

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Scienze Statistiche
SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE STATISTICHE
CICLO XXII

**Selezione e stima di dipendenze
multivariate: il caso della *pair-copula
construction* con una sua applicazione
in campo assicurativo**

Direttore della Scuola: Ch.ma Prof.ssa ALESSANDRA SALVAN

Supervisore: Ch.mo Prof. MARCO FERRANTE

Dottorando: DANIELE FRANCESCHI

Riassunto

L'oggetto di questa tesi è l'analisi di una recente tecnica di modellizzazione della dipendenza tra le variabili di un vettore aleatorio che va sotto il nome di *pair-copula construction* (PCC). La peculiarità di questa tecnica è che permette di costruire strutture multivariate anche molto complesse con l'impiego esclusivo di copule bivariate semplici oppure condizionate. Per organizzare le varie decomposizioni si ricorre a delle particolari strutture grafiche dette *Regular vine* (R-vine), composte da una successione di alberi. In particolare in questo lavoro si è privilegiato l'impiego e lo studio dei *Drawable vine* (D-vine) i cui alberi hanno la particolare struttura di cammino. Sono stati discussi sia aspetti teorici legati alla forma semplificata della PCC, sia aspetti pratici come l'implementazione di algoritmi per la stima di tali modelli.

I modelli studiati sono quindi stati applicati ad un problema pratico in campo assicurativo, la determinazione del capitale di rischio attraverso un possibile modello interno alternativo alla formula standard proposta dalla nuova Direttiva comunitaria Solvency II. E' stato quindi costruito un dataset di dati reali multivariati, su cui sono state testate diverse strutture di dipendenza, tra le quali diversi possibili modelli di D-vine. Sono infine stati posti a confronto i requisiti di capitale relativi così ottenuti.

Abstract

The subject of this thesis is the analysis of a recent technique, the pair-copula construction (PCC), for modeling dependence between variables of a random vector. The main property of this technique is that it allows us to build even very complex multivariate structures only using bivariate copulas with conditional cumulative distribution functions as arguments. There is a large number of such decompositions, so we use particular graphical structures (made by a sequence of trees and called regular vines) to organize them. In particular, this work has favored the use and the study of drawable vines (D-vine), that have path structure. We discuss theoretical aspects related to the simplified form of the PCC, as well as the implementation of algorithms for estimating these models.

The developed models are applied to a practical problem in the insurance field. The determination of the risk capital through an internal model can be an alternative to the standard formula proposed by the new EU Directive Solvency II. We construct a multivariate dataset of real data, on which we test different structures of dependence, including several possible models of D-vine. Finally, we compare the risk based capital obtained with these models.

Indice

1	Introduzione	1
1.1	Overview	1
1.2	Principali contributi della tesi	4
2	Copule multivariate	7
2.1	Definizione e proprietà base	7
2.1.1	Le copule	8
2.1.2	Teorema di Sklar	9
2.1.3	I limiti di Frechét-Hoeffding	11
2.1.4	Altre proprietà	13
2.2	Misure di dipendenza	15
2.2.1	Il coefficiente di correlazione lineare	16
2.2.2	Misure di correlazione di rango	17
2.2.3	La dipendenza di coda	21
2.3	Famiglie di copule	23
2.3.1	Le copule ellittiche	23
2.3.2	Le copule Archimedee	29
3	La <i>pair-copula construction</i>	35
3.1	La decomposizione in <i>pair-copule</i>	36
3.2	Proprietà della PCC semplificata	39
3.3	I <i>vines</i>	41
3.3.1	Concetti introduttivi	41
3.3.2	Il legame tra R-vine e PCC	46
3.4	Il caso dei D-vine	47
3.4.1	Simulazione da D-vine	47
3.4.2	Selezione di un modello di D-vine	48
3.5	Simulazione di un D-vine con <i>pair-copule</i> Gaussiane	52
3.5.1	PCC semplificata	54
3.5.2	PCC generale	55
3.5.3	Conclusioni	56

4	L'avvento di Solvency II	57
4.1	La nuova direttiva	57
4.2	La formula <i>standard</i>	60
4.2.1	La formula <i>standard</i> per il <i>Non-Life</i>	62
4.3	Un modello interno per il rischio di sottoscrizione Danni	65
5	Applicazione in campo assicurativo	67
5.1	Analisi preliminari dataset	68
5.1.1	Presentazione dataset	68
5.1.2	Verifica ipotesi di indipendenza	73
5.2	Scelta dei modelli distributivi	74
5.2.1	Il numero degli eventi: <i>frequency</i>	75
5.2.2	Il costo del singolo evento: <i>severity</i>	76
5.2.3	Convoluzione di Montecarlo	81
5.3	Stima degli indicatori di dipendenza	81
5.4	Aggregazione delle cinque classi di rischio	83
5.4.1	Indipendenza	87
5.4.2	Correlazione lineare	88
5.4.3	Struttura D-vine	88
5.4.4	Confronto tra ipotesi di dipendenza	92
5.5	Conclusioni	94
A	Funzioni utili per PCC	97
A.1	Copula Gaussiana bivariata	97
A.2	Copula di sopravvivenza di Clayton bivariata	98
	Bibliografia	99

Elenco delle figure

2.1	Grafici delle copule: a) $W(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$; b) $M(u_1, u_2) = \min(u_1, u_2)$; c) $\Pi(u_1, u_2) = u_1 u_2$	12
2.2	<i>Contour plot</i> delle copule: a) $M(u_1, u_2) = \min(u_1, u_2)$; b) $\Pi(u_1, u_2) = u_1 u_2$; c) $W(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$	13
2.3	1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula Gaussiana e differenti valori di ρ	25
2.4	1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula t con 1 grado di libertà e differenti valori di ρ	28
2.5	1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula t con 10 gradi di libertà e differenti valori di ρ	28
2.6	1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula di Gumbel e differenti valori di θ (2, 5, 10 relativi a τ pari a 0,5, 0,8 e 0,9).	31
2.7	1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula di Clayton e differenti valori di θ (2, 8, 18 relativi a τ pari a 0,5, 0,8 e 0,9).	32
2.8	1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula di Frank e differenti valori di θ (-10, 3, 20).	33
2.9	1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con le copule di Gumbel, di Clayton e di Frank con i valori di θ relativi ai τ di Kendall compatibili con $\rho = 0, 8$ e $\rho = 0, 1$ della copula Gaussiana.	34
3.1	La condizione di prossimità: un arco attorno ad un nodo.	42
3.2	Un esempio di R-vine con 5 variabili. Per ogni arco $e = \{a, b\} \in E_i$, a destra del segno ' ' si trovano gli elementi dell'insieme D_e , mentre a sinistra il primo termine è $C_{e,a}$ ed il secondo è $C_{e,b}$	43
3.3	Un esempio di D-vine con 5 variabili.	44

3.4	Un esempio di C-vine con 5 variabili.	45
3.5	D-vine con 4 variabili e <i>pair-copule</i> Gaussianhe: sotto ad ogni arco è riportato il parametro della copula a cui è associato. . .	53
4.1	Solvency II: struttura a pilastri della normativa.	58
5.1	Ammontare degli eventi avvenuti in ordine cronologico nel periodo 2005-2007	70
5.2	Ammontare degli eventi avvenuti nel periodo 2005-2007 divisi per zone geografiche	71
5.3	Contributo di ciascuna classe di rischio sul costo totale degli eventi atmosferici negli anni 2005-2007	72
5.4	Istogrammi delle distribuzioni del costo degli eventi per singola classe di rischio: a) Agricolo, b) Civile, c) Commercio, d) PMI, e) Religioso	77
5.5	Eventi 2005-2007: grafici di dispersione delle classi di rischio .	82
5.6	Società Cattolica di Assicurazione, Ramo Incendio 2005-2007: <i>Combined Ratio</i> ed <i>Expense Ratio</i>	85
5.7	<i>Risk Based Capital</i> per classe di rischio	87
5.8	a) Ordinamento delle classi di rischio che massimizza la somma dei tau di Kendall; b) Ordinamento delle classi di rischio che minimizza la somma dei tau di Kendall.	89
5.9	<i>Risk Based Capital</i> : confronto tra diverse ipotesi di aggregazione	93

Elenco delle tabelle

3.1	Matrice di Varianza-Covarianza Σ	53
3.2	D-vine con 4 variabili, copule Gaussiane e sotto l'ipotesi di PCC semplificata: confronto tra i parametri delle copule condizionate determinati con la stima sequenziale e con il calcolo esatto.	54
3.3	Matrice di Varianza-Covarianza di 100.000 osservazioni simulate da un D-vine con 4 variabili, copule Gaussiane e sotto l'ipotesi di PCC semplificata.	55
3.4	Matrice di Varianza-Covarianza di 100.000 osservazioni simulate da un D-vine con 4 variabili, copule Gaussiane e senza l'ipotesi di PCC semplificata.	56
4.1	Formula standard QIS5: Matrice di correlazione tra i fattori di rischio.	62
4.2	Formula standard QIS5: Matrice di correlazione dei sottomoduli di rischio del SCR_{nl}	63
5.1	Indici utilizzati per riportare gli importi monetari passati al 2009	69
5.2	Statistiche descrittive: eventi atmosferici 2005-2007	73
5.3	Test delle differenze di segno per testare l'indipendenza	74
5.4	Numero di eventi osservati per classi di rischio negli anni 2005-2007	76
5.5	Distribuzione del costo del singolo evento per classe di rischio: ordine di preferenza dei diversi modelli relativamente alla bontà di adattamento espressa dai test del χ^2 e di Kolmogorov-Smirnov.	80
5.6	Costo degli eventi per classi di rischio: parametri della distribuzione Lognormale	80
5.7	Indice di correlazione lineare e τ di Kendall per classi di rischio	83
5.8	Momenti del costo aggregato e requisiti di capitale relativi per classi di rischio	86

5.9	Parametri delle <i>pair-copule</i> Gaussianie relativi alle strutture di D-vine A e B.	91
5.10	Parametri delle <i>pair-copule</i> di sopravvivenza di Clayton relativi alle strutture di D-vine A e B.	91
5.11	Momenti della distribuzione del costo aggregato degli eventi secondo diverse strutture di dipendenza ($A=REL-CIV-AGR-PMI-COM$, $B=AGR-COM-REL-PMI-CIV$)	92

Capitolo 1

Introduzione

1.1 Overview

Nel campo assicurativo la teoria classica del rischio ignorava elementi come l'interesse e le spese e si assumeva come unica variabile aleatoria il costo dei sinistri. La moderna teoria del rischio ha introdotto e sviluppato modelli di simulazione della gestione di una compagnia di assicurazioni, i cosiddetti *management simulation models*, in grado di rappresentare le molteplici interdipendenze esistenti e valutare, a seguito di reiterate simulazioni, non solo il valore atteso, ma anche la variabilità, e quindi la rischiosità, del processo economico e patrimoniale legato all'impresa.

Attualmente si stanno sviluppando a livello internazionale i nuovi *framework* di vigilanza (tra cui Solvency II), che tendono verso la definizione di *ratio* patrimoniali calibrati sul profilo di rischio delle Compagnie (*Risk-Based Capital requirement*) e i nuovi principi contabili internazionali¹ che contribuiscono a modificare concetti consolidati con l'introduzione ad esempio delle riserve sinistri stocastiche, del *fair value*, ecc.

In tale contesto in continuo mutamento, il progetto Solvency II ha come obiettivo la riforma delle attuali regole di solvibilità nell'Unione Europea. Lo scopo è l'introduzione di un regime di solvibilità che non si limiti all'analisi dei soli rischi tecnici, ma che consideri l'intero *risk profile*, che sia flessibile, non eccessivamente complesso e compatibile con i metodi di controllo delle autorità di vigilanza. La sua importanza è accentuata dal fatto che i nuovi standard contabili IAS diminuiranno il livello di prudenza in molte nazioni.

Rispetto all'attuale normativa vigente, Solvency II chiederà alle compagnie di assicurazione europee la determinazione di un requisito di capitale più

¹Le nuove regole di contabilità internazionale vanno sotto il nome di principi IAS, che sta per *International Accounting Standards*.

aderente alle caratteristiche di ciascuna impresa, desumibile da due possibili approcci: uno basato su una formula standard e uno più avanzato che utilizza dei modelli interni.

Con Solvency II il business *non-life* di un'impresa di assicurazione viene gestito principalmente tramite 12 *Line of Business* (LoB), citiamo a titolo di esempio Incendio e altri danni a beni, R.C. Auto, R.C. Generale... Il tipo di aggregazione delle LoB proposto dalla formula standard è quello della correlazione lineare. In ottica di un possibile modello interno, che colga maggiormente le peculiarità delle dipendenze tra LoB, si possono utilizzare degli strumenti più sofisticati e in particolare in questo lavoro si è ricorso all'utilizzo della *pair-copula construction* (di seguito PCC).

La teoria della PCC è stata introdotta circa dieci anni fa con i lavori di Bedford e Cooke (2001, 2002), che a loro volta si basavano su Joe (1996), ma è in anni più recenti che ha suscitato maggior interesse. Si inserisce infatti tra le possibili soluzioni alla sempre più crescente richiesta di modelli multivariati non-gaussiani proveniente da diversi campi scientifici, come per esempio la finanza (Cherubini et al. 2004), le scienze ambientali (Salvadori et al. (2007)) e la medicina (Beaudoin e Lakhel-Chaieb (2008)).

La PCC riguarda la costruzione di strutture multivariate attraverso il solo impiego di copule bivariate, le cosiddette *pair-copule*. Il modello si basa su una decomposizione della densità multivariata in una serie di *pair-copule*, che possono essere applicate alle variabili originali o alle loro funzioni di distribuzione condizionate o non condizionate.

L'ipotesi fondamentale che sta alla base della PCC è che tutte le copule condizionate coinvolte siano in una forma 'semplificata' e cioè che dipendano dalle variabili condizionanti esclusivamente attraverso le due distribuzioni condizionate che hanno come argomento. Tuttavia anche quando l'assunzione di forma semplificata è lontano dall'essere rispettata è stato verificato da Hobaek Haff et al. (2010) che la PCC rappresenta comunque una buona approssimazione della distribuzione.

Per distribuzioni n -variate il numero di possibili decomposizioni può essere considerevole e a tale proposito Bedford e Cooke (2001, 2002) hanno introdotto una struttura grafica utile a organizzarle, i *regular vine* (R-vine). Questi sono formati da una successione ordinata di alberi, dove a ciascun arco è assegnata una diversa *pair-copula*; in questo modo ogni modello corrisponde ad una diversa decomposizione della distribuzione, che è completamente definita dagli $n(n-1)/2$ archi e dalle densità marginali di ciascuna variabile.

La classe dei R-vine è ancora molto vasta e per questo, almeno inizialmente (Aas et al. (2009)), sono state studiate due sottoclassi di essi: i

canonical vine (C-vine) dalla caratteristica struttura a stella degli alberi e i *drawable vine* (D-vine) i cui alberi hanno una struttura di cammino (*path*).

Implementare un R-vine è una procedura che consiste essenzialmente di tre punti: scegliere la struttura del R-vine, definire le famiglie di *pair-copule* da utilizzare e infine stimarne i parametri.

Focalizzando maggiormente la nostra attenzione sulla stima di un D-vine, definirne la struttura significa decidere l'ordine delle variabili dell'albero di primo livello, perchè gli altri sono tutti subordinati. Intuitivamente la scelta dovrebbe ricadere sulla combinazione che massimizza le dipendenze tra le variabili e con questo intento è stato recentemente proposto in Dissmann et al. (2011) un possibile metodo di selezione per la classe generica dei R-vine. Per un D-vine determinare l'ordinamento ottimo equivale a risolvere il problema del Commesso Viaggiatore per il quale non esiste un algoritmo efficiente, quindi l'unico metodo di soluzione è rappresentato dall'enumerazione totale.

Per quanto riguarda la selezione delle famiglie di *pair-copule*, queste sono scelte indipendentemente l'una dall'altra e può essere per esempio utilizzato il test di verifica delle informazioni di Akaike (Akaike (1973)).

Infine per la stima dei parametri delle *pair-copule* il metodo maggiormente utilizzato in letteratura è quello della massima verosimiglianza. Siccome il numero di variabili rispetto cui ottimizzare è come minimo $n(n - 1)/2$, è importante dotarsi di buoni valori di partenza per l'ottimizzazione. A questo scopo risulta utile la procedura della stima sequenziale suggerita in Aas et al. (2009).

In letteratura il metodo della PCC non è l'unico a fondarsi sulla teoria delle copule bivariate (per esempio in Embrechts et al. (2003) e Savu e Trede (2006) si propone un'estensione multivariata delle copule Archimedee bivariate); tuttavia, riteniamo che la PCC rappresenti attualmente il modo più flessibile e intuitivo di estendere le copule bivariate alle dimensioni più elevate.

Per l'applicazione della teoria della PCC ai dati reali è necessario disporre di numerose osservazioni congiunte per poter coglierne le dipendenze e per questo motivo in letteratura si trovano soprattutto studi relativi a serie storiche finanziarie che sono facilmente reperibili. Ha rappresentato un grande sforzo per questa tesi il reperimento dei dati provenienti dal mondo assicurativo e adeguati all'implementazione della PCC.

Presentiamo in breve la struttura della tesi. Nel capitolo 2 viene introdotto il concetto di copula e se ne esplorano le caratteristiche principali ponendo particolare attenzione al legame con i vari indicatori di dipendenza. Il capitolo 3 è dedicato alla metodologia della PCC abbinata alle diverse strutture grafiche di R-vine; viene inoltre discussa l'ipotesi fondamentale sottostante

l'utilizzo della PCC, cioè l'assunzione di forma semplificata delle *pair-copule*. Il capitolo 4 fornisce una panoramica generale sulla direttiva Solvency II, focalizzando l'attenzione sul calcolo dei requisiti di capitale, presentando anche gli aspetti teorici di un possibile modello interno per il business Danni. Nel capitolo 5 viene presentato il lavoro svolto per costruire un dataset di dati reali provenienti dal mondo assicurativo; per tali dati è stata quindi implementata la PCC al fine di valutare il requisito di capitale nell'ottica di un possibile modello interno.

1.2 Principali contributi della tesi

Di seguito vengono riassunti per punti i principali risultati ottenuti in questo lavoro:

- In Aas et al. (2009) è proposto un algoritmo per la simulazione di un campione da un D-vine avente le variabili distribuite come delle uniformi, in questo lavoro l'algoritmo viene generalizzato per la simulazione di un campione da un D-vine avente distribuzioni marginali qualsiasi.
- Implementando l'algoritmo descritto al punto precedente con dati simulati da una normale multivariata di dimensione 4 è stato possibile verificare l'ipotesi fondamentale sottostante l'utilizzo della PCC e cioè che le copule condizionate sono indipendenti dalle variabili condizionanti eccetto che attraverso le distribuzioni condizionate, cioè le variabili condizionanti non compaiono come un parametro esplicito delle copule. Si può concludere ciò osservando che una calibrazione empirica dei parametri di queste copule condizionate risulta indipendente dai valori assunti dalle variabili condizionanti.
- Relativamente al problema assicurativo descritto precedentemente, al fine di confrontare diversi metodi di modellizzazione delle dipendenze è necessario disporre di informazioni relative a eventi che hanno colpito simultaneamente diverse LoB. Data l'estrema difficoltà di reperire dati di questo tipo per line of business, ci siamo concentrati su una sottoclasse della LoB Incendio e Altri Danni ai Beni. In particolare è stato costruito un dataset di dati reali provenienti da una Compagnia assicurativa italiana (Società Cattolica di Assicurazione) contenente il costo degli eventi atmosferici (trombe d'aria, grandinate, bufere) al lordo della riassicurazione e relativi agli anni 2005-2009 con vista al 31.12.2009. Per le successive analisi sono stati tuttavia ritenuti stabili i soli dati relativi al periodo 2005-2007. Il dataset è stato suddiviso in classi di

rischio identificate in base al tipo di prodotto assicurativo sottostante (Agricolo, Civile, Commercio, Piccola e Media Impresa, Religioso).

- Dopo aver fatto una serie di analisi preliminari sul dataset in oggetto, abbiamo stimato il costo aggregato degli eventi per ciascuna classe di rischio e il relativo capitale di rischio. Con l'obiettivo di proporre un modello interno alternativo alla formula standard, si è scelto di confrontare il capitale di rischio aggregato ottenuto applicando diverse strutture di dipendenza ed in particolare sono state testate l'ipotesi di indipendenza, l'aggregazione mediante una matrice di correlazione lineare (come proposto dalla formula standard di Solvency II) e la struttura di D-vine con diversi ordini delle variabili.

Capitolo 2

Copule multivariate

La distribuzione congiunta di ogni vettore aleatorio contiene implicitamente sia una descrizione del comportamento marginale di ogni singola variabile sia una descrizione della loro struttura di dipendenza. In un certo senso lo scopo delle copule è quello di isolare ed identificare la descrizione della struttura di dipendenza. Sicuramente in ambito multivariato non è l'unico modo di indagare sulla dipendenza, ma a nostro avviso è forse quello più naturale quando si opera in un contesto di distribuzioni statiche piuttosto che in un ambiente dinamico.

Nel paragrafo 2.1 forniremo una descrizione matematica delle copule ponendo particolare attenzione nella discussione al teorema di Sklar, che rappresenta uno dei risultati più importanti in questo contesto dato il largo uso che se ne fa nelle applicazioni.

L'oggetto di discussione del paragrafo 2.2 saranno tre diverse misure di dipendenza legate alle copule. In particolare analizzeremo il coefficiente di correlazione lineare, gli indicatori di correlazione di rango e i coefficienti di dipendenza di coda.

Verranno infine presentate nel paragrafo 2.3 due particolari famiglie di copule: le copule ellittiche e le copule Archimedee. Completeranno le osservazioni analitiche alcuni grafici di simulazioni maggiormente esplicativi.

Per maggiori dettagli riguardo alle copule e per le dimostrazioni complete delle proprietà di seguito presentate rimandiamo a Nelsen (1999) e McNeil et al. (2006).

2.1 Definizione e proprietà base

Prima di definire matematicamente cosa sia una copula, diamo un rapido accenno di quale sia il nostro obiettivo in pratica. Consideriamo un vettore

aleatorio (X_1, X_2, \dots, X_n) con distribuzioni marginali $F(x_1), F(x_2), \dots, F(x_n)$ e distribuzione congiunta $H(x_1, x_2, \dots, x_n)$, dove $F(x_i) = P[X_i \leq x_i]$ e $H(x_1, x_2, \dots, x_n) = P[X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n]$. A ciascuna n -upla di numeri reali (x_1, x_2, \dots, x_n) possiamo associare $n + 1$ numeri, $F(x_1), F(x_2), \dots, F(x_n)$ e $H(x_1, x_2, \dots, x_n)$, ciascuno appartenente all'intervallo $I = [0, 1]$. Si crea quindi una corrispondenza tra un punto $(F(x_1), F(x_2), \dots, F(x_n))$ in I^n e un numero $H(x_1, x_2, \dots, x_n)$ in I . Questo legame che assegna un valore di distribuzione congiunta a una n -upla ordinata di valori di distribuzioni univariate è in realtà una funzione. Le funzioni di questo tipo sono le copule.

2.1.1 Le copule

Definizione 2.1. *Una copula n -dimensionale è una funzione di distribuzione definita su $[0, 1]^n$ con distribuzioni marginali uniformi standard. Quindi una funzione $C(\mathbf{u}) = C(u_1, \dots, u_n)$ è una copula n -dimensionale se soddisfa le seguenti proprietà:*

1. Per ogni $(u_1, \dots, u_n) \in [0, 1]^n$ si ha che $C(u_1, \dots, u_n) = 0$ se almeno una coordinata è nulla;
2. $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ per ogni $i \in \{1, \dots, n\}$ e $u_i \in [0, 1]$;
3. Vale la disuguaglianza triangolare. Per ogni $(a_1, \dots, a_n), (b_1, \dots, b_n) \in [0, 1]^n$ con $a_i \leq b_i$ si ha

$$\sum_{i_1=1}^2 \cdots \sum_{i_n=1}^2 (-1)^{i_1 + \dots + i_n} C(u_{1i_1}, \dots, u_{ni_n}) \geq 0, \quad (2.1)$$

dove $u_{j1} = a_j$ e $u_{j2} = b_j$ per ogni $j \in \{1, \dots, n\}$.

Osservazione 2.1. *La condizione che una copula $C(\mathbf{u})$ sia crescente rispetto a ciascuna sua componente implica $C(\mathbf{u}) = 0$ se almeno una coordinata di \mathbf{u} è nulla.*

Un'utile proprietà delle copule riguarda le sue marginali:

Osservazione 2.2. *Se C è una copula n -dimensionale, allora per ogni $2 \leq k < n$ le marginali k -dimensionali di C sono a loro volta copule.*

Presentiamo infine un utile teorema riguardante le derivate parziali delle copule. Lo enunciamo per copule bivariate, ma la proprietà è facilmente estendibile al caso n -dimensionale. Per una dimostrazione completa si veda McNeil et al. (2006).

Teorema 2.1. *Sia $C(u_1, u_2)$ una copula, allora per ogni $u_1 \in [0, 1]$ la derivata parziale $\partial C/\partial u_2$ esiste quasi per tutti gli u_2 , e per tale u_1 e u_2 vale*

$$0 \leq \frac{\partial}{\partial u_2} C(u_1, u_2) \leq 1.$$

Analogamente per ogni $u_2 \in [0, 1]$ la derivata parziale $\partial C/\partial u_1$ esiste quasi per tutti gli u_1 , e per tale u_1 e u_2 vale

$$0 \leq \frac{\partial}{\partial u_1} C(u_1, u_2) \leq 1.$$

Inoltre le funzioni $u_1 \mapsto \partial C(u_1, u_2)/\partial u_2$ e $u_2 \mapsto \partial C(u_1, u_2)/\partial u_1$ sono definite e non-decrescenti quasi ovunque su $[0, 1]$.

2.1.2 Teorema di Sklar

Il teorema di Sklar rappresenta probabilmente il risultato più importante riguardo alle copule ed è usato praticamente in tutte le loro applicazioni. Questo elegante teorema dimostra innanzitutto che le distribuzioni multivariate contengono copule ed inoltre che le copule possono essere utilizzate congiuntamente a distribuzioni univariate per costruire distribuzioni multivariate.

Teorema 2.2 (Teorema di Sklar). *Sia H una funzione di distribuzione n -dimensionale con marginali F_1, \dots, F_n . Allora esiste una n -copula $C : [0, 1]^n \rightarrow [0, 1]$ tale che per ogni $x_1, \dots, x_n \in \bar{\mathbb{R}} = [-\infty, \infty]$,*

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)). \quad (2.2)$$

Se le marginali F_1, \dots, F_n sono tutte continue, allora C è unica; altrimenti C è univocamente determinata su $\text{Ran}F_1 \times \dots \times \text{Ran}F_n$, dove $\text{Ran}F_i = F_i(\bar{\mathbb{R}})$ rappresenta il codominio di F_i . Al contrario, se C è una n -copula e F_1, \dots, F_n sono funzioni di distribuzione, allora la funzione H definita sopra è una funzione di distribuzione n -dimensionale di marginali F_1, \dots, F_n .

Quindi il teorema di Sklar dimostra che per le distribuzioni multivariate è possibile separare le marginali univariate dalla struttura di dipendenza multivariata, quest'ultima espressa mediante una copula.

Notiamo che la formulazione del teorema è invertibile e a tale scopo definiamo l'inversa generalizzata di una funzione di distribuzione F come $F^{-1}(t) = \inf\{x \in \mathbb{R} | F(x) \geq t\}$ per ogni $t \in [0, 1]$, dove per convenzione $\inf \emptyset = -\infty$.

Corollario 2.3. *Sia H una funzione di distribuzione n -dimensionale con distribuzioni marginali continue F_1, \dots, F_n e copula C (dove C soddisfa la condizione (2.2)). Allora per ogni $\mathbf{u} \in [0, 1]^n$ risulta*

$$C(u_1, \dots, u_n) = H(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)). \quad (2.3)$$

Esempio 2.1. *Siano rispettivamente Φ e Φ_R^n le funzioni di distribuzione di una normale standard e di una normale standard multivariata con matrice di correlazione lineare R . Allora*

$$C(u_1, \dots, u_n) = \Phi_R^n(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))$$

è la n -copula normale o Gaussiana.

Le formule (2.2) e (2.3) sono molto importanti quando si lavora con le copule. Mentre la prima, come abbiamo già notato, dimostra che una distribuzione congiunta H si forma raggruppando insieme le distribuzioni marginali attraverso una copula C , la seconda mostra che le copule possono essere estratte dalle distribuzioni multivariate con marginali continue. Inoltre la (2.3) mostra come le copule esprimono la dipendenza in una scala di quantili, infatti $C(u_1, \dots, u_n)$ è la probabilità congiunta che X_i stia sotto al suo u_i -quantile, per ogni $i = 1, \dots, n$.

Un'ulteriore conseguenza del teorema di Sklar è che si può definire in modo naturale il concetto di copula di una distribuzione, nel caso in cui le marginali siano continue.

Definizione 2.2. *Sia \mathbf{X} un vettore aleatorio con funzione di distribuzione congiunta H e distribuzioni marginali F_1, \dots, F_n continue, allora la copula di H (o di \mathbf{X}) è la funzione di distribuzione C di $(F_1(X_1), \dots, F_n(X_n))$.*

Concludiamo questa sezione presentando l'utile proprietà che una copula di una distribuzione risulta invariante rispetto a trasformazioni delle marginali strettamente crescenti.

Teorema 2.4. *Sia \mathbf{X} un vettore aleatorio n -dimensionale con distribuzioni marginali continue e copula C , siano poi T_1, \dots, T_n delle funzioni strettamente crescenti. Allora il vettore aleatorio $(T_1(X_1), \dots, T_n(X_n))$ ha anch'esso copula C .*

Questa proprietà di invarianza unitamente col teorema di Sklar permette di interpretare la copula di una distribuzione come un modo molto naturale di rappresentare la struttura di dipendenza di una distribuzione multivariata, ovviamente sotto l'ipotesi di marginali continue.

2.1.3 I limiti di Fréchet-Hoeffding

La disuguaglianza nota come limiti di Fréchet-Hoeffding è una delle proprietà fondamentali delle copule.

Consideriamo le funzioni M^n , Π^n e W^n definite su $[0, 1]^n$ nel seguente modo:

$$M^n(\mathbf{u}) = \min(u_1, \dots, u_n),$$

$$\Pi^n(\mathbf{u}) = u_1 \cdots u_n,$$

$$W^n(\mathbf{u}) = \max(u_1 + \cdots + u_n - 1, 0).$$

Le funzioni M^n e Π^n sono n -copule per ogni valore $n \geq 2$, mentre W^n lo è solo nel caso $n = 2$, come dimostrato in Embrechts et al. (2003).

Il seguente teorema è noto in letteratura come limiti di Fréchet-Hoeffding¹.

Teorema 2.5. *Sia C una qualsiasi n -copula, allora per ogni $\mathbf{u} \in [0, 1]^n$*

$$W^n(\mathbf{u}) \leq C(\mathbf{u}) \leq M^n(\mathbf{u}). \quad (2.4)$$

W^n e M^n si definiscono, rispettivamente, *limite inferiore* e *limite superiore di Fréchet-Hoeffding*.

Come conseguenza del teorema di Sklar la disuguaglianza (2.4) può essere espressa anche in termini delle funzioni di distribuzione.

Corollario 2.6. *Siano X_1, \dots, X_n variabili aleatorie con funzione di distribuzione congiunta H e marginali F_1, \dots, F_n , allora per ogni $\mathbf{x} = (x_1, \dots, x_n) \in \bar{\mathbb{R}}^n$*

$$\max(F_1(x_1) + \cdots + F_n(x_n) - 1, 0) \leq H(\mathbf{x}) \leq \min(F_1(x_1), \dots, F_n(x_n)).$$

Sebbene W^n non sia una copula per $n \geq 3$, il limite inferiore della disuguaglianza (2.4) è il ‘migliore possibile’ nel senso che per ogni $n \geq 3$ e per ogni $\mathbf{x} \in [0, 1]^n$, esiste una copula C tale che $C(\mathbf{x}) = W^n(\mathbf{x})$ (Nelsen (1999) pag. 42).

I limiti di Fréchet-Hoeffding ci permettono inoltre di introdurre i concetti di co-monotonia e contro-monotonia.

Definizione 2.3. *Due variabili aleatorie X_1 e X_2 si dicono co-monotone o contro-monotone se il vettore (X_1, X_2) ha come copula, rispettivamente, $M(x_1, x_2)$ o $W(x_1, x_2)$.*

¹Una dimostrazione completa la si può trovare in Nelsen (1999).

La copula Π^n è detta *copula prodotto* e la sua peculiarità può essere riassunta nel seguente risultato:

Teorema 2.7. *Sia (X_1, \dots, X_n) un vettore di variabili aleatorie con funzioni di distribuzione marginali continue e copula C , allora X_1, \dots, X_n sono indipendenti se e solo se $C = \Pi^n$.*

La dimostrazione è una diretta conseguenza del Teorema di Sklar e del fatto che X_1, \dots, X_n sono indipendenti se e solo se la loro distribuzione congiunta si fattorizza nel prodotto delle marginali.

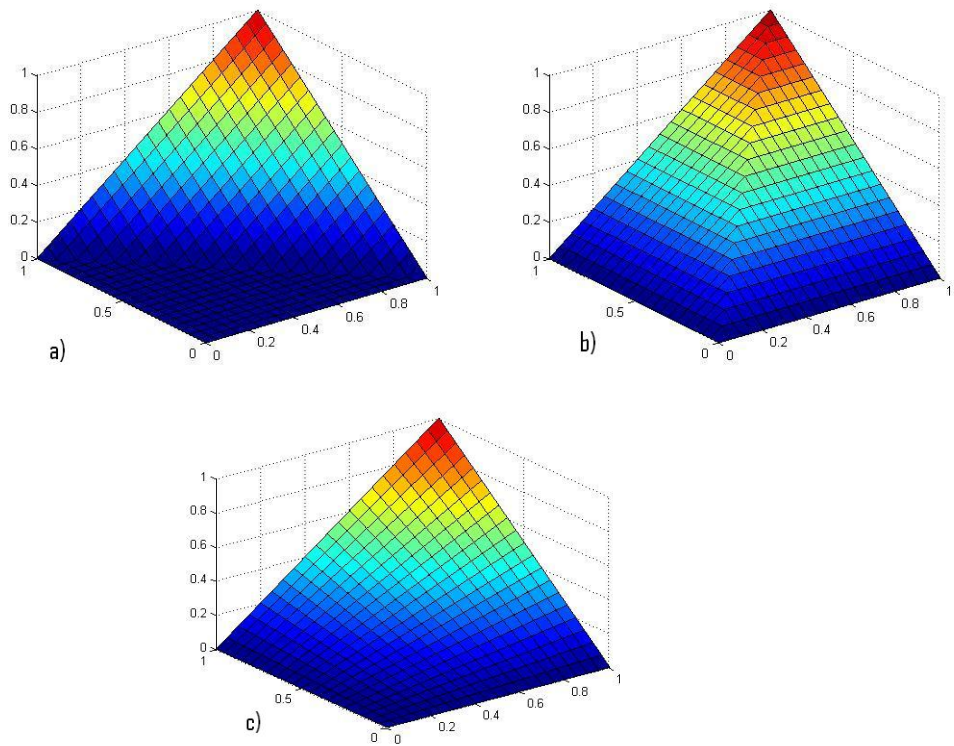


Figura 2.1: Grafici delle copule: a) $W(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$; b) $M(u_1, u_2) = \min(u_1, u_2)$; c) $\Pi(u_1, u_2) = u_1 u_2$.

Vogliamo dare ora una descrizione della forma del grafico di una copula, limitandoci a discutere il caso bivariato. In generale, il grafico di una qualsiasi copula bivariata è una superficie contenuta nel cubo unitario $[0, 1]^3$ il cui bordo è un quadrilatero non piano di vertici $(0, 0, 0)$, $(1, 0, 0)$, $(1, 1, 1)$ e $(0, 1, 0)$. Dal teorema 2.5 consegue che il grafico giace tra i grafici dei limiti di Fréchet-Hoeffding e a tal fine abbiamo riportato in figura 2.1 i grafici

delle copule $M(u_1, u_2)$ e $W(u_1, u_2)$, a cui accostiamo il grafico della copula prodotto $\Pi(u_1, u_2)$ che è una porzione del paraboloide iperbolico $u_3 = u_1 u_2$.

Un modo alternativo di presentare il grafico di una copula è attraverso il suo *contour plot*, cioè tramite il grafico delle sue curve di livello. Un tale grafico è dato dall'insieme in $[0, 1]^2$ delle linee $C(u_1, u_2) = k$, per certi valori della costante $k \in [0, 1]$. In figura 2.2 presentiamo i *contour plot* delle copule M , W e Π , per i quali facciamo notare la non necessità di indicare il valore della curva di livello, poichè le condizioni sul bordo $C(1, t) = C(t, 1) = t$ permettono già di individuare il valore della costante k .

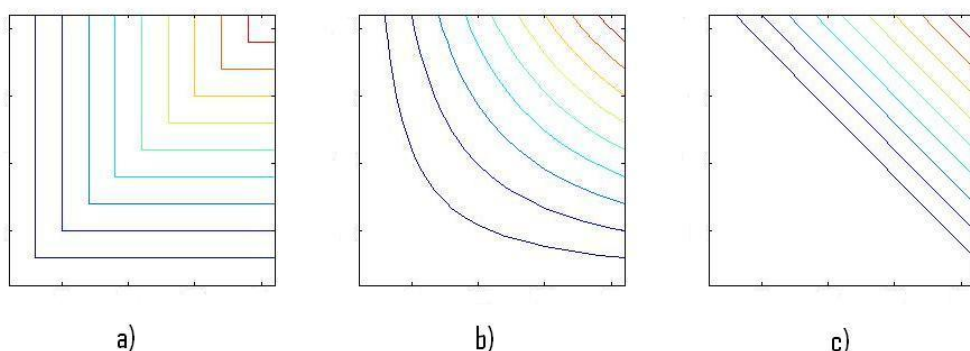


Figura 2.2: *Contour plot* delle copule: a) $M(u_1, u_2) = \min(u_1, u_2)$; b) $\Pi(u_1, u_2) = u_1 u_2$; c) $W(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$.

2.1.4 Altre proprietà

Ricordiamo brevemente che data una funzione di distribuzione H , si definisce funzione di distribuzione di sopravvivenza la funzione \bar{H} per la quale vale $\bar{H} = 1 - H$. Premesso ciò, è possibile estendere la relazione (2.2) del teorema di Sklar anche alle funzioni di distribuzione di sopravvivenza.

Sia \mathbf{X} un vettore aleatorio n -dimensionale con funzione di distribuzione congiunta H , distribuzioni marginali H_1, \dots, H_n e distribuzioni marginali di sopravvivenza $\bar{H}_1, \dots, \bar{H}_n$. Definiamo *copula di sopravvivenza* la copula \hat{C} che soddisfa la relazione

$$\bar{H}(x_1, \dots, x_n) = \hat{C}(\bar{H}_1(x_1), \dots, \bar{H}_n(x_n)). \quad (2.5)$$

Nel caso in cui le H_1, \dots, H_n siano continue, allora la (2.5) diventa

$$\begin{aligned} \bar{H}(x_1, \dots, x_n) &= P(X_1 > x_1, \dots, X_n > x_n) \\ &= P(1 - H_1(X_1) \leq \bar{H}_1(x_1), \dots, 1 - H_n(X_n) \leq \bar{H}_n(x_n)), \end{aligned}$$

dalla quale segue che \widehat{C} è la funzione di distribuzione di $\mathbf{1}-\mathbf{U}$, dove $\mathbf{U} := (H_1(X_1), \dots, H_n(X_n))$. Per questo motivo, in generale, la copula di sopravvivenza di una copula C indica la distribuzione di $\mathbf{1}-\mathbf{U}$ quando C è la distribuzione di \mathbf{U} .

Bisogna prestare attenzione a non confondere la copula di sopravvivenza \widehat{C} con la funzione di sopravvivenza di una copula, che peraltro non è nemmeno una copula. Dato che una copula C non è altro che una distribuzione multivariata, si può definire la sua funzione di sopravvivenza, che indichiamo con \bar{C} , e per la quale vale

$$\begin{aligned}\bar{C}(u_1, \dots, u_n) &= P(U_1 > u_1, \dots, U_n > u_n) \\ &= P(1 - U_1 \leq 1 - u_1, \dots, 1 - U_n \leq 1 - u_n) \\ &= \widehat{C}(1 - u_1, \dots, 1 - u_n).\end{aligned}$$

Sussiste un'utile relazione tra una copula e la sua copula di sopravvivenza, che nel caso bivariato è

$$\widehat{C}(1 - u_1, 1 - u_2) = 1 - u_1 - u_2 + C(u_1, u_2). \quad (2.6)$$

Oltre alla copula di sopravvivenza si possono definire altre funzioni a partire dalle copule, come la co-copula e il duale di una copula, di cui forniamo una definizione nel caso bivariato.

Definizione 2.4. *Assegnata una copula C e un vettore $\mathbf{u} \in [0, 1]^n$, si definiscono co-copula la funzione*

$$C^*(\mathbf{u}) = 1 - C(\mathbf{1}-\mathbf{u})$$

e duale di una copula la funzione

$$\tilde{C}(\mathbf{u}) = \sum_{i=1}^n u_i - C(\mathbf{u}).$$

Nè la co-copula nè il duale di una copula sono copule, infatti entrambe non soddisfano tutte le proprietà richieste. Per esempio per la co-copula bivariata si ha che in generale

$$C^*(u_1, 0) = 1 - C(1 - u_1, 1) = u_1 \neq 0,$$

per cui non risulta soddisfatta la prima proprietà della definizione 2.1. Invece per una prova che il duale di una copula non è una copula rimandiamo al lemma 6.4.2 di Schweizer e Sklar (1983).

Vediamo ora un caso in cui si possono utilizzare le funzioni di sopravvivenza delle copule per definire una regola di ordinamento.

Definizione 2.5. Siano C_1 e C_2 due copule, diciamo che C_1 è minore di C_2 ($C_1 \prec C_2$) se

$$C_1(\mathbf{u}) \leq C_2(\mathbf{u}) \quad e \quad \bar{C}_1(\mathbf{u}) \leq \bar{C}_2(\mathbf{u})$$

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

Nel caso bivariato queste due condizioni si equivalgono, infatti

$$\begin{aligned} \bar{C}_1(u_1, u_2) \leq \bar{C}_2(u_1, u_2) &\Leftrightarrow 1 - u_1 - u_2 + C_1(u_1, u_2) \leq 1 - u_1 - u_2 + C_2(u_1, u_2) \\ &\Leftrightarrow C_1(u_1, u_2) \leq C_2(u_1, u_2). \end{aligned}$$

Siccome non tutte le copule sono confrontabile secondo questo metodo, l'ordinamento che ne deriva è un ordinamento parziale che è chiamato *ordinamento di concordanza*.

Un'ultima considerazione riguarda i limiti di Fréchet-Hoeffding, infatti rispetto a questo ordinamento il limite inferiore W^2 è minore di tutte le altre copule bivariate, mentre tutte le copule n -dimensionali sono minori del limite superiore M^n .

2.2 Misure di dipendenza

In questa sezione analizzeremo tre diversi tipi di misure di dipendenza: il coefficiente di correlazione lineare, la correlazione di rango e il coefficiente di dipendenza di coda. Sebbene la natura e le proprietà di ciascuna di queste dipendenze siano differenti tra loro, ognuna di esse offre una misura scalare di dipendenza per una generica coppia di variabili aleatorie.

La correlazione è sicuramente uno strumento importante ed utile, ma si dimostra appropriato principalmente in ambito di modelli gaussiani multivariati o più in generale per distribuzioni ellittiche².

Gli altri due tipi di misure di dipendenza sono misure 'basate sulle copule', nel senso che queste, diversamente dalla correlazione lineare, dipendono direttamente dalla forma della distribuzione bivariata e non dal comportamento delle rispettive marginali; questa caratteristica le rende particolarmente adatte alla parametrizzazione delle copule.

²Fanno parte della famiglia delle distribuzioni ellittiche la distribuzione Normale multivariata e la t di Student multivariata. Per maggiori dettagli su questa classe di distribuzioni rimandiamo alla sezione 2.3.1.

2.2.1 Il coefficiente di correlazione lineare

Definizione 2.6. Sia (X, Y) un vettore di variabili aleatorie di varianza positiva e finita. Si definisce coefficiente di correlazione lineare la quantità

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}, \quad (2.7)$$

dove $Cov(X, Y) = E(XY) - E(X)E(Y)$ è la covarianza di (X, Y) , e $Var(X)$ e $Var(Y)$ sono le varianze di X e di Y .

La correlazione lineare è una misura di dipendenza lineare che assume valori nell'intervallo $[-1, 1]$. È uno strumento molto popolare, ma spesso porta a interpretazioni errate della misura di dipendenza. Il suo ampio utilizzo è dovuto principalmente alla sua facilità di calcolo, ma se escludiamo le distribuzioni ellittiche, il coefficiente di correlazione evidenzia diversi limiti che ora andremo a discutere.

Innanzitutto se due variabili aleatorie X_1 e X_2 sono indipendenti, allora risulta $\rho(X_1, X_2) = 0$, tuttavia il viceversa non vale in generale se non nel caso gaussiano.

Esempio 2.2. Le variabili $X_1 \sim N(0, 1)$ e $X_2 = (X_1)^2$ hanno coefficiente di correlazione lineare nullo, ma sono chiaramente strettamente dipendenti.

Un'altra semplice, ma rilevante, osservazione è che il coefficiente di correlazione è definito solo per variabili aleatorie aventi varianza finita. Questa restrizione non è sicuramente ideale per una misura di dipendenza per esempio nel caso in cui si considerino distribuzioni con code pesanti, per le quali la varianza è invece infinita.

Un'altra proprietà del coefficiente di correlazione è che per $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, con $\alpha, \gamma > 0$ risulta

$$\rho(\alpha X + \beta, \gamma Y + \delta) = \rho(X, Y),$$

per cui la correlazione è invariante rispetto a trasformazioni lineari strettamente crescenti. Un altro limite che il coefficiente di correlazione presenta è che non è invariante rispetto a trasformazioni non lineari strettamente crescenti $T : \mathbb{R} \rightarrow \mathbb{R}$, per le quali in generale abbiamo

$$\rho(T(X), T(Y)) \neq \rho(X, Y).$$

Oltre agli aspetti fino a qui considerati riteniamo che i due aspetti maggiormente rilevanti dal punto di vista applicativo che sostanzialmente limitano il campo di applicazione del coefficiente di correlazione lineare alle distribuzioni ellittiche sono i seguenti³:

- **Limite 1:** assegnate le distribuzioni marginali e la matrice di correlazione lineare di un generico vettore aleatorio, non è possibile determinare in modo univoco la sua distribuzione congiunta;
- **Limite 2:** assegnate le distribuzioni marginali non è possibile costruire una distribuzione congiunta per qualsiasi valore del coefficiente di correlazione compreso in $[-1, 1]$.

Riguardo al secondo punto esiste un teorema dimostrato inizialmente da Hoeffding e successivamente anche da Fréchet che permette di determinare nel caso generale i valori estremi dell'intervallo tra cui può variare il coefficiente di correlazione.

Teorema 2.8. *Sia (X, Y) un generico vettore aleatorio con distribuzioni marginali a varianza finita F_X e F_Y , mentre sia incognita la distribuzione congiunta, allora valgono:*

1. *I valori correlazione ammissibili formano un intervallo chiuso $[\rho_{\min}, \rho_{\max}]$ con $\rho_{\min} < 0 < \rho_{\max}$.*
2. *Il valore minimo di correlazione $\rho = \rho_{\min}$ è raggiunto se e solo se X e Y sono contro-monotoniche, mentre il valore massimo $\rho = \rho_{\max}$ è ottenuto se e solo se X e Y sono co-monotoniche.*
3. *Gli estremi ρ_{\min} e ρ_{\max} raggiungono i valori -1 e 1 se sono della stesso tipo⁴ rispettivamente le coppie X e $-Y$ o X e Y .*

2.2.2 Misure di correlazione di rango

Le misure di correlazione di rango (*rank correlation*) sono semplicemente delle misure di dipendenza che dipendono esclusivamente dalla distribuzione congiunta delle variabili e non dal loro comportamento marginali, diversamente dal caso della correlazione lineare che dipende invece da entrambi. Vedremo che gli stimatori empirici standard della correlazione di rango possono essere

³Per maggiori dettagli ed esempi riguardanti questi due ultimi aspetti rimandiamo a Embrechts et al. (2002).

⁴Due variabili aleatorie X e Y sono dello stesso tipo se esistono $a, b \in \mathbb{R}$ con $a > 0$ tali che $Y = aX + b$.

calcolati osservando gli ordinamenti (*ranks*) dei dati, da cui appunto il nome. In altre parole, per questo tipo di indicatori è sufficiente conoscere l'ordinamento del campione di ciascuna variabile di interesse e non l'effettivo valore numerico.

La ragione dello studio di questo tipo di indicatori va ricercata principalmente nella loro grande utilità nella pratica, trovano infatti largo impiego nella calibrazione delle copule a dati reali. Inoltre dal punto di vista teorico, essendo legati direttamente alle funzioni di copula, presentano proprietà più interessanti rispetto alla correlazione lineare, come avremo modo di discutere tra poco.

Prima di analizzare i diversi indicatori di correlazione di rango è necessario introdurre i concetti di concordanza e discordanza tra osservazioni. Intuitivamente una coppia di variabili aleatorie è concordante se 'grandi' valori di una sono associati a 'grandi' valori dell'altra, e 'piccoli' valori di una a 'piccoli' valori dell'altra.

Definizione 2.7. *Due osservazioni (x_1, y_1) e (x_2, y_2) di un vettore aleatorio (X, Y) si dicono concordanti se $x_1 < x_2$ e $y_1 < y_2$ oppure se $x_1 > x_2$ e $y_1 > y_2$. Analogamente (x_1, y_1) e (x_2, y_2) si dicono discordanti se $x_1 < x_2$ e $y_1 > y_2$ oppure se $x_1 > x_2$ e $y_1 < y_2$.*

Attraverso la concordanza e la discordanza di una popolazione definiamo ora i principali indicatori di correlazione di rango: il τ di Kendall e il ρ di Spearman.

Considerato un campione $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ di n osservazioni da un vettore (X, Y) di variabili aleatorie, esistono $\binom{n}{2}$ possibili coppie di osservazioni, di cui c concordanti e d discordanti. Il τ di Kendall per un tale campione è definito come

$$\tau = \frac{c - d}{c + d} = \binom{n}{2}^{-1} (c - d).$$

In modo equivalente, τ si può definire come la differenza tra la probabilità di concordanza e la probabilità di discordanza per una coppia di osservazioni (x_i, y_i) e (x_j, y_j) scelte casualmente. Analogamente si definisce il τ di Kendall per un vettore di variabili aleatorie continue (X, Y) .

Definizione 2.8. *Il τ di Kendall per il vettore aleatorio (X, Y) è definito come*

$$\tau(X, Y) = P((X - \tilde{X})(Y - \tilde{Y}) > 0) - P((X - \tilde{X})(Y - \tilde{Y}) < 0),$$

dove (\tilde{X}, \tilde{Y}) è una copia indipendente e identicamente distribuita (i.i.d.) di (X, Y) .

Nel teorema 2.9 esplicitiamo il legame tra l'indicatore $\tau(X, Y)$ e la distribuzione congiunta del vettore (X, Y) .

Teorema 2.9. *Siano X e Y due variabili aleatorie continue la cui copula è C , allora il τ di Kendall di X e Y è dato da*

$$\tau(X, Y) = 4 \iint_{[0,1]^2} C(u, v) dC(u, v) - 1. \quad (2.8)$$

Notando che l'integrale in (2.8) può essere interpretato come il valore atteso della funzione $C(U, V)$ con U e V uniformi standard e distribuzione congiunta $C(U, V)$, allora la formula (2.8) si può riscrivere nel seguente modo

$$\tau(X, Y) = 4E(C(U, V)) - 1.$$

Avendo a disposizione n coppie i.i.d. di un vettore aleatorio (X, Y) , allora uno stimatore non parametrico del τ di Kendall è

$$\tau(X, Y) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sign}((X_i - X_j)(Y_i - Y_j)). \quad (2.9)$$

Definiamo ora un secondo indicatore di correlazione di rango, il ρ di Spearman che come il τ di Kendall può essere definito in funzione delle probabilità di concordanza e discordanza.

Definizione 2.9. *Siano (X_1, Y_1) , (X_2, Y_2) e (X_3, Y_3) tre vettori aleatori indipendenti, ma con la stessa distribuzione congiunta H e distribuzioni marginali F e G . Il ρ di Spearman è una quantità proporzionale alla differenza tra la probabilità di concordanza e la probabilità di discordanza dei vettori aleatori (X_1, Y_1) e (X_2, Y_3) (o equivalentemente (X_3, Y_2)), che in formule diventa*

$$\rho_S = 3(P((X_1 - X_2)(Y_1 - Y_3) > 0) - P((X_1 - X_2)(Y_1 - Y_3) < 0)).$$

Facciamo notare che la distribuzione congiunta di (X_1, Y_1) è $H(x, y)$, mentre quella di (X_2, Y_3) è $F(x)G(y)$ dato che X_2 e Y_3 sono scelti indipendenti.

In analogia con il teorema 2.9 per il τ di Kendall, il teorema 2.10 stabilisce il legame tra ρ di Spearman e copule⁵.

Teorema 2.10. *Siano X e Y due variabili aleatorie continue la cui copula è C , allora il ρ di Spearman di X e Y è dato da*

$$\rho_S(X, Y) = 12 \iint_{[0,1]^2} uv dC(u, v) - 3 = 12 \iint_{[0,1]^2} C(u, v) dudv - 3. \quad (2.10)$$

⁵Per una dimostrazione dei Teoremi 2.9 e 2.10 rimandiamo a McNeil et al.(2006), pag.207.

Ricordando che F e G sono le distribuzioni marginali rispettivamente di X e di Y , se definiamo $U = F(X)$ e $V = G(Y)$, allora la (2.10) diventa

$$\begin{aligned}\rho_S(X, Y) &= 12 \iint_{[0,1]^2} uv dC(u, v) - 3 = 12E(UV) - 3 = \\ &= \frac{E(UV) - 1/4}{1/12} = \frac{E(UV) - E(U)E(V)}{\sqrt{Var(U)}\sqrt{Var(V)}} = \\ &= \frac{Cov(U, V)}{\sqrt{Var(U)}\sqrt{Var(V)}} = \rho(F(X), G(Y)).\end{aligned}$$

In altre parole, il ρ di Spearman è semplicemente la correlazione lineare delle variabili trasformate $U = F(X)$ e $V = G(Y)$.

Il τ di Kendall e il ρ di Spearman presentano diverse proprietà comuni. Innanzitutto sono entrambe misure di dipendenza simmetriche con valori appartenenti all'intervallo $[-1, 1]$, in particolare il valore -1 è relativo a variabili contro-monotone, mentre 1 indica variabili co-monotone. Variabili aleatorie indipendenti hanno un valore di τ e ρ_S nullo, tuttavia non vale il viceversa, cioè una correlazione di rango pari a 0 non implica necessariamente l'indipendenza tra le variabili. Dalle relazioni (2.9) e (2.10) segue che il τ di Kendall e il ρ di Spearman sono entrambi funzioni crescenti del valore della copula sottostante, da cui deriva che essi sono anche crescenti rispetto all'ordinamento descritto nella Definizione 2.5.

Relativamente ai due limiti evidenziati per l'indice di correlazione lineare, il primo rimane ancora valido: note le distribuzioni marginali e gli indicatori di correlazione di rango non è possibile determinare in modo univoco la distribuzione congiunta. Diversamente il secondo si può ritenere superato se si considerano il τ di Kendall o il ρ di Spearman: assegnate due qualsiasi distribuzioni marginali, è possibile individuare una distribuzione congiunta relativa ad un qualsiasi valore di correlazione di rango appartenente all'intervallo $[-1, 1]$. Un possibile modo di fare questo è considerare una combinazione convessa del tipo

$$H(x, y) = \lambda W(F(x), G(y)) + (1 - \lambda)M(F(x), G(y)),$$

dove W e M sono i limiti di Fréchet-Hoeffding relativi rispettivamente alle situazioni di contro-monotonia e co-monotonia. Per il vettore aleatorio (X, Y) avente questa distribuzione congiunta risulta

$$\tau(X, Y) = \rho_S(X, Y) = 1 - 2\lambda,$$

che possono assumere un qualsiasi valore in $[-1, 1]$ per appropriate scelte di $\lambda \in [0, 1]$.

Nonostante entrambi questi indicatori misurino la probabilità di concordanza tra variabili aleatorie con una certa copula, i loro valori sono spesso diversi. Una relazione universale tra τ e ρ_S è espressa dalla disuguaglianza (2.11) contenuta nel seguente teorema.

Teorema 2.11. *Sia (X, Y) un vettore di variabili aleatorie continue e siano τ e ρ_S il τ di Kendall e il ρ di Spearman rispettivamente. Allora*

$$-1 \leq 3\tau - 2\rho_S \leq 1. \quad (2.11)$$

Il τ di Kendall e il ρ di Spearman non sono gli unici indicatori di *rank correlation*, ma sono certamente quelli più studiati in virtù della loro utilità in pratica nella parametrizzazione delle copule. Citiamo comunque a titolo di esempio anche il coefficiente di correlazione mediano β introdotto da Blomqvist e l'indice di cogradazione quadratico di Gini⁶.

2.2.3 La dipendenza di coda

I coefficienti di dipendenza di coda (*tail dependence*), proprio come quelli di correlazione di rango, sono delle misure di dipendenza che dipendono esclusivamente dalla copula di una coppia di variabili aleatorie X e Y , per tanto ne ereditano la proprietà di invarianza rispetto a trasformazioni strettamente crescenti. Il motivo per cui si studiano questi indicatori è che permettono di valutare il grado di dipendenza sugli estremi o, in altre parole, misurano il livello di dipendenza nelle code della distribuzione bivariata, aspetto molto importante soprattutto in campo finanziario o attuariale.

Definizione 2.10. *Siano X e Y due variabili aleatorie continue con funzioni di distribuzioni F e G . Si definisce coefficiente di dipendenza di coda superiore ('upper tail dependence')*

$$\lambda_U := \lim_{u \rightarrow 1^-} P(Y > G^{-1}(u) | X > F^{-1}(u)),$$

a condizione che esista il limite $\lambda_U \in [0, 1]$. Se $0 < \lambda_U \leq 1$, si afferma che X e Y presentano dipendenza sulla coda superiore o che sono asintoticamente dipendenti sulla coda superiore; se invece $\lambda_U = 0$, X e Y sono asintoticamente indipendenti sulla coda superiore. In modo del tutto analogo si definisce il coefficiente di dipendenza di coda inferiore ('lower tail dependence')

$$\lambda_L := \lim_{u \rightarrow 0^+} P(Y \leq G^{-1}(u) | X \leq F^{-1}(u)),$$

a condizione che esista il limite $\lambda_L \in [0, 1]$.

⁶Maggiori dettagli su questi indicatori si possono trovare in Nelsen (1999).

Se F e G sono distribuzioni continue, allora è possibile esprimere λ_U e λ_L in funzione dell'unica copula C della distribuzione bivariata, confermando quindi il fatto che il coefficiente di dipendenza di coda esprime effettivamente una proprietà delle copule.

Utilizzando la (2.5) per λ_U otteniamo che

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{\widehat{C}(1-u, 1-u)}{1-u} = \lim_{u \rightarrow 0^+} \frac{\widehat{C}(u, u)}{u}, \quad (2.12)$$

dove \widehat{C} è la copula di sopravvivenza di C definita in (2.6). Per λ_L usiamo la proprietà della distribuzione condizionata e la (2.3) ottenendo

$$\begin{aligned} \lambda_L &= \lim_{u \rightarrow 0^+} \frac{P(Y \leq G^{-1}(u), X \leq F^{-1}(u))}{P(X \leq F^{-1}(u))} \\ &= \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}. \end{aligned} \quad (2.13)$$

Ricordando che il legame tra la copula di sopravvivenza \widehat{C} e la funzione di sopravvivenza di una copula \bar{C} è

$$\bar{C}(u, v) = 1 - u - v + C(u, v) = \widehat{C}(1-u, 1-v),$$

possiamo ricavare il seguente risultato che lega i coefficienti di dipendenza di coda superiore e inferiore

$$\begin{aligned} \lambda_{U,C} &= \lim_{u \rightarrow 1^-} \frac{\bar{C}(u, u)}{1-u} \\ &= \lim_{u \rightarrow 1^-} \frac{\widehat{C}(1-u, 1-u)}{1-u} \\ &= \lim_{u \rightarrow 0^+} \frac{\widehat{C}(u, u)}{u} \\ &= \lambda_{L,\widehat{C}}. \end{aligned} \quad (2.14)$$

La (2.14) dimostra che il coefficiente di dipendenza di coda superiore della copula C è equivalente al coefficiente di dipendenza di coda inferiore della copula di sopravvivenza \widehat{C} . In modo del tutto simile si dimostra lo stesso legame λ_L della copula C e λ_U di \widehat{C} .

Un'ultima osservazione riguarda le copule a simmetria radiale, per le quali vale $\lambda_U = \lambda_L$, in virtù del fatto che $C = \widehat{C}$ e cioè che la copula di sopravvivenza equivale alla copula stessa.

2.3 Famiglie di copule

Esistono vari metodi utilizzati per creare tipologie differenti di copule capaci di descrivere molteplici forme di dipendenza. In questa sezione verranno presentate due particolari famiglie di copule: le copule ellittiche, caratterizzate da un importante legame con il coefficiente di correlazione lineare e basate su distribuzioni ellittiche, e le copule Archimedee, particolarmente utili grazie alla loro maneggevolezza matematica.

2.3.1 Le copule ellittiche

Introduciamo innanzitutto i concetti di distribuzione sferica e distribuzione ellittica.

Definizione 2.11. *Un vettore aleatorio $\mathbf{X} = (X_1, X_2, \dots, X_n)$ ha una distribuzione sferica se risulta $\mathbf{Z}\mathbf{X} = \mathbf{X}$, cioè se è ortogonale per ogni \mathbf{Z} tale che $\mathbf{Z}^T \mathbf{Z} = \mathbf{Z}\mathbf{Z}^T = \mathbf{I}_n$.*

Le distribuzioni sferiche sono distribuzioni simmetriche di vettori incorrelati di media nulla, due rappresentanti di questa famiglia sono per esempio la distribuzione normale multivariata e la distribuzione logistica standard.

Definizione 2.12. *Un vettore aleatorio $\mathbf{X} = (X_1, X_2, \dots, X_n)$ ha una distribuzione ellittica, $\mathbf{X} \sim E_n(\mu, \Sigma, \phi)$ se $\mathbf{X} =_d \mu + A\mathbf{Y}$, dove μ e A sono, rispettivamente, un vettore e una matrice di costanti e \mathbf{Y} è una variabile aleatoria con distribuzione sferica. Inoltre ϕ e Σ sono tali per cui*

$$\varphi_{\mathbf{X}-\mu}(\mathbf{t}) = \phi(\mathbf{t}^T \Sigma \mathbf{t}),$$

ovvero che la funzione caratteristica di $\mathbf{X} - \mu$ è funzione della forma quadratica $\mathbf{t}^T \Sigma \mathbf{t}$. La matrice A è completamente definita dagli altri parametri della distribuzione.

Le distribuzioni ellittiche hanno varie proprietà di cui ricordiamo le principali⁷:

- Le distribuzioni marginali di una distribuzione ellittica sono a loro volta ellittiche dello stesso tipo, cioè con lo stesso generatore ϕ .
- Combinazioni lineari di vettori aleatori indipendenti, ma con distribuzioni ellittiche aventi la stessa matrice di dispersione Σ (eventualmente riscalate per una costante di proporzionalità positiva) sono a loro volta ellittiche.

⁷Per maggiori dettagli sulle distribuzioni ellittiche e le dimostrazioni delle loro proprietà si veda Embrechts et al. (2003).

- Il coefficiente di correlazione lineare è la naturale misura di dipendenza tra variabili aleatorie con distribuzione congiunta ellittica. Risulta infatti che per $\mathbf{X} \sim E_n(\mu, \Sigma, \phi)$ con $0 < Var(X_i), Var(X_j) < \infty$ si ha

$$\rho_{ij} = \frac{Cov(X_i, X_j)}{\sqrt{Var(X_i)}\sqrt{Var(X_j)}} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}}\sqrt{\Sigma_{jj}}}.$$

Fanno parte della classe delle distribuzione ellittiche varie distribuzioni multivariate accomunate dalla stessa trattabilità della distribuzione normale multivariata, ma che allo stesso tempo permettono di modellizzare eventi estremi e forme di dipendenza diverse dalla normale.

A questo punto possiamo definire le copule ellittiche in modo molto semplice, sono infatti le copule delle distribuzioni ellittiche. I principali esponenti di questa classe di copule sono la copula Gaussiana e la copula t .

Copula Gaussiana

La copula Gaussiana è la copula corrispondente ad una distribuzione normale multivariata con matrice di correlazione lineare R . Se indichiamo con Φ_R^n la funzione di distribuzione della normale standard n -variata e con Φ^{-1} l'inversa della distribuzione normale univariata, ricaviamo facilmente dal teorema di Sklar che l'espressione della generica copula Gaussiana n -dimensionale risulta

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

La copula Gaussiana non ha una forma chiusa semplice, ma può essere espressa come integrale della distribuzione congiunta. Nel caso bivariato e per un coefficiente di correlazione lineare $|\rho| < 1$, si può riscrivere nel seguente modo

$$C_R^{Ga}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left\{-\frac{s_1^2 - 2\rho s_1 s_2 + s_2^2}{2(1-\rho^2)}\right\} ds_1 ds_2.$$

Una prima peculiarità della copula Gaussiana è che per essa sussistono delle relazioni che legano direttamente il coefficiente di correlazione lineare ρ con il τ di Kendall e il ρ di Spearman⁸:

$$\tau = \frac{2}{\pi} \arcsin \rho$$

⁸Per una dimostrazione di queste relazioni si veda McNeil et al.(2006).

$$\rho_S = \frac{6}{\pi} \arcsin \frac{1}{2} \rho$$

Per quanto riguarda invece la dipendenza di coda, la copula Gaussiana è sprovvista sia del coefficiente di dipendenza di coda superiore che inferiore per valori di $|\rho| < 1$.

Per verificare l'effetto derivante dalla variazione del coefficiente di correlazione lineare sulla struttura di dipendenza, abbiamo rappresentato in figura 2.3 alcune simulazioni che prevedono la stessa struttura di copula Gaussiana applicata a delle marginali uniformi $U(0,1)$, ma relative a differenti valori di ρ .

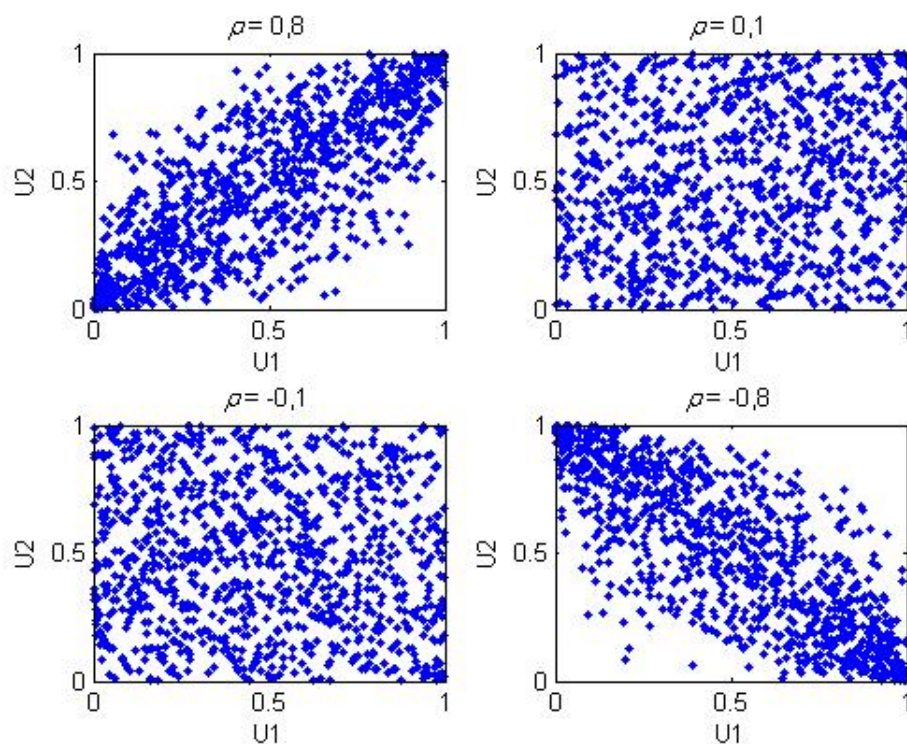


Figura 2.3: 1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula Gaussiana e differenti valori di ρ .

Per simulare osservazioni da una copula Gaussiana C_R^{Ga} , il metodo più comune si basa sulla decomposizione di Cholesky⁹ della matrice R e prevede le seguenti operazioni:

⁹La decomposizione di Cholesky di R è quell'unica matrice triangolare inferiore A per cui $R = AA^T$. L'algoritmo per determinare i termini di tale matrice è stato proposto

- Determinare la decomposizione di Cholesky A di R ;
- Simulare n realizzazioni indipendenti $\mathbf{z} = (z_1, \dots, z_n)$ da una distribuzione normale standard $N(0, 1)$;
- Porre $\mathbf{x} = A\mathbf{z}$;
- Porre $u_i = \Phi(x_i)$ per $i = 1, \dots, n$;
- $(u_1, \dots, u_n) \sim C_R^{Ga}$ per distribuzioni marginali uniformi $U(0, 1)$.

Se la distribuzione marginale è diversa dalla distribuzione uniforme standard, allora è prevista un'ultima operazione:

- $y_i = F_Y^{-1}(u_i)$ per $i = 1, \dots, n$, dove F_Y^{-1} è l'inversa della distribuzione marginale F_Y .

Copula t

Nello stesso modo con cui abbiamo estratto una copula dalla distribuzione normale multivariata, è possibile ricavare una copula da una qualsiasi altra distribuzione con marginali continue. Esaminiamo ora il caso della copula t che assume la seguente forma

$$C_{\nu,R}^t(\mathbf{u}) = t_{\nu,R}^n(t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_n)),$$

dove con $t_{\nu,R}^n$ indichiamo la distribuzione congiunta t con ν gradi di libertà e matrice di correlazione lineare R e con t^{-1} l'inversa della distribuzione t univariata.

Come per la copula Gaussiana bidimensionale, esiste una formulazione analoga anche per la copula t bivariata che è

$$C_{\nu,R}^t(u, v) = \int_{-\infty}^{t_{\nu}^{-1}(u)} \int_{-\infty}^{t_{\nu}^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ 1 + \frac{s_1^2 - 2\rho s_1 s_2 + s_2^2}{\nu(1-\rho^2)} \right\}^{-\frac{\nu+2}{2}} ds_1 ds_2,$$

solo nel caso in cui $\nu > 2$, diversamente infatti la matrice di correlazione lineare R non è definita.

dallo stesso Cholesky e prevede che $a_{i,j} = \frac{\rho_{i,j} - \sum_{s=1}^{j-1} a_{i,s} a_{j,s}}{\sqrt{1 - \sum_{s=1}^{j-1} a_{j,s}^2}}$ per $1 \leq j \leq i \leq n$ dove per convenzione $\sum_{s=1}^0(\cdot) = 0$. Nel caso bivariato è semplicemente $A = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}$.

L'algoritmo utilizzato per generare osservazioni da una copula t si basa sulla nota relazione che se $\mathbf{X} \sim t_{\nu,R}^n$, allora vale

$$\mathbf{X} =_d \mu + \frac{\sqrt{\nu}}{\sqrt{S}} \mathbf{Z},$$

dove $\mu \in \mathbb{R}^n$, $S \sim \chi_{\nu}^2$ e \mathbf{Z} , che si distribuisce come una normale n -variata $N(0, R)$, sono indipendenti. Per simulare da una copula $t_{\nu,R}^n$ le operazioni da eseguire sono quindi le seguenti:

- Determinare la decomposizione di Cholesky A di R ;
- Simulare n osservazioni indipendenti $\mathbf{z} = (z_1, \dots, z_n)$ da una distribuzione normale standard;
- Simulare una variabile casuale s da una distribuzione χ^2 con ν gradi di libertà e indipendente da \mathbf{z} ;
- Porre $\mathbf{a} = A\mathbf{z}$;
- Porre $\mathbf{x} = \frac{\sqrt{\nu}}{\sqrt{s}} \mathbf{a}$;
- Porre $u_i = t_{\nu}(x_i)$ per $i = 1, \dots, n$;
- $(u_1, \dots, u_n) \sim C_{\nu,R}^t$ per distribuzioni marginali uniformi $U(0, 1)$.

Anche in questo caso, se la distribuzione marginale è diversa dalla distribuzione uniforme standard, c'è un ultimo passaggio che consiste in:

- $y_i = F_Y^{-1}(u_i)$ per $i = 1, \dots, n$, dove F_Y^{-1} è l'inversa della distribuzione marginale F_Y .

Dall'analisi della figura 2.4 si può osservare che la copula t per un basso numero di gradi di libertà evidenzia un doppio trend, uno positivo e uno negativo, anche nel caso di correlazione lineare nulla.

Una caratteristica che distingue la copula Gaussiana dalla copula t è che per la prima esistono delle relazioni analitiche che legano il τ di Kendall e il ρ di Spearman al coefficiente di correlazione lineare, mentre per la seconda questo non accade, tuttavia la t copula è fornita di dipendenza di coda superiore e inferiore, diversamente dalla copula Gaussiana. Come tutte le copule ellittiche anche la copula t gode di simmetria radiale, perciò i coefficienti di dipendenza di coda superiore e di dipendenza di coda inferiore coincidono ed il loro valore è

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

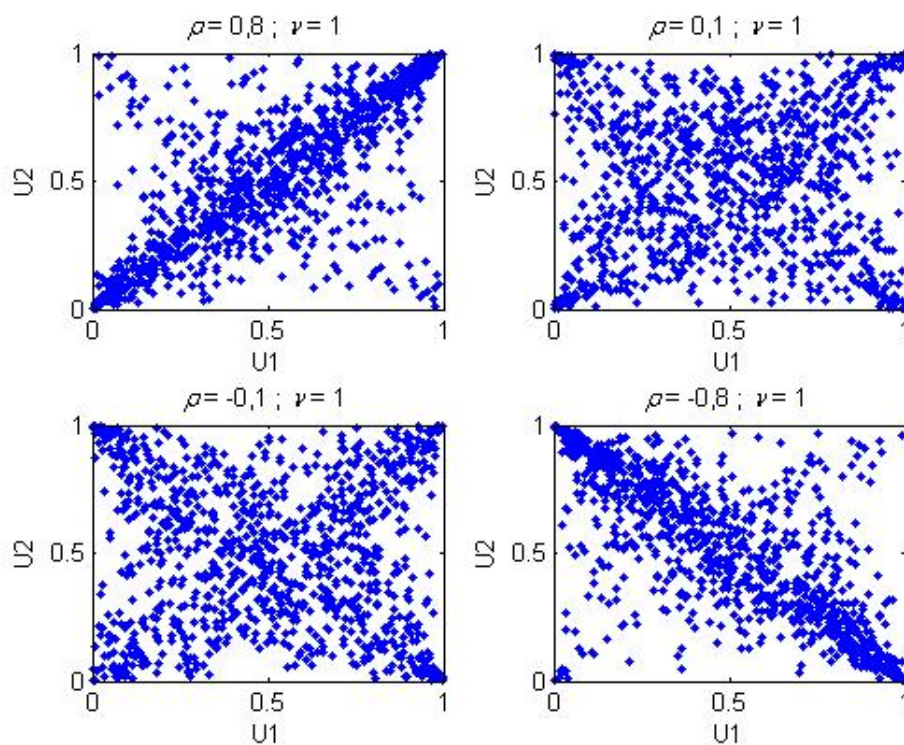


Figura 2.4: 1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula t con 1 grado di libertà e differenti valori di ρ .

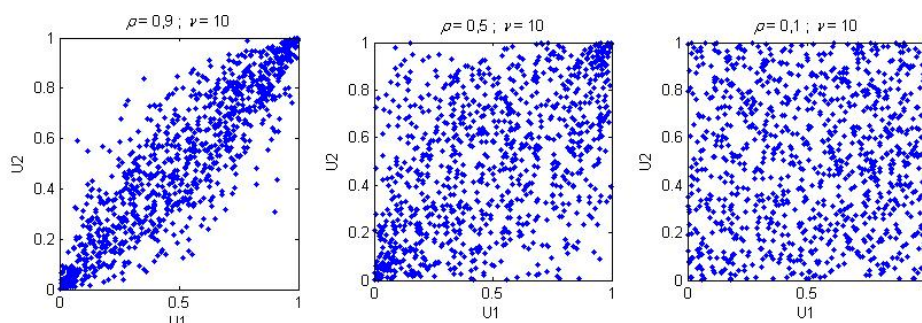


Figura 2.5: 1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula t con 10 gradi di libertà e differenti valori di ρ .

Da questo si può chiaramente vedere che la dipendenza di coda (sia superiore sia inferiore) della copula t è crescente rispetto a ρ , mentre è decrescente rispetto a ν . Quest'ultimo aspetto era del tutto prevedibile dato che al crescere dei gradi di libertà la copula t si avvicina alla forma della copula Gaussiana, come rappresentato in figura 2.5.

2.3.2 Le copule Archimedee

Fanno parte della famiglia delle copule Archimedee numerose classi di copule parametriche che forniscono una vasta gamma di differenti strutture di dipendenza. Un primo aspetto che accomuna tutte queste copule e le differenzia da quelle ellittiche è che le copule Archimedee possono essere espresse attraverso una forma chiusa. Diversamente dalle copule esaminate sino ad ora le copule Archimedee non sono ricavate da distribuzioni multivariate attraverso il teorema di Sklar, da cui consegue che l'estensione della copula bivariata al caso n -dimensionale non è sempre scontata, ma anzi richiede la verifica di alcune condizioni tecniche. Tuttavia noi ci limiteremo ad esaminare il caso delle copule Archimedee bivariate rimandando a Nelsen (1999) per maggiori dettagli circa le estensioni in dimensioni superiori.

La costruzione delle copule Archimedee si basa sulla possibilità di scindere la distribuzione congiunta H tra le due marginali (F e G) tramite una funzione $\lambda(t)$ che deve essere positiva nell'intervallo $[0, 1]$, cioè $\lambda(H(x, y)) = \lambda(F(x)) \cdot \lambda(G(y))$. Ponendo poi $\varphi(t) = -\ln(\lambda(t))$ la relazione si trasforma in $\varphi(H(x, y)) = \varphi(F(x)) + \varphi(G(y))$ ed in termini di copule $\varphi(C(u, v)) = \varphi(u) + \varphi(v)$. Utilizzando la funzione pseudo-inversa¹⁰ $\varphi^{[-1]}$, una condizione necessaria e sufficiente affinché

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)). \quad (2.15)$$

sia una copula è che φ sia una funzione convessa.

Le copule che possono essere espresse come in (2.15) si definiscono copule Archimedee e la funzione φ si chiama generatore. Nel caso in cui $\varphi(0) = \infty$ si dice che φ è un generatore in senso stretto, inoltre se $\varphi^{[-1]} = \varphi^{-1}$ e quindi se $C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$ si dice che la copula è Archimedea in senso stretto.

¹⁰La funzione pseudo-inversa di una funzione $\varphi : [0, 1] \rightarrow [0, \infty]$ continua e strettamente decrescente con $\varphi(1) = 0$ e $\varphi(0) \leq \infty$ è una funzione definita in $[0, \infty]$ tale che

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

Sono numerose le proprietà matematiche che caratterizzano questa classe di copule, e tra questi ricordiamo:

- C è simmetrica, ovvero $C(u, v) = C(v, u)$ per ogni $u, v \in [0, 1]$;
- C è associativa, cioè per ogni $u, v, w \in [0, 1]$ vale $C(C(u, v), w) = C(u, C(v, w))$;
- se φ è un generatore, allora lo è anche $c\varphi$ per una qualsiasi costante c positiva.

Per simulare da una copula Archimedeana con funzione generatrice φ si può seguire il seguente procedimento:

- Simulare due osservazioni indipendenti s e q da una distribuzione uniforme $U(0, 1)$;
- Porre $t = K_C^{-1}(q)$, dove K_C è la funzione di distribuzione di $C(U, V)$;
- Porre $u = \varphi^{[-1]}(s\varphi(t))$ e $v = \varphi^{[-1]}((1-s)\varphi(t))$;
- $(u, v) \sim C^\varphi$.

Qualora le distribuzioni marginali non dovessero essere delle Uniformi $(0,1)$, allora

- $x = F^{-1}(u)$ e $y = G^{-1}(v)$, dove F^{-1} e G^{-1} sono le inverse delle funzioni di distribuzione marginali F e G .

Un'utile relazione che caratterizza le copule Archimedee è quella che lega il τ di Kendall con la funzione generatrice φ . Se X e Y sono variabili aleatorie continue con unica funzione di copula Archimedeana C di generatore φ , allora

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

Esistono inoltre dei teoremi che permettono di esprimere anche i coefficienti di *tail dependence* in funzione del generatore. In particolare si può dimostrare che se una copula Archimedeana in senso stretto con generatore φ ammette dipendenza di coda superiore allora $\varphi^{-1'}(0) = -\infty$ e

$$\lambda_U = 2 - 2 \lim_{s \rightarrow 0^+} \frac{\varphi^{-1'}(2s)}{\varphi^{-1'}(s)}, \quad (2.16)$$

mentre se ammette dipendenza di coda inferiore risulta

$$\lambda_L = 2 \lim_{s \rightarrow \infty} \frac{\varphi^{-1'}(2s)}{\varphi^{-1'}(s)}. \quad (2.17)$$

Le copule Archimedee maggiormente studiate e impiegate nella pratica sono quelle appartenenti alle famiglie di Gumbel, di Clayton e di Frank, di cui forniremo una descrizione qui di seguito.

Copula di Gumbel

Il generatore di questa classe di copule è $\varphi(t) = (-\ln t)^{-\theta}$ con $\theta \geq 1$. Dalla (2.15) si ricava facilmente che

$$C_\theta^{Gu}(u, v) = \exp \left\{ -[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta} \right\}.$$

La copula di Gumbel è Archimedea in senso stretto e il τ di Kendall vale $1 - 1/\theta$. Facciamo ora notare che per i valori estremi del parametro otteniamo

$$C_1^{Gu}(u, v) = u \cdot v = \Pi(u, v) \quad \text{e} \quad \lim_{\theta \rightarrow \infty} C_\theta^{Gu}(u, v) = \min(u, v) = M(u, v).$$

Questa famiglia di copule ammette coefficiente di dipendenza di coda superiore, infatti $\varphi^{-1'}(s) = -s^{1/\theta-1} \exp\{-s^{1/\theta}\}/\theta$ con $\varphi^{-1'}(0) = -\infty$ e quindi utilizzando la (2.16) si ottiene

$$\lambda_U = 2 - 2 \lim_{s \rightarrow 0^+} \frac{\varphi^{-1'}(2s)}{\varphi^{-1'}(s)} = 2 - 2^{1/\theta} \lim_{s \rightarrow 0^+} \frac{\exp\{-(2s)^{1/\theta}\}}{\exp\{-s^{1/\theta}\}} = 2 - 2^{1/\theta}.$$

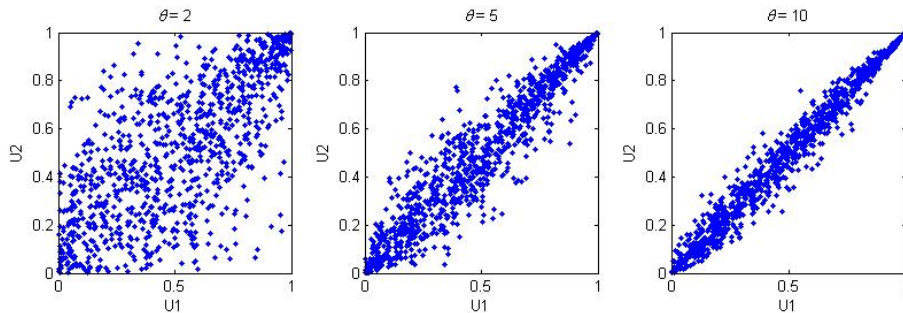


Figura 2.6: 1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula di Gumbel e differenti valori di θ (2, 5, 10 relativi a τ pari a 0,5, 0,8 e 0,9).

Nella figura 2.6 sono riportate alcune simulazioni da una copula di Gumbel con diversi valori del parametro e come è logico aspettarsi al crescere di θ le osservazioni risultano progressivamente meno disperse.

Concludiamo la presentazione di questa famiglia di copule, ricordando che per la loro caratteristica di esaltare le dipendenze sulla coda destra, la copula di Gumbel risulta particolarmente indicata per descrivere fenomeni legati a rischi catastrofali.

Copula di Clayton

Il generatore di questa famiglia di copule è $\varphi(t) = (t^{-\theta} - 1)/\theta$ con $\theta \geq -1$ e $\theta \neq 0$. Il suo funzionale è

$$C_{\theta}^{Cl}(u, v) = \max([u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0).$$

Non è Archimeda in senso stretto per qualsiasi valore del parametro, ma solo per $\theta > 0$, nel qual caso la sua scrittura si semplifica in

$$C_{\theta}^{Cl}(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}.$$

Esaminando le copule che si ottengono per i valori estremi di θ si trova $C_{-1}^{Cl}(u, v) = W(u, v)$, $\lim_{\theta \rightarrow 0} C_{\theta}^{Cl}(u, v) = \Pi(u, v)$ e $\lim_{\theta \rightarrow \infty} C_{\theta}^{Cl}(u, v) = M(u, v)$.

L'espressione del τ di Kendall è $\theta/(\theta + 2)$ e restringendosi alla situazione $\theta > 0$ la copula di Clayton, a differenza di quella di Gumbel, ammette coefficiente di dipendenza di coda inferiore. Si ha $\varphi^{-1'}(s) = (-1/\theta)(1 - \theta s)^{-1/\theta - 1}$ e quindi per la (2.17) ricaviamo che

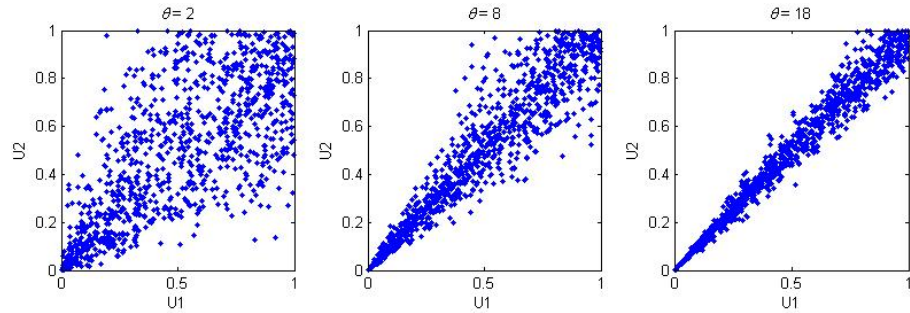


Figura 2.7: 1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula di Clayton e differenti valori di θ (2, 8, 18 relativi a τ pari a 0,5, 0,8 e 0,9).

$$\lambda_L = 2 \lim_{s \rightarrow \infty} \frac{\varphi^{-1'}(2s)}{\varphi^{-1'}(s)} = 2 \lim_{s \rightarrow \infty} \left[\frac{(1 + s\theta s)^{-1/\theta - 1}}{(1 + \theta s)^{-1/\theta - 1}} \right] = 2 \cdot 2^{-1/\theta - 1} = 2^{-1/\theta}.$$

Osservando i grafici riportati in figura 2.7 si può notare come a parità di τ di Kendall la copula di Clayton, oltre a identificare una coda sinistra particolarmente accentuata, risulti più dispersa rispetto a quella di Gumbel.

Copula di Frank

La terza copula Archimedeica che presentiamo è la copula di Frank che è definita nel seguente modo

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

a partire dal generatore $\varphi(t) = -\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$ con $\theta \in \mathbb{R} \setminus \{0\}$.

In modo molto simile a quanto osservato per la copula di Clayton si ha che $\lim_{\theta \rightarrow -\infty} C_{\theta}^{Fr}(u, v) = W(u, v)$, $\lim_{\theta \rightarrow 0} C_{\theta}^{Fr}(u, v) = \Pi(u, v)$ e $\lim_{\theta \rightarrow \infty} C_{\theta}^{Fr}(u, v) = M(u, v)$.

La copula di Frank è Archimedeica in senso stretto ed inoltre è l'unica della famiglia delle copule Archimedee per la quale vale l'uguaglianza $C_{\theta}^{Fr}(u, v) = \widehat{C}_{\theta}^{Fr}(u, v)$, da cui consegue la proprietà di simmetria radiale, come si può per esempio vedere nei grafici riportati in figura 2.8.

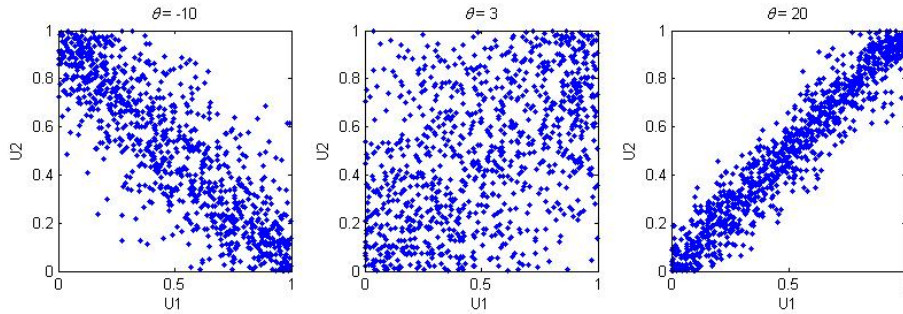


Figura 2.8: 1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con la copula di Frank e differenti valori di θ (-10, 3, 20).

Si può dimostrare (vedasi per esempio Genest (1987)) che l'espressione per il τ di Kendall è $\tau = 1 - 4\theta^{-1}(1 - D_1(\theta))$, dove $D_1(x)$ è la funzione di Debye $D_1(\theta) = \theta^{-1} \int_0^{\theta} x/(\exp(x) - 1) dx$.

Per quanto riguarda infine i coefficienti di dipendenza di coda si ha che

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

da cui

$$\varphi^{-1'}(0) = -\frac{e^{\theta} - 1}{\theta} \neq -\infty,$$

perciò la copula di Frank non ammette il coefficiente di coda superiore. Inoltre, combinando la (2.14) con il fatto che è una copula simmetrica, si deduce che non possiede neanche il coefficiente di coda inferiore.

Concludiamo la trattazione delle copule Archimedee con un confronto grafico tra queste e la copula Gaussiana, che ricordiamo essere il rappresentante principale della classe delle copule ellittiche. Volendo fare un confronto omogeneo con i dati simulati dalla copula Gaussiana in figura 2.3, abbiamo innanzitutto determinato il τ di Kendall relativo ai valori del coefficiente di correlazione lineare ρ dalla relazione che li lega nell'ipotesi di copula Gaussiana (che ricordiamo essere $\tau = \frac{2}{\pi} \arcsin \rho$); successivamente abbiamo utilizzato questi valori di τ per stimare i parametri delle copule Archimedee. I grafici con i dati così simulati sono riportati in figura 2.9 che, se confrontati con la figura 2.3, mettono in evidenza tutte le caratteristiche distintive circa le diverse strutture di dipendenza che le rispettive copule vanno a descrivere.

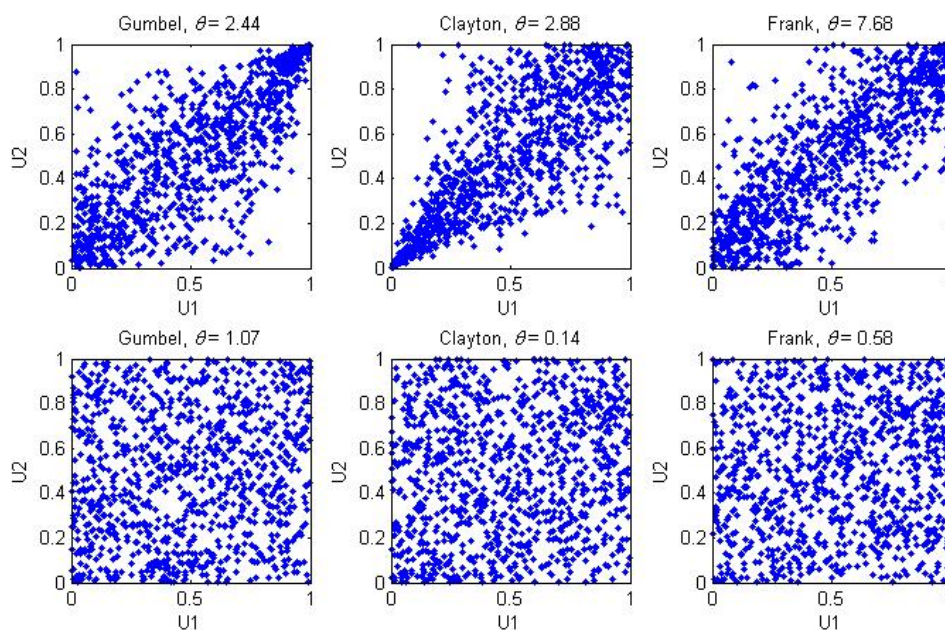


Figura 2.9: 1000 coppie di vettori simulati ipotizzando distribuzioni marginali Uniformi (0,1) con le copule di Gumbel, di Clayton e di Frank con i valori di θ relativi ai τ di Kendall compatibili con $\rho = 0,8$ e $\rho = 0,1$ della copula Gaussiana.

Capitolo 3

La pair-copula construction

L'interesse verso modelli multivariati è cresciuto in modo considerevole nell'ultimo decennio grazie soprattutto alla spinta data dalla rivoluzione tecnologica e le copule rappresentano uno strumento molto potente. Esse sono ormai uno strumento ampiamente utilizzato in diversi campi scientifici, per citarne alcuni: la finanza, le scienze attuariali e le scienze ambientali. Nonostante esista una vasta gamma di possibili copule bivariate e di relative strutture di dipendenza, la scelta si restringe notevolmente per dimensioni maggiori. Per questo motivo si sono sviluppate varie strutture basate sulle copule, tra le quali si inserisce la *pair-copula construction* (di seguito PCC). La teoria della PCC è stata introdotta dal lavoro di Joe (1996), successivamente sviluppato da Bedford e Cooke (2001, 2002) e in studi più recenti quali Kurowicka e Cooke (2006), Aas et al. (2009) e Dissman et al. (2011).

Il lavoro pionieristico di Joe (1996), ripreso poi nei lavori di Bedford e Cooke (2001, 2002) e inserito in un contesto inferenziale da Aas et al. (2009) sono alla base di tutta la teoria che si è sviluppata successivamente. La tecnica della PCC è una costruzione basata su strutture ad albero, costruite a partire da copule bivariate (le *pair-copule*) i cui argomenti sono funzioni di distribuzione condizionate¹. Nonostante si presenti in forma molto semplice, la PCC è uno strumento molto flessibile e ha le potenzialità per descrivere strutture anche piuttosto complesse.

In letteratura il metodo della PCC non è l'unico a fondarsi sulla teoria delle copule bivariate e per esempio in Embrecht et al. (2003) e Savu e Trede (2006) sono state proposte delle estensioni multivariate delle copule Archimedee bivariate. Costruzioni di questo tipo presentano tuttavia delle limitazioni, una delle più stringenti è che si possono modellare solo variabili che esibiscono dipendenza positiva.

¹Il numero delle variabili condizionanti è zero nel livello base e crescente di uno per ciascun livello successivo.

Nel paragrafo 3.1 ci occuperemo di determinare la fattorizzazione di una distribuzione multivariata attraverso copule bivariate condizionate e non condizionate, ovvero tramite la PCC.

Si è soliti formulare l'ipotesi che le copule condizionate non dipendano dalle variabili condizionanti se non attraverso le distribuzioni condizionate che hanno per argomento. Questa assunzione porta ad una forma semplificata della PCC che tratteremo nel paragrafo 3.2.

Nel successivo paragrafo 3.3 descriveremo in generale le strutture grafiche denominate R-vine e il loro legame con la PCC, mentre tratteremo nel paragrafo 3.4 il caso specifico dei D-vine in termini di procedimento di stima e di selezione di un modello.

Per finire (paragrafo 3.5) dimostreremo la validità della forma semplificata della PCC simulando un D-vine con *pair-copule* Gaussiane.

3.1 La decomposizione in *pair-copule*

Consideriamo n variabili aleatorie (X_1, \dots, X_n) con funzione di distribuzione congiunta $F_{1\dots n}(x_1, \dots, x_n)$. Assumendo che $F_{1\dots n}(x_1, \dots, x_n)$ sia assolutamente continua con marginali continue e strettamente crescenti, $F_i(x_i)$ con $i = 1 \dots, n$, allora la corrispondente densità congiunta può essere fattorizzata nel seguente modo

$$f_{1\dots n}(x_1, \dots, x_n) = f_1(x_1) \cdot f_{2|1}(x_2|x_1) \cdot \dots \cdot f_{n|1\dots n-1}(x_n|x_1, \dots, x_{n-1}). \quad (3.1)$$

La decomposizione in (3.1) è unica a meno di una permutazione delle variabili.

In un certo senso ogni funzione di distribuzione congiunta contiene implicitamente sia una descrizione del comportamento marginale di ciascuna variabile sia una descrizione della loro struttura di dipendenza. Come abbiamo già avuto modo di discutere nel capitolo precedente, la peculiarità delle copule è quella di riassumere le informazioni sulla loro struttura di dipendenza.

Attraverso il teorema di Sklar (cfr. paragrafo 2.1.2) si può ottenere una fattorizzazione alternativa della medesima densità congiunta

$$f_{1\dots n}(x_1, \dots, x_n) = c_{12\dots n}(F_1(x_1), \dots, F_n(x_n)) \cdot f_1(x_1) \cdots f_n(x_n), \quad (3.2)$$

dove $c_{12\dots n}(\cdot)$ è una qualche densità di copula n -variata identificata in modo unico.

Ora il nostro scopo è quello di semplificare l'espressione (3.1) con le informazioni contenute in (3.2) e per fare questo iniziamo considerando il caso $n = 3$, per il quale la (3.1) è semplicemente

$$f_{123}(x_1, x_2, x_3) = f_1(x_1)f_{2|1}(x_2|x_1)f_{3|12}(x_3|x_1, x_2). \quad (3.3)$$

In questo caso $f_{12}(x_1, x_2)$ per la (3.2) diventa

$$f_{12}(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2),$$

da cui si ricava $f_{2|1}(x_2|x_1)$ dividendo ambo i membri per la densità marginale $f_1(x_1)$, ottenendo

$$f_{2|1}(x_2|x_1) = c_{12}(F_1(x_1), F_2(x_2))f_2(x_2).$$

Il terzo fattore a secondo membro della (3.3) può essere espresso in termini di c_{23} e $c_{23|1}$, dove quest'ultima è la densità della copula relativa a $F_{23|1}(x_2, x_3|x_1)$.

Partendo da

$$c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1); x_1) = \frac{f_{23|1}(x_2, x_3|x_1)}{f_{2|1}(x_2|x_1)f_{3|1}(x_3|x_1)}$$

possiamo ricavare che

$$\begin{aligned} f_{3|12}(x_3|x_1, x_2) &= \frac{f_{23|1}(x_2, x_3|x_1)}{f_{2|1}(x_2|x_1)} \\ &= c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1); x_1)f_{3|1}(x_3|x_1) \\ &= c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1); x_1)c_{13}(F_1(x_1), F_3(x_3))f_3(x_3). \end{aligned}$$

Sostituendo tutto in (3.3), si ottiene la versione completa della PCC nel caso $n = 3$, che è

$$\begin{aligned} f_{123}(x_1, x_2, x_3) &= c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1); x_1)c_{12}(F_1(x_1), F_2(x_2)) \times \\ &\quad \times c_{13}(F_1(x_1), F_3(x_3))f_1(x_1)f_2(x_2)f_3(x_3). \end{aligned} \quad (3.4)$$

Precisiamo che la (3.4) è solo una delle tre possibili decomposizioni in dimensione 3. Il numero delle decomposizioni cresce molto velocemente all'aumentare della dimensione e per esempio per $n = 5$ si possono individuare 240 diverse rappresentazioni². Al fine di organizzarle Bedford e Cooke (2001,

²Per la generica dimensione n non è ancora stata individuata una formula che determina il numero esatto delle diverse decomposizioni, per cui l'unico metodo per individuarle è l'enumerazione.

2002) hanno introdotto delle strutture grafiche denominate *vine*, di cui ci occuperemo nella sezione 3.3.

Si noti inoltre che in generale la copula $c_{23|1}$ dipende dalla variabile condizionante x_1 , non solo attraverso i suoi argomenti $F_{2|1}(x_2|x_1)$ e $F_{3|1}(x_3|x_1)$, ma anche direttamente da x_1 . Su questo aspetto torneremo comunque tra poco, dopo aver esteso la (3.4) alla generica dimensione n attraverso una possibile fattorizzazione della densità condizionata.

Per comodità di scrittura omettiamo gli argomenti delle funzioni e consideriamo le densità di copula condizionate $c_{i,j|i_1,\dots,i_k}$ per $i < j$ e $i_1 < \dots < i_k$ dove $k \in [1, n-1]$, con la quale possiamo per esempio riscrivere $f_{t|1\dots t-1}$ come

$$\begin{aligned} f_{t|1\dots t-1} &= c_{1,t|2,\dots,t-1} \cdot f_{t|2\dots t-1} = \\ &= \left[\prod_{s=1}^{t-2} c_{s,t|s+1,\dots,t-1} \right] \cdot c_{t-1,t} \cdot f_t \end{aligned} \quad (3.5)$$

Ricaviamo l'espressione della PCC in dimensione n utilizzando la (3.1) e sostituendo nella (3.5) $s = 1$ e $t = i + j$, ottenendo

$$f_{1\dots n} = \left[\prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{i,i+j|v_{ij}} \right] \cdot \prod_{k=1}^n f_k, \quad (3.6)$$

dove v_{ij} è l'insieme degli indici $\{i+1, \dots, i+j-1\}$. Teniamo comunque a precisare che la (3.6) non è l'unica fattorizzazione possibile³, ma è solo una di quelle ammissibili, perchè l'espressione finale è necessariamente condizionata dalla scelta fatta per scomporre $f_{t|1\dots t-1}$.

Gli elementi fondamentali della PCC sono quindi le *pair-copule*, ovvero le copule bivariate che hanno per argomenti le distribuzioni condizionate, fatta eccezione per quelle del livello iniziale che non sono condizionate. Il numero di variabili condizionanti aumenta di livello in livello da 1 a $n-2$, per esempio in dimensione 5 le variabili condizionanti relative all'ultimo livello saranno 3.

Considerare la dipendenza da tutte queste variabili condizionanti può risultare molto difficile quando ci si occupa di inferenza in campo parametrico, ma può diventare addirittura impossibile in ambito non-parametrico. Per questo motivo si richiede l'assunzione che le *pair-copule* non dipendano dalle variabili condizionanti, eccetto che attraverso le distribuzioni condizionate che hanno come argomento. Sotto questa condizione si ottiene quella che viene comunemente definita PCC semplificata, che nel caso tridimensionale diventa

³Vedremo nella sezione 3.5 che la (3.6) corrisponde alla densità di un D-vine.

$$f_{123}(x_1, x_2, x_3) = c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1))c_{12}(F_1(x_1), F_2(x_2)) \times \\ \times c_{13}(F_1(x_1), F_3(x_3))f_1(x_1)f_2(x_2)f_3(x_3).$$

3.2 Proprietà della PCC semplificata

In questa sezione cercheremo di caratterizzare le distribuzioni per le quali la decomposizione semplificata della PCC è ammissibile oppure no, per fare questo ci baseremo sul lavoro di Hobaek Haff et al. (2010), che è un buon riferimento per la maggior parte dei risultati proposti in questa sezione.

Nonostante il coefficiente di correlazione lineare sia la misura di dipendenza maggiormente utilizzata⁴, non lo si ritiene in generale adeguato a descrivere le condizioni sotto le quali vale la forma semplificata. Infatti gli si preferiscono gli indicatori di correlazione di rango, come il τ di Kendall e il ρ di Spearman, o i coefficienti di dipendenza di coda che dipendono direttamente dalle distribuzioni bivariate delle variabili.

Prendiamo n variabili aleatorie X_1, \dots, X_n aventi funzione di densità congiunta $f_{1\dots n}(x_1, \dots, x_n)$ e consideriamo la generica decomposizione in (3.6), allora valgono le seguenti proprietà:

1. Se per ogni copula $C_{ij|v_{ij}}$ i relativi τ di Kendall (rispettivamente ρ di Spearman) sono tutti funzioni delle variabili condizionanti, allora la decomposizione non può essere della forma semplificata.
2. Anche se nessun τ di Kendall (rispettivamente ρ di Spearman) relativo alle varie $C_{ij|v_{ij}}$ dipende dalle variabili condizionanti, può comunque non valere la forma semplificata.
3. Se per ogni copula $C_{ij|v_{ij}}$ i relativi coefficienti di dipendenza di coda superiore (rispettivamente coefficienti di dipendenza di coda inferiore) sono tutti funzioni delle variabili condizionanti, allora la decomposizione non può essere della forma semplificata.
4. Anche se nessun coefficiente di dipendenza di coda superiore (rispettivamente coefficiente di dipendenza di coda inferiore) relativo alle varie $C_{ij|v_{ij}}$ dipende dalle variabili condizionanti, può comunque non valere la forma semplificata.

⁴Abbiamo visto nel capitolo precedente che è particolarmente legato alle distribuzioni ellittiche.

Le proprietà 1 e 3 sono dirette conseguenze delle definizioni degli indicatori considerati. Facendo infatti riferimento alle espressioni (2.9), (2.10), (2.12) e (2.13) vediamo che τ , ρ_S , λ_U e λ_L dipendono direttamente dalla copula di riferimento, pertanto ne ereditano il comportamento per quanto riguarda le variabili condizionanti. Le proprietà 2 e 4 possono essere verificate con semplici controesempi⁵.

Le condizioni viste fino ad ora descrivono tutte dei casi in cui non vale la forma semplificata, si può tuttavia dimostrare che sotto alcune condizioni una particolare famiglia di distribuzione ammette la PCC semplificata.

Teorema 3.1. *Siano X_1, \dots, X_n variabili aleatorie con densità congiunta $f_{1\dots n}(x_1, \dots, x_n)$, si consideri la decomposizione in (3.6) e si assuma inoltre che tutte le funzioni di distribuzione presenti in tale decomposizione appartengano alla famiglia delle distribuzioni di posizione-scala con i parametri di posizione e di scala dipendenti dalle variabili condizionanti, ovvero che siano della forma*

$$F_{k|v_{ij}}(x_k|\mathbf{x}_{v_{ij}}) = g_{k|v_{ij}} \left(\frac{x_k - a_{k|v_{ij}}(\mathbf{x}_{v_{ij}})}{b_{k|v_{ij}}(\mathbf{x}_{v_{ij}})} \right), \quad k = i, i + j, \quad (3.7)$$

dove $\mathbf{x}_{v_{ij}}$ è il vettore delle variabili condizionanti, $a_{k|v_{ij}}$ e $b_{k|v_{ij}}$ (rispettivamente il parametro di posizione e il parametro di scala) sono due funzioni continue a valori reali, con $b_{k|v_{ij}} > 0$, e $g_{k|v_{ij}}$ è una certa funzione di distribuzione su \mathbb{R} . Sotto queste ipotesi una condizione necessaria e sufficiente affinché la decomposizione sia della forma semplificata è che le funzioni di densità condizionate corrispondenti alle pair-copule presenti nella decomposizione assumano la seguente forma

$$f_{i,i+j|v_{ij}}(x_i, x_{i+j}|\mathbf{x}_{v_{ij}}) = \frac{h_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\mathbf{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\mathbf{x}_{v_{ij}}))}{b_{i|v_{ij}}(\mathbf{x}_{v_{ij}})b_{i+j|v_{ij}}(\mathbf{x}_{v_{ij}})}$$

dove $h_{i,i+j|v_{ij}}$ è una qualche funzione continua e non-negativa su $[0, 1]^2$.

Una conseguenza molto utile ed importante di questo teorema è che una distribuzione ellittica $E_n(\mu, \Sigma, \phi)$, con matrice di scala Σ definita positiva, può essere rappresentata da una decomposizione di PCC semplificata.

Concludiamo quindi dicendo che non tutte le distribuzioni multivariate ammettono la versione semplificata della PCC, tuttavia questa può essere ritenuta in generale una buona approssimazione, anche in quei casi in cui la forma semplificata sembra molto lontana dall'essere ammissibile. A tal proposito alcuni utili esempi sono presentati in Hobaek Haff et al. (2010).

⁵Si vedano per esempio le situazioni proposte in Hobaek Haff et al. (2010)

3.3 I vines

In questa sezione ci occuperemo di descrivere dei modelli grafici che aiutano ad organizzare le possibili decomposizioni in *pair-copule*. Questo oltre ad essere molto utile è effettivamente necessario soprattutto per distribuzioni in dimensioni elevate, infatti, come abbiamo accennato nel paragrafo 3.1, già in dimensione 5 ci sono 240 diverse PCC. I modelli grafici di cui ci occuperemo sono stati introdotti da Bedford e Cooke (2001,2002) che li hanno denominati *regular vine*.

3.3.1 Concetti introduttivi

Per definire un *vine* è prima necessario definire il concetto di albero.

Definizione 3.1. Un albero $T = \{N, E\}$ è un grafo aciclico dove con N ed E si indicano rispettivamente l'insieme dei suoi nodi e l'insieme dei suoi archi⁶.

Per un albero con n nodi ci sono al massimo $n - 1$ archi e perchè un albero sia completamente determinato è sufficiente che siano specificati tutti i suoi archi.

Possiamo ora definire le strutture di un *vine* e di un *regular vine* (R-vine).

Definizione 3.2. \mathcal{V} è un *vine* di n elementi se

1. $\mathcal{V} = (T_1, \dots, T_{n-1})$.
2. T_1 è un albero con $N_1 = \{1, \dots, n\}$ nodi e un insieme di archi E_1 ; per $i = 2, \dots, n - 1$, T_i è un albero con nodi $N_i = E_{i-1}$ con $\#N_i = n - (i - 1)$.

\mathcal{V} è un *regular vine* (R-vine) di n elementi se

3. Condizione di prossimità: per $i = 2, \dots, n - 1$, se $\{a, b\} \in E_i$, dove $a = \{a', a''\}$ e $b = \{b', b''\}$, allora è vero che $\#a \cap b = 1$.

In altre parole, un R-vine di n elementi è un insieme ordinato di $n - 1$ alberi tali che gli archi dell'albero j diventano i nodi dell'albero $j + 1$. La condizione di prossimità (illustrata in figura 3.1) assicura che due nodi nell'albero $j + 1$ sono uniti da un unico arco se questi nodi condividono un nodo comune nell'albero j .

Gli insiemi dei nodi e degli archi presentano delle strutture particolari, per esempio si osserva che nel primo albero l'insieme dei nodi contiene tutti

⁶Un arco è una coppia non ordinata di nodi.

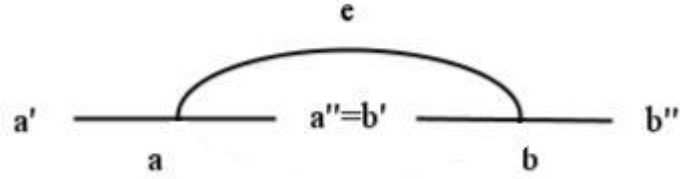


Figura 3.1: La condizione di prossimità: un arco attorno ad un nodo.

gli indici $1, \dots, n$, mentre l'insieme degli archi è un insieme di $n - 1$ coppie di questi indici. Nel secondo albero l'insieme dei nodi è costituito da coppie di indici e l'insieme degli archi da coppie di coppie di indici e così via per gli alberi di ordine superiore.

Per approfondire le proprietà dei *vine* abbiamo bisogno di definire degli insiemi associati ai suoi archi.

Definizione 3.3. Per ogni arco $e_i \in E_i$ si definisce unione completa di e_i il sottoinsieme $U_{e_i} = \{n \in N_1 | \exists e_j \in E_j, j = 1, \dots, n - 1, \text{ con } n \in e_1 \in e_2 \in \dots \in e_{i-1} \in e_i\} \subset N_1$. Per $e_i = \{a, b\} \in E_i$ con $a, b \in E_{i-1}$ $i = 1, \dots, n - 1$ si definisce insieme condizionante dell'arco e_i l'insieme $D_{e_i} = U_a \cap U_b$, mentre si definiscono insiemi condizionati di e_i $C_{e_i,a} = U_a \setminus D_{e_i}$, $C_{e_i,b} = U_b \setminus D_{e_i}$ e $C_{e_i} = C_{e_i,a} \cup C_{e_i,b} = U_a \Delta U_b$ dove $A \Delta B := (A \setminus B) \cup (B \setminus A)$ indica la differenza simmetrica di due insiemi.

I tre insiemi che abbiamo appena definito sono l'unione completa di un arco, cioè l'insieme di tutti gli indici che quell'arco contiene. Se a e b sono due nodi uniti da un arco, allora l'insieme condizionato e l'insieme condizionante sono rispettivamente la differenza simmetrica e l'intersezione dell'unione completa di a e di b .

Diamo un esempio di questi insiemi riferendoci alla figura 3.2. L'unione completa dell'arco a che unisce (23) e (24) nell'albero T_2 è l'insieme di indici $\{2, 3, 4\}$, mentre l'unione completa dell'arco b che nello stesso albero unisce (24) e (45) è l'insieme $\{2, 4, 5\}$. Inoltre gli insiemi condizionati e condizionanti dell'arco che unisce a e b sono rispettivamente gli insiemi $\{3, 5\}$ e $\{2, 4\}$.

Gli insiemi condizionati e condizionanti di tutti gli archi di \mathcal{V} sono raggruppati in un altro insieme denominato *insieme dei vincoli (constraint set)*. Ogni elemento di questo insieme è costituito da una coppia di indici corrispondente ad un insieme condizionato e ad un insieme contenente indici relativi all'insieme condizionante.

Definizione 3.4. L'insieme dei vincoli di un R -vine \mathcal{V} è l'insieme

$$\mathcal{CV} = \{(\{C_{e,a}, C_{e,b}\}, D_e) | e \in E_i, e = \{a, b\}, i = 1, \dots, n - 1\},$$

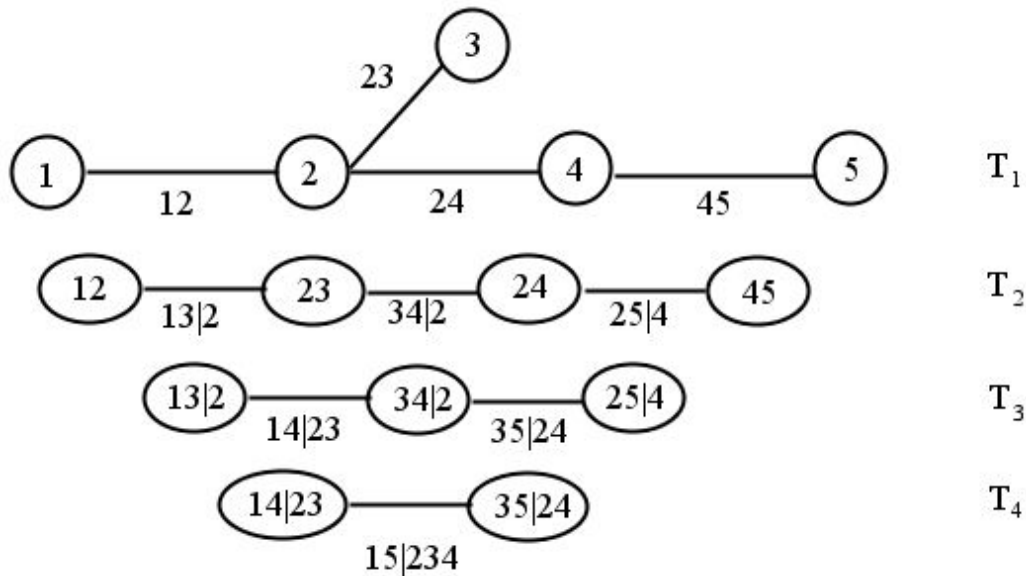


Figura 3.2: Un esempio di R-vine con 5 variabili. Per ogni arco $e = \{a, b\} \in E_i$, a destra del segno '|' si trovano gli elementi dell'insieme D_e , mentre a sinistra il primo termine è $C_{e,a}$ ed il secondo è $C_{e,b}$.

che identifica in modo univoco il R-vine a cui è associato.

Il seguente risultato è stato dimostrato in Bedford e Cooke (2002) e riguarda la cardinalità degli insiemi sin qui visti.

Teorema 3.2. *Se \mathcal{V} è un R-vine su n elementi allora per ogni $k = 1, \dots, n-1$ e per ogni $e \in E_k$ gli insiemi condizionati associati ad e contengono un unico elemento, cioè $\#C_{e,a} = 1$. Inoltre si ha che $\#U_e = k + 1$ e $\#D_e = k - 1$.*

Siccome in un R-vine gli insiemi condizionati sono formati da un unico indice, risulta conveniente contrassegnare i nodi degli alberi tramite i suoi insiemi condizionati e condizionanti. In particolare è comodo riferirsi all'arco $e = \{a, b\}$ come se fosse $a, b|D_e$.

I *regular vine* sono caratterizzati da molte altre interessanti proprietà che si possono trovare in Bedford e Cooke (2002):

- Ogni coppia di indici compare una e una sola volta come insieme condizionato di un arco in un R-vine
- Se gli insiemi condizionati di due archi a e b in un R-vine sono uguali, allora $a = b$.

La classe dei R-vine è molto ampia e al suo interno si possono identificare diverse strutture, in particolare due famiglie di *vine* sono state studiate in letteratura: i D-vine e i C-vine⁷.

Definizione 3.5. *Un regular vine è chiamato:*

- D-vine se ogni nodo nell'albero T_1 ha grado al massimo 2;
- C-vine se in ogni albero T_i c'è un unico nodo di grado $n - i$.

In un D-vine i nodi del primo albero hanno tutti grado uno o due, caratteristica che gli conferisce la classica struttura di cammino, come rappresentato in figura 3.3.

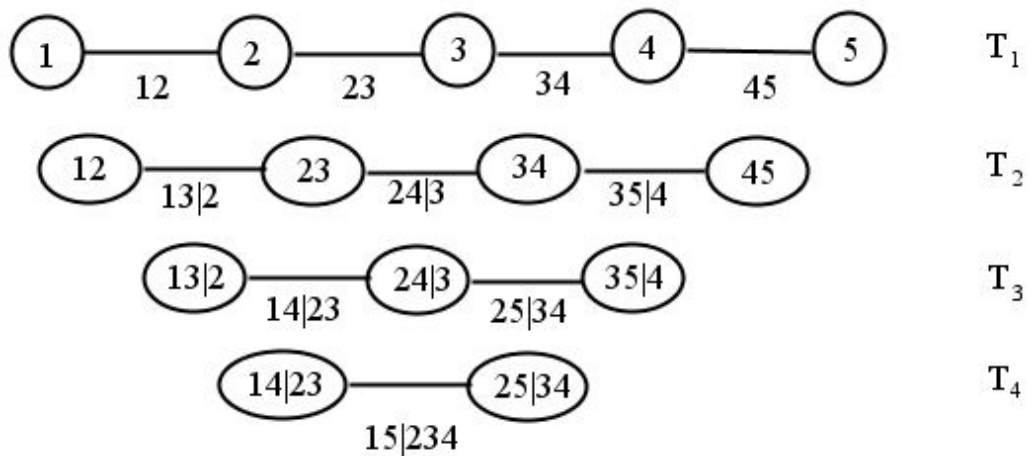


Figura 3.3: Un esempio di D-vine con 5 variabili.

L'albero di ogni C-vine ha un unico nodo di grado massimo a cui sono legati tutti gli altri nodi, la sua struttura tipica è quella di stella, come riportato in figura 3.4.

Per concludere questa sezione, facciamo notare che lavorare con queste due famiglie di R-vine è spesso conveniente perchè la loro struttura è completamente determinata dall'albero di primo livello (D-vine) e dalla gerarchia di condizionamento delle variabili (C-vine). Queste proprietà ci permettono anche di calcolare il numero dei diversi D-vine e C-vine che si possono ottenere con n variabili.

⁷Il termine D-vine deriva dall'abbreviazione del termine *drawable vine*, mentre C-vine sta per *canonical vine*.

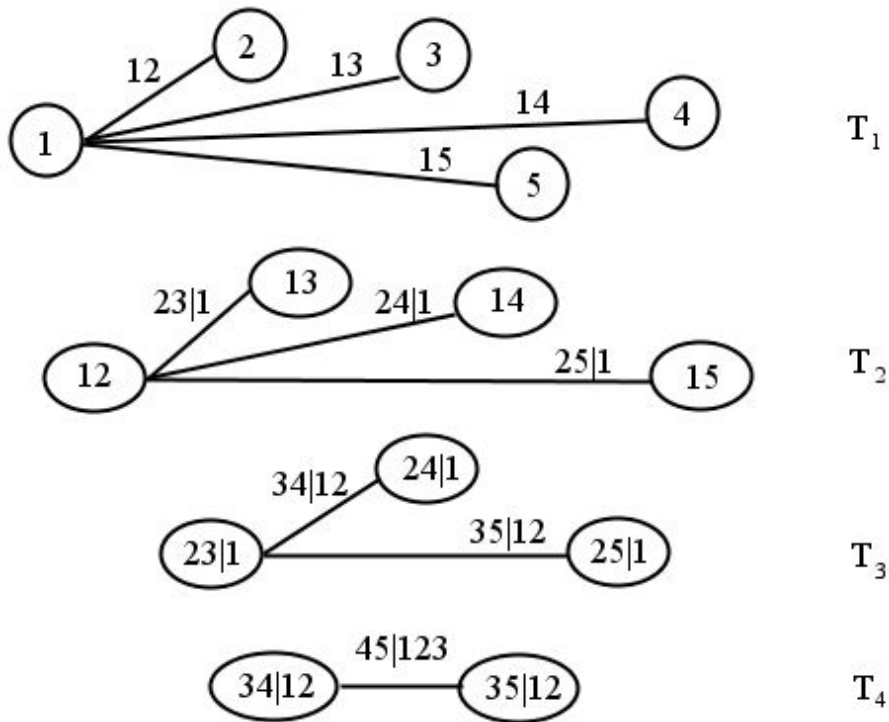


Figura 3.4: Un esempio di C-vine con 5 variabili.

Per un D-vine n -dimensionale ci sono $n!$ diversi ordinamenti per le variabili in T_1 , tuttavia gli archi di un generico *vine* sono archi non orientati⁸, perciò l'ordine dei nodi nel primo albero di un D-vine può essere invertito senza cambiare il tipo di *vine*. Quindi il numero di combinazioni che porta a D-vine effettivamente diversi si dimezza riducendosi a $\frac{n!}{2}$.

Nei C-vine l'insieme condizionante è lo stesso per tutti gli archi di uno stesso albero, per esempio in figura 3.4 $\{1\}$ e $\{1, 2\}$ sono gli insiemi condizionanti di tutti gli archi, rispettivamente, degli alberi T_2 e T_3 . In generale, per la scelta dell'unico elemento dell'insieme condizionante nell'albero T_2 ci sono n scelte possibili, per l'albero T_3 sono $n - 1$ e così via fino all'albero T_{n-1} per il quale rimangono 3 possibilità di scelta. Quindi in totale ci sono $n(n - 1) \cdots 3 = \frac{n!}{2}$ diversi C-vine con n variabili.

⁸Si dice che un arco è non orientato se $a, b|D_e = b, a|D_e$, dove (a, b) è una qualsiasi coppia di indici e D_e un insieme condizionante.

3.3.2 Il legame tra R-vine e PCC

Come abbiamo già detto introducendo i R-vine, il loro scopo è quello di fornire un aiuto nell'organizzazione delle varie decomposizioni in PCC, in particolare il legame tra queste strutture grafiche e la PCC avviene attraverso l'abbinamento degli archi dei R-vine con le *pair-copule*.

Ad ogni R-vine corrisponde una *R-vine copula specification* che ora definiamo formalmente come in Bedford e Cooke (2002).

Definizione 3.6. $(\mathbf{F}, \mathcal{V}, B)$ è una R-vine copula specification se

1. $\mathbf{F} = (F_1, \dots, F_n)$ è un vettore di funzioni di distribuzioni continue e invertibili,
2. \mathcal{V} è un R-vine di n elementi,
3. $B = \{B_e | i = 1, \dots, n-1; e \in E_i\}$ è un insieme di copule dove B_e è una copula bivariata, una cosiddetta pair-copula.

Una distribuzione congiunta F di un vettore aleatorio (X_1, \dots, X_n) si dice che realizza una *R-vine copula specification* $(\mathbf{F}, \mathcal{V}, B)$ se per ogni arco $e = \{a, b\} \in E_i$, con $i = 1, \dots, n-1$, allora B_e è la copula bivariata di $X_{C_{e,a}}$ e di $X_{C_{e,b}}$ condizionata da $\mathbf{X}_{D_e} = \{X_i | i \in D_e\}$. Dovrà inoltre essere che F_j è la distribuzione marginale di X_j per $j = 1, \dots, n$. Indichiamo infine la densità della copula B_e per l'arco $e = \{a, b\}$ con $c_{C_{e,a}, C_{e,b} | D_e}$. Si può anche semplicemente dire che una distribuzione F con queste caratteristiche è la distribuzione di un R-vine.

Se consideriamo il R-vine riportato in figura 3.2 abbiamo bisogno di definire le copule non condizionate $c_{12}, c_{23}, c_{24}, c_{45}$ per l'albero T_1 , le copule condizionate $c_{13|2}, c_{34|2}$ e $c_{25|4}$ per l'albero T_2 e così via. La flessibilità di questa struttura permette di scegliere le varie copule anche da famiglie differenti e di stimarne i parametri in modo quasi del tutto indipendente. Non c'è una perfetta indipendenza perchè le scelte delle copule di un albero incidono sulle variabili condizionate degli alberi successivi.

Il risultato principale dimostrato in Bedford e Cooke (2001, 2002) è che la densità di un R-vine può essere fattorizzata nel prodotto delle copule bivariate condizionate e non condizionate associate a ciascuno dei suoi archi.

Teorema 3.3. *Se $(\mathbf{F}, \mathcal{V}, B)$ è una R-vine copula specification su n elementi, allora è unica la funzione di distribuzione del R-vine e la sua funzione di densità è*

$$f_{1..n}(\mathbf{x}) = \prod_{k=1}^n f_k(x_k) \prod_{i=1}^{n-1} \prod_{e \in E_i} c_{C_{e,a}, C_{e,b} | D_e}(F_{C_{e,a} | D_e}(x_{C_{e,a}} | \mathbf{x}_{D_e}), F_{C_{e,b} | D_e}(x_{C_{e,b}} | \mathbf{x}_{D_e})), \quad (3.8)$$

dove $\mathbf{x} = (x_1, \dots, x_n)$, $e = \{a, b\}$ e \mathbf{x}_{D_e} è il vettore delle variabili in D_e , cioè $\mathbf{x}_{D_e} = \{x_i | i \in D_e\}$. Inoltre le f_i sono le densità marginali corrispondenti alle distribuzioni marginali F_i con $i = 1, \dots, n$.

E' interessante notare che le copule presenti in (3.8) sono indicizzate dagli elementi dell'insieme dei vincoli \mathcal{CV} che, come sappiamo, è sufficiente a identificare in modo univoco un R-vine.

In Aas et al. (2009) è stata introdotta un'espressione delle distribuzioni marginali condizionate $F_{C_{e,a} | D_e}(x_{C_{e,a}} | \mathbf{x}_{D_e})$ e $F_{C_{e,b} | D_e}(x_{C_{e,b}} | \mathbf{x}_{D_e})$ che risulterà molto utile in campo inferenziale. Se indichiamo con $e = \{a, b\} \in E_i$, $a = \{a_1, a_2\}$, $b = \{b_1, b_2\}$ l'arco che unisce $C_{e,a}$ con $C_{e,b}$ date le variabili D_e , allora

$$\begin{aligned} F_{C_{e,a} | D_e}(x_{C_{e,a}} | \mathbf{x}_{D_e}) &= \frac{\partial C_{C_a | D_a}(u, v)}{\partial v} \Big|_{(u,v)=(F_{C_{a,a_1} | D_a}(x_{C_{a,a_1}} | \mathbf{x}_{D_a}), F_{C_{a,a_1} | D_a}(x_{C_{a,a_2}} | \mathbf{x}_{D_a}))} \\ &:= h(F_{C_{a,a_1} | D_a}(x_{C_{a,a_1}} | \mathbf{x}_{D_a}), F_{C_{a,a_1} | D_a}(x_{C_{a,a_2}} | \mathbf{x}_{D_a})), \end{aligned} \quad (3.9)$$

dove le distribuzioni $F_{C_{e,a_1} | D_e}(x_{C_{e,a_1}} | \mathbf{x}_{D_e})$ e $F_{C_{e,a_2} | D_e}(x_{C_{e,a_2}} | \mathbf{x}_{D_e})$ si possono ricavare ricorsivamente. In modo del tutto analogo si ottiene $F_{C_{e,b} | D_e}(x_{C_{e,b}} | \mathbf{x}_{D_e})$.

Per finire, le distribuzioni $F_{C_{e,a} | D_e}(x_{C_{e,a}} | \mathbf{x}_{D_e})$ e $F_{C_{e,b} | D_e}(x_{C_{e,b}} | \mathbf{x}_{D_e})$ vengono chiamate *variabili trasformate*.

3.4 Il caso dei D-vine

Dopo una panoramica generale sulla PCC e i *regular vine*, in questa sezione focalizzeremo la nostra attenzione sulla classe dei D-vine che è sicuramente una delle più studiate anche in virtù della sua maggiore adattabilità a modellare situazioni reali.

In particolare presenteremo la tecnica di simulazione da un D-vine e il procedimento per la selezione del modello migliore per un dataset assegnato.

3.4.1 Simulazione da D-vine

Metodi di simulazione di *vine* sono stati brevemente discussi in Bedford e Cooke (2001), successivamente ripresi da Aas et al. (2009) per quanto riguarda i C-vine e i D-vine e da Dissman et al. (2011) per una generalizzazione

ai R-vine. Abbiamo ripreso il lavoro di Aas et al. (2009) sui D-vine, che era tuttavia limitato al caso di distribuzioni marginali uniformi, e lo abbiamo esteso al caso di distribuzioni marginali qualsiasi.

Iniziamo innanzitutto presentando la tecnica su cui si basa l'algoritmo di simulazione di un campione (x_1, \dots, x_n) con distribuzioni marginali uniformi e struttura di dipendenza descritta da un D-vine. Vanno estratti indipendentemente w_i con $i = 1, \dots, n$ da una uniforme $[0, 1]$ e poi si pone

- $x_1 = w_1$
- $x_2 = F_{2|1}^{-1}(w_2|x_1)$
- $x_3 = F_{3|12}^{-1}(w_3|x_1, x_2)$
- $\dots = \dots$
- $x_n = F_{n|1\dots n-1}^{-1}(w_n|x_1, \dots, x_{n-1})$.

Le distribuzioni condizionate $F_{i|1,\dots,i-1}$ si ricavano utilizzando le funzioni h definite dalla relazione (3.9), che nel caso dei D-vine sono del seguente tipo⁹

$$F_{i|1,\dots,i-1} = \frac{\partial C_{1,i|2,\dots,i-1}(u, v)}{\partial v} \Big|_{(u,v)=(F_{i|2,\dots,i-1}, F_{1|2,\dots,i-1})}$$

La procedura per simulare da un D-vine con distribuzioni marginali qualsiasi è descritta nell'Algoritmo 1, che è formato da un ciclo *for* contenente altri due cicli *for*, uno che restituisce i valori campionati e un altro che fornisce le distribuzioni condizionate necessarie per i calcoli. I valori delle prime due variabili sono però determinati esternamente da questi cicli. Le distribuzioni condizionate $v_{ij} = F_{i|1,\dots,j-1}$ sono ricavate utilizzando ricorsivamente le funzioni h i cui argomenti sono le variabili trasformate e l'insieme $\Theta_{j,i}$ contenente i valori dei parametri della *pair-copula* di densità $c_{j,j+i|1,\dots,j-1}$.

3.4.2 Selezione di un modello di D-vine

Per adattare un modello di D-vine ad un dataset reale, i passaggi richiesti sono i seguenti:

- (a) Selezionare una specifica struttura (fattorizzazione) di D-vine, cioè determinare quali coppie di variabili condizionate e non condizionate considerare.

⁹Per semplicità di notazione sono stati omessi gli argomenti delle funzioni di distribuzione.

Algoritmo 1**OUTPUT:** Campione (x_1, \dots, x_n) da un D-vine di marginali F_i

```
1 Estrarre  $w_1$  dalla distribuzione  $F_1$ 
2 Estrarre  $w_i$  indipendenti da un'uniforme  $[0, 1]$ ;  $i = 2, \dots, n$ 
3  $x_1 = v_{1,1} = w_1$ 
4  $x_2 = v_{2,1} = F_2^{-1}(h^{-1}(w_2, F_1(v_{1,1}), \Theta_{1,1}))$ 
5  $v_{2,2} = h(F_1(v_{1,1}), F_2(v_{2,1}), \Theta_{1,1})$ 
6 for  $i = 3, 4, \dots, n$ 
7    $v_{i,1} = w_i$ 
8   for  $k = i - 1, i - 2, \dots, 2$ 
9      $v_{i,1} = h^{-1}(v_{i,1}, v_{i-1,2k-2}, \Theta_{k,i-k})$ 
10  end for
11   $v_{i,1} = F_i^{-1}(h^{-1}(v_{i,1}, F_{i-1}(v_{i-1}), \Theta_{1,i-1}))$ 
12   $x_i = v_{i,1}$ 
13  if  $i = n$  then
14    STOP
15  end if
16   $v_{i,2} = h(F_{i-1}(v_{i-1,1}), F_i(v_{i,1}), \Theta_{1,i-1})$ 
17   $v_{i,3} = h(F_i(v_{i,1}), F_{i-1}(v_{i-1,1}), \Theta_{1,i-1})$ 
18  if  $i > 3$  then
19    for  $j = 2, 3, \dots, i - 2$ 
20       $v_{i,2j} = h(v_{i-1,2j-2}, v_{i,2j-1}, \Theta_{j,i-j})$ 
21       $v_{i,2j+1} = h(v_{i,2j-1}, v_{i-1,2j-2}, \Theta_{j,i-j})$ 
22    end for
23  end if
24   $v_{i,2i-2} = h(v_{i-1,2i-4}, v_{i-1,2j-2}, \Theta_{i-1,1})$ 
25 end for
```

(b) Per ciascuna coppia definita in (a), identificare la classe della copula da utilizzare.

(c) Stimare i parametri relativi a ciascuna copula selezionata.

Ora analizziamo nel dettaglio ciascuna di queste operazioni separatamente, ma in seguito vedremo come si possono combinare insieme per determinare il D-vine ‘migliore’ relativamente ad un dataset reale.

(a) Selezione struttura D-vine :

Selezionare un D-vine per un dataset assegnato significa decidere per quali coppie di variabili definire delle copule. A tale proposito, descrivendo i diversi modelli di *vine* abbiamo già avuto modo di discutere

l'utile proprietà dei D-vine riguardo alla loro struttura, che risulta completamente determinata dall'ordine delle variabili nell'albero di primo livello, che definisce in modo univoco le strutture degli alberi successivi.

Per un D-vine si tratta quindi di selezionare l'albero T_1 in modo tale che le coppie di variabili scelte esprimano i livelli più alti di dipendenza. Come misura di dipendenza una scelta adeguata ci sembra il τ di Kendall, dato che dipende esclusivamente dalla distribuzione bivariata ed è indipendente dal comportamento marginale delle variabili.

Il metodo basato sul τ di Kendall prevede di calcolare i valori empirici di questo indicatore per tutte le possibili coppie di variabili (i, j) , con $1 \leq i \leq j \leq n$, e di selezionare l'albero che massimizza la somma dei valori assoluti dei τ di Kendall. Si considerano i valori assoluti anziché i valori con segno per evitare che dipendenze di segno opposto, magari anche molto forti, si compensino tra loro.

Siccome gli alberi dei D-vine hanno la forma di un cammino, trovare la soluzione ottima equivale a risolvere il problema del Commesso Viaggiatore¹⁰. Tuttavia per tale problema non esiste un algoritmo risolutivo efficiente, quindi un possibile metodo di soluzione è rappresentato dall'enumerazione totale di tutti i cammini possibili, che per un D-vine con n variabili ricordiamo essere $\frac{n!}{2}$.

(b) Scelta delle copule :

Una volta che sono state individuate le coppie di variabili condizionate e non condizionate, è necessario scegliere per ciascuna di esse la famiglia di copule da utilizzare.

Le *pair-copule* non devono necessariamente appartenere tutte alla stessa classe, ma per ogni coppia andrebbe scelta quella che meglio si adatta ai dati. Per fare questo può essere utile ricorrere al test di verifica delle informazioni di Akaike (AIC test, Akaike (1973)) che corregge la log-verosimiglianza di una copula con il numero dei suoi parametri. Siccome l'AIC test prevede di scegliere la soluzione con il valore del test inferiore, possono risultare penalizzate quelle famiglie di copule che dipendono da più parametri, come per esempio la copula t che è definita da due parametri. Un altro criterio che è si potrebbe utilizzare è il test di *goodness-of-fit* (GOF), che per le strutture di dipendenza è sostanzialmente un caso speciale del più generale test per le strutture multivariate (per una sua applicazione si veda Aas et al. (2009)).

¹⁰In teoria dei grafi il problema del Commesso Viaggiatore si formula nel seguente modo: dato un grafo completo pesato, trovare il cammino di peso minimo toccando tutti i nodi una sola volta e tornando al nodo di partenza.

(c) Stima dei parametri delle copule :

Il metodo maggiormente utilizzato in letteratura per la stima dei parametri delle copule di una PCC è denominato *stima sequenziale* che permette di ottenere dei buoni valori di partenza per la stima di massima verosimiglianza. Come metodo è stato proposto da Aas et al. (2009) e le sue proprietà asintotiche sono state studiate da Hobaek Haff (2010).

Ipotizziamo di disporre di n variabili aleatorie osservate in T istanti temporali differenti e che quindi il dataset di riferimento sia formato da vettori del tipo $\mathbf{x}_i = (x_{i1}, \dots, x_{iT})$ per $i = 1, \dots, n$.

Omettendo per semplicità di notazione gli argomenti delle distribuzioni condizionate, la funzione di log-verosimiglianza per un D-vine è data da

$$\sum_{j=1}^{n-1} \sum_{i=1}^{n-j} \sum_{t=1}^T \log \left(c_{i,i+j|i+1,\dots,i+j-1} \left(F_{i_t|(i+1)_t,\dots,(i+j-1)_t}, \right. \right. \\ \left. \left. F_{(i+j)_t|(i+1)_t,\dots,(i+j-1)_t}; \Theta_{j,i} \right) \right),$$

dove $\Theta_{j,i}$ è l'insieme dei parametri della copula $c_{i,i+j|i+1,\dots,i+j-1}$.

La funzione di log-verosimiglianza deve essere massimizzata numericamente e i valori di partenza dei parametri possono essere determinati in modo sequenziale (da qui il nome di *stima sequenziale*):

1. Stimare i parametri delle copule dell'albero T_1 utilizzando le osservazioni originali.
2. Determinare le variabili trasformate, che ricordiamo essere le distribuzioni condizionate, per l'albero T_2 utilizzando le funzioni h e i parametri stimati al punto 1.
3. Stimare i parametri delle copule dell'albero T_2 utilizzando i valori delle variabili trasformate del punto 2.
4. Ripetere le operazioni dei punti 2 e 3 per tutti gli alberi successivi fino all'ultimo albero T_{n-1} .

Dopo aver analizzato più nel dettaglio tutte le operazioni necessarie per determinare un D-vine, è chiaro che per individuare il modello 'migliore' sarebbe necessario realizzare e porre a confronto i risultati dei punti (b) e (c) per

tutte le possibili costruzioni di D-vine e per le diverse famiglie di copule che si intendesse testare.

Per superare questo problema in Dissman et al. (2011) è stato proposto un metodo euristico per selezionare un R-vine e che noi abbiamo adattato al caso del D-vine nel seguente modo:

1. Calcolare i valori empirici del τ di Kendall per tutte le possibili coppie di variabili dell'albero di primo livello (i, j) , con $1 \leq i \leq j \leq n$.
2. Scegliere il cammino che massimizza la somma dei valori assoluti dei τ di Kendall calcolati in 1, cioè

$$\max \sum_{e=\{i,j\}} |\hat{\tau}_{ij}|.$$

Il cammino 'migliore' corrisponde all'albero T_1 che in questo modo identifica anche la struttura di tutti gli alberi successivi.

3. Per ciascuna delle coppie individuate selezionare la copula e stimarne i relativi parametri.
4. Calcolare le variabili trasformate utilizzando le copule e i parametri stimati al punto 3 insieme con le funzioni h .
5. Ripetere le operazioni dal punto 3 fino all'ultimo albero T_{n-1} .

Torneremo su questo nel capitolo 5 quando cercheremo di adattare il modello di D-vine ad un dataset reale.

3.5 Simulazione di un D-vine con *pair-copule* Gaussiane

In questa sezione simuleremo da un D-vine utilizzando l'Algoritmo 1 presentato nella sezione 3.4.1 per dimostrare la validità della forma semplificata della PCC nel caso di 4 variabili aleatorie e di *pair-copule* Gaussiane. Per fare questo sfruttiamo il fatto che la distribuzione di un D-vine, applicata a variabili aleatorie distribuite marginalmente come delle normali e abbinata ad opportune scelte di *pair-copule* Gaussiane, coincide con una distribuzione normale multivariata.

Per il nostro esempio abbiamo considerato un vettore aleatorio normale multivariato $\mathbf{X} = (X_1, X_2, X_3, X_4)$ con vettore delle medie nullo e con i valori della matrice di Varianza-Covarianza Σ riportati in tabella 3.1.

Σ	X_1	X_2	X_3	X_4
X_1	1,00	0,10	0,66	0,90
X_2	0,10	0,25	0,33	0,50
X_3	0,66	0,33	2,00	0,20
X_4	0,90	0,50	0,20	5,00

Tabella 3.1: Matrice di Varianza-Covarianza Σ .

Abbiamo quindi creato un dataset di dati osservati simulando 500.000 valori da tale distribuzione¹¹.

La struttura di D-vine che abbiamo deciso di utilizzare è quella riportata in figura 3.5, la cui funzione di densità congiunta¹² è del tipo della (3.6)

$$f_{1234} = c_{14|23}(F_{1|23}, F_{4|23})c_{13|2}(F_{1|2}, F_{3|2})c_{24|3}(F_{2|3}, F_{4|3}) \times c_{12}(F_1, F_2)c_{23}(F_2, F_3)c_{34}(F_3, F_4)f_1f_2f_3f_4, \quad (3.10)$$

in cui ci sono tre copule condizionate e altrettante non condizionate.

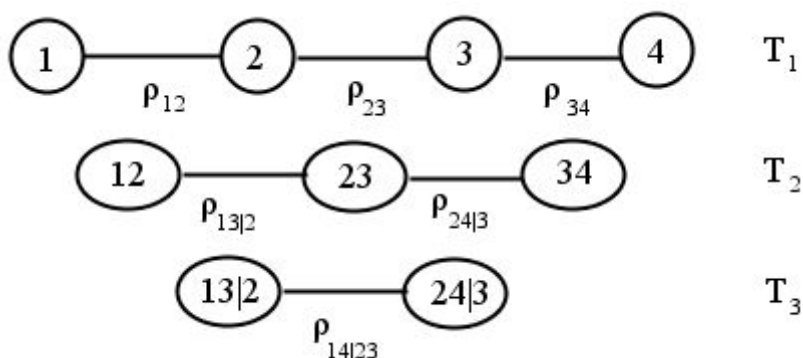


Figura 3.5: D-vine con 4 variabili e *pair-copule* Gaussianie: sotto ad ogni arco è riportato il parametro della copula a cui è associato.

Tutte le *pair-copule* presenti in (3.10) appartengono alla famiglia Gaussiana il cui unico parametro è il coefficiente di correlazione lineare ρ .

Per poter effettuare un confronto tra i risultati della PCC semplificata e di quella generale abbiamo voluto mettere a confronto le matrici di Varianza-Covarianza di 100.000 osservazioni simulate rispettivamente da un D-vine con

¹¹Il metodo che abbiamo utilizzato per simulare da una distribuzione normale multivariata è quello basato sulla decomposizione di Cholesky della matrice Σ . Per maggiori dettagli si veda McNeil et al. (2006).

¹²Per comodità di notazione sono stati omessi gli argomenti delle funzioni di distribuzione e di densità.

copule condizionate non dipendenti direttamente dalle variabili condizionanti e da un D-vine con copule il cui parametro è invece funzione dei valori delle variabili condizionanti.

3.5.1 PCC semplificata

Con riferimento alla figura 3.5 l'ipotesi di PCC semplificata riguarda le copule degli alberi T_2 e T_3 , cioè $c_{13|2}$, $c_{24|3}$ e $c_{14|23}$, i cui parametri sono rispettivamente le correlazioni condizionate $\rho_{13|2}$, $\rho_{24|3}$ e $\rho_{14|23}$.

Un metodo alternativo alla stima sequenziale presentata nella sezione precedente sfrutta una particolare proprietà della distribuzione normale multivariata che permette di calcolare le correlazioni condizionate in modo ricorsivo a partire dalle correlazioni semplici. Come dimostrato in Morales et al. (2006) nel caso di normali multivariate le correlazioni parziali coincidono con le correlazioni condizionate¹³, pertanto nel nostro caso abbiamo che

$$\rho_{13|2} = \rho_{13;2} = \frac{\rho_{13} - \rho_{12}\rho_{23}}{\sqrt{(1 - \rho_{12}^2)(1 - \rho_{23}^2)}}, \quad \rho_{24|3} = \rho_{24;3} = \frac{\rho_{24} - \rho_{23}\rho_{34}}{\sqrt{(1 - \rho_{23}^2)(1 - \rho_{34}^2)}},$$

$$\rho_{14|23} = \rho_{14;23} = \frac{\rho_{14|3} - \rho_{12|3}\rho_{24|3}}{\sqrt{(1 - \rho_{12|3}^2)(1 - \rho_{24|3}^2)}}.$$

Nella tabella 3.2 riportiamo i valori dei parametri delle pair-copule condizionate ottenuti con il calcolo esatto attraverso le correlazioni parziali e con il metodo della stima sequenziale. Notiamo che questi ultimi sottostimano leggermente i valori esatti, che infatti abbiamo preferito utilizzare per le simulazioni successive.

	$\rho_{13 2}$	$\rho_{24 3}$	$\rho_{14 23}$
Stima sequenziale	0,413	0,458	0,477
Calcolo esatto	0,432	0,474	0,493

Tabella 3.2: D-vine con 4 variabili, copule Gaussian e sotto l'ipotesi di PCC semplificata: confronto tra i parametri delle copule condizionate determinati con la stima sequenziale e con il calcolo esatto.

Potendo quindi calcolare tutti i parametri delle *pair-copule* e utilizzando la funzione h (calcolata nel caso Gaussiano in Appendice A.1), sono state

¹³La correlazione parziale si indica con $\rho_{ij;v_{-ij}}$, dove v_{-ij} è l'insieme $\{1, \dots, n\} \setminus \{i, j\}$. Per le definizioni di correlazione condizionate e di correlazione parziale rimandiamo a Kendall e Stuart (1967).

simulate 100.000 osservazioni utilizzando l'Algoritmo 1. Nella tabella 3.3 riportiamo i valori di Σ^* , matrice di Varianza-Covarianza dei valori simulati in ipotesi di PCC semplificata.

Σ^*	X_1^*	X_2^*	X_3^*	X_4^*
X_1^*	0,990	0,099	0,664	0,885
X_2^*	0,099	0,251	0,326	0,503
X_3^*	0,664	0,326	1,994	0,189
X_4^*	0,885	0,503	0,189	5,000

Tabella 3.3: Matrice di Varianza-Covarianza di 100.000 osservazioni simulate da un D-vine con 4 variabili, copule Gaussianhe e sotto l'ipotesi di PCC semplificata.

3.5.2 PCC generale

Per la PCC generale le copule condizionate $c_{13|2}$, $c_{24|3}$ e $c_{14|23}$ dipendono direttamente dai valori delle variabili condizionanti, pertanto non sono valori costanti le correlazioni condizionate $\rho_{13|2}$, $\rho_{24|3}$ e $\rho_{14|23}$.

Il metodo empirico che abbiamo utilizzato per adattare l'Algoritmo 1 prevede di inserire prima delle righe 9, 20, 21 e 24 (ovvero tutte le volte in cui sono richieste le correlazioni condizionate) un controllo che permetta di aggiornare i parametri con i valori correnti delle variabili condizionanti.

Prendiamo per esempio il parametro $\rho_{13|2}$, per il quale abbiamo costruito il vettore di controllo $\rho_{X_1 X_3 | X_2 = x_2}$ utilizzando le osservazioni congiunte di (X_1, X_2, X_3) suddivise in sezioni; ciascuna sezione era individuata da intervalli di ampiezza 0,5 dei valori assunti dalla variabile X_2 . Qualora si fosse presentata la necessità di utilizzare $\rho_{13|2}$ sarebbe stato sufficiente considerare la correlazione tra i valori di X_1 e di X_3 contenuti nella sezione individuata dal valore x_2 della variabile X_2 .

In modo del tutto analogo abbiamo trattato $\rho_{24|3}$, mentre per $\rho_{14|23}$ abbiamo considerato le osservazioni congiunte di (X_1, X_2, X_3, X_4) suddividendole prima in sezioni rispetto a X_2 e successivamente ciascuna di queste sezioni è stata suddivisa in sottosezioni in base ai valori di X_3 .

Con queste modifiche all'Algoritmo 1 abbiamo potuto simulare 100.000 osservazione da un D-vine con copule Gaussianhe, ma senza sfruttare l'ipotesi di PCC semplificata. Nella tabella 3.4 riportiamo i valori di Σ^{**} , matrice di Varianza-Covarianza dei valori simulati.

Σ^{**}	X_1^{**}	X_2^{**}	X_3^{**}	X_4^{**}
X_1^{**}	1,004	0,102	0,668	0,865
X_2^{**}	0,102	0,251	0,326	0,503
X_3^{**}	0,668	0,326	1,993	0,187
X_4^{**}	0,865	0,503	0,187	5,025

Tabella 3.4: Matrice di Varianza-Covarianza di 100.000 osservazioni simulate da un D-vine con 4 variabili, copule Gaussianhe e senza l'ipotesi di PCC semplificata.

3.5.3 Conclusioni

I valori di Σ^* e di Σ^{**} (riportati nelle tabelle 3.3 e 3.4) sono molto prossimi tra loro e approssimano in maniera quasi perfetta i valori reali della matrice di Varianza-Covarianza Σ (tabella 3.1). Questo conferma quanto discusso nella sezione 3.2 circa l'ammissibilità della forma semplificata della PCC per la distribuzione Gaussiana che è di tipo ellittico.

Come considerazione conclusiva vogliamo far notare che il metodo da noi utilizzato nel caso di PCC generale, ovvero quello di considerare sezioni delle osservazioni congiunte, è una tecnica che si può utilizzare in dimensioni basse. Per esempio in dimensione 5 si presenta la necessità di condizionare rispetto a 3 variabili (per la copula dell'albero T_4) e questo richiederebbe la disponibilità di un gran numero di osservazioni congiunte (X_1, X_2, X_3, X_4, X_5) in modo tale che all'interno di ogni 'sotto-sottosezione' ci siano sufficienti valori per stimare i parametri.

Capitolo 4

L'avvento di Solvency II

L'obiettivo di questo capitolo è quello approfondire il problema assicurativo che abbiamo presentato brevemente nell'introduzione e che affronteremo dal punto di vista pratico nel capitolo 5.

Nel paragrafo 4.1 forniremo una panoramica della nuova direttiva comunitaria Solvency II, che è stata studiata dall'*European Insurance and Occupational Pensions Authority* (EIOPA) al fine di uniformare le discipline normative riguardanti le attività delle imprese di assicurazione e riassicurazione del territorio europeo.

Uno dei punti principali di Solvency II riguarda la quantificazione del capitale di rischio che un'impresa di assicurazione deve detenere e che può essere determinato tramite due approcci: attraverso una formula *standard* oppure un modello interno.

Presenteremo poi brevemente la struttura generale della formula *standard* con un focus particolare sul modulo di rischio riguardante il rischio di sottoscrizione del business Danni (paragrafo 4.2). Discuteremo infine nel paragrafo 4.3 la costruzione di un possibile modello interno adeguato a descrivere il profilo di rischio di una compagnia Danni.

4.1 La nuova direttiva

A fine 2009 è diventato legge comunitaria il nuovo regime di vigilanza prudenziale per le imprese di assicurazione che entrerà in vigore nel 2013. Esso è destinato a portare un profondo cambiamento nel settore assicurativo per quanto riguarda i criteri e le metriche di valutazione, condizionando di fatto le dinamiche per la creazione di valore all'interno delle compagnie.

Nella lettera del 26 gennaio 2010 l'ISVAP comunicava:

Il nuovo requisito patrimoniale di solvibilità corrisponde al capitale economico che un'impresa deve detenere per limitare la probabilità di rovina allo 0,5 per cento.

Da questo breve messaggio era già possibile percepire la portata del cambiamento che Solvency II avrebbe portato: il livello di patrimonializzazione di un'impresa non sarebbe più dipeso direttamente dal volume dei premi e delle riserve, ma dalla loro volatilità, cioè dallo specifico profilo di rischio dell'impresa stessa.

In analogia con quanto fatto nel comparto bancario da Basilea II, anche Solvency II fa ricorso alla metafora dei 'tre pilastri' per riassumere i punti portanti della nuova normativa; ne proponiamo una rappresentazione grafica in figura 4.1.

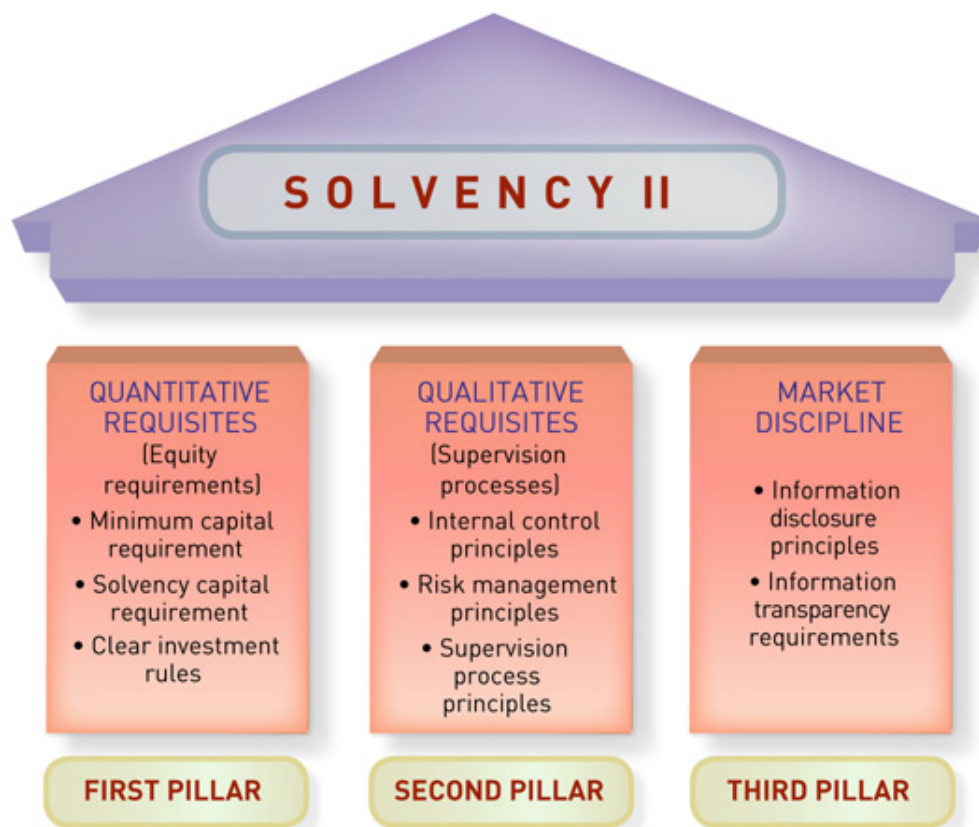


Figura 4.1: Solvency II: struttura a pilastri della normativa.

Al *primo pilastro* sono dedicati gli articoli della proposta di Direttiva dal numero 73 al numero 132, nei quali viene discusso il calcolo dei requisiti

quantitativi ed in particolare sono contenute disposizioni riguardanti:

- i criteri di valutazione delle attività e delle passività;
- la determinazione dei fondi propri e i criteri per la loro ammissibilità ai fini della copertura dei requisiti patrimoniali;
- il calcolo del *Solvency Capital Requirement* (SCR) attraverso due approcci
 - formula standard;
 - modello interno;
- il calcolo del *Minimum Capital Requirement* (MCR);
- gli investimenti a copertura delle riserve tecniche e del SCR;
- il calcolo dei requisiti patrimoniali per i gruppi assicurativi.

Il principale requisito patrimoniale del nuovo regime è il *Solvency Capital Requirement* che è definito come il *Value at Risk*¹ (VaR) dei fondi propri dell'impresa con riferimento ad un livello di confidenza del 99,5 per cento e ad un orizzonte temporale di un anno. Lo scopo del SCR è coprire la perdita inattesa e deve risultare calibrato in modo da considerare tutti i rischi quantificabili a cui l'impresa è esposta.

Nella proposta di Direttiva vengono descritti i cosiddetti *risk modules*, cioè le componenti del SCR, ossia i rischi che concorrono alla determinazione del suo ammontare, senza tuttavia precisarne le modalità di calcolo che vengono invece approfondite nelle *Technical Specifications* (TS).

Per determinare il SCR sono ammessi due approcci: tramite una formula *standard* definita dalla normativa e comune a livello europeo, oppure tramite un modello interno sviluppato dalle singole imprese di assicurazione. Riguardo alla formula *standard*, le TS descrivono oltre alle modalità di calcolo e ai parametri da utilizzare per i singoli moduli di rischio, le semplificazioni di calcolo adottate per alcune categorie di rischio e i casi in cui i parametri *standard* possono essere sostituiti con valori stimati dall'impresa. Per quanto riguarda l'uso dei modelli interni (che potranno essere totali o parziali), il loro utilizzo è subordinato all'autorizzazione dell'Autorità di vigilanza.

Un'ulteriore misura di rischio derivata dal SCR è il *Minimum Capital Requirement* che rappresenta la soglia patrimoniale minima al di sotto della

¹Definita una distribuzione delle perdite, si definisce VaR la massima perdita realizzabile rispetto ad un fissato intervallo di confidenza in un determinato orizzonte temporale.

quale il livello di rischio per gli assicurati diventa inaccettabile. E' calibrato con un livello di confidenza inferiore, ma è comunque compreso tra il 25% e il 45% del SCR. In questo contesto quindi anche la soglia di allerta non è più soltanto funzione del requisito di capitale individuato, ma ha essa stessa una natura di tipo stocastica.

Il *secondo pilastro* è disciplinato dagli articoli della proposta dal 27 al 49 e descrive i requisiti di tipo qualitativo e l'attività di vigilanza. Sono dunque individuabili due componenti: una prima componente riguardante le norme che concernono il *risk management*, la *governance* e il controllo interno dell'impresa, e una seconda che riguarda la disciplina delle attività, degli strumenti e dei poteri di vigilanza.

Per quanto riguarda il primo aspetto, è sicuramente rilevante osservare che nell'ambito del proprio sistema di *risk management* l'impresa deve effettuare con regolarità una propria valutazione dei rischi e della situazione di solvibilità. Tale attività nella terminologia del legislatore va sotto al nome di *Own Risk and Solvency Assesment* (ORSA).

Del secondo aspetto è invece importante il potere attribuito ai supervisori che possono imporre alla singola impresa dei *capital add-ons*, ovvero dei requisiti patrimoniali aggiuntivi rispetto a quelli determinati in base alla formula standard o all'eventuale modello interno.

Infine negli articoli della proposta 35 e dal 50 al 55 vengono descritti i contenuti del *terzo pilastro* che disciplina l'informativa e la reportistica alla vigilanza e al mercato.

Esso prevede, oltre ad una serie di norme riguardanti i contenuti e le modalità dell'informativa alla vigilanza, gli obblighi informativi verso il mercato, in termini di *governance*, *risk management*, grado di solvibilità ed eventualmente il mancato rispetto dei requisiti di capitale.

Ci interessa far notare che una separazione tecnica tra i tre pilastri potrebbe condurre a fraintendimenti riguardo ad una separazione di pertinenze nel trattamento dei rischi. Per esempio il primo e secondo pilastro rappresentano semplicemente due diverse visuali dello stesso complesso in cui l'aspetto quantitativo riveste il ruolo di 'protocollo standard' grazie al quale il giudizio qualitativo e l'intervento preventivo si fondano su un confronto rigoroso e oggettivo con le dinamiche del mercato.

4.2 La formula *standard*

Come abbiamo già detto nel paragrafo precedente, la proposta di Direttiva non contiene indicazioni su come il requisito di capitale vada calcolato, perchè la forma definitiva della formula *standard* non è ancora stata stabilita. A tale

proposito sono stati proposti negli anni alle imprese di assicurazione diversi esercizi (i cosiddetti *Quantitative Impact Study*, QIS) allo scopo di valutare quale potesse essere l'effetto dei capitali richiesti dal nuovo sistema sul proprio patrimonio. Attualmente sono stati svolti cinque diversi studi di questo tipo, l'ultimo risalente al luglio 2009 e di cui ora descriviamo brevemente la struttura della formula *standard*².

Il *Solvency Capital Requirement* secondo il QIS5 è calcolato nel seguente modo³

$$SCR = BSCR - ADJ + SCR^{OP}$$

dove *ADJ* indica gli aggiustamenti per l'assorbimento del rischio associato alle riserve tecniche e alle imposte differite, *SCR^{OP}* indica i rischi operativi che possono scaturire da inefficienze nei sistemi o nei processi interni, mentre *BSCR* indica il *Basic Solvency Capital Requirement* e viene calcolato come

$$BSCR = \sqrt{\sum_{r \times c} Corr SCR_{r,c} \cdot SCR_r \cdot SCR_c} + SCR_{intangibles}$$

dove i requisiti di capitale dei moduli di rischio

- *Market* (legato alla volatilità e al livello degli strumenti finanziari),
- *Default* (legato al rischio di insolvenza delle controparti),
- *Life* (legato alla gestione dei contratti sulla vita),
- *Health* (legato alla gestione dei contratti sulla salute),
- *Non-Life* (legato alla gestione dei contratti dei rami danni)

vengono aggregati mediante la matrice di correlazione lineare riportata in tabella 4.1 a cui viene aggiunto il requisito di capitale associato al rischio *intangibles*, che riguarda gli *asset intangibili*, a loro volta legati ai rischi di mercato (perdita di valore) e ai rischi interni (dipendenti dalle finalità degli asset).

²Per maggiori dettagli sui QIS precedenti un utile riferimento è Savelli e Clemente (2009b).

³EIOPA (2010)

	SCR_{mkt}	SCR_{def}	SCR_{life}	SCR_{health}	SCR_{nl}
SCR_{mkt}	1				
SCR_{def}	0,25	1			
SCR_{life}	0,25	0,25	1		
SCR_{health}	0,25	0,25	0,25	1	
SCR_{nl}	0,25	0,5	0	0	1

Tabella 4.1: Formula standard QIS5: Matrice di correlazione tra i fattori di rischio.

4.2.1 La formula *standard* per il *Non-Life*

Come già nel QIS4⁴ anche nel QIS5⁵ l'EIOPA ha segmentato il business non vita in 12 *Line of Business* (LoB) che corrispondono a singoli rami oppure ad aggregazioni dei tradizionali rami ministeriali. Il rischio di sottoscrizione per le assicurazioni *Non-Life* è il rischio tecnico specifico originato dalle obbligazioni derivanti dai contratti assicurativi non vita, relativamente agli eventi coperti e ai processi utilizzati per la gestione del business.

I sottomoduli di rischio relativi al *Non-Life Underwriting Risk* sono il *Premium-Reserve*, il *CAT* e il *Lapse Risk*.

- Il *Premium Risk* (rischio di tariffazione) nasce al momento della stipula della polizza e prima che qualsiasi sinistro sia accaduto ed è il rischio che il costo dato dalle spese e dall'ammontare delle perdite dovute ai sinistri sia maggiore dei premi incassati; è riconducibile ai concetti di sottotariffazione e sovrasinistralità.
- Il *Reserve Risk* (rischio di riservazione) è il rischio che si verifichi uno scostamento negativo tra la previsione di costo rappresentata dalle riserve tecniche e l'effettivo costo che si realizzerà negli anni successivi con riferimento alle predette riserve.
- Il *CAT Risk* (rischio catastrofe) è dato da eventi estremi o irregolari che non sono sufficientemente catturati dai rischi di assunzione e di riservazione; si tratta di eventi con bassissima probabilità di verificarsi, ma con grandissimo impatto e di difficile previsione.
- Il *Lapse Risk* (rischio di riscatto, introdotto per la prima volta proprio nel QIS5) è legato alle eventuali opzioni lasciate agli assicurati circa il riscatto dei contratti assicurativi non vita.

⁴EIOPA (2007)

⁵EIOPA (2010)

Per calcolare i rischi tecnici *non-life*, vengono innanzitutto valutati insieme il rischio di tariffazione e il rischio di riservazione per determinare il *Premium-Reserve Risk* che successivamente è aggregato al *CAT Risk* e al *Lapse Risk* secondo la seguente formula

$$SCR_{nl} = \sqrt{\sum_{r \times c} CorrNL_{r,c} \cdot NL_r \cdot NL_c},$$

dove i valori della matrice di correlazione *CorrNL* sono riportati nella tabella 4.2.

	NL_{pr}	NL_{lapse}	NL_{CAT}
NL_{pr}	1		
NL_{lapse}	0	1	
NL_{CAT}	0,25	0	1

Tabella 4.2: Formula standard QIS5: Matrice di correlazione dei sottomoduli di rischio del SCR_{nl} .

Un'ulteriore modifica introdotta nel QIS5 rispetto agli esercizi precedenti riguarda il valore della correlazione tra il *Premium-Reserve* e il *CAT Risk*, che ora non è più nullo, ma pari a 0,25.

Il requisito di capitale per il rischio di tariffazione e riservazione è calcolato attraverso la formula

$$NL_{pr} = \rho(\sigma) \cdot V$$

dove:

- V è il volume complessivo relativo ai premi e alla riserva sinistri⁶;
- σ è la deviazione standard, che si ottiene:
 - dall'aggregazione a livello di singola LoB delle deviazioni standard relative al rischio di tariffazione e al rischio di riservazione⁷;
 - dalla successiva riaggregazione delle deviazioni standard per LoB attraverso una matrice di correlazione lineare fissata dall'EIOPA:

$$\sigma = \sqrt{\frac{1}{V^2} \sum_{r,c} CorrLoB_{r,c} \cdot \sigma_r \cdot V_r \cdot \sigma_c \cdot V_c}$$

⁶Il volume complessivo comprende anche un eventuale beneficio di diversificazione geografica dei contratti assicurativi propri dell'impresa.

⁷L'aggregazione avviene ipotizzando una correlazione di 0,5 tra tariffazione e riservazione.

- $\rho(\sigma)$ è una funzione della deviazione standard aggregata.

La logica sottostante alla formula di $\rho(\sigma)$ è semplicemente quella di far dipendere il requisito di capitale da una misura di dispersione, cioè di variabilità nel tempo, dell'ammontare dei premi e delle riserve. In particolare è strutturata in modo che, ipotizzando che la distribuzione del rischio sottostante sia una lognormale, produca un fabbisogno di capitale coerente con una misura di rischio di VaR con un livello di confidenza del 99,5 per cento. La relazione che la descrive è quindi

$$\rho(\sigma) = \frac{\exp\left(N_{0,995} \cdot \sqrt{\log(\sigma^2 + 1)}\right)}{\sqrt{\sigma^2 + 1}} - 1$$

dove $N_{0,995}$ indica il quantile di ordine 99,5 per cento di una distribuzione normale standard. Approssimativamente vale che $\rho(\sigma) \approx 3\sigma$.

Tornando a considerare la volatilità relativa di ogni LoB, essa può essere determinata tramite due modalità definite *market wide approach* e *undertaking specific approach*. Il primo prevede semplicemente di utilizzare dei coefficienti *volatility factor* fissati direttamente dall'autorità di vigilanza⁸, mentre il secondo permette - previa l'autorizzazione dell'autorità di vigilanza - l'utilizzo di parametri *undertaking specific* ottenuti integrando i parametri di mercato ($\sigma_{LoB,M}$) con stime interne della volatilità basate su dati relativi alla propria attività⁹ ($\sigma_{LoB,US}$). Tale ponderazione avviene tramite dei 'coefficienti di credibilità' c_{LoB} diversificati tra *Line of Business* e saranno tanto maggiori quanto più si estendono nel passato le serie storiche da cui si derivano le stime interne di volatilità. In formule

$$\sigma_{LoB} = c_{LoB} \cdot \sigma_{LoB,US} + (1 - c_{LoB}) \cdot \sigma_{LoB,M}.$$

In particolare l'EIOPA ha stabilito che i parametri stimati internamente risultano completamente attendibili, cioè il fattore di credibilità è 100%, quando le serie storiche utilizzate non sono inferiori a 15 anni per i rami RCG, RCA, Credito e Cauzione, mentre per gli altri la soglia scende a 10 anni. Il presupposto che sta alla base di un tale schema di ponderazione è che l'affidabilità della stima cresce al crescere dell'orizzonte storico su cui è calibrata; tuttavia è da sottolineare che tale presupposto risulta valido nel

⁸E' tuttavia prevista la possibilità di correggere a livello di singola LoB la volatilità del *premium risk* con un fattore che tenga conto degli eventuali trattati di riassicurazione non proporzionale.

⁹Per il dettaglio dei metodi ammessi al calcolo degli *Undertaking Specific Parameter* rimandiamo a EIOPA (2010).

caso in cui si ipotizza la volatilità stazionaria nel tempo, e tale ipotesi è tanto più debole quanto aumenta l'arco temporale di valutazione.

Siccome il valore della volatilità è di particolare impatto sul calcolo del requisito di solvibilità, la sua corretta stima è stata, almeno fino ad oggi, oggetto di continue revisioni metodologiche da parte di diversi contributi alla normativa. La linea adottata dall'EIOPA nel fissare i fattori di volatilità è comunque di tipo prudenziale, comportando quindi un maggior assorbimento di capitale per quelle compagnie che adotteranno il metodo *standard*; questo si traduce anche in un incentivo all'impiego di modelli interni e quindi alla stima di requisiti patrimoniali più in linea con il reale profilo di rischio aziendale.

4.3 Un modello interno per il rischio di sottoscrizione Danni

Come abbiamo già avuto modo di discutere l'entità del rischio in Solvency II è espressa da una misura di VaR (*Value at Risk*) ed infatti il rischio di sottoscrizione non rappresenta un'eccezione a questo principio.

E' possibile definirlo come la massima perdita realizzabile entro un dato orizzonte temporale per un determinato intervallo di confidenza, in formule

$$VaR(0, t)_{0,995} = -U_{0,995}(t)$$

dove:

$-U_{0,995}(t)$ è il percentile di ordine 99,5% dell'ammontare della *risk reserve* $U(t)$. Siccome l'arco temporale considerato in Solvency II è un anno, il requisito di capitale dato dal modello interno sarà semplicemente

$$VaR(0, 1)_{0,995} = -U_{0,995}(1)$$

Di seguito presentiamo la formulazione più generale della *risk reserve*¹⁰ della quale tuttavia adotteremo una versione semplificata nel prossimo capitolo, quando affronteremo un caso reale. Nella letteratura della teoria del rischio la riserva di rischio al termine dell'anno t è data da:

$$\tilde{U}_t = (1+j) \cdot \tilde{U}_{t-1} + \left[(B_t - \tilde{X}_t - E_t) - (B_t^{RE} - \tilde{X}_t^{RE} - C_t^{RE}) \right] \cdot (1+j)^{1/2} \quad (4.1)$$

dove:

¹⁰Si veda Savelli (2002).

- B_t = premi di tariffa al lordo della riassicurazione nell'anno t ;
- \tilde{X}_t = importo aggregato del costo dei sinistri nell'anno t ;
- E_t = spese generali sostenute nell'anno t ;
- B_t^{RE} = premi di tariffa ceduti ai riassicuratori nell'anno t ;
- \tilde{X}_t^{RE} = importo dei premi coperti dai riassicuratori nell'anno t ;
- C_t^{RE} = importo delle commissioni di riassicurazione nell'anno t ;
- j = tasso annuo di interessi *risk free*.

Sotto l'ipotesi di un flusso costante di entrate ed uscite nell'arco dell'anno, l'elevamento alla 1/2 equivale ad affermare che esse si concentrino a metà anno.

I premi di tariffa al lordo della riassicurazione B_t sono composti dai premi puri $P_t = E(\tilde{X}_t)$, dal caricamento di sicurezza applicato, ottenuto come percentuale fissa λ dei premi puri, e dal caricamento per le spese, dato da una quota fissa c degli stessi premi di tariffa lordi:

$$B_t = P_t + \lambda \cdot P_t + c_t \cdot B_t.$$

Formulando le ipotesi che tutti i sinistri siano reciprocamente indipendenti, che il loro costo unitario appartenga alla stessa famiglia di distribuzione e che il numero dei sinistri e la variabile del costo del singolo sinistro siano tra loro indipendenti, allora è possibile definire il costo aggregato dei sinistri \tilde{X}_t con la formula

$$\tilde{X}_t = \sum_{i=1}^{\tilde{K}_t} \tilde{Z}_{i,t} \quad (4.2)$$

dove \tilde{K}_t indica la frequenza di accadimento dei sinistri e $\tilde{Z}_{i,t}$ l'importo del singolo sinistro.

Capitolo 5

Applicazione in campo assicurativo

In questo capitolo applicheremo la metodologia della *pair-copula construction* a dati reali provenienti dal mondo assicurativo e cercheremo di fornire degli strumenti per valutare il requisito di capitale in termini di un possibile modello interno alternativo alla formula standard di Solvency II¹.

Nel paragrafo 5.1 presenteremo il lavoro fatto per costruire il dataset di riferimento, fornendone successivamente una descrizione tramite alcune statistiche descrittive; verificheremo infine attraverso un test non parametrico l'ipotesi di indipendenza dei dati.

La determinazione delle distribuzioni marginali del costo aggregato degli eventi di ciascuna classe di rischio è l'oggetto di studio del paragrafo 5.2. Andremo a descrivere tramite opportuni modelli stocastici la distribuzione del numero degli eventi (*frequency*) e il costo del singolo evento (*severity*) e presenteremo un metodo che permette di combinare questi due aspetti.

Descritto il comportamento individuale di ciascuna classe di rischio, ci occuperemo successivamente nel paragrafo 5.3 della stima delle loro dipendenze in termini di correlazione lineare e τ di Kendall.

Successivamente, nel paragrafo 5.4, calcoleremo il requisito di capitale individuale di ciascuna classe di rischio e successivamente confronteremo i requisiti aggregati secondo diverse strutture di dipendenza. Tra queste, di particolare interesse, presenteremo il modello di D-vine combinato con diversi ordini delle variabili e l'utilizzo di varie famiglie di copule.

¹Le analisi presentate in questo capitolo sono ispirate dal lavoro di Moraldi (2011).

5.1 Analisi preliminari dataset

In questa sezione presentiamo il dataset che verrà utilizzato per le successive analisi fornendo le regole con le quali è stato costruito e alcune statistiche descrittive. I dati di partenza sono stati messi a disposizione da Società Cattolica di Assicurazione, una compagnia assicurativa di medie dimensioni che opera sul mercato italiano.

5.1.1 Presentazione dataset

Come abbiamo già avuto modo di discutere nel capitolo precedente, secondo Solvency II i rischi tecnici *non life* contemplano 12 LoB, quindi per poter confrontare in modo omogeneo i risultati della formula standard con quelli prodotti da un potenziale modello interno dovremmo disporre di osservazioni congiunte tra le diverse LoB. Siccome informazioni di questo tipo sono realisticamente molto difficili da recuperare dal mondo assicurativo/riassicurativo, abbiamo concentrato la nostra attenzione su alcune sottoclassi della LoB Incendio e Altri Danni ai Beni.

Al fine di confrontare gli assorbimenti di capitale dovuti a diverse strutture di dipendenza abbiamo identificato all'interno della LoB Incendio e Altri Danni ai Beni cinque diverse classi di rischio (Agricolo, Civile, Commercio, PMI, Religioso) relative ai prodotti che forniscono protezione contro gli eventi atmosferici quali le grandinate, le trombe d'aria e le bufere.

Di seguito forniamo una breve descrizione delle garanzie offerte dalle classi di rischio considerate:

- **Agricolo:** comprende le coperture per tutti i beni mobili e immobili utilizzati da un'azienda agricola, come per esempio serre, fienili, stalle; oltre ai fabbricati copre anche i beni contenuti come trattori e altri macchinari ad uso agricolo. Generalmente viene estesa anche all'abitazione civile associata al contesto agricolo;
- **Civile:** riguardano questa classe di rischio le coperture alle costruzioni ad uso civile quali per esempio le abitazioni private, gli ospedali, le banche, i comuni, i musei, le caserme, gli uffici pubblici e privati, le palestre;
- **Commercio:** con queste coperture si proteggono i beni mobili e immobili relativi ad edifici adibiti ad uso commerciale, oltre alle strutture in sé come i negozi, la copertura è estesa anche ai contenuti;

- **Piccola e Media Impresa (PMI):** i prodotti di questa classe proteggono le attività artigianali e industriali di piccole e medie dimensioni relativamente ai fabbricati e ai beni contenuti, come per esempio le merci;
- **Religioso:** sono coperti da questi prodotti i beni mobili e immobili legati al mondo religioso tra i quali per esempio le chiese, le canoniche, gli oratori, i musei vaticani.

Ripercorriamo ora le operazioni fatte per costruire il dataset su cui sono basate le successive analisi.

Innanzitutto sono stati estratti dai sistemi gestionali della Società i sinistri relativi ai prodotti sopra descritti accaduti in Italia negli anni dal 2005 al 2009 alle varie viste del 31.12 2005-2009 di ciascun anno. Oltre al prodotto assicurativo a cui è stata associata una classe di rischio, sono state ricavate le informazioni sull'importo pagato, sull'importo posto a riserva, sulla data di accadimento del sinistro e sull'ubicazione del rischio (espressa attraverso una suddivisione in quattro macro-zone: Nord-Est, Nord-Ovest, Centro e Sud).

Le viste ai vari anni sono utili perchè permettono di rilevare oltre ai sinistri accaduti e denunciati nell'anno di accadimento, anche i sinistri accaduti, ma non ancora denunciati, i cosiddetti IBNR (*Incurred But Not Reported*).

Considerando che le nostre valutazioni avranno come data di riferimento il 2009 e siccome la durata del ciclo liquidativo dei sinistri di queste classi di rischio si attesta intorno ai tre anni, abbiamo ritenuto maggiormente attendibili, e quindi utilizzabili, i dati relativi ai soli sinistri con anni di accadimenti 2005, 2006 e 2007.

Per determinare il costo di ciascun sinistro, definito come la somma degli importi pagati negli anni più il valore posto a riserva nell'anno di valutazione, abbiamo inflazionato le somme pagate riportandole al 2009 (per fare questo ci siamo serviti degli indici riassunti nella tabella 5.1²) e abbiamo considerato solo l'eventuale riserva ancora presente alla data di valutazione (in questo caso il 31.12.2009).

	2005	2006	2007	2008	2009
Indice con base 2009	107,3	105,8	104,0	100,7	100

Tabella 5.1: Indici utilizzati per riportare gli importi monetari passati al 2009

²I dati sono stati ricavati dal sito dell'ISTAT <http://www.istat.it/prezzi/precon/dati> relativamente all'inflazione passata dell'indice FOI famiglie operai e impiegati - indice generale senza tabacchi - media anno

La successiva operazione compiuta sui dati è stata quella di eliminare i sinistri chiusi senza seguito, determinati in base a delle soglie minime di costo diverse per classe di rischio.

L'oggetto delle nostre analisi sono i sinistri che hanno colpito simultaneamente diverse classi di rischio e che definiamo eventi. Un evento è un fenomeno che per un massimo di 72 ore ha insistito su uno stesso territorio. Incrociando le informazioni della data di accadimento e dell'ubicazione del rischio è stato possibile assegnare a ciascun sinistro un unico evento. Infine quantifichiamo il costo di un evento con la somma dei costi dei sinistri che lo compongono.

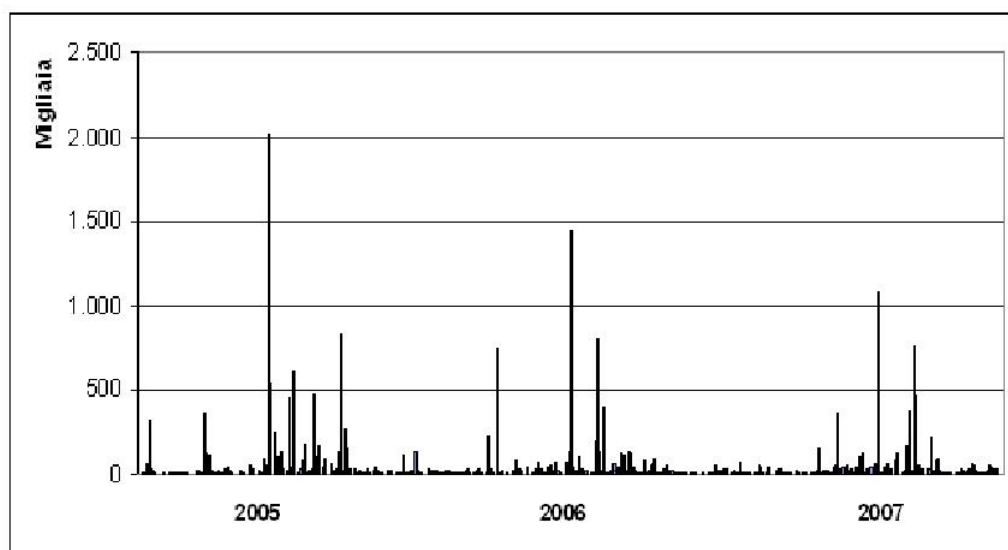


Figura 5.1: Ammontare degli eventi avvenuti in ordine cronologico nel periodo 2005-2007

Nel periodo 2005-2007 sono stati osservati 770 eventi i cui ammontari complessivi (espressi in migliaia di euro) sono rappresentati in ordine cronologico in figura 5.1. Notiamo che solo 9 eventi (poco più dell'1%) superano la soglia di 500 mila euro, il più grave dei quali ha causato da solo circa 2 milioni di danni. Nonostante questi eventi decisamente importanti, il 94% di essi (721 eventi) non supera i 100 mila euro. In termini più analitici, il costo medio di un evento ammonta a 34 mila euro, mentre la deviazione standard è di poco superiore ai 125 mila euro, evidenziando quindi una forte variabilità all'interno del campione osservato.

Rispetto alla rappresentazione degli eventi data dal grafico in figura 5.1, la figura 5.2 fornisce un dettaglio maggiore circa la ripartizione degli stessi eventi sul territorio italiano. Ovviamente questa rappresentazione è fortemente

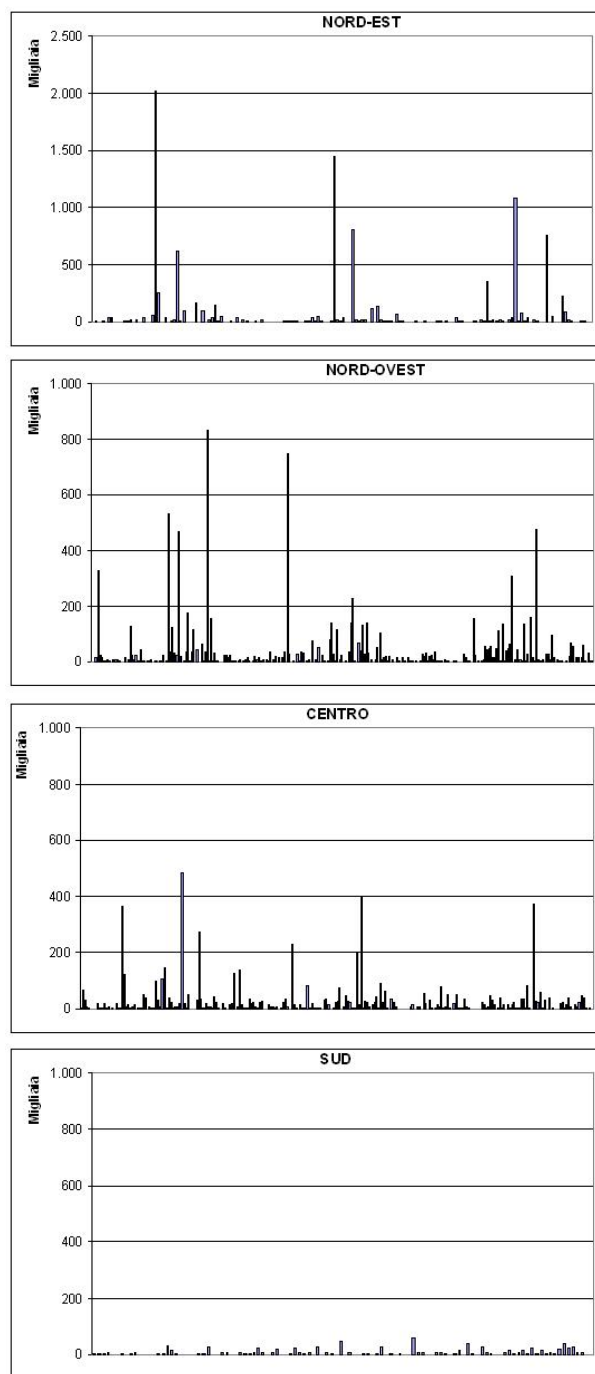


Figura 5.2: Ammontare degli eventi avvenuti nel periodo 2005-2007 divisi per zone geografiche

condizionata dalla distribuzione territoriale del portafoglio della compagnia di assicurazione di riferimento. Fatta questa premessa, notiamo che quasi tutti gli eventi più gravi sono avvenuti nella zona del Nord-Est, mentre al Sud non si registrano episodi di particolare rilievo.

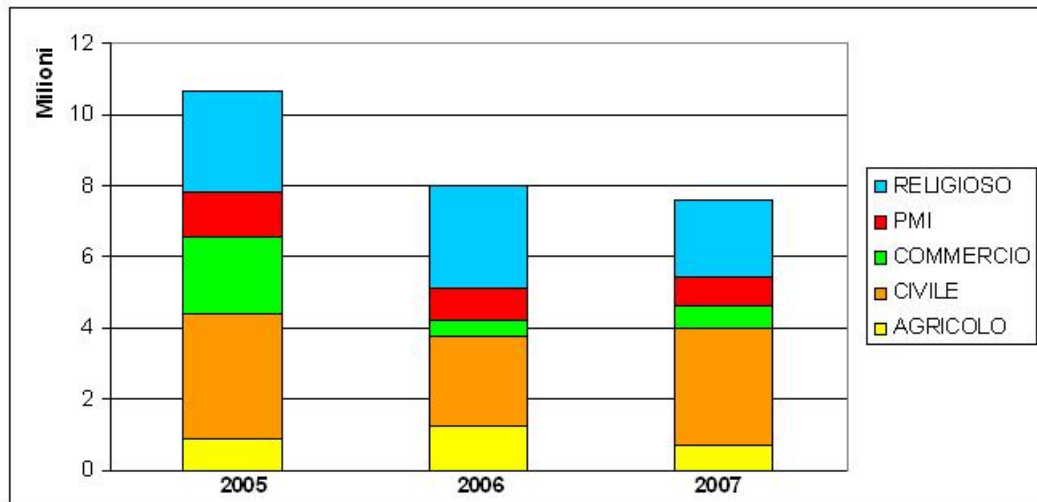


Figura 5.3: Contributo di ciascuna classe di rischio sul costo totale degli eventi atmosferici negli anni 2005-2007

Commentiamo ora l'impatto delle singole classi di rischio (Agricolo, Civile, Commercio, PMI e Religioso) sull'ammontare totale degli eventi, mostrato in figura 5.3. Avendo a disposizione una serie storica di solo 3 anni non è possibile rilevare un effettivo trend di crescita o diminuzione. Le due classi di rischio che risultano predominanti sulle altre sono quelle del Civile e del Religioso, soprattutto perchè nel portafoglio della compagnia sono presenti prodotti di queste classi in numero maggiore rispetto alle altre. Le percentuali di composizione non presentano rilevanti variazioni nel tempo, tranne che per la classe di rischio del Commercio che sembra calare nel tempo, ma questo è spiegato da un grosso evento di punta superiore al milione di euro avvenuto nel 2005.

Concludiamo questa sezione presentando nella tabella 5.2 le principali statistiche di base utili a comprendere la composizione del dataset³.

Notiamo che quasi ogni volta che si verifica un evento atmosferico che genera danni, come una grandinata o una tromba d'aria, vengono colpiti edifici ad uso civile, come è logico attendersi. Come abbiamo già avuto modo di verificare sugli ammontari complessivi, la stragrande maggioranza degli eventi

³Laddove si tratti di importi monetari, essi sono espressi in migliaia di euro.

	Complessivo	AGR	CIV	COM	PMI	REL
n. eventi	770	190	603	175	225	425
n. eventi <100	721	185	588	167	219	408
% eventi <100	94%	97%	98%	95%	97%	96%
Minimo	0,15	0,22	0,15	0,16	0,31	0,31
Massimo	2.019	215	503	1.012	195	712
1° quartile	2,1	1,4	1,2	0,9	2,1	2,2
Mediana	7,1	4,6	3,3	3,2	5,1	6,0
3° quartile	21,4	13,7	8,7	9,3	12,5	14,3
Media	34,1	15,2	15,3	18,8	13,2	18,6
Deviazione std.	125,6	29,5	51,4	81,1	25,0	49,6
Asimmetria	9,45	4,14	7,00	10,76	4,13	8,20
Curtosi	115	20	55	130	20	95

Tabella 5.2: Statistiche descrittive: eventi atmosferici 2005-2007

è inferiore ai 100 mila euro e questo in misura ancora maggiore nelle singole classi di rischio. Nonostante vi siano punte decisamente elevate, come già sottolineato commentando figura 5.1, la media dei singoli eventi non arriva a 35 mila euro. Una caratteristica importante che accomuna tutte le classi di rischio è che la media risulta sempre maggiore, e di parecchio, della mediana e, cosa più sorprendente, anche del terzo quartile. La spiegazione è da ricercarsi negli elevati valori di asimmetria, dovuti alla presenza di valori delle perdite estremi. La distribuzione del costo degli eventi è infatti asimmetrica positiva, il che implica una coda destra, dove si manifestano eventi con bassa frequenza ma ad alto impatto, particolarmente spesso. Questo comportamento lo evidenziamo in particolare per la classe di rischio Commercio, nella quale abbiamo già fatto notare la presenza di un grosso evento estremo. Considerazioni del tutto analoghe si possono fare analizzando i valori della curtosi, in particolare la distribuzione risulta chiaramente leptocurtica, con coda destra molto pesante e la maggior parte delle osservazioni concentrata intorno al valore medio.

5.1.2 Verifica ipotesi di indipendenza

Prima di continuare con le prossime analisi riteniamo utile verificare la validità dell'ipotesi di variabili indipendenti che si è soliti formulare in queste circostanze. Lo scopo di tali valutazioni è di poter lavorare sui dati con una certa libertà, senza dover tener conto di dipendenze tra valori riferiti ad istanti temporali diversi.

Uno strumento presente in letteratura e che utilizziamo per verificare l'ipotesi di indipendenza è il *test delle differenze di segno* che fa parte del gruppo dei cosiddetti test di casualità, per i quali un utile riferimento è Brockwell e Davis (1991).

- *Test delle differenze di segno.* Siano S il numero di i per i quali risulta $y_i > y_{i-1}$ e n la dimensione del campione, allora sotto l'ipotesi nulla di indipendenza si dimostra che

$$\mu_S = E(S) = \frac{n-1}{2} \quad , \quad \sigma_S^2 = VAR(S) = \frac{n+1}{12}$$

e che la statistica test

$$\frac{S - \frac{n-1}{2}}{\sqrt{\frac{n+1}{12}}}$$

si distribuisce come una Normale standard.

Zona	S	n. eventi	Statistica test	Ipotesi accettata
Nord-Est	74	160	-1,50	H ₀
Nord-Ovest	130	262	-0,11	H ₀
Centro	113	238	-1,23	H ₀
Sud	51	110	-1,15	H ₀
Totale	381	770	-0,43	H ₀

Tabella 5.3: Test delle differenze di segno per testare l'indipendenza

Avendo a disposizione anche la collocazione geografica per macro-area degli eventi è stato possibile testare l'ipotesi nulla di indipendenza oltre che sul totale degli eventi, anche per zona territoriale. I risultati del test sono riportati nella tabella 5.3 e ci portano ad accettare l'ipotesi nulla di indipendenza per ciascuna segmentazione territoriale.

5.2 Scelta dei modelli distributivi

In questa sezione ci occuperemo della stima della distribuzione del costo aggregato degli eventi per classe di rischio (secondo la definizione data in (4.2)) e più in particolare, come anticipato nel paragrafo 4.3, andremo a determinare le distribuzioni del numero di eventi atteso in un anno (*frequency*) e la distribuzione del costo del singolo evento (*severity*). Presenteremo infine un metodo che permette di integrare queste due distribuzioni al fine di

ottenere la stima del costo complessivo degli eventi. Riguardo a queste tematiche si possono trovare utili riferimenti di letteratura in Klugman et al. (1998) e Hajek (2011).

5.2.1 Il numero degli eventi: *frequency*

La distribuzione maggiormente impiegata per modellare la frequenza delle perdite è sicuramente la distribuzione di Poisson che appare come la scelta più naturale dato che nasce proprio come limite della distribuzione binomiale nei casi in cui si vuole fare inferenza sulla probabilità del verificarsi nel tempo di eventi indipendenti. Trova infatti largo impiego nel campo industriale in quanto è utile alla produzione di statistiche circa l'analisi del numero di pezzi difettosi in un dato intervallo di tempo.

Malgrado ciò la distribuzione di Poisson non sempre risulta essere la migliore scelta per modellare la numerosità dei sinistri e questo perchè è una distribuzione monoparametrica con media e varianza coincidenti. Una distribuzione che permette invece di trattare la varianza indipendentemente dalla media è quella della binomiale negativa, che per questa sua peculiarità è maggiormente indicata a trattare campioni sovradispersi, dove cioè la varianza eccede la media.

Non avendo a disposizione serie storiche⁴ sufficientemente lunghe per evidenziare un fenomeno di sovradisersione nel numero degli eventi, riteniamo quindi soddisfacente l'utilizzo della distribuzione di Poisson, la cui funzione di densità è data da

$$f(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!},$$

dove x è il valore della variabile aleatoria e λ è il parametro. Come accennato sopra i valori di media e varianza coincidono ed in particolare sono uguali al parametro stesso λ .

Utilizziamo il metodo della massima verosimiglianza per la stima del parametro. Il principio su cui si basa tale metodo è quello di individuare la distribuzione teorica sottostante ad un campione osservato, utilizzando i dati osservati perchè ritenuti i più probabili, in quanto osservati. Questo si traduce nel determinare i parametri della distribuzione in modo che assuma il valore massimo in corrispondenza dei dati osservati.

Una volta calcolata la funzione di log-verosimiglianza per una Poisson

⁴Vogliamo precisare che nonostante ci riferiamo alla serie degli eventi come ad una serie storica, essa non lo è in senso stretto in quanto le osservazioni non sono equispaziate nel tempo.

$$\log L(\lambda) = \sum_{i=1}^n \log \left(\frac{e^{-\lambda} \lambda^{x_i}}{x_i!} \right) = -n\lambda + \left(\sum_{i=1}^n x_i \right) \log(\lambda) - \log(x_i!),$$

ne imponiamo la derivata rispetto al parametro uguale a zero

$$\frac{\partial}{\partial \lambda} \log L(\lambda) = -n + \left(\sum_{i=1}^n x_i \right) \frac{1}{\lambda} = 0,$$

ottenendo in questo modo che lo stimatore di massima verosimiglianza (MLE) è la semplice media dei dati osservati

$$\hat{\lambda}_{MLE} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Stimando il parametro con il metodo dei momenti avremmo ottenuto lo stesso risultato in quanto λ rappresenta anche la media.

Nella tabella 5.4 sono riportati per ciascuna classe di rischio i numeri degli eventi osservati negli anni 2005-2007. Siccome una semplice media dei valori non avrebbe permesso di cogliere l'evoluzione nel tempo del volume di business, abbiamo preferito considerare come parametro delle distribuzioni di Poisson i valori dell'anno 2005, che riteniamo maggiormente rappresentativi.

anno	Agricolo	Civile	Commercio	PMI	Religioso
2005	64	209	70	84	152
2006	63	199	60	67	133
2007	63	195	45	74	140

Tabella 5.4: Numero di eventi osservati per classi di rischio negli anni 2005-2007

5.2.2 Il costo del singolo evento: *severity*

Esistono varie funzioni di ripartizione utili a modellare l'entità economica dei sinistri, ma la selezione della funzione più adatta a tale fine dipende dalle caratteristiche specifiche dei dati osservati; attraverso test statistici dedicati è poi possibile valutare la bontà di adattamento del modello al campione considerato.

Mentre per la stima dei parametri della *frequency* abbiamo ritenuto di considerare un unico anno di osservazione, per la stima della *severity* utilizzeremo l'intera serie storica a nostra disposizione, come anche suggerito in I.A.A. (2010).

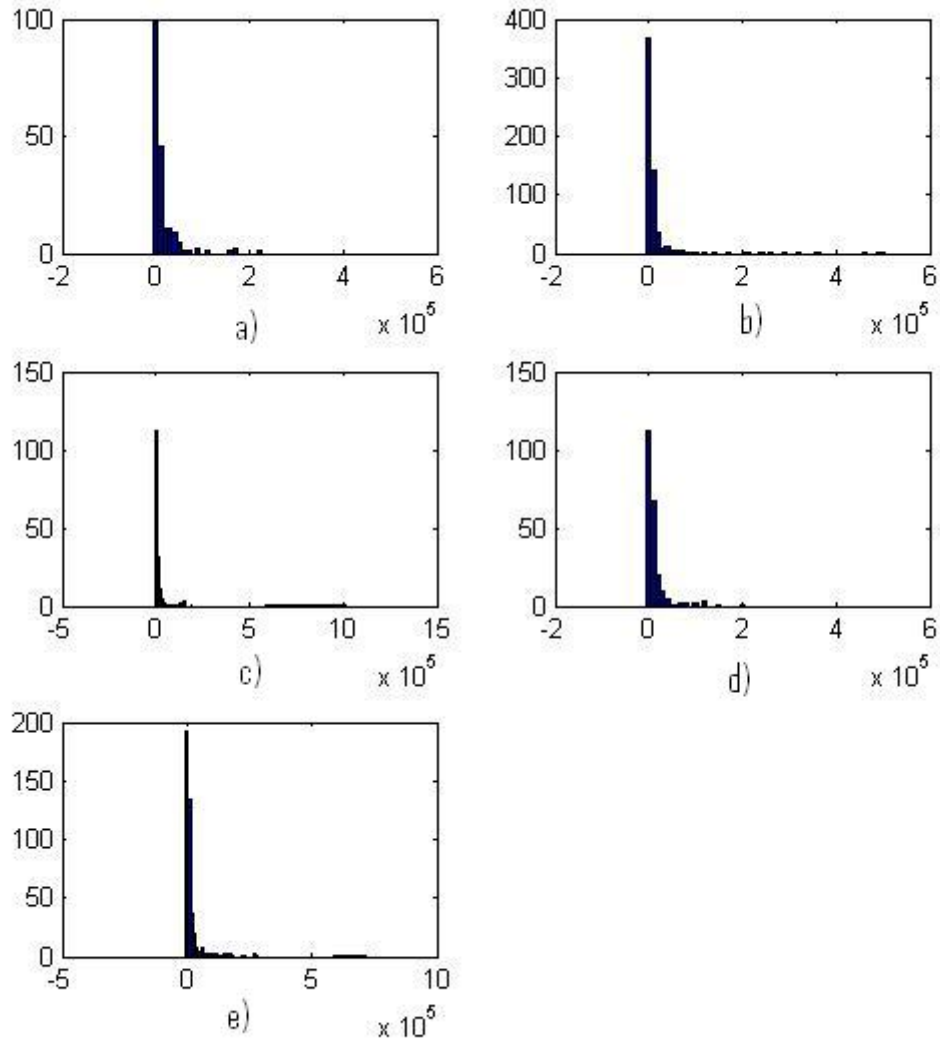


Figura 5.4: Istogrammi delle distribuzioni del costo degli eventi per singola classe di rischio: a) Agricolo, b) Civile, c) Commercio, d) PMI, e) Religioso

Iniziamo questa sezione presentando innanzitutto i dati per i quali cercheremo la distribuzione più appropriata. In figura 5.4 sono riportati gli istogrammi del costo degli eventi delle varie classi di rischio. Possiamo notare come tutte le distribuzioni siano fortemente asimmetriche, con code piuttosto pesanti e manifestano tutte leptocurtosi. Nonostante queste caratteristiche comuni, osserviamo che la coda della classe Civile oltre che essere spessa è

anche omogenea, a differenza delle code del Commercio e del Religioso per le quali si rilevano dei ‘salti’ tra i valori più vicini alla media e quelli più estremi.

Fatte queste premesse, la condizione di ammissibilità della funzione a modellare un campione è la conformità del modello al campione stesso, che nel caso dell’analisi del costo dei sinistri viene espressa come vincolo di positività dell’indice di asimmetria o *skewness*. Tale indice è definito nel modo seguente:

$$\gamma_1 = \frac{m_3}{m_2^{3/2}},$$

dove m_3 e m_2 sono rispettivamente i momenti centrali terzo e secondo della distribuzione.

Sono diverse le distribuzioni che soddisfano tale requisito, di seguito ricordiamo le principali:

- *Distribuzione Lognormale*

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}x} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}},$$

definita per $x > 0$ e avente media e varianza

$$E(X) = e^{\mu + \frac{\sigma^2}{2}}, \quad Var(X) = E(X)^2(e^{\sigma^2} - 1);$$

- *Distribuzione Gamma*

$$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x},$$

definita per $x > 0$, dove $\Gamma(k) = \int_0^\infty t^{k-1} e^{-t} dt$ è la funzione Gamma. Media e varianza sono

$$E(X) = \frac{\alpha}{\beta}, \quad Var(X) = \frac{\alpha}{\beta^2};$$

- *Distribuzione di Weibull*

$$f(x; \lambda, k) = \frac{k}{\lambda^k} x^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k},$$

definita per $x > 0$, $\lambda > 0$ e $k > 0$. Media e varianza sono

$$E(X) = \lambda \Gamma\left(1 + \frac{1}{k}\right), \quad Var(X) = \lambda^2 \Gamma\left(1 + \frac{2}{k}\right) - \lambda^2 \Gamma\left(1 + \frac{1}{k}\right)^2,$$

dove $\Gamma(k)$ è la funzione Gamma definita sopra;

- *Distribuzione Esponenziale*

$$f(x; \lambda) = \lambda e^{-\lambda x},$$

definita per $x > 0$ e $\lambda > 0$, con media e varianza

$$E(X) = \frac{1}{\lambda}, \quad Var(X) = \frac{1}{\lambda^2};$$

- *Distribuzione di Pareto*

$$f(x; \alpha, \phi) = \frac{\alpha \phi^\alpha}{x^{\alpha+1}},$$

dove $0 < \phi \leq x < +\infty$ e $\alpha > 0$. Media e varianza sono

$$E(X) = \frac{\alpha \phi}{\alpha - 1}, \quad Var(X) = \frac{\alpha \phi^2}{(\alpha - 1)(\alpha - 2)}.$$

Una metodologia forse più sofisticata prevederebbe di descrivere il costo del singolo evento attraverso una mistura di distribuzioni, una per la stima degli eventi *attritional* e una per gli eventi *large* (o di punta), dove per *attritional* si intendono quegli eventi al di sotto di una soglia predefinita, mentre quelli che stanno sopra a tale soglia vengono definiti *large*. In particolare è la teoria dei valori estremi, nota anche con l'abbreviazione EVT, che in letteratura si occupa dello studio degli eventi di punta. Non abbiamo ritenuto di applicare modelli di tipo EVT nel presente lavoro soprattutto per il tipo di dataset a nostra disposizione, che non presenta infatti una quantità sufficiente di osservazioni per permettere delle analisi consistenti sui sinistri di tipo *large*.

L'unico accorgimento che abbiamo apportato al dataset al fine di una miglior stima della *severity* riguarda la classe Commercio, dalla quale è stato escluso l'evento più grave perchè ritenuto un outlier; senza di esso media, deviazione standard e asimmetria si riducono in modo considerevole, in particolare l'asimmetria passa da un valore di 10,7 a 3,8.

Occupiamoci ora di determinare quale delle distribuzioni sopra descritte meglio si adatta ai dati di ciascuna classe di rischio. A tale scopo abbiamo utilizzato il tool statistico *BestFit*, una componente del prodotto *@Risk*, che rappresenta un utile strumento nelle analisi di rischio. Il metodo impiegato da *BestFit* per stimare i parametri delle diverse distribuzioni è quello dei minimi quadrati, una tecnica di ottimizzazione che permette di individuare la funzione che più si avvicina ai dati del campione. In particolare la funzione che si ottiene è quella che minimizza la somma dei quadrati dei residui,

ovvero le distanze tra i dati osservati e quelli della curva che rappresenta la funzione stessa.

Per verificare l'adeguatezza dei modelli stimati col metodo dei minimi quadrati, *BestFit* permette di calcolare i valori di alcuni test di *goodness-of-fit*, che sono i test non parametrici del χ^2 e di Kolmogorov-Smirnov. In termini di bontà di adattamento, lo strumento abbina a ciascuno dei modelli stimati un ordine di preferenza basato sui risultati dei diversi test, tali valori sono riportati nella tabella 5.5.

	Agricolo χ^2 ; K-S	Civile χ^2 ; K-S	Commercio χ^2 ; K-S	PMI χ^2 ; K-S	Religioso χ^2 ; K-S
Lognormale	I ; I	I ; I	I ; I	I ; I	I ; I
Gamma	V ; V	V ; V	V ; V	V ; IV	V ; V
Weibull	IV ; IV	IV ; IV	IV ; IV	IV ; V	IV ; IV
Esponenziale	II ; II	III ; III	III ; III	II ; II	II ; III
Pareto	III ; III	II ; II	II ; II	III ; III	III ; II

Tabella 5.5: Distribuzione del costo del singolo evento per classe di rischio: ordine di preferenza dei diversi modelli relativamente alla bontà di adattamento espressa dai test del χ^2 e di Kolmogorov-Smirnov.

In base ai risultati di entrambi i test e per ciascuna classe di rischio la distribuzione che risulta preferita rispetto alle altre è la distribuzione Lognormale, mentre quelle che si adattano meno al dataset sono le distribuzioni Gamma e Weibull.

Classe di rischio	Parametri Lognormale
Agricolo	$\mu = 8,81, \sigma = 1,21$
Civile	$\mu = 8,38, \sigma = 1,58$
Commercio	$\mu = 8,55, \sigma = 1,36$
PMI	$\mu = 8,72, \sigma = 1,23$
Religioso	$\mu = 8,78, \sigma = 1,44$

Tabella 5.6: Costo degli eventi per classi di rischio: parametri della distribuzione Lognormale

Abbiamo osservato che per ciascuna classe di rischio le distribuzioni lognormali stimate con il metodo dei minimi quadrati portavano ad una generale sovrastima dei momenti principali ed in particolare la varianza e l'asimmetria risultavano decisamente maggiori rispetto ai valori empirici. Per tale motivo abbiamo ritenuto di mantenere il modello lognormale, ma di stimarne i parametri con il metodo dei momenti, in questo modo abbiamo potuto

garantire l'uguaglianza tra media e varianza empiriche con i rispettivi valori teorici, e quindi la correttezza del coefficiente di variazione. I valori di asimmetria risultano ancora sovrastimati, anche se in modo meno marcato del precedente. I parametri stimati sono infine riassunti nella tabella 5.6.

5.2.3 Convoluzione di Montecarlo

Individuate le distribuzione del numero di eventi attesi in un anno e del costo del singolo evento, è possibile ricomporre le serie di *frequency* e di *severity* tramite la convoluzione di Montecarlo, che permette di simulare un elevato numero di possibili esercizi annuali attraverso i seguenti punti:

1. Fissare il numero n di scenari/annualità che si desidera simulare.
2. Estrarre casualmente un valore dalla distribuzione della frequenza, questo rappresenterà il numero di eventi nell'ipotetica annualità.
3. Estrarre casualmente dalla distribuzione del costo del singolo evento un numero di importi pari al valore considerato al punto precedente e sommare tutte le realizzazioni.
4. Reiterare i punti 2 e 3 per n volte.

5.3 Stima degli indicatori di dipendenza

In questa sezione andremo ad investigare le eventuali dipendenze presenti tra le classi di rischio e per fare questo utilizzeremo alcuni indicatori di dipendenza già introdotti nel capitolo 2. In particolare considereremo il coefficiente di correlazione lineare ρ e l'indicatore di correlazione di rango τ di Kendall.

Per avere una prima indicazione sul tipo di dipendenza reciproca esistente tra le varie classi di rischio, riteniamo sia utile dare uno sguardo alla figura 5.5 in cui abbiamo riportato i grafici a dispersione delle varie classi di rischio, dall'analisi dei quali notiamo una certa correlazione positiva tra le diverse variabili. In questi grafici rileviamo inoltre la presenza di un evento particolarmente grave nella classe Commercio che tende leggermente ad appiattirne la forma.

Abbiamo discusso degli aspetti teorici, dei pregi e delle eventuali limitazioni del coefficiente di correlazione lineare e del τ di Kendall nel paragrafo 2.2, vediamo ora i valori che questi hanno assunto nel caso del dataset degli eventi di cui disponiamo.

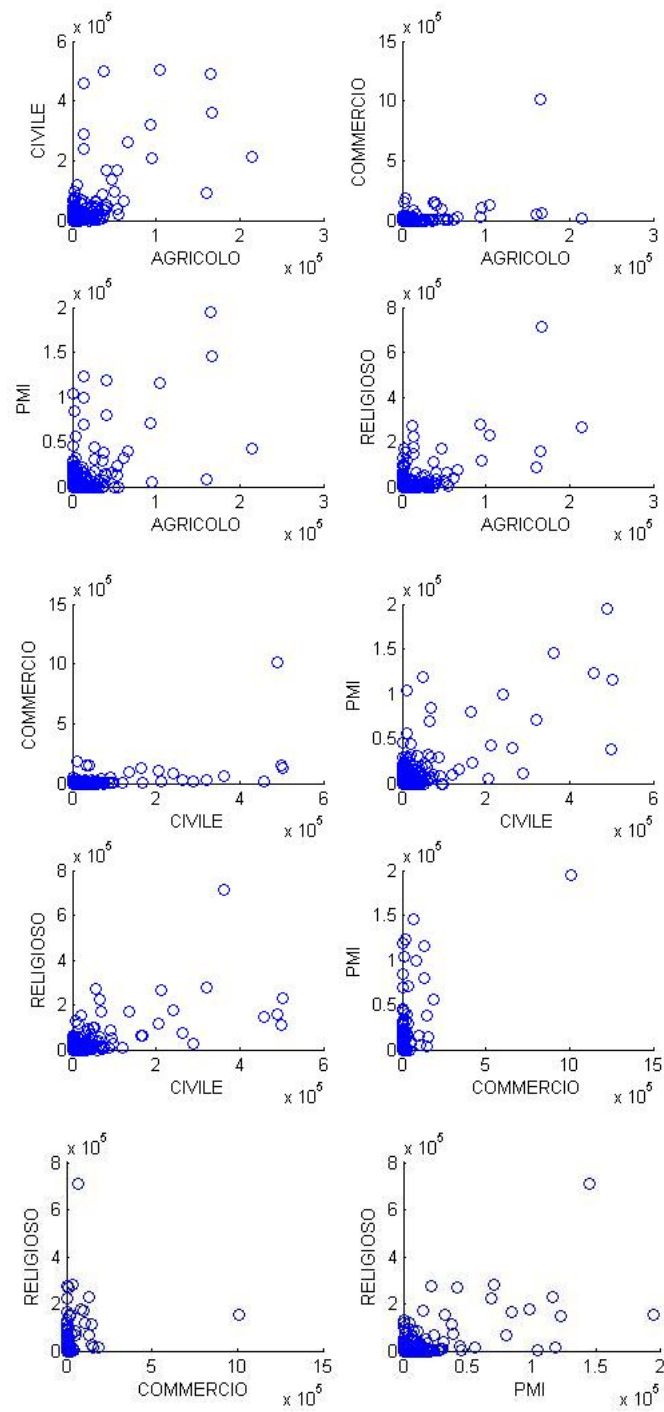


Figura 5.5: Eventi 2005-2007: grafici di dispersione delle classi di rischio

Nella tabella 5.7 sono riportate le stime di ρ e di τ per ciascuna coppia di classi di rischio che abbiamo ottenuto utilizzando, rispettivamente, le formule (2.7) e (2.9).

Classi di rischio	ρ	τ
AGR-CIV	0,6777	0,2222
AGR-COM	0,4849	0,1177
AGR-PMI	0,5732	0,1276
AGR-REL	0,6456	0,1307
CIV-COM	0,5494	0,1657
CIV-PMI	0,7305	0,1239
CIV-REL	0,6615	0,1818
COM-PMI	0,5868	0,1325
COM-REL	0,2753	0,1155
PMI-REL	0,6274	0,0972

Tabella 5.7: Indice di correlazione lineare e τ di Kendall per classi di rischio

Osservando i valori del coefficiente di correlazione lineare notiamo l'effettiva presenza di una correlazione positiva tra tutte le coppie di classi di rischio. In particolare ρ non risulta mai inferiore a 0,4 (fatta eccezione per la coppia Commercio-Religioso, che manifesta la minor dipendenza) raggiungendo il suo valore massimo per la coppia Civile-PMI.

Prendiamo infine in considerazione i valori assunti dal τ di Kendall che, come il coefficiente di correlazione lineare, assume sempre valori positivi. Tuttavia, diversamente da ρ , il range di valori assunti da τ è piuttosto ristretto (tra 0,09 e 0,22) denotando quindi tra le diverse classi di rischio una minore variabilità del grado di dipendenza espresso da questo indicatore.

5.4 Aggregazione delle cinque classi di rischio

Nelle sezioni precedenti abbiamo individuato dei modelli probabilistici adatti a descrivere il costo aggregato per ognuna delle cinque classi di rischio considerate; in questa sezione andremo inizialmente a calcolare i requisiti di capitale per ciascuna di queste classi individualmente e successivamente le aggregeremo secondo diverse strutture di dipendenza al fine di evidenziare come e quanto vari il requisito di capitale necessario a svolgere l'attività assicurativa.

Abbiamo descritto nella sezione 4.3 la struttura completa di un possibile modello interno in termini di riserva di rischio, ma per le nostre analisi ne utilizzeremo una forma semplificata che presentiamo di seguito.

Indichiamo con \tilde{X} il costo aggregato di ciascuna classe di rischio che definiamo come

$$\tilde{X} = \sum_{i=1}^{\tilde{K}} \tilde{Z}_i,$$

dove \tilde{K} indica la frequenza di accadimento degli eventi e \tilde{Z} l'importo del singolo evento. Abbiamo quindi effettuato 100.000 simulazioni da una distribuzione di Poisson, in base al valore estratto abbiamo simulato da delle distribuzioni Lognormali e infine sommato i valori⁵.

Il passo successivo è stato quello di calcolare per le cinque classi di rischio la distribuzione della riserva di rischio su un orizzonte annuale utilizzando la relazione

$$\tilde{U} = B - \tilde{X} - E, \quad (5.1)$$

dove ricordiamo che B sono i premi di tariffa ed E le spese sostenute. Abbiamo ricavato la (5.1) dalla formula più generale (4.1) per un orizzonte temporale annuo e assumendo inoltre di non considerare il contributo della riassicurazione e di non disporre di un capitale iniziale.

Definiti i premi di tariffa B come

$$B = P + \lambda \cdot P + c \cdot B,$$

dove P sono i premi puri, mentre λ e c sono rispettivamente il caricamento di sicurezza applicato e il caricamento per le spese; facciamo le ulteriori ipotesi che non vi sia margine di guadagno sulle spese e che i premi puri coincidano col valore atteso della variabile di costo aggregato. In formule

$$c \cdot B = E, \quad P = E(\tilde{X}) = n \cdot m,$$

dove la seconda relazione sfrutta l'indipendenza tra il numero ed il costo degli eventi, infatti n e m sono rispettivamente il valore atteso del numero di eventi e il costo medio di un singolo evento.

Sotto queste condizioni è possibile calcolare per ciascuna classe di rischio la riserva di rischio definita in (5.1) che ora assume la seguente forma

$$\tilde{U} = (1 + \lambda) \cdot E(\tilde{X}) - \tilde{X}.$$

⁵I parametri delle Poisson e delle Lognormali variano a seconda della classe di rischio considerata e sono i valori determinati nelle sezioni precedenti.

Il caricamento di sicurezza λ può essere inteso come il complemento a 1 del *Combined Ratio*⁶, dato che risulta il margine di guadagno che rimane alla compagnia. Per la sua quantificazione confrontiamo i *Combined Ratio* del ramo Incendio di Cattolica Assicurazioni nel periodo 2005-2007 riportati in figura 5.6⁷ e osserviamo che il valore medio di λ è 25,3% rispetto ai premi di tariffa B, ma siccome a noi serve un coefficiente da applicare ai premi puri, riteniamo ragionevole utilizzare il valore 18%.

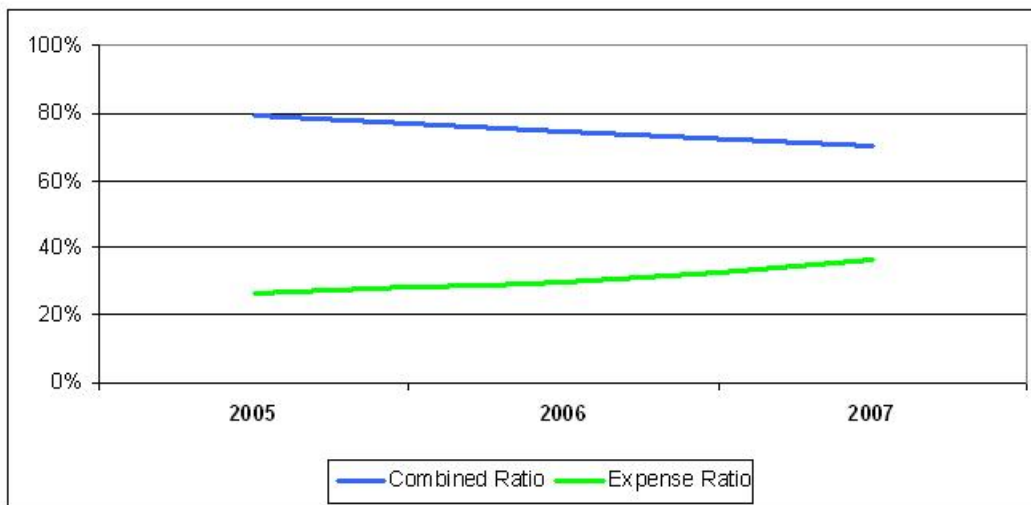


Figura 5.6: Società Cattolica di Assicurazione, Ramo Incendio 2005-2007: *Combined Ratio* ed *Expense Ratio*

Ricaviamo inoltre la stima di c , il caricamento per le spese, osservando l'andamento dell'*Expense Ratio*⁸ riportato sempre in figura 5.6; il valore medio che utilizzeremo è 30,9%. Abbiamo assunto costanti i valori di λ e di c tra le varie classi di rischio, non avendo motivi per ipotizzare una diversificazione.

Quantificata la riserva di rischio è possibile stimare il capitale di rischio, inteso come la massima perdita realizzabile in un anno per un livello di

⁶Il *Combined Ratio* è uno dei principali indicatori utilizzati nel campo assicurativo, esso infatti valuta la performance di una Compagnia di assicurazione. Un valore inferiore a 100% significa che la Compagnia sta traendo profitto dalla sua gestione tecnica, al contrario un valore superiore a 100% rivela che la Compagnia sta pagando più sinistri rispetto ai premi raccolti.

⁷I dati sono stati ricavati dal sito dell'ANIA attraverso il software *Infobila*.

⁸L'*Expense Ratio* è l'indicatore tecnico che valuta le spese e i costi fissi rapportati ai premi di competenza.

confidenza pari a 99,5%. E' infatti una misura di tipo *VaR* che possiamo definire come

$$U_r = \tilde{X}_{0,995} - (1 + \lambda) \cdot E(\tilde{X}).$$

Per confrontare infine i requisiti di capitale tra le varie classi di rischio o tra le diverse ipotesi di dipendenza è utile definire il requisito di capitale relativo (*Risk Based Capital* o *RBC ratio*) introdotto in Daykin et al. (1984). Esso è definito come il rapporto tra il capitale di rischio e i premi di tariffa B , che in formule diventa

$$RBC = \frac{\tilde{X}_{0,995} - (1 + \lambda) \cdot n \cdot m}{B}.$$

Nella tabella 5.8 riportiamo per ciascuna classe di rischio i requisiti di capitale appena definiti e i principali momenti del costo aggregato degli eventi, i cui importi monetari sono espressi in milioni di euro. Come completamento delle analisi forniamo anche una rappresentazione grafica dei valori di RBC in figura 5.7.

	Agricolo	Civile	Commercio	PMI	Religioso
n	64	209	70	84	152
m	0,014	0,015	0,013	0,013	0,018
<i>PremiPuri</i>	0,90	3,21	0,91	1,11	2,82
$E(\tilde{X})$	0,90	3,21	0,91	1,11	2,82
$\sigma(\tilde{X})$	0,23	0,79	0,27	0,26	0,65
<i>CoV</i>	26,2%	24,7%	30,1%	23,5%	23,1%
$\gamma(\tilde{X})$	1,49	5,12	1,76	1,26	1,61
$\frac{VaR_{0,995} - E(\tilde{X})}{\sigma(\tilde{X})}$	3,53	3,87	3,84	3,52	3,82
RBC	43,7%	45,4%	57,3%	37,8%	41,0%

Tabella 5.8: Momenti del costo aggregato e requisiti di capitale relativi per classi di rischio

Innanzitutto notiamo che per ciascuna classe di rischio i premi puri, ottenuti come prodotto del numero di eventi e del costo medio del singolo evento, corrispondono esattamente al valore atteso del costo aggregato, denotando quindi un buon adattamento del modello ai dati.

La classe di rischio Commercio presenta il maggiore requisito di capitale relativo, seguita dalle classi Civile, Agricolo e Religioso, mentre il minor requisito lo osserviamo per la classe delle Piccole-Medie Imprese. Un risultato

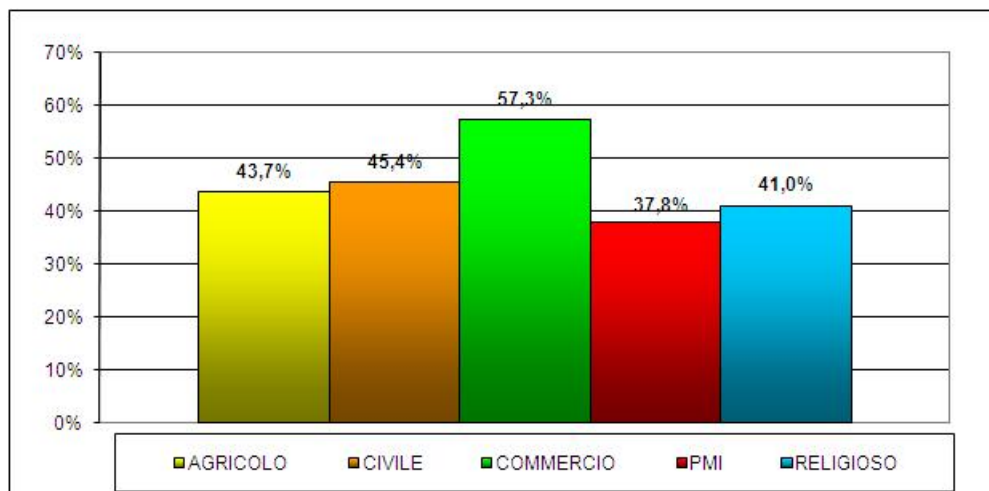


Figura 5.7: *Risk Based Capital* per classe di rischio

di questo tipo è sicuramente legato all'indice di asimmetria che, fatta eccezione per il valore eccezionalmente alto della classe Civile, risulta molto elevato per il Commercio e questo fatto si riflette anche in un alto valore del coefficiente di variazione del costo aggregato; tale indicatore è infatti minore per tutte le altre classi, anche se presenta sempre valori molto elevati e mai inferiori a 23%. Possiamo inoltre notare come, una volta presa la media, il moltiplicatore da applicare alla deviazione standard per ottenere il $VaR_{0,995}$ risulta anch'esso influenzato dagli alti valori di asimmetria; i suoi valori sono infatti sempre superiori a 3,5 in ogni classe di rischio, quando un valore di 2,58 indicherebbe una distribuzione simmetrica. Per ulteriori confronti di questi valori con il mercato italiano, si può far riferimento a Savelli e Clemente (2009c) in cui vengono utilizzati diversi modelli simulativi diversificati per ramo di bilancio.

5.4.1 Indipendenza

Iniziamo in questa sezione a introdurre gli aspetti teorici per il calcolo del requisito di capitale aggregato secondo diverse ipotesi di dipendenza. La prima e più semplice tipologia di aggregazione è rappresentata dall'ipotesi di indipendenza che per le analisi fatte finora riconosciamo non realistica, infatti le classi di rischio hanno tutte manifestato dipendenza positiva tra di loro.

Il requisito di capitale nell'ipotesi di indipendenza è

$$RBC_{IND} = \frac{\tilde{X}_{IND;0,995} - (1 + \lambda)E(\tilde{X}_{IND})}{B},$$

dove i premi di tariffa B e la distribuzione congiunta del costo aggregato sono dati da

$$B = B_{AGR} + B_{CIV} + B_{COM} + B_{PMI} + B_{REL},$$

$$\tilde{X}_{IND} = \tilde{X}_{AGR} + \tilde{X}_{CIV} + \tilde{X}_{COM} + \tilde{X}_{PMI} + \tilde{X}_{REL}.$$

Quindi B è semplicemente la somma dei premi di tariffa delle singole classi precedentemente calcolati, mentre la media del costo aggregato sotto l'ipotesi di indipendenza è la somma delle medie dei costi aggregati costruiti sulle varie classi.

5.4.2 Correlazione lineare

La seconda ipotesi che andiamo a considerare è quella sottostante la formula standard di Solvency II e cioè l'aggregazione tramite matrice di correlazione lineare.

Troviamo in Savelli e Clemente (2009a) la formulazione del *Risk Based Capital* relativa a questa tipologia di aggregazione, che è

$$RBC_{CORR} = \sqrt{\sum_{i=1}^5 \sum_{j=1}^5 Corr_{i,j} \cdot CC_i \cdot CC_j} - \lambda \sum_{i=1}^5 P_i,$$

dove con CC_i si intende il *Capital Charge* di ciascuna classe di rischio definito nel seguente modo

$$CC_i = \tilde{X}_{0,995} - E(\tilde{X}_i).$$

Infine come matrice di correlazione utilizziamo quella stimata dai dati reali, cioè i valori della colonna ρ nella tabella 5.7.

5.4.3 Struttura D-vine

Con l'obiettivo di valutare in modo alternativo la dipendenze tra la variabili, in questa sezione testiamo la struttura di D-vine per diversi ordini delle classi di rischio, a cui assoceremo la *pair-copula construction* valutata con differenti famiglie di copule.

Abbiamo già avuto modo di discutere nel capitolo 3 il processo di stima di un D-vine dato un dataset reale, ma ricordiamo qui brevemente quali sono i punti principali di tale operazione: selezione dell'ordinamento delle variabili nell'albero di primo livello, scelta delle copule bivariate e infine stima dei corrispondenti parametri.

Come criterio di discernimento tra i possibili ordinamenti⁹ abbiamo deciso di utilizzare quello che prevede di assegnare a ciascuna combinazione un valore sintetico del livello di dipendenza tra le variabili rispetto a tale ordinamento. In particolare a ciascuna struttura di D-vine è stata associata la somma dei τ di Kendall relativi ai diversi ordini delle classi; con riferimento ai valori di τ riportati nella tabella 5.7 abbiamo individuato e rappresentato in figura 5.8 gli ordinamenti che massimizzano e minimizzano¹⁰ tale somma.

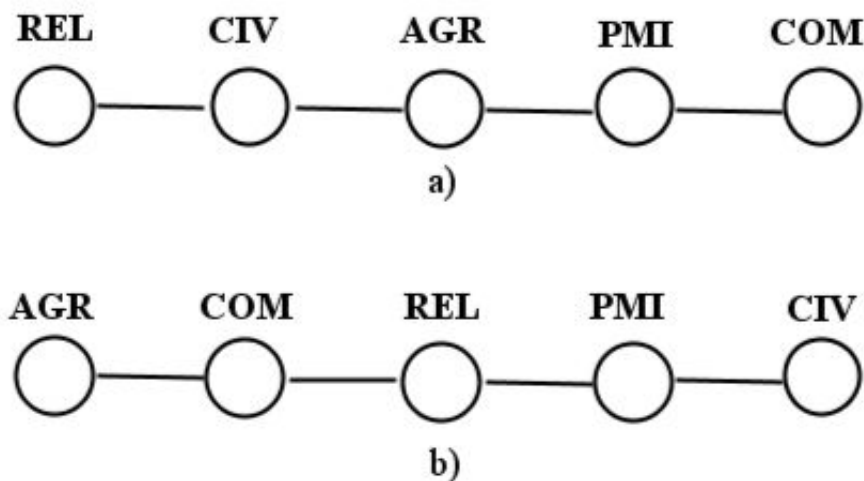


Figura 5.8: a) Ordinamento delle classi di rischio che massimizza la somma dei tau di Kendall; b) Ordinamento delle classi di rischio che minimizza la somma dei tau di Kendall.

Come conseguenza dei valori di τ di Kendall piuttosto simili tra le diverse coppie di variabili, anche i valori associati agli ordinamenti A e B non sono molto distanti, infatti risultano rispettivamente 0,6641 e 0,4543.

Avendo deciso di confrontare gli ordinamenti che massimizzano e minimizzano le dipendenze (identificati dal criterio sopra descritto), il passo successivo consiste nello scegliere la famiglia di *pair-copule* da implementare. Dato

⁹Essendo 5 le classi di rischio, le possibili rappresentazioni di diversi D-vine tra cui scegliere sono $5!/2 = 60$.

¹⁰Per comodità ci riferiremo agli ordinamenti che massimizzano e minimizzano la somma dei τ di Kendall rispettivamente con l'ordine A e B.

che le due principali classi di copule sono quelle ellittiche e quelle Archimedee, abbiamo ritenuto interessante un confronto tra un esponente di ciascuna delle due famiglie; in particolare abbiamo voluto mettere a confronto una copula simmetrica e una che presenta dipendenza sulla coda destra (dove risiedono gli eventi più gravi) scegliendo quindi la copula Gaussiana e la copula di sopravvivenza di Clayton¹¹.

L'ultimo step per l'implementazione dei D-vine è rappresentato dalla stima dei parametri delle *pair-copule*. Abbiamo già avuto modo di osservare nel capitolo 2 che sia la copula Gaussiana sia quella di sopravvivenza di Clayton sono copule monoparametriche; in particolare il parametro della prima è rappresentato dal coefficiente di correlazione lineare ρ , mentre il parametro della seconda, θ , lo otteniamo attraverso la seguente trasformazione del τ di Kendall

$$\theta = \frac{2\tau}{1 - \tau},$$

per il quale è richiesto che sia $\tau > 0$. I nostri dati rispettano tale vincolo, infatti i valori riportati in tabella 5.7 sono tutti positivi.

Per ciascuna configurazione di D-vine abbiamo stimato i parametri delle copule del primo livello attraverso i dati osservati, mentre per i parametri dei livelli successivi abbiamo adottato il metodo della stima sequenziale descritto nel paragrafo 3.4.2 e che prevede il calcolo di ρ e di θ sulle variabili trasformate (definite attraverso l'utilizzo delle funzioni h). Questi valori sono stati infine utilizzati come parametri iniziali per massimizzare la funzione di log-verosimiglianza¹².

Nelle tabelle 5.9 e 5.10 riportiamo rispettivamente i risultati delle stime dei parametri per le *pair-copule* Gaussiane e di sopravvivenza di Clayton. Dando uno sguardo ai valori dei coefficienti di correlazione lineare nella tabella 5.9 notiamo che sia per la configurazione A che B i parametri relativi alle *pair-copule* di livello superiore al primo sono mediamente più bassi dei parametri di primo livello; un comportamento di questo tipo non si osserva invece per i parametri delle *pair-copule* di sopravvivenza di Clayton in tabella 5.10.

¹¹Non abbiamo utilizzato la copula di Gumbel, tradizionalmente impiegata in questi studi per la sua dipendenza su coda superiore, perchè per questa famiglia di copule non esiste una forma chiusa della funzione h che è necessaria per l'implementazione della PCC. Dato che la copula Archimedeica di Clayton presenta dipendenza sulla coda sinistra e ammette una formulazione in forma chiusa della funzione h , abbiamo ritenuto una valida alternativa per i nostri scopi implementare la sua copula di sopravvivenza. Il calcolo di h è riportato in Appendice A.2.

¹²Quest'operazione è stata fatta attraverso l'utilizzo della funzione `fminunc` di *Matlab*

	A		B
$\rho_{REL,CIV}$	0,661	$\rho_{AGR,COM}$	0,485
$\rho_{CIV,AGR}$	0,678	$\rho_{COM,REL}$	0,275
$\rho_{AGR,PMI}$	0,573	$\rho_{REL,PMI}$	0,627
$\rho_{PMI,COM}$	0,587	$\rho_{PMI,CIV}$	0,730
$\rho_{REL,AGR CIV}$	0,339	$\rho_{AGR,REL COM}$	0,222
$\rho_{CIV,PMI AGR}$	0,134	$\rho_{COM,PMI REL}$	0,344
$\rho_{AGR,COM PMI}$	0,319	$\rho_{REL,CIV PMI}$	0,241
$\rho_{REL,PMI CIV,AGR}$	0,174	$\rho_{AGR,PMI COM,REL}$	0,243
$\rho_{CIV,COM AGR,PMI}$	0,216	$\rho_{COM,CIV REL,PMI}$	0,263
$\rho_{REL,COM CIV,AGR,PMI}$	0,162	$\rho_{AGR,CIV COM,REL,PMI}$	0,333

Tabella 5.9: Parametri delle *pair-copule* Gaussianne relativi alle strutture di D-vine A e B.

	A		B
$\theta_{REL,CIV}$	0,444	$\theta_{AGR,COM}$	0,267
$\theta_{CIV,AGR}$	0,571	$\theta_{COM,REL}$	0,261
$\theta_{AGR,PMI}$	0,293	$\theta_{REL,PMI}$	0,215
$\theta_{PMI,COM}$	0,305	$\theta_{PMI,CIV}$	0,283
$\theta_{REL,AGR CIV}$	0,481	$\theta_{AGR,REL COM}$	0,346
$\theta_{CIV,PMI AGR}$	0,164	$\theta_{COM,PMI REL}$	1,293
$\theta_{AGR,COM PMI}$	0,523	$\theta_{REL,CIV PMI}$	0,494
$\theta_{REL,PMI CIV,AGR}$	0,521	$\theta_{AGR,PMI COM,REL}$	0,920
$\theta_{CIV,COM AGR,PMI}$	0,351	$\theta_{COM,CIV REL,PMI}$	0,524
$\theta_{REL,COM CIV,AGR,PMI}$	0,528	$\theta_{AGR,CIV COM,REL,PMI}$	0,408

Tabella 5.10: Parametri delle *pair-copule* di sopravvivenza di Clayton relativi alle strutture di D-vine A e B.

Concludiamo questa sezione definendo il requisito di capitale nell'ipotesi della struttura di D-vine. La variabile di costo aggregato \tilde{X}_{DV} è ottenuta come somma delle cinque variabili simulate con la PCC ed il requisito di capitale si calcola in modo sostanzialmente analogo con quanto visto nell'ipotesi di indipendenza e sarà

$$RBC_{DV} = \frac{\tilde{X}_{DV;0,995} - (1 + \lambda)E(\tilde{X}_{DV})}{B}.$$

5.4.4 Confronto tra ipotesi di dipendenza

In questa ultima sezione ci occupiamo infine di commentare i risultati ottenuti in relazione alle diverse ipotesi di dipendenza formulate precedentemente. Per le aggregazioni che richiedevano l'utilizzo della PCC sono state simulate 10.000 osservazione dai rispettivi D-vine.

Iniziamo dunque andando ad esaminare la tabella 5.8 in cui sono stati riportati i principali momenti e indicatori delle varie distribuzioni che abbiamo ottenuto¹³.

	Indip.	PCC A Gauss	PCC B Gauss	PCC A s. Clayton	PCC B s. Clayton
$E(\tilde{X})$	8,95	8,95	8,95	8,95	8,95
$\sigma(\tilde{X})$	1,11	1,76	1,73	1,58	1,52
CoV	12,4%	19,7%	19,3%	17,7%	17,0%
$\gamma(\tilde{X})$	2,23	1,04	1,26	1,71	1,73

Tabella 5.11: Momenti della distribuzione del costo aggregato degli eventi secondo diverse strutture di dipendenza ($A=REL-CIV-AGR-PMI-COM$, $B=AGR-COM-REL-PMI-CIV$)

In relazione ai risultati specifici di ciascuna classe di rischio (per i valori si veda la tabella 5.8), notiamo innanzitutto che formulare una qualsiasi delle ipotesi di dipendenza ha avuto l'effetto di ridurre mediamente i coefficienti di variazione e, seppure in maniera minore, anche gli indicatori di asimmetria (su questo aspetto torneremo fra poco), fatto che si spiega come una diretta manifestazione del beneficio di diversificazione.

Analizzando più nel dettaglio i coefficienti di variazione, notiamo una certa omogeneità di risultati tra alcuni tipi di aggregazioni. Questo particolare comportamento è dovuto al fatto che tutte le ipotesi di aggregazione preservano il valore atteso della distribuzione, mentre per quanto riguarda la deviazione standard si rilevano valori sostanzialmente simili solo tra combinazioni di D-vine che utilizzano la stessa famiglia di copule, in particolare risultano leggermente più contenuti e vicini al valore di indipendenza quelli relativi alle copule di sopravvivenza di Clayton.

Andando ora ad esaminare l'indice di asimmetria, il primo aspetto curioso da notare è che il valore maggiore è raggiunto nell'ipotesi di indipendenza; riteniamo che questo fatto sia dovuto all'alto valore di asimmetria osservato per la classe Civile e che questo influenzi in modo diretto la distribuzione aggregata sotto l'ipotesi di indipendenza. Diversamente da quanto osservato

¹³Qualora si tratti di importi monetari, essi sono espressi in milioni di euro.

per il coefficiente di variazione, i valori di asimmetria tendono ad essere più elevati, ma simili, nei D-vine con copule di sopravvivenza di Clayton rispetto a quelli con copule Gaussiane; a giustificazione di tale risultato ricordiamo che le copule Gaussiane sono simmetriche.

In seguito a queste considerazioni giungiamo infine a presentare i risultati ottenuti in termini di capitale richiesto relativo, riassunti nella figura 5.9.

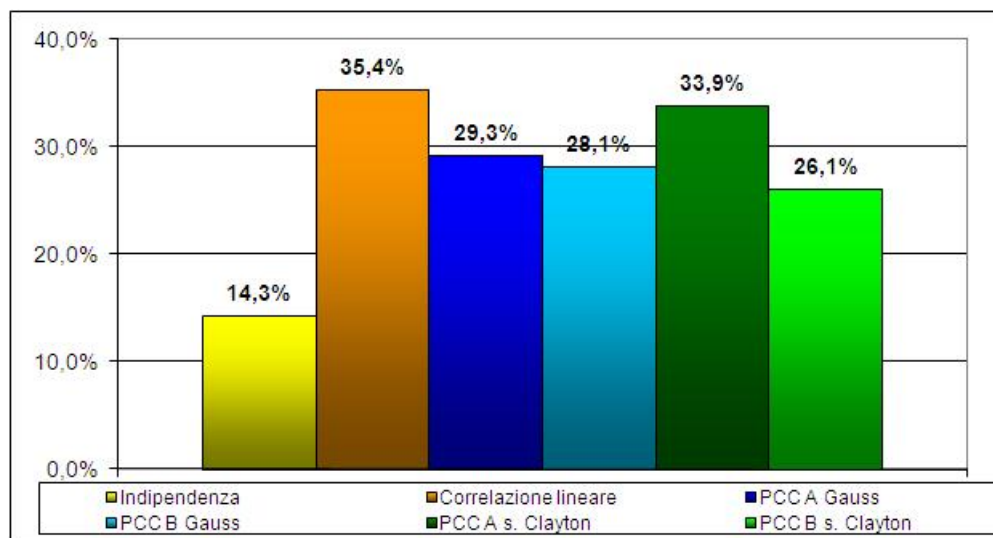


Figura 5.9: *Risk Based Capital*: confronto tra diverse ipotesi di aggregazione

Il tipo di aggregazione che rivela il requisito di capitale minore è ovviamente quella di indipendenza, ma come abbiamo già detto precedentemente quest'ipotesi non è realistica in relazione al nostro dataset, per il quale una dipendenza positiva tra le classi di rischio è piuttosto manifesta.

Al contrario, la correlazione lineare e la PCC abbinata all'ordinamento A con copule di sopravvivenza di Clayton riportano i requisiti più elevati e tutto sommato anche abbastanza simili tra loro.

Dopo queste prime osservazioni, quello che ci sembra interessante far notare è l'effetto che hanno avuto sui RBC le strutture di D-vine con i diversi ordinamenti delle classi di rischio combinati con le due famiglie di copule considerate. Innanzitutto, come era logico aspettarsi, i capitali richiesti relativi alle strutture che massimizzano le informazioni sulla dipendenza negli alberi di primo livello (struttura A) risultano maggiori rispetto ai requisiti legati ai D-vine costruiti con lo stesso tipo di copule, ma con l'ordine che esprime il livello minore di dipendenza (struttura B)¹⁴. Tuttavia l'uso di una famiglia di *pair-copule* piuttosto di un'altra ha un effetto diverso sulla distanza tra

¹⁴A tale proposito abbiamo verificato che nel caso di copule di sopravvivenza di Clayton

il capitale richiesto con la struttura A e quello con la struttura B ; in particolare tale distanza risulta più contenuta nel caso di copule gaussiane e più ampia se si utilizzano copule di sopravvivenza di Clayton. Riteniamo che la giustificazione di un tale comportamento sia una combinazione di più fattori: la specificità del dataset in oggetto e il tipo di dipendenza che esprime la classe di copule considerata.

Oltre al fatto che le dipendenze (esprese in termini di τ di Kendall) assumono valori abbastanza simili per la maggior parte delle coppie di variabili osservate, va ricordato che la copula gaussiana caratterizza una distribuzione a simmetria radiale che non descrive in modo molto accurato la dipendenza sulle code (dove però si calcola il VaR) e questo fa sì che l'ordinamento delle classi di rischio non incida in modo determinante sul capitale richiesto finale. Diversamente dalla copula gaussiana, quella di sopravvivenza di Clayton coglie le dipendenze sulla coda destra ed infatti i RBC calcolati sullo stesso dataset risultano maggiormente influenzati dall'ordine delle variabili nel D-vine.

5.5 Conclusioni

In questo capitolo abbiamo evidenziato le potenzialità della tecnica della *pair-copula construction* nella ricerca di un modello adatto a descrivere una situazione reale in campo assicurativo, un campo diverso da quelli in cui si trova solitamente utilizzata in letteratura (come può essere quello finanziario). In particolare, abbiamo combinato la metodologia della PCC con la ricerca di un modello che possa quantificare i rischi a cui una compagnia di assicurazione è esposta, in termini del suo grado di solvibilità.

Il dataset che abbiamo costruito ci ha permesso di implementare un modello con cinque variabili aleatorie, ma se avessimo avuto a disposizione maggiori informazioni sarebbe stato facilmente adattabile per affrontare problemi in dimensioni maggiori. La PCC può infatti descrivere anche dipendenze multivariate con un alto grado di complessità.

In questo lavoro abbiamo privilegiato il confronto tra diverse strutture di dipendenza, al fine di testare le rispettive ipotesi sottostanti, piuttosto che concentrarci sulla ricerca del miglior modello in relazione al dataset disponibile. Questo potrebbe tuttavia rappresentare un obiettivo futuro per il quale si potrebbe ricercare la famiglia di copule che meglio si adatta a ciascuna

alcuni ordinamenti delle variabili, relativi ad un grado di dipendenza intermedio tra le situazioni A e B come per esempio AGR-PMI-CIV-COM-REL, portano a dei valori di RBC intermedi tra i casi di massima dipendenza e di minima dipendenza.

coppia di variabili. Un'ulteriore sofisticazione del modello potrebbe essere quella di considerare gli effetti di eventuali coperture riassicurative.

Concludiamo dicendo che l'importanza di questo lavoro non è solo legata all'identificazione di un idoneo requisito di capitale che soddisfi la direttiva Solvency II, ma può anche rappresentare per una compagnia di assicurazione un'opportunità di porre maggior attenzione ed eventualmente anche rivedere la propria politica di valutazione e gestione dei rischi.

Appendice A

Funzioni utili per PCC

Presentiamo una breve rassegna delle funzioni che abbiamo utilizzato in questo lavoro per implementare la *pair-copula construction*.

A.1 Copula Gaussiana bivariata

Innanzitutto abbiamo che la funzione di distribuzione della copula Gaussiana è

$$C_{1,2}(u_1, u_2; \rho) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\} dx dy,$$

dove ρ è il parametro della copula. Si ricava la funzione di densità che risulta

$$c_{1,2}(u_1, u_2; \rho) = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{\rho^2(z_1^2 + z_2^2) - 2\rho z_1 z_2}{2(1-\rho^2)}\right\},$$

dove per comodità abbiamo posto $z_1 = \Phi^{-1}(u_1)$ e $z_2 = \Phi^{-1}(u_2)$, con $\Phi^{-1}(\cdot)$ l'inversa della funzione di distribuzione di una Normale standard.

Per questa copula la funzione h^1 è data da

$$h_{1,2}(u_1, u_2; \rho) = \Phi\left(\frac{\Phi^{-1}(u_1) - \rho\Phi^{-1}(u_2)}{\sqrt{1-\rho^2}}\right),$$

da cui ricaviamo che la sua inversa è

$$h_{1,2}^{-1}(u_1, u_2; \rho) = \Phi\left(\Phi^{-1}(u_1)\sqrt{1-\rho^2} + \rho\Phi^{-1}(u_2)\right).$$

¹Per il dettaglio dei calcoli si veda Aas et al. (2009)

A.2 Copula di sopravvivenza di Clayton bivariata

La funzione di distribuzione della copula di Clayton è (Venter, (2001))

$$C(u_1, u_2; \theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}},$$

dove $0 < \theta < \infty$ è il parametro che controlla la dipendenza. L'indipendenza è raggiunta per $\theta \rightarrow 0$, mentre $\theta \rightarrow \infty$ implica perfetta dipendenza. Ricaviamo la distribuzione della copula di sopravvivenza di Clayton, che è

$$\widehat{C}(u_1, u_2; \theta) = u_1 + u_2 - 1 + [(1 - u_1)^{-\theta} + (1 - u_2)^{-\theta} - 1]^{-\frac{1}{\theta}}.$$

Calcoliamo la funzione h

$$\begin{aligned} \widehat{h}_{1,2}(u_1, u_2; \theta) &= \widehat{F}_{1|2}(u_1, u_2; \theta) = \\ &= \frac{\partial}{\partial u_2} \widehat{C}_{1,2}(u_1, u_2; \theta) = \\ &= \frac{\partial}{\partial u_2} \left(u_1 + u_2 - 1 + [(1 - u_1)^{-\theta} + (1 - u_2)^{-\theta} - 1]^{-\frac{1}{\theta}} \right) = \\ &= 1 - (1 - u_2)^{-\theta-1} [(1 - u_1)^{-\theta} + (1 - u_2)^{-\theta} - 1]^{-\frac{1}{\theta}-1}, \end{aligned}$$

e ricaviamo la sua inversa

$$\widehat{h}_{1,2}^{-1}(u_1, u_2; \theta) = 1 - \left([(1 - u_1)(1 - u_2)^{1+\theta}]^{-\frac{\theta}{1+\theta}} + 1 - (1 - u_2)^{-\theta} \right)^{-\frac{1}{\theta}}.$$

Determiniamo infine la funzione di densità della copula di sopravvivenza di Clayton

$$\begin{aligned} \widehat{c}_{1,2}(u_1, u_2; \theta) &= \frac{\partial^2}{\partial u_1 \partial u_2} \widehat{C}_{1,2}(u_1, u_2; \theta) = \\ &= \frac{\partial}{\partial u_1} \left(1 - (1 - u_2)^{-\theta-1} [(1 - u_1)^{-\theta} + (1 - u_2)^{-\theta} - 1]^{-\frac{1}{\theta}-1} \right) = \\ &= (1 + \theta) [(1 - u_1)(1 - u_2)]^{-\theta-1} [(1 - u_1)^{-\theta} + (1 - u_2)^{-\theta} - 1]^{-2-\frac{1}{\theta}}. \end{aligned}$$

Bibliografia

- [1] Aas,K., Czado,C., Frigessi,A., Bakken,H. (2009). Pair-copula construction of multiple dependence *Insurance, Mathematics and Economics* 44(2), 182-198.
- [2] Akaike,H. (1973). Information theory and an extension of the likelihood ratio principle. In B.N. Petrov (Ed.), *Proceedings of the Second International Symposium of the Information Theory*, 257-281. Budapest: Akademiai Kiado.
- [3] Beaudoin,D., Lakhel-Chaieb,L. (2008). Archimedian copula model selection under dependent Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence* 32, 245-268.
- [4] Bedford,T., Cooke,R.M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence* 32, 245-268.
- [5] Bedford,T., Cooke,R.M. (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics* 30(4), 1031-1068.
- [6] Brockwell,P.J, Davis,R.A. (1991). *Times Series: Theory and Method*, Sprmnger-Verlag, New York.
- [7] Cherubini,U., Luciano,E., Vecchiato,W. (2004). *Copula Methods in Finance*, Chichester: Wiley.
- [8] Dissmann,J., Brechmann,E.C., Czado,C., Kurowicka,D. (2011). Selecting and estimating regular vine copulae and application to financial returns. *Submitted for publication*.
- [9] Daykin,C., Pentikainen,T., Pesonen,M. (1984) *Practical risk theory for actuaries*. Chapman Hall.

- [10] EIOPA (2007). QIS 4 Technical Specifications. Relazione tecnica EIOPA.
- [11] EIOPA (2010). QIS 5 Technical Specifications. Relazione tecnica EIOPA.
- [12] Embrechts,P., Lindskog,F., McNeil,A. (2003). *Modelling dependence with copulas and applications to risk management*. In S.T.Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance*. North-Holland: Elsevier.
- [13] Embrechts,P., McNeil,A.J., Straumann,D. (2002). *Correlation and dependency in risk management: properties and pitfalls.*, in *Risk Management: Value at Risk and Beyond*. Cambridge University Press.
- [14] Genest,C. (1987). Frank's family of bivariate distributions *Biometrika* 14 549-555.
- [15] Hajek,S. (2011) *Solvency II. Strumenti per il Risk Management delle aziende assicurative*. Ed. EGEA.
- [16] Hobaek Haff,I. (2010). Parameter estimation for pair-copula constructions. *Submitted for publication*.
- [17] Hobaek Haff,I., Aas,K., Frigessi,A. (2010). On the simplified pair-copula construction - Simply useful or too simplistic? *Journal of Multivariate Analysis* 101(5) 1296-1310.
- [18] International Actuarial Association (I.A.A.) (2010) *Stochastic Modeling - Theory and Reality from an Actuarial Perspective*.
- [19] Joe,H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In L. Rueschendorf and B. Schweizer and M.D. Taylor (Ed.), *Distributions with fixed marginals related Topics*.
- [20] Kendall,M., Stuart,A. (1967) *The Advanced Theory of Statistics, Vol. 2. Inference and relationship, 2nd ed.* Griffin, London.
- [21] Klugman,S.A, Panjer,H.H., Willmot,G.E. (1998) *Loss Models: from Data to Decisions*, Wiley.
- [22] Kurowicka,D., Cooke,R.M. (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. Chichester: Wiley.

- [23] McNeil,A.J., Frey,R., Embrechts,P. (2006) *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.
- [24] Moraldi,F. (2011) *La modellizzazione del costo sinistri del ramo property e la valutazione del Risk capital*. Tesi di laurea, Università Cattolica del Sacro Cuore, Milano.
- [25] Morales,O., Kurowicka,D., Roelen,A. (2006). Elicitation procedures for conditional and unconditional rank correlations *Resources for the future, Expert Judgement Policy Symposium and Technical Workshop, Washington DC, March 13-14*.
- [26] Nelsen,R.B. (1999) *An introduction to copulas*. Springer, New York.
- [27] Salvadori,G., De Michele,C., Kottegoda,N.T., Rosso,R. (2007). *Extremes in Nature: An Approach using Copulas*. Dordrecht: Springer.
- [28] Savelli,N. (2002) Solvency and traditional reinsurance for non-life insurance. *6th International Congress on IME (Insurance: Mathematics and Economics) (Lisbon)*.
- [29] Savelli,N., Clemente,G. (2009a) Hierarchical structures in the aggregation of premium risk for insurance underwriting. *39th ASTIN Colloquium (Helsinki)*.
- [30] Savelli,N., Clemente,G. (2009b) Il progetto solvency 2. *Dispensa di Teoria del Rischio II, UCSC*.
- [31] Savelli,N., Clemente,G. (2009c) Modelling aggregate non-life underwriting risk: standard formula vs internal model. *Giornale dell'Istituto Italiano degli Attuari*.
- [32] Savu,C., Trede,M. (2006). Hierarchical Archimidean copulas. In *International Conference on High Frequency Finance, Konstanz, Germany, May*.
- [33] Schweizer,B., Sklar,A. (1983). *Probabilistic Metric Spaces*. New York: North Holland.
- [34] Venter,G.G. (2001) Tails of copulas. In *Proceedings ASTIN Washington, USA*, pp 68-113.

Daniele Franceschi

CURRICULUM VITAE

Contact Information

Department of Statistics, University of Padova,
Via Cesare Battisti, 241-243 - 35121 Padova (Italy)

e-mail: franceschi@stat.unipd.it

Current Position

Since January 2007; (expected completion: January 2012)

PhD Student in Statistical Sciences, University of Padova.

Thesis title: Selezione e stima di dipendenze multivariate: il caso della *pair-copula construction* con una sua applicazione in campo assicurativo

Supervisor: Prof. Marco Ferrante

Research interests

- Multivariate distribution
- Copula functions
- Vine and Pair-Copula Construction
- Solvency II and Internal Model for non-life insurance business

Education

October 2000 – July 2005

Laurea quadriennale in Matematica.

University of Padova, Faculty of Science MM.FF.NN.

Title of dissertation: “Proprietà delle code della distribuzione stazionaria per processi stocastici nonlineari a tempo discreto ”

Supervisor: Prof. Marco Ferrante

Final mark: 101/110

Further education

October 2005 – April 2006

Course name: Corso di Alta Formazione in Finanza Matematica

Organizing Institution: University of Bologna, Department of Mathematics

Organizer: prof. A. Pascucci, prof. S. Polidoro

Work experience

July 2008 – present

Società Cattolica di Assicurazione.

Employed on permanent contract by Risk Management office in the Non-life business area.

January 2008 – July 2008

Società Cattolica di Assicurazione.

Stage by Risk Management office

May 2006 – August 2006

Unipol Banca.

Stage by Fund Management office with focus on analysis of Unipol Funds' performance.

Computer skills

- Matlab
- C++ (basic)
- LaTeX
- R (basic)
- Microsoft Office Package: Word, Excel, Access, Powerpoint

Language skills

Italian (native), English (good), German (moderate/basic).