# Mục lục

# Chương 1

# Kiến thức chuẩn bị

## 1.1 Copula cho phân phối nhiều chiều và sự phụ thuộc

### 1.1.1 Giới thiệu về copula

**Định nghĩa 1.1.1.** Giả sử $S_1, \ldots, S_n$ là các tập con không rỗng của $\overline{\mathbb{R}}$, ở đây $\overline{\mathbb{R}}$ được ký hiệu là đường thẳng thực mở rộng $[-\infty, +\infty]$. Giả sử $H$ là một hàm thực với $n$ biến trên miền xác định $DomH = S_1 \times \cdots \times S_n$ và cho $\mathbf{a} \leq \mathbf{b}$ (trong đó $\mathbf{a} = (a_1, \ldots, a_n)$, $\mathbf{b} = (b_1, \ldots, b_n)$ và $a_k \leq b_k$ với mọi $k = \overline{1, n}$), giả sử $B = [\mathbf{a}, \mathbf{b}] = [a_1, b_1] \times \cdots \times [a_n, b_n]$ là một $n$-hộp có các đỉnh trong $DomH$. Ta định nghĩa thể tích-$H$ của $B$ là:

$$V_H(B) = \sum sgn(c) H(c)$$

trong đó $c$ chạy trên các đỉnh của B, và $sgn(c)$ là dương nếu trong các tọa độ của nó có một số chẵn các $c_k = a_k$ với mọi $k$, và là âm trong trường hợp ngược lại.

Nói cách khác, thể tích-$H$ của một n-hộp $B = [\mathbf{a}, \mathbf{b}]$ là:

$$V_H(B) = \Delta_{\mathbf{a}}^{\mathbf{b}} H(t) = \Delta_{a_n}^{b_n} \ldots \Delta_{a_1}^{b_1} H(t)$$

trong đó

$$\Delta_{a_k}^{b_k} H(t) = H(t_1, \ldots, t_{k-1}, b_k, t_{k+1}, \ldots, t_n) - H(t_1, \ldots, t_{k-1}, a_k, t_{k+1}, \ldots, t_n).$$

Chẳng hạn, nếu $H(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n)$ là hàm phân phối xác suất đồng thời của một bộ $n$ biến ngẫu nhiên $X_1, \ldots, X_n$ thì ta có:

$$V_H(B) = P(a_1 \leq X_1 \leq b_1, \ldots, a_n \leq X_n \leq b_n).$$

**Định nghĩa 1.1.2.** Hàm thực $H$ của $n$ biến được gọi là $n$-tăng nếu $V_H(B) \geq 0$ cho mọi $n$-hộp $B$ có các đỉnh nằm trong miền xác định $DomH$ hay có thể nói thể tích của nó trên hộp bất kỳ là không âm.

Giả sử miền xác định của một hàm thực $H$ với $n$ biến xác định bởi $DomH = S_1 \times \cdots \times S_n$, ở đó mỗi $S_k$ có một phần tử nhỏ nhất $a_k$. Chúng ta nói rằng $H$ có đáy (ground) nếu $H(t) = 0$ với mọi $t$ trong $DomH$ sao cho $t_k = a_k$ tại số $k$ bé nhất. Nếu mỗi $S_k$ khác rỗng và có phần tử lớn nhất $b_k$, thì $H$ có phân phối biên duyên và phân phối biên duyên một chiều của $H$ là hàm $H_k$ với $DomH_k = S_k$ và với $H_k(x) = H(b_1, \ldots, b_{k-1}, x, b_{k+1}, \ldots, b_n)$ với $x \in S_k$, phân phối biên duyên một-chiều gọi là phân phối biên duyên.

**Định lý 1.1.1.** Giả sử $S_1, \ldots, S_k$ là các tập khác rỗng của $\overline{\mathbb{R}}$, và giả sử hàm $H$ có đáy $n$-tăng với miền xác định $S_1 \times \cdots \times S_n$. Thì $H$ là tăng theo mỗi đối số, tức là nếu $(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_n)$ và $(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_n)$ nằm trong miền xác định $DomH$ và $x \leq y$, thì $H(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_n) \leq H(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_n)$.

**Định lý 1.1.2.** Giả sử $S_1, \ldots, S_k$ là các tập khác rỗng của $\overline{\mathbb{R}}$, và giả sử hàm $H$ có đáy $n$-tăng với phân phối biên duyên và miền xác định $S_1 \times \cdots \times S_n$. Khi đó, nếu bất kỳ điểm $x = (x_1, \ldots, x_n)$ và $y = (y_1, \ldots, y_n)$ trong $S_1 \times \ldots \times S_n$ thì:

$$|H(x) - H(y)| \leq \sum_{k=1}^{n} |H_k(x_k) - H_k(y_k)|.$$

**Định nghĩa 1.1.3.** Một hàm phân phối $n$-chiều là một hàm $H$ với miền xác định $\overline{\mathbb{R}}^n$ sao cho $H$ có đáy, n-tăng và $H(+\infty, \ldots, +\infty) = 1$.

### 1.1.2 Một vài định nghĩa và tính chất của copula

Các copula $N$ chiều là các hàm $N$ biến từ $[0,1]^N$ vào $[0,1]$, và hàm copula là một hàm phân phối nhiều chiều (multivariate distribution function) xác định trên hình lập phương đơn vị $\mathbf{I}^N = [0,1]^N$, nó thể hiện sự phụ thuộc vào nhau của một bộ $N$ biến ngẫu nhiên. Các copula là các hàm đặc biệt với nhiều tính chất thú vị, và khi ta biết copula thì cũng có thể tính toán được sự phụ thuộc của các biến ngẫu nhiên trên hiệp phương sai (covariance) và sự tương quan (correlation).

**Định nghĩa 1.1.4.** (Nelsen (1998), trang 39) Một Copula $N$-chiều là một hàm $C$ với những tính chất sau:

1. Miền xác định (domain) của hàm $C$: $DomC = \mathbf{I}^N = [0,1]^N$;
2. Hàm $C$ có đáy (ground) và hàm $N$-tăng;
3. Hàm $C$ có các biên duyên $C_n$, thỏa mãn $C_n(u) = C(1, \ldots, 1, u, 1, \ldots, 1) = u, \forall u \in \mathbf{I}$.

Chú ý rằng bất kỳ $N$-copula $C$, $N \geq 3$, mỗi phân phối biên duyên $k$-chiều của $C$ là $k$-copula. Nói một cách khác, một $N$-copula là một hàm $C$ từ $[0,1]^N$ vào $[0,1]$ với những tính chất sau:

1. Mỗi $u \in [0,1]^N$, $C(u) = 0$ nếu một tọa độ nhỏ nhất của $u$ là 0 và $C(u) = u_k$ nếu tất cả các tọa độ của $u$ bằng 1 trừ ra $u_k$.

2. Mỗi $a$ và $b$ trong $[0,1]^N$ sao cho $a_i \leq b_i$ với mọi $i$, $V_C([a,b]) \geq 0$.

**Định lý 1.1.3.** (định lý Sklar [1959]) Cho $F$ là một hàm phân phối $N$-chiều với phân phối biên duyên liên tục $F_1, \dots, F_N$. Khi đó $F$ có một biểu diễn copula duy nhất:

$$F(x_1, \dots, x_n, \dots, x_N) = C(F_1(x_1), \dots, F_n(x_n), \dots, F_N(x_N)). \tag{1.1}$$

**Định lý 1.1.4.** Giả sử $H$ là hàm phân phối $n$-chiều với phân phối biên duyên liên tục $F_1, \dots, F_n$ và copula $C$, (ở đây $C$ thỏa mãn điều kiện (1.2) ). Khi đó bất kỳ $u$ trong $[0,1]^n$

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

**Ví dụ 1.1.1.** Chúng ta sử dụng dữ liệu của thị trường trao đổi kim loại London: LME (London Metal Exchange) và chúng ta xét giá hiện có của hàng hóa hợp kim nhôm (Al); đồng (Cu); ni-ken (Ni); chì (Pb) và giá của hợp kim nhôm 15 tháng trước (Al - 15), vào tháng 1 năm 1988. Chúng ta giả thiết rằng phân phối lợi suất của các tài sản là phân phối gauss. Trong trường hợp này, ma trận tương quan được cho bởi bảng 1.1.

| | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| Al | 1.00 | 0.82 | 0.44 | 0.36 | 0.33 |
| Al-15 | | 1.00 | 0.39 | 0.34 | 0.30 |
| Cu | | | 1.00 | 0.37 | 0.31 |
| Ni | | | | 1.00 | 0.31 |
| Pb | | | | | 1.00 |

Bảng 1.1: Ma trận tương quan $\rho$ của dữ liệu LME

### 1.1.3 Các hàm phân phối đồng thời Fréchet-Hoeffding

Chúng ta xét các hàm $C^-$ , $C^+$ và $C^\perp$ xác định trên $[0,1]^N$ như sau:

$$C^-(u_1, \dots, u_n, \dots, u_N) = \max(\sum_{n=1}^{N} u_n - N + 1, 0) \tag{1.2}$$

$$C^+(u_1, \dots, u_n, \dots, u_N) = \min(u_1, \dots, u_n, \dots, u_N) \tag{1.3}$$

$$C^\perp(u_1, \dots, u_n, \dots, u_N) = \prod_{n=1}^{N} u_n \tag{1.4}$$

Hàm phân phối đồng thời $C^-$, $C^+$ gọi là các cận dưới và cận trên Fréchet-Hoeffding, còn hàm $C^\perp$ gọi là tích copula. Hàm $C^+$, $C^\perp$ là $N$- copula với mọi $N \geq 2$, trong khi hàm $C^-$ không là copula cho bất kỳ $N \geq 3$, xét ví dụ:

**Định nghĩa 1.1.5.** Chúng ta nói rằng copula $C_1$ là nhỏ hơn copula $C_2$ (hoặc $C_2$ là lớn hơn $C_1$) và viết là $C_1 \prec C_2$ (hoặc $C_2 \succ C_1$) nếu:

$$\forall (u_1, \ldots, u_n, \ldots, u_N) \in \mathbf{I}^N, \; C_1(u_1, \ldots, u_n, \ldots, u_N) \leq C_2(u_1, \ldots, C_n, \ldots, C_N) \quad (1.5)$$

hay

$$\forall (u_1, \ldots, u_n, \ldots, u_N) \in \mathbf{I}^N, \; \overline{C}_1(u_1, \ldots, u_n, \ldots, u_N) \leq \overline{C}_2(u_1, \ldots, u_n, \ldots, u_N)$$

Chẳng hạn trường hợp hai chiều:

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

Chúng ta ký hiệu $\overline{C}$ là hàm sống sót đồng thời (joint survival function) cho $n$ biến ngẫu nhiên với hàm phân phối đồng thời $C$, tức là nếu $(U_1, \ldots, U_n)^T$ có hàm phân phối $C$, thì $\overline{C}(u_1, \ldots, u_n) = \mathbf{P}\{U_1 > u_1, \ldots, U_n > u_n\}$.

**Định lý 1.1.5.** Bất kỳ $n \geq 3$ và mọi $u \in [0,1]^n$ có $n$-copula $C$ (mà phụ thuộc trên $u$) thì:

$$C(u) = C^-(u)$$

Hàm mật độ $c$ liên kết với copula được xác định bởi :

$$c(u_1, \ldots, u_n, \ldots, u_N) = \frac{\partial C(u_1, \ldots, u_n, \ldots, u_N)}{\partial u_1 \ldots \partial u_n \ldots \partial u_N}.$$

Để đạt được hàm mật độ $f$ của phân phối $N$-chiều $F$, chúng ta sử dụng hệ thức sau:

$$f(x_1, \ldots, x_n, \ldots, x_N) = c(F_1(x_1), \ldots, F_n(x_n), \ldots, F_n(x_N)) \prod_{n=1}^{N} f_n(x_n)$$

trong đó $f_n$ là hàm mật độ của hàm phân phối biên duyên $F_n$.

## 1.1.4 Copula và biến ngẫu nhiên

Giả sử $X_1, \ldots, X_n$ tương ứng là các biến ngẫu nhiên với các hàm phân phối liên tục $F_1, \ldots, F_n$, và hàm phân phối đồng thời $H$. Thì $(X_1, \ldots, X_n)^T$ có duy nhất một copula $C$, ở đây $C$ xác định bởi (1.2). Copula biểu diễn phân phối của véctơ ngẫu nhiên $(X_1, \ldots, X_n)^T$ trở thành:

$$H(x_1, \ldots, x_n) = \mathbf{P}\{X_1 \leq x_1, \ldots, X_n \leq x_n\} = C(F_1(x_1), \ldots, F_n(x_n))$$

Phép biến đổi $X_i \mapsto F(X_i)$ sử dụng trong biểu diễn trên thường là phép biến đổi xác suất và tạo thành công cụ chuẩn trong phương pháp mô phỏng.

Hơn nữa $X_1, \ldots, X_n$ là độc lập nếu và chỉ nếu $H(x_1, \ldots, x_n) = F_1(x_1) \ldots F_n(x_n)$ với mọi $x_1, \ldots, x_n$ trong $\overline{\mathbb{R}}$.

**Định lý 1.1.6.** Giả sử $(X_1, \ldots, X_n)^T$ là một véctơ của biến ngẫu nhiên liên tục với copula $C$, khi đó $X_1, \ldots, X_n$ là độc lập nếu và chỉ nếu $C = C^{\perp}$.

Một tính chất đẹp của copula chính là phép biến đổi đơn điệu ngặt của biến ngẫu nhiên, copula cũng là bất biến. Nếu hàm phân phối của biến ngẫu nhiên $X$ là liên tục và $\alpha$ là một hàm đơn điệu ngặt mà miền xác định chứa miền giá trị $RanX$ (miền giá trị :*Range*), thì hàm phân phối của biến ngẫu nhiên $\alpha(X)$ cũng là liên tục.

**Định lý 1.1.7.** Giả sử $(X_1, \ldots, X_n)^T$ là một véctơ của biến ngẫu nhiên liên tục với copula $C_{X_1,\ldots,X_n}$. Giả sử $\alpha_1, \ldots, \alpha_n$ tương ứng đơn điệu ngặt trên $RanX_1, \ldots, RanX_n$, và giả sử $(\alpha_1(X_1), \ldots, \alpha_n(X_n))^T$ có copula $C_{\alpha_1(X_1),\ldots,\alpha_n(X_n)}$. Hơn nữa $\alpha_k$ giảm ngặt cho một vài $k$ (không mất tính tổng quát giả sử $k = 1$). Khi đó:

$$\begin{aligned} C_{\alpha_1(X_1),\ldots,\alpha_n(X_n)}(u_1, u_2, \ldots, u_n) = {} & C_{\alpha_2(X_2),\ldots,\alpha_n(X_n)}(u_2, \ldots, u_n) \\ & - C_{X_1,\alpha_2(X_2),\ldots,\alpha_n(X_n)}(1 - u_1, u_2, \ldots, u_n). \end{aligned}$$

Chú ý rằng hàm sống sót đồng thời với $n$ biến ngẫu nhiên $U(0,1)$ có hàm phân phối đồng thời là copula $C$ thì $\overline{C}(u_1, \ldots, u_n) = \hat{C}(1 - u_1, \ldots, 1 - u_n)$.

## 1.2 Các khái niệm sự phụ thuộc

### 1.2.1 Độ đo sự tương thích

Giả sử $(x, y)^T$ và $(\tilde{x}, \tilde{y})^T$ là hai quan sát từ một véctơ ngẫu nhiên liên tục $(X, Y)^T$. Khi đó $(x, y)^T$ và $(\tilde{x}, \tilde{y})^T$ là tương thích nếu $(x - \tilde{x})(y - \tilde{y}) > 0$, và không tương thích nếu $(x - \tilde{x})(y - \tilde{y}) < 0$.

**Định lý 1.2.1.** Giả sử $(X, Y)^T$ và $(\tilde{X}, \tilde{Y})^T$ là các véc tơ độc lập của các biến ngẫu nhiên liên tục với hàm phân phối đồng thời $H$ và $\tilde{H}$ với biên duyên chung $F$ (của $X$ và $\tilde{X}$) và $G$ (của $Y$ và $\tilde{Y}$). Giả sử $C$ và $\tilde{C}$ được ký hiệu là copula của $(X, Y)^T$ và $(\tilde{X}, \tilde{Y})^T$, tương ứng, sao cho $H(x, y) = C(F(x), G(y))$ và $\tilde{H}(x, y) = \tilde{C}(F(x), G(y))$, $Q$ ký hiệu là hiệu số giữa hai xác suất sự tương thích và sự không tương thích của $(X, Y)^T$ và $(\tilde{X}, \tilde{Y})^T$, tức là :

$$Q = \mathbf{P}\{(X - \tilde{X})(Y - \tilde{Y}) > 0\} - \mathbf{P}\{(X - \tilde{X})(Y - \tilde{Y}) < 0\}$$

thì

$$Q = Q(C, \tilde{C}) = 4 \iint_{[0,1]^2} \tilde{C}(u, v) dC(u, v) - 1.$$

**Định nghĩa 1.2.1.** (Nelsen (1998), trang 136) Một độ đo $\kappa$ liên kết giữa hai biến ngẫu nhiên liên tục $X_1$ và $X_2$ của copula $C$ là một độ đo sự tương thích nếu nó thỏa mãn các tính chất sau:

1, $\kappa$ được định nghĩa cho mọi cặp $X_1, X_2$ của biến ngẫu nhiên liên tục;
2, $-1 = \kappa_{X,-X} \le \kappa_C \le \kappa_{X,X} = 1$;
3, $\kappa_{X_1,X_2} = \kappa_{X_2,X_1}$;
4, Nếu $X_1, X_2$ là độc lập, thì $\kappa_{X_1,X_2} = \kappa_{C^{\perp}} = 0$;

5, $\kappa_{-X_1,X_2} = \kappa_{X_2,-X_1} = -\kappa_{X_1,X_2}$;
6, Nếu $C_1 \prec C_2$, thì $\kappa_{C_1} \leq \kappa_{C_2}$;
7, Nếu $\{\ (X_{1,n}; X_{2,n}\ )\}$ là một dãy biến ngẫu nhiên liên tục với copula $C_n$, và nếu $\{C_n\}$ hội tụ theo từng điểm tới $C$, thì $\lim_{n\to\infty} \kappa_{C_n} = \kappa_C$.

**Chú ý 1.2.1.** Một tính chất quan trọng khác của $\kappa$ trở thành hàm copula của các biến ngẫu nhiên $(X_1, ..., X_n, ..., X_N)$ là bất biến dưới, tăng thực sự qua phép ánh xạ

$$C_{X_1,...,X_n,...,X_N} = C_{h_1(X_1),...,h_n(X_n),...,h_N(X_N)} \text{ nếu } \partial_x h_n(x) > 0. \tag{1.6}$$

Chúng ta đưa ra một vài hệ thức giữa độ đo $\tau$ và $\rho$ (Nelsen [1998]), nó có thể được biểu diễn bằng một miền bị chặn. Chẳng hạn, nếu không phân tích được biểu thức, ta sẽ tính toán với công thức tương đương:

$$\tau = 1 - 4 \iint_{\mathbf{I}^2} \partial_{u_1} C(u_1, u_2) \partial_{u_2} C(u_1, u_2) du_1 du_2 \tag{1.7}$$

$$\rho = 12 \iint_{\mathbf{I}^2} C(u_1, u_2) du_1 du_2 - 3 \tag{1.8}$$

$$\gamma = 4 \int_{\mathbf{I}} (C(u, u) + C(u, 1 - u) - u) du \tag{1.9}$$

### 1.2.2 Độ đo sự phụ thuộc

**Định nghĩa 1.2.2.** (Nelsen (1998), trang [170]) Một độ đo $\delta$ liên kết giữa hai biến ngẫu nhiên liên tục $X_1$ và $X_2$ của copula là độ đo sự phụ thuộc nếu nó thỏa mãn các tính chất sau:
1, $\delta$ được định nghĩa cho mọi cặp $X_1$, $X_2$ của biến ngẫu nhiên liên tục;
2. $0 = \delta_{C^\perp} \leq \delta_C \leq \delta_{C^+} = 1$;
3, $\delta_{X_1,X_2} = \delta_{X_2,X_1}$;
4, $\delta_{X_1,X_2} = \delta_{C^\perp} = 0$ nếu và chỉ nếu $X_1$ và $X_2$ độc lập;
5, $\delta_{X_1,X_2} = \delta_{C^+} = 1$ nếu và chỉ nếu $X_1$ và $X_2$ là hàm đơn điệu ngặt hầu chắc chắn của các biến khác;
6, Nếu $h_1$ và $h_2$ là hàm đơn điệu ngặt hầu chắc chắn trên $ImX_1$ và $ImX_2$ thì :

$$\delta_{h_1(X_1).h_2(X_2)} = \delta_{X_1,X_2};$$

7, Nếu $\{(X_{1,n}, X_{2,n})\}$ là một dãy biến ngẫu nhiên liên tục bởi copula $C_n$ hội tụ theo từng điểm đến $C$, thì $\lim_{n\to\infty} \delta_{C_n} = \delta_C$.

### 1.2.3 Những khái niệm phụ thuộc khác

Có nhiều khái niệm phụ thuộc khác nhau, chúng được sử dụng nhiều trong tài chính. Chẳng hạn, $X_1$ và $X_2$ phụ thuộc góc phần tư dương (positive quadrant dependent: PQD) nếu:

$$\mathbf{P}\{X_1 > x_1, X_2 > x_2\} \geq \mathbf{P}\{X_1 > x_1\}\mathbf{P}\{X_2 > x_2\}. \tag{1.10}$$

Giả sử rằng $X_1$ và $X_2$ là hai biến ngẫu nhiên thay cho hai tổn thất tài chính. Xác suất đồng thời với tổn thất lớn thì biến ngẫu nhiên phụ thuộc lớn hơn biến ngẫu nhiên độc lập một lần.

$$C \succ C^{\perp}. \tag{1.11}$$

Ta đưa ra định nghĩa sự phụ thuộc tiệm cận đuôi như sau:

**Định nghĩa 1.2.3.** Nếu cho một hàm copula $C$ hai chiều cho bởi:

$$\lim_{u \to 1} \frac{\overline{C}(u,u)}{1-u} = \lambda \tag{1.12}$$

thì $C$ có sự phụ thuộc tiệm cận đuôi trên với $\lambda \in (0,1]$ và không phụ thuộc tiệm cận đuôi trên với $\lambda = 0$, $\lambda$ gọi là hệ số phụ thuộc tiệm cân đuôi trên.

Ta nhớ lại rằng hàm copula sống sót $\hat{C}$ của hai biến ngẫu nhiên với copula $C$ xác định bởi:

$$\hat{C}(u_1,u_2) = u_1 + u_2 - 1 + C(1-u_1, 1-u_2) \tag{1.13}$$

và hàm sống sót đồng thời $\overline{C}$ cho hai biến ngẫu nhiên $U(0,1)$ có hàm phân phối đồng thời là $C$ cho bởi:

$$\overline{C}(u_1,u_2) = 1 - u_1 - u_2 + C(u_1,u_2) \tag{1.14}$$

khi đó:

$$\overline{C}(u_1,u_2) = \hat{C}(1-u_1, 1-u_2). \tag{1.15}$$

Hơn nữa:

$$\lim_{u \to 1} \frac{\overline{C}(u,u)}{1-u} = \lim_{u \to 1} \frac{\hat{C}(1-u,1-u)}{1-u} = \lim_{u \to 0} \frac{\hat{C}(u,u)}{u}.$$

## 1.3 Sơ lược về các hàm copula

### 1.3.1 Phân phối elliptic

**Định nghĩa 1.3.1.** Nếu $X$ là véc tơ ngẫu nhiên $n$-chiều, $\mu \in \mathbb{R}^n$ và ma trận đối xứng $\Sigma$ cấp $n \times n$ xác định không âm, hàm đặc trưng $\varphi_{X-\mu}(t)$ của $X - \mu$ là một hàm có dạng toàn phương $t^T \Sigma t$, $\varphi_{X-\mu}(t) = \phi(t^T \Sigma t)$, chúng ta nói rằng $X$ có phân phối elliptical với tham số $\mu, \Sigma$ và $\phi$ và chúng ta viết $X \sim E_n(\mu, \Sigma, \phi)$.

Khi $n = 1$, lớp các phân phối elliptic trùng với lớp các phân phối đối xứng 1-chiều. Hàm $\phi$ trong định nghĩa trên được gọi là hàm sinh đặc trưng.

**Định lý 1.3.1.** $X \sim E_n(\mu, \Sigma, \phi)$ với $rank(\Sigma) = k$ nếu và chỉ nếu tồn tại một biến ngẫu nhiên $R \geq 0$ độc lập với $U$, một véctơ ngẫu nhiên $k$-chiều phân phối đều trên siêu cầu đơn vị $\{z \in \mathbb{R}^k | z^T z = 1\}$, và một ma trận $A$ cỡ $n \times k$ với $AA^T = \Sigma$ sao cho

$$X =_d \mu + RAU.$$

**Ví dụ 1.3.1.** Giả sử véctơ ngẫu nhiên $n$-chiều $X \sim \mathcal{N}_n(0, \mathbf{I}_n)$, với thành phần $X_i \sim \mathcal{N}(0,1), i = 1, \ldots, n$, là độc lập và hàm đặc trưng của $X_i$ là $\exp(-t_i^2/2)$, hàm đặc trưng của $X$ sẽ là

$$\exp\{-\frac{1}{2}(t_1^2 + \cdots + t_n^2)\} = \exp\{-\frac{1}{2}t^T t\}.$$

Từ định lý trên thấy rằng $X \sim E_n(0, \mathbf{I}_n, \phi)$, ở đây $\phi(u) = \exp(-u/2)$.

Để tìm được biểu diễn sao cho $Cov(X) = \Sigma$, chúng ta sử dụng định lý trên thì thu được:

$$Cov(X) = Cov(\mu + RAU) = A\mathbf{E}(R^2)Cov(U)A^T$$

với điều kiện là $\mathbf{E}(R^2) < \infty$. Giả sử $Y \sim \mathcal{N}_n(0, \mathbf{I}_n)$. Khi đó $Y =_d ||Y||U$, ở đây$||Y||$ và $U$ là độc lập. Hơn nữa $||Y||^2 \sim \chi_n^2$, $\mathbf{E}(||Y||^2) = n$. Từ đó $Cov(Y) = \mathbf{I}_n$, chúng ta thấy rằng nếu $U$ là phân phối đều trên siêu cầu đơn vị trong $\mathbb{R}^n$, thì $Cov(U) = \mathbf{I}_n/n$, do đó $Cov(X) = AA^T\mathbf{E}(R^2)/n$. Chọn hàm sinh đặc trưng $\phi^*(s) = \phi(s/c)$, ở đây $c = \mathbf{E}(R^2)/n$, chúng ta có $Cov(X) = \Sigma$. Từ đó phân phối elliptic được mô tả đầy đủ bởi $\mu, \Sigma$ và $\phi$, ở đây $\phi$ có thể chọn sao cho $Cov(X) = \Sigma$ (nếu $Cov(X)$ được xác định ). Nếu $Cov(X)$ thu được như trên thì phân phối của $X$ xác định duy nhất bởi $\mathbf{E}(X)$, $Cov(X)$ và phân phối của nó là biên duyên một chiều, phân phối chuẩn hoặc phân phối student bậc 4.

**Định lý 1.3.2.** Giả sử $X \sim E_n(\mu, \Sigma, \phi)$, giả sử $B$ là một ma trận cỡ $q \times n$ và $\mathbf{b} \in \mathbb{R}^n$. Khi đó:

$$\mathbf{b} + BX \sim E_q(\mathbf{b} + B\mu, B\Sigma B^T, \phi).$$

Chứng minh: Từ định lý 1.3.1, $\mathbf{b} + BX$ có biểu diễn ngẫu nhiên:

$$\mathbf{b} + BX =_d \mathbf{b} + B\mu + RBAU.$$

Giả sử ta có sự phân hoạch $X, \mu$ và $\Sigma$ như sau:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

ở đây $X_1$ và $\mu_1$ là các véctơ cỡ $r \times 1$ và $\Sigma_{11}$ là một ma trận cỡ $r \times r$. Ta có định lý sau:

**Định lý 1.3.3.** Giả sử $X \sim E_n(\mu, \Sigma, \phi)$. Khi đó

$$X_1 \sim E_r(\mu_1, \Sigma_{11}, \phi), \quad X_2 \sim E_{n-r}(\mu_2, \Sigma_{22}, \phi).$$

Từ đó các phân phối biên duyên của phân phối elliptic là elliptic.

### 1.3.2 Copula liên quan đến phân phối elliptic

**Định nghĩa 1.3.2.** (Copula gauss nhiều chiều-MVN) Cho $\rho$ là một ma trận đối xứng, xác định dương với đường chéo chính $diag\rho = 1$ và hàm tiêu chuẩn của phân phối chuẩn nhiều chiều $\Phi_\rho$ với ma trận tương quan $\rho$. Copula gauss nhiều chiều được định nghĩa như sau:

$$C(u_1, \ldots, u_n, \ldots, u_N, \rho) = \Phi_\rho(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n), \ldots, \Phi^{-1}(u_N)). \tag{1.16}$$

Nhớ rằng chúng ta có phân phối chuẩn nhiều chiều là:

$$\frac{1}{(2\pi)^{\frac{N}{2}}|\rho|^{\frac{1}{2}}}\exp\Big(-\frac{1}{2}x^T\rho^{-1}x\Big) = c(\Phi^{-1}(x_1), \ldots, \Phi^{-1}(x_n), \ldots, \Phi^{-1}(x_N))\times \Big(\prod_{n=1}^{N}\frac{1}{\sqrt{2\pi}}\exp\big(-\frac{1}{2}x_n^2\big)\Big). \tag{1.17}$$

Chúng ta cũng suy ra:

$$c(u_1, \ldots, u_n, \ldots, u_N) = \frac{\frac{1}{(2\pi)^{\frac{N}{2}}|\rho|^{\frac{1}{2}}}\exp\big(-\frac{1}{2}\varsigma^T\rho^{-1}\varsigma\big)}{\frac{1}{(2\pi)^{\frac{N}{2}}}\exp\big(-\frac{1}{2}\varsigma^T\varsigma\big)}. \tag{1.18}$$

Áp dụng công thức trên thì mật độ tương ứng của công thức (1.26) là:

$$c(u_1, ..., u_n, ..., u_N, \rho) = \frac{1}{|\rho|^{\frac{1}{2}}}\exp\Big(-\frac{1}{2}\varsigma^T(\rho^{-1} - \mathbf{I})\Big) \tag{1.19}$$

trong đó $\varsigma_n = \Phi^{-1}(u_n)$.

Chúng ta xét việc tính toán đến mật độ có điều kiện. Giả sử ta ký hiệu một véctơ $U = [U_1^T, U_2^T]^T$ của biến ngẫu nhiên đều. Xét dạng phân hoạch, chúng ta có:

$$\rho = \begin{bmatrix}\rho_{11} & \rho_{12} \\ & \rho_{22}\end{bmatrix}. \tag{1.20}$$

Chúng ta cũng chứng tỏ rằng mật độ có điều kiện của $U_2$ cho giá trị của $U_1$ là

$$c(U_2|U_1; \rho) = \frac{1}{|\rho\_22 - \rho_{12}^T\rho_{11}^{-1}\rho_{12}|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\varsigma\Big([\rho_{22} - \rho_{12}^T\rho_{11}^{-1}\rho_{12}]^{-1} - \mathbf{I}\Big)\varsigma\right) \tag{1.21}$$

với $\rho = \Phi^{-1}(U_2) - \rho_{12}^T\rho_{11}^{-1}\Phi^{-1}(U_1)$. Sử dụng công thức này và nếu phân phối biên duyên được chỉ rõ, chúng ta có thể biểu diễn điểm hồi quy phân vị hoặc tính toán giá trị khác như là giá trị kỳ vọng.

**Định nghĩa 1.3.3.** (Copula student nhiều chiều-MVT) Cho $\rho$ là ma trận đối xứng, xác định dương với đường chéo chính $dig\rho = 1$ và hàm tiêu chuẩn của phân phối student nhiều chiều $T_{\rho,\nu}$ với bậc tự do $\nu$ và ma trận tương quan $\rho$. Khi đó copula student nhiều chiều được định nghĩa là:

$$C(u_1, ..., u_n, ..., u_N; \rho, \nu) = T_{\rho,\nu}\big(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_n), ..., t_\nu^{-1}(u_N)\big) \tag{1.22}$$

với nghịch đảo $t_\nu^{-1}$ là phân phối student một chiều. Hàm mật độ tương ứng sẽ là:

$$c(u_1, ..., u_n, ..., u_N; \rho) = |\rho|^{-\frac{1}{2}}\frac{\Gamma\big(\frac{\nu+N}{2}\big)\big[\Gamma\big(\frac{\nu}{2}\big)\big]^N(1 + \frac{1}{\nu}\varsigma^T\rho^{-1}\varsigma)^{-\frac{\nu+N}{2}}}{\big[\Gamma\big(\frac{\nu+1}{2}\big)\big]^N\Gamma\big(\frac{\nu}{2}\big)\prod_{n=1}^{N}\big(1 + \frac{\varsigma_n^2}{\nu}\big)^{-\frac{\nu+1}{2}}} \tag{1.23}$$

với $\varsigma_n = t_\nu^{-1}(u_n)$.

### 1.3.3 Copula Archimedean

Chúng ta định nghĩa copula Archimedean (Genest và Mackay [1986]) như sau:

$$C(u_1, ..., u_n, ..., u_N) = \begin{cases} \varphi^{-1}\big(\varphi(u_1) + ... + \varphi(u_n) + ... + \varphi(u_N)\big) & \text{nếu } \sum_{n=1}^{N} \varphi(u_n) \leq \varphi(0) \\ \\ 0 & \text{(ngược lại)} \end{cases} \tag{1.24}$$

với $\varphi(u)$ là hàm $\mathcal{C}^2; \varphi(1) = 0, \varphi'(u) < 0$ và $\varphi''(u) > 0$ với mọi $0 \leq u \leq 1$. $\varphi(u)$ gọi là hàm sinh của copula. Các copula Archimedean đóng vai trò quan trọng, bởi vì chúng đòi hỏi một vài tính chất ($C$ là đối xứng, kết hợp,...). Tuy nhiên copula Archimedean tính toán đơn giản.

**Ví dụ 1.3.2.** Xét độ đo sự tương thích $\tau$ của Kendall được cho bởi:

$$\tau = 1 + 4\int_0^1 \frac{\varphi(u)}{\varphi'(u)}\, du. \tag{1.25}$$

Chúng ta có kết quả trong bảng copula Archimedean hai chiều là:

| copula | $\varphi(u)$ | $C(u_1, u_2)$ |
|---|---|---|
| $C^{\perp}$ | $-\ln u$ | $u_1 u_2$ |
| Gumbel | $(-\ln u)^{\alpha}$ | $\exp\Big(-\big(\tilde{u}_1^{\alpha} + \tilde{u}_2^{\alpha}\big)^{\frac{1}{\alpha}}\Big)$ |
| Joe | $\big(-\ln 1 - (1-u)^{\alpha}\big)$ | $1 - \big(\tilde{u}_1^{\alpha} + \tilde{u}_2^{\alpha} - \tilde{u}_1^{\alpha}\tilde{u}_2^{\alpha}\big)^{\frac{1}{\alpha}}$ |
| Kimeldorf-Sampson | $u^{-\alpha} - 1$ | $\big(\tilde{u}_1^{-\alpha} + \tilde{u}_2^{-\alpha} - 1\big)^{-\frac{1}{\alpha}}$ |

Giả sử $\overline{w}(u) = \exp\big(-\varphi(u)\big)$, chúng ta chú ý rằng phương trình (1.39) được viết dưới dạng:

$$\overline{w}\big(C(u_1, \ldots, u_n, \ldots, u_N)\big) = \prod_{n=1}^{N} \overline{w}(u_n). \tag{1.26}$$

Áp dụng $\varphi$ cho cả hai phân phối đồng thời và phân phối biên duyên, các phân phối trở nên độc lập. Chúng ta cũng chú ý rằng các copula Archimedean thì liên quan đến sự tạo thành của nhiều phân phối nhiều chiều. Giả sử $\gamma$ là một véctơ chứa tham số tạo thành bởi một hàm phân phối đồng thời $\Gamma$ với phân phối biên duyên $\Gamma_n$ và $\mathbf{H}$ là phân phối nhiều chiều. Chúng ta ký hiệu $F_1, \ldots, F_N$ $N$-phân phối một chiều, ta chứng tỏ rằng:

$$F(x_1, ..., x_n, ..., x_N) = \int \cdots \int \mathbf{H}(H_1^{\gamma_1}(x_1), ..., H_n^{\gamma_n}(x_n), ..., H_N^{\gamma_N}(x_N)\, d\Gamma(\gamma_1, ..., \gamma_n, ..., \gamma_N)) \tag{1.27}$$

là một phân phối nhiều chiều với biên duyên $F_1, \ldots, F_N$, chúng ta có:

$$H_n(x_n) = \exp\big(-\psi_n^{-1}(F_n(x_n))\big) \tag{1.28}$$

với phép biến đổi Laplace $\psi_n$ của phân phối biên duyên $\Gamma_n$. Cho $\psi$ là phép biến đổi Laplace của phân phối đồng thời $\Gamma$, biểu thức (1.42) được viết một dạng khác:

$$F(x_1,...,x_n,...,x_N) = \psi\big(\psi_1^{-1}(F_1(x_1)),...,\psi_n^{-1}(F_n(x_n)),...,\psi_N^{-1}(F_N(x_N))\big). \quad (1.29)$$

Nếu các biên duyên $\Gamma_n$ là như nhau, $\Gamma$ là cận dưới Fréchet và $\mathbf{H}$ là tích các copula. Ta chú ý rằng:

$$F(x_1,...,x_n,...,x_N) = \psi\big(\psi^{-1}(F_1(x_1))+...+\psi^{-1}(F_n(x_n))+...+\psi^{-1}(F_N(x_N))\big). \quad (1.30)$$

Ánh xạ Laplace nghịch đảo $\psi^{-1}$ là một hàm sinh cho các copula Archimedean.

### 1.3.4 Giá trị cực trị các copula

Một giá trị cực trị của copula $C$ tại $(u_1,...,u_n,...,u_N)$ thỏa mãn hệ thức sau đây:

$$C(u_1^t,...,u_n^t,...,u_N^t) = C^t(u_1,...,u_n,...,u_N)\ \forall t > 0. \quad (1.31)$$

**Ví dụ 1.3.3.** Xét copula Gumbel là một giá trị cực trị copula, thật vậy:

$$\begin{aligned}
C(u_1^t,u_2^t) &= \exp\Big(-\big[(-\ln u_1^t)^\alpha + (-\ln u_2^t)^\alpha\big]^{\frac{1}{\alpha}}\Big) \\
&= \exp\Big(-\big(t^\alpha\big[(-\ln u_1^t)^\alpha + (-\ln u_2^t)^\alpha\big]\big)^{\frac{1}{\alpha}}\Big) \\
&= \Big[\exp\Big(-\big[(-\ln u_1)^\alpha + (-\ln u_2)^\alpha\big]^{\frac{1}{\alpha}}\Big)\Big]^t \\
&= C^t(u_1,u_2).
\end{aligned}$$

# Chương 2

# Các kết luận thống kê về copula

## 2.1 Kỹ thuật mô phỏng

Mô phỏng đóng một vai trò rất quan trọng trong những kết luận về thống kê. Đặc biệt chúng giúp cho việc nghiên cứu các tính chất về công thức ước lượng. Cho ví dụ sau, giả sử ta tạo thành một phân phối 5-chiều với một copula MVT, hai phân phối biên duyên suy rộng-pareto và ba phân phối biên duyên suy rộng-beta. Ta phải thực hiện mô phỏng khái niệm dạng phân phối.

Mô phỏng các biến ngẫu nhiên với phân phối đều cho copula C có thể được thực hiện với thuật toán tổng quát:

1. Tạo thành N biến ngẫu nhiên với phân phối đều độc lập $(v_1, ..., v_n, ..., v_N)$;

2. Tạo thành N biến ngẫu nhiên đệ quy:

$$u_n = C^{-1}_{(u_1,...,u_{n-1})}(v_n) \tag{2.1}$$

với

$$\begin{aligned} C_{(u_1,...,u_{n-1})}(u_n) &= P\{U_n \leq u_n | (U_1, ..., U_{n-1}) = (u_1, .., u_{n-1})\} \\ &= \frac{\partial^{n-1}_{(u_1,...,u_{n-1})}\mathbf{C}(u_1, ..., u_n, 1, ..., 1)}{\partial^{n-1}_{(u_1,...,u_{n-1})}\mathbf{C}(u_1, ..., u_{n-1}, 1, ..., 1)}. \end{aligned} \tag{2.2}$$

Ý tưởng chính của thuật toán này là mô phỏng mỗi $u_n$ bởi phân phối điều kiện của nó. Trong trường hợp 2-chiều, hệ thức (2.1) trở thành

$$\begin{aligned} u_1 &= v_1 \\ \partial_{u_1} C(u_1, u_2) &= v_2 \end{aligned}$$

bởi vì $\partial_{u_1} C(u_1, 1) = 1$. Tuy nhiên, chúng ta chú ý rằng biến ngẫu nhiên sinh ra cần một phương pháp giải nghiệm. Cho lớp các copula, tồn tại thuật toán đặc trưng mạnh. Đây là trường hợp của Kimeldorf-Sampson, Gumbel, Marshall-Olkin,..., và tổng quát nhất là trường hợp của copula Archimedean, chúng ta có:

$$C_{(u_1,...,u_{n-1})}(u_n) = \frac{\varphi_{(n-1)}(\varphi(u_1) + ... + \varphi(u_n))}{\varphi_{(n-1)}(\varphi(u_1) + ... + \varphi(u_n - 1))} \tag{2.3}$$

với

$$\varphi_{(n)} = \frac{\partial^n}{\partial u^n}\varphi^{-1}. \tag{2.4}$$

Cho copula MVN hoặc copula MVT, biến ngẫu nhiên phân phối đều thu được từ sự tạo thành các số ngẫu nhiên tương ứng phân phối MVN hoặc MVT.

**Chú ý 2.1.1.** Để thu được các số ngẫu nhiên nếu phân phối biên duyên không là phân phối đều, chúng ta sử dụng lớp các phương pháp nghịch đảo.

# 2.2 Ước lượng không tham số.

## 2.2.1 Copula thực nghiệm.

Cho $\chi = \{(x_1^t, ..., x_N^t)\}_{t=1}^T$ được ký hiệu là một mẫu. Phân phối copula thực nghiệm được cho bởi (Deheuvel [1979]):

$$\hat{C}\Big(\frac{t_1}{T}, ..., \frac{t_n}{T}, ..., \frac{t_N}{T}\Big) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}_{\big[x_1^t \le x_1^{(t_1)}, ..., x_n^t \le x_n^{(t_n)}, ..., x_N^t \le x_N^{(t_N)}\big]} \tag{2.5}$$

ở đây $x_n^{(t)}$ là thống kê thứ tự và $1 \le t_1, ..., t_N \le T$. Tần số copula thực nghiệm tương ứng để $\hat{c}\big(\frac{t_1}{T}, ..., \frac{t_n}{T}, ..., \frac{t_N}{T}\big) = \frac{1}{T}$ nếu $(x_1^{(t_1)}, ..., x_N^{(t_N)})$ thuộc về $\chi$ hoặc 0 nếu trái lại. Hệ thức giữa phân phối copula thực nghiệm và tần số là:

$$\hat{C}\Big(\frac{t_1}{T}, ..., \frac{t_n}{T}, ..., \frac{t_N}{T}\Big) = \sum_{i_1=1}^{t_1} \cdots \sum_{i_N=1}^{t_N} \hat{c}\Big(\frac{i_1}{T}, ..., \frac{i_n}{T}, ..., \frac{i_N}{T}\Big) \tag{2.6}$$

và

$$\hat{c}\Big(\frac{t_1}{T}, \ldots, \frac{t_n}{T}, \ldots, \frac{t_N}{T}\Big) = \sum_{i_1=1}^{2} \cdots \sum_{i_N=1}^{2} (-1)^{i_1+\cdots+i_N} \times \hat{C}\Big(\frac{t_1 - i_1 + 1}{T}, \ldots, \frac{t_n - i_n + 1}{T}, \ldots, \frac{t_N - i_N + 1}{T}\Big). \tag{2.7}$$

Sử dụng hệ thức $\rho = 2sin(\frac{\pi}{6}\varrho)$ giữa hai tham số tương quan của copula gauss và Spearman, chúng ta kết luận ước lượng ma trận tương quan của tài sản cho copula gauss (Bảng 2.1). Nếu chúng ta so sánh nó với tương quan cho trong Bảng 1.1, chúng ta thấy rằng nó đóng nhưng khác nhau.

## 2.2.2 Phép đồng nhất của copula Archimedean

Chúng ta sẽ phát triển phương pháp thực nghiệm để đồng nhất hóa copula trong trường hợp Archimedean. Giả sử $X$ là một véctơ với $N$ biến ngẫu nhiên, copula $C$ liên kết với hàm sinh $\varphi$ và hàm $K$ định nghĩa bởi (Genest và Rivest [1993]):

$$K(u) = P\{C(U_1, ..., U_N) \le u\} \tag{2.8}$$

| | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| Al | 1.00 | 0.87 | 0.52 | 0.41 | 0.36 |
| Al-15 | | 1.00 | 0.45 | 0.36 | 0.32 |
| Cu | | | 1.00 | 0.43 | 0.39 |
| Ni | | | | 1.00 | 0.34 |
| Pb | | | | | 1.00 |

Bảng 2.1: Ma trận tương quan $\rho$ của dữ liệu LME

ta chứng tỏ được rằng:

$$K(u) = u + \sum_{n=1}^{N} (-1)^n \frac{\varphi^n(u)}{n!} \varkappa_{n-1} \tag{2.9}$$

với $\varkappa_n(u) = \frac{\partial_n \varkappa_{n-1}(u)}{\partial_n \varphi(u)}$ và $\varkappa_0(u) = \frac{1}{\partial_n \varphi(u)}$. Trong trường hợp hai chiều, công thức này được rút gọn:

$$K(u) = u - \frac{\varphi(u)}{\varphi'(u)}.$$

Một ước lượng không tham số của $K$ được cho bởi:

$$\hat{K}(u) = \frac{1}{T} \sum_{t=1}^{T} \mathbf{1}_{[\vartheta_i \leq u]} \tag{2.10}$$

với

$$\vartheta_i = \frac{1}{T-1} \sum_{t=1}^{T} \mathbf{1}_{[x_1^t < x_1^i, \ldots, x_N^t < x_N^i]}. \tag{2.11}$$

Ý tưởng là từ $\hat{K}$ thích hợp chọn được một tham số copula trong họ copula Archimedean.

## 2.3 Ước lượng tham số

### 2.3.1 Ước lượng hợp lý cực đại (Maximum likelihood estimation: ML)

Cho tham số $\theta$ là $K \times 1$ véctơ để ước lượng tham số và $\Theta$ là không gian tham số. Ước lượng hợp lý cho quan sát t là hàm mật độ xác suất cho quan sát t, coi như là một hàm của $\theta$, ký hiệu $L_t(\theta)$. Giả sử $l_t(\theta)$ là hàm log-hợp lý (log - likelihood) của $L_t(\theta)$. Cho $T$ quan sát, chúng ta có:

$$l(\theta) = \sum_{t=1}^{T} l_t(\theta) \tag{2.12}$$

là hàm log-hợp lý. Và $\hat{\theta}_{ML}$ là gọi ước lượng hợp lý cực đại (Maximum likelihood estimation: ML) nếu:

$$l(\hat{\theta}_{ML}) \geq l(\theta) \quad \forall \theta \in \Theta. \tag{2.13}$$

Chúng ta có thể chứng tỏ rằng $\hat{\theta}_{ML}$ có tính chất tiệm cận chuẩn tắc và chúng ta có:

$$\sqrt{T}(\hat{\theta}_{ML} - \theta_0) \longrightarrow \mathcal{N}(0, \mathcal{J}^{-1}(\theta_0)) \tag{2.14}$$

$$var\big(\hat{\theta}_{ML}\big) = \big(-H_{\hat{\theta}_{ML}}\big)^{-1} \tag{2.15}$$

hoặc nghịch đảo của ước lượng OPG

$$var\big(\hat{\theta}_{ML}\big) = \big(J^T_{\hat{\theta}_{ML}} J_{\hat{\theta}_{ML}}\big)^{-1} \tag{2.16}$$

ước lượng khác gọi là ước lượng White (hoặc " sandwich") được định nghĩa bởi:

$$var\big(\hat{\theta}_{ML}\big) = \big(-H_{\hat{\theta}_{ML}}\big)^{-1}\big(J^T_{\hat{\theta}_{ML}} J_{\hat{\theta}_{ML}}\big)\big(-H_{\hat{\theta}_{ML}}\big)^{-1}. \tag{2.17}$$

Biểu thức hàm log-hợp lý trở thành

$$l(\theta) = \sum_{t=1}^{T} \ln c\big(F_1(x_1^t), ..., F_n(x_n^t), ..., F_N(x_N^t)\big) + \sum_{t=1}^{T}\sum_{n=1}^{N} \ln f_n(x_n^t). \tag{2.18}$$

Nếu chúng ta giả thiết các phân phối biên duyên là phân phối đều, ta có:

$$l(\theta) = \sum_{t=1}^{T} \ln c\big(u_1^t, ..., u_n^t, ..., u_N^t\big). \tag{2.19}$$

Trong trường hợp copula gauss, ta có:

$$l(\theta) = -\frac{T}{2}\ln|\rho| - \frac{1}{2}\sum_{t=1}^{T} \varsigma_t^T(\rho^{-1} - \mathbf{I})\varsigma_t \tag{2.20}$$

với $\varsigma_t = (\Phi^{-1}(u_1^t), \ldots, \Phi^{-1}(u_n^t), \ldots, \Phi^{-1}(u_N^t))$ và ước lượng ML của $\rho$ cũng vậy

$$\hat{\rho}_{ML} = \frac{1}{T}\sum_{t=1}^{T} \varsigma_t^T \varsigma_t. \tag{2.21}$$

Cho copula gauss khác, ước lượng tham số copula khác có thể đòi hỏi sự tối ưu hóa cho hàm log-hợp lý. Trường hợp này copula student là:

$$l(\theta) = \sum_{t=1}^{T} \ln c\big(F_1(x_1^t;\theta_1), ..., F_n(x_n^t;\theta_n), ..., F_N(x_N^t;\theta_N);\alpha\big) + \sum_{t=1}^{T}\sum_{n=1}^{N} \ln f_n(x_n^t;\theta_n) \tag{2.22}$$

với $\theta = (\theta_1, \ldots, \theta_N, \alpha)$. Chúng ta cũng có thể thực hiện ước lượng của phân phối biên duyên một chiều trong bước thứ nhất

$$\hat{\theta}_n = argmax \sum_{t=1}^{T} \ln f_n(x_n^t; \theta_n) \tag{2.23}$$

và ước lượng $\alpha$ trong bước thứ hai được cho bởi ước lượng trước

$$\hat{\alpha} = argmax \sum_{t=1}^{T} \ln c\big(F_1(x_1^t; \hat{\theta}_1), ..., F_n(x_n^t; \hat{\theta}_n), ..., F_N(x_N^t; \hat{\theta}_N); \alpha\big). \tag{2.24}$$

Phương pháp 2 - bước này được gọi là phương pháp hàm suy luận cho phân phối biên duyên hoặc phương pháp IFM, nói chung chúng ta có:

$$\hat{\theta}_{EML} \neq \hat{\theta}_{IFM}. \tag{2.25}$$

Ta thấy rằng "phương pháp IFM là phương pháp hiệu quả nhất để so sánh với phương pháp ML". Sử dụng một ý tưởng gần phương pháp IFM, chúng ta chú ý rằng véctơ tham số $\alpha$ của copula có thể không phải chỉ rõ ước lượng phân phối biên duyên. Phương pháp bao gồm biến đổi dữ liệu $(x_1^t, ..., X_N^t)$ vào biến ngẫu nhiên với phân phối đều $(\hat{u}_1^t, \ldots, \hat{u}_N^t)$ và ước lượng tham số bằng cách sau:

$$\hat{\alpha} = argmax \sum_{t=1}^{T} \ln c\big(\hat{u}_1^t, ..., \hat{u}_n^t, ..., \hat{u}_N^t; \alpha\big). \tag{2.26}$$

Trong trường hợp này, $\hat{\alpha}$ có thể được coi như dạng ước lượng ML cho phân phối biên duyên được quan sát( không phải giả thiết trên dạng tham số của phân phối biên duyên). Bởi vì nó là cơ sở trong phân phối thực nghiệm, chúng ta gọi nó là phương pháp hợp lý cực đại chính tắc hoặc CML. Chúng ta thực hiện một bài tập Monte Carlo để nghiên cứu tính chất của ba phương pháp (các giá trị tham số là $\lambda = 2,\ \gamma = 1.5,\ \rho = 0.5$), chúng ta có mô tả phân phối của ước lượng $\hat{\rho}_{EML},\ \hat{\rho}_{IFM}$ và $\hat{\rho}_{CML}$ cho các cỡ mẫu khác $T = 100$, $T = 500, T = 1000$, và $T = 2500$. Chúng ta chú ý rằng mật độ của ba ước lượng là rất gần.

**Chú ý 2.3.1.** Để ước lượng tham số $\rho$ của copula gauss với phương pháp CML, chú ý rằng cách tiếp cận này tương đương tính toán sự tương quan Spearman và được sử dụng hệ thức:

$$\rho = 2sin\Big(\frac{\pi}{6}\varrho\Big) \tag{2.27}$$

chúng ta thực hiện như sau:

1. Biến đổi dữ liệu gốc vào dữ liệu gauss:

   (a) Ước lượng các hàm phân phối thực nghiệm, sử dụng thống kê thứ tự.

   (b) Các giá trị gauss tạo thành khi thực hiện nghịch đảo của phân phối chuẩn từ hàm phân phối thực nghiệm.

2. Tính toán sự tương quan của biến đổi dữ liệu.

Chúng ta sử dụng lại ví dụ thị trường trao đổi kim loại London: LME. Ước lượng ma trận tương quan với phương pháp CML được cho bởi Bảng 2.2.

| | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| Al | 1.00 | 0.85 | 0.49 | 0.39 | 0.35 |
| Al-15 | | 1.00 | 0.43 | 0.35 | 0.32 |
| Cu | | | 1.00 | 0.41 | 0.36 |
| Ni | | | | 1.00 | 0.33 |
| Pb | | | | | 1.00 |

Bảng 2.2: Ma trận tương quan $\hat{\rho}_{CML}$ của copula gaussian cho dữ liệu LME

### 2.3.2 Phương pháp moment

Chúng ta xét moment thực nghiệm $h_{t,i}(\theta)$ phụ thuộc trên $K \times 1$ véctơ tham số $\theta$. $T$ là số lần quan sát và $m$ là số điều kiện hoặc moment. Xét véctơ hàng $h_t(\theta)$ mà các phần tử $h_{t,1}(\theta), ..., h_{t,m}(\theta)$ và $H(\theta)$ là ma trận $T \times m$ với phần tử $h_{t,i}(\theta)$. Giả sử $g(\theta)$ là một véctơ $m \times 1$ cho bởi:

$$g_i(\theta) = \frac{1}{T} \sum_{t=1}^{T} h_{t,i}(\theta). \tag{2.28}$$

Tổng quát hóa phương pháp moment GMM (Generalized Method of Moment), hàm tiêu chuẩn $Q(\theta)$ định nghĩa bởi:

$$Q(\theta) = g(\theta)^T W^{-1} g(\theta) \tag{2.29}$$

với W là một ma trận $m \times m$ đối xứng, xác định dương. Ước lượng GMM $\hat{\theta}_{GMM}$ tương ứng là:

$$\hat{\theta}_{GMM} = \underset{\theta \in \Theta}{argmin}\, Q(\theta). \tag{2.30}$$

Giống như ước lượng ML, ta có thể thấy rằng $\hat{\theta}_{GMM}$ có tính chất của chuẩn tắc tiệm cận và chúng ta có:

$$\sqrt{T}\left(\hat{\theta}_{GMM} - \theta_0\right) = \mathcal{N}(0, \Sigma). \tag{2.31}$$

Trong trường hợp tối ưu trọng số ($W$ là ma trận covariance $\Phi$ của $H(\theta)$ ) chúng ta có:

$$var\left(\hat{\theta}_{GMM}\right) = \frac{1}{T}\left[D^T \hat{\Phi}^{-1} D\right]^{-1} \tag{2.32}$$

với $D$ là ma trận Jacobian cấp $m \times K$ của $g(\theta)$ đã tính toán cho ước lượng $\hat{\theta}_{GMM}$

# Chương 3

# Ứng dụng copula trong đo lường rủi ro tài chính

Một trong những ứng dụng có hiệu lực mạnh nhất của copula là liên quan đến "Phép đo lường rủi ro". Chúng ta xét ba bài toán: tổn thất tổng hợp và phân tích giá trị rủi ro, giá trị cực trị nhiều chiều và rủi ro thị trường, tần số tương quan và tính toán giá trị rủi ro.

## 3.1 Tổn thất tổng hợp và phân tích giá trị rủi ro

Trong phần này, chúng ta sử dụng copula như là tập hợp các phân phối tổn thất và tính toán giá trị rủi ro (Value-at-Risk: VaR)

### 3.1.1 Trường hợp rời rạc

Xét mô hình rủi ro cá thể, mà thường được sử dụng trong bảo hiểm, Wang [1999] và Marceau, Cossette, Gaillardetz và Rioux [1999]. Giả sử $X_n$ và $X$ là $n^{th}$ rủi ro và tập hợp tổn thất. $X_n$ được định nghĩa như sau:

$$X_n = \begin{cases} A_n & \text{nếu } B_n = 1 \\ 0 & \text{nếu } B_n = 0 \end{cases} \tag{3.1}$$

ở đây $B_n$ là một biến ngẫu nhiên Bernoulli với tham số $p_n$ (nó phụ thuộc vào phân phối đồng thời của biến ngẫu nhiên $B_n$). Giả sử $F$ tương ứng là hàm phân phối tích lũy. Chúng ta có:

$$F(e_1, \dots, e_N) = C(B_{p_1}(e_1), \dots, B_{p_N}(e_N)). \tag{3.2}$$

Khi đó hàm sinh moment (mgf: moment generating function) của $(X_1, \dots, X_N)$ là:

$$\mathcal{M}_{(X_1,\dots,X_N)}(t_1, \dots, t_N) = \sum_{e_1,\dots,e_N \in \{0,1\}} c(e_1, \dots, e_N) \prod_{n=1}^{N} [\mathcal{M}_{A_n}(t_n)]^{e_n} \tag{3.3}$$

ở đây $\mathcal{M}_{A_n}(t) = E[e^{tA_n}]$ được biểu diễn là hàm sinh moment của biến ngẫu nhiên $A_n$. Phân phối $G$ của $X = \sum X_n$ thu được bởi nghịch đảo $\mathcal{M}_{(X_1,\ldots,X_N)}(t,\ldots,t)$ (bởi vì $\mathcal{M}_{(X_1,\ldots,X_N)}(t,\ldots,t)$) là hàm sinh moment của tổng trên.

### 3.1.2 Trường hợp liên tục

**a, Các copula đơn giản cho số chiều lớn**

**Định lý 3.1.1.** Trong trường hợp của copula student, ma trận $\rho$ có thể được sử dụng ước lượng bởi thuật toán sau:

1. Cho $\hat{\rho}_0$ là ước lượng ML của ma trận $\rho$ cho coupla gauss;
2. $\hat{\rho}_{m+1}$ thu được từ phương trình sau:

$$\hat{\rho}_{m+1} = \frac{1}{T}\Big(\frac{\nu+N}{\nu}\Big)\sum_{t=1}^{T}\frac{\varsigma_t^T\varsigma_t}{1+\frac{1}{\nu}\varsigma_t^T\hat{\rho}_m^{-1}\varsigma_t}; \tag{3.4}$$

3. Lặp lại bước thứ hai cho đến khi hội tụ $\hat{\rho}_{m+1} = \hat{\rho}_m (:= \hat{\rho}_\infty)$;
4. Ước lượng CML (hoặc IFM) của ma trận $\rho$ cho copula student là $\hat{\rho}_{CML} = \hat{\rho}_\infty$.

Chúng ta nhớ lại rằng copula student cho bởi biểu thức (1.33), nó trở thành hàm log-hợp lý:

$$\begin{aligned} l(\theta) &= T\Big(\ln\Gamma\Big(\frac{\nu+N}{\nu}\Big) - \Gamma\big(\frac{\nu}{2}\big)\Big) - NT\Big(\Gamma\Big(\frac{\nu+1}{2}\Big) - \Gamma\big(\frac{\nu}{2}\big)\Big) \\ &- \frac{T}{2}\ln|\rho| - \Big(\frac{v+N}{2}\Big)\sum_{t=1}^{T}\ln\Big(1+\frac{1}{v}\varsigma_t^T\rho^{-1}\varsigma_t\Big) + \Big(\frac{v+1}{2}\Big)\sum_{t=1}^{T}\sum_{n=1}^{N}\ln\Big(1+\frac{\varsigma_n^2}{v}\Big). \end{aligned} \tag{3.5}$$

Hàm log-hợp lý tập trung là:

$$\frac{\partial l(\theta)}{\partial\rho^{-1}} = -\frac{T}{2}\rho - \Big(\frac{\nu+N}{\nu}\Big)\sum_{t=1}^{T}\frac{\frac{1}{\nu}\varsigma_t^T\varsigma_t}{1+\frac{1}{\nu}\varsigma_t^T\rho^{-1}\varsigma_t}. \tag{3.6}$$

Nó cũng trở thành ước lượng ML nếu thỏa mãn phương trình ma trận phi tuyến sau:

$$\hat{\rho}_{ML} = \frac{1}{T}\Big(\frac{\nu+N}{\nu}\Big)\sum_{t=1}^{T}\frac{\varsigma_t^T\varsigma_t}{1+\frac{1}{\nu}\varsigma_t^T\hat{\rho}_{ML}^{-1}\varsigma_t}. \tag{3.7}$$

**b, Minh họa đầu tiên với phân phối biên duyên gauss**

Chúng ta nhớ lại rằng ma trận tương quan với ước lượng phương pháp CML cho copula gauss được xác định bởi Bảng 3.1. Trong trường hợp copula student, các ước lượng là khác nhau (ta thấy ở Bảng 3.2). Chúng ta tính toán vốn đầu tư cho các phương án đầu tư khác nhau. Trong những bảng sau, chúng ta chứng tỏ hợp thành của các danh mục đầu tư - một số âm tương ứng với một vị trí ngắn.

| | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| Al | 1.000 | 0.850 | 0.492 | 0.386 | 0.354 |
| Al-15 | | 1.000 | 0.429 | 0.346 | 0.316 |
| Cu | | | 1.000 | 0.409 | 0.359 |
| Ni | | | | 1.000 | 0.333 |
| Pb | | | | | 1.000 |

Bảng 3.1: Ma trận tương quan $\hat{\rho}_{CML}$ của copula gauss cho dữ liệu LME

| | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| $P_1$ | 1 | 1 | 1 | 1 | 1 |
| $P_2$ | -1 | -1 | -1 | -1 | -1 |
| $P_3$ | 2 | 1 | -3 | 4 | 5 |

| $\nu = 1$ | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| Al | 1.000 | 0.820 | 0.326 | 0.254 | 0.189 |
| Al-15 | | 1.000 | 0.271 | 0.223 | 0.164 |
| Cu | | | 1.000 | 0.267 | 0.219 |
| Ni | | | | 1.000 | 0.195 |
| Pb | | | | | 1.000 |

| $\nu = 2$ | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| Al | 1.000 | 0.875 | 0.429 | 0.351 | 0.282 |
| Al-15 | | 1.000 | 0.371 | 0.314 | 0.253 |
| Cu | | | 1.000 | 0.364 | 0.302 |
| Ni | | | | 1.000 | 0.278 |
| Pb | | | | | 1.000 |

| $\nu = 3$ | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| Al | 1.000 | 0.888 | 0.466 | 0.387 | 0.321 |
| Al-15 | | 1.000 | 0.410 | 0.349 | 0.291 |
| Cu | | | 1.000 | 0.399 | 0.336 |
| Ni | | | | 1.000 | 0.312 |
| Pb | | | | | 1.000 |

| $\nu = 5$ | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| Al | 1.000 | 0.893 | 0.493 | 0.410 | 0.350 |
| Al-15 | | 1.000 | 0.437 | 0.373 | 0.320 |
| Cu | | | 1.000 | 0.423 | 0.362 |
| Ni | | | | 1.000 | 0.337 |
| Pb | | | | | 1.000 |

| $\nu = 25$ | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| Al | 1.000 | 0.878 | 0.503 | 0.408 | 0.364 |
| Al-15 | | 1.000 | 0.446 | 0.370 | 0.331 |
| Cu | | | 1.000 | 0.425 | 0.371 |
| Ni | | | | 1.000 | 0.345 |
| Pb | | | | | 1.000 |

| $\nu = 100$ | Al | Al-15 | Cu | Ni | Pb |
|---|---|---|---|---|---|
| Al | 1.000 | 0.861 | 0.496 | 0.394 | 0.358 |
| Al-15 | | 1.000 | 0.436 | 0.354 | 0.322 |
| Cu | | | 1.000 | 0.415 | 0.363 |
| Ni | | | | 1.000 | 0.337 |
| Pb | | | | | 1.000 |

Bảng 3.2: Ma trận tương quan $\hat{\rho}_{CML}$ của copula student cho dữ liệu LME

Chúng ta tính toán VaR bởi giả thiết phân phối biên duyên là phân phối gauss nhưng với các hàm copula khác nhau. Kết quả là tóm tắt Bảng 3.3 và 3.4. Chúng ta chú ý rằng vốn kinh tế với mức tin cậy 99.9% thì copula gauss thấp hơn copula student. Điều đó chứng tỏ rằng cấu trúc phụ thuộc-hoặc hàm copula-có ảnh hưởng lớn trong tính toán VaR.

| | 90% | 95% | 99% | 99.5% | 99.9% |
|---|---|---|---|---|---|
| $P_1$ | 7.265 | 9.297 | 13.15 | 14.59 | 17.54 |
| $P_2$ | 7.263 | 9.309 | 13.17 | 14.59 | 17.65 |
| $P_3$ | 13.92 | 17.92 | 25.24 | 27.98 | 33.40 |

Bảng 3.3: Đo lường vốn kinh tế với copula gauss

**c, Minh họa thứ hai với phân phối biên duyên phần đuôi dư**

Chúng ta xét trường hợp của phân phối biên duyên phần đuôi dư, thực tế xảy ra nhiều trong tài chính. Chúng ta giả thiết rằng cấu trúc phụ thuộc được xác định một copula student với 2-bậc tự do. Trong phần trước, phân phối biên duyên của tiêu chuẩn lợi suất tài sản với là gauss. Chúng ta có thể giả sử rằng phân phối của tiêu chuẩn lợi suất cho một tài sản là student với $\nu$ - bậc tự do - các phân phối còn lại khác là gauss. Với tài sản của Al, ta thu được góc phần tư bên trái. Góc phần tư ở giữa và bên phải tương ứng là tài sản của Cu và Pb. Chúng ta cũng chú ý phần đuôi dư có tác dụng lớn trên VaR cho giá trị của mức tin cậy cao.

Trong thực tế phân phối biên duyên được dùng để mô tả lợi suất của các tài sản. Chúng ta xét ví dụ phân phối hyperbolic nói chung. Hàm mật độ tương ứng được xác

| $\nu = 1$ | 90% | 95% | 99% | 99.5% | 99.9% |
|---|---|---|---|---|---|
| $P_1$ | 5.696 | 7.9.77 | 13.22 | 15.46 | 20.16 |
| $P_2$ | 5.709 | 7.980 | 13.19 | 15.46 | 20.15 |
| $P_3$ | 13.43 | 19.34 | 34.32 | 40.65 | 55.12 |

| $\nu = 2$ | 90% | 95% | 99% | 99.5% | 99.9% |
|---|---|---|---|---|---|
| $P_1$ | 6.476 | 8.751 | 13.79 | 15.90 | 20.26 |
| $P_2$ | 6.470 | 8.753 | 13.70 | 15.77 | 20.12 |
| $P_3$ | 13.20 | 18.27 | 30.53 | 36.00 | 49.47 |

| $\nu = 3$ | 90% | 95% | 99% | 99.5% | 99.9% |
|---|---|---|---|---|---|
| $P_1$ | 6.834 | 9.077 | 13.85 | 15.73 | 20.31 |
| $P_2$ | 6.812 | 9.082 | 13.84 | 15.83 | 20.15 |
| $P_3$ | 13.18 | 17.86 | 28.69 | 33.56 | 45.01 |

| $\nu = 5$ | 90% | 95% | 99% | 99.5% | 99.9% |
|---|---|---|---|---|---|
| $P_1$ | 7.081 | 9.283 | 13.88 | 15.68 | 19.58 |
| $P_2$ | 7.109 | 9.351 | 13.94 | 15.72 | 19.57 |
| $P_3$ | 13.31 | 17.63 | 27.18 | 31.18 | 41.78 |

| $\nu = 25$ | 90% | 95% | 99% | 99.5% | 99.9% |
|---|---|---|---|---|---|
| $P_1$ | 7.318 | 9.453 | 13.56 | 15.10 | 18.18 |
| $P_2$ | 7.308 | 9.401 | 13.37 | 14.94 | 18.41 |
| $P_3$ | 13.64 | 17.60 | 25.38 | 28.36 | 34.90 |

| $\nu = 100$ | 90% | 95% | 99% | 99.5% | 99.9% |
|---|---|---|---|---|---|
| $P_1$ | 7.293 | 9.387 | 13.27 | 14.79 | 17.56 |
| $P_2$ | 7.297 | 9.427 | 13.42 | 14.78 | 17.56 |
| $P_3$ | 13.85 | 17.81 | 25.32 | 28.11 | 34.07 |

Bảng 3.4: Đo lường vốn kinh tế với copula student

định:

$$f(x) = a(\lambda, \alpha, \beta, \delta)\big(\delta^2 + (x-\mu)^2\big)^{\frac{1}{4}(2\lambda-1)} \times K_{\lambda-\frac{1}{2}}\Big(\alpha\sqrt{\delta^2 + (x-\mu)^2}\Big) \exp\big(\beta(x-\mu)\big) \quad (3.8)$$

ở đây $K$ được ký hiệu hàm Bessel và:

$$a(\lambda, \alpha, \beta, \delta) = \frac{(\alpha^2 - \beta^2)^{\frac{1}{2}\lambda}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^{\lambda}K_{\lambda}\big(\delta\sqrt{\alpha^2-\beta^2}\big)}. \quad (3.9)$$

Mô hình hyperbolic được ứng dụng vào quản lý thị trường rủi ro. Hơn nữa ta chỉ xét trường hợp một chiều. Phân phối của hyperbolic tổng quát nhiều chiều có dạng sau:

$$f(x) = A(\lambda,\alpha,\beta,\delta,\rho)\big(\delta^2 + (x-\mu)^T\rho^{-1}(x-\mu)\big)^{\frac{1}{4}(2\lambda-N)} \\ \times K_{\lambda-\frac{N}{2}}\Big(\alpha\sqrt{\delta^2+(x-\mu)^T\rho^{-1}(x-\mu)}\Big)\exp\big(\beta^T(x-\mu)\big) \quad (3.10)$$

ở đây

$$A(\lambda,\alpha,\beta,\delta,\rho) = \frac{(\alpha^2-\beta^T\rho\beta)^{\frac{1}{2}\lambda}}{(2\pi)^{\frac{N}{2}}\alpha^{\lambda-\frac{N}{2}}\delta^\lambda K_\lambda\Big(\delta\sqrt{\alpha^2-\beta^T\rho\beta}\Big)}. \quad (3.11)$$

Tuy nhiên, phân phối này dẫn tới vấn đề tính toán phức tạp cho ước lượng và các bước lặp lại, ta không nên dùng cái này. Nếu chúng ta xây dựng một phân phối nhiều chiều với phân phối hyperbolic một chiều và một copula. Với copula gauss, hàm mật độ là:

$$f(x) = \frac{1}{|\rho|^{\frac{1}{2}}}A(\lambda,\alpha,\beta,\delta,\rho)\Big[\prod_{n=1}^{N}\big(\delta_n^2+(x_n-\mu_n)^2\big)^{\frac{1}{4}(2\lambda_n-1)}\Big] \\ \times\Big[\prod_{n=1}^{N}K_{\lambda_n-\frac{1}{2}}\big(\alpha_n\sqrt{\delta_n^2+(x_n-\mu_n)^2}\big)\Big]\exp\big(\beta^T(x-\mu)-\frac{1}{2}\varsigma^T(\rho^{-1}-\mathbf{I})\varsigma\big) \quad (3.12)$$

ở đây

$$A(\lambda,\alpha,\beta,\delta,\rho) = \frac{1}{(2\pi)^{\frac{N}{2}}}\prod_{n=1}^{N}\frac{(\alpha_n^2-\beta_n^2)^{\frac{1}{2}\lambda_n}}{\alpha_n^{\lambda_n-\frac{1}{2}}\delta_n^{\lambda_n}K_{\lambda_n}\Big(\delta_n\sqrt{\alpha_n^2-\beta_n^2}\Big)} \quad (3.13)$$

và $\varsigma = (\varsigma_1,\dots,\varsigma_N)^T$, $\varsigma_n = \Phi^{-1}(F_n(x_n))$ và $F_n$ phân phối $GH$ một chiều. Ta thấy rằng phân phối này có $3(N-1)$ tham số để so sánh với phân phối hyperbolic nhiều chiều tổng quát.

## 3.2 Giá trị cực trị nhiều chiều và rủi ro thị trường

### 3.2.1 Lý thuyết giá trị cực trị

#### a, Trường hợp một chiều

Đầu tiên ta xét $m$ biến ngẫu nhiên độc lập $X_1,\dots,X_k,\dots,X_m$ với cùng hàm xác suất $F$. Phân phối cực trị $\chi_m^+ = \Big(\bigwedge_{k=1}^m X_k\Big)$ được xác định bởi định lý Fisher - Tippet.

**Định lý 3.2.1.** Nếu các tồn tại hằng số $a_m$, $b_m$ và một phân phối giới hạn không suy biến $G$ sao cho:

$$\lim_{m\to\infty} P\Big\{\frac{\chi_m^+-b_m}{a_m}\le x\Big\} = G(x),\ \forall x\in\mathbb{R} \quad (3.14)$$

thì $G$ là một trong những phân phối sau đây:

$$\text{(Fréchet) } G(x) = \Upsilon_\alpha(x) = \begin{cases} 0 & x \le 0 \\ \exp(-x^{-\alpha}) & x > 0 \end{cases} \quad g(x) = \begin{cases} 0 \\ \alpha x^{-(1+\alpha)} \exp(-x^{-\alpha}) \end{cases}$$

$$\text{(Weibull) } G(x) = \Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha) & x \le 0 \\ 1 & x > 0 \end{cases} \quad g(x) = \begin{cases} \alpha(-x)^{\alpha-1} \exp(-(-x)^\alpha) \\ 0 \end{cases}$$

$$\text{(Gumbel) } G(x) = \Lambda(x) = \exp(-e^{-x}) \quad x \in \mathbb{R} \qquad g(x) = \exp(-x - e^{-x})$$

Trong trường hợp này chúng ta nói $F$ thuộc về miền cực đại hấp dẫn (maximum domain of attraction :MDA) của $G - F \in MDA(G)$.

Hệ thức (3.14) có thể được viết dưới dạng sau:

$$F^m(a_m x + b_m) \longrightarrow G(x). \tag{3.15}$$

Chúng ta chú ý rằng nếu ta chỉ rõ $\{a_m\}$ và $\{b_m\}$ sao cho:

$$\lim_{m\to\infty} m[1 - F(a_m x + b_m)] = -\ln G(x) \tag{3.16}$$

thì $F \in MDA(G)$.

**b, Trường hợp nhiều chiều**

Trường hợp nhiều chiều, chúng ta quan tâm đến mô tả phân phối của cực trị $\chi_m^+$:

$$\chi_m^+ = \left(\chi_{1,m}^+, \ldots, \chi_{n,m}^+, \ldots, \chi_{N,m}^+\right) = \Big(\bigwedge_{k=1}^m X_{1,k}, \ldots, \bigwedge_{k=1}^m X_{n,k}, \ldots, \bigwedge_{k=1}^m X_{N,k}\Big). \tag{3.17}$$

Giống như trường hợp một chiều, giới hạn của cực trị chuẩn :

$$\lim_{m\to\infty} P\Big\{\frac{\chi_{1,m}^+ - b_{1,m}}{a_{1,m}} \le x_1, \ldots, \frac{\chi_{n,m}^+ - b_{n,m}}{a_{n,m}} \le x_n, \ldots, \frac{\chi_{N,m}^+ - b_{N,m}}{a_{N,m}} \le x_N\Big\} = G(x),$$
$$\forall (x_1, \ldots, x_n, \ldots, x_N) \in \mathbb{R}^N. \tag{3.18}$$

**Định lý 3.2.2.** Lớp các phân phối cực trị nhiều chiều là lớp các hàm phân phối max-ổn định với phân phối biên duyên không suy biến.

**Định lý 3.2.3.** Lớp các phân phối giá trị cực trị nhiều chiều là lớp các copula cực trị với phân phối biên duyên không suy biến.

Chúng ta nhớ lại rằng, copula cực trị thỏa mãn:

$$C(u_1^t, \ldots, u_n^t, \ldots, u_N^t) = C^t(u_1, \ldots, u_n, \ldots, u_N), \forall t > 0. \tag{3.19}$$

Giả sử $C$ là một cực trị copula với phân phối biên duyên là phân phối cực trị một chiều. Trường hợp này, $F(x_1, \ldots, x_n, \ldots, x_N) = C(F_1(x_1). \ldots, F_n(x_n), \ldots, F_N(x_N))$ là một phân phối MEV. Từ lý thuyết cực trị một chiều, $C^k(F_1(x_1)\dots, F_n(x_n), \dots, F_N(x_N))$ và $C(F_1^k(x_1)\dots, F_n^k(x_n), \dots, F_N^k(x_N))$ phải có phân phối giới hạn giống nhau cho mỗi số nguyên k. Chúng ta có:

$$C(u_1^k, \dots, u_n^k, \dots, u_N^k) = C^k(u_1, \dots, u_n, \dots, u_N), \forall k \in \mathbb{N}.$$

Để thu được phép biểu diễn Pickands của phân phối MEV, chúng ta giả sử $D$ là một phân phối nhiều chiều với số mũ đơn vị và một copula cực trị $C$. Ta sử dụng hệ thức sau:

$$\begin{aligned} C(u_1, \dots, u_n, \dots, u_N) &= C(e^{-\tilde{u}_1}, \dots, e^{-\tilde{u}_n}, \dots, e^{-\tilde{u}_N}) \\ &= D(\tilde{u}_1, \dots, \tilde{u}_n, \dots, \tilde{u}_N) \end{aligned} \tag{3.20}$$

chúng ta có

$$D^t(\tilde{u}) = D(t\tilde{u}) \tag{3.21}$$

và $D$ là một phân phối mũ nhiều chiều min-ổn định MSMVE (min-stable multivariate exponential: MSMVE)

**Định lý 3.2.4.** (Biểu diễn Pickands của phân phối MSMVE) Cho $D(\tilde{u})$ là hàm sống sót với phân phối biên duyên là phân phối mũ, $D$ thỏa mãn:

$$-\ln D(t\tilde{u}) = -t \ln D(\tilde{u}), \quad \forall t > 0 \tag{3.22}$$

nếu và chỉ nếu biểu diễn của $D$ là:

$$-\ln D(\tilde{u}) = \int \cdots \int_{S_n} \max_{1 \le n \le N} (q_n \tilde{u}_n) dS(q), \quad \tilde{u} \ge 0 \tag{3.23}$$

ở đó $S$ là độ đo hữu hạn trên $S_n$. (Joe [1997])

Thông thường biểu diễn Pickands một hàm phụ thuộc $B(w)$ được xác định bởi:

$$D(\tilde{u}) = \exp\Big[ -\Big( \sum_{n=1}^{N} \tilde{u}_n \Big) B(w_1, \dots, w_n, \dots, w_N) \Big] \tag{3.24}$$

$$B(w) = \int \cdots \int_{S_n} \max_{1 \le n \le N} (q_n w_n) dS(q) \tag{3.25}$$

với $w_n = \tilde{u}_n / \sum_{n=1}^{N} \tilde{u}_n$. $B$ là hàm lồi và

$$\max(w_1, \dots, w_n, \dots, w_N) \le B(w_1, \dots, w_n, \dots, w_N) \le 1. \tag{3.26}$$

Ta cần thiết kiểm tra nó là một copula

$$C^{\perp} \prec C \prec C^{+}. \tag{3.27}$$

Giả sử $F$ là hàm phân phối có $N$-biến ngẫu nhiên với phân phối biên duyên $F_1, \dots, F_N$ và liên kết với một copula $C$. Chúng ta giả định rằng tồn tại giới hạn phân phối và $F$ cũng thuộc về miền cực đại hấp dẫn của một phân phối $G$. Copula sẽ giúp chúng ta giải quyết bài toán mô tả phân phối $G$. Chúng ta sẽ kí hiệu phân phối biên duyên $G_1, \dots, G_N$ của $G$ và $C_\star$ tương ứng là một copula.

**Định lý 3.2.5.** $F \in MDA(G)$ nếu và chỉ nếu

1. $F_n \in MDA(G_n), \forall n = 1, \ldots, N$;
2. $C \in MDA(C_\star)$.

**Định lý 3.2.6.** $C \in MDA(C_\star)$ nếu và chỉ nếu:

$$\lim_{m\to\infty} \frac{1 - C\big((1-u)^{w_1}, \ldots, (1-u)^{w_n}, \ldots, (1-u)^{w_N}\big)}{u} = B(w_1, \ldots, w_n, \ldots, w_N). \tag{3.28}$$

Định lý này rất quan trọng, bởi vì phân phối MEV chỉ ra các hàm phụ thuộc.

**c, Trường hợp hai chiều**

Trong trường hợp hai chiều, lý thuyết cực trị thì dễ dàng tính toán bởi vì tính lồi và tính chất (3.25) trở nên cần thiết và là điều kiện đủ cho (3.23). Chúng ta có:

$$\begin{aligned} C(u_1, u_2) &= D(\tilde{u}_1, \tilde{u}_2) \\ &= \exp\Big[-(\tilde{u}_1 + \tilde{u}_2) B\Big(\frac{\tilde{u}_1}{\tilde{u}_1 + \tilde{u}_2}, \frac{\tilde{u}_2}{\tilde{u}_1 + \tilde{u}_2}\Big)\Big] \\ &= \exp\Big[\ln(u_1 u_2) B\Big(\frac{\ln u_1}{\ln(u_1 u_2)}, \frac{\ln u_2}{\ln(u_1 u_2)}\Big)\Big] \\ &= \exp\Big[\ln(u_1 u_2) A\Big(\frac{\ln u_1}{\ln(u_1 u_2)}\Big)\Big] \end{aligned} \tag{3.29}$$

với $A(w) = B(w, 1-w)$, tuy nhiên $A$ là lồi với $A(0) = A(1) = 1$ và thử lại $\max(w, 1-w) \leq A(w) \leq 1$.

Chúng ta xét copula Gumbel định nghĩa ở phần trước, chúng ta có $-\ln D(\tilde{u}) = (\tilde{u}_1^\alpha + \tilde{u}_2^\alpha)^{\frac{1}{\alpha}}$, $B(w_1, w_2) = \tilde{u}_1^\alpha + \tilde{u}_2^\alpha)^{\frac{1}{2}} / (\tilde{u}_1 + \tilde{u}_2) = (w_1^\alpha + w_2^\alpha)^{\frac{1}{2}}$ và $A(w) = [w^\alpha + (1-w)^{\frac{1}{\alpha}}]$. Để chỉ ra cực trị copula, ta sử dụng phụ thuộc độ đo $A(w)$ đơn giản:

$$\begin{aligned} \tau &= 4 \int_{\mathbf{I}} (1-w) d \ln A(w) \\ \varrho &= 12 \int_{\mathbf{I}} \frac{1}{[A(w)+1]^2} dw - 3. \end{aligned} \tag{3.30}$$

Khi các cực trị độc lập, chúng ta có $A(w) = 1$, ($\alpha = 1$ cho copula Gumbel), và $\tau = \varrho = 0$. Trường hợp phụ thuộc đầy đủ tương ứng $A(w) = \max(w, 1-w)$ dẫn đến copula Fréchet trên hay $C(u_1, u_2) = \exp\Big[\ln(u_1 u_2) \max\Big(\frac{\ln u_1}{\ln(u_1 u_2)}, \frac{\ln u_2}{\ln(u_1 u_2)}\Big)\Big] = \min(u_1, u_2)$. Phương trình (3.28) được sử dụng để kiểm tra độc lập của cực trị. Ví dụ chúng ta xét copula Kimeldorf - Sampson $C(u_1, u_2) = (u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-\frac{1}{\alpha}}$ với $\alpha \geq 0$, chúng ta có $\lim\limits_{u\to 0} \frac{1 - C\left((1-u)^w, (1-u)^{1-w}\right)}{u} = \lim\limits_{u\to 0} \frac{1 - \left((1-u)^{-\alpha w} + (1-u)^{-\alpha(1-w)} - 1\right)^{-\frac{1}{\alpha}}}{u} = \lim\limits_{u\to 0} \frac{1 - (1 + \alpha u + 0(u))^{-\frac{1}{\alpha}}}{u} = \lim\limits_{u\to 0} \frac{u + 0(u)}{u} 1$. A(u) sẽ dần đến 1, và copula Kimeldorf - Sampson thuộc về miền $C^\perp$

### 3.2.2 Các phương pháp ước lượng

**a, Phép xấp xỉ không tham số**

Xét một phân phối $Z = \frac{\ln U_1}{\ln U_1 + \ln U_2}$ thỏa mãn $\mathbf{F}(z) = z + (1-z)A^{-1}(z)\partial_z A(z)$, ở đây phân phối đồng thời của $U_1$ và $U_2$ xác định bởi cực trị copula $C_\star(u_1, u_2) = \exp\left[\ln(u_1u_2)A\left(\frac{\ln u_1}{\ln(u_1u_2)}\right)\right]$. Bởi vì $A(0) = A(1) = 1$, nó trở thành $A(w) = \exp\int\limits_0^w \frac{\mathbf{F}(z)-z}{1-z}\,dz$ hoặc $A(w) = \exp\int\limits_w^1 \frac{\mathbf{F}(z)-z}{1-z}\,dz$. Chúng ta cũng thu được hai ước lượng không tham số của $A(w)$ bởi phân phối thực nghiệm $\hat{\mathbf{F}}$ trong phân phối lý thuyết $\mathbf{F}$. Sử dụng hệ thức (3.28), xét các ước lượng khác dựa trên cơ sở mẫu của biến ban đầu $(X_1, X_2)$, không trên mẫu của biến cực đại $(X_1^+, X_2^+)$.

**b, Phép xấp xỉ tham số**

Trong các dạng toán cực đại theo từng điểm, phân phối $G$ được ứng dụng bởi phương pháp ML

$$\mathbf{G}(\chi_1^+, \dots, \chi_n^+, \dots, \chi_N^+) = C_\star\big(\mathbf{G}_1(\chi_1^+), \dots, \mathbf{G}_n(\chi_n^+), \dots, \mathbf{G}_N(\chi_N^+)\big) \tag{3.31}$$

ở đây $C_\star$ là một copula cực trị và $\mathbf{G}_n$ là GEV phân phối $\mathcal{GEV}(\mu, \sigma, \xi)$ xác định bởi:

$$\mathbf{G}(x) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{\frac{1}{\xi}}\right\} \tag{3.32}$$

xác định trên giá $\Delta = \{x : 1 + \xi(\frac{x-\mu}{\sigma})\}$. Ba kiểu phân phối một chiều không tham số là tổ hợp các họ giá trị cực trị tổng quát và chúng ta có tương ứng $\xi = \alpha^{-1} > 0$, $\xi = -\alpha^{-1} < 0$ và $\xi \longrightarrow 0$ cho các phân phối liên quan Fréchet, Weibull, Gumbel. Phương pháp IFM hoặc CML có thể được sử dụng ước lượng tham số (3.28) để rút gọn thời gian tính toán. Trong trường hợp này, hợp lý - log của biên duyên một chiều là:

$$l(\chi_n^+; \theta) = -\ln\sigma - \left(\frac{1+\xi}{\xi}\right)\ln\left(1 + \xi\left(\frac{\chi_n^+ - \mu}{\sigma}\right)\right) - \left[1 + \xi\left(\frac{\chi_n^+ - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}. \tag{3.33}$$

Một điểm rất quan trọng liên quan đến giá trị ban đầu của ước lượng copula. Ta thu được bởi nghịch đảo sự phụ thuộc tiệm cận đuôi trên hoặc sự tương thích độ đo như $\tau$ của Kendall.

## 3.3 Tần số tương quan và tính toán rủi ro

Phương pháp luận để tính toán rủi ro xác định trong dữ kiện sau:

- Giả sử $\zeta$ là biến ngẫu nhiên mô tả tổn thất. Chúng ta cũng định nghĩa $\zeta^k(t)$ là quá trình ngẫu nhiên của $\zeta$ cho mỗi tính toán rủi ro $k$ $(k = 1, \dots K)$

- Mỗi rủi ro, chúng ta giả thiết số các biến cố tại thời gian $t$ là một biến ngẫu nhiên $N_k(t)$

- Quá trình tổn thất $\varrho(t)$ được định nghĩa như sau:

$$\begin{aligned}\varrho(t) &= \sum_{k=1}^{K} \varrho^k(t) \\ &= \sum_{k=1}^{K} \sum_{j=1}^{N_k(t)} \zeta_j^k(t).\end{aligned} \tag{3.34}$$

- Vốn tài chính với mức tin cậy $\alpha$ thường được xác định như sau:

$$EC = \mathbf{F}^{-1}(\alpha). \tag{3.35}$$

Nếu phương pháp luận là đơn giản, thì nhiều vấn đề sẽ nảy sinh trong thực tiễn. Một vấn đề khác nữa của tính toán rủi ro: tương quan các tần số của các kiểu rủi ro là khác nhau

$$\mathbf{E}[N_{k_1}(t)N_{k_2}(t)] \neq \mathbf{E}[N_{k_1}(t)] \times \mathbf{E}[N_{k_2}(t)] \tag{3.36}$$

$N_k(t)$ là giả thiết chung từ biến ngẫu nhiên Poisson $\mathcal{P}$ với giá trị trung bình $\lambda_k$. Ý tưởng cũng là mở rộng phân phối Poisson nhiều chiều (Johnson, Kotz và Balakrishnan [1997]). Giả sử $N_{11}, N_{12}$ và $N_{22}$ là ba biến ngẫu nhiên Poisson độc lập với giá trị trung bình $\lambda_{11}, \lambda_{12}$ và $\lambda_{22}$. Trong trường hợp hai chiều, phân phối đồng thời dựa trên biến $N_1 = N_{11} + N_{12}$ và $N_2 = N_{22} + N_{12}$. Chúng ta có $N_1 \sim \mathcal{P}(\lambda_1 = \lambda_{11} + \lambda_{12})$ và $N_2 \sim \mathcal{P}(\lambda_2 = \lambda_{22} + \lambda_{12})$. Tuy nhiên xác suất đồng thời sẽ là:

$$P(N_1 = n_1, N_2 = n_2) = \sum_{n=0}^{\min(n_1,n_2)} \frac{\lambda_{11}^{n_1-n}\lambda_{12}^{n_2-n}\lambda_{12}^{n}e^{-(\lambda_{11}+\lambda_{22}+\lambda_{12})}}{(n_1-n)!(n_2-n)!n!}. \tag{3.37}$$

# KẾT LUẬN

Trong luận văn này em đã trình bày những khái niệm, định nghĩa, định lý về copula và một vài kết quả chính về ứng dụng của copula trong tài chính. Copula thật ra là mô hình phụ thuộc cấu trúc. Copula bộc lộ là một công cụ rất mạnh trong tài chính, đặc biệt trong các mô hình lợi suất và quản lý rủi ro. Những đóng góp chính của luận văn bao gồm:

1. Những kiến thức chung về xác suất, thống kê và copula
2. Các kết luận chính của thống kê về copula.
3. Ứng dụng trong tài chính về phép đo lường rủi ro

Tuy nhiên do thời gian không nhiều, nên luận văn còn có những chỗ thiếu sót và hạn chế. Em rất mong nhận được sự đóng góp, phản biện của quý thầy cô bạn đọc để luận văn được hoàn thiện tốt hơn.

# Tài liệu tham khảo


[1] Abramowitz, M. và I.A. Stegun *Handbook of Mathematical Function, ninth edition, Dover* [1997].

[2] Davidson, R. và J. Mackinnon *Estimation and inference in Econometrics* Oxford University Press, Oxford [1993].

[3] Eric Bouye *Copulas for Finance A Reading Guide and Some Application* London [2000]

[4] Genest, C. và J. Mackay *The joy of copulas: Bivariate distributions with uniform marginals*, American Statistician,40, 280-283 [1986].

[5] Genz, A. và F. Bretz *Numerical computation of multivariate t-probabilities*, Departement of Mathematics, Washington State University, Working Paper [1999a, 1999b].

[6] Joe, H. *Multivariate Models and Dependence Concepts, Monographs on Statistics and Applied Probability*, 73, Chapmann and Hall, London [1997].

[7] Jorgensen, B. *The theory of dispersion models*, Chapman and Hall, London [1997].

[8] Jorgensen, B.và S.L. Lauritzen *Multivariate dispersion models, Departement of Statistics and Demography*, Odense University, Research Report [1998].

[9] Legras, J. *The 1998 financial crisis: how serious was it?* Crédit Commercial de France, Working Paper [1999].

[10] Marceau, E., H. Cossette, P. Gaillardetz và J. Rioux *Dependence in the individual risk model* Université de Laval, École d'Atuariat, Rapport Technique, 99-01 [1999].

[11] Nelsen, R.B *An Introduction to Copula, Lectures Notes in Statistics, 139* Springer Verlag, New York [1998].

[12] Paul Embrechts, Filip Lindskog và Alexander McNeil *Modelling Dependence with Copulas and Application to Risk Mângement* Switzerland [2001]

[13] Schweitzer,B. và Wolff *On nonparametric measures of dependence for radom variables* Annals of Statistics, 9, 879-885 [1981].

[14] Sklar, A *Fonction de repartition à n dimension et leurs marges*, Publications de l'Institut de Statistique de l'Université de Paris, 8, 229-231 [1959].

[15] Song, P *Multivariate dispersion models genrated from Gaussian copula* forthcoming in Scandi-navian Journal of Statistics Zürich [2000].

[16] Wang, S.S. *Aggregation of correlated risk portfolios: models and algorithms* CAS Committee on Theory of Risk, preprint [1999].