

극단치 분포와 Copula함수를 이용한 주식시장간 극단적 의존관계 분석

김용현 · 배석주[†]

한양대학교 산업공학과

The Analysis of Tail Dependence Between Stock Markets Using Extreme Value Theory and Copula Function

Yong Hyun Kim · Suk Joo Bae

Department of Industrial Engineering, Hanyang University, Seoul 133-791, Korea

This article suggests the methods to investigate adverse movement across global stock markets arising from insolvency of subprime mortgage in U.S. Our application deals with asymptotic tail dependence of daily stock index returns (KOSPI, DJIA, Shanghai Composite) of three countries; Korea, U.S., and China, over specific period via extreme value theory and copula functions. Daily stock index returns among three countries show higher extremal dependence during the period exposed to systematic shock. We confirm that extreme value theory and copula functions have potential to well describe the extreme dependence between three countries' daily stock index returns.

Keywords: Extreme Value Theory, GPD, Copulas, Tail Dependence

1. 서론

금융 산업에 노출되어 있는 위험요소들은 무수히 많다. 개별 금융기관의 입장에서는 수익을 위해서 위험을 보유하고 있어야 하지만 대부분의 경우에는 위험을 회피해야 할 대상으로 인식하고 그를 관리하고자 한다. 국제결제은행(BIS : Bank for International Settlements)은 위험가중자산대비 보유해야할 자본의 크기를 최소 8% 이상이 되도록 하는 최저자기자본규제(Minimum Required Capital)를 규정함으로써 위험 자산의 측정 및 관리를 강조하고 있는데, 실제로 위험을 측정하는데 있어서 가장 중요한 것은 위험자산 수익률에 대한 변동성과 분포 함수에 대한 적절한 가정이라고 할 수 있다. 현실적으로는 이론적 어려움과 계산상 번거로움으로 인해 위험 자산의 수익률이 일정한 변동성을 갖고 정규분포를 따른다는 가정을 하지만 여러 실증 연구들에 따르면 변동성은 시간에 따라 변하며 수

익률의 분포는 비대칭적이고 꼬리가 두터운 분포를 하는 것으로 알려져 있다(Embréchts, 1997). 따라서 정규분포 가정은 위험을 과소평가하는 경향을 띠게 된다.

뿐만 아니라 금융기관이 보유하고 있는 자산들은 여러 개의 위험 자산으로 구성되어 있는 포트폴리오임을 감안할 때, 위험 자산 간의 의존관계를 반영한 자산전체의 결합분포함수를 결정하는 것이 중요하다고 할 수 있다. 선형상관관계만을 고려한 포트폴리오 이론은 비체계적 위험(기업 고유의 위험)을 최소화시키는데 목적이 있으므로(Markowitz, 1952) 최근 불거진 미국의 서브프라임 모기지부실로 인한 신용경색과 같은 체계적 위험(시장전체 위험)이 발생한 경우에는 분산효과를 발휘할 수 없게 된다.

따라서 위험관리자의 입장에서는 '시장 전체적인 체계적 위험'과 '발생빈도는 낮지만 손실금액이 큰 사건'에 관심을 가져야 하며 이를 측정하기 위하여 극단치 이론과 Copula함수를

이 논문은 2007년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음(KRF-2007-331-D00539).

[†] 연락저자 : 배석주 교수, 133-791 서울 성동구 행당동 17번지 한양대학교 산업공학과, Tel : 02-2220-0473, Fax : 02-2296-0417,

E-mail : sjbae@hanyang.ac.kr

2007년 10월 접수; 2007년 11월 수정본 접수; 2007년 11월 게재 확정.

이용한 극단적 사건들 간 의존관계의 모형화가 중요하다.

$$= 1 - \exp\left(-\frac{y}{\beta(u)}\right) \quad \text{for } \xi = 0$$

$$y = X - u > 0, \beta(u) = \sigma + \xi(u - \mu)$$

2. 극단치 이론(Extreme Value Theory)

여러 실증 연구에서 장기적으로 자산가격의 예측이 가능할 수 있음이 제기 되었으며 주가와 같은 자산수익률이 보여주는 실제분포는 정규분포보다 꼬리부분에 더 많은 관측치가 존재한다는 것이 일반적으로 인정되고 있다. 위험을 측정하고 관리해야 하는 관리자의 입장에서는 금융자산수익률을 정규분포로 가정하는 경우 위험을 과소평가 하게 되는 오류를 범하게 된다. 극단치는 발생확률이 낮은, 즉 빈번하지 않은 사건으로서 그 관측치가 소수이므로 정확하게 실증적인 분포를 찾기가 어렵다. 따라서 이론적인 분포함수가 실제 분포를 근사화하는데 사용된다. 이러한 극단치의 점근적(asymptotic) 분포는, 기초가 되는 분포에 대한 잘못된 가정으로 발생할 수 있는 오류를 줄일 수 있게 해준다.

극단치 분포는 극단치를 어떻게 정의하느냐에 따라 두 가지로 나누어진다. 첫째는 전체구간을 몇 개의 세부구간(block)으로 나눈 후 각 구간에서 발생하는 최대값을 극단치로 정의하는 Block Maxima 방법이고, 둘째는 임계값(threshold)을 초과하는 값들을 극단치로 정의하는 POT(Peak Over Threshold) 방법이다. Block Maxima는 일반화된 극단치 분포(GEV : Generalized Extreme Value)로, POT는 일반화된 파레토 분포(GPD : Generalized Pareto Distribution)로 각각 수렴하는 것으로 알려져 있다(Ebrecht, 2004). 그러나 Block Maxima 방법의 경우에는 구간 설정에 따른 결과의 차이가 민감하고 극단치 사건들의 군집현상(clustering)을 제대로 반영하지 못하는 이유로 본 논문에서는 POT 방법을 고려하기로 한다.

2.1 GPD(Generalized Pareto Distribution)

X_1, \dots, X_n 을 서로 동일하고 독립적인 손실분포함수 F 의 확률변수라고 하자. 극단 사건(Extreme Events)의 측정은 임계치 u 를 초과하는 X_i 를 관찰함으로써 얻을 수 있다. 임계치 u 에 대한 X_i 의 초과 정도를 다음의 조건부 확률을 이용하여 분포함수로 표현할 수 있다.

$$F_u(y) = \Pr\{X - u \leq y | X > u\}$$

$$= \frac{F(y+u) - F(u)}{1 - F(u)}, y > 0$$

임계치 u 가 추정되면 조건부 확률분포 $F_u(y)$ 의 극한분포는 GPD로 수렴하며 아래와 같은 극한확률밀도함수를 갖게 된다. [정리 : Balkema-De Haan-Pickands Theorem](1975)

$$G_{\xi, \beta(u)}(y) = 1 - \left(1 + \left(\xi \frac{y}{\beta(u)}\right)^{-1/\xi}\right) \quad \text{for } \xi \neq 0$$

μ 는 위치모수(location parameter), σ 는 척도모수(scale parameter)이며 ξ 는 분포의 형태를 결정짓는 형태모수(shape parameter)이다. 만약 $\xi > 0$ 이면 Type I 파레토 분포, $\xi < 0$ 이면 Type II 파레토분포, $\xi = 0$ 이면 지수분포 형태를 띤다. 임계치 u 가 충분히 크다면, Balkema-De Haan-Pickands Theorem에 의해 임계치를 초과하는 x 에 대한 손실분포의 꼬리 부분은 다음과 같이 근사화할 수 있다.

$$F_u(y) \approx G_{\xi, \beta(u)}(y)$$

따라서 손실분포는 꼬리부분과 중심부분을 분해한 반모수적(Semiparametric)인 분포로 생각할 수 있다(Carmona, 2004).

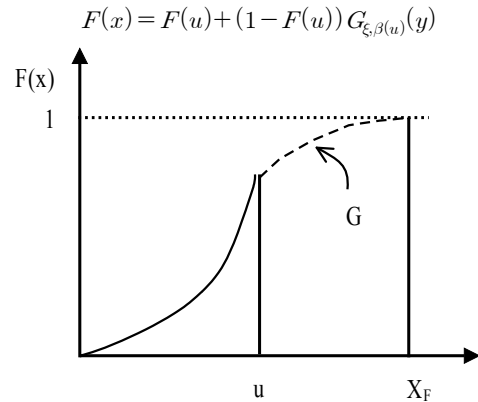


Figure 1. 분포함수 $F(x)$ 의 분해

2.2 평균초과함수 (Mean Excess Function)

평균초과함수는 임계치 이상의 관측치들에 대한 초과정도를 임계치에 대한 함수로 표현한다.

$$e(u) = E[X - u | X > u] = \frac{\beta(u)}{1 - \xi}$$

전체 표본이 N 개인 자료에서 평균초과함수는 다음과 같이 계산할 수 있다.

$$e(u) = \frac{\sum_{i=1}^N (X_i - u)^+}{\sum_{i=1}^N I(X_i > u)}$$

$$(X_i - u)^+ = \max(X_i - u, 0)$$

$$\text{여기서, } I(X_i > u) = \begin{cases} 1, & X_i > u \\ 0, & X_i \leq u \end{cases}$$

X 가 모수 λ 인 지수분포를 따른다면, 무기역성질에 의하여 $e(u) = \lambda^{-1}$, 즉 상수가 되지만 GPD의 경우에는 다음과 같이 임계치에 선형비례하는 함수로 표현된다.

$$e(u) = \frac{1}{1-F(x)} \int_u^{x_F} (x-u)dF(x) = \frac{\beta + \xi u}{1-\xi}$$

x_F : 유한 끝점 (Finite end point)

임계치 u 에 대한 평균초과함수의 기울기는 두터운 꼬리를 가지는 경우 $(\frac{1}{2} < \xi < 1)$ 다음과 같이 양의 값을 갖는다.

$$\frac{de(u)}{du} = \frac{\xi}{1-\xi} u > 0$$

그리고 임계치 u 가 커짐에 따라서도 평균초과함수는 양의 선형관계를 가지게 된다. 따라서 $[u, e(u)]$ 의 집합이 우상향하는 형태로 나타난다면 관측치들이 GPD분포를 따른다고 가정할 수 있다.

2.3 GPD의 추정

GPD의 모수를 추정하는 방법으로 확률가중모멘트법, 최우추정법 및 베이지언 시뮬레이션 방법을 적용할 수 있다. 본고에서는 100개 이상의 관측치가 있을 때 많이 사용되는 최우추정법을 사용하도록 한다. 최우추정치는 다음과 같은 우도함수를 최대화함으로써 구할 수 있다.

$$l(\xi, \beta(u)) = -k \ln(\beta(u)) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^k \ln\left(1 + \xi \frac{y_i}{\beta(u)}\right)$$

$\hat{\Theta} = \operatorname{argmax} l(\xi, \beta(u)), \Theta$: parameter space

3. Copula and Sklar's Theorem

3.1 Copula함수의 정의

Copula함수는 단일변량 한계분포와 결합분포를 연결시키는 함수로서 다음의 식을 만족한다.

$$H(x_1, x_2) = C_\theta(F_1(x_1), F_2(x_2)) \equiv C_\theta(u_1, u_2)$$

[정의 : Copula] n -차원 Copula C 는 다음과 같은 특성을 만족하는 함수이다(Sklar, 1959).

- (1) $C: [0, 1]^n \rightarrow [0, 1]$
- (2) C 는 "grounded" 함수이고 n -증가함수이다.
- (3) C 는 모든 $u \in [0, 1]^n$ 에 대하여 다음을 만족하는 한계분포함수 $C_i (i = 1, 2, \dots, n)$ 을 갖는다.

$$C_i(u) = C(1, \dots, 1, u, 1, \dots, 1) = u$$

위의 세 가지 특성을 만족하는 copula함수가 결합분포와 같다는 것은 다음과 같은 Sklar정리에 근거한다.

[정리 : Sklar's Theorem] H 가 결합분포함수를 나타내고 이의 연속인 한계분포함수가 F_1, \dots, F_n 이라고 하자. 이 경우 모든 실수 x_1, \dots, x_n 에 대하여 다음과 같은 관계를 만족하는 copula함수 C 가 유일하게 존재한다.

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

이는 결합분포를 각 확률변수들의 한계분포들과, 의존관계를 나타내는 Copula함수 C 로 분해할 수 있음을 나타낸다.

[보조정리] H 가 n -차원 결합분포함수를 나타내고 이의 연속인 한계분포함수가 F_1, \dots, F_n 이며 copula함수 C 를 갖는다고 하자.

모든 $u = (u_1, \dots, u_n) \in [0, 1]^n$ 에 대하여 다음이 성립한다.

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

3.2 상호의존관계의 측정

기존 선형 상관계수의 단점을 보완하고자 Kendall, Spearman은 사전적인 확률분포에 의존하지 않고 순위(rank)에 의존하는 비모수적 추정인 순위상관계수(rank correlation coefficient)를 제안하였다. 이 장에서는 순위상관계수와 더불어 극단적인 움직임의 동조화(concordance)의 정도를 표현한 꼬리 의존지수(tail dependence index)를 소개한다.

3.2.1 Kendall's tau

두 확률변수 X 와 Y 의 Kendall's tau는 아래와 같이 정의한다.

$$\begin{aligned} \tau &= \Pr[(X_1 - X_2)(Y_1 - Y_2) > 0] \\ &\quad - \Pr[(X_1 - X_2)(Y_1 - Y_2) < 0] \\ &= 4 \int \int_{\mathcal{F}^2} C \, dC - 1 \\ &= 4 \int \int_{\mathcal{F}^2} C(u, v) c(u, v) \, dudv - 1 \end{aligned}$$

비모수적으로 Kendall's tau를 추정하는 방법은 다음과 같이 방향성을 고려하여 모든 쌍의 개수에 대한 의존정도를 표현함으로써 정의할 수 있다.

$$\hat{\tau} = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \operatorname{sign}((x_i - x_j)(y_i - y_j))$$

3.2.2 Spearman's Rho

두 확률변수 X 와 Y Spearman's rho는 각 확률변수에 대한

한계분포인 $F_X(x)$ 와 $F_Y(y)$ 사이의 상관계수로 정의할 수 있다.

$$\rho_s(X, Y) = 12 \int_0^1 \int_0^1 C(u, v) du dv - 3$$

즉, Spearman's rho와 선형상관계수 사이에는 다음과 같은 관계가 있음을 알 수 있다.

$$\rho_s(X, Y) = \text{corr}(F(X), F(Y))$$

그리고 비모수적으로는 다음과 같이 추정된다.

$$\hat{\rho}_s = \frac{12}{n(n^2-1)} \sum_{i=1}^n \left(\text{rank}(x_i) - \frac{n+1}{2} \right) \left(\text{rank}(y_i) - \frac{n+1}{2} \right)$$

τ 와 ρ_s 는 선형상관계수와 달리, X 와 Y 사이의 증가함수로의 변환에 의해서 그 값이 변하지 않는 성질을 가지고 있다. 하지만 (linear or rank) correlation과 같이 한가지의 숫자를 갖는 측정치보다 자세한 정보를 반영할 수 있는 새로운 종속성 구조가 선택되어야 하며, 꼬리의존지수(Tail Dependence Measure)를 통해 표현할 수 있다(Embréchts, 1999).

3.2.3 Tail Dependence Measures

꼬리의존지수(Tail dependence index)는 이변량 분포의 upper quadrant tail과 lower quadrant tail에서 의존정도를 측정할 수 있다.

$$\lambda_u(X, Y) = \lim_{q \rightarrow 1} P(Y > F_Y^{-1}(q) | X > F_X^{-1}(q)) = \lim_{q \rightarrow 1} \frac{1 - 2q + C(q, q)}{1 - q}$$

$$\lambda_l(X, Y) = \lim_{q \rightarrow 0} P(Y \leq F_Y^{-1}(q) | X \leq F_X^{-1}(q)) = \lim_{q \rightarrow 0} \frac{C(q, q)}{q}$$

$\lambda_u(\lambda_l) \in (0, 1]$ 이면 upper tail (lower tail)에서 의존관계가 존재하고, $\lambda_u(\lambda_l) = 0$ 이면 upper tail (lower tail)에서 근사적으로 독립이라고 한다. 꼬리의존지수 역시 각 확률변수들의 한계분포와 상관없이 Copula에 의해서만 구해지는 값들이다. 뿐만 아니라 X 와 Y 의 증가함수로의 변환에 의해서도 변하지 않는 특성이 있다(Embréchts, 2001).

3.3 Copula 함수의 종류

3.3.1 Elliptical Copulas

Gaussian copula와 Student's t-copula는 대표적인 타원형분포(Elliptical distribution)와 관련된 copula로서 결합분포함수로부터

유도된다. 이변량인 경우 Gaussian copula C^{Gauss} 는 다음과 같이 정의한다.

$$C^{Gauss}(u_1, u_2) = \int_{-\infty}^{\Phi_1^{-1}(u_1)} \int_{-\infty}^{\Phi_2^{-1}(u_2)} f(x_1, x_2) dx_2 dx_1$$

$\Phi_1, \Phi_2 \in [0, 1]$ 은 1차원 누적정규분포밀도함수(한계분포)를 나타낸다.

[다변량 Gaussian copula] Σ 는 주대각선 원소들이 1을 갖는 PSD(Positive Symmetric Definite)행렬이고, Φ_Σ 는 상관계수행렬 Σ 를 갖는 다변량 표준정규분포함수라고 한다면 다변량 Gaussian Copula는 다음과 같이 정의한다.

$$C(u_1, \dots, u_n; \Sigma) = \Phi_\Sigma(\Phi_1^{-1}(u_1), \dots, \Phi_n^{-1}(u_n))$$

[다변량 Student's t-copula] Σ 는 주대각선 원소들이 1을 갖는 PSD(Positive Symmetric Definite)행렬이고, $t_{\Sigma, \nu}$ 는 상관계수행렬 Σ 와 자유도 ν 를 갖는 다변량 Student's t-분포함수라고 하자.

$$t_{\Sigma, \nu}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} \frac{\Gamma\left(\frac{\nu+n}{2} |\Sigma|^{-\frac{1}{2}}\right)}{\Gamma\left(\frac{\nu}{2}\right) (\nu\pi)^{\frac{n}{2}}} \left(1 + \frac{1}{\nu} x^T \Sigma^{-1} x\right)^{-\frac{\nu+n}{2}} dx$$

다변량 Student's t-Copula는 다음과 같다.

$$C(u_1, \dots, u_n; \Sigma, \nu) = t_{\Sigma, \nu}(t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_n))$$

3.3.2 Archimedean copula

Archimedean copula는 다음과 같이 정의한다(Genest and MacKay, 1986).

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

$\phi(u)$ 는 copula의 생성자(generator)라고 부르는 함수로 모든 $0 \leq u \leq 1$ 에 대하여 $\phi(1) = 0$, $\phi'(u) < 0$, 그리고 $\phi''(u) > 0$ 을 만족한다.

Table 1. Archimedean Copula

Family	Generator	Parameter	tail index
$\frac{[Frank]}{(1979)}$	$\ln \frac{e^{\alpha u} - 1}{e^\alpha - 1}$	$(-\infty, \infty)$	no tail index
$\frac{[Clayton]}{(1978)}$	$u^{-\alpha} - 1$	$\alpha > 1$	$\lambda_l = 2^{-1/\alpha}$
$\frac{[Gumbel]}{(1960)}$	$(-\ln u)^\alpha$	$\alpha \geq 1$	$\lambda_u = 2 - 2^{1/\alpha}$

3.3.3 Extreme Value Copula

극단값 copula C 는 다음과 같은 관계를 만족한다.

$$C(u_1^t, \dots, u_n^t) = C^t(u_1, \dots, u_n), \quad \forall t > 0, (u, v) \in [0, 1]^2$$

극단값 Copula C 에서 서로독립인 확률변수 X_i 와 Y_i 의 ($i = 1, 2, \dots, n$) 순서쌍(X_i, Y_i)를 추출했다고 가정한다. 그리고 각 확률변수들의 최대값을 $M_n = \max(X_1, \dots, X_n)$, $N_n = \max(Y_1, \dots, Y_n)$ 으로 정의하면, 극단값 Copula C 는 임의 순서쌍(M_n, N_n)과 관련된 Copula 함수가 된다. 이는 max-stability 성질이라고 하며 다음의 식으로 표현될 수 있다(Joe, 1997, p.175).

$$C(u, v) = \exp \left\{ \ln(uv) A \left(\frac{\ln(u)}{\ln(uv)} \right) \right\}$$

$A(\cdot)$ 는 의존함수(dependence function)이며 $[0, 1]$ 에 속하는 모든 t 에 대해서 $\max(t, t-1) \leq A(t) \leq 1$ 을 만족하는 Convex function을 의미한다. 극단값 Copula C 는 아래와 같이 $A(\cdot)$ 에 따라 정의되어진다. 그리고 $\lambda_u = 2 - 2^{1/\alpha}$ 인 위쪽 꼬리의존지수만 존재한다.

[Gumbel copula (1960)]

$$A(t) = (t^\alpha + (1-t)^\alpha)^{1/\alpha}, \quad \alpha \geq 1$$

[Galambos copula (1975)]

$$A(t) = 1 - (t^{-\alpha} + (1-t)^{-\alpha})^{-1/\alpha}, \quad 0 \leq \alpha < \infty$$

[Husler and Reiss copula (1987)]

$$A(t) = t\Phi \left[\frac{1}{\alpha} + \frac{1}{2}\alpha \ln \left(\frac{t}{1-t} \right) \right] + (1-t)\Phi \left[\frac{1}{\alpha} - \frac{1}{2}\alpha \ln \left(\frac{t}{1-t} \right) \right]$$

[BB5 copula (Joe, 1997)]

$$A(t) = \left[t^\theta + (1-t)^\theta - (t^{-\alpha\theta} + (1-t)^{-\alpha\theta})^{-1/\alpha} \right]^{1/\theta}$$

$\alpha > 0, \theta \geq 1$

3.4 Copula 함수의 추정

Copula 함수의 추정은 각 한계분포함수의 모수벡터 ϕ_i ($i = 1, 2, \dots, n$) 및 의존관계를 나타내는 Copula 함수 모수벡터 θ 를 추정하는 것으로서 일반적으로는 두 모수벡터들을 동시에 추정하는 방법(MLE : Maximum Likelihood Estimation)을 고려해볼 수 있으나, 추정해야할 모수의 개수가 많아지는 경우 계산 상 효율적이지 못하다는 지적 때문에 2단계 추정법(IFM : Inference Functions for Margins)을 많이 사용한다. 먼저 Marginal 분포 F_i 의 모수벡터(ϕ_i)를 최우추정한 후,

$$\hat{\phi}_i = \operatorname{argmax} l^i(\phi_i) = \operatorname{argmax} \sum_{t=1}^T \ln f_i(x_i^t; \phi_i)$$

($t = 1, \dots, T$: i 번째 한계분포를 추정하는데 사용되는 표본)

주어진 $\hat{\phi}_i$ 하에서 의존관계를 나타내는 Copula 모수벡터(θ)를 최우추정한다.

$$\hat{\theta} = \operatorname{argmax} l^i(\theta) = \operatorname{argmax} \sum_{t=1}^T \ln c(F_1(x_1^t; \hat{\phi}_1), \dots, F_n(x_n^t; \hat{\phi}_n); \theta)$$

일반적으로 MLE와 IFM 추정치는 서로 같지 않지만 IFM은 MLE와 비교할 때 효율성을 갖는 것으로 알려져 있다(Joe, 1997). 추정해야할 모수의 개수가 증가할수록 많은 추정시간을 소요하는 MLE보다는 IFM의 유용성이 크게 된다.

3.5 적정 Copula 함수의 선택

적정 Copula는 실증 Copula와 모수적 Copula의 차이를 제공했을 때, 그 값이 최소화되는 Copula를 적정 Copula로 선택하는 방법이다. 여기서 실증 Copula는 다음과 같이 정의한다(Deheuvels, 1978).

$u_{(1)} \leq u_{(2)} \leq \dots \leq u_{(n)}$, $v_{(1)} \leq v_{(2)} \leq \dots \leq v_{(n)}$ 인 순서통계량에 대하여 실증 copula \hat{C} 는 점 $\left(\frac{i}{n}, \frac{j}{n}\right)$ 에서 n 이 무한히 커질 때, 다음과 같이 C 에 수렴한다($i, j = 1, 2, \dots, n$).

$$\hat{C} \left(\frac{i}{n}, \frac{j}{n} \right) = \frac{1}{n} \sum_{k=1}^n 1_{[u_k \leq u_{(i)}, v_k \leq v_{(j)}]} \rightarrow C$$

뿐만 아니라, Bayesian 방법에 근거한 정보기준을 이용하여 적정한 Copula를 선택하기도 한다(Ralph S. Silva and Hedibert F. Lopes, 2007).

step1 : 한계분포함수(F, G)를 이용하여 관측치들(X_i, Y_i)을 균일분포로 변형시킨다.

$$F(X_i) = U_i, \quad G(Y_i) = V_i$$

step2 : 다음의 우도함수를 최대화하는 Copula 함수의 모수벡터를 찾는다.

$$\hat{\theta} = \operatorname{argmax} L(\theta; U, V) = \prod_{i=1}^n C_\theta(U_i, V_i)$$

step3 : 최소가 되는 AIC를 선택한다.

$$AIC = 2p - 2 \times (\log\text{-likelihood}), \quad p : \text{모수의 개수}$$

실제로는 위에서 언급하지 않은, 여러 가지 적정 Copula 검정 방법들이 존재한다. 하지만 각 방법들마다 장·단점들이

뚜렷하고 더욱이 검정 결과들이 실증분석에 사용되는 자료에 크게 의존하기 때문에 어떤 검정 방법이 통계적으로 가장 오차가 적고, 검정력이 가장 높은지 완벽한 해결책을 제시해 주지는 못한다. 이에 본 연구에서는 두 번째 방법으로 제시한 정보기준(AIC)에 의한 선택방법을 적정 Copula 함수의 선택기준으로 사용한다.

4. 실증 분석

본 연구에서는 한국 종합주가지수(KOSPI)가 미국 다우존스산업지수(DJIA), 그리고 중국 상해종합지수(Shanghai Composite : S.C.)와 갖는 극단적 의존관계를 분석하기 위해 626거래일(2003년 1월 3일~2007년 9월 20일)동안의 일별 증가를 수집하였다. 그리고 KOSPI와 DJIA, KOSPI와 Shanghai Composite를 차례대로 동일한 방법으로 분석하였다.

우선 체계적인 위험으로 변동성이 커지는 관심구간을 설정하였다(2007년 7월 25일~2009년 20일). 하지만 이에 해당하는 시계열자료들을 이용하는 경우 자료의 개수가 충분히 크지 않기 때문에 추정오차가 크게 발생하고 적합도가 낮아진다는 단점이 있다. 따라서 관심 구간 전까지의 극단적 의존관계와 관심 구간을 포함하는 현재까지의 극단적 의존관계를 비교함으로써 관심 구간에서의 영향을 간접적으로 파악하고자 한다.

	03.01.03~07.07.24	03.01.03~07.09.20
KOSPI DJIA	case 1~1	case 1~2*
KOSPI S.C.	case 2~1	case 2~2

(주) 이하는 case 1~2*를 중심으로 분석하였으며 다른 case 역시 동일하게 수행하였음.

4.1 자료의 기초통계량 분석

각 지수의 일별 증가는 불안정적 시계열이므로 로그차분을 통해서 안정적 시계열로 바꾸었다.

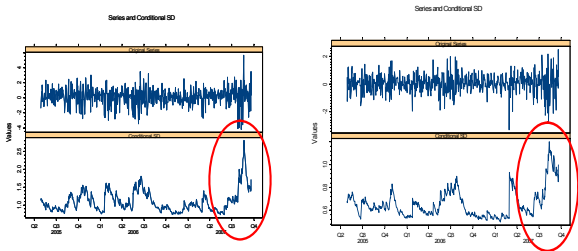


Figure 2. 일별수익률 및 변동성(KOSPI, DJIA)

체계적인 위험으로 인하여 변동성이 극단적으로 높아지는 구간을 찾기 위해 수익률의 변동성을 GARCH(1, 1)으로 모형

화하여 비교하였다. <그림 2>에서 볼 수 있듯이 2007년 3분기 이후에서 상호간의 변동성이 급격하게 증폭되는 것을 확인할 수 있다.

실제 자료들이 따르는 분포 모양을 추정하기에 앞서 정규성 검정을 시행한다. 실제 자료들은 비대칭적이고 두터운 꼬리를 가진다고 알려져 있다(Embrechts, 1999). 이는 아래의 그림에서 눈으로 직접 확인할 수 있다.

Table 2. 기초통계량(KOSPI, DJIA)

일 별	KOSPI	DJIA
평 균	0.1218%	0.0517%
표준편차	1.1417%	0.6769%
왜 도	-0.3538	-0.3911
첨 도	5.0812	5.3050
JB통계량 (p-value)	122.7424* (0.0000)	150.6315* (0.0000)

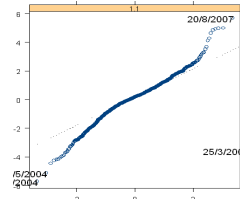


Figure 3. QQ-Plot(KOSPI, DJIA)

* H_0 : Data is normally distributed, H_1 : not H_0

$$JB = \frac{T}{6} \left(\widehat{skew}^2 + \frac{(\widehat{kurt} - 3)^2}{4} \right) \sim \chi^2(2)$$

\widehat{skew} : sample skewness, \widehat{kurt} : sample kurtosis

<그림 3>은 QQ-Plot으로서 이론적 정규분포의 분위수(x축)에 대한 실제 자료의 분위수(y축)를 비교하는 방식인데, 실제 자료들이 45°선 주변에 가깝게 분포할수록 정규분포에 적합하다는 것을 의미한다. 그런데 실제로는 양쪽 꼬리부분으로 갈수록 실제 자료와 45°선의 이탈 정도가 높아지는 것을 확인할 수 있고 따라서 전통적인 방법, 즉 개별 자산의 수익률 분포를 정규분포로 가정하는 것은 위험을 측정하는데 큰 오차가 발생할 수 있음을 암시한다.

4.2 POT-Copula 분석

4.2.1 POT 분석

개별 자산 수익률의 극단치 분포를 모형화하기 위해 POT방

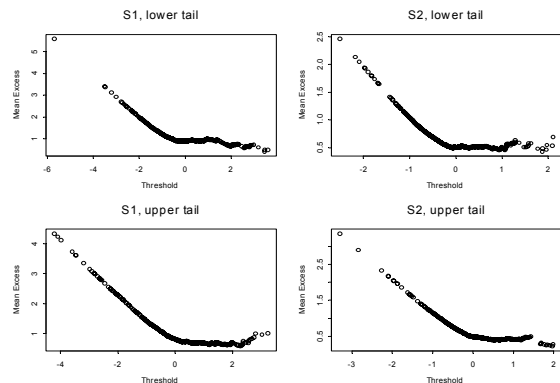


Figure 4. 평균초과함수(KOSPI, DJIA)

법을 적용한다. 여기서는 임계치 u 를 설정하는 것이 중요하다. 임계치를 추정하는 다양한 방법의 선행 연구들에서는 주로 평균초과 함수와 Hill 그래프를 이용한다. 본 연구에서도 두 가지의 방법을 조화시켜 임계치 u 를 설정하였다. 평균초과 함수의 경우 양(+)¹⁾의 기울기를 가질 때 GPD와 같은 극단치 분포를 따른다고 볼 수 있다.

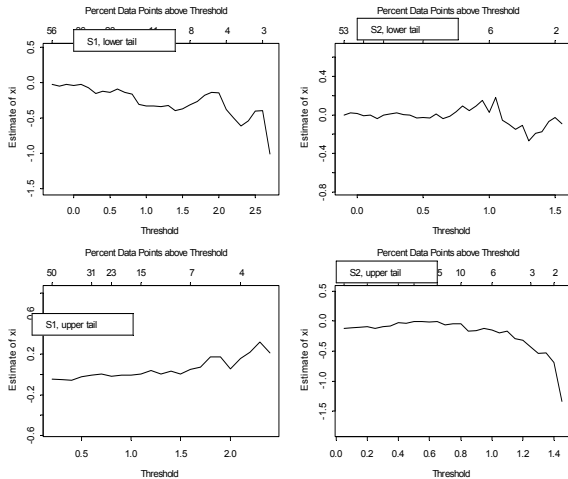


Figure 5. Hill 그래프(KOSPI, DJIA)

하지만 <그림 4>를 보면 임계치를 명확하게 정의하기 힘들다. 따라서 <그림 5>를 통해서 꼬리모수(ξ)가 안정적으로 움직이기 시작하는 부분을 동시에 고려하여 <표 3>과 같이 임계치 u 를 정하도록 한다.

Table 3. 임계치의 결정

	KOSPI		DJIA	
	upper	lower	upper	lower
threshold	+ 1.21%	-0.61%	+ 0.82%	-0.71%
%of data	13.16%	20.06%	9.47%	10.75%

임계치를 반영하여 반모수적방법으로 GPD의 꼬리부분을 추정하고 적합시킨 결과는 다음의 <표 4>와 같다. GPD모수를 아래와 같이 추정하고 실제 자료를 추정된 분포함수에 적합시킨 결과 대부분의 자료들은 <그림 6>에서 보듯이 실선에 잘 분포하고 있음을 알 수 있다.

Table 4. GPD 모수의 추정

	KOSPI		DJIA	
	upper	lower	upper	lower
ξ	0.0342	-0.1447	-0.1254	-0.0240
$\beta(u)$	0.6338	1.0438	0.4767	0.5222

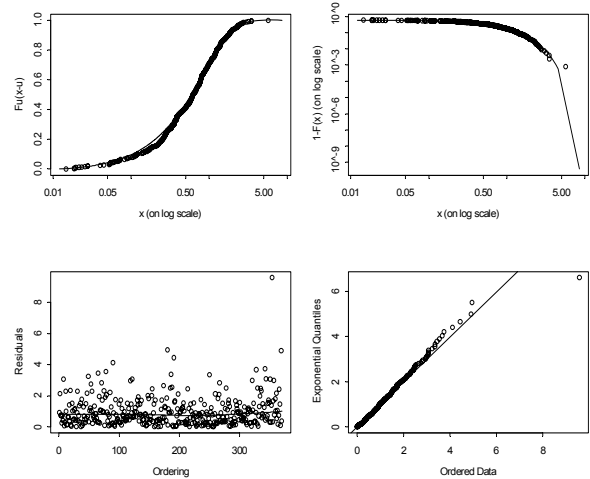


Figure 6. $x-u$ 의 분포함수(左·上)와 밀도함수(右·上) 잔차에 대한 산점도(左·下)와 QQ-plot(右·下)

4.2.2 Copula분석

<그림 7>은 각 국가 간 지수 수익률을 산점도로 표시한 것이다.

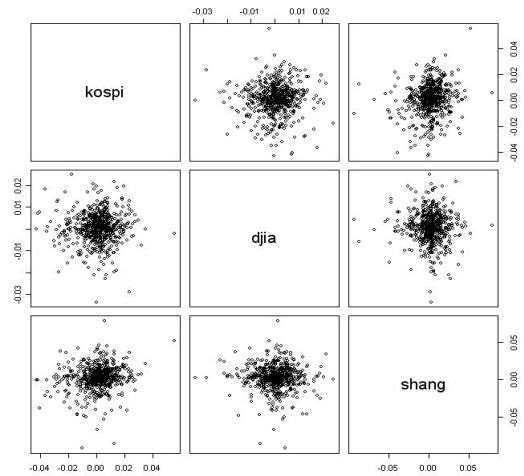


Figure 7. 수익률 산점도(KOSPI, DJIA, S.C.)

각 수익률 간 산점도는 극단적인 양의 수익률 부분과 왼쪽 하단의 극단적 음의 수익률 부분이 일반적인 타원형분포의 모습과는 상이한 것을 확인할 수 있다. 이는 극단적 움직임들이 상호 의존관계를 갖고 있음을 암시한다. 따라서 POT방법으로 추정된 각각의 한계분포들과 의존관계를 반영하는 Copula함수를 이용하여 결합분포함수를 추정할 것이다.

Table 5. 실증 분석에 사용할 Copula 함수의 종류

Elliptical	Archimedean	Extreme Value
Gaussian	(1) Frank	(4) Galambos
Student's t	(2) Clayton	(5) Husler-Reiss
	(3) Gumbel	(6) BB5

Gaussian Copula 함수의 경우에는 모형에 적용하기는 쉬우나 꼬리의존지수가 0으로 측정되기 때문에 적합하지 않다. Student's t-Copula의 경우에는 자유도 ν 가 작을수록 두터운 꼬리를 표현할 수 있지만 왼쪽과 오른쪽 꼬리의 비대칭성을 반영하지 못하는 단점이 있다. 하지만 Archimedean Copula 함수는 위의 Elliptical Copula 함수에 비해 훨씬 복잡하지만 생성자 함수 $\phi(u)$ 와 Copula 모수에 따라 비대칭성과 상호 간의 의존 정도를 반영할 수 있는 다양한 형태를 가진다. 그리고 Extreme Value Copula는 극단치분포를 설명하는데 유용하다고 알려져 있다. 따라서 본 연구에서는 실증 Copula 함수를 잘 설명할 수 있는 적정 Copula를 선정하기 위해서 <표 5>에서 제시한 Archimedean Copula 함수(1)~(3)와 Extreme Value Copula 함수(3)~(6)만을 고려하고자 한다.

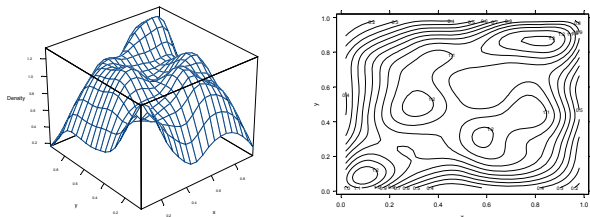


Figure 8. 실제자료에 근거한 결합밀도 함수의 조감도 및 등고선(KOSPI, DJIA)

<그림 8>은 한국과 미국지수에 대한 결합밀도 함수와 그의 등고선을 나타낸 그림이다. 양쪽 꼬리 부분의 봉우리가 높은 것은 상호 간 극단적인 움직임이 높음을 의미하는 것이며 선형 상관관계를 통한 의존관계의 설명이 적절하지 않음을 알 수 있다. 따라서 위의 결합밀도 함수를 가장 잘 적합할 수 있는 Copula 함수를 찾아내고 꼬리의존지수를 분석함으로써 체계적인 위험에 대한 극단적 의존관계의 영향을 파악할 수 있다.

4.2.3 결과 분석

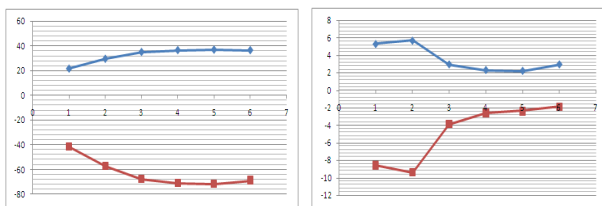


Figure 9. KOSPI/DJIA(左), KOSPI/S.C.(右)의 최우추정치(上), AIC검정값(下)

KOSPI와 DJIA는 Extreme Value Copula를 선택한 경우 (3)~(6), KOSPI와 S.C.는 Archimedean Copula를 선택한 경우 (1), (2) 최우추정치의 값이 높아지고 AIC값이 낮아진다. 즉 KOSPI와 DJIA는 극단적인 움직임에 의해서, KOSPI와 S.C.는 비대칭적인 움직임에 의해서 결합밀도 함수가 잘 표현되어 있음을 의미한다. 특히 각각은 Husler-Reiss Copula(5), Clayton Copula(2)가

적정 Copula로 선정되었는데 실제로 KOSPI와 S.C.의 가장 적합한 Copula 함수로 선택된 Clayton Copula는 비대칭적이고 음(-)의 극단적 의존관계가 양의 극단적 의존관계보다 큰 성질을 갖는 것으로 알려져 있다(Kjersti Aas, 2004).

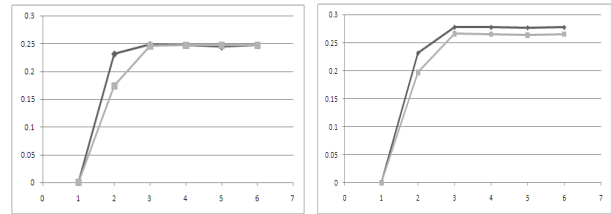


Figure 10. KOSPI/DJIA 꼬리의존관계의 변화 (검은 실선 : (-)꼬리의존지수, 회색실선 : (+)꼬리의존지수)

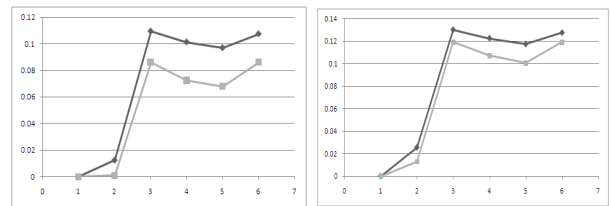


Figure 11. KOSPI/S.C. 꼬리의존관계의 변화

KOSPI와 DJIA간 음의 꼬리의존지수는 체계적 위험이 발생하기 이전의 경우보다 전체 구간에 대한 꼬리 의존지수가 약 3.06%P 증가하였고, 양의 꼬리의존지수는 약 1.85%P 증가한 것으로 나타났다. 하지만 KOSPI와 S.C.간 음의 꼬리의존지수는 1.30%P, 양의 꼬리의존지수 1.21%P 증가한 결과를 나타냈다. 실제로 체계적 위험이 발생한 관심 구간의 극단적 의존관계는 기간을 반영한 선형보간법을 이용하여 간접적인 방법으로 구하였다. 그 결과는 다음의 <표 6>과 같다.

Table 6. 기간 별 꼬리의존지수의 추정

		2003.01.03 ~2007.07.24	2003.01.03 ~2007.09.20	2007.07.25 ~2007.09.20
KOSPI DJIA	λ_l	24.6476%	27.7087%	75.0741%
	λ_u	24.6415%	26.4876%	55.0550%
KOSPI S.C.	λ_l	1.2525%	2.5451%	22.5460%
	λ_u	0.0869%	1.2970%	20.0211%

5. 결론

위험관리자는 시장 충격으로 인한 양 시장의 영향 관계를 검증하기 위해서 전체 자료에 대한 선형 상관정도가 아닌 극단적인 값들 간 의존 정도를 파악하는 것이 중요하다. 체계적인 위험이 발생했을 경우에는 수익률의 변동성이 증폭되고 꼬리 의존 정도 역시 상당히 강해지는데 특히 음의 수익률 간 꼬리

의존관계가 양의 수익률 간 꼬리의존관계보다 높게 측정된다. 이러한 현상은 한국과 중국시장 간의 관계에서 더욱 명확하게 나타나고 Clayton Copula 함수를 통해 모형화가 가능하였다. 반면에 한국과 미국시장은 Husler-Reiss Copula에 의해 잘 설명되어짐을 알 수 있다.

꼬리의존지수는 극단적 수익률이 같은 방향으로 움직일 극한확률, 즉 동조화의 정도로 해석할 수 있으므로 한국과 중국시장의 동조화는 미국과 동조화할 확률보다 낮은 수준으로 측정되긴 하였으나 증가폭은 미국과의 경우보다 훨씬 크다고 결론을 내릴 수 있다. 하지만 이를 근거로 중국과의 동조화가 강해지고 미국과의 탈동조화가 진행되고 있다고 주장하는 것은 무리가 있다. 동조화의 신호는 극단적 수익률의 움직임 뿐만 아니라 펀더멘털(기본적인 내재 가치를 나타내는 국가 기초 경제 여건)측면에서 고려해야할 사항이 더 많기 때문이다. 즉 펀더멘털에 기초한 극단적 움직임의 동조화를 모형화할 수 있다면 효율적인 자산배분을 통한 포트폴리오를 구성할 수 있고 체계적인 위험을 관리할 수 있을 것이다.

참고 문헌

- Beatriz Vaz de Melo Mendes. (2006), "A Bayesian analysis of clusters of extreme losses", *Applied Stochastic models in business and industry*, **22** 155-167.
- Carmona (2004), "Statistical Analysis of Financial Data in S-Plus", Springer.
- Chan, Ngai Hang and Hoi Ying Wong. (2005), "Simulation Techniques in Financial Risk Management", Wiley Inter Science.
- Edward W. Frees and Emiliano A. Valdez (1997), "Understanding Relationships Using Copulas", 32nd Actuarial Research Conference.
- Eric Zivot, and Jiahui, Wang. (2006), "Modelling Financial Time Series with S-PLUS 2nd edition", Springer.
- George Casella, and Roger, L. Berger. (2002), "Statistical Inference 2nd edition", Duxbury.
- Jun Yan. (2006), "Enjoy the Joy of Copulas", Preprint submitted to Elsevier Science.
- Kim, M.-J. and Shin, S. H. (2003). "Estimation and Simulation of Copula Function : An Application to Daily Korean Treasury and A-Rate Corporate Spot Rates", *Korean Journal of Futures and Options* **11**(2), pp. 103-131.
- Kim, T.-H. and Zhao, Huijing. (2006). "Extremal Dependence in Asia Pacific Exchange Markets", *The Korean Journal of Financial Management* **23**(1), pp.193-225.
- Kjersti Aas (2004.12), "Modelling the dependence structure of financial assets : A survey of four copulas", *Norwegian Computing Center* 19.
- Mario R. Melchiori. (2003), "Which Archimedean Copula is the right one?", (www.YieldCurve.com).
- Oh, Sekyung. and Moon, Seong Ju. (2006.4), Comparative Analysis of Portfolio Risk Measures based on EVT-Copula Approach during Financial Crises *증권학회지* 제35권 3호.
- Paul Embrechts, Alexander Mcneil, and Daniel, Straumann. (1999,7), "Correlation and Dependence in Risk Management : Properties and Pitfalls", Working paper, Risklab, ETH Zurich.
- Paul Embrechts, Claudia, Kluppelberg. and Thomas, Mikosch. (1997), "Modelling Extremal Events for Insurance and Finance", Springer.
- Paul Embrechts, F. Lindskog, and McNeal, A. J. (2001), "Modelling Dependence with Copulas and Applications to Risk Management", Working paper, Risklab, ETH Zurich.
- Roger B. Nelson. (2006), "An Introduction to Copula 2nd edition", Springer.