응용통계연구 (2011)
24(3), 523–533

DOI: 10.5351/KJAS.2011.24.3.523

Copula 함수를 이용한 이변량분포의 VaR 추정

홍종선[1] · 이재형[2]

[1]성균관대학교 통계학과, [2]성균관대학교 응용통계연구소

(2011년 1월 접수, 2011년 3월 채택)

요 약

위험관리수단으로 시장위험을 정확하게 측정하는 방법 중의 하나로 VaR를 선호한다. 현실생활에서는 단일분포가 아닌 두 개 이상의 다변량분포에 대한 VaR를 추정해야 하는 경우가 많다. 이런 경우에는 VaR를 추정하기 위해 다변량분포를 고려해야 한다. 본 연구는 확률변수들의 종속적 구조를 파악하고 비정규성의 특성을 갖는 다변량 분포함수를 생성하기 위하여 Copula 함수를 사용한다. 여러 산업의 수익률분포에 적합한 Clayton, Gumbel, Frank Copula 함수가 포함된 Archimedean Copula 함수를 추정하여 다변량 수익률 분포함수를 결정하고 이에 대응하는 VaR를 유도한다. 국내의 두 산업체의 자료를 실증예제로 하여 세 종류의 Copula 함수의 모수를 추정하고 이에 대응하는 이변량 분포로부터 VaR와 각각의 주변 분포의 VaR를 구한다. 실제의 VaR를 기준으로 기존 방법으로 구한 VaR와 비교 분석하여 추정의 정확성을 토론한다.

주요용어: 생성함수, 수익률, 신뢰수준, 위험, 종속성.

1. 서론

1990년대 이후 각국의 금융 및 자본시장이 개방되고 자유화됨에 따라 국제금융시장이 빠르게 통합되고 있다. 특히 파생상품(derivative product)시장이 급속도로 성장하고 외환위기의 영향으로 자산가격이 급격하게 변동함에 따라 시장위험을 어떻게 관리할 것인가에 대한 관심이 세계적으로 고조되고 있다. 이에 따라 국제결제은행(Bank for International Settlements; BIS)에서도 신 BIS 기준을 통해 금융기관이 독자적으로 개발한 내부모형을 사용하는 것을 허용함으로써 금융기관들로 하여금 시장위험을 정확하게 측정하려는 노력을 활성화시키고 있고, 이러한 위험관리수단으로 금융기관들이 가장 선호하는 기법이 VaR(Value at Risk; 리스크값)이다.

VaR는 금리, 주가, 환율 등의 기초적 시장가격들에 대한 주어진 신뢰수준에서 목표기간에 걸쳐 발생할 수 있는 최대손실금액(largest losses amount)을 말한다 (Jorion, 1997). 현재까지 VaR를 추정하기 위하여 많은 방법론과 다양한 기법들이 제시되었다. VaR를 추정하는 방법에는 세 가지의 측면으로 나누어 생각할 수 있다. 첫 번째는 모수적 방법과 비모수적 방법으로 나누고, 두 번째는 전통적인 방법과 극단치 이론(extreme value theory)을 이용한 방법으로 나눈다. 세 번째는 미래 현금흐름(cash flow)의 분포 측면에서 부분가치 평가법(local valuation)과 완전가치 평가법(full valuation)으로 나누어 볼 수 있다.

수익률분포의 VaR를 추정하기 위해 실제 수익률이 가지고 있는 분포를 통계적인 밀도함수의 분포로 적합시킬 필요가 있다. 즉 임계값을 산출하기 위해 적합시키려는 분포 중 어떤 것이 가장 좋은 분포인지

[1]교신저자: (110-745) 서울시 종로구 명륜동 3-53, 성균관대학교 경제학부 통계학전공, 교수.
E-mail: cshong@skku.ac.kr

알아보려는 노력이 계속 논의되어 왔다. 특히 지금까지의 방법은 금융자산의 수익률에 대한 분포를 정규분포로 가정하여 VaR를 추정하는 모수적 방법에 크게 의존해 왔으나 VaR 실증분석의 결과에 의하면 금융자산의 수익률분포는 정규분포보다 다른 분포들에 더 적합하다는 주장이 제기되었다. 예를 들어 Zangari (1996)는 미국의 주식을 이용한 금융자산들에 대한 수익률분포의 경우 꼬리부분이 더 두꺼운 것으로 보고하고 있으며, Li (1999)는 외환시장에서 이루어지는 환거래가 정규분포를 따르지 않는다는 사실을 보고한 바 있다. 따라서 다른 여러가지 분포들 예를 들어 치우친 t(skewed t) 분포와 치우친 라플라스(skewed laplace) 분포, 정규혼합(normal mixture) 분포 등의 대체분포를 통하여 실제 수익률분포를 설명하려 했다. 그리고 홍종선과 권태완 (2010)은 실제 수익률분포에 정규혼합분포가 높은 적합도를 나타내는 결과를 보였다.

현실생활에서는 어떤 특정한 회사의 VaR를 추정하는 것 보다는 그 회사가 포함된 몇 개의 산업 내지는 국내 전체의 산업의 VaR를 추정하는 경우가 많다. 즉 단일분포의 VaR가 아닌 두 개 이상의 다변량 분포에 대한 VaR를 추정해야한다.

Copula 함수는 여러 확률변수들 사이의 복잡한 종속성(dependency) 구조를 파악하기 위한 방안으로서 Sklar (1959)에 의해 제시되었고, 개별 밀도함수 사이의 관계와 결합분포와의 관계를 설명하는 방법으로 확률변수들의 상관성을 설명할 수 있는 확률 도구로 활용되었다. Copula 함수가 금융분야에 응용되기 시작한 것은 1999년이 되어서 활용되었다. 재무금융 분야에서 복잡한 종속성 구조를 고려하기 위해 많이 활용되며 최근에 활발히 발전되었다 (Nelson, 2006; Umberto와 Walter, 2004; 등 참조).

Copula 함수는 대칭형 Copula 함수와 비대칭형 Copula 함수로 구분되는데, 대칭형 함수로는 Gaussian Copula, student t Copula, Frank Copula 함수가 있으며, 비대칭형 함수로는 Clayton Copula와 Gumbel Copula 함수가 대표적이다. 또한 Copula 함수의 종류를 타원형(Elliptical) Copula, 아르키메디안(Archimedian) Copula로 구분되는데, 타원형 Copula 에서는 상관계수행렬을 모수로 가지는 Gaussian Copula와 상관계수행렬과 자유도를 모수로 가지는 student t Copula 함수가 있다. 그리고 아르키메디안 Copula에는 Clayton Copula, Gumbel Copula, Frank Copula 함수가 종속적 구조를 표현하는 생성함수(generator)를 갖고 있다. 아르키메디안 Copula는 분석이 쉬울 뿐만 아니라 여러가지 생성함수로 표현되기 때문에 Copula 함수 중에서 중요하다. 재무금융자료에 대해 Copula 함수의 적용 방법에 있어서 어떠한 Copula 함수를 적용하느냐에 따라 시뮬레이션 결과가 다르게 나올 수 있으므로 Copula 함수의 선택은 매우 중요하다. Copula 함수를 이용한 VaR 추정 연구로는 Longin (2001), Bae와 Karolyi (2003), Breymann (2003) 등의 연구 외에 한국의 주가 수익률에 대한 연구로는 student t Copula를 이용한 김명직과 신성환 (2003)과 Gaussian과 student t Copula 함수를 이용한 황수영 (2005) 등이 있다.

본 연구에서는 다변량 수익률자료에 적합한 세 종류의 아르키메디안 Copula 함수를 추정하고, 추정한 Copula 함수를 이용하여 이변량 자료의 VaR를 추정한다. Copula 함수를 이용해 VaR를 추정하면 여러 종목을 통합한 산업의 VaR를 측정할 수 있고, 추정된 VaR를 통해 투자자는 다양한 정보를 가지고 투자를 할 수 있다. 또한 Copula 함수를 이용하여 통합적 산업에 대한 VaR 뿐만 아니라 각각의 개별 산업 수익률분포에 대한 VaR를 추정할 수 있다.

본 연구의 2절에서는 Copula 함수의 기본적인 설명과 아르키메디안 Copula 함수에 포함된 Clayton, Gumbel, Frank Copula 함수와 각각의 생성함수에 대하여 소개하고 함수에 포함된 모수 추정방법에 대하여 설명한다. 3절에서는 Copula 함수를 이용한 수익률분포의 적합과 VaR 추정방법을 제안하는데 우선 이변량인 경우에 Copula 함수의 확률난수 추출방법과 이변량 분포로부터 VaR 그리고 이에 대응하는 단일 분포의 VaR를 추정하는 방법을 소개한다. 4절에서는 한국의 주식수익율 자료를 수집한 실증예제를 통해 세 종류의 Copula 함수의 모수를 추정하고 이에 대응하는 이변량 분포로부터 VaR와 각각의

주변 분포의 VaR를 추정하여, 실제의 VaR를 기준으로 기존의 다른 방법으로 구한 VaR들과 비교 분석한다. 마지막으로 5절에서 결론을 유도한다.

2. Copula 함수와 모수추정

2.1. Copula 함수

Copula 함수는 여러 확률변수들 사이의 종속성 구조를 고려하면서 다변량 누적분포함수를 추정하는데 유용한 방법이다. 대부분의 재무자료의 경우 분포 형태가 비대칭적이면서 꼬리부분이 두터운 비정규성의 형태를 가지며 또한 극단값에 동일한 방향을 나타내는 경향이 강하다. 따라서 이러한 분포의 꼬리 부분을 다루는 VaR의 경우에 상관계수에 의해 종속성 구조를 반영하기 보다는 Copula 함수를 사용하여 그 종속성 구조를 파악하는 것이 타당하다.

Copula 함수는 균일분포(Uniform distribution)를 따르는 확률변수들의 결합분포함수로 정의한다. 즉 확률변수 $U_1, \ldots, U_n$이 구간 $[0,1]$에서의 균일분포를 따른다면, Copula 함수는 다음과 같다.

$$C(u_1, \ldots, u_n) = P(U_1 \leq u_1, \ldots, U_n \leq u_n). \tag{2.1}$$

Copula 모형을 연구하는데 핵심적인 역할을 하는 Sklar (1959)의 정리는 다음과 같다. 확률변수 $X_1, \ldots, X_n$의 주변분포함수 $F_1(x), \ldots, F_n(x)$가 존재할 때, 확률변수들의 누적분포함수 $F(x_1, \ldots, x_n)$에 대하여 n차원의 Copula 함수 C가 존재하며 다음이 성립된다.

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \tag{2.2}$$

그리고 $F_1(x), \ldots, F_n(x)$이 연속형인 경우에는 Copula 함수 C는 유일하게 존재하며 다음과 같이 주어진다.

$$C(u_1, \ldots, u_n) = C\left(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)\right), \tag{2.3}$$

여기서 $u_i \in [0,1]$이고 $F_i^{-1}(\cdot)$은 $F_i(\cdot)$의 역함수이다 $(i = 1, \ldots, n)$.

Sklar의 정리를 통해서 결합분포함수는 연속적인 주변분포함수와 종속구조로 분리될 수 있고, 종속구조는 Copula 함수로 표현된다는 것을 알 수 있다. 즉 Copula 함수는 주변분포함수들과 결합분포함수를 연결시키는 역할을 하며, 변수간의 의존구조에 관한 모든 정보를 포함하고 있다. Copula 함수를 사용하면, 여러 종목의 금융자산 수익률들의 결합확률분포를 모형하는데 훨씬 쉽고 편리하며 다변량 확률분포를 구현할 수 있는 장점이 있다.

2.2. Copula 함수의 종류

본 연구에서 사용하는 아르키메디안 Copula 함수는 다음과 같이 정의한다 (Genest와 MacKay, 1986).

$$C(u_1, \ldots, u_n) = \begin{cases} \varphi^{-1}[\varphi(u_1) + \cdots + \varphi(u_n)], & \text{만약 } \sum_{i=1}^n \varphi(u_i) \leq \varphi(0), \\ 0, & \text{그외}, \end{cases} \tag{2.4}$$

여기서 $\varphi(u)$는 Copula의 생성자(generator)라는 함수로 모든 $0 \leq u \leq 1$에 대하여 $\varphi(1) = 0$, $\varphi'(u) < 0$, 그리고 $\varphi''(u) > 0$를 만족하는 연속함수이다. 이변량 $u_1, u_2 \in [0,1]$를 가정하고 하나의 모수 θ를 갖는 아르키메디안 Copula 중에서 본 연구에서 사용할 Copula 함수와 이에 대응하는 생성함수는 표 2.1과 같다.

표 **2.1.** 아르키메디안 Copula 함수의 종류

종류	$C_\theta(u_1,u_2)$	θ의 범위	생성함수 $\varphi(u)$
Clayton	$\max\left(\left[u_1^{-\theta}+u_2^{-\theta}-1\right]^{-\frac{1}{\theta}},0\right)$	$[-1,\infty)$	$u^{-\theta}-1$
Gumbel	$\exp\left\{-\left[(-\ln u_1)^\theta+(-\ln u_2)^\theta\right]^{\frac{1}{\theta}}\right\}$	$[1,\infty)$	$(-\ln u)^\theta$
Frank	$-\frac{1}{\theta}\ln\left(1+\frac{(e^{-\theta u_1}-1)(e^{-\theta u_2}-1)}{e^{-\theta}-1}\right)$	$[1,\infty)$	$\ln\left(\frac{\exp(\theta u)-1}{\exp(\theta)-1}\right)$

표 **2.2.** Copula 함수의 모수 추정식

종류	생성함수 $\varphi(u)$	모수 추정식
Clayton	$u^{-\theta}-1$	$\tau_\theta=\frac{\theta}{\theta+2}$
Gumbel	$(-\ln u)^\theta$	$\tau_\theta=\frac{\theta-1}{\theta}$
Frank	$\ln\left(\frac{\exp(\theta u)-1}{\exp(\theta)-1}\right)$	$\tau_\theta=1+\frac{4(D_1(\theta)-1)}{\theta}$

여기서 $D_1(\theta)=1/\theta\int_0^\theta u/\{\exp(u)-1\}du$.

2.3. Copula 함수의 모수추정

일반적으로 Copula 함수의 모수 추정방법으로는 모수적 추정법과 비모수적 추정법이 있다. 모수적 추정법에 대해서 살펴보면 최대가능도추정법을 기반으로 세가지 방법이 있다. 첫 번째 방법은 MLE(Maximum Likelihood Method) 방법인데 Copula 함수의 모수를 주변확률분포의 모수와 함께 추정한다. 두 번째 방법은 IFM(Inference Function for Margins) 방법으로 주변확률분포의 모수와 Copula 함수의 모수를 분리하여 두 단계에 걸쳐서 추정하는 방법이다. 세 번째 방법은 CML(Canonical Maximum Likelihood) 방법으로 IFM 방법과 다른 점은 각각 주변확률분포함수에 대해 모수적 분포함수를 가정하지 않고 경험적 분포함수로 추정한 후에 Copula 함수의 모수벡터를 추정한다. 이 방법은 개별 주변확률분포의 형태를 모를 경우에 유용한 방법이 될 수 있다. 또한 비모수적으로는 순서통계량을 사용하는 방법이 있다 (김명직과 신성환, 2003).

Copula 함수의 모수 θ의 추정은 매우 복잡한 과정을 갖고 있으나 아르키메디언 Copula의 경우에는 종속성 측도로 잘 알려진 비모수 통계량인 켄달(Kendall)의 τ_θ를 이용하여 Copula 함수의 모수를 추정하는 방법인 calibration 방법을 이용하면 쉽게 추정할 수 있다. 켄달의 τ_θ와 아르키메디언 Copula의 생성함수와의 관계는 다음과 같다 (Nelson, 2006).

$$\tau_\theta=1+4\int_0^1\frac{\varphi(u)}{\varphi'(u)}du. \tag{2.5}$$

아르키메디언 Copula 함수의 종류에 따라 대응하는 생성함수와 켄달의 τ_θ값과의 관계식을 정리한 표 2.2를 이용하여 표본을 통하여 켄달의 τ_θ값을 얻은 후 모수 θ를 추정한다.

3. Copula 함수를 이용한 이변량분포의 적합과 VaR

3.1. 이변량 Copula 함수의 확률난수 추출방법

Copula 함수를 이용한 VaR를 추정하기 위해선 먼저 추정한 모수의 정보를 포함하는 Copula 함수의 난수들을 추출하는 시뮬레이션 과정이 필요하다. 일반적인 알고리즘은 확률변수 $\mathbf{U}=(U_1,\ldots,U_n)$의 조

표 3.1. Copula 함수의 모수 추정식

종류	$C_{2\|1}(u_1,u_2)$	$u_2 = C^{-1}_{2\|1}(v_2,u_1)$
Clayton	$\left(\dfrac{u_1^{-\theta}u_2^{-\theta}-1}{u_1^{-\theta}}\right)^{-\frac{1}{\theta}-1}$	$\left\{v_1^{-\theta}\left(v_2^{-\frac{\theta}{\theta+1}}-1\right)+1\right\}^{-\frac{1}{\theta}}$
Gumbel	$\dfrac{\exp\left(-\left((-\ln u_1)^{\theta}+(-\ln u_2)^{\theta}\right)^{\frac{1}{\theta}}\right)}{\exp\left(\left(-(-\ln u_1)^{\theta}\right)^{\frac{1}{\theta}}\right)}$	$\exp\left(\left((\ln v_1 v_2)^{\theta}-(\ln v_1)^{\theta}\right)^{\theta}\right)^{\frac{1}{\theta}}$
Frank	$\dfrac{\exp(-\theta u_1)(\exp(-\theta u_2)-1)}{\exp(-\theta)-1+(\exp(-\theta u_1)-1)(\exp(-\theta u_2)-1)}$	$-\dfrac{1}{\theta}\ln\left\{1+\dfrac{v_2(1-e^{-\theta})}{v_2(e^{-\theta v_1}-1)-e^{-\theta v_1}}\right\}$

건부 분포함수와 그의 역함수를 사용한다. 예를 들어 이변량($n=2$)인 경우에 $\mathbf{U}=(U_1,U_2)$의 확률난수는 다음과 같이 추출한다.

1. 상호 독립인 균일분포로부터 확률난수(random number) v_1과 v_2를 추출한다.
2. $u_1 = v_1$으로 둔다.
3. $u_2 = C^{-1}_{2|1}(v_2,u_1)$을 계산한다, 여기서 $C_{2|1}(u_1,u_2) \equiv \Pr(U_2 \le u_2 | U_1 = u_1) = \partial C(u_1,u_2)/\partial u_1$.

반복추출한 (u_1,u_2)는 Copula 함수로부터의 이변량 자료가 된다. 세 종류의 Copula 조건부 분포함수와 이에 대응하는 역함수는 표 3.1과 같다.

3.2. 이변량 분포의 VaR 추정

두 개의 금융자산 A와 B로 구성된 포트폴리오를 고려하자. 확률변수 X와 Y는 금융자산 A와 B의 투자기간 동안의 각 수익률을 나타낸다고 하자. p_0는 초기투자가치이고 p_1은 투자기간 말의 투자가치라고 하면 다음의 관계가 성립한다.

$$p_1 = p_0\left(w_1 e^{X} + w_2 e^{Y}\right), \tag{3.1}$$

여기서 w_1과 w_2는 금융자산 A와 B의 각 투자비중을 나타낸다. 따라서 투자기간 동안에 포트폴리오 X와 Y의 수익률은 다음과 같이 구한다 (여성칠, 2006).

$$R_p = \ln\left(\frac{p_1}{p_0}\right) = \ln\left(w_1 e^{X} + w_2 e^{Y}\right). \tag{3.2}$$

신뢰수준 $100(1-\alpha)\%$에서 VaR를 이변량 자료의 수익률 분포함수의 경우로 표현하면 다음과 같이 주어진다.

$$\mathrm{VaR}_{\alpha} = p_0 r_{\alpha}, \tag{3.3}$$

여기서 $\alpha = P(R_p \le -r_{\alpha})$, 즉 $r_{\alpha} = -F^{-1}_{R_p}(\alpha)$이다.

X와 Y의 수익률 결합확률밀도함수를 $f(x,y)$ 그리고 $F_X(x)$, $F_Y(y)$는 X와 Y 각각의 주변분포함수라고 할 때, 이변량 U_1과 U_2의 Copula 함수를 이용해 α번째 분위수 $(u_{1\alpha},u_{2\alpha})$와의 관계는 다음과 같다.

$$\begin{aligned}\alpha &= C(u_{1\alpha},u_{2\alpha}) = P\left[U_1 \le u_{1\alpha}, U_2 \le u_{2\alpha}\right] \\ &= P\left(R_p \le -r_{\alpha}\right)\end{aligned} \tag{3.4}$$

$$= \iint_{\left\{(x,y): w_1 e^X + w_2 e^Y \le e^{-r_\alpha}\right\}} f(x,y)dxdy$$
$$= \iint_{\left\{(u_1,u_2): w_1 e^{F_X^{-1}(u_{1\alpha})} + w_2 e^{F_Y^{-1}(u_{2\alpha})} \le e^{-r_\alpha}\right\}} c(u_1,u_2)du_1du_2,$$

여기서 $u_1 = F_X(x), u_2 = F_Y(y)$이다. $(u_{1\alpha}, u_{2\alpha})$는 다음을 만족하므로

$$w_2 e^{F_Y^{-1}(u_{2\alpha})} = e^{-r_\alpha} - w_1 e^{F_X^{-1}(u_{1\alpha})}, \tag{3.5}$$

X와 Y의 수익률 분포함수에서 α번째 분위수 r_α는 다음과 같이 구하며 식 (3.3)을 이용하여 VaR_α를 계산한다.

$$r_\alpha = -\ln\left(w_1 e^{F_X^{-1}(u_{1\alpha})} + w_2 e^{F_Y^{-1}(u_{2\alpha})}\right). \tag{3.6}$$

만약 $w_1 = w_2 = 0.5$라면, $r_\alpha = -\ln(e^{F_X^{-1}(u_{1\alpha})} + e^{F_Y^{-1}(u_{2\alpha})}) + \ln 2$이다.

따라서 본 연구에서는 위에서 제시한 방법을 이용하여 세 종류의 아르키메디안 Copula 함수를 이용하여 Copula 함수의 모수를 추정한 뒤, 제시한 몬테칼로 방법을 통해 다양한 분위수 1%, 2.5%, 5%, 7.5%, 10%의 VaR를 산출한다. 세 종류의 Clayton, Gumbel, Frank Copula 함수마다의 밀도함수의 특성에 따라 난수를 추출하는 과정이 각각 다르나, 난수를 추출한 후 VaR를 구하는 과정은 세가지 Copula 함수 모두 동일하다.

3.3. 단일 분포의 VaR 추정

다변량 Copula 함수를 이용하여 주변(margin) Copula 함수를 구할 수 있고 이를 통해 개별 수익률분포의 VaR를 추정할 수 있다. 이변량 U_1과 U_2의 Copula 분포함수 $C(u_1,u_2)$와 밀도함수 $c(u_1,u_2)$ 그리고 주변 Copula 분포함수 $C(u)$와 밀도함수 $c(u)$의 관계는 다음과 같다.

$$c(u_1) = \frac{d}{u_1}C(u_1), \quad \text{여기서} \quad C(u_1) = C(u_1, 1),$$
$$= \int_{u_2} c(u_1,u_2)du_2, \quad \text{여기서} \quad c(u_1,u_2) = \frac{d^2}{du_1du_2}C(u_1,u_2).$$

확률변수 U_1의 주변 Copula 함수값이 α일 때의 $u_{1\alpha}$는 $C(u_{1\alpha}) = \alpha$로 표현되므로

$$\alpha = C(u_{1\alpha}) = P(R_p \le -r_\alpha) \tag{3.7}$$
$$= \int_{\left\{(u_1): e^{F_X^{-1}(u_{1\alpha})} \le e^{-r_\alpha}\right\}} c(u_1)du_1$$
$$= C(F_X(-r_\alpha)).$$

따라서 r_α는 다음과 같이 구한다.

$$r_\alpha = -F_X^{-1}(u_{1\alpha}). \tag{3.8}$$

다음으로 확률변수 U_2의 주변 Copula 함수 $C(u_{2\alpha}) = \alpha$를 만족하는 $u_{2\alpha}$에 대하여

$$\alpha = C(u_{2\alpha}) = P(R_p \le -r_\alpha) \tag{3.9}$$
$$= \int_{\left\{(u_2): e^{F_Y^{-1}(u_{2\alpha})} \le e^{-r_\alpha}\right\}} c(u_2)du_2$$
$$= C(F_Y(-r_\alpha)).$$

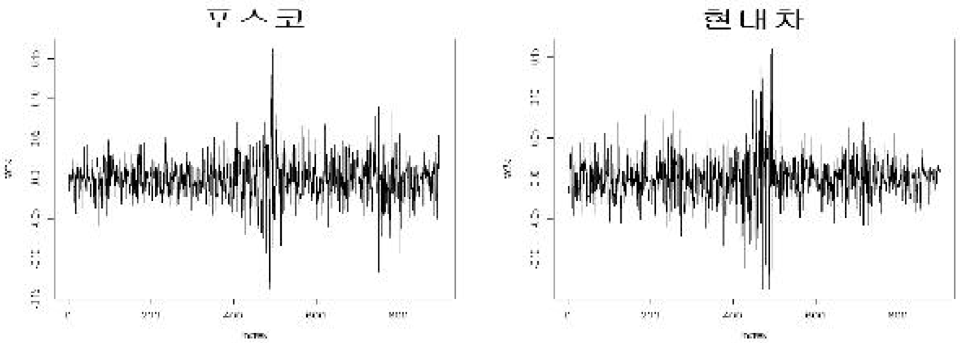

그림 4.1. 포스코와 현대차의 시계열 도표

표 4.1. 수익률자료의 기초통계량

기초통계량	포스코	현대차
평균	−0.0003732	−0.0008855
표준편차	0.02791	0.02978
왜도	0.175	0.274
첨도	3.529	4.461

그러므로 r_α는 다음과 같이 구하며, 식 (3.1)을 이용하여 개별 수익률에 대한 VaR$_\alpha$를 계산한다.

$$r_\alpha = -F_Y^{-1}(u_{2\alpha}). \tag{3.10}$$

3.2절에서 세 종류의 이변량 아르키메디안 Copula 함수의 다양한 분위수를 계산하여 구한 VaR와 이에 대응하는 주변 Copula 함수로부터 VaR를 산출하여 비교 분석한다.

4. 실증분석

4.1. 수익률 분포의 기초 통계량

실증분석에서 사용한 자료는 2007월 10월 1일부터 2010년 9월 30일까지 총 899개의 일일 주식 종가자료를 사용하였다. 그림 4.1은 포스코와 현대차의 일별 로그수익률에 대한 시계열을 보여주고 있다. 로그수익률 자료의 기초통계량을 정리한 표 4.1을 살펴보면, 평균이 모두 0에 가까우며, 표준편차는 각각 0.02791, 0.02978로 다소 적은 폭의 변화가 있음을 알 수 있다. 포스코 자료 수익률의 왜도는 0.175를 가지며 첨도는 3.529를 가지고 있음을 알 수 있고 현대차의 경우 현대차 자료 수익률의 왜도는 0.274이고 첨도는 4.461로 나타났음을 알 수 있다. 두 수익률 분포 모두 정규분포에 비하여 약간 오른쪽으로 치우쳐져 있으며 조금 높은 초과첨도를 가지고 있음을 보여준다. 이는 포스코와 현대차 자료의 수익률이 비대칭적이며 다소 두꺼운 꼬리분포를 가지고 있음을 나타낸다. 또한 두 분포의 정규성 검정을 실시해 본 결과 두 분포 모두 정규성을 만족하지 못하는 것으로 나타났다 (Jarque-Berra 그리고 Shapiro-Wilk 검정통계량의 p-값 ≈ 0). 따라서 이 자료에 정규분포를 가정한 기존의 VaR 추정방법은 적절하지 못하다. 그러므로 4.2절에서는 본 연구에서 제안한 방법을 사용한다.

4.2. 이변량의 VaR 추정

4.1절에서의 실증예제에 대하여 아르키메디안 Copula 함수에 속하는 Clayton, Gumbel, Frank Copula 함수를 사용한다. Copula 함수의 모수를 추정하기 위하여 3절에서 제시한 calibration 방법을 사용하

표 **4.2.** Copula 함수의 모수 추정값

Copula 함수	θ
Clayton	0.538
Gumbel	1.269
Frank	1.980

표 **4.3.** 이변량 VaR

	VaR_1	$\mathrm{VaR}_{2.5}$	VaR_5	$\mathrm{VaR}_{7.5}$	VaR_{10}
분산-공분산	11.320	7.513	5.961	5.407	4.627
Clayton	11.431	9.023	6.567	5.701	4.554
Gumbel	9.536	7.749	5.829	5.688	5.102
Frank	9.277	7.417	7.016	4.970	4.113

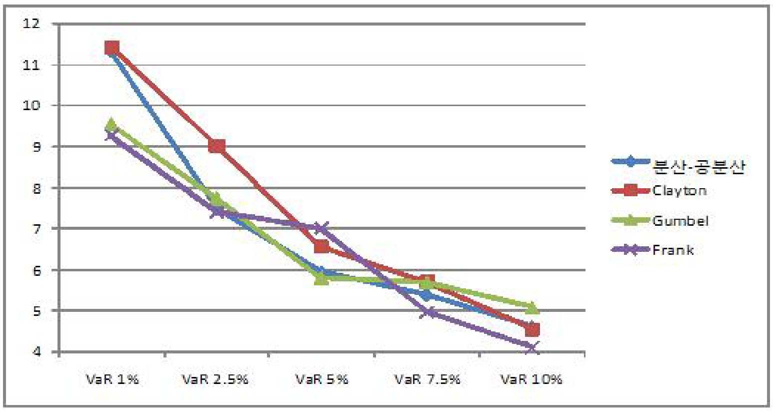

그림 **4.2.** 이변량 VaR 비교

며, 개별 주변분포함수는 student t 분포를 가정한다. Copula 함수를 이용하여 추정한 VaR와 많이 사용하는 기존의 방법인 분산-공분산 방법으로 산출한 VaR를 비교하고자 한다. 분산-공분산 방법을 사용한 이변량 VaR는 다음과 같이 추정한다 (Jorion, 1997).

$$\mathrm{VaR}_\alpha = \sqrt{\mathrm{VaR}_{\alpha x}^2 + \mathrm{VaR}_{\alpha y}^2 + 2\rho_{x,y}\mathrm{VaR}_{\alpha x}\mathrm{VaR}_{\alpha y}}, \tag{4.1}$$

여기서 $\mathrm{VaR}_{\alpha x}$, $\mathrm{VaR}_{\alpha y}$는 각각 신뢰수준 $100(1-\alpha)\%$에서 수익률 확률변수 X와 Y의 VaR로 $\mathrm{VaR}_{\alpha x} = p_0^x(E(R_p^x) + r_\alpha^x)$와 $\mathrm{VaR}_{\alpha y} = p_0^y(E(R_p^y) + r_\alpha^y)$으로 정의한다. 포스코와 현대차의 수익률변수의 종속성을 나타내는 τ_θ값이 0.212이므로 식 (2.5)를 사용하여 추정한 세 종류의 Copula 함수의 모수값은 표 4.2와 같다. 여기서 모수 θ와 τ_θ의 값은 작지 않은 값으로 포스코와 현대차의 수익률 사이에 종속관계를 갖고 있다고 판단할 수 있다.

추정된 모수를 바탕으로 각각의 Copula 함수에 대하여 3.1절에서 제안한 방법으로 표본크기가 899인 확률난수를 생성하여 α가 1%, 2.5%, 5%, 7.5%, 10%일 때 $C(u_{1\alpha}, u_{2\alpha}) = \alpha$를 만족하는 $(u_{1\alpha}, u_{2\alpha})$를 구한다. 그리고 식 (3.6)을 이용하여 이에 대응하는 r_α값을 계산하고 식 (3.3)의 VaR를 추정한다. 세 종류의 Copula 함수를 이용하여 추정한 VaR의 결과는 표 4.3에, α에 대응하는 VaR를 그림 4.2에 나타내었다.

표 **4.4.** 원자료와 정규혼합분포를 이용한 VaR

	분포	VaR_1	$\text{VaR}_{2.5}$	VaR_5	$\text{VaR}_{7.5}$	VaR_{10}
포스코	원자료	7.360	4.941	4.012	3.650	3.172
	정규혼합	7.936	5.490	4.229	3.584	3.134
현대차	원자료	8.349	5.546	4.336	3.885	3.324
	정규혼합	9.970	7.064	4.380	3.174	2.599

표 **4.5.** Copula 함수를 이용한 VaR

	Copula 함수	VaR_1	$\text{VaR}_{2.5}$	VaR_5	$\text{VaR}_{7.5}$	VaR_{10}
포스코	Clayton	5.174	4.084	2.972	2.580	2.061
	Gumbel	4.316	3.507	2.638	2.574	2.309
	Frank	4.199	3.357	3.176	2.249	1.861
현대차	Clayton	5.520	4.357	3.171	2.753	2.199
	Gumbel	4.605	3.742	2.815	2.747	2.464
	Frank	4.480	3.581	3.388	2.400	1.986

표 4.3과 그림 4.2를 통하여 본 연구에서 제안한 Copula 함수를 이용하는 방법과 분산-공분산 방법으로 추정한 VaR를 비교한 결과, 수익률분포의 극단치 부분인 α가 1%, 2.5% 부분에서는 차이가 있었지만, α가 커질수록 차이를 보이지 않고 비슷한 값으로 수렴되는 것을 알 수 있다. 이것은 분산-공분산 방법이 α가 작을 때 Copula 함수를 이용할 때 보다 위험을 과대 추정한다는 결과를 확인할 수 있다 (여성칠, 2006). 또한 각각의 α에 대한 VaR의 변동의 추이도 분산-공분산 방법보다 Copula 함수를 이용한 VaR가 안정적인 값을 갖는다는 것을 파악할 수 있다.

본 연구에서 토론한 분산-공분산 방법과 Copula 함수를 이용하는 방법은 모두 이변량 이상의 다변량 분포에서 VaR를 추정하는 방법이며 원자료인 다변량에서 진실(true)의 VaR는 알려져 있지 않으므로 어느 방법이 우월하다고 판단하기 어렵다. 그러나 단변량 분포에서는 비교 가능함으로 4.3절에서는 이변량으로부터 주변분포를 유도한 다음 개별 VaR를 추정하고 진실 VaR와 비교분석한다.

4.3. 단변량 VaR 추정

본 연구에서 채택한 원자료와 정규혼합분포에서 현대차와 포스코의 VaR를 산출한 값을 표 4.4에 나타내었다. 정규혼합분포의 경우 포스코, 현대차의 수익률을 정규혼합분포에 적합시켜 모수를 추정하고 정규혼합분포의 임계값과 표준편차와의 곱으로 개별 수익률의 VaR를 구하였다 (홍종선과 권태완, 2010).

각 주변분포의 Copula 함수를 이용하는 방법에서는 신뢰수준 α에서 각각 1%, 2%, 5%, 7.5%, 10%에 대응하는 $C(u_{1\alpha}) = \alpha$를 만족하는 $u_{1\alpha}$에 대하여 임계값을 $r_\alpha = -F_X^{-1}(u_{1\alpha})$를 산출하여 표준편차와의 곱으로 수익률의 VaR를 구하고, 동일한 방법으로 $u_{2\alpha}$에 대하여 임계값 $r_\alpha = -F_Y^{-1}(u_{2\alpha})$를 산출하여 표준편차와의 곱으로 수익률의 VaR를 추정한 결과를 표 4.5에 정리하였다.

본 논문에서 제안한 Copula 함수를 이용하여 추정한 VaR를 정규혼합분포를 이용하여 추정한 VaR와 원자료의 VaR를 비교한 결과 Copula 함수로 추정한 개별 VaR는 진실의 VaR보다 위험을 과소추정한 것으로 나타났다. 여기서 구한 단변량에 대한 VaR 추정결과를 이변량으로 확대 해석이 가능하지 않다. 그러나 표 4.4와 4.5를 살펴보면, 세 종류의 Copula 함수를 이용하여 추정한 VaR는 진실의 VaR와 큰 차이를 보이지 않고 매우 유사함을 파악할 수 있다. 그러므로 본 연구에서 이용한 세 종류의 이변량 Copula 함수를 이용하여 VaR를 추정할 수 있다.

5. 결론

VaR를 추정하는데 중요한 문제는 실제 수익률분포의 비대칭성 및 두꺼운 꼬리와 같은 비정규성과 관련된 문제들을 해결하는 것이다. 그리고 단일분포의 VaR가 아닌 다변량 분포에 대한 VaR가 현실적으로 적절하게 사용될 것이다. 따라서 이러한 VaR를 추정하기 위해 비정규성을 따르는 다변량 자료를 고려해야 한다.

본 논문에서는 최근 재무 금융분야에서 활발히 응용되고 있는 Copula 함수를 이용하여 실증 예제의 수익률분포에 적합시켜 이변량 분포의 VaR를 산출하였다. 그리고 기존의 모수적 방법인 분산-공분산 방법과 비교를 통하여 Copula 함수를 이용하여 추정한 이변량 분포의 VaR의 효용성을 확인하였다. 또한 이를 바탕으로 주변분포의 VaR도 추정하여 실제분포에 적합한 정규혼합분포를 이용하여 추정한 VaR와 비교를 통해 얼마나 실제 VaR와 근사하는지를 탐색하였다.

단일 변량인 경우에는 Copula 함수를 이용하여 추정한 VaR보다 정규혼합분포를 이용하여 추정한 VaR가 실제 VaR의 차이가 적었음을 파악할 수 있다. 그러나 이변량 VaR를 추정하는 것이 본 연구의 연구 목적이며, Copula 함수가 가지는 장점인 종속성이 반영된 VaR를 산출할 수 있다는 장점이 있다. 그러므로 본 연구는 복잡한 상호작용이 항상 존재하는 실제 금융시장의 자산들간에서 비정규성의 문제를 해결하기 위하여 Copula 함수의 적용과 이변량 분포의 VaR를 추정방법을 제안하였고 Copula 함수의 효용성을 살펴보았다.

다변량 Copula 함수를 이용한 VaR 추정은 함수의 특성상 두 확률변수에 대한 VaR의 추정만이 가능하다. 따라서 이차원 이상의 다차원 Copula 함수를 이용하여 다변량에 대한 VaR 추정은 두 변수의 쌍에 대한 VaR의 추정량들로 구성된 행렬형태의 VaR 추정량들로 설명되어야 하며 이에 대하여는 향후 연구과제로 남겨둔다.

참고문헌

김명직, 신성환 (2003). Copula 함수의 추정과 시뮬레이션, <선물연구>, **11**, 103–131.

여성칠 (2006). 코퓰러와 극단치이론을 이용한 위험척도의 추정 및 성과분석, <응용통계연구>, **19**, 481–504.

황수영 (2005). <Copula 함수와 극단치 이론을 이용한 Value at Risk 측정에 관한 실증연구>, 한국과학기술원, 박사학위논문.

홍종선, 권태완 (2010). 수익률분포의 적합과 리스크값 추정, <한국데이터정보과학회지>, **21**, 219–229.

Bae, K. H. and Karolyi, A. (2003). A new approach to measuring financial contagion, *Review of Financial Studies*, **16**, 717–763.

Breymann, W. (2003). Dependence structures for multivariate high-frequency data in finance, *Quantitative Finance*, **3**, 1–14.

Genest, C. and MacKay, J. (1986). The joy of copulas: Bivariate distributions with Uniform marginals, *The American Statisticain*, **40**, 280–283.

Jorion, P. (1997). *Value at Risk*, McGraw-Hill, New York.

Li, D. X. (1999). *Value at Risk based on the Volatility Skewness and Kurtosis*, RiskMetrics Group.

Longin, S. (2001). Extreme correlation of international equity markets, *The Journal of Finance*, **2**, 649–676.

Nelson, R. B. (2006). *An Introduction to Copula*, 6th Edition, Springer.

Sklar, A. (1959). Fonctions de repartition a n dimensions et leurs marges, *l'Institut de Statistique de L'Universite de Paris*, **8**, 229–231.

Umberto, C. L. and Walter, V. (2004). *Copula Methods in Finance*, Wiley.

Zangari, P. (1996). An Improved Methodology for Measuring VaR, *RiskMetrics Monitor*, **2**, 7–25.

VaR Estimation of Multivariate Distribution Using Copula Functions

Chong Sun Hong[1] · Jae Hyung Lee[2]

[1]Department of Statistics, Sungkyunkwan University
[2]Research Institute of Applied Statistics, Sungkyunkwan University

(Received January 2011; accepted March 2011)

Abstract

Most financial preference methods for market risk management are to estimate VaR. In many real cases, it happens to obtain the VaRs of the univariate as well as multivariate distributions based on multivariate data. Copula functions are used to explore the dependence of non-normal random variables and generate the corresponding multivariate distribution functions in this work. We estimate Archimedian Copula functions including Clayton Copula, Gumbel Copula, Frank Copula that are fitted to the multivariate earning rate distribution, and then obtain their VaRs. With these Copula functions, we estimate the VaRs of both a certain integrated industry and individual industries. The parameters of three kinds of Copula functions are estimated for an illustrated stock data of two Korean industries to obtain the VaR of the bivariate distribution and those of the corresponding univariate distributions. These VaRs are compared with those obtained from other methods to discuss the accuracy of the estimations.

Keywords: Confidence level, dependence, earnings rate, generator, risk.

[1]Corresponding author: Professor, Department of Statistics, Sungkyunkwan University, 3-53, Myungryun-Dong, Jongro-Gu, Seoul 110-745, Korea. E-mail: cshong@skku.ac.kr