

일반화 파레토 분포에서 임계치 결정에 대한 대안적 연구

윤정연¹ · 조재범² · 정병철³

¹한국지식재산연구원(KIIP), ²(주) 에프원컨설팅, ³서울시립대학교 통계학과

(2010년 12월 접수, 2011년 7월 채택)

요약

일반적으로 일반화 파레토 분포(Generalized Pareto Distribution; GPD)에서 임계치를 결정하는 방법으로는 MEF-그래프나 Hill-그래프를 통한 주관적인 판단을 이용한다는 약점이 존재한다. 본 연구에서는 이와 같은 기존 방법의 약점을 해결하기 위하여 GPD에서 임계치를 결정하는 방법으로 로버스트 추정량을 이용하는 새로운 접근 방법을 제안하였다. 더불어 1987년 1월 5일부터 2009년 8월 3일까지 공시된 KOSPI지수의 일별수익률의 손실부분에 해당하는 왼쪽꼬리부분을 이용하여 실증분석을 실시하였다. 실증분석은 기존의 그래프를 이용한 임계치 결정방법과 본 연구에서 제안한 방법에서 계산된 VaR이 어떤 차이가 존재하는가를 알아보는 방법으로 실시되었다. 분석 결과 본 논문에서 제안한 임계치 결정방법에 의하여 계산된 VaR값들은 기존 방법의 VaR과 큰 차이를 보이지 않았다. 아울러 본 연구에서 제안한 임계치 결정방법의 안정성을 파악한 결과 기존 방법과 큰 차이를 보이지 않았다. 이와 같은 결과들을 토대로 본 연구에서 제안한 로버스트 추정량을 이용한 임계치 결정방법은 기존의 그래프를 이용한 주관적인 임계치 결정방법에 대한 대안적인 방법으로 충분히 고려될 수 있을 것으로 생각된다.

주요어: 일반화 파레토 분포, 임계치, MEF-그래프, Hill-그래프, 로버스트 추정법, VaR.

1. 서론

일반적으로 금융시장에서 정의하는 리스크는 운영(operation), 시장(market), 신용(credit)리스크로 분류된다 (윤평식과 김철중, 2000). 그 중 운영리스크는 수익과의 연관성이 낮아 리스크 관리를 통해 손실을 줄일 수 있는 특징을 가진다. 최근 금융위기로 인해 금융시장에서 발생빈도가 낮은 반면 거대한 손실이 발생할 위험이 커지고 있으며 이는 회사의 손실에 큰 영향을 주기 때문에 이 위험을 관리하고 측정하는 것이 중요하게 여겨지고 있다. 운영리스크를 측정하기 위해 손실분포접근법(Loss Distribution Approach; LDA)을 주로 이용하며 이를 측정하기 위해서는 손실량과 손실빈도에 대한 확률분포의 추정이 필요하다 (McNeil, 1999). 손실량 분포를 모형화하기 위한 연속형 확률분포 중 극단치 분포가 상대적으로 두꺼운 꼬리를 갖고 있다. 즉, 발생빈도는 낮지만 손실의 정도가 큰 위험을 측정하기 위해서는 극단치 이론을 사용하는 것이 바람직하다. 극단치 이론은 극단치를 구하는 방법에 따라 일반화 극단치 분포(Generalized Extreme Value; GEV)와 일반화 파레토 분포(Generalized Pareto Distribution; GPD)로 구분된다 (McNeil, 1999). 일반화 극단치 분포는 일정한 기간마다 극대 값을 구해 손실데이터의 확률분포를 모형화하는 방법이며 일반화 파레토 분포는 일정한 임계치를 초과하는 극단치에 대해 확률분포를 모형화하는 방법이다 (McNeil, 1999).

본 논문에서 일반화 파레토 분포에 기본을 두고 분석하고자 한다. 일반화 파레토 분포에서 임계치를 결정할 때 QQ-plot과 MEF(Mean Excess Function; MEF)-그래프 및 Hill-그래프를 이용한다 (Beirlant

³교신저자: (130-743) 서울시 동대문구 전농동, 서울시립대학교 통계학과, 부교수. E-mail: bcjung@uos.ac.kr

등, 1996). 하지만 기존의 임계치를 결정하는 방법은 판단자의 기준에 따라 임계치가 다르게 결정될 수 있으며 추정된 임계치를 이용하여 위험성 측정을 할 경우 객관성의 문제가 제시될 수 있다. 따라서 본 연구에서는 이러한 임계치 결정에 대한 기존 방법의 문제점을 보완하기 위한 방법을 제안하고자 한다. 본 연구에서 제안하는 방법은 임계치 결정과정에 정규성 가정을 할 수 없는 데이터(이상치가 있거나 두꺼운 꼬리를 갖거나 또는 데이터가 한쪽으로 치우친 경우)에서 사용되는 로버스트 추정법을 바탕으로 하는 방법이다.

일반적으로 시장의 위험성은 VaR(Value at Risk)를 사용하여 측정한다. 본 논문에서는 이러한 임계치 결정방법의 효율성을 평가하기 위하여 VaR을 사용하고자 한다. 하지만 GPD에서 VaR은 임계치 u 가 어떤 값인가에 따라 의존하게 된다. 임계치가 잘못 결정되면 시장의 위험을 나타내는 VaR도 왜곡될 수 있으므로, GPD를 이용하는 경우 임계치 u 의 결정방법은 매우 중요한 문제라 할 수 있다. 따라서 본 연구에서는 임계치 결정 방법의 효율성을 VaR값을 이용하여 파악하고자 한다. 본 논문에서는 1987년 1월 5일부터 2009년 8월 3일까지 공시된 KOSPI 일별주가 수익률 자료에 GPD를 가정한 후 기존의 임계치 결정방법과 제안된 임계치 결정방법에 의하여 계산된 VaR을 이용하여 임계치 결정방법의 효율성을 살펴보고자 한다. 더불어 임계치 결정방법의 안정성을 파악하기 위하여 전체 데이터를 2개월씩 이동시켜가며 총 24개의 분할자료를 만든 후, 각 임계치 결정방법에 따른 VaR값들의 평균과 표준편차를 비교하고자 한다. 이와 같은 방법을 통하여 본 연구에서 제안하는 임계치 결정방법의 안정성을 보이고자 한다.

2. 일반화 파레토 분포(Generalized Pareto Distribution; GPD)

극단치 이론은 임의의 분포를 가지는 확률변수가 취할 수 있는 여러 값 중, 양쪽부분에 치우친 값들의 분포이다. 즉, 분포의 꼬리부분에 대한 통계적 이론으로 전체 분포가 아닌 양쪽 꼬리 부분을 강조함으로써 금융시장의 위험관리에서 극단적 수익률을 관찰하는데 매우 유용한 분포이다. 극단치 이론은 표본자료에서 극단치를 구하는 방법의 따라 크게 GEV 분포와 GPD 분포로 구분되는데, 본 연구에서는 GPD에 초점을 맞추고자 한다.

GPD는 특정 임계치(threshold)를 넘는 극단치들의 극한분포로 정의된다. 서로 동일하고 독립적인 극단치 X_1, X_2, \dots, X_n 이 특정분포 F 에서 추출되었다고 할 때, $y = X - u$ 를 특정 임계치(u)를 넘는 값으로 정의한다면, y 의 조건부 확률분포는 다음과 같이 나타낼 수 있다.

$$F_u(y) = P(X - u \leq y | X > u) = \frac{F(u + y) - F(u)}{1 - F(u)}, \quad 0 \leq y < \infty. \quad (2.1)$$

식 (2.1)에 나타난 $F_u(y)$ 는 Balkema-de Hann-Pickands 정리에 의하면 다음의 GPD 분포함수로 수렴하게 된다 (Jenkinson, 1955).

$$F_{\xi, \sigma}(x) = \begin{cases} 1 - \left\{ 1 + \xi \frac{(x-u)}{\sigma} \right\}^{-\frac{1}{\xi}}, & \xi \neq 0, \\ 1 - \exp\left(-\frac{x-u}{\sigma}\right), & \xi = 0. \end{cases} \quad (2.2)$$

식 (2.2)에 나타난 GPD는 형태모수 ξ 의 크기에 따라 다음의 세 가지 형태로 구분된다.

- 1) $\xi > 0$: Pareto Type I (Pareto, Cauchy, t -분포 등).
- 2) $\xi \rightarrow 0$: 지수분포.
- 3) $\xi < 0$: Pareto Type II (Uniform, Beta 분포 등).

식 (2.2)의 GPD에서 각 모수는 임계치 u 를 초과하는 자료들로부터 최대우도(Maximum Likelihood; ML) 추정방법을 이용하여 추정할 수 있다. ML 추정에 필요한 로그우도함수는 다음과 같다.

$$\log L(\xi, \sigma) = \begin{cases} -n \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^n \log \left(1 + \frac{\xi}{\sigma}(x_i - u)\right), & \xi \neq 0, \\ -n \log \sigma - \frac{1}{\sigma} \sum_{i=1}^n (x_i - u), & \xi = 0. \end{cases} \quad (2.3)$$

3. 임계치 결정

GPD는 특정 임계치 u 를 초과하는 극단값들의 확률분포이므로 임계치 u 를 어떤 값으로 정의하느냐에 따라 극단값들의 수가 달라지며, 이를 이용한 모수 추정치들도 달라지게 된다. 그러므로 임계치 u 의 결정이 매우 중요한 문제이다. 기존 연구에서 임계치의 결정은 보통 POT(Peak Over Threshold)방식을 이용한다. POT방식에서 임계치의 결정은 MEF-그래프를 이용하는 방법과 Hill-그래프를 이용하는 방법이 존재한다 (McNeil, 1999).

3.1. MEF-그래프

MEF-그래프는 임계치 u 를 초과하는 평균초과함수를 이용하는 방법으로, GPD에서 MEF는 다음과 같이 정의된다.

$$e(u) = E(X - u | X > u), \quad 0 \leq u \leq x_F, \quad (3.1)$$

여기서 x_F 는 주어진 데이터의 최대값을 나타낸다. 그러므로 전체 표본의 개수가 n 개인 자료에서 평균 초과함수의 추정치는 다음과 같이 계산된다.

$$\hat{e}(u) = \frac{\sum_{i=1}^n (X_i - u)}{\sum_{i=1}^n I(X_i > u)}. \quad (3.2)$$

즉, MEF-그래프는 임계치 u 를 초과하는 값들의 합을 임계치(u)를 초과하는 자료의 개수로 나눈 평균 값을 의미한다 (Seluck와 Ramazan, 2001). MEF-그래프 방법에 의한 임계치 결정방법은 모든 자료에 대한 MEF-그래프가 어떤 특정한 임계치(u)값 이상에서 양의 기울기를 가지면 이 특정 u 값을 임계치로 결정하는 방식이다.

3.2. Hill-그래프

Hill 추정치는 n 개의 표본 중에서 k 개의 극단적인 값들의 평균과 k 번째로 큰 값의 차이를 의미한다. 즉, 각각의 자료를 크기 순서에 따라 $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ 로 정렬하고 다음과 같이 Hill 추정치를 구한다.

$$\xi_{k,n} = \left[\frac{1}{k} \sum_{i=1}^n \log(X_{(i)}) \right] - \log(X_{(k)}), \quad k \geq 2. \quad (3.3)$$

Hill 그래프의 형태모수 $\xi_{k,n}$ 의 값이 어떤 특정한 값 $X_{(i)}$ 이후로 안정화를 보이면 그 값을 임계치로 결정한다 (Embrechts 등, 1997).

4. 새로운 임계치 결정방법

일반적으로 GPD에서 임계치 결정시 MEF-그래프 및 Hill-그래프의 방법을 사용한다. 하지만 위의 방법은 그래프의 형태에 따라 연구자의 결정에 의해 값이 정해지기 때문에 어느 정도 연구자의 주관성이 개입될 여지가 충분히 존재한다. 본 논문에서는 이러한 주관성의 문제를 해결하기 위한 방법으로 위치모수에 대한 로버스트 추정량을 이용하는 새로운 임계치 결정방법으로 제안하고자 한다. 본 논문에서 제안하는 임계치 결정방법은 자료의 형태가 극단치들을 포함하고 있기 때문에 위치모수에 대한 로버스트 추정량인 중위수(median)를 이용하여 다음의 같은 방식을 통하여 임계치를 결정하고자 한다.

$$\text{threshold} = \text{median} - k \times s, \quad k = 2.5, 3.0, 3.5, 4.0, \quad (4.1)$$

여기서 s 는 σ 에 대한 로버스트 추정량으로 $s = 1.483\text{MAD} = 1.483 \text{ med}_i \{|x_i - \text{med}_j(x_i)|\}$ 를 사용하고자 한다. 이때 $s = 1.483\text{MAD}$ 에서 1.483은 정규분포에서 s 가 σ 의 일치추정량이 되도록 하는 상수이다 (송문섭, 1996). 아울러 k 는 임계치를 결정하기 위한 상수로서 k 에 따라 임계치가 다르게 결정되므로 0.5의 차이를 두고자 한다. k 값이 커질수록 극단치로 고려하는 임계치는 커지게 되며 임계치를 초과하는 값들의 개수는 작아지게 된다.

5. VaR(Value at Risk)의 측정

일반적으로 시장의 위험성은 VaR(Value at Risk)을 이용하여 측정한다. VaR은 $c\%$ 의 확신을 갖고 정상적인 시장여건 하에서 어떤 개별 포지션 또는 포트폴리오 포지션의 목표기간 동안에 발생할 최대 손실 금액 X 로 정의된다. GPD하에서 VaR은 조건부 확률분포를 이용하여 계산된다 (윤평식과 김철중, 2000). 이때 GPD분포 하에서 조건부 확률분포는 다음과 같다 (McNeil, 1999).

$$F_u(y) = P(X - u \leq y | X > u) = \frac{F(u+y) - F(u)}{1 - F(u)}, \quad 0 \leq y < \infty. \quad (5.1)$$

위의 식을 정리하면 다음의 식으로 정리된다.

$$F(u+y) = F_u(y)[1 - F(u)] + F(u), \quad (5.2)$$

여기서 임계치(u)를 초과하는 확률 $P(X > u)$ 는 $1 - F(u)$ 이므로 이는 $1 - F(u) = n_u/N$ 으로 쓸 수 있다. 이때 N 은 전체 관측값 수를 나타내고, n_u 는 임계치 u 를 넘는 극단값의 수를 나타낸다. 따라서 위의 식을 정리하면, 다음과 같은 식을 얻을 수 있다.

$$1 - F(u+y) = P(X > u+y) = \frac{n_u}{N} \left(1 + \xi \frac{x-u}{\sigma}\right)^{-\frac{1}{\xi}}. \quad (5.3)$$

식 (5.3)을 이용하면 GPD분포에서 q 분위수 값은 다음과 같이 구할 수 있다.

$$x_q = u + \frac{\beta}{\xi} \left[\left(\frac{N}{n_u} (1-q) \right)^{-\xi} - 1 \right], \quad (5.4)$$

여기서 β 는 시장의 규모를 나타낸다. 그러므로 식 (5.4)에 나타난 q 분위수 값이 이 시장의 VaR을 나타낸다.

이와 같은 VaR은 GPD에서 임계치 u 가 어떤 값인가에 따라 의존하게 된다. 임계치가 잘못 결정되면 시장의 위험을 나타내는 VaR도 왜곡될 수 있으므로, GPD를 이용하는 경우 임계치 u 의 결정방법은 매우 중요한 문제라 할 수 있다. 따라서 본 연구에서는 임계치 결정 방법의 효율성을 여러 임계치 결정 방법에 따라 얻어진 임계치들을 이용하여 VaR값을 계산한 후 이들 값의 비교를 통해 실시하였다.

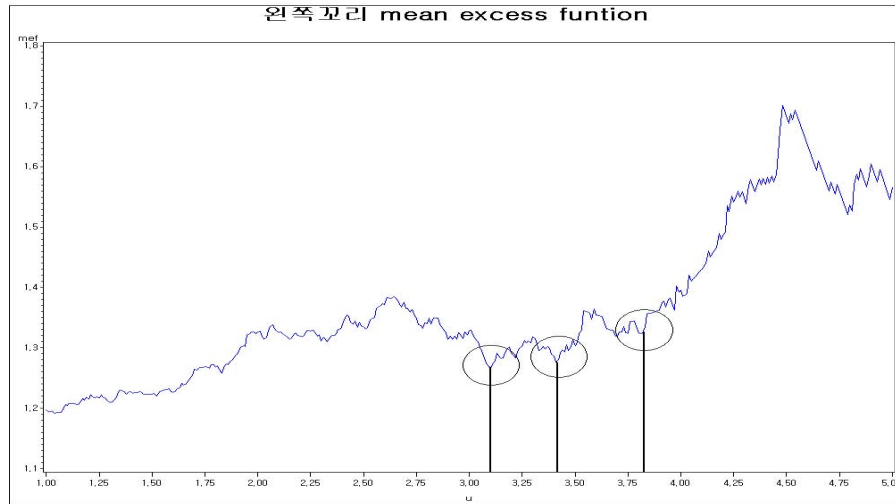


그림 6.1. 왼쪽꼬리 MEF-그래프

6. 자료 분석

본 논문은 1987년 1월 5일부터 2009년 8월 3일까지 공시된 6129개의 KOSPI지수의 일별 주가지수 수익률 자료에서 손실부분에 해당하는 왼쪽 꼬리부분을 사용하였다. 이러한 일별 수익률 자료를 바탕으로 앞 절에서 설명한 MEF-그래프, Hill-그래프 및 본 연구에서 제안한 임계치 결정방법에 의하여 임계치를 결정하였다. 이때 표현과정에서 손실률은 음수의 형태로 표현되고, 임계치도 음수 값을 갖게 되지만 표현상 이들 값에 -1 을 곱하여 양수 값을 갖도록 표기하였다. 결정된 임계치를 초과하는 데이터를 이용해 서 모수를 추정하고 추정된 모수를 통해서 VaR를 계산한 후 이들 값을 비교하였다.

6.1. 임계치 결정

■ MEF-그래프를 이용하는 방법

다음 그림 6.1은 KOSPI 일별 주가지수익률의 손실부분에 대한 MEF-그래프를 나타낸다. MEF-그래프를 이용한 임계치 결정 방식은 MEF-그래프가 증가하는 지점을 임계치로 결정하는 방식으로 그림 6.1에서 동그라미로 표시된 부분이 이에 해당한다. 이를 통하여 3.125, 3.450, 3.824를 임계치로 결정하였다. 이들 임계치들을 초과하는 표본수는 각각 230개, 177개와 129개로 나타났다.

■ Hill-그래프를 이용하는 방법

다음 그림 6.2는 KOSPI 일별 주가지수익률의 손실부분에 대한 Hill-그래프를 나타낸다. 일반적으로 Hill-그래프에 의한 임계치 결정방식은 Hill-그래프에서 기울기가 완만해 지는 지점을 임계치로 결정한다(그림 6.2에서 동그라미 표시 부분). 본 논문에서는 각 점의 기울기를 구하고 기울기가 10^{-1} , 10^{-2} 및 10^{-3} 보다 작게 되는 시점을 임계치로 결정하였다. 위의 기준에 의해서 임계치를 3.980, 2.790, 2.160으로 결정하였다. 이들 임계치들을 초과하는 표본수는 각각 110개, 262개와 461개로 나타났다.

■ 로버스트 추정량을 이용하는 방법

본 연구에서 제안한 임계치 결정방법인 로버스트 추정량을 이용하는 방법을 적용하기 위하여 필요한 손

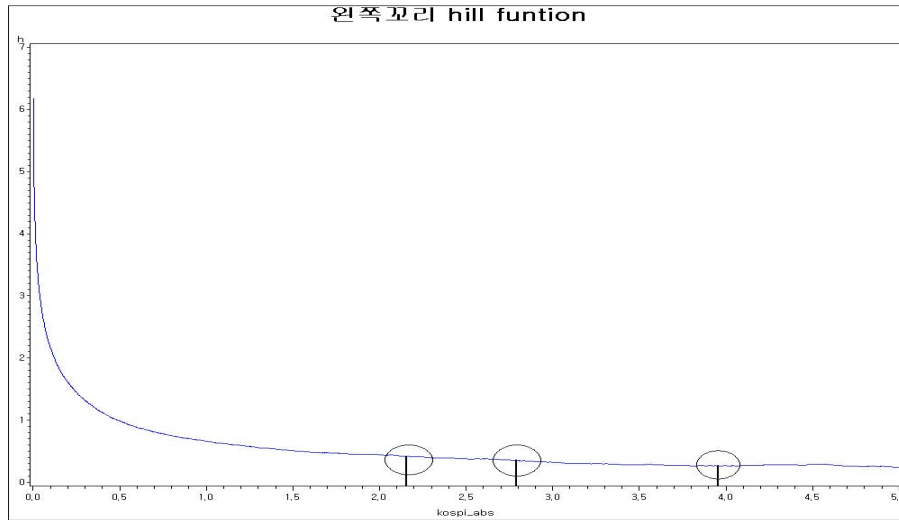


그림 6.2. 왼쪽꼬리 Hill-그래프

표 6.1. 로버스트 추정량을 이용하여 결정된 임계치

k	2.0	2.5	3.0	3.5	4.0	4.5	5.0
임계치	2.456	2.861	3.265	3.670	4.075	4.480	5.289

실부분의 증위수에 대한 추정치는 0.83666으로 나타나고, $s = 0.909536$ 으로 추정되었다. 이를 이용하여 다양한 k 에 따라 얻어지는 임계치들은 표 6.1에 나타나 있다. 표 6.1에서 k 값이 2인 경우와 4.5이상인 경우 임계치를 초과하는 표본수가 너무 많거나 너무 적어서 이 부분은 분석에서 제외하였다.

6.2. 모수추정 및 VaR 추정

GPD에서 결정된 임계치(u)를 초과하는 데이터들을 이용하여 VaR을 구하기 위해서는 GPD의 모수를 추정해야 한다. 본 논문에서는 식 (2.3)에 나타난 로그최대우도함수를 최대화하는 ML추정법을 이용하여 해당 모수를 추정하고, 이를 바탕으로 식 (5.4)에 의하여 VaR값을 계산하였다. 다음 표 6.2는 3가지 임계치 결정방법에 따라 결정된 극단치를 이용한 모수추정 결과와 이에 바탕을 둔 VaR을 나타낸다.

표 6.2에 나타난 3가지 방법에 의해 구한 VaR를 비교해 보면, 계산된 VaR값들은 각 임계치 결정방법에 따라 큰 차이가 없음을 알 수 있다. 특히 본 연구에서 제안한 로버스트 추정량을 이용하여 임계치 결정방법을 이용한 VaR값은 임계치에 크게 영향을 받지 않고 안정적인 모습을 보이고 있다. 따라서 기존의 MEF-그래프와 Hill-그래프에 의한 임계치 결정 방법이 갖는 연구자의 주관성의 문제를 로버스트 추정량을 이용하면 해결할 수 있을 것으로 생각한다. 더불어 제안된 방법에 따라 계산된 VaR값도 기존 방법에 의해 계산된 VaR과 큰 차이가 없으므로 GPD에서 임계치 결정방법으로 본 연구에서 제안한 로버스트 추정량을 사용하는 것이 기존 방법들의 약점을 해결하기 위한 충분한 대안이 될 것이라 생각한다.

6.3. VaR의 안정성 비교

앞 절에서는 전체 데이터를 이용해 추정된 VaR값의 비교를 통해 제안된 로버스트 추정량을 이용한 임계치 결정방법이 MEF-그래프와 Hill-그래프를 이용하는 방법에 대한 대안적인 방법이 될 수 있음을 보

표 6.2. 왼쪽꼬리의 모수 추정값 및 계산된 VaR값

임계치 결정방법	k	임계치	표본수	ξ	σ	VaR
MEF	-	3.125	230개	0.115	1.134	3.627
		3.450	177개	0.144	1.111	3.641
		3.824	129개	-0.329	3.227	3.341
Hill	-	2.160	461개	0.045	1.256	3.613
		2.790	262개	0.119	1.173	3.886
		3.980	110개	0.112	1.248	3.605
로버스트 추정량	2.5	2.861	269개	0.047	1.276	3.623
	3.0	3.265	201개	0.099	1.185	3.622
	3.5	3.670	141개	0.136	1.152	3.604
	4.0	4.075	102개	0.113	1.255	3.607

표 6.3. 24개의 분할자료에 대한 임계치와 VaR의 평균과 표준편차

임계치 결정방법	k	임계치		VaR	
		평균	표준편차	평균	표준편차
MEF	-	3.410	0.229	3.838	0.150
		3.689	0.168	3.855	0.126
Hill	-	4.080	0.173	3.826	0.147
로버스트 추정량	2.5	3.005	0.193	3.848	0.139
	3.0	3.429	0.227	3.851	0.135
	3.5	3.853	0.260	3.829	0.100
	4.0	4.278	0.292	3.773	0.111

였다. 이번 절에서는 전체데이터를 2개월의 간격을 두고 15년씩 나누어 총 24개의 자료로 분할한 후, 분할된 자료에서 얻어진 임계치와 VaR값을 이용하여 제안하는 임계치 결정방법의 안정성을 보이고자 한다. 이를 위하여 분할된 24개의 자료에 각 방법을 이용하여 임계치를 결정하고, 이를 바탕으로 총 24개의 VaR값을 계산하였다. 이와 같이 계산된 24개의 VaR값의 변동이 크다면 임계치 결정과정의 안정성에 문제가 있는 것으로 파악할 수 있다.

분할된 24개의 자료를 통해서 3가지 임계치 결정방법에 의하여 결정된 임계치의 평균과 표준편차 및 이를 이용하여 계산된 VaR의 평균과 표준편차는 다음 표 6.3에 나타나 있다.

표 6.3를 살펴보면, 먼저 24개의 분할된 자료에서 각 임계값에 의하여 구해진 VaR의 평균들이 비슷함을 알 수 있다. 이는 VaR이 임계값에는 크게 의존하지는 않는다는 것을 의미한다. 더불어 본 연구에서 제안한 방법을 이용한 VaR의 표준편차의 경우에도 기존 방법에 의하여 계산된 VaR의 표준편차와 비슷하거나 작게 나타나 VaR을 계산하는 데 있어서 안정적인 방법으로 나타났다.

이와 같은 결과들을 토대로 연구자의 주관에 의존하지 않고 객관적으로 임계치를 결정할 수 있는 본 연구에서 제안한 로버스트 추정량을 이용한 임계치 결정방법이 기존의 그래프를 이용한 임계치 결정방법과 더불어 효과적인 VaR 추정을 위한 임계치 결정방법으로 사용될 수 있을 것으로 판단된다.

7. 결론

본 연구에서는 GPD분포에서 임계치를 결정하는 방법으로 로버스트 추정량을 이용하는 새로운 접근 방법을 제안하였다. 더불어 1987년 1월 5일부터 2009년 8월 3일까지 공시된 KOSPI지수의 일별수익률의 손실부분에 해당하는 왼쪽꼬리부분을 이용하여 기존의 그래프를 이용한 임계치 결정방법과 본 연구

에서 제안한 방법의 효율성을 VaR을 이용하여 파악하였다. 아울러 제안된 임계치 결정방법의 안정성을 파악하기위해 원 자료를 24개의 분할 자료로 나누고 이들 자료의 VaR값의 평균과 표준편차를 통해 VaR값들의 안정성을 살펴보았다. 실증분석 결과, 그래프를 이용하는 기존의 임계치 결정방법과 본 연구에서 제안한 임계치 결정방법에 의하여 계산된 VaR값들은 서로 큰 차이를 보이지 않았다. 아울러 24개의 분할자료에 대한 제안된 방법을 이용하여 계산한 VaR의 평균과 표준편차도 기존 방법을 이용한 결과와 큰 차이를 보이지 않았다. 이와 같은 결과들을 토대로 연구자의 주관에 의존하지 않고 객관적으로 임계치를 결정할 수 있는 본 연구에서 제안한 로버스트 추정량을 이용한 임계치 결정방법이 기존의 그래프를 이용한 주관적인 임계치 결정방법에 대한 대안적인 방법으로 충분히 고려될 수 있을 것으로 생각된다.

참고문헌

- 송문섭 (1996). <로버스트 통계>, 자유아카데미, 서울.
- 윤평식, 김철중 (2000). <금융기관 시장위험관리>, 한국금융연수원, 서울.
- Beirlant, J., Teugels, J. L. T. and Vynckier, P. (1996). *Practical Analysis of Extreme Values*, Leuven University Press, Leuven.
- Journal of the American Statistical Association*, **92**, 1609-1620.
- Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer, New York.
- Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements, *Quarterly Journal of the Royal Meteorological Society*, **81**, 158-171.
- McNeil, A. J. (1999). Extreme value theory for risk managers, In *Internal Modelling and CAD II*, pp. 93-113, Risk books.
- Selcuk, F. and Ramazan, G. (2001). *Overnight Borrowing, Interest Rates and Extreme Value Theory*, Departmental Working Papers 0103, Department of Economics, Bilkent University,

An Alternative Study of the Determination of the Threshold for the Generalized Pareto Distribution

Jeongyeon Yoon¹ · JaeBeom Cho² · Byoung Cheol Jung³

¹Korea Institute of Intellectual Property(KIIP); ²F1consulting Corporation

³Department of Statistics, University of Seoul

(Received December 2010; accepted July 2011)

Abstract

In practice, thresholds are determined by the two subjective assessment methods in a generalized pareto distribution of mean extreme function(MEF-graph) or Hill-graph. To remedy the problem of subjectiveness of these methods, we propose an alternative method to determine the threshold based on the robust statistics. We compared the MEF-graph, Hill-graph and our method through VaRs on the Korean stock market data from January 5, 1987 to August 3, 2009. As a result, the VaR based on the proposed method is not much different from the existing methods, and the standard deviation of VaR for our method was the smallest. The results show that our method can be a promising alternative to determine thresholds of the generalized pareto distributions.

Keywords: Generalized pareto distribution, threshold, MEF-graph, Hill-graph, robust estimation, VaR.

³Corresponding author: Associate professor, Department of Statistics, University of Seoul, Jeonngong-Dong 90, Dongdaemun-Gu, Seoul 136-743, Korea. E-mail: bcjung@uos.ac.kr