Introduzione alle Copule

Rostagno

September 13, 2024

Le copule sono strumenti matematici che permettono di modellare e stimare la dipendenza tra diverse variabili casuali. Sono particolarmente utili in finanza, dove la dipendenza tra i rendimenti degli asset, i tassi di interesse e i tempi di default sono fattori cruciali per la valutazione del rischio e la determinazione del prezzo di strumenti finanziari complessi. **L'importanza delle copule risiede nella loro capacità di separare la modellazione delle distribuzioni marginali delle singole variabili dalla modellazione della loro struttura di dipendenza** .
In altre parole, invece di dover specificare una funzione di distribuzione congiunta per tutte le variabili, è possibile utilizzare una copula per combinare le distribuzioni marginali di ciascuna variabile in una distribuzione congiunta che rifletta la dipendenza desiderata. Questo approccio offre grande flessibilità nella modellazione, poiché consente di scegliere le distribuzioni marginali e la copula in modo indipendente, a seconda delle caratteristiche specifiche dei dati e del problema in esame.
Ad esempio, si potrebbe utilizzare una distribuzione t di Student per modellare i rendimenti degli indici azionari, che spesso presentano code più spesse rispetto alla distribuzione normale, e quindi utilizzare una copula di Gumbel per rappresentare la dipendenza asimmetrica tra i mercati, con una maggiore probabilità di movimenti congiunti al rialzo rispetto a quelli al ribasso.
La teoria delle copule si basa sul teorema di Sklar, che afferma che ogni funzione di distribuzione congiunta può essere espressa in termini di una copula e delle distribuzioni marginali delle variabili. Il teorema di Sklar garantisce l'esistenza e l'unicità della copula nel caso di variabili casuali continue.
Esistono diverse famiglie di copule, ciascuna con proprietà specifiche in termini di dipendenza di coda, simmetria e altre caratteristiche. Alcune delle famiglie di copule più utilizzate in finanza includono la copula gaussiana, la copula t di Student, le copule Archimedee (come la copula di Gumbel, la copula di Clayton e la copula di Frank) e la copula di Marshall-Olki.
La scelta della copula più adatta dipende dalla natura del problema e dalle

caratteristiche della dipendenza che si desidera modellare. Ad esempio, la copula t di Student è spesso preferita alla copula gaussiana quando si vogliono modellare dipendenze di coda più elevate, mentre le copule Archimedee consentono di modellare diversi tipi di dipendenza asimmetrica.
Le copule sono strumenti potenti che trovano applicazione in diversi ambiti della finanza, tra cui:

- Pricing di opzioni multivariate e altri derivati: le copule possono essere utilizzate per modellare la dipendenza tra i sottostanti di un'opzione basket, un'opzione rainbow o altri derivati multi-asset , consentendo una valutazione più accurata del prezzo di questi strumenti.

- Gestione del rischio: le copule sono ampiamente utilizzate nella modellazione del rischio di credito, dove consentono di stimare la probabilità di default congiunta di diverse attività o controparti. Le copule sono anche utilizzate nella stima del Value at Risk (VaR) di portafogli contenenti attività con distribuzioni non normali e dipendenze complesse.

- Calibrazione e simulazione: la flessibilità delle copule consente di calibrare i modelli ai dati di mercato in modo efficiente e di simulare scenari di mercato realistici che tengano conto della dipendenza tra le variabili.

In sintesi, le copule rappresentano uno strumento matematico versatile e potente per la modellazione della dipendenza in finanza, con un ampio spettro di applicazioni pratiche nella valutazione del rischio, nel pricing di derivati e nella gestione del portafogli.

Fondamenti Matematici delle Copule

Le copule sono strumenti matematici che permettono di modellare e rappresentare la dipendenza tra variabili casuali. A differenza di misure di dipendenza tradizionali come la correlazione lineare, le copule catturano la dipendenza in modo più completo, includendo la dipendenza nelle code delle distribuzioni e non limitandosi a relazioni lineari.
Ecco una spiegazione delle formule e delle proprietà chiave:

Definizione di Copula

Una d-copula è una funzione $C : [0, 1]^d \to [0, 1]$, dove $d \geq 2$ (numero di variabili, nelle proprietà seguenti consideriamo le copule bivariate), che soddisfa le seguenti proprietà:

1. **Groundedness:**
   $C(u,0) = C(0,v) = 0$ per ogni $u, v \in [0,1]^2$. Ciò significa che la copula è zero se una delle variabili è zero.

2. **Marginalità:**
   $C(u,1) = u$, $C(1,v) = v$ per ogni $u, v \in [0,1]^2$. Questa proprietà assicura che la copula sia coerente con le distribuzioni marginali, ovvero che quando una delle variabili assume il suo valore massimo, la copula coincide con la funzione di ripartizione dell'altra variabile.

3. **2-crescita (o 2-increasing):**
   $C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0$ per ogni $u_1 \leq u_2$, $v_1 \leq v_2$ in $[0,1]^2$. Questa proprietà assicura che la copula sia non decrescente in entrambe le variabili, il che è necessario affinché la copula rappresenti una dipendenza positiva o negativa tra le variabili.

Teorema di Sklar

Questo teorema, centrale nella teoria delle copule, stabilisce un legame tra le copule e le funzioni di distribuzione congiunta.
In breve, il teorema afferma che:
Data una funzione di distribuzione congiunta $F(x,y)$ con marginali $F_1(x)$ e $F_2(y)$, esiste una copula C tale che:

$$F(x,y) = C(F_1(x), F_2(y))$$

Inoltre, se $F_1(x)$ e $F_2(y)$ sono continue, allora la copula C è unica.

Conseguenze del Teorema di Sklar:

- Costruzione di modelli di dipendenza: Permette di costruire una funzione di distribuzione congiunta a partire da distribuzioni marginali arbitrarie e da una copula che ne modella la dipendenza. Questa proprietà è particolarmente utile per modellare dati reali, dove spesso si conoscono le distribuzioni marginali ma non la struttura di dipendenza.

- Separazione tra marginali e dipendenza: Mette in luce come la struttura di dipendenza tra le variabili sia completamente catturata dalla copula, indipendentemente dalle distribuzioni marginali.

Famiglie di copule

Esistono diverse famiglie di copule, classificate in base alla loro struttura o ai metodi utilizzati per la loro costruzione. Di seguito, vengono elencate alcune delle principali famiglie:

Fréchet-Hoeffding: Questa famiglia include le copule che rappresentano i limiti inferiore (W) e superiore (M) della dipendenza tra due variabili casuali. La copula W rappresenta la perfetta dipendenza negativa, mentre la copula M rappresenta la perfetta dipendenza positiva

Cuadras-Augé: Questa famiglia di copule è costruita come una media geometrica ponderata delle copule M e P, dove P rappresenta l'indipendenza tra le variabili

Marshall-Olkin: Questa famiglia di copule è spesso utilizzata per modellare la dipendenza tra variabili casuali che rappresentano tempi di vita.

Archimedee: Queste copule sono generate da una funzione detta "generatore". Le copule Archimedee sono popolari per la loro flessibilità e la relativa facilità di utilizzo

Proprietà delle copule

Le copule posseggono diverse proprietà che le rendono utili per la modellazione della dipendenza. Alcune di queste proprietà sono:

- **Invarianza rispetto a trasformazioni monotone crescenti:** Le copule sono invarianti rispetto a trasformazioni strettamente crescenti delle variabili marginali.
- **Misure di concordanza:** Diverse misure di concordanza come la rho di Spearman e la tau di Kendall possono essere espresse in termini di copule.
- **Dipendenza di coda:** Le copule possono catturare la dipendenza tra le code delle distribuzioni marginali, ovvero la tendenza delle variabili ad assumere valori estremi congiuntamente.

Possiamo quindi affermare che le copule offrono un approccio potente e flessibile per la modellazione della dipendenza tra variabili casuali. La loro capacità di separare la struttura di dipendenza dalle distribuzioni marginali, la loro invarianza rispetto a trasformazioni monotone crescenti e la loro capacità di catturare la dipendenza di coda, le rendono strumenti preziosi in molte applicazioni pratiche.

Applicazioni Finanziarie dei Modelli di Copula

Superiorità delle Copule sulla Correlazione Lineare con Distribuzioni Non Normali

L'assunzione di normalità dei rendimenti, tipico di modelli come quello di Black-Scholes, è spesso disatteso nei mercati finanziari. Le serie storiche di strumenti come azioni ed obbligazioni dimostrano la presenza di code più pesanti rispetto a quanto previsto dalla distribuzione normale e la diffusione di prodotti derivati con payoff non lineari amplifica ulteriormente questo fenomeno.
In questo contesto, la **correlazione lineare**, misurata ad esempio con il coefficiente di Pearson, si dimostra uno strumento limitato. Essa cattura solo le dipendenze lineari tra le variabili, mentre le relazioni tra gli asset finanziari possono assumere forme ben più complesse. La correlazione lineare è efficace solo quando le variabili sono legate da relazioni lineari. Tuttavia, in presenza di legami non lineari, la correlazione lineare potrebbe essere fuorviante. Ad esempio, una variabile con distribuzione chi-quadrato è perfettamente correlata al suo quadrato, che ha distribuzione normale, ma la correlazione lineare non sarebbe in grado di rappresentare correttamente questa relazione.

Spieghiamo il significato del coefficente di Pearson.

Definizione: Il coefficiente di correlazione lineare, noto anche come correlazione di Pearson, è una misura della dipendenza lineare tra due variabili casuali che assumono valori reali e che hanno varianza finita. È definito come la covarianza delle due variabili divisa per il prodotto delle loro deviazioni standard. Formalmente, per due variabili casuali non degeneri X e Y appartenenti a L^2 , il coefficiente di correlazione lineare ρ_{XY} è:

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

Il coefficiente di correlazione di Pearson assume valori compresi tra -1 e +1, dove:

- +1 indica una perfetta correlazione lineare positiva: all'aumentare di una variabile, l'altra aumenta in modo perfettamente lineare.

- -1 indica una perfetta correlazione lineare negativa: all'aumentare di una variabile, l'altra diminuisce in modo perfettamente lineare.
- 0 indica l'assenza di correlazione lineare: non c'è una relazione lineare tra le due variabili.

È importante sottolineare che il coefficiente di correlazione di Pearson misura solo la dipendenza lineare. Due variabili possono essere fortemente dipendenti in modo non lineare e avere comunque un coefficiente di correlazione di Pearson pari a zero.

Le **copule**, invece, offrono un approccio più flessibile per modellare la dipendenza tra variabili casuali, anche in presenza di distribuzioni non normali. Il vantaggio principale delle copule risiede nella loro capacità di separare la struttura di dipendenza dalle distribuzioni marginali. Questo permette di combinare diverse distribuzioni marginali, capaci di cogliere la non-normalità dei rendimenti (ad esempio la distribuzione t di Student o distribuzioni asimmetriche), con una vasta gamma di copule che descrivono la struttura di dipendenza.
Il Teorema di Sklar, fondamento della teoria delle copule, afferma che ogni funzione di distribuzione congiunta può essere espressa in termini di distribuzioni marginali e di una copula. Questa proprietà permette di costruire modelli di dipendenza altamente flessibili, adatti a rappresentare le complesse relazioni tra gli asset finanziari. Ad esempio, è possibile utilizzare una copula gaussiana per modellare la struttura di dipendenza, pur mantenendo distribuzioni marginali non gaussiane per i singoli asset. In questo modo, si ottiene un modello in grado di catturare sia la non-normalità dei rendimenti che la struttura di dipendenza tra gli stessi.
In definitiva, le copule rappresentano uno strumento più completo e affidabile rispetto alla correlazione lineare per modellare le dipendenze tra variabili casuali, soprattutto in presenza di distribuzioni non normali. La loro flessibilità e capacità di rappresentare accuratamente le complesse relazioni tra gli asset le rendono essenziali per una corretta valutazione del rischio, un pricing accurato e una migliore comprensione delle dinamiche dei mercati finanziari.

Dipendenze di Coda e la loro Importanza nel VaR e nell'Expected Shortfall

La dipendenza di coda si riferisce alla tendenza di due o più variabili casuali a muoversi insieme in modo più estremo nelle code delle loro distribuzioni, rispetto a quanto previsto da una distribuzione normale con la stessa correlazione lineare. In altre parole, la dipendenza di coda misura la probabilità che si verifichino eventi estremi congiuntamente.

Sappiamo che la non normalità a livello univariato è associata al cosiddetto problema della *fat-tail*. In un contesto multivariato, il problema della *fat-tail* può essere riferito sia alle distribuzioni marginali univariate che alle distribuzioni congiunte di probabilità di grandi movimenti di mercato. Questo concetto è chiamato *tail dependence*. L'uso di funzioni copula ci permette di modellare separatamente queste due caratteristiche. Per rappresentare la dipendenza dalla coda consideriamo la probabilità che un evento con probabilità inferiore a v si verifichi nella prima variabile, dato che un evento con probabilità inferiore a v si verifica nella seconda. In concreto, ci chiediamo quale sia la probabilità di osservare, ad esempio, un crollo con probabilità inferiore di v=1% nell'indice Nikkei 225, dato che nell'indice S&P 500 si è verificato un crollo con probabilità inferiore all'1%. Si ha

$$\begin{aligned}\lambda(v) &\equiv \Pr(Q_{NKY} \leq v \mid Q_{SP} \leq v) \\ &= \frac{\Pr(Q_{NKY} \leq v, Q_{SP} \leq v)}{\Pr(Q_{SP} \leq v)} \\ &= \frac{C(v,v)}{v}\end{aligned}$$

Tipi di dipendenza di coda:

Dopo che abbiamo calcolato il nostro $\lambda(v)$, possiamo dividere la dipendenza di coda in due tipi principali:

- **Dipendenza di coda inferiore:** misura la probabilità che entrambe le variabili assumano valori estremamente bassi contemporaneamente.

$$\lambda_L \equiv \lim_{v\to 0^+} \frac{C(v,v)}{v}$$

- **Dipendenza di coda superiore:** misura la probabilità che entrambe le variabili assumano valori estremamente alti contemporaneamente.

$$\begin{aligned}\lambda_U = \lim_{v\to 1^-} \lambda_v &\equiv \lim_{v\to 1^-} \frac{\Pr(\overline{Q}_{NKY} > v, \overline{Q}_{SP} > v)}{\Pr(\overline{Q}_{SP} > v)} \\ &= \lim_{v\to 1^-} \frac{1-2v+C(v,v)}{1-v}\end{aligned}$$

Importanza nel VaR e nell'Expected Shortfall:

La dipendenza di coda ha un impatto significativo sul calcolo del VaR e dell'Expected Shortfall, due misure di rischio ampiamente utilizzate nella gestione del rischio finanziario.

- **VaR (Value at Risk):** rappresenta la perdita massima stimata che un portafoglio potrebbe subire in un determinato orizzonte temporale e con un dato livello di confidenza.
- **Expected Shortfall:** rappresenta la perdita media attesa in caso di superamento del VaR.

In presenza di dipendenza di coda, il VaR e l'Expected Shortfall calcolati assumendo una distribuzione normale tendono a sottostimare il rischio effettivo del portafoglio. Questo perché la distribuzione normale non riesce a catturare adeguatamente la probabilità di eventi estremi congiunti. Mentre utilizzando le copule per modellare la dipendenza tra gli asset di un portafoglio, è possibile ottenere una stima più accurata del VaR e dell'Expected Shortfall, tenendo conto della probabilità di eventi estremi congiunti. Quindi grazie alle copule si ottiene una stima più accurata del rischio di portafoglio e si possono prendere decisioni di investimento più consapevoli.

Tariffare le opzioni

La valutazione di opzioni multivariate, come le opzioni basket o rainbow, che dipendono da più attività sottostanti, rappresenta una sfida significativa in finanza. Le copule forniscono un potente strumento per affrontare questo problema.

- In sostanza, una copula viene utilizzata per costruire la distribuzione congiunta dei prezzi delle attività sottostanti alla scadenza.
- Questa distribuzione viene quindi utilizzata per calcolare il valore atteso del payoff dell'opzione in base a tutti i possibili risultati dei prezzi delle attività sottostanti.
- Attualizzando questo valore atteso al tasso privo di rischio, si ottiene il prezzo dell'opzione.

Nei mercati incompleti, dove non esiste una misura di probabilità unica priva di arbitraggio, le copule sono fondamentali per derivare strategie di super-replicazione. Queste strategie mirano a creare un portafoglio di attività negoziabili che replichi il payoff dell'opzione in tutte le possibili situazioni future, garantendo così l'assenza di opportunità di arbitraggio. Le copule consentono di determinare i limiti superiori e inferiori del prezzo dell'opzione in base alle diverse ipotesi sulla struttura di dipendenza tra le attività sottostanti. Mostriamo ora degli esempi:

- **Opzioni arcobaleno:** queste opzioni, che dipendono dal minimo o dal massimo di un paniere di attività, possono essere valutate utilizzando le copule per catturare la dipendenza tra i rendimenti delle attività. Le fonti dimostrano come le copule possano essere utilizzate per derivare strategie di super-replicazione per le opzioni arcobaleno, fornendo limiti superiori e inferiori al prezzo.

- **Opzioni barriera:** per le opzioni in cui l'esercizio è condizionato al fatto che il prezzo dell'attività sottostante raggiunga o meno una determinata barriera, le copule possono essere utilizzate per modellare la dipendenza tra il percorso del prezzo dell'attività e l'evento di attivazione della barriera.

Gestione dei rischi

Le copule possono essere utilizzate per modellare la dipendenza tra diversi tipi di rischio, come rischio di mercato, rischio di credito e rischio operativo. Ciò è particolarmente utile per le istituzioni finanziarie che sono esposte a più tipi di rischio, in quanto consente loro di valutare il rischio complessivo a cui sono esposte. Ad esempio, una banca può utilizzare le copule per modellare la dipendenza tra le insolvenze sui prestiti e i movimenti dei tassi di interesse, consentendo loro di valutare il rischio del proprio portafoglio prestiti in diversi scenari economici.
In particolare, nella gestione del rischio di credito, le copule vengono utilizzate nella valutazione di strumenti di debito strutturati come le obbligazioni garantite da crediti (CDO). Le CDO sono obbligazioni garantite da un pool di attività sottostanti, come mutui o prestiti alle imprese. Il rischio di credito di una CDO dipende dalla dipendenza tra le insolvenze delle attività sottostanti. Le copule forniscono un modo flessibile per modellare questa dipendenza, consentendo agli investitori di valutare il rischio e il rendimento delle CDO in modo più accurato.

Tipi di modelli di copula

- Copula gaussiana: descrive la dipendenza tra variabili casuali utilizzando la distribuzione normale multivariata. Non è in grado di catturare la dipendenza dalla coda. È definita come la funzione di distribuzione congiunta di un vettore normale multivariato standard, dove ogni variabile marginale è stata trasformata nella sua forma uniforme standard utilizzando la funzione di distribuzione normale standard inversa.
$$C_R^{Ga}(u,v) = \Phi_R(\Phi^{-1}(u), \Phi^{-1}(v))$$
$$= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho_{XY}^2}} \exp\left(\frac{2\rho_{XY} st - s^2 - t^2}{2(1-\rho_{XY}^2)}\right) ds\, dt$$

- Copula t di Student: Questa copula può catturare la dipendenza dalla coda e viene spesso utilizzata per modellare i rendimenti degli asset che mostrano code pesanti. La copula t di Student bivariata è data dalla seguente formula:
$C_v^t(u,v) = \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(v)} \frac{\Gamma(\frac{v+2}{2})}{\Gamma(\frac{v}{2})\pi v\sqrt{1-\rho^2}} (1 + \frac{x^2 - 2\rho xy + y^2}{v(1-\rho^2)})^{-\frac{v+2}{2}} dxdy$ dove t_v^{-1} è l'inversa della funzione di distribuzione t di Student univariata con v gradi di libertà e ρ è il coefficiente di correlazione.

- Copula di Clayton: questa copula mostra una forte dipendenza dalla coda inferiore, il che significa che le variabili hanno maggiori probabilità di assumere insieme valori estremi bassi.È esaustiva e fornisce la copula del prodotto se $\alpha = 0$, il limite inferiore di Frechet $max(v+z-1, 0)$ quando $\alpha = -1$ e quello superiore per $\alpha \to +\infty$. È definita dalla seguente formula: $C(v,z) = max[(v^{-\alpha} + z^{-\alpha} - 1)^{-\frac{1}{\alpha}}, 0]$.

- Copula di Gumbel: questa copula mostra una forte dipendenza dalla coda superiore, indicando che le variabili hanno maggiori probabilità di assumere insieme valori estremi elevati. Fornisce la copula del prodotto se $\alpha = 1$ e il limite superiore di Frechet $min(v,z)$ per $\alpha \to +\infty$. È definita dalla seguente formula: $C(v,z) = exp{-}[(-\ln v)^{\alpha} + (-\ln z)^{\alpha}]^{\frac{1}{\alpha}}$.

- Copula di Frank: questa copula è una copula simmetrica che può catturare sia la dipendenza dalla coda superiore che quella dalla coda inferiore. Si riduce alla copula del prodotto se $\alpha = 0$ e raggiunge i limiti inferiore e superiore di Frechet rispettivamente per $\alpha \to -\infty$ e $\alpha \to +\infty$. È definita dalla seguente formula: $C(v,z) = -\frac{1}{\alpha}\ln(\frac{1+(e^{-\alpha v}-1)(e^{-\alpha z}-1)}{e^{-\alpha}-1})$.