

UNIVERSITA' DEGLI STUDI DELLA CALABRIA

25 MAGGIO 2007

## **Metodologie Montecarlo nella riassicurazione dei rami danni**

Marco Pirra

Docente di Teoria del rischio per le assicurazioni contro i danni  
Università degli Studi del Sannio

## 1. Capitolo 1

A Global Framework for Insurer Solvency Assessment: un'analisi critica

## 2. Capitolo 2

Le copule come strumento di analisi delle dipendenze dei rischi

## 3. Capitolo 3

Un modello multivariato

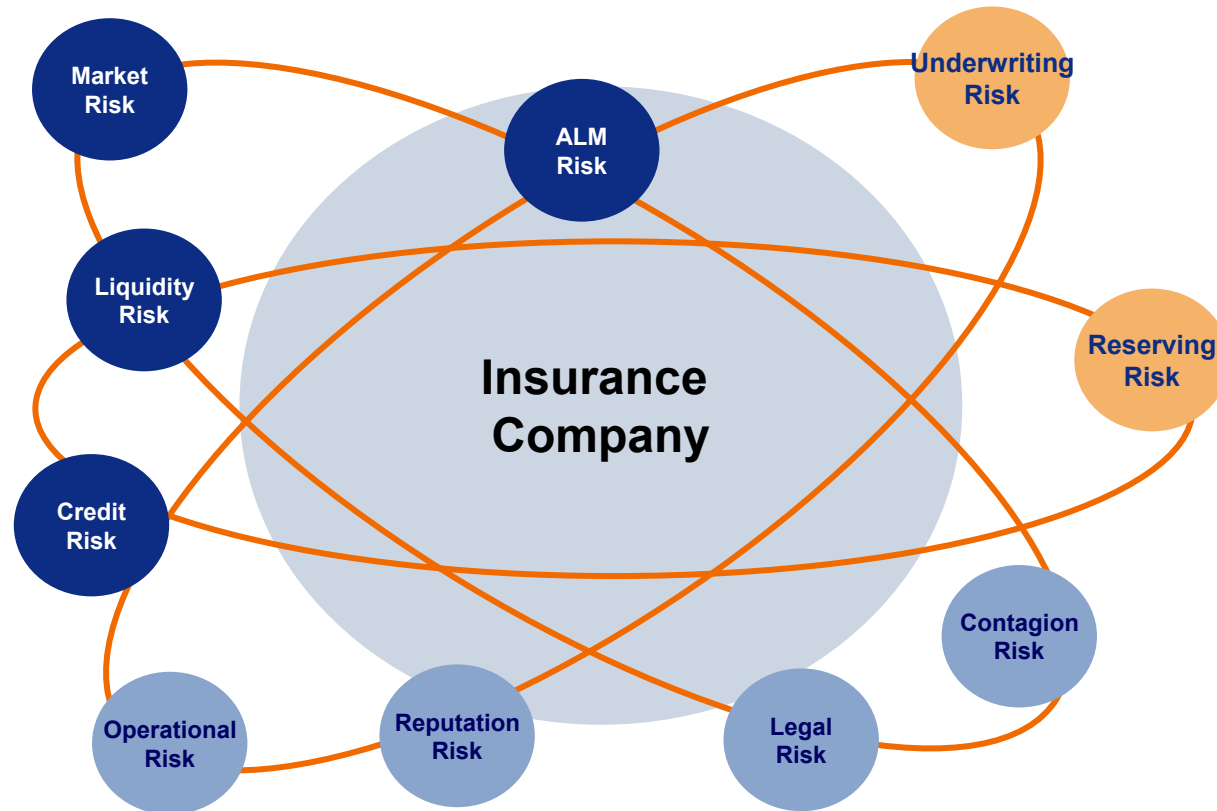
## 4. Capitolo 4

La variabilità del parametro di dipendenza e l'esborso stocastico del riassicuratore

# Capitolo 1: A Global Framework - IAA

- ✓ Il documento **A Global Framework for Insurer Solvency Assessment**, diffuso nel 2004, è stato predisposto da un gruppo di lavoro appositamente costituito dalla Associazione Internazionale degli Attuari (IAA).
- ✓ L'obiettivo prefissato è il **supporto** della attività della **Associazione Internazionale delle Autorità di Vigilanza delle Assicurazioni (IAIS)** per l'analisi degli elementi richiesti ai fini della determinazione di standard di capitale universali, per poter fornire inoltre un approccio, riassumendo le best practices in uso, fruibile per tutte le autorità.
- ✓ Il documento approfondisce i metodi che le autorità di vigilanza potrebbero utilizzare per determinare non solo la **situazione finanziaria** corrente di un assicuratore, ma anche per comprendere il range delle possibili posizioni future.

## ✓ Una possibile classificazione dei rischi



## Capitolo 2: le copule

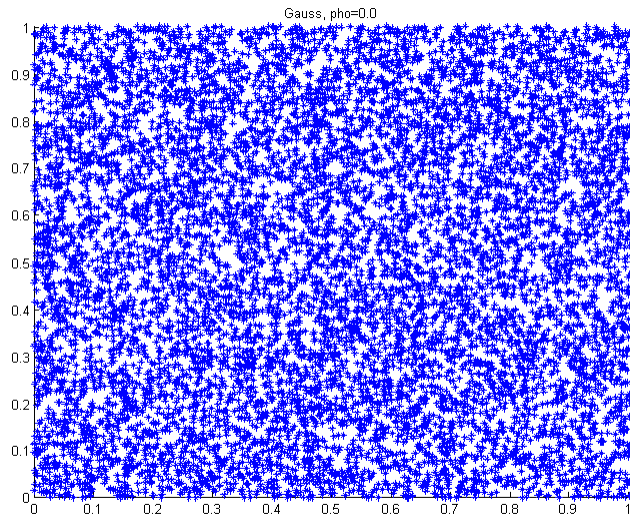
- ✓ T. Mikosch: *“When I started writing the paper in 2003 a Google search of the word “copula” gave 10.000 responses. In September 2005 the same search gives 650.000 responses. There is an explosion of activity. What is going on?”* – **Maggio 2007: 950.000 risultati**
- ✓ Un punto centrale nella analisi del rischio è la costruzione di un modello statistico che descriva e tenga conto della **variabilità** insita nella realtà oggetto di analisi
- ✓ Per lungo tempo la modellistica statistica in ambito attuariale si è basata su ipotesi semplificatrici. La **distribuzione normale** ha dominato gli studi sulle distribuzioni multivariate: veniva assunta solitamente come punto di partenza e raramente messa in discussione.
- ✓ La gestione del rischio assicurativo al giorno d'oggi si trova a far fronte ad una **molteplicità di fattori di rischio**. Pertanto la necessità di modellare la dipendenza al di là della “semplice” normale multivariata è diventata una questione cruciale.

## Capitolo 2: le copule

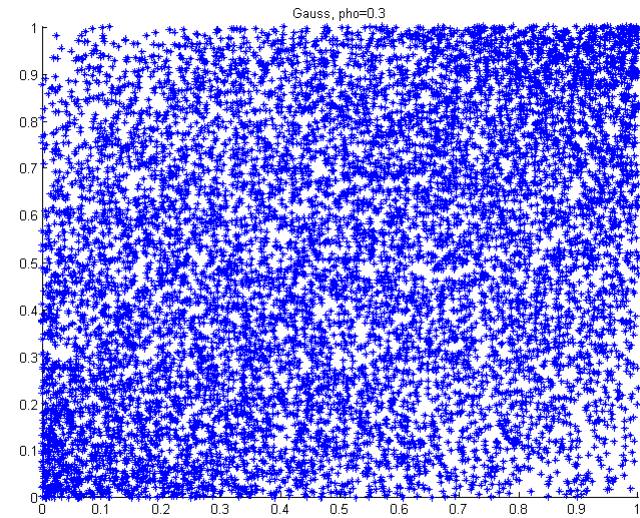
- ✓ L'idea di base delle copule è quella di separare la dipendenza e le distribuzioni marginali da una distribuzione multivariata
- ✓ 1940: Hoeffding studia le proprietà delle distribuzioni multivariate
- ✓ 1959: compare per la prima volta il termine copula (Sklaar)
- ✓ 1998: letteratura accademica sull'uso delle copule in ambito risk management
- ✓ 2004: alcune compagnie di assicurazione ed istituzioni finanziarie iniziano ad usare le copule come strumento di risk management

# Copula Gaussiana

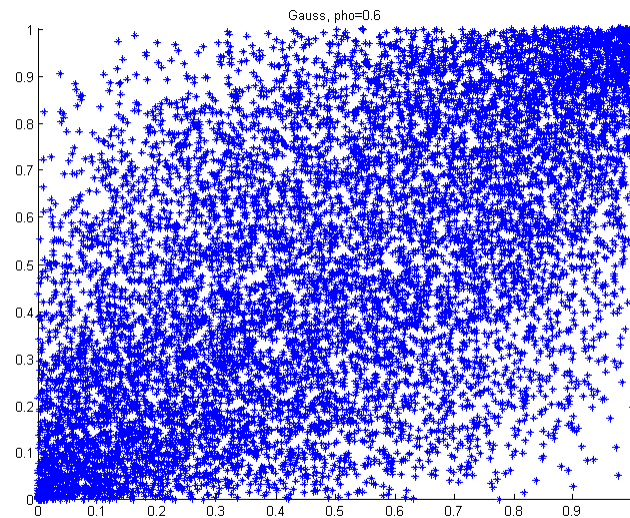
$$\rho = 0.0$$



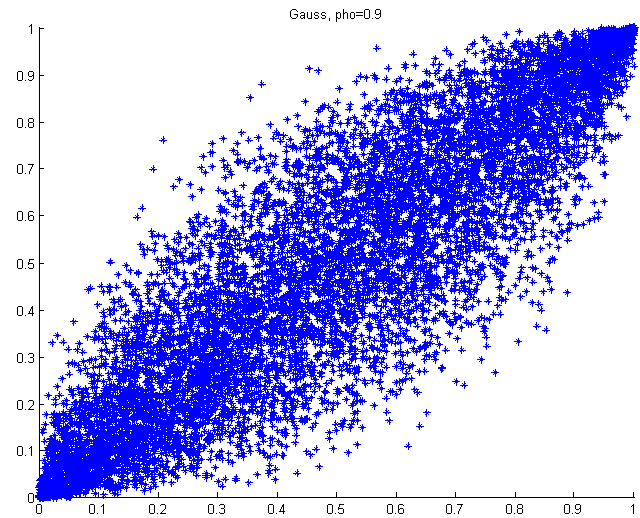
$$\rho = 0.3$$



$$\rho = 0.6$$

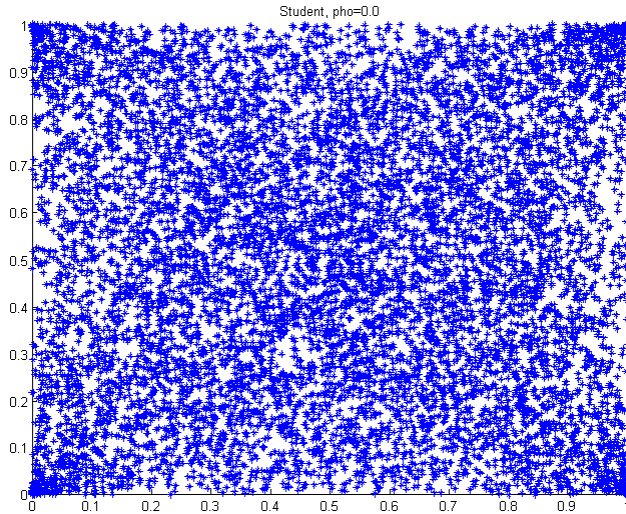


$$\rho = 0.9$$

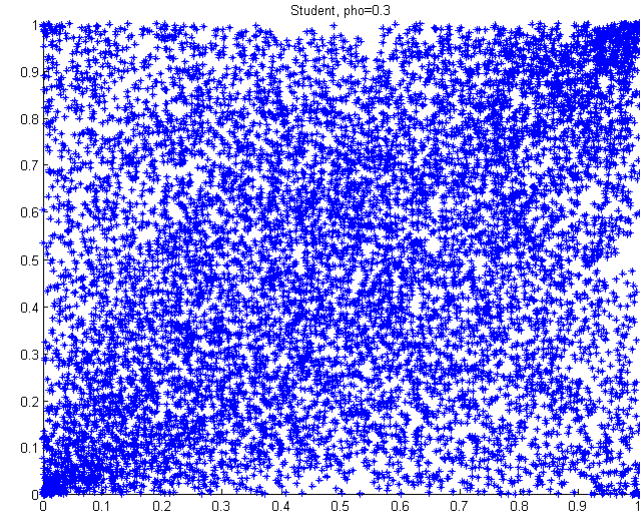


# Copula Student 3 g.d.l.

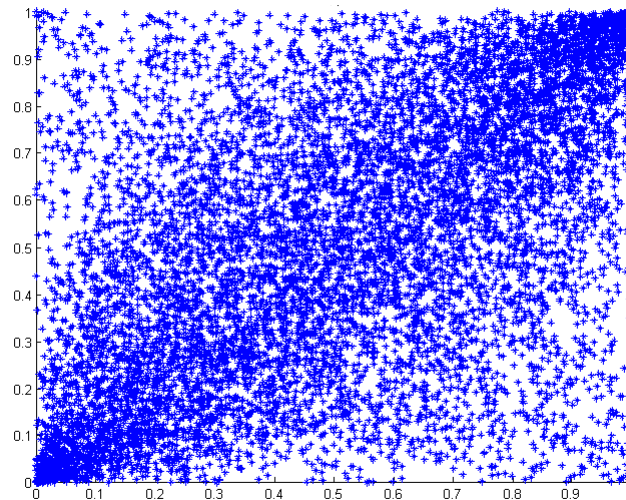
$\rho = 0.0$



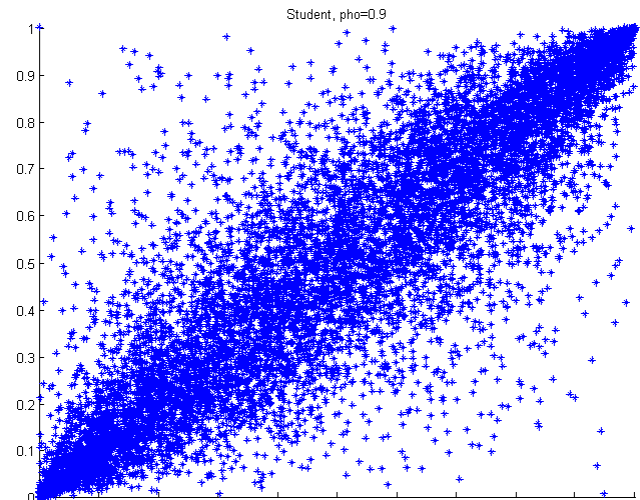
$\rho = 0.3$



$\rho = 0.6$



$\rho = 0.9$





## Capitolo 3: un modello multivariato

- ✓ Obiettivo del lavoro è la valutazione di un **trattato di riassicurazione excess of loss** per il quale l'esborso atteso del riassicuratore è dato da:

$$1 - \quad XRiass = E\left[g(X, Y, S)\right]$$

con

$$2 - \quad g(X, Y, S) = \begin{cases} 0 & \text{se } (X + Y) < k \\ X + Y - k + \frac{X + Y - k}{X + Y} S & \text{se } (X + Y) \geq k \end{cases}$$

- ✓ Denuit, Kaas et al. (2005)

$$3 - \quad g(L, A) = \begin{cases} 0 & \text{se } L < r \\ L - r + \frac{L - r}{L} A & \text{se } r \leq L < \omega \\ \omega - r + \frac{\omega - r}{\omega} A & \text{se } L \geq \omega \end{cases}$$

- ✓ La differenza fondamentale che salta all'occhio è **la scomposizione della variabile aleatoria danno in due componenti** (nella fattispecie danno a cose e danno a persone) : il problema diventa un problema multivariato!

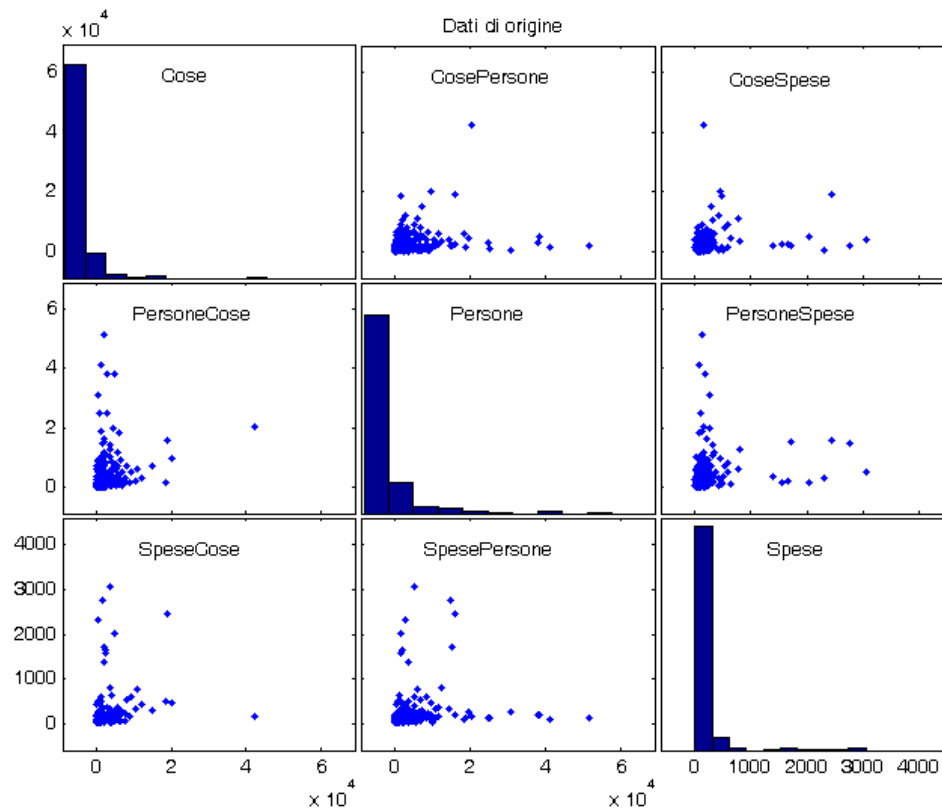
## Capitolo 3: le copule utilizzate

- ✓ Per la implementazione delle **Copule Archimedee (Clayton, Gumbel, Frank)** di dimensione superiore a due si è fatto ricorso alla variabile frailty così come descritta da Marshall e Olkin, (1988), e da Joe, (1997).
- ✓ E' bene sottolineare come le Copule Archimedee di dimensione pari a tre (o superiore a tre) ammettano soltanto dipendenza positiva
- ✓ Oltre a queste Copule Archimedee sono state considerate anche le **Copule Ellittiche (la copula Gaussiana e la copula t di Student)**, nel caso particolare in cui risultino definite attraverso una matrice di dispersione generica:

$$4 - \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{pmatrix}$$

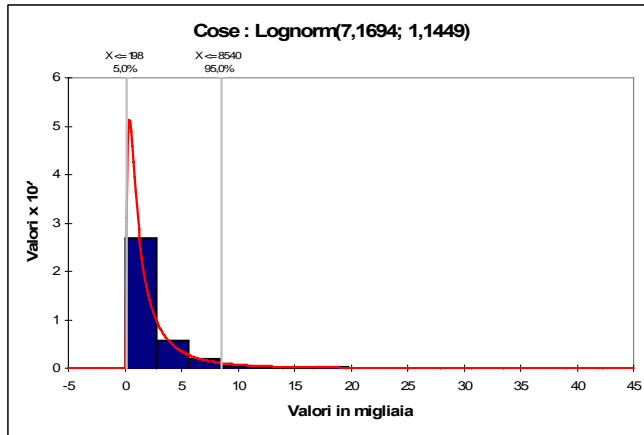
## Capitolo 3: lo studio delle marginali

- ✓ La distribuzione di partenza dei dati analizzati (rappresentati da una matrice di sinistri che hanno dato luogo sia ad un danno a cose sia ad un danno a persone e per i quali sono state pagate delle spese) può essere sinteticamente rappresentata dal seguente grafico

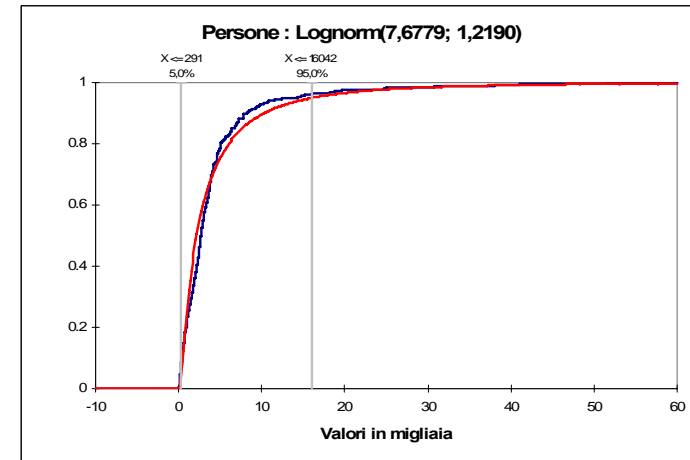


# Capitolo 3: distribuzioni marginali

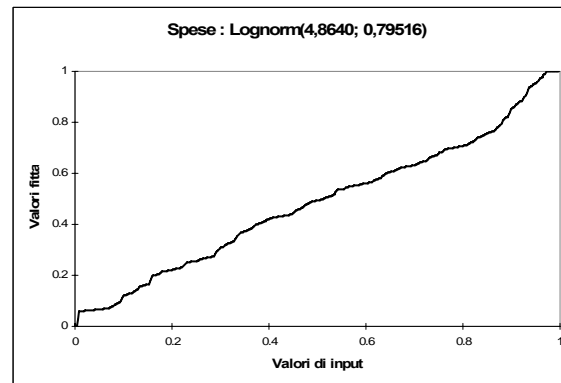
## ✓ Danno a cose



## ✓ Danno a persone



## ✓ Spese



## Capitolo 3: distribuzione congiunta

- ✓ Per quanto riguarda lo studio della distribuzione congiunta, una volta costruita la funzione di verosimiglianza, si è proceduto alla massimizzazione di questa funzione andando a ricavare gli stimatori dei parametri e l'errore della stima.

$$5 - l(\theta) = \sum_{i=1}^n \log c \{ F_1(X_{i1}; \lambda), \dots, F_k(X_{ik}; \lambda); \alpha \} + \sum_{i=1}^n \sum_{j=1}^k \log f_i(X_{ij}; \lambda)$$

$$6 - \hat{\theta}_{MAXV} = \arg \max_{\theta \in \Theta} l(\theta)$$

# Capitolo 3: risultati stima (1/2)

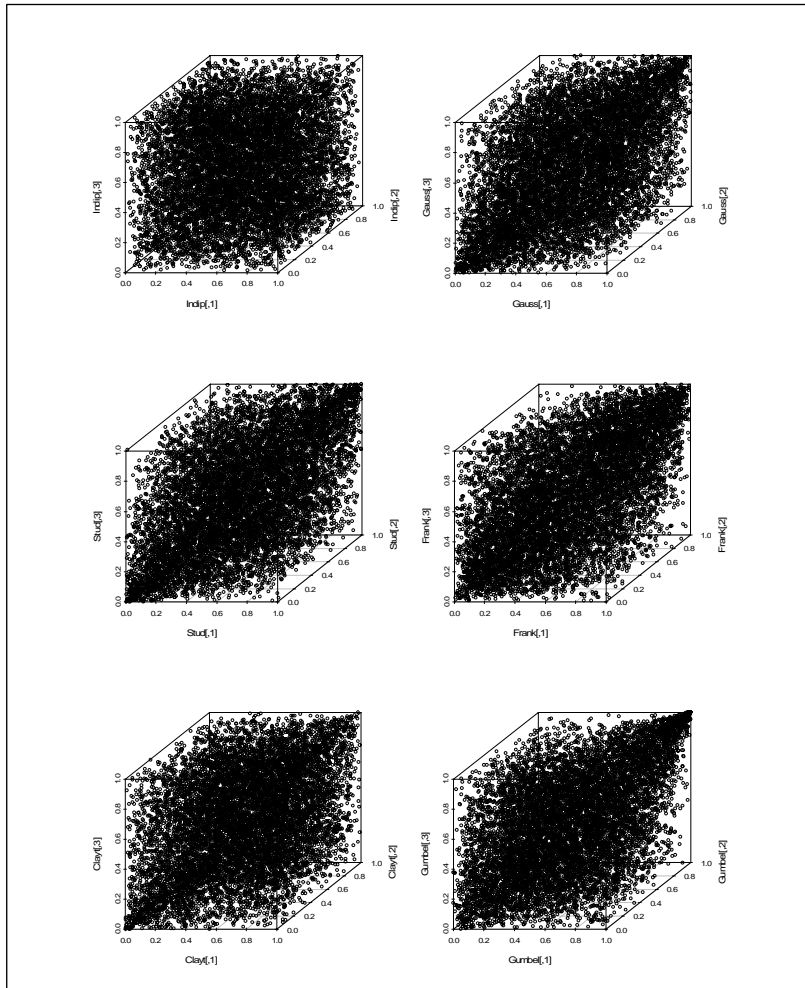
	GAUSSIANA		STUDENT		FRANK	
Max LogVerosimiglianza	-7.610,906		-7.610,418		-7.608,614	
	Stima	Std.Err.	Stima	Std.Err.	Stima	Std.Err.
<i>Marginale Cose</i>						
$\mu$	7,169408	0,06420161	7,170424	0,06416636	<b>7,162167</b>	<b>0,06634803</b>
$\Sigma$	1,144877	0,04549729	1,145622	0,04559453	<b>1,182203</b>	<b>0,04868928</b>
<i>Marginale Persone</i>						
$\mu$	7,678005	0,06835586	7,677498	0,0686053	<b>7,694094</b>	<b>0,06965384</b>
$\Sigma$	1,218961	0,04833482	1,225368	0,04928326	<b>1,244579</b>	<b>0,05029681</b>
<i>Marginale Spese</i>						
$\mu$	4,864091	0,04459069	4,8607288	0,04484499	<b>4,8516812</b>	<b>0,04405146</b>
$\Sigma$	0,795162	0,03153035	0,7995199	0,03215744	<b>0,8017709</b>	<b>0,0317579</b>
<i>Dipendenza</i>						
$\rho1/ \text{alfa}$	0,2850225	0,05152136	0,297131	0,05422855	<b>2,40091</b>	<b>0,280595</b>
$\rho2$	0,3230299	0,05022575	0,3394071	0,05406583	<i>n.d.</i>	<i>n.d.</i>
$\rho3$	0,3469253	0,04932792	0,3589707	0,05167475	<i>n.d.</i>	<i>n.d.</i>
<i>g.d.l.</i>	<i>n.d.</i>	<i>n.d.</i>	42,7928507	33,79621312	<i>n.d.</i>	<i>n.d.</i>

## Capitolo 3: risultati stima (2/2)

	CLAYTON		GUMBEL	
Max LogVerosimiglianza	-7.633,144		-7.620,542	
	Stima	Std.Err.	Stima	Std.Err.
<i>Marginale Cose</i>				
$\mu$	7,175667	0,06560731	7,206565	0,06530276
$\Sigma$	1,173862	0,0471881	1,18581	0,04770228
<i>Marginale Persone</i>				
$\mu$	7,68121	0,06795253	7,697965	0,06876125
$\Sigma$	1,217149	0,04759162	1,237879	0,04954977
<i>Marginale Spese</i>				
$\mu$	4,8786243	0,04373949	4,8810487	0,04591241
$\Sigma$	0,7839897	0,03037782	0,8328449	0,03388796
<i>Dipendenza</i>				
$\rho1/ \text{alfa}$	0,243012	0,05071211	1,268017	0,04681707
$\rho2$	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>
$\rho3$	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>
<i>g.d.l.</i>	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>



## Capitolo 3: simulazione terne



- ✓ Generare variabili da una copula generica può essere fatto attraverso il condizionamento iterativo [Bouyè et al., 2000]
- ✓ Per alcune copule Archimedee con dimensione superiore a due esiste un algoritmo quando la funzione inversa del generatore è la trasformata di Laplace di una qualche variabile aleatoria positiva (la frailty).

## Capitolo 3: esborso atteso

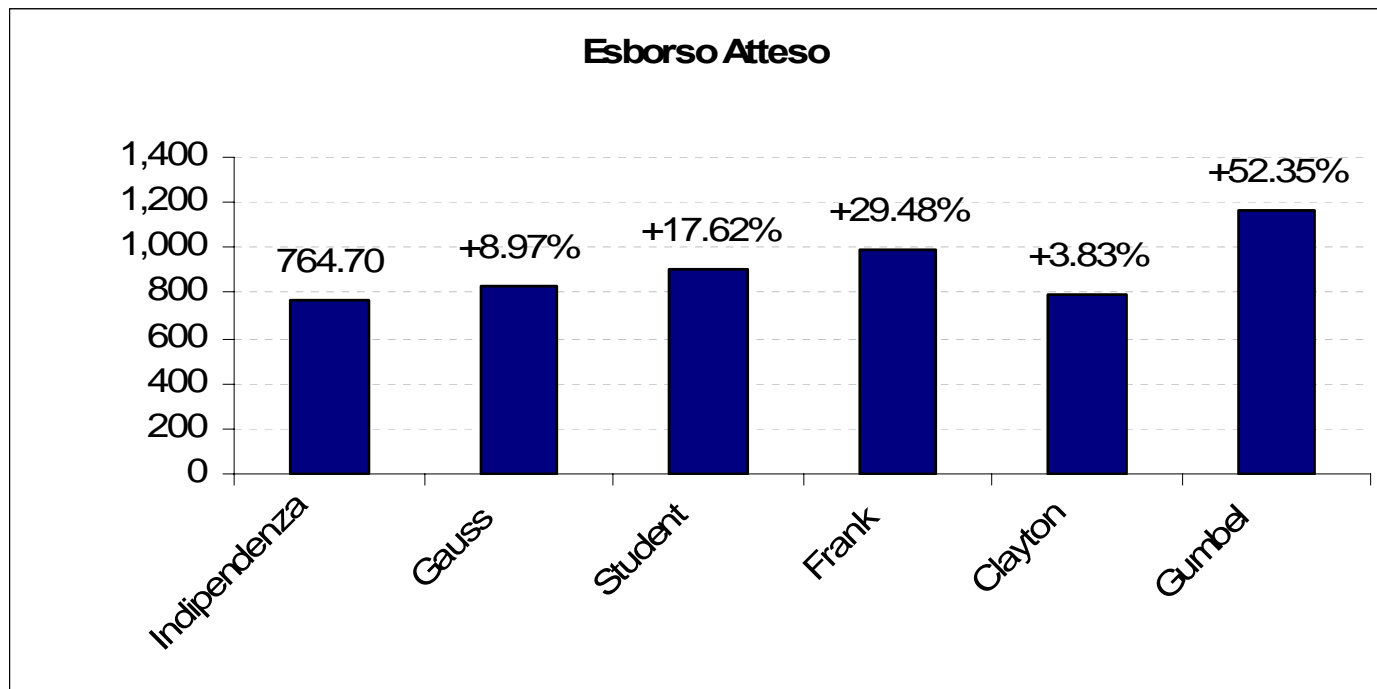
- ✓ I risultati relativi all'esborso atteso del riassicuratore e alla variabilità dello stesso sono i seguenti:

	Media Cose	Media Persone	Media Spese	Esborso Atteso	Varianza Esborso	Delta Esborso Atteso (*)	Delta Varianza Esborso(*)
<i>Indipendenza</i>	2.513,90	4.578,90	177,86	764,70	6.196,30	-----	-----
<i>Gauss</i>	2.502,90	4.540,40	177,87	833,32	6.139,10	8,97%	-0,92%
<i>Student</i>	2.515,90	4.585,20	179,04	899,41	6.383,50	17,62%	3,02%
<i>Frank</i>	<b>2.407,60</b>	<b>4.831,90</b>	<b>175,84</b>	<b>990,11</b>	<b>8.112,50</b>	<b>29,48%</b>	<b>30,92%</b>
<i>Clayton</i>	2.600,90	4.550,30	179,43	793,98	6.057,30	3,83%	-2,24%
<i>Gumbel</i>	2.721,30	4.703,70	188,18	1.165,00	9.324,30	52,35%	50,48%

(\*) Rispetto al caso di indipendenza

## Capitolo 3: esborso atteso

- ✓ Se si concentra l'attenzione sull'esborso atteso, prendendo a riferimento il valore nel caso di indipendenza:



## Capitolo 3: considerazioni

- ✓ Al fine di dare un **significato** alla scelta iniziale da cui è partita l'analisi, vale a dire la rappresentazione dell'esborso del riassicuratore nella forma (2), **i risultati ottenuti vanno evidentemente confrontati con i risultati che si otterrebbero rappresentando l'esborso del riassicuratore nella forma (3).**
- ✓ Considerando un'unica variabile aleatoria *danno*

	Media Danno	Media Spese	Esborso Atteso	Varianza Esborso	Delta Esborso Atteso (*)	Delta Varianza Esborso(*)
<i>Indipendenza</i>	7.074,30	178,48	805,69	5.630,10	-----	-----
<i>Frank</i>	7.327,02	177,68	870,62	6.257,81	8,06%	11,15%

(\*) Rispetto al caso di indipendenza

## Capitolo 3: considerazioni

✓ **Ma:**

	<b>Esborso Atteso</b>	<b>Varianza Esborso</b>	<b>Delta Esborso Atteso</b>	<b>Delta Varianza Esborso</b>
<i>Frank 3 v.a.</i>	990,11	8.112,50	-----	-----
<i>Frank 2 v.a.</i>	870,62	6.257,81	-12,07%	-22,86%

- ✓ Solitamente i sinistri più onerosi sono quelli che provocano sia un danno a cose sia un danno a persone; tra queste due componenti del danno c'è solitamente una dipendenza positiva (a danno a cose di grande entità corrisponde danno a persone di grande entità, qualora il sinistro abbia provocato entrambe le tipologie) non trascurabile; dipendenza positiva c'è infine tra le componenti del danno e le spese (la logica porta a pensare che a sinistri che hanno provocato danni ingenti siano associate le spese maggiori, visto che l'evidenza empirica peraltro dimostra come i sinistri onerosi siano solitamente gli ultimi ad essere liquidati, dopo lunghe cause legali o perizie di più tecnici).

## Capitolo 4: considerazioni

- ✓ Il modello proposto, così come del resto i modelli sviluppati in letteratura presenta un **limite: la stima del parametro (parametri) che caratterizza la struttura di dipendenza, qualsiasi essa sia, è una stima puntuale;**
- ✓ In altri termini, l'esborso atteso del riassicuratore che si riesce a calcolare, nonostante sia calcolato attraverso un approccio simulativo e non in forma chiusa, è un valore "deterministico".
- ✓ L'obiettivo del capitolo è quello di studiare la funzione di distribuzione del parametro di dipendenza, che non ha una forma chiusa o riconducibile a funzioni statistiche note, al fine di evitare una sottostima del rischio oggetto di analisi. Il fine è quello di **pervenire ad un esborso atteso del riassicuratore "stocastico", che tenga conto dell'incertezza insita nel parametro.**

## Capitolo 4: l'approccio seguito

- ✓ L'approccio seguito sfrutta le potenzialità delle tecniche inferenziali Bayesiane congiuntamente alle tecniche simulative basate su metodo Monte Carlo e Catene di Markov (meglio note in letteratura come Markov Chain Monte Carlo techniques, o più semplicemente MCMC) [Gilks, 1996]
- ✓ Esistono anche altre metodologie, una tra queste di relativo successo tra gli attuari, vale a dire il bootstrapping:
  - ✓ l'idea alla base del bootstrapping è che, qualora non si abbiano informazioni sufficienti sulla popolazione da cui proviene il campione oggetto di analisi, i valori del campione siano la migliore guida alla distribuzione reale.
  - ✓ Lo svantaggio di questo metodo è che i valori del parametro sono ridotti ad un intervallo e si possono ottenere distribuzioni multimodali.

## Capitolo 4: MCMC

- ✓ Le tecniche Bayesiane rappresentano l'incertezza dei valori di parametri ignoti attraverso funzioni di probabilità, e considerano quindi i parametri variabili aleatorie.
- ✓ Sia  $D$  (nel nostro caso  $X, Y$  ed  $S$ ) l'insieme di dati osservati a disposizione e  $\Theta$  il vettore dei parametri del modello: per ottenere qualsiasi informazione inferenziale, dovremmo conoscere la distribuzione di probabilità congiunta  $P(D, \theta)$

$$7 - P(D, \theta) = P(\theta)P(D|\theta) \quad \longleftarrow \quad P(D|\theta) = L(\theta; D)$$

$$8 - P(\theta|D) = \frac{P(\theta)P(D|\theta)}{\int P(\theta)P(D|\theta)d\theta} \quad \longrightarrow \quad P(\theta|D) \propto P(\theta)P(D|\theta) = P(\theta)L(\theta; D)$$



- ✓ Una delle critiche solitamente mosse alla analisi Bayesiana è il fatto che la **distribuzione a posteriori dipende dalla scelta della distribuzione a priori**. In un lavoro di Forte e Pirra (2006) viene analizzato l'effetto della scelta di distribuzioni a priori differenti e verificato **l'impatto marginale** che tale scelta ha se il numero delle simulazioni è elevato.
  
- ✓ Altro importante concetto derivante dalla analisi Bayesiana è quello della **distribuzione predittiva**, vale a dire la distribuzione di una nuova osservazione condizionata ai dati storici. La distribuzione predittiva può essere calcolata come mistura della distribuzione a posteriori e della distribuzione del modello. **Sarà proprio la distribuzione predittiva quella utilizzata per il calcolo dei valori relativi all'esborso del riassicuratore.**

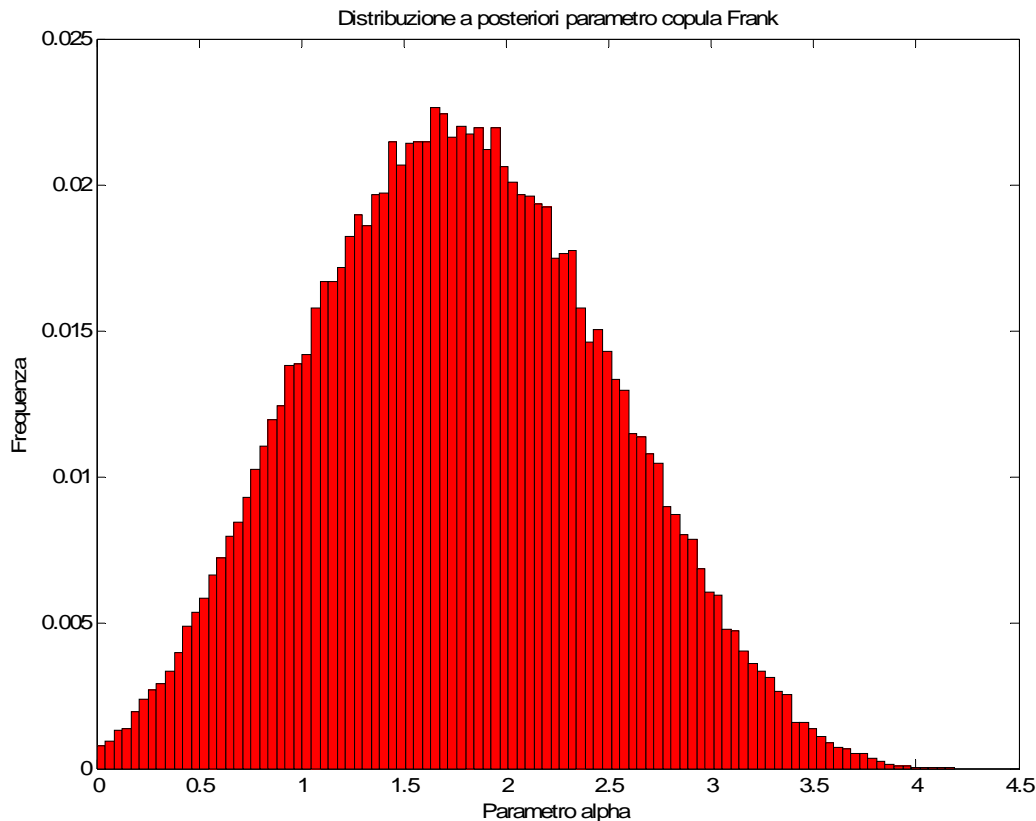
## Capitolo 4: la variabilità dei parametri nelle copule

- ✓ Le distribuzioni a posteriori dei parametri delle copule non sono distribuzioni standard e risulta pertanto complicato estrarre dei campioni da queste utilizzando delle metodologie standard (come ad esempio l'inversione della funzione di ripartizione).
- ✓ Tuttavia l'algoritmo **Metropolis-Hastings** è una valida soluzione per superare questo ostacolo.
- ✓ Nel caso di distribuzioni a priori uniformi la funzione di densità della distribuzione a posteriori è data dalla funzione di verosimiglianza:

$$9 - L(\boldsymbol{\theta}; \mathbf{X}) = \prod_{i=1}^n f(X_i; \boldsymbol{\theta})$$

## Capitolo 4: copula Frank

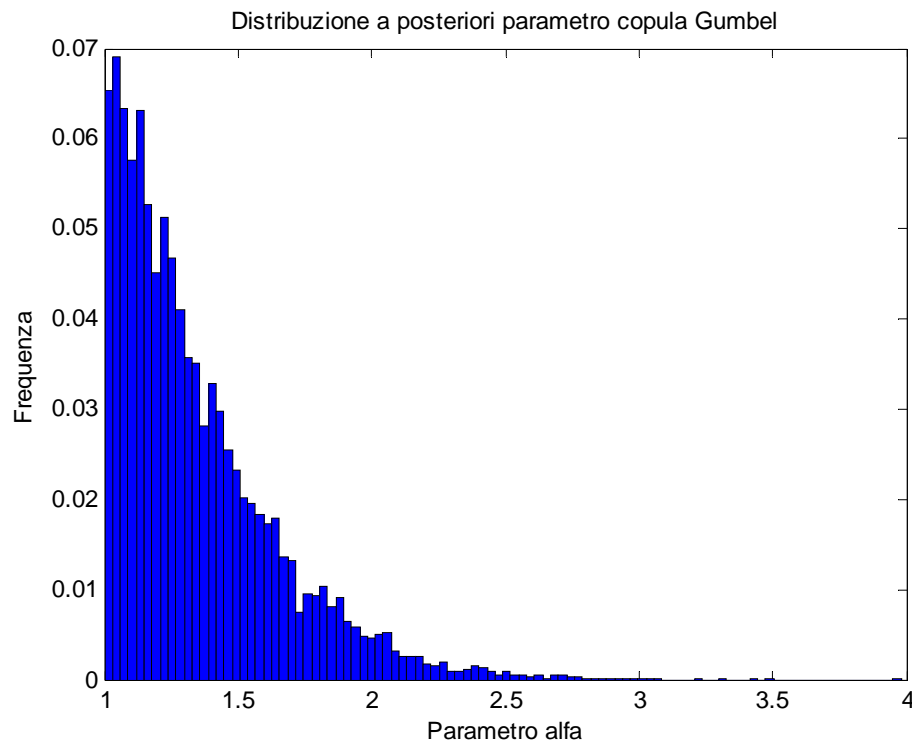
- ✓ La distribuzione simulata del parametro della copula Frank a tre variabili, data la matrice di dati di partenza, è la seguente:



Frank	MCMC	MAX. VER.
<i>Stima del parametro</i>	2,0214	1,9126
<i>Standard error</i>	0,7122	0,2285
<i>Estremo Inf Intervallo 95%</i>	0,4493	1,4648
<i>Estremo Sup Intervallo 95%</i>	3,1816	2,1353
<i>Indice di asimmetria</i>	0,1061	----

## Capitolo 4: copula Gumbel

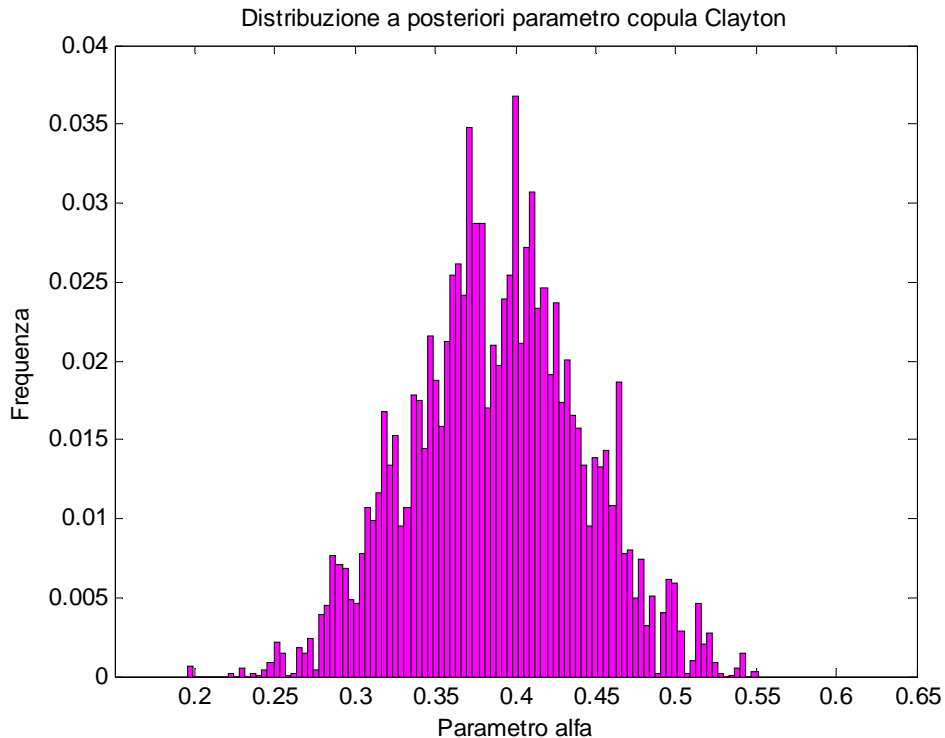
- ✓ La distribuzione simulata del parametro della copula Gumbel a tre variabili, data la matrice di dati di partenza, è la seguente:



Gumbel	MCMC	MAX. VER.
<i>Stima del parametro</i>	1,3429	1,2404
<i>Standard error</i>	0,3008	0,0350
<i>Estremo Inf Intervallo 95%</i>	1,0102	1,1719
<i>Estremo Sup Intervallo 95%</i>	2,1052	1,2746
<i>Indice di asimmetria</i>	1,4713	----

## Capitolo 4: copula Clayton

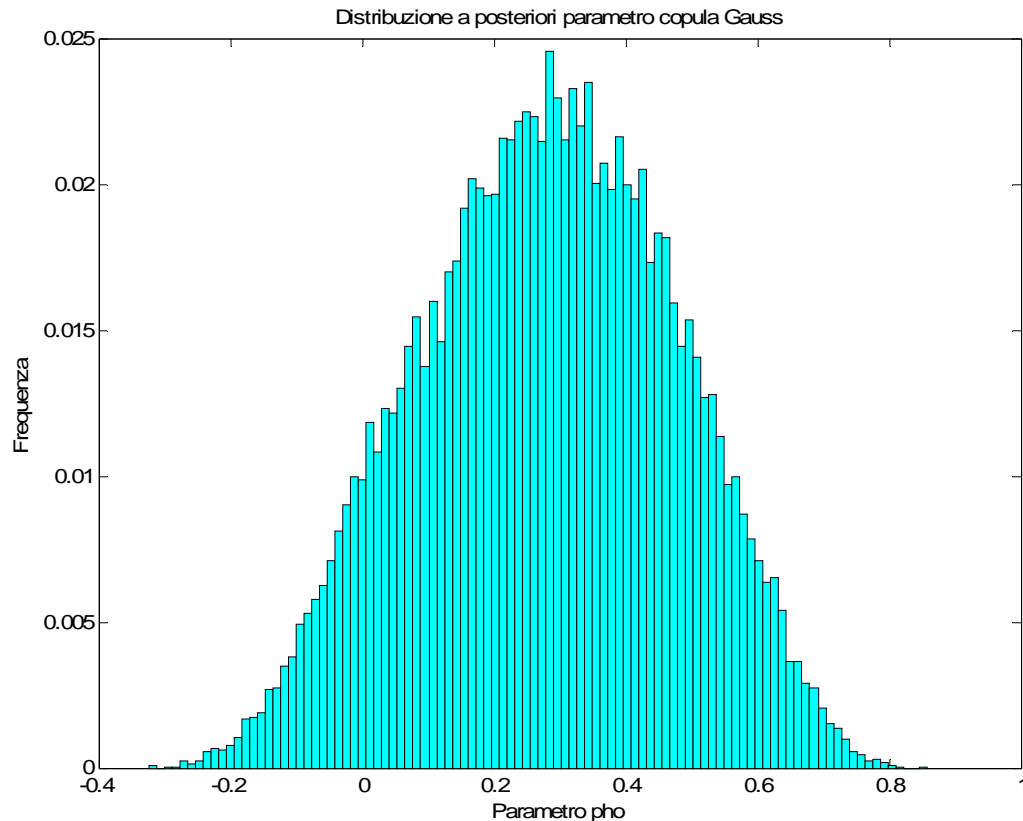
- ✓ La distribuzione simulata del parametro della copula Clayton a tre variabili, data la matrice di dati di partenza, è la seguente:



Clayton	MCMC	MAX.VER.
<i>Stima del parametro</i>	0,3886	0,3877
<i>Standard error</i>	0,0533	0,0538
<i>Estremo Inf Intervallo 95%</i>	0,2867	0,2823
<i>Estremo Sup Intervallo 95%</i>	0,4968	0,4401
<i>Indice di asimmetria</i>	0,0094	----

## Capitolo 4: copula Gauss

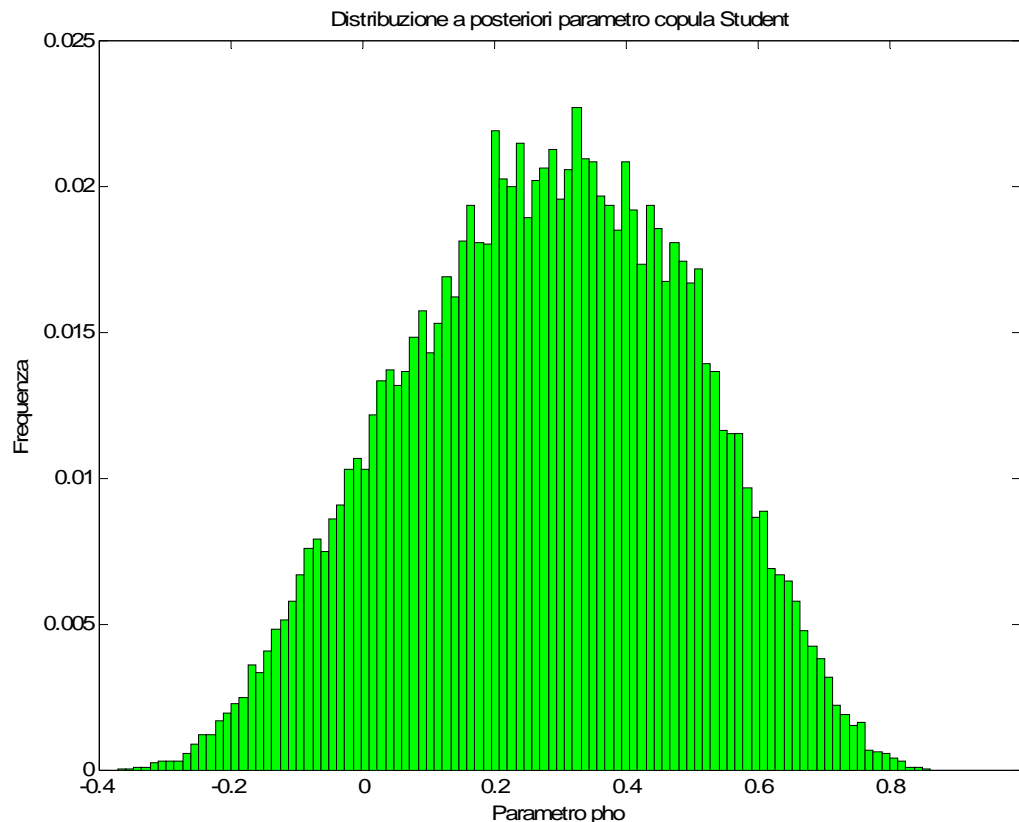
- ✓ La distribuzione simulata del parametro della copula Gauss a tre variabili, data la matrice di dati di partenza, è la seguente:



Gaussiana	MCMC	MAX.VER.
<i>Stima del parametro</i>	0,2800	0,3473
<i>Standard error</i>	0,1928	0,0329
<i>Estremo Inf Intervallo 95%</i>	-0,0930	0,2829
<i>Estremo Sup Intervallo 95%</i>	0,6343	0,3794
<i>Indice di asimmetria</i>	-0,0879	----

## Capitolo 4: copula Student

- ✓ La distribuzione simulata del parametro della copula Student a tre variabili, data la matrice di dati di partenza, è la seguente:



Student	MCMC	MAX. VER.
<i>Stima del parametro</i>	0,2793	0,3481
<i>Standard error</i>	0,2134	0,0349
<i>Estremo Inf Intervallo 95%</i>	-0,1357	0,2796
<i>Estremo Sup Intervallo 95%</i>	0,6667	0,3821
<i>Indice di asimmetria</i>	-0,1157	----

## Capitolo 4: esborso stocastico

- ✓ Nella logica bayesiana la distribuzione più efficace per rappresentare le dipendenze tenendo in considerazione la incertezza insita nel parametro è la distribuzione predittiva, che è che una mistura della distribuzione a posteriori dei valori del parametro e la distribuzione copula.
- ✓ Nel nostro contesto attuariale, la distribuzione predittiva risulta fondamentale per la rappresentazione dell'esborso del riassicuratore come intervallo di confidenza (intervallo di confidenza della funzione di distribuzione del parametro).
- ✓ Sfruttando le risultanze dell'analisi Bayesiana, riportate nel paragrafo precedente, misturando la logica MCMC con la funzione copula, si ottengono nuovi valori relativi all'esborso del riassicuratore.



## Capitolo 4: risultati

- ✓ In questi ricalcoli, rispetto a quelli svolti nel capitolo precedente, cambiano i valori del parametro di dipendenza, non più inteso come stima di massima verosimiglianza, ma **estratto dalla sua distribuzione**.

Struttura di dipendenza	Esborso Atteso Media Parametro	Esborso Atteso Confidenza 95%		Esborso Atteso Massima Verosimiglianza
		Limite inferiore	Limite Superiore	
<i>Frank</i>	€ 931	€ 842	€ 976	€ 899
<i>Gumbel</i>	€ 1.218	€ 847	€ 1.419	€ 1.159
<i>Clayton</i>	€ 823	€ 800	€ 859	€ 815
<i>Gaussiana</i>	€ 870	€ 725	€ 1.054	€ 874
<i>Student</i>	€ 895	€ 727	€ 1.091	€ 903

## Considerazioni conclusive

- ✓ In questo lavoro di ricerca si è cercato di avvalorare, matematicamente e computazionalmente, una tesi ben definita
- ✓ Questo però non è stato un punto di arrivo della ricerca, ma un nuovo punto di partenza per analisi approfondite: **si è infatti dimostrato come sia limitativo sintetizzare gli effetti della dipendenza in un unico valore del parametro** che la rappresenta e quali effetti abbia la variabilità implicita nel parametro, ovvero la variabilità della sua distribuzione.
- ✓ **Il ragionamento esplicito nei capitoli terzo e quarto rientra nella logica ispiratrice del progetto Solvency II**, che come analizzato nel primo capitolo del lavoro, persegue l'obiettivo di non trascurare ma anzi quantificare tutti i rischi che caratterizzano il mondo delle assicurazioni

## Bibliografia essenziale

1. International Actuarial Association (IAA), (2004), *A Global Framework for Insurer Solvency Assessment*, Research Report of the Insurer Solvency Assessment Working Party.
2. Nelsen R.B., (1999), *An Introduction to Copulas*, Lecture Notes in Statistics No. 139, New York: Springer-Verlag.
3. Denuit M., Dhaene J., Goovaerts M., Kaas R., (2005), *Actuarial Theory for Dependent Risks*, Wiley & Sons.
4. Gilks W.R., Richardson S., Spiegelhalter D.J., (1996), *Markov Chain Monte Carlo in Practice*, London, Chapman and Hall.