

Il VaR è stato calcolato facendo uso della copula Gaussiana, di quella Student-t, di due copule archimediane (Frank e Gumbel) e della copula di Pluckett.

Per quanto riguarda le marginali, si è considerato un primo caso in cui si assume che esse siano distribuite come student-t con v gradi di libertà. Successivamente le marginali sono state valutate come normali di media zero e varianza unitaria

I parametri utilizzati:

- per la copule vere e proprie (cioè le strutture di dipendenza), sono quelli stimati col metodo IFM per ogni coppia di indice azionario.
- per le marginali sono i gradi di libertà della distribuzione t-student, sempre stimati col metodo IFM

Si è proceduto inoltre a calcolare il VaR per ogni portafoglio nella maniera tradizionale, ovvero supponendo una distribuzione congiunta dei rendimenti normale bivariata.

I dati usati si riferiscono alle serie storiche settimanali dei rendimenti dei 5 indici menzionati sopra, a partire da ottobre 1994 fino all'ultimo dato a disposizione che si riferiva ad Aprile 2005, per un totale di 550 osservazioni.

Per ogni portafoglio $Z_i = \beta X + (1 - \beta)Y_i$, con $i=1, \dots, 4$ e X e Y_i che rappresentano le variabili casuali dei rendimenti degli indici azionari, si è fatto variare β , cioè il peso di ogni indice dentro al portafoglio Z_i , per studiare come cambia il VaR rispetto alla variabile β .

L'analisi del VaR per ognuno dei 4 portafogli, così come è stata impostata, ha il duplice scopo di:

- esaminare come la struttura di dipendenza tra le serie dei rendimenti influisce sulla misura del VaR
- analizzare l'andamento della misura del rischio al variare della composizione dei pesi all'interno del portafoglio.

L'indagine viene replicata per tutti i portafogli per vedere se i risultati si ripetono o meno in tutte e 4 le circostanze.

Di seguito riportiamo le tabelle che raccolgono i risultati del calcolo del VaR per i portafogli e per le diverse combinazioni di copule e marginali.

Tabella 7 VaR calcolati con livello di confidenza del 5% e vari pesi β per il portafoglio Mibtel-CaC40. Si considerano marginali distribuite come student-t. (tra parentesi gli standard error)

$v_{\text{mibtel}}=18.92$ (10.62) , $v_{\text{CaC40}}=130.72$ (122.69)							
COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho=0.30$ (0.044)	-1.51	-1.29	-1.21	-1.33	-1.46	-1.58
Student-t	$\rho =0.30$ (0.046) $v=8.11$ (4.014)	-0.51	-0.47	-0.45	-0.47	-0.52	-0.55
Gumbel	$p=1.22$ (0.054)	-1.54	-1.37	-1.32	-1.40	-1.60	-1.73
Franck	$p=1.95$ (0.321)	-1.55	-1.40	-1.36	-1.43	-1.61	-1.73
Plackett	$p=2.84$ (0.470)	-1.55	-1.41	-1.38	-1.45	-1.62	-1.73
Tradizionale	$\rho=0.27$	-2.38	-1.81	-1.64	-1.87	-2.51	-2.98

Tabella 8 VaR calcolati con livello di confidenza del 5% e vari pesi β per il portafoglio Mibtel-CaC40. Si considerano marginali distribuite come $N(0,1)$. (tra parentesi gli standard error)

COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.30$ (0.044)	-1.50	-1.27	-1.18	-1.27	-1.50	-1.66
Gumbel	$\rho =1.22$ (0.054)	-1.54	-1.41	-1.36	-1.41	-1.55	-1.64
Franck	$p=1.95$ (0.321)	-1.53	-1.38	-1.32	-1.38	-1.54	-1.64
Plackett	$p=2.84$ (0.470)	-1.52	-1.34	-1.28	-1.34	-1.52	-1.64
Tradizionale	$\rho=0.27$	-2.38	-1.81	-1.64	-1.87	-2.51	-2.98

Tabella 9 VaR calcolati con livello di confidenza del 1% e vari pesi β per il portafoglio Mibtel-CaC40. Si considerano marginali distribuite come student-t. (tra parentesi gli standard error)

$v_{mibtel}=18.92$ (10.62) , $v_{CaC40}=130.72$ (122.69)							
COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.30$ (0.044)	-2.15	-1.82	-1.73	-1.92	-2.31	-2.55
Student-t	$\rho =0.30$ (0.046) $v=8.11$ (4.014)	-1.17	-1.05	-1.01	-1.05	-1.17	-1.25
Gumbel	$p=1.22$ (0.054)	-2.26	-2.15	-2.14	-2.24	-2.42	-2.54
Franck	$p=1.95$ (0.321)	-2.19	-1.95	-1.89	-2.04	-2.35	-2.54
Plackett	$p=2.84$ (0.470)	-2.17	-1.91	-1.85	-2.00	-2.33	-2.54
Tradizionale	$\rho=0.27$	-4.63	-3.54	-3.20	-3.64	-4.87	-5.77

Tabella 10 VaR calcolati con livello di confidenza del 1% e vari pesi β per il portafoglio Mibtel-CaC40. Si considerano marginali distribuite come $N(0,1)$. (tra parentesi gli standard error)

COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.30$ (0.044)	-2.12	-1.78	-1.66	-1.78	-2.12	-2.34
Gumbel	$p =1.22$ (0.054)	-2.22	-2.10	-2.06	-2.10	-2.23	-2.33
Franck	$p=1.95$ (0.321)	-2.15	-1.90	-1.81	-1.90	-2.16	-2.33
Plackett	$p=2.84$ (0.470)	-2.14	-1.87	-1.77	-1.87	-2.14	-2.33
Tradizionale	$\rho=0.27$	-4.63	-3.54	-3.20	-3.64	-4.87	-5.77

Tabella 11 VaR calcolati con livello di confidenza del 5% e vari pesi β per il portafoglio Mibtel-DAX30. Si considerano marginali distribuite come student-t. (tra parentesi gli standard error)

$v_{\text{mibtel}}=18.92 (10.62) , v_{\text{DAX30}}=34.06 (32.71)$							
COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.34 (0.044)$	-1.55	-1.32	-1.23	-1.34	-1.59	-1.75
Student-t	$\rho =0.36 (0.045)$ $v=7.76 (3.831)$	-0.48	-0.45	-0.43	-0.45	-0.52	-0.55
Gumbel	$p=1.29 (0.061)$	-1.57	-1.41	-1.36	-1.43	-1.61	-1.73
Franck	$p=2.41 (0.335)$	-1.59	-1.45	-1.40	-1.47	-1.62	-1.73
Plackett	$p=3.49 (0.528)$	-1.59	-1.46	-1.42	-1.48	-1.63	-1.73
Tradizionale	$\rho=0.31$	-2.43	-1.86	-1.68	-1.90	-2.52	-2.98

Tabella 12 VaR calcolati con livello di confidenza del 5% e vari pesi β per il portafoglio Mibtel-DAX30. Si considerano marginali distribuite come $N(0,1)$. (tra parentesi gli standard error)

COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.34 (0.044)$	-1.51	-1.27	-1.18	-1.27	-1.51	-1.67
Gumbel	$p=1.29 (0.061)$	-1.55	-1.43	-1.39	-1.43	-1.55	-1.64
Franck	$p=2.41 (0.335)$	-1.54	-1.40	-1.35	-1.40	-1.54	-1.64
Plackett	$p=3.49 (0.528)$	-1.53	-1.37	-1.31	-1.37	-1.53	-1.64
Tradizionale	$\rho=0.31$	-2.43	-1.86	-1.68	-1.90	-2.52	-2.98

Tabella 13 VaR calcolati con livello di confidenza del 1% e vari pesi β per il portafoglio Mibtel-DAX30. Si considerano marginali distribuite come student-t. (tra parentesi gli standard error)

$v_{\text{mibtel}}=18.92$ (10.62) , $v_{\text{DAX30}}=34.06$ (32.71)							
COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.34$ (0.044)	-2.23	-1.88	-1.76	-1.93	-2.32	-2.56
Student-t	$\rho =0.36$ (0.045) $v=7.76$ (3.831)	-1.17	-1.06	-1.03	-1.06	-1.17	-1.25
Gumbel	$p=1.29$ (0.061)	-2.37	-2.27	-2.26	-2.32	-2.45	-2.54
Franck	$p=2.41$ (0.335)	-2.27	-2.03	-1.95	-2.08	-2.36	-2.54
Plackett	$p=3.49$ (0.528)	-2.26	-2.00	-1.92	-2.05	-2.35	-2.54
Tradizionale	$\rho=0.31$	-4.71	-3.60	-3.27	-3.70	-4.89	-5.78

Tabella 14 VaR calcolati con livello di confidenza del 1% e vari pesi β per il portafoglio Mibtel-DAX30. Si considerano marginali distribuite come $N(0,1)$. (tra parentesi gli standard error)

COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.34$ (0.044)	-2.12	-1.79	-1.66	-1.79	-2.12	-2.34
Gumbel	$p=1.29$ (0.061)	-2.27	-2.17	-2.15	-2.17	-2.26	-2.33
Franck	$p=2.41$ (0.335)	-2.17	-1.93	-1.84	-1.93	-2.17	-2.33
Plackett	$p=3.49$ (0.528)	-2.15	-1.90	-1.81	-1.90	-2.15	-2.33
Tradizionale	$\rho=0.31$	-4.71	-3.60	-3.27	-3.70	-4.89	-5.78

Tabella 15 VaR calcolati con livello di confidenza del 5% e vari pesi β per il portafoglio Mibtel-Dow-Jones. Si considerano marginali distribuite come student-t. (tra parentesi gli standard error)

$V_{\text{mibtel}}=18.92$ (10.62) , $V_{\text{Dow-Jones}}=34.06$ (32.71)							
COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho = 0.59$ (0.029)	-1.64	-1.39	-1.30	-1.39	-1.64	-1.80
Student-t	$\rho = 0.58$ (0.029) $v=8.87$ (4.259)	-0.52	-0.49	-0.48	-0.49	-0.52	-0.54
Gumbel	$p=1.62$ (0.077)	-1.70	-1.69	-1.68	-1.69	-1.71	-1.73
Franck	$p=4.14$ (0.320)	-1.65	-1.55	-1.51	-1.55	-1.65	-1.73
Plackett	$p=6.31$ (0.718)	-1.63	-1.51	-1.46	-1.51	-1.64	-1.73
Tradizionale	$\rho=0.56$	-2.72	-2.16	-1.95	-2.09	-2.59	-2.97

Tabella 16 VaR calcolati con livello di confidenza del 5% e vari pesi β per il portafoglio Mibtel-Dow-Jones. Si considerano marginali distribuite come $N(0,1)$. (tra parentesi gli standard error)

COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho = 0.59$ (0.029)	-1.55	-1.32	-1.23	-1.32	-1.55	-1.71
Gumbel	$p=1.62$ (0.077)	-1.62	-1.60	-1.59	-1.61	-1.63	-1.64
Franck	$p=4.14$ (0.320)	-1.57	-1.47	-1.43	-1.47	-1.57	-1.64
Plackett	$p=6.31$ (0.718)	-1.55	-1.43	-1.39	-1.43	-1.56	-1.64
Tradizionale	$\rho=0.56$	-2.72	-2.16	-1.95	-2.09	-2.59	-2.97

Tabella 17 VaR calcolati con livello di confidenza del 1% e vari pesi β per il portafoglio Mibtel-Dow-Jones. Si considerano marginali distribuite come student-t (tra parentesi gli standard error)

$v_{\text{mibtel}}=18.92$ (10.62) , $v_{\text{Dow-Jones}}=34.06$ (32.71)							
COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho = 0.59$ (0.029)	-3.14	-2.61	-2.40	-2.61	-3.15	-3.48
Student-t	$\rho = 0.58$ (0.029) $v=8.87$ (4.259)	-1.23	-1.15	-1.13	-1.15	-1.22	-1.26
Gumbel	$p=1.62$ (0.077)	-3.31	-3.24	-3.22	-3.24	-3.31	-3.36
Franck	$p=4.14$ (0.320)	-3.14	-2.84	-2.73	-2.84	-3.16	-3.36
Plackett	$p=6.31$ (0.718)	-3.12	-2.80	-2.68	-2.80	-3.14	-3.36
Tradizionale	$\rho = 0.56$	-5.20	-4.13	-3.76	-4.05	-5.03	-5.78

Tabella 18 VaR calcolati con livello di confidenza del 1% e vari pesi β per il portafoglio Mibtel-Dow-Jones. Si considerano marginali distribuite come $N(0,1)$. (tra parentesi gli standard error)

COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho = 0.59$ (0.029)	-2.16	-1.82	-1.70	-1.82	-2.16	-2.37
Gumbel	$p=1.62$ (0.077)	-2.36	-2.26	-2.25	-2.26	-2.30	-2.33
Franck	$p=4.14$ (0.320)	-2.19	-2.00	-1.93	-2.00	-2.20	-2.33
Plackett	$p=6.31$ (0.718)	-2.18	-1.97	-1.90	-1.97	-2.18	-2.33
Tradizionale	$\rho = 0.56$	-5.20	-4.13	-3.76	-4.05	-5.03	-5.78

Tabella 19 VaR calcolati con livello di confidenza del 5% e vari pesi β per il portafoglio Mibtel-FTSE100. Si considerano marginali distribuite come student-t (tra parentesi gli standard error)

$v_{\text{mibtel}}=18.92$ (10.62) , $v_{\text{FTSE100}}=27.27$ (18.03)							
COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.27$ (0.049)	-1.56	-1.31	-1.23	-1.33	-1.58	-1.74
Student-t	$\rho =0.27$ (0.049) $v=5.91$ (1.956)	-0.48	-0.41	-0.39	-0.41	-0.46	-0.49
Gumbel	$p=1.21$ (0.052)	-1.62	-1.53	-1.50	-1.54	-1.65	-1.73
Franck	$p=1.77$ (0.336)	-1.59	-1.42	-1.37	-1.43	-1.61	-1.73
Plackett	$p=2.62$ (0.439)	-1.57	-1.38	-1.32	-1.40	-1.59	-1.73
Tradizionale	$\rho =0.24$	-2.43	-1.86	-1.68	-1.90	-2.52	-2.98

Tabella 20 VaR calcolati con livello di confidenza del 5% e vari pesi β per il portafoglio Mibtel-FTSE100. Si considerano marginali distribuite come $N(0,1)$ (tra parentesi gli standard error)

COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.27$ (0.049)	-1.50	-1.26	-1.18	-1.26	-1.50	-1.66
Gumbel	$p=1.21$ (0.052)	-1.56	-1.48	-1.44	-1.58	-1.57	-1.64
Franck	$p=1.77$ (0.336)	-1.53	-1.37	-1.31	-1.37	-1.53	-1.64
Plackett	$p=2.62$ (0.439)	-1.51	-1.33	-1.26	-1.33	-1.52	-1.64
Tradizionale	$\rho =0.24$	-2.43	-1.86	-1.68	-1.90	-2.52	-2.98

Tabella 21 VaR calcolati con livello di confidenza del 1% e vari pesi β per il portafoglio Mibtel-FTSE100. Si considerano marginali distribuite come student-t (tra parentesi gli standard error)

$v_{\text{mibtel}}=18.92$ (10.62) , $v_{\text{FTSE100}}=27.27$ (18.03)							
COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.27$ (0.049)	-3.05	-2.53	-2.32	-2.53	-3.06	-3.39
Student-t	$\rho =0.27$ (0.049) $v=5.91$ (1.956)	-1.09	-1.00	-0.96	-1.00	-1.12	-1.20
Gumbel	$p=1.21$ (0.052)	-3.18	-2.93	-2.83	-2.93	-3.19	-3.36
Franck	$p=1.77$ (0.336)	-3.08	-2.68	-2.52	-2.68	-3.10	-3.36
Plackett	$p=2.62$ (0.439)	-3.07	-2.63	-2.46	-2.63	-3.08	-3.36
Tradizionale	$\rho =0.24$	-4.76	-3.58	-3.20	-3.63	-4.86	-5.78

Tabella 22 VaR calcolati con livello di confidenza del 1% e vari pesi β per il portafoglio Mibtel-FTSE100. Si considerano marginali distribuite come $N(0,1)$ (tra parentesi gli standard error)

COPULA	Stima del parametro della copula	β_{mibtel} 0.1	β_{mibtel} 0.3	β_{mibtel} 0.5	β_{mibtel} 0.7	β_{mibtel} 0.9	β_{mibtel} 1
Gaussian	$\rho =0.27$ (0.049)	-2.12	-1.78	-1.66	-1.78	-2.12	-2.34
Gumbel	$p=1.21$ (0.052)	-2.25	-2.14	-2.11	-2.14	-2.24	-2.33
Franck	$p=1.77$ (0.336)	-2.15	-1.89	-1.80	-1.89	-2.15	-2.33
Plackett	$p=2.62$ (0.439)	-2.13	-1.85	-1.75	-1.86	-2.14	-2.33
Tradizionale	$P =0.24$	-4.76	-3.58	-3.20	-3.63	-4.86	-5.78

Come si può notare, manca la misura del VaR per la copula Student-t nei casi in cui le marginali siano considerate distribuite come $N(0,1)$, in quanto la funzione usata per misurare il VaR con il software R non è riuscita a calcolarlo.

Naturalmente ad un livello di significatività più basso, aumenta in valore assoluto il valore della misura del rischio di perdita, poiché consideriamo valori della distribuzione sempre più lontani dal centro della distribuzione.

Per ogni portafoglio e per entrambi i livelli di significatività, laddove si valutano marginali “studentizzate”, la copula Student-t ha un valore del VaR minore in valore assoluto rispetto alle altre copule. Si può dire quindi che tale copula sottovaluta la perdita di capitale che potrebbe esserci in caso di sorte avversa.

Le altre copule considerate hanno valori che si avvicinano tra loro anche se la copula Gumbel (una delle 2 copule archimediane) sembra avere, in tutti i casi, valori maggiori in termini assoluti delle altre copule.

Rispetto alla variabile rappresentata dal peso β , si nota in ogni caso come il VaR diminuisca man mano che ci si avvicina ad un valore del β pari a 0.5 (50%). Ciò sta ad indicare che quanto più il portafoglio è bilanciato tra i 2 titoli azionari, tanto è minore il rischio di incorrere in perdite, poiché siamo nel caso di massima diversificazione.

Al contrario man mano che i pesi sono squilibrati (β che tende a 0 o a 1), il rischio di incorrere in perdita di capitale aumenta, poiché “investendo” tutto in un unico indice, entra in gioco il rischio specifico dell’indice in cui puntiamo tutto.

L’aumento del rischio è abbastanza simmetrico nei 2 casi contrapposti che si pesi di più un indice del portafoglio rispetto all’altro o viceversa. Tuttavia il VaR sarà maggiore, nei casi estremi di $\beta=0$ o $\beta=1$, per quell’indice con volatilità (standard error) maggiore.

Sempre rispetto al variare della composizione dei portafogli, si riscontra nella copula Gaussiana la variazione più elevata della misura del Value at Risk tra valori estremi e centrali del parametro β .

Da notare inoltre che le copule considerate (ad esclusione di quella student) tendono a coincidere per i valori estremi di β .

Se confrontiamo i risultati ottenuti assumendo distribuzioni diverse per le marginali, si nota come:

- per il livello di confidenza del 5% i valori nei due casi siano abbastanza vicini, anche se nel caso di marginali “studentizzate” il valore del VaR in termini assoluti è maggiore.
- Per il livello di confidenza del 1% , la differenza tra i due casi aumenta , perché si sente maggiormente il fatto che le marginali distribuite come una student t hanno code più pesanti. Questa differenza è più grande nel caso in cui si stimi un valore basso dei gradi di libertà per la marginale T-student, perché al variare del parametro dei gradi di libertà all’infinito la marginale ha un comportamento sempre più simile a quello di una normale.

Se confrontiamo il VaR calcolato nella maniera tradizionale (ultima riga di ogni tabella) , cioè supponendo una distribuzione congiunta delle osservazioni normale multivariata, con quello valutato dalle copule, ci accorgiamo come in ogni caso considerato la misura della perdita potenziale di capitale sia “sovrastimata”.

Ad una banca che calcoli il VaR per informare i clienti circa il rischio di perdita potenziale di capitale di un portafoglio di titoli, non consigliamo di usare il metodo tradizionale di misurazione del VaR.

Tra i vari portafogli quelli che hanno la maggiore probabilità di incorrere in perdite sono quelli composti dall’indice Mibtel con l’indice Dow-Jones e dal Mibtel con il FTSE100 inglese. Questi portafogli , infatti, hanno valori del VaR in ogni caso maggiori rispetto agli altri 2 portafogli.

5. IMPLEMENTAZIONE IN R: La funzione **copula**.

Il programma, scritto col software R , ha lo scopo di tradurre in pratica il metodo di stima IFM (spiegato nel capitolo 3) per i parametri delle copule e delle marginali. Inoltre vengono calcolati gli standard error dei parametri a partire dalla matrice d'informazione di Godambe (3.8).

Da ultimo viene presentata la procedura tecnica per il calcolo del VaR per mezzo delle copule.

Questo capitolo raccoglie i comandi più importanti del programma che viene riportato per intero nell'appendice.

Le librerie d'interesse per il programma sono:

```
library(ts)  
library(tseries)
```

Il loro uso è limitato alla fase di analisi delle serie di input, alla stima del modello Garch(1,1) e al calcolo dei residui standardizzati del modello.

I dati di input sono le osservazioni di due serie storiche di indici azionari.

```
serie<-read.table("C:\\WINNT\\Profiles\\zilio\\Desktop\\  
serietutte.txt", header=T)
```

```
copula<-function(dati,pesi)  
{
```

La funzione **copula**, che dipende dalle variabili **dati** (dati di input) e **pesi** (pesi del portafoglio bivariato su cui è calcolato il VaR) , ha al suo interno tutte le procedure per analizzare e stimare un modello Garch(1,1) per le serie dei rendimenti degli indici azionari dati come input, per stimare i parametri delle copule e i relativi standard error , per calcolare i test di significatività sui parametri (e i p-value) e per valutare numericamente il VaR per il portafoglio formato dai 2 indici azionari considerati.

In conclusione, basta assegnare, alla funzione `copula`, i dati di input e i pesi del portafoglio, per ottenere in output ciò che si desidera studiare.

```
lserie<-log(dati)
rserie<-matrix(0,549,2)
rserie[,1]<-diff(lserie[,1],1)
rserie[,2]<-diff(lserie[,2],1)
```

Questi comandi calcolano i rendimenti logaritmici per entrambe gli indici azionari.

Successivamente per le serie dei rendimenti viene stimato un modello Garch(1,1)

```
unigarch=function(y)
{
garch.y=garch(y,order=c(1,1),trace=F)
  resid.stand=residuals(garch.y,standardize=T)
  return(resid.stand)
}
parg<-array(0,c(2,3))
se.g<-array(0,c(2,3))
rgserie<-array(0,c(549,2))
for (i in 1:2)
{
parg[i,]<-garch(rserie[,i],order=c(1,1))$coef
se.g[i,]<-garch(rserie[,i],order=c(1,1))$asy.se.coef
test.g<-parg/se.g
p.value.g<-pnorm(test.g)
rgserie[,i]<-unigarch(rserie[,i])
rgserie[1,]<-0
}
```

Per i 2 modelli Garch(1,1) (uno per serie) , vengono stimati con il metodo della massima verosimiglianza, mediante il comando `Garch`, i parametri del modello e i relativi standard error. A partire da queste informazioni si ottengono i test di significatività dei parametri e i rispettivi p-value.

La funzione **unigarch** restituisce, per entrambe le serie, i residui standardizzati del modello.

```
par(mfrow = c(2, 1))
hist(rgserie[,1],freq=F,right=F,main="Residui standardizzati
del Garch(1,1) della 1° serie",ylab="Density" ,
xlab="quantili")
hist(rgserie[,2],freq=F,right=F,main="Residui standardizzati
del Garch(1,1) della 2° serie",ylab="Density" ,
xlab="quantili")
```

Col comando **hist** vengono disegnati gli istogrammi dei residui standardizzati del Garch(1,1) per le due serie. Si vuole vedere se tali residui hanno un'andamento della distribuzione normale (N(0,1) visto che sono standardizzati) o se, data la pesantezza delle cose, è preferibile stimare per essi una distribuzione t-student.

```
jbt<-rep(0,2)
p.val<-rep(0,2)
for (i in 1:2)
{
jjbt<-jarque.bera.test(rgserie[,i])
jbt[i]<-jjbt$statistic
p.val[i]<-(1-pchisq(jbt[i],2))
}
```

Con il comando **jarque.bera.test** si testa l'ipotesi di normalità (tramite l'omonimo test) della distribuzione dei residui standardizzati delle due serie. La statistica test sotto l'ipotesi di normalità ha distribuzione χ_2^2 . Il comando **p.val** restituisce i p.value del test.

Si procede ora alla stima dei gradi di libertà, ipotizzando per le 2 serie dei residui una distribuzione t-student.

```
pr<-array(0,2)
stderr.t<-c(0,0)
```

```

for ( i in 1:2)
{
tstudent<-function(p)
sum(log(dt(rgserie[,i],p)))
pri<-
optim(p=10,tstudent,gr=NULL,method=c("BFGS"),control=list(fn
scale=-1),hessian=TRUE)
pr[i]<-pri$par
stderr.t[i]<-sqrt(-solve(pri$hessian))
}

```

Il comando `optim` calcola il valore della variabile `p`, per cui la funzione (funzione di verosimiglianza della distribuzione t-student) ammette valore massimo. Inoltre viene restituito il valore dell'hessiano, il quale invertito e cambiato di segno rappresenta la matrice d'informazione osservata, che è una stima della varianza del parametro stimato col metodo della massima verosimiglianza. Come abbiamo visto questa stima non è efficiente.

```

se<-array(0,c(549,2))
for ( i in 1:2)
{
se[,i]<-pt(rgserie[,i],pr[i])
}

```

Le due serie dei residui vengono trasformate col comando `pt` nelle rispettive serie delle probabilità cumulate delle funzioni di distribuzione student-t stimate.

D'ora in poi queste serie rappresenteranno le “marginali” delle funzioni copula.

```

parametri<-function(dati1,dati2)
{

```

La funzione `parametri` raccoglie i comandi per la stima dei parametri delle copule e dei corrispondenti standard error nel modo tradizionale (non efficiente nel nostro caso) e mediante il calcolo della matrice di Godambe. Inoltre questa funzione restituisce i valori del VaR calcolato con le copule per il portafoglio formato dai due indici azionari.


```

n<-length(dati1)
cor<-c(0,0)
stderr.stud<-c(0,0)
{
studcopbiv<-function(p)
(-(n/2)*log(1-p[1]^2)+n*lgamma((p[2]+2)/2)+n*lgamma(p[2]/2)-
2*n*lgamma((p[2]+1)/2)-((p[2]+2)/2)*sum(log(1+
(qt(dati1,p[2])^2+qt(dati2,p[2])^2-2*p[1]*qt(dati1,p[2])*
qt(dati2,p[2]))/(p[2]*(1-p[1]^2))))+(p[2]+1)/2)*sum
(log((1+(qt(dati1,p[2])^2)/p[2])*(1+(qt(dati2,p[2])^2)/p[2])
)))
cor<-optim(p=c(0.5,3),studcopbiv,gr=NULL,method=c("BFGS"),
control=list(fnscale=-1),hessian=TRUE)
par<-cor$par
stderr.stud[1]<-sqrt(-solve(cor$hessian[1,1]))
stderr.stud[2]<-sqrt(-solve(cor$hessian[2,2]))
}

```

Come prima cosa mediante `optim` calcoliamo i parametri della copula e il loro hessiano, che serve per stimare le rispettive varianze. La funzione da massimizzare è quella di verosimiglianza della copula d'interesse.

Nel seguito, per comodità, riportiamo i comandi solo per la copula t-student, ma le stesse procedure devono essere svolte anche per le altre copule.

La funzione `parametri` ha al suo interno anche i comandi per il calcolo del VaR del portafoglio composto dai due indici azionari. Esso può essere calcolato per diverse composizioni del portafoglio, date dalla combinazioni di pesi assegnati a ciascun indice (per vari valori della variabile pesi). Da notare che per il calcolo del VaR si ricorre ad un doppio `integrate`, come suggerito dall'espressione 4.1.

```

ro <-par[1]
v<-pr[1]
v1<-pr[2]
v2<-par[2]
prob<-function(z)

```

```

{
f3 <- function(y)
{
integrate(function(x)(1/sqrt(1-
ro^2))*gamma((v2+2)/2)*gamma(v2/2)/ gamma
((v2+1)/2)^2*((1+(qt(pt(x,v),v2)^2+qt(pt(y,v1),v2)^2-
2*ro*qt(pt(x,v),v2)*qt(pt(y,v1),v2))/(v2*(1-ro.s^2)))^-
(v2+2)/2)/(((1+(qt(pt(x,v),v2)^2)/v2)^(-(v2+1)/2))*((1+
(qt(pt(y,v1),v2)^2)/v2)^(-(v2+1)/2)))*dt(x,v),lower=-Inf,
upper=(1/pesi)*z-((1-pesi)/pesi)*y)$value
}
integrate(function(y)      sapply(y,f3)*      dt(y,v1),lower=-
Inf,upper=Inf)$value-0.01
}
st<-uniroot(prob,0,-4)

```

Poiché la funzione che definisce il VaR dipende da un parametro (z), si deve trovare il valore di questo parametro per il quale la funzione valga 0. il comando `uniroot` serve a tale proposito.

Il passo successivo fatto dalla funzione parametri è quello di determinare in maniera efficiente gli standard error, partendo dalla rappresentazione della matrice di Godambe (3.8)

Si definisce per prima la matrice V , ovvero la matrice attesa dei prodotti delle funzioni punteggio (una per la copula e due per le marginali). Al posto del valore atteso che definisce V , usiamo una stima per esso, come nella 3.12

Di seguito vengono calcolate le derivate prime “campionarie” delle funzioni di verosimiglianza delle marginali e della copula (come già detto, riportiamo solo quella della copula t-student). Le derivate sono calcolate a partire dalla definizione di derivata come rapporto incrementale di una data funzione. L’ approssimazione data (ϵ) varia da 10^{-4} a 10^{-5} .

Derivata prima della funzione di verosimiglianza della marginale rispetto al parametro dei gradi di libertà.

```
f=function(p) log(dt(rgserie[,1],p))
```

```

dermarg1 <- function(p,eps=1e-5) (f(p+eps)-f(p))/eps
marg1<-(dermarg1(pr[1]))

```

```

f=function(p) log(dt(rgserie[,2],p))
dermarg2 <- function(p,eps=1e-5) (f(p+eps)-f(p))/eps
marg2<-(dermarg2(pr[2]))

```

Derivata prima della funzione di verosimiglianza della copula rispetto ai parametri della copula

```

f=function(ro)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2) +lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+
(qt(se[,1],v)^2+qt(se[,1+1],v)^2-2*ro*qt(se[,1],v)*
qt(se[,1+1],v))/(v*(1-ro^2)))) +((v+1)/2)*(log((1+(qt
(se[,1],v)^2)/v)*(1+(qt(se[,1+1],v)^2)/v))))
dersturo1<-function(p,eps=2e-5) (f(p+eps)-f(p))/eps
student<-dersturo1(par[1])

```

```

ro<-par[1]
f=function(v)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(se[,1],v)^2
+qt(se[,1+1],v)^2-2*ro*qt(se[,1],v)*qt(se[,1+1],v))/(v*(1-
ro^2)))) +((v+1)/2)*(log((1+(qt(se[,1],v)^2)/v)*
(1+(qt(se[,1+1],v)^2)/v))))
dersturo1<-function(p,eps=2e-5) (f(p+eps)-f(p))/eps
student1<-dersturo1(par[2])

```

Il secondo passo è il calcolo della matrice D, ovvero la matrice delle derivate seconde e delle derivate seconde miste delle funzioni di verosimiglianza rispetto a tutti i parametri.

Come prima, al posto del valore atteso si è proceduto alla stima della matrice delle derivate seconde “campionarie” (3.13).

Derivata seconda della funzione di verosimiglianza della marginale rispetto al parametro dei gradi di libertà.

```

f=function(p) log(dt(rgserie[,1],p))
dermarg11 <- function(p,eps=1e-4){
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
marg11<-dermarg11(pr[1])

```

```

f=function(p) log(dt(rgserie[,2],p))
dermarg22 <- function(p,eps=1e-4)
{
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
marg22<-dermarg22(pr[2])

```

Derivata seconda della funzione di verosimiglianza della copula rispetto ai parametri della copula

```

f=function(v)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-2*lgamma
((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),v)^2
+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1),v)*
qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2)))))+(v+1)/2*(log((1+
(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)
^2)/v))))
der2 <- function(p,eps=1e-5)
{
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
student3<-(der2(par[2]))

f=function(ro)

```

```

(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1)
,v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1)
,v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2))))+((v+1)/2)
*(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,
2],p2),v)^2)/v))))
der2 <- function(p,eps=1e-5){
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
student7<-(der2(par[1]))

```

Derivata seconda mista della funzione di verosimiglianza della copula rispetto a ciascun parametro della copula

```

p1<-pr[1]
p2<-pr[2]
f=function(v,ro)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1)
,v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1)
,v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2))))+((v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2]
],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
student2<-(der2(par[2],par[1]))

```

```

f=function(v,p1)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),
v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1)
,v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2))))+((v+1)/2)*

```

```

(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
student4<-(der2(par[2],pr[1]))

f=function(v,p2)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1),v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2)))))+((v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
student5<-(der2(par[2],pr[2]))

f=function(ro,v)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1),v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2)))))+((v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}

```

```

student6<-(der2(par[1],par[2]))

f=function(ro,p1)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),
v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1),
v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2))))+((v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
student8<-(der2(par[1],pr[1]))

f=function(ro,p1)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),
v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1),
v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2))))+((v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
student9<-(der2(par[1],pr[2]))

```

Il passo finale è ricomporre le due matrici secondo la 3.14 in maniera da trovare la matrice di Godambe per il calcolo degli errori standard efficienti per i parametri stimati.

```

vstudent<-0
for (i in 1:n)

```

```

{
dstudent<-
dstudent+(rbind(marg1[i],marg2[i],student[i],student1[i]))**
cbind(marg1[i],marg2[i],student[i],student1[i]))/n
}
o<-0; p<-0;q<-0;r<-0
for (i in 1:n)
{
o<-o+(cbind(marg11[i],0,0,0))/n
p<-p+(cbind(0,marg22[i],0,0))/n
q<-
q+(cbind(student8[i],student9[i],student7[i],student6[i]))/n
r<-
r+(cbind(student4[i],student5[i],student2[i],student3[i]))/n
}
dstudent<-rbind(o,p,q,r)
gstudent<-
sqrt((solve(dstudent**solve(vstudent)**t(dstudent))/n))

```

Infine si sono usati gli standard error efficienti appena stimati per calcolare test di significatività sui parametri della copula e delle marginali, calcolando anche il rispettivo p-value facendo leva sulla distribuzione campionaria dei parametri (3.7).

```

t.Student1<-par[1]/gstudent[3,3]
t.Student2<-par[2]/gstudent[4,4]
t.marg1<-pr[1]/gstudent[1,1]
t.marg2<-pr[2]/gstudent[2,2]
p.Student1<-1-pnorm(t.Student1)
p.Student2<-1-pnorm(t.Student2)
p.marg1<-1-pnorm(t.marg1)
p.marg2<-1-pnorm(t.marg2)
}
}

```


6. CONCLUSIONI

Il presente lavoro ha tra i suoi scopi quello di presentare il concetto di copula , la quale possiede la proprietà di legare tra loro singole funzioni di distribuzione marginali, secondo una struttura di dipendenza, per formare una funzione di distribuzione multivariata. E' una peculiarità propria delle copule quella di essere "scomponibile" in queste due parti.

Dopo averne dato una definizione dal punto di vista teorico-matematico , ed averne considerato i legami con le misure di associazione , si è proceduto all'analisi inferenziale della copula, che può essere vista come un modello multivariato.

Successivamente la nozione di copula è stata applicata al campo dell'analisi finanziaria ed in particolare al calcolo del Value at Risk.

L'analisi svolta dal punto di vista inferenziale e per la misura del VaR è stata implementata con il software statistico R.

Il risultato più importante della teoria delle copule è senz'altro il teorema di Sklar il quale garantisce che ogni copula è una funzione di distribuzione congiunta, se i suoi argomenti sono funzioni di distribuzione marginali, e vale anche l'opposto, cioè ogni funzione di distribuzione congiunta può essere estesa a copula e se le marginali sono continue l'estensione è unica.

Il teorema di Sklar permette di scrivere la probabilità congiunta come funzione delle marginali cumulative e viceversa :

$$F(x) = C(F_1(x_1), \dots, F_n(x_n)).$$

Questa possibilità viene chiamata **interpretazione probabilistica di base delle copule**.

Un'implicazione del teorema è quella di permettere alla copula di essere trattata come un modello multivariato e di ottenere quindi la relativa funzione di densità, che può essere scritta in forma canonica, come segue, separando la densità della copula vera e propria (struttura di dipendenza) dalle densità marginali

$$f(x_1, \dots, x_n) = c(F_1(x_1), \dots, F_n(x_n)) \cdot \prod_{j=1}^n f_j(x_j)$$

Le principali famiglie di copule sono : la Gaussiana, la student-t (che derivano dalle omonime distribuzioni multivariate) , le copule archimediane (Franch, Gumbel) e la copula di Plackett.

Dal punto di vista inferenziale esistono due metodi di stima per i parametri delle copule : il classico metodo di verosimiglianza (detta anche esatta) ed il metodo IFM (inference for the margins).

Il secondo metodo parte dalla rappresentazione canonica, e stima (col metodo della verosimiglianza) prima i parametri delle marginali e successivamente quelli della struttura di dipendenza. Questo metodo può essere definito, impropriamente, come verosimiglianza “a due passi”.

Il metodo IFM è preferibile al primo perché è computazionalmente più semplice, specie nei casi di grandi dimensioni, dato che non è necessario stimare congiuntamente i parametri delle marginali e quelli della struttura di dipendenza.

Tuttavia gli standard error calcolati come semplice inversione della matrice d’informazione di Fisher (hessiano della funzione di verosimiglianza calcolato nel punto di stima del parametro) non sono efficienti e si deve ricorrere al calcolo della matrice di Godambe (3.8).

Gli standard error così ottenuti possono essere usati per calcolare test di significatività sui parametri delle copule e delle marginali sfruttando la distribuzione asintotica 3.7 dei parametri.

Il metodo IFM e la stima della matrice di Godambe, confrontata con l’inversione dell’ hessiano della funzione di verosimiglianza, sono stati applicati ai dati settimanali di due indici azionari (DAX30 e Dow-Jones).

Si nota come nel secondo caso gli standard error risultino sottostimati rispetto al primo caso.

Il metodo delle copule viene applicato alla misura del Value at Risk tramite il concetto di distribuzione condizionale via copula $C_k(u_k | u_1, \dots, u_{k-1})$ ed in particolare tramite l’espressione 4.1.

L’utilizzo delle copule per il calcolo del VaR è interessante per vedere come la struttura di dipendenza, scelta per “legare” le marginali, influenzi il valore del Value at Risk. Allo stesso modo si può studiare come le ipotesi sulla distribuzione marginale influenzi tale misura.

L'analisi è stata svolta per 4 portafogli di indici azionari (Mibtel-Cac, Mibtel-Dax, Mibtel-Dow-Jones e Mibtel-FTSE) su dati settimanali, con livelli di significatività dell'1 e del 5%, facendo uso delle 5 principali famiglie di copule.

Le marginali sono state considerate, in un primo caso, distribuite come student-t con v gradi di libertà, successivamente distribuite come normali con media zero e varianza unitaria.

Da ultimo, sono stati fatti variare i pesi attribuiti ai portafogli per misurare la sensibilità del valore del VaR a tale parametro.

La copula Student-t è risultata avere un valore del VaR minore in valore assoluto rispetto alle altre copule, mentre dal punto di vista dei pesi attribuiti ai portafogli si nota come il VaR diminuisca man mano che ci si avvicina ad un valore del β pari a 0.5 , poiché siamo nel caso di massima diversificazione.

APPENDICE

```
library(ts)
library(tseries)
library(sn)

serie<-read.table("A:\\ seriedati.txt",header=T)
serie1<-cbind(serie[,3],serie[,11])

copula<-function(dati, pesi)
{
  lserie<-log(dati)
  rserie<-matrix(0,549,2)
  rserie[,1]<-diff(lserie[,1],1)
  rserie[,2]<-diff(lserie[,2],1)
  unigarch=function(y)
  {
    garch.y=garch(y,order=c(1,1),trace=F)
    resid.stand=residuals(garch.y,standardize=T)
    return(resid.stand)
  }
  parg<-array(0,c(2,3))
  se.g<-array(0,c(2,3))
  rgserie<-array(0,c(549,2))
  for (i in 1:2)
  {
    parg[i,]<-garch(rserie[,i],order=c(1,1))$coef
    se.g[i,]<-garch(rserie[,i],order=c(1,1))$asy.se.coef
    test.g<-parg/se.g
    p.value.g<-pnorm(test.g)
    rgserie[,i]<-unigarch(rserie[,i])
    rgserie[1,]<-0
  }

  par(mfrow = c(2, 1))
```

```

hist(rgserie[,1],freq=F,right=F,main="Residui standardizzati
del          Garch(1,1)          della          1°
serie",ylab="Density",xlab="quantili")
hist(rgserie[,2],freq=F,right=F,main="Residui standardizzati
del          Garch(1,1)          della          2°
serie",ylab="Density",xlab="quantili")

jbt<-rep(0,2)
p.val<-rep(0,2)
for (i in 1:2)
{
jjbt<-jarque.bera.test(rgserie[,i])
jbt[i]<-jjbt$statistic
p.val[i]<-(1-pchisq(jbt[i],2))
}

pr<-array(0,2)
stderr.t<-c(0,0)
for ( i in 1:2)
{
tstudent<-function(p)
sum(log(dt(rgserie[,i],p)))
pri<-
optim(p=10,tstudent,gr=NULL,method=c("BFGS"),control=list(fn
scale=-1), hessian=TRUE)
pr[i]<-pri$par
stderr.t[i]<-sqrt(-solve(pri$hessian))
}

se<-array(0,c(549,2))
for ( i in 1:2)
{
se[,i]<-pt(rgserie[,i],pr[i])
}

parametri<-function(dati1,dati2)
{

```

```

n<-length(dati1)
cor<-c(0,0)
stderr.stud<-c(0,0)
studcopbiv<-function(p)
  (-(n/2)*log(1-p[1]^2)+n*lgamma((p[2]+2)/2)+n*lgamma(p[2]/2)-
  2*n*lgamma((p[2]+1)/2)-((p[2]+2)/2)*sum(log(1+
  (qt(dati1,p[2])^2+qt(dati2,p[2])^2-2*p[1]*qt(dati1,p[2])*qt
  (dati2,p[2]))/(p[2]*(1-p[1]^2))))+((p[2]+1)/2)*sum(log((1+
  (qt(dati1,p[2])^2)/p[2])*(1+(qt(dati2,p[2])^2)/p[2])))
  cor<-optim(p=c(0.5,3),studcopbiv,gr=NULL,method=c("BFGS"),
  control=list(fnscale=-1),hessian=TRUE)
  par<-cor$par
  stderr.stud[1]<-sqrt(-solve(cor$hessian[1,1]))
  stderr.stud[2]<-sqrt(-solve(cor$hessian[2,2]))

gaussbiv<-function(p)
  (-n/2*log(1-p^2)+sum((qnorm(dati1)^2+qnorm(dati2)^2)/2+
  (2*p*qnorm(dati1)*qnorm(dati2)-qnorm(dati2)^2-
  qnorm(dati1)^2)/(2*(1-p^2))))
  corr<-optim(p=0.1,gaussbiv,gr=NULL,method=c("BFGS"),
  control=list(fnscale=-1),hessian=TRUE)
  par1<-corr$par
  stderr.gauss<-sqrt(-solve(corr$hessian))

plackett<-function(p)
  (n*log(p)+sum(log(1+(dati1-2*dati1*dati2+dati2)*(p-1)))-
  1.5*sum(log((1+(p-1)*(dati1+dati2))^2-4*dati1*dati2*p*(p-
  1))))
  pl<-optim(p=0.1,plackett,gr=NULL,method=c("BFGS"),
  control=list(fnscale=-1),hessian=TRUE)
  par2<-pl$par
  stderr.pl<-sqrt(-solve(pl$hessian))

Franck<-function(p)
  sum(log(-(1/p)*((exp(-p*dati1))*p*(exp(-p*dati2)*p)/(exp(-
  p)-1)/(1+((exp(-p*dati1))-1)*(exp(-p*dati2)-1)/(exp(-p)-1))-
  ((exp(-p*dati1))*p*(exp(-p*dati2)-1)/(exp(-p)-1))*((exp(-

```

```

p*dati1))-1)*(exp(-p*dati2)*p)/(exp(-p)-1))/((1+((exp(-
p*dati1))-1)*(exp(-p*dati2)-1)/(exp(-p)-1))^2))))
  fr<-optim(p=0.1,Franck,gr=NULL,method=c("BFGS"),
control=list(fnscale=-1),hessian=TRUE)
par3<-fr$par
stderr.fr<-sqrt(-solve(fr$hessian))

gumbel<-function(p)
(sum(-((-log(dati1))^p+(-log(dati2))^p)^(1/p))-sum(log(dati1
*dati2)))+(p-1)*sum(log(-log(dati1)))+(p-1)*sum(log(-log
(dati2)))+(1-2*p)/p)*sum(log((-log(dati1))^p+(-log (dati2
))^p))+sum(log((-log(dati1))^p+(-log(dati2))^p)^(1/p) +p-
1)))
gm<-optim(p=1,gumbel,gr=NULL,method=c("BFGS"),
control=list(fnscale=-1),hessian=TRUE)
par4<-gm$par
stderr.gm<-sqrt(-solve(gm$hessian))

p<-par4
v<-pr[1]
v1<-pr[2]
prob<-function(z)
{
f3 <- function(y)
{
integrate(function(x) exp(-((-log(pt(x,v))))^p+(-
log(pt(y,v1)))^p)^(1/p)) *(1/pt(x,v))*(1/pt(y,v1))*(-
log(pt(x,v)))^(p-1)*(-log(pt(y,v1)))^(p-1)* ((-
log(pt(x,v)))^p+(-log(pt(y,v1)))^p)^(1-2*p)/p) *((( -
log(pt(x,v)))^p+(-log(pt(y,v1)))^p)^(1/p)+p-1)* dt(x,v)
,lower=-Inf,upper=(1/pesi)*z-((1-pesi)/pesi)*y)$value
}
integrate(function(y) sapply(y,f3)*dt(y,v1) ,lower=-
Inf,upper=Inf)$value-0.01
}
gu<-uniroot(prob,0,-4)

p<-par2

```



```

prob<-function(z)
{
f3 <- function(y)
{
  integrate(function(x) p*(1+(pt(x,v)-
2*pt(x,v)*pt(y,v1)+pt(y,v1))*(p-1))/ ((1+(p-
1)*(pt(x,v)+pt(y,v1)))^2-4*pt(x,v)*pt(y,v1)*p*(p-1))^1.5*
dt(x,v),lower=-Inf,upper=(1/pesi)*z-((1-pesi)/pesi)*y)$value
}
integrate(function(y) sapply(y,f3)*dt(y,v1),lower=-
Inf,upper=Inf)$value-0.01
}
pl<-uniroot(prob,0,-4)

```

```

p<-par3
prob<-function(z)
{
f3 <- function(y)
{
  integrate(function(x) dt(x,v)*-( (1/p) * ((exp(-p *
pt(x,v))) * p * ( exp(-p * pt(y,v1)) * p)/(exp(-p ) - 1)/(1
+ (( exp(-p * pt(x,v))) - 1) * ( exp(-p * pt(y,v1)) -
1)/(exp(-p ) - 1))- (( exp(-p * pt(x,v))) * p * ( exp(-p *
pt(y,v1)) - 1)/(exp(-p ) - 1)) * ((( exp(-p * pt(x,v))) -
1) * ( exp(-p * pt(y,v1)) * p)/(exp(-p ) - 1)))/((1 + ((
exp(-p * pt(x,v))) - 1) * ( exp(-p *pt(y,v1)) - 1)/(exp(-p
) - 1))^2)),lower=-Inf,upper=(1/pesi)*z-((1-
pesi)/pesi)*y)$value
}
integrate(function(y) sapply(y,f3)*dt(y,v1),lower=-
Inf,upper=Inf)$value-0.01
}
fr<-uniroot(prob,0,-4)

```

```

ro.s<-par[1]
v2<-par[2]
prob<-function(z)
{

```

```

f3 <- function(y)
{
integrate(function(x) (1/sqrt(1-ro.s^2))*gamma((v2+2)/2)*
gamma(v2/2) /gamma((v2+1)/2)^2*((1+(qt(pt(x,v),v2)^2+
qt(pt(y,v1),v2)^2-2*ro.s* qt(pt(x,v),v2)*qt(pt(y,v1),
v2))/(v2*(1-ro.s^2)))^-(v2+2)/2)/(((1+(qt(pt(x,v),v2)^2)
/v2)^(-v2+1)/2))*((1+(qt(pt(y,v1),v2)^2)/v2)^(-
(v2+1)/2)))*dt(x,v),lower=-Inf,upper=(1/pesi)*z-((1-
pesi)/pesi)*y)$value
}
integrate(function(y) sapply(y,f3)* dt(y,v1),lower=-
Inf,upper=Inf)$value-0.01
}
st<-uniroot(prob,0,-4)

ro<-par1
prob<-function(z)
{
f3 <- function(y)
{
integrate(function(x) 1/sqrt(1-ro^2)*exp(((qnorm(pt(x,v)))^2
+ (qnorm(pt(y,v1)))^2)/2+(2*ro*(qnorm(pt(x,v))))
*(qnorm(pt(y,v1)))-(qnorm(pt(y,v1)))^2-(qnorm(pt(x,v)))^2)
/(2*(1-ro^2)))*dt(x,v),lower=-Inf, upper=(1/pesi)*z-((1-
pesi)/ pesi)*y)$value
}
integrate(function(y) sapply(y,f3)* dt(y,v1),lower=-
Inf,upper=Inf)$value-0.01
}
ga<-uniroot(prob,0,-4)

a<-cbind(0,0)
b<-cbind(0,0)
for (i in 1:2)
{
a[i]<-mean(rgserie[,i])
b[i]<-sqrt(var(rgserie[,i]))
}

```

```

rot<-cor(rgserie)[1,2]
vartrad<--(pesi^2*(-a[1]-qnorm(0.99,a[1],b[1]^2)*b[1])^2+(1-
pesi)^2*(-a[2]-qnorm(0.99,a[2],b[2]^2)*b[2])^2+pesi*(1-
pesi)*ro[1,2]*(-a[1]-qnorm(0.99,a[1],b[1]^2)*b[1])*(-a[2]-
qnorm(0.99,a[2],b[2]^2)*b[2]))

```

```

f=function(p) log(dt(rgserie[,1],p))
dermarg1 <- function(p,eps=1e-5) (f(p+eps)-f(p))/eps
marg1<-(dermarg1(pr[1]))

```

```

f=function(p) log(dt(rgserie[,2],p))
dermarg2 <- function(p,eps=1e-5) (f(p+eps)-f(p))/eps
marg2<-(dermarg2(pr[2]))

```

```

f=function(p) log(dt(rgserie[,1],p))
dermarg11 <- function(p,eps=1e-4)
{
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}

```

```

marg11<-dermarg11(pr[1])
f=function(p) log(dt(rgserie[,2],p))
dermarg22 <- function(p,eps=1e-4)
{
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
marg22<-dermarg22(pr[2])

```

```

f=function(p)
log(-(1/p) * ((exp(-p * se[,1])) * p * ( exp(-p * se[,2])
* p)/(exp(-p ) - 1)/(1 + (( exp(-p * se[,1])) - 1) * (
exp(-p * se[,2]) - 1)/(exp(-p ) - 1))- (( exp(-p * se[,1]))
* p * ( exp(-p * se[,2]) - 1)/(exp(-p ) - 1)) * ((( exp(-p
* se[,1])) - 1) * ( exp(-p * se[,2]) * p)/(exp(-p ) -

```

```

1)))/((1 + (( exp(-p * se[,1])) - 1) * ( exp(-p * se[,2]) -
1)/(exp(-p ) - 1))^2))))
derfral<-function(p,eps=1e-5) (f(p+eps)-f(p))/eps
franck1<-(derfral(par3))

f=function(p)
  log(-( (1/p) * ((exp(-p * se[,1])) * p * ( exp(-p *
se[,1+1]) * p)/(exp(-p ) - 1)/(1 + (( exp(-p * se[,1])) -
1) * ( exp(-p * se[,1+1]) - 1)/(exp(-p ) - 1))- (( exp(-p *
se[,1])) * p * ( exp(-p * se[,1+1]) - 1)/(exp(-p ) - 1)) *
((( exp(-p * se[,1])) - 1) * ( exp(-p * se[,1+1]) *
p)/(exp(-p ) - 1)))/((1 + (( exp(-p * se[,1])) - 1) * (
exp(-p * se[,1+1]) - 1)/(exp(-p ) - 1))^2))))
derfra2 <- function(p,eps=1e-4)
{
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
franck2<-derfra2(par3)

p2<-pr[2]
f=function(p,p1)  log(-( (1/p) * ((exp(-p *
pt(rgserie[,1],p1))) * p * ( exp(-p * pt(rgserie[,2],p2)) *
p)/(exp(-p ) - 1)/(1 + (( exp(-p * pt(rgserie[,1],p1))) -
1) * ( exp(-p * pt(rgserie[,2],p2)) - 1)/(exp(-p ) - 1))-
(( exp(-p * pt(rgserie[,1],p1))) * p * ( exp(-p *
pt(rgserie[,2],p2)) - 1)/(exp(-p ) - 1)) * ((( exp(-p *
pt(rgserie[,1],p1))) - 1) * ( exp(-p * pt(rgserie[,2],p2)) *
p)/(exp(-p ) - 1)))/((1 + (( exp(-p * pt(rgserie[,1],p1))) -
1) * ( exp(-p * pt(rgserie[,2],p2)) - 1)/(exp(-p ) -
1))^2))))
derfrap1 <- function(p,p1,eps=1e-4){
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
franckp1<-derfrap1(par3,pr[1])

```

```

p2<-pr[1]
f=function(p,p1)
log(-(1/p)*((exp(-p*pt(rgserie[,1],p1))) * p * (exp(-
p * pt(rgserie[,2],p2)) * p)/(exp(-p) - 1)/(1 + ((exp(-p
* pt(rgserie[,1],p1))) - 1) * (exp(-p *
pt(rgserie[,2],p2)) - 1)/(exp(-p) - 1))- ((exp(-p *
pt(rgserie[,1],p1))) * p * (exp(-p * pt(rgserie[,2],p2)) -
1)/(exp(-p) - 1)) * (((exp(-p * pt(rgserie[,1],p1))) - 1)
* (exp(-p * pt(rgserie[,2],p2)) * p)/(exp(-p) - 1))/(1 +
((exp(-p * pt(rgserie[,1],p1))) - 1) * (exp(-p *
pt(rgserie[,2],p2)) - 1)/(exp(-p) - 1))^2))))
derfrap2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
franckp2<-derfrap2(par3,pr[2])

f=function(p)
log(p)+(log(1+(se[,1]-2*se[,1+1]*se[,1]+se[,1+1])*(p-1)))-
1.5*(log((1+(p-1)*(se[,1]+se[,1+1]))^2-
4*se[,1]*se[,1+1]*p*(p-1)))
derpla1<-function(p,eps=1e-5) (f(p+eps)-f(p))/eps
plackett1<-derpla1(par2)

f=function(p)
log(p)+(log(1+(se[,1]-2*se[,1+1]*se[,1]+se[,1+1])*(p-1)))-
1.5*(log((1+(p-1)*(se[,1]+se[,1+1]))^2-
4*se[,1]*se[,1+1]*p*(p-1)))
derpla2 <- function(p,eps=1e-4)
{
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
plackett2<-derpla2(par2)

```

```

p2<-pr[2]
f=function(p,p1)
log(p)+(log(1+(pt(rgserie[,1],p1)-2*pt(rgserie[,2],p2)*
pt(rgserie[,1],p1) +pt(rgserie[,2],p2))*(p-1)))-
1.5*(log((1+(p-1)*(pt(rgserie[,1],p1)+ pt(rgserie[,2],p2)
))^2-4*pt(rgserie[,1],p1)*pt(rgserie[,2],p2)*p*(p-1)))
derplap1 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
plackettp1<-derplap1(par2,pr[1])

p1<-pr[1]
f=function(p,p1)
log(p)+(log(1+(pt(rgserie[,1],p1)-2*pt(rgserie[,2],p2)*
pt(rgserie[,1],p1) +pt(rgserie[,2],p2))*(p-1)))-
1.5*(log((1+(p-1)*(pt(rgserie[,1],p1)+ pt(rgserie[,2],p2)
))^2-4*pt(rgserie[,1],p1)*pt(rgserie[,2],p2)*p*(p-1)))
derplap2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
plackettp2<-derplap2(par2,pr[2])

f=function(p)
(-((-log(se[,1]))^p+(-log(se[,1+1]))^p)^(1/p))-
(log(se[,1]*se[,1+1]))+(p-1)*(log(-log(se[,1])))+(p-
1)*(log(-log(se[,1+1])))+((1-2*p)/p)*(log((-
log(se[,1]))^p+(-log(se[,1+1]))^p)))+(log(((
log(se[,1]))^p+(-log(se[,1+1]))^p)^(1/p)+p-1))
dergum1<-function(p,eps=1e-4) (f(p+eps)-f(p))/eps
gumbell1<-dergum1(par2)

```

```

f=function(p)
  (-((-log(se[,1]))^p+(-log(se[,1+1]))^p)^(1/p))-
  (log(se[,1]*se[,1+1]))+(p-1)*(log(-log(se[,1])))+(p-
  1)*(log(-log(se[,1+1])))+((1-2*p)/p)*(log((-
  log(se[,1]))^p+(-log(se[,1+1]))^p))+log(((
  log(se[,1]))^p+(-log(se[,1+1]))^p)^(1/p)+p-1))
dergum2 <- function(p,eps=1e-5)
{
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
gumbel2<-(dergum2(par4))

p2<-pr[2]
f=function(p,p1)
  (-((-log(pt(rgserie[,1],p1)))^p+(-log(pt(rgserie[,2],
p2))))^p)^(1/p))-(log(pt(rgserie[,1],p1)*pt(rgserie[,2],p2)
))+(p-1)*(log(-log(pt(rgserie[,1],p1))))+(p-1)*(log(-
log(pt(rgserie[,2],p2))))+((1-2*p)/p)*(log((-log(pt
(rgserie[,1],p1)))^p+(-log(pt(rgserie[,2],p2))))^p))+log(((
log(pt(rgserie[,1],p1)))^p+(-log(pt(rgserie[,2],
p2))))^p)^(1/p)+p-1))
dergump1 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
gumbel1<-dergump1(par4,pr[1])

p1<-pr[1]
f=function(p,p1)
  (-((-log(pt(rgserie[,1],p1)))^p+(-log(pt(rgserie[,2],
p2))))^p)^(1/p))-(log(pt(rgserie[,1],p1)*pt(rgserie[,2],
p2)))+(p-1)*(log(-log(pt(rgserie[,1],p1))))+(p-1)*(log(-
log(pt(rgserie[,2],p2))))+ ((1-2*p)/p)*(log((-log(pt
(rgserie[,1],p1)))^p +(-log(pt(rgserie[,2],p2))))^p))+

```

```

(log((( -log(pt(rgserie[,1],p1)))^p+(-log(pt(rgserie[,2],
p2))))^p)^(1/p)+p-1))
dergump2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
gumbelp2<-dergump1(par4,pr[2])

f=function(p)
-1/2*log(1- p^2)+((qnorm(se[,1])^2+qnorm(se[,2])^2)/2+(2*p
*qnorm(se[,1])* qnorm(se[,2])-qnorm(se[,2])^2-
qnorm(se[,1])^2)/(2*(1-p^2)))
dergau1<-function(p,eps=1e-5) (f(p+eps)-f(p))/eps
gaussian1<-dergau1(par1)

f=function(p)
-1/2*log(1-p^2)+((qnorm(se[,1])^2+qnorm(se[,2])^2)/2+(2*p*
qnorm(se[,1])* qnorm(se[,2])-qnorm(se[,2])^2-
qnorm(se[,1])^2)/(2*(1-p^2)))
dergau2 <- function(p,eps=1e-5)
{
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
gaussian2<-(dergau2(par1))

p2<-pr[2]
f=function(p,p1)
-1/2*log(1-p^2)+((qnorm(pt(rgserie[,1],p1))^2+qnorm(pt
(rgserie[,2],p2))^2) /2+(2*p*qnorm(pt(rgserie[,1],p1))*
qnorm(pt(rgserie[,2],p2))-qnorm(pt(rgserie[,2],p2))^2-
qnorm(pt(rgserie[,1],p1))^2)/(2*(1-p^2)))
dergaup1 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps

```



```

deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
gaussianp1<-dergump1(par1,pr[1])

p1<-pr[1]
f=function(p,p1)
-1/2*log(1-p^2)+((qnorm(pt(rgserie[,1],p1))^2+qnorm
(pt(rgserie[,2],p2))^2)/2 +(2*p*qnorm(pt(rgserie[,1],p1))*
qnorm(pt(rgserie[,2],p2))-qnorm(pt(rgserie[,2],p2))^2-
qnorm(pt(rgserie[,1],p1))^2)/(2*(1-p^2)))
dergaup2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
gaussianp2<-dergaup2(par1,pr[2])

v<-par[2]
f=function(ro)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(se[,1],v)^2+
qt(se[,1+1],v)^2-2*ro*qt(se[,1],v)* qt(se[,1+1],v)))/(v*(1-
ro^2))))+((v+1)/2)*(log((1+(qt(se[,1],v)^2)/v)
*(1+(qt(se[,1+1],v)^2)/v))))
dersturo1<-function(p,eps=2e-5) (f(p+eps)-f(p))/eps
student<-dersturo1(par[1])

ro<-par[1]
f=function(v)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(se[,1],v)^2+
qt(se[,1+1],v)^2-2*ro*qt(se[,1],v)* qt(se[,1+1],v)))/(v*(1-
ro^2))))+((v+1)/2)*(log((1+(qt(se[,1],v)^2)/v)*
(1+(qt(se[,1+1],v)^2)/v))))
dersturo1<-function(p,eps=2e-5) (f(p+eps)-f(p))/eps
student1<-dersturo1(par[2])

```

```

p1<-pr[1]
p2<-pr[2]
f=function(v,ro)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),
v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],
p1),v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2))))+((v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
student2<-(der2(par[2],par[1]))

```

```

ro<-par[1]
f=function(v)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),
v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1),
v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2))))+((v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,eps=1e-5)
{
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
student3<-(der2(par[2]))

```

```

f=function(v,p1)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-2*
lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),
v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1)

```

```

,v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2)))+(v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
student4<-(der2(par[2],pr[1]))

f=function(v,p2)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),
v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1],p1),
v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2)))+(v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
student5<-(der2(par[2],pr[2]))

f=function(ro,v)
(-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],
p1),v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt
(pt(rgserie[,1],p1),v)*qt(pt(rgserie[,2],p2),v))/(v*(1-
ro^2)))+(v+1)/2)* (log((1+(qt(pt(rgserie[,1],p1),
v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}

```

```

}
student6<-(der2(par[1],par[2]))

f=function(ro)
  (-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
  2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1)
,v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1]
,p1),v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2))))+((v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2]
],p2),v)^2)/v))))
der2 <- function(p,eps=1e-5)
{
deriv1 <- function(y) (f(y+eps)-f(y))/eps
deriv <- (deriv1(p+eps)-deriv1(p))/eps
deriv
}
student7<-(der2(par[1]))

f=function(ro,p1)
  (-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
  2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],p1),
v)^2+qt(pt(rgserie[,2],p2),v)^2-2*ro*qt(pt(rgserie[,1]
,p1),v)*qt(pt(rgserie[,2],p2),v))/(v*(1-ro^2))))+((v+1)/2)*
(log((1+(qt(pt(rgserie[,1],p1),v)^2)/v)*(1+(qt(pt(rgserie[,2]
],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
student8<-(der2(par[1],pr[1]))

f=function(ro,p1)
  (-(1/2)*log(1-ro^2)+lgamma((v+2)/2)+lgamma(v/2)-
  2*lgamma((v+1)/2)-((v+2)/2)*(log(1+(qt(pt(rgserie[,1],
p1),v)^2+qt(pt(rgserie[,2],p2),v)^2-
  2*ro*qt(pt(rgserie[,1],p1),v)*qt(pt(rgserie[,2],p2),v))/(v*(

```

```

1-ro^2)))+(v+1)/2)* (log((1+(qt(pt(rgserie[,1],p1),
v)^2)/v)*(1+(qt(pt(rgserie[,2],p2),v)^2)/v))))
der2 <- function(p,p1,eps=1e-4)
{
deriv1 <- function(y) (f(p+eps,y)-f(p,y))/eps
deriv <- (deriv1(p1+eps)-deriv1(p1))/eps
deriv
}
student9<-(der2(par[1],pr[2]))

vgaussian<-0;vfranck<-0;vplackett<-0;vgumbel<-0;vstudent<-0
for (i in 1: n)
{
vgaussian<-vgaussian+(rbind(marg1[i],marg2[i],gaussian1[i])
%%cbind (marg1[i],marg2[i],gaussian1[i]))/n
vfranck<-vfranck+(rbind(marg1[i],marg2[i],franck1[i])
%%cbind (marg1[i],marg2[i],franck1[i]))/ n
vplackett<-vplackett+(rbind(marg1[i],marg2[i],plackett1[i])
%%cbind (marg1[i],marg2[i],plackett1[i]))/ n
vgumbel<-vgumbel+(rbind(marg1[i],marg2[i],gumbel1[i])
%%cbind (marg1[i],marg2[i],gumbel1[i]))/ n
vstudent<-
vstudent+rbind(marg1[i],marg2[i],student[i],student1[i])%%c
bind(marg1[i],marg2[i],student[i],student1[i]))/n
}

a<-0; b<-0;c<-0;d<-0; e<-0;f<-0;g<-0; h<-0;i1<-0;l<-0;m<-
0;n<-0;o<-0; p<-0 ;q<-0;r<-0
for (i in 1: n)
{
a<-a+(cbind(marg11[i],0,0))/ n
b<-b+(cbind(0,marg22[i],0))/ n
c<-c+(cbind(franckp1[i],franckp2[i],franck2[i]))/ n
d<-d+(cbind(marg11[i],0,0))/ n
e<-e+(cbind(0,marg22[i],0))/ n
f<-f+(cbind(gumbelp1[i],gumbelp2[i],gumbel2[i]))/ n
g<-g+(cbind(marg11[i],0,0))/ n
h<-h+(cbind(0,marg22[i],0))/ n

```

```

i1<-i1+(cbind(plackettp1[i],plackettp2[i],plackett2[i]))/ n
l<-l+(cbind(marg1l[i],0,0))/ n
m<-m+(cbind(0,marg22[i],0))/ n
n1<-n1+(cbind(gaussianp1[i],gaussianp2[i],gaussian2[i]))/ n
o<-o+(cbind(marg1l[i],0,0,0))/ n
p<-p+(cbind(0,marg22[i],0,0))/ n
q<-q+(cbind(student8[i],student9[i],student7[i]
,student6[i]))/n
r<-r+(cbind(student4[i],student5[i],student2[i]
,student3[i]))/ n
}
dfranck<-rbind(a,b,c)
dgumbel<-rbind(d,e,f)
dplackett<-rbind(g,h,i1)
dgaussian<-rbind(l,m,n1)
dstudent<-rbind(o,p,q,r)

gstudent<-
sqrt((solve(dstudent%%solve(dstudent))%%t(dstudent))/ n))
ggaussian<-sqrt((solve(dgaussian%%solve(vgaussian)
%%t(dgaussian))/ n))
gfranck<-sqrt((solve(dfranck%%solve(vfranck))%%t(dfranck))/
n))
gplackett<-sqrt((solve(dplackett%%solve(vplackett)
%%t(dplackett))/ n))
ggumbel<-sqrt((solve(dgumbel%%solve(vgumbel))%%t(dgumbel))/
n))

t.Gaussian<-par1/ggaussian[3,3]
t.Student1<-par[1]/gstudent[3,3]
t.Student2<-par[2]/gstudent[4,4]
t.Franck<-par3/gfranck[3,3]
t.Gumbel<-par4/ggumbel[3,3]
t.Plackett<-par2/gplackett[3,3]
t.marg1<-pr[1]/ggaussian[1,1]
t.marg2<-pr[2]/ggaussian[2,2]
p.Gaussian<-1-pnorm(t.Gaussian)
p.Student1<-1-pnorm(t.Student1)

```

```

p.Student2<-1-pnorm(t.Student2)
p.Franck<-1-pnorm(t.Franck)
p.Gumbel<-1-pnorm(t.Gumbel)
p.Plackett<-1-pnorm(t.Plackett)
p.marg1<-1-pnorm(t.marg1)
p.marg2<-1-pnorm(t.marg2)

print("stime e standard error dei parametri delle copule")
print(paste("  Gaussian  ",round(par1,2),"
",round(stderr.gauss[1],4)))
print(paste("  Student1  ",round(par[1],2),"
",round(stderr.stud[1],4)))
print(paste("  Student2  ",round(par[2],2),"
",round(stderr.stud[2],4)))
print(paste("  Franck    ",round(par3,2),"
",round(stderr.fr,4)))
print(paste("  Gumbel    ",round(par4,2),"
",round(stderr.gm,4)))
print(paste("  Plackett  ",round(par2,2),"
",round(stderr.pl,4)))
print(" ")
print("stime e nuovi standard error dei parametri delle
copule e delle marginali")
print(paste("  Gaussian  ",round(par1,2),"
",round(ggaussian[3,3],4)))
print(paste("  Student1  ",round(par[1],2),"
",round(gstudent[3,3],4)))
print(paste("  Student2  ",round(par[2],2),"
",round(gstudent[4,4],4)))
print(paste("  Franck    ",round(par3,2),"
",round(gfranck[3,3],4)))
print(paste("  Gumbel    ",round(par4,2),"
",round(ggumbel[3,3],4)))
print(paste("  Plackett  ",round(par2,2),"
",round(gplackett[3,3],4)))
print(paste("  1° marginale  ",round(pr[1],2),"
",round(ggaussian[1,1],4)))

```

```

print(paste(" 2° marginale ",round(pr[2],2),"
",round(ggaussian[2,2],4)))
print(" ")
print("Test di significatività e relativi p-value per i
parametri delle copule e delle marginali")
print(paste(" Gaussian ",round(t.Gaussian,2),"
",round(p.Gaussian,)))
print(paste(" Student1 ",round(t.Student1,2),"
",round(p.Student1,2)))
print(paste(" Student2 ",round(t.Student2,2),"
",round(p.Student2,2)))
print(paste(" Franck ",round(t.Franck,2),"
",round(p.Franck,2)))
print(paste(" Gumbel ",round(t.Gumbel,2),"
",round(p.Gumbel,2)))
print(paste(" Plackett ",round(t.Plackett,2),"
",round(p.Plackett,2)))
print(paste(" 1° marginale ",round(t.marg1,2),"
",round(p.marg1,2)))
print(paste(" 2° marginale ",round(t.marg1,2),"
",round(p.marg1,2)))
print(" ")
print(paste(" Calcolo del VaR per le diverse copule con
peso",round(pesi,3)))
print(paste(" Gaussian ",round(ga$root,3)))
print(paste(" Student ",round(st$root,3)))
print(paste(" Franck ",round(fr$root,3)))
print(paste(" Gumbel ",round(gu$root,3)))
print(paste(" Pluckett ",round(pl$root,3)))
print(paste("Tradizionale ",round(vartrad,3),"
coeff.correlazione ",round(rot,2)))
}
print(paste(" Stima e standard error per i parametri del
Garch(1,1) della 1° serie "))
print(paste(" alfa0 ",round(parg[1,1],3),"
",round(se.g[1,1],6)," ",round(test.g[1,1],2),"
",round(p.value.g[1,1],2)))

```



```

print(paste("  alfa1  ",round(parg[1,2],3),"
",round(se.g[1,2],6)," ",round(test.g[1,2],2),"
",round(p.value.g[1,2],2)))
print(paste("  beta1  ",round(parg[1,3],3),"
",round(se.g[1,3],6)," ",round(test.g[1,3],2),"
",round(p.value.g[1,3],2)))
print(paste("  "))
print(paste(" Stima e standard error per i parametri del
Garch(1,1) della 2° serie "))
print(paste("  alfa0  ",round(parg[2,1],3),"
",round(se.g[2,1],6)," ",round(test.g[2,1],2),"
",round(p.value.g[2,1],2)))
print(paste("  alfa1  ",round(parg[2,2],3),"
",round(se.g[2,2],6)," ",round(test.g[2,2],2),"
",round(p.value.g[2,2],2)))
print(paste("  beta1  ",round(parg[2,3],3),"
",round(se.g[2,3],6)," ",round(test.g[2,3],2),"
",round(p.value.g[2,3],2)))
print(paste("  "))
print(paste("Risultati del test di Jarque-Bera sui residui
del Garch(1,1)"))
print(paste(" 1° serie  "," statistica di Jarque-Bera
",round(jbt[1],6)," p-value ",round(p.val[1],6)))
print(paste(" 2° serie  "," statistica di Jarque-Bera
",round(jbt[2],6)," p-value ",round(p.val[2],6)))
print(paste("  "))
print(paste("Stima e standard error per i gradi di libertà
delle 2 marginali"))
print(paste(" 1° marginale  "," v1 ",round(pr[1],6),"
",round(stderr.t[1],6)))
print(paste(" 2° marginale  "," v2 ",round(pr[2],6),"
",round(stderr.t[2],6)))
print(paste("  "))
parametri(se[,1],se[,2])
print(paste("  "))
}

```


Bibliografia

Cherubini, U., Luciano, E. & Vecchiato, W. (2004) *Copula methods in finance*. John Wiley & Sons, New York.

Durrleman, V., Nikeghbali, A., & Roncalli, T. (2000) *Copulas for Finance: A Reading Guide and Some Applications*. Groupe de recherche operationelle, Credit Lyonnais, working paper.

Embrechts, P., McNeil, A., Straumann, D. (1999) *Correlation and Dependency in risk management: Properties and Pitfalls*. Department of Mathematik, ETHZ, Zurich, working paper.

Embrechts, P., McNeil, A., Lidskog, F. (2001) *Modelling dependence with Copulas and Applications to risk Management*. Department on Mathematik, ETHZ, Zurich, working paper.

Feller, W. (1971) *An introduction to Probability Theory and Its Applications*, Vol. II. John Wiley & Sons, New York.

Hoeffding, W (1940) "Masstabinvariante Korrelationstheorie", *Schriften des mathematischen Instituts und des Instituts fur Angewandte Mathematik der universitat Berlin*, 5, 179-233.

Joe, H. & Xu, J.J. (1996) *The Estimation Method of Inference Functions for Margins for Multivariate Models*. Dept. of Statistics University of British Columbia, Tech. Rept. 166.

Joe, H. (1997) *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

Kendall, M.G. (1938) "A new measure of rank correlation", *Biometrika*, 30, 81-93

- Kimberling, C.H. (1974) "A probabilistic interpretation of complete monotonicity", *Aequationes Math.*, 10, 152-164.
- Lehmann, E (1966) "Some concepts of dependence", *Ann. Math. Statist.*, 37, 1137-1153.
- Luciano, E. & Marena, M. (2003) "Copulae as a new tool in financial modelling", *Operational Research: An International Journal*, 2, 139-155.
- Nelsen, R.B. (1991) "Copulas and association", in G. Dall'Aglio, S. Kotz & G. Salinetti (eds), *Advances in Probability Distributions with Given Marginals*. Kluwer Academic Publishers, Dordrecht, pp. 51-74.
- Nelsen, R.B. (1999) "An introduction to copulas", *Lecture Notes in Statistics*. Springer-verlag, New York.
- Plackett, R.L. (1965) "A class of bivariate distributions", *J. Amer. Statist. Assoc.*, 60, 516-522.
- Rockinger, M. & Jondeau, E. (2001) *Conditional Dependency of Financial Series: an Application of Copulas*. Working paper.
- Scarsini, M. (1984) "On measures of concordance", *Stochastica*, 8, 201-218
- Schweizer, B. & Wolff, E. (1981) "On non-parametric measures of dependence for random variables", *Ann. Statist.*, 9, 879-885.
- Serfling, R.J. (1980) *Approximation theorems of Mathematical Statistics*. John Wiley & Sons, New York.
- Shao, J. (1999) *Mathematical Statistics*. Springer-Verlag, New York.
- Sklar, A. (1959) "Functions de repartition à n-dimensions et leurs marges", *Pub. Inst. Statist. Univ. Paris*, 8, 229-231.

Sklar, A. (1996) “Random variables, distribution functions, and copulas – a personal look backward and forward”, in L. Rüschendorf, B. Schweizer & M.D. Taylor (eds), *Distribution with fixed marginals and Related topics*. Institute of Mathematical Statistics, Hayward, pp. 1-14.