

## 극단치이론을 이용한 보험사 위험자본의 추정

여성철<sup>1)</sup> 장동한<sup>2)</sup> 이병모<sup>3)</sup>

### 요약

전 세계적으로 금융시장에서는 예측할 수 없는 대형 사건들이 지속적으로 일어나고 있으며, 특히 보험시장의 경우에는 대재해성(catastrophe)손실 등을 포함한 극단적 사건에 대한 예측이 날이 갈수록 어려워지고 있는바 극단적 위험관리에 대한 필요성이 증대되고 있다. 극단적 위험관리에 있어 분포의 꼬리영역만을 분리하여 그 정보를 최대로 이용하는 방법이 필요한데, 이러한 문제들을 해결하기 위해 극단치들의 움직임을 모형화하는 소위 극단치 이론(Extreme Value Theory: EVT)을 이용하는 것이 요구된다. 극단치 이론은 현재 여러 분야에서 활용되고 있는데, 특히 금융시장에서는 극단적 변화가 미치는 영향을 분석하기 위해서 극단치 이론을 이용한 금융위험분석을 실시하고 있다. 본 연구에서는 위험관리에 있어서 극단치 이론의 중요성을 검토하고 보험사의 위험자본에 초점을 맞추어 손실 발생의 극단적 위험을 측정하고 이에 대비한 위험자본의 적정규모를 측정하여 보았다.

주요용어: 위험관리, 손해심도, 극단치이론, 위험자본.

### 1. 서론

금융기관이 직면하는 위험은 크게 시장위험, 신용위험, 운영위험으로 나눌 수 있는데 금융기관 위험관리자의 입장에서 볼 때 금융기관의 위험은 두 가지로 나눌 수 있다. 하나는 금융기관의 일상적인 활동에서 발생할 수 있는 손실의 위험이고 다른 하나는 극단적인 사건이 발생하여 금융기관의 존립 자체가 위태로워질 위험이다. 극단적인 위험의 발생은 금융시장에서 일상적인 가격의 움직임과 구조적으로 다른 메커니즘에 의해서 자산가격의 급격한 움직임이 일어날 가능성이 있기 때문이다. 금융시장의 증대한 불안정성을 관리하기 위하여 이와 같이 드물게 일어나긴 하지만 막대한 영향을 미치는 사건들을 다루어야 한다. 예를 들어 외환위기와 같이 예기치 못한 사건의 발생이나 투기로 인한 자산가격의 급격한 움직임을 다루어야 하는 것이다. 멕시코경제위기(1994), 아시아금융위기(1997), 러시아금융위기(1998) 등을 포함하는 극단적인 사건들의 경험은 극단적 위험관리의 중요성을 일깨워 주었다.

1) (143-701) 서울시 광진구 화양동 1번지, 건국대학교 상경대학 응용통계학과, 교수

E-mail: scyeo@konkuk.ac.kr

2) (143-701) 서울시 광진구 화양동 1번지, 건국대학교 상경대학 국제무역학과, 교수

E-mail: dhchang@konkuk.ac.kr

3) (135-929) 서울시 강남구 역삼동 790-1, 동양파이낸설(주) 업무혁신팀, 멤버

E-mail: cutyboy77@nate.com

일반적으로 위험을 측정하는 것이란 위험의 분포를 위험척도라고 알려져 있는 수치로 요약하는 것이다. 전통적인 모수/비모수적 위험측정 방법은 다수의 관측치가 존재하는 실증분포의 분석을 통해 위험을 측정하지만, 극단적인 위험에 대한 정보를 포함하고 있지 않는 문제가 있다.

극단적 위험관리의 경우 일반적으로 자료에서 직접적으로 눈에 띄지 않는 분위수와 꼬리 확률에 대한 추정을 요구하므로 전통적인 위험측정 방법은 분명 적절하지 않다. 예를 들어 전체의 분포를 이용하여 VaR(value at risk)를 추정하게 되면 모수의 추정에 있어서 분포에서 가장 많은 자료가 존재하는 중심부분에 큰 영향을 미치고, 분포의 꼬리부분 자료가 가지고 있는 정보는 추정에 적게 반영된다. 따라서 전통적인 방법으로 VaR를 추정할 경우, 주 관심 대상인 분포의 꼬리부분을 정확히 추정할 수 없으므로 정확한 VaR를 추정할 수 없게 되는 것이다. 따라서 극단적 위험관리에 있어 분포의 꼬리영역만을 분리하여 그 정보를 최대로 이용하는 방법이 필요한데 이러한 문제들을 해결하는 데는 극단값의 움직임을 모형화 하는 극단치 이론(Extreme Value Theory: EVT)을 이용하는 것이 타당하다.

극단치 이론이란 관측된 전체 데이터의 분포를 모형화한 결과로서 꼬리부분의 확률을 얻는 것 대신에 분포의 꼬리부분에 속하는 극단의 데이터들을 다른 영역의 데이터들과 분리하여 이들만의 점근적 분포형태를 제공하는 확률이론이라고 할 수 있다. 따라서 극단치 이론을 이용하면 꼬리부분의 확률을 추정하는데 중심부의 자료가 미치는 영향을 줄일 수 있으며 또한 원래의 데이터의 확률분포에 상관없이 분포의 꼬리모양을 보다 정확하게 파악할 수 있다. 또한 금융데이터의 특징으로 잘 알려져 있는 분포의 비대칭성과 두터운 꼬리형태를 반영할 뿐만 아니라 실증분포의 꼬리를 부드럽게 만들어서 과거에 발생한 손실을 초과하는 극단적 상황의 발생가능성을 반영하도록 하여준다. 따라서 극단치 이론은 금융기관이나 감독당국이 관심을 갖는 대재해성 시장위험(catastrophic market risk)과 이들 위험에 대비한 자본금의 적정성을 파악하는 데에도 유용하다. 극단치 이론에 관한 문헌으로는 Embrechts 등(1997), Reiss와 Thomas(2001), Coles(2001) 등이 있고, 특히 보험사 위험자본의 추정과 관련한 논문으로는 McNeil(1997), Corradin(2003), Hürlimann(2004) 등이 있다.

극단치 이론은 현재 여러 분야에서 활용되고 있는데, 특히 금융시장에서는 극단적 변화가 미치는 영향을 분석하기 위해서 극단치 이론을 이용한 위험 분석을 실시하고 있다. 통화위기나 주식시장 폭락, 신용불이행, 그리고 손해보험에서 손실규모가 일정규모 이상으로 매우 큰 경우에 어느 정도의 범위를 일정규모 이상으로 보고 위험관리를 하는 것이 바람직한지에 대해 금융기관이나 보험회사의 관심이 높아지고 있는 것이다. 본 연구에서는 그 대표적인 예로서 보험사의 위험관리를 검토해 보고자 한다.

보험사의 자본은 보유자본과 경영활동을 위해 실질적으로 필요한 요구자본으로 구분할 수 있고, 보유자본은 회계적자본과 가용자본으로, 요구자본은 위험자본과 규제자본으로 구성되어진다. 보유자본 중 회계적자본(accounting capital)이란 현재 B/S상 보험회사가 가지고 있는 자기자본인 순자산가치를 말하고, 가용자본(available capital)이란 예상치 못한 손실이 발생하는 경우 계약자나 채권자에 대한 약속을 이행하기 위한 비상금(risk buffer)을 포함한다. 요구자본 중 위험자본(risk capital)이란 일정기간동안 일정 신뢰구간 내에서 측

정된 각종 리스크로부터 발생할 수 있는 최대손실에 대비하여 보험회사가 도산하지 않기 위해 가지고 있어야 할 자본 또는 정상적인 영업활동을 위해 실질적으로 필요한 자본을 의미한다. 규제자본(regulatory capital)이란 감독규정 준수를 위해 필요한 자본을 말하는데 지급여력비를 산출시의 지급여력기준 금액이 이에 해당된다.

본 논문의 목적은 위험관리에 있어서 극단치 이론의 중요성을 검토하고 극단치 이론이 가장 합리적인 근거를 제시하는 실례를 분석하는 데에 있다. 구체적으로는, 보험사의 자본 중 위험자본에 초점을 맞추어 손실 발생의 극단적 위험을 측정하여 극단적 위험에 대비한 위험자본을 측정하고자 하였다.

## 2. 위험자본

### 2.1. 금융기관의 위험관리와 자본배분

사업의 성격상 금융기관은 위험을 보유하고 위험을 관리하는 대가로 부가가치를 창출하는 기관이다. 오늘날 금융시장의 불확실성 증대, 정보기술의 발달에 따른 신용 위험의 출현 등으로 금융기관 위험관리의 중요성이 더해지고 있으나 근본적으로는 금융기관의 생산성 향상 차원에서 위험관리는 더 없이 중요하다. 위험관리는 수익과 위험을 동시에 고려하는 체계적이고 합리적인 의사결정 시스템을 구축함으로써 경영효율성 제고에 기여하게 된다. 동일한 예상수익 하에서 낮은 위험을 나타내는(또는 동일한 예상위험 하에서 높은 수익을 나타내는) 분야에 자본을 적정 배분함으로써 장기적으로 수익성을 제고하고 기업가치를 극대화할 수 있다.

한편, 금융기관의 위험관리 기능은 각종 영업활동에 대한 위험분석을 통해 잠재적 위기 상황에 대처할 수 있게끔 경영전략을 수립할 수 있고, 나아가 합리적인 성과 측정을 가능하게 함으로써 자원의 최적 배분을 가능하게 한다. 최근 들어 위험을 감안한 성과측정(RAPM: Risk Adjusted Performance Measurement)이 금융기관 경영의 중요한 수단으로 부각되고 있는데 그 이유는 금융기관의 자본을 가장 우수한 성과를 보이는 분야에 우선적으로 배분함으로써 경영효율성을 제고할 수 있기 때문이다.

위험관리에 있어 자본은 매우 중요한 의미를 갖는데 이는 자본배분이 특정 자산의 운용 과정에서 얼마나 손실을 입을 수 있는가를 따지는 개념이기 때문이다. 어떤 금융기관이 갑작스런 손실을 입고 채권자들이 부채의 상환을 요구할 때 이 금융기관이 파산을 면하기 위해서는 부채를 상환한 후에 남는 자기자본이 손실액을 상회하여야 한다. 따라서 금융기관의 위험관리 측면에서 볼 때 중요한 것은 자산의 가치하락으로 금융기관이 손실을 보게 되었을 때 부채의 상환이 불가능하게 되지 않을 만큼의 충분한 자본을 보유하고 있는지의 여부라고 할 수 있다.

### 2.2. 보험사의 위험자본

보험사의 위험은 크게 언더라이팅(underwriting)위험, 자산(asset)위험, 신용(credit)위험 및 부외(off-balance)위험으로 분류된다. 언더라이팅위험은 보험료위험과 준비금위험으로 구분되는데 보험료 산출과 준비금 산정 오류에 따른 위험을 뜻한다. 각 업종별로 수년

간의 손해액 및 손해를 진전계수와 투자수익 및 사업비율 등을 근거로 언더라이팅위험을 계수화 한 것이다. 자산위험이란 보유자산의 시장가치 감소나 채무불이행에 따른 위험으로 예를 들면 주식의 경우 주가 하락에 따른 위험, 채권의 경우 원금이나 이자를 받지 못하는 위험이다. 신용위험이란 재보험 거래시의 미수금과 기타 미수금에 따른 위험을 말하는데 기타 미수금이란 미수 이자, 미수 임대료, 계열사로부터의 미수금을 뜻한다. 부외위험(off-balance risk) 또는 관계사투자위험이란 자회사에 대한 투자, 관계사에 대한 채무보증, 통제불능 자산, 재보험처리 되지 않은 거대재해 위험 등 재무제표 상에 나타나지 않는 위험을 측정하는 것이다.

위와 같은 여러 위험에 기인하여 보험사의 손실 규모는 다분히 불규칙적으로 발생하게 된다. 특정 해의 보험사의 발생손해액 규모는 지급보험금과 준비금의 합으로 계산되며, 미래의 보험사 손실규모는 보상청구건수 및 개별보상액의 추정치에 근거하여 산정되는바 실제 손실과 예상 손실 사이에 큰 차이가 생기는 경우가 적지 않다. 여기서, 보험사의 위험자본(economic risk capital)이란 미래의 불리한 상황의 전개를 상정하여 헷징 차원에서 보유하여야 하는 자본의 크기를 의미한다. 최악의 경우까지를 고려할 때 보험사는 일상적인 규모의 위험자본에 더하여 새로운 자본의 확보가 요구된다. 비상사태가 이미 발생하고 난 시점에서선 견잡을 수 없는 현금유출의 압박이 생길 수 있으며, 투자자들은 용자 조건에 과도한 리스크 프리미엄을 요구하게 될 것이다.

이러한 최악의 경우에 대비하여 보험사들은 통상적으로 재보험에 가입하여 인수 리스크의 일부를 재보험사에 전가하게 된다. 따라서 재보험 가입의 최우선 목표는 예기치 못했던 대형 손실의 발생이 위험자본에 미치는 영향을 최소화하는 데에 있는바 이런 관점에서 보험사의 재보험 가입은 자본배분의 중요한 축을 이룬다고 할 수 있다. 인수 리스크의 일부를 재보험사에 전가함으로써 보험사는 위험자본의 규모를 줄일 수 있는데 여기서 재보험료 지급은 그 대가로 간주되는 것이 옳을 것이다.

이와 같이 보험사 자본배분 차원에서나 재보험관리의 경제성 차원에서 위험자본은 더 없이 중요한 관리 대상이 되는데, 보다 구체적으로는 보험사의 위험자본 규모를 여하히 측정하느냐가 관건이 된다. 보험업과 같이 사업의 위험성이 적지 않은 경우에 있어 기말 영업손실이 위험자본을 초과할 확률을 보험사가 감내할만 하게 충분히 작은 수준에서 억제할 수 있다면 그 사업은 성공적이라고 할 수 있겠다.

위험자본의 규모를 측정하는 방법으로는 통상 VaR(Value at Risk)를 많이 사용하지만 본 논문에서는 보다 실무적인 산정방법을 사용하고자 하는데, 보험사의 위험자본을 예상 손실과 S&P A rating 에 해당하는 99.93%의 분위수 손실과의 차이로 정의하여 측정하고자 한다.

### 2.3. 손해심도모형

보험사의 손실을 확률변수  $X_1, X_2, \dots$  라고 하고, 이들 확률변수들이 독립이고 항등적인 분포함수  $F(x) = P(X \leq x), x > 0$  을 갖는다고 가정한다. 그리고 재보험의 하한점점(lower attachment point)  $r$ 과 상한점점(upper attachment point)  $R$ 이 주어졌을 때 보험사의 손실

$X_i$ 에 대한 재보험사가 지급하는 보험금  $Y_i$ 는 다음과 같이 주어진다.

$$Y_i = \begin{cases} 0, & 0 < X_i < r, \\ X_i - r, & r \leq X_i < R, \\ R - r, & R \leq X_i < \infty. \end{cases} \quad (2.1)$$

이 때, 특정한 빈도로 발생하는 특정 손실보다 더 큰 손실에 대해 많은 보험금의 지출이 발생하게 될 경우에는  $r$ 을 어떻게 선택해야 할 것인가에 대한 문제가 제기될 수 있는데, 이 문제의 답을 찾기 위해서는 보험기간을 고정 시키고 그 기간 동안 피보험자에 의해 발생하게 될 손실의 빈도에 대해 알 필요가 있다. 보험기간 동안의 손실발생회수를  $N$ 이라고 하면 손실은  $Z = \sum_{i=1}^N Y_i$ 가 된다.

위의 점점에 관한 문제를 공식화 하는 데에는 다음과 같은 방법을 생각해 볼 수 있다. 우선 정해진 값  $p$ 에 대해  $P(Z > 0) < p$ 를 만족하는  $r$ 을 선택한다. 즉,  $r$ 이 결정되면 보험기간 내 총 보험금이 0이 아닐 확률이  $p$ 보다 작다는 것을 의미한다. 따라서  $r$ 과  $R$ 을 선택하는 문제는 손해심도분포  $F_X(x)$ 의 높은 분위수(quantile)의 추정을 요약한 것이라 할 수 있다.

일반적으로 수집 가능한 손실 데이터의 과거 자료는 일정규모  $\delta$ 를 초과하는 손실들로 이루어져 있다. 왜냐하면 모든 손실 데이터를 수집하는 것은 현실적으로 불가능하고 작은 손실 데이터는 그 중요성이 떨어지기 때문이다. 보험도 일반적으로 중요한 손실에 대한 보상을 해주며, 피보험자들도 보통 작은 손실에 대해서는 스스로 처리를 하고 보험사에 알리지 않는다. 그러므로 데이터는 하한점점  $r$ 보다 작은  $\delta$ 에서 절단된 확률변수들로 간주된다.  $\delta$ 에서 절단된 손실들의 분포함수는 다음과 같이 정의된다.

$$F_{X\delta}(x) = \Pr[X \leq x | X > \delta] = \begin{cases} 0, & x \leq \delta, \\ \frac{F_X(x) - F_X(\delta)}{1 - F_X(\delta)}, & x > \delta. \end{cases} \quad (2.2)$$

이제 과거의 손실 데이터가  $\delta$ 에서 절단된 독립이고 항등적인 확률변수들이라고 가정하고, 이를 바탕으로  $\delta$ 에서 절단된 손해심도분포  $F_{X\delta}(x)$ 의 추정치  $\hat{F}_{X\delta}(x)$ 를 찾고자 한다. 이를 위한 한 가지 방법은 데이터에 모수적 모형을 적합 시킨 후, 최대우도와 같은 접합 기준을 효과적으로 활용하여 모수 추정치를 얻는 것이다.

### 3. 극단치이론

극단치이론은 극단적 사건에 대한 확률적 성질을 고찰하는 이론이다. 본 절의 EVT에 관한 내용은 Yeo(2006)에서 자세히 논의하였으므로 여기서는 편의상 EVT에 대해 간략히 요약하기로 한다. 일반적으로 EVT의 결과로 얻어지는 확률분포에는 일반화 극단치 분포(generalized extreme value distribution: GEV)모형과 일반화 파레토분포(generalized Pareto distribution: GPD)모형의 두 가지 타입이 있다. 먼저 GEV모형에 관해 살펴보기로 한다.

확률변수들  $X_1, X_2, \dots, X_n$ 이 독립이고 항등적(IID)이며 분포함수  $F$ 를 가진다고 하자. 그리고 이 확률변수들 중 최대확률변수를  $M_n$ , 즉  $M_n = \max(X_1, \dots, X_n)$ 이라고 하자.

Fisher-Tippett(1928)정리에 의하면 어떤 상수들  $a_n \in R^+, b_n \in R$ 에 의해 표준화된 최대확률변수  $\tilde{M}_n = (M_n - b)/a_n$ 은  $n \rightarrow \infty$ 일 때 비퇴화분포(non-degenerate distribution)로 수렴한다면 아래의 표준극단치분포들(standard extreme value distributions)이라고 부르는 세가지 유형의 분포들 중 하나로 수렴한다.

$$\text{Fréchet: } \Phi_\alpha = \begin{cases} 0, & x \leq 0, \\ \exp\{-x^{-\alpha}\}, & x > 0, \end{cases} \quad \alpha > 0, \quad (3.1)$$

$$\text{Weibull: } \Psi_\alpha = \begin{cases} \exp\{-(-x)^\alpha\}, & x \leq 0, \\ 1, & x > 0, \end{cases} \quad \alpha > 0, \quad (3.2)$$

$$\text{Gumbell: } \Lambda(x) = \exp\{-e^{-x}\}, \quad x \in R. \quad (3.3)$$

한편, Jenkinson(1955)과 von Mises(1936)는  $n \rightarrow \infty$ 일 때  $\tilde{M}_n$ 의 수렴분포로 위의 세가지 유형의 표준극단치분포들로 나타내는 대신 일반화극단치분포(generalized extreme value distribution: GEV)라고 부르는 분포함수를 제시하였다. 즉,

$$\lim_{n \rightarrow \infty} P(\tilde{M}_n \leq x) = \lim_{n \rightarrow \infty} P\left(\frac{\tilde{M}_n - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H_\xi(x) \quad (3.4)$$

가 된다. 여기서  $H_\xi(x)$ 는 GEV분포함수로서 다음과 같이 주어진다.

$$H_\xi(x) = \begin{cases} \exp\{-(1 + \xi x)^{-\frac{1}{\xi}}\}, & 1 + \xi x > 0, \xi \neq 0 \\ \exp(-e^{-x}), & x \in R, \xi = 0 \end{cases} \quad (3.5)$$

여기서  $\xi$ 는 GEV의 형태를 결정짓는 모수로서 형태모수(shape parameter) 또는 꼬리지수(tail index)라고 부르는데,  $\xi = \alpha^{-1} > 0, \xi = -\alpha^{-1} > 0, \xi = 0$ 을 취할 때  $H_\xi$ 는 각각 Fréchet 분포족, Weibull 분포족, Gumbell 분포족에 속하게 된다. 특히  $\xi = 0$ 인 경우는  $\xi \rightarrow 0$ 일 때의  $H_\xi$ 의 극한분포함수  $H_0$ 로 볼 수 있다.

위의 식 (3.4)에서  $n \rightarrow \infty$ 일 때  $M_n$ 의 위치모수(location parameter)와 척도모수(scale parameter)를 각각  $\mu$ 와  $\sigma$ 라고 하면  $n$ 이 충분히 클 때  $M_n$ 의 분포함수는 근사적으로 다음과 같이 주어진다.

$$P(M_n \leq x) \approx H_\xi\left(\frac{\mu - \sigma}{\sigma}\right) = H_{\xi, \mu, \sigma}(x). \quad (3.6)$$

다음으로 GPD모형에 관해 살펴보기로 한다. 확률변수  $X$ 의 분포함수를  $F$ 라 하고  $F$ 의 오른쪽 끝점을  $x_F$ (즉,  $x_F = \sup\{x \in R : F(x) \leq 1\} \leq \infty$ )라고 하자. 이 때 임의의 높은 임계치  $u$ 에 대해서 임계치초과분포함수(excess distribution over the threshold)를 다음과 같이 정의한다.

$$F_u(x) = P(X - u \leq x | X > u) = \frac{F(u+x) - F(u)}{1 - F(u)}, \quad 0 \leq x < x_F - u. \quad (3.7)$$

즉, 임계치초과분포함수  $F_u$ 는 확률변수  $X$ 가 주어진 임계치  $u$ 를 초과할 때 초과크기에 대한 조건부분포함수를 나타낸다. Balkema와 de Haan(1974)과 Pickands(1975)에 의하면  $X$ 의 분

포함수  $F$ 가 GEV분포의 MDA에 속하는 경우(즉,  $F \in MDA(H_\xi), \xi \in R$ )가 되면  $u \rightarrow x_F$ 일 때  $F_u(x)$ 는 일반화파레토분포(generalized Pareto distribution: GPD)함수  $G_{\xi, \sigma_n}(x)$ 로 수렴한다고 알려져 있다. 여기서

$$G_{\xi, \sigma_u}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\sigma_u}\right)^{-\frac{1}{\xi}}, & x \geq 0, 1 + \frac{\xi x}{\sigma_u} > 0, \xi \neq 0, \\ 1 - \exp\left(-\frac{x}{\sigma_u}\right), & x \geq 0, \xi = 0, \end{cases} \quad (3.8)$$

$\sigma_u = \sigma + \xi u$ 이다. 따라서 실제 응용의 경우  $n$ 이 충분히 클 때 최대확률변수  $M_n$ 의 분포가 Fisher-Tippett정리에 의해 근사적으로 GEV분포함수  $H_{\xi, \mu, \sigma}$ 를 따르고 이 때  $x_F$ 에 충분히 가까운 높은 임계치  $u$ 를 잡으면 임계치 초과분포함수  $F_u$ 는 근사적으로 GPD함수에 접근한다.

GPD함수에 내포된 두 개의 모수 중  $\sigma_u (> 0)$ 는 임계치  $u$ 에 따라 변하는 척도모수이며  $\xi$ 는 임계치  $u$ 와 무관한 모수로서 GEV분포에서  $\xi$ 와 같이 GPD의 형태를 결정짓는 형태모수(shape parameter)또는 꼬리지수(tail index)라고 부른다. 그런데  $\xi > 0$ 이면  $G_{\xi, \sigma_n}$ 는 파레토 타입분포가 되고,  $\xi = 0$ 이면  $G_{\xi, \sigma_n}$ 는 지수분포가 되며  $\xi < 0$ 인 경우에는  $G_{\xi, \sigma_n}$ 는 제 II 파레토 타입분포가 된다. 특히  $\xi > 0$ 일 때 GPD함수  $G_{\xi, \sigma_n}$ 는 두터운 꼬리모양을 가지므로 금융데이터의 분석에 적합한 분포라고 할 수 있다.

일반적으로 GEV분포함수는 세 개의 모수  $\xi, \mu, \sigma$ 를 내포하는데 비해 GPD함수는 두 개의 모수  $\xi, \sigma_u$ 를 가지므로 일반적으로 GPD함수를 이용하는 것이 좀더 효율적이라고 볼 수 있다. 또한 대부분의 금융데이터는 GEV모형에서 정의하듯이 극단치가 일정한 주기를 가지고 발생하지 않는다. 따라서 본 논문에서는 극단치 분포의 두가지 타입 중에서 GPD모형에 대해서만 논의하기로 한다.

어떤 기간동안 관찰된  $n$ 개의 표본데이터  $x_1, x_2, \dots, x_n$ 이 분포함수  $F$ 를 따르고  $F$ 의 오른쪽 끝점을  $x_F$ 라고 하자. 그리고  $x_F$ 에 충분히 가까운 임계치  $u$ 가 주어졌을 때  $n$ 개의 표본데이터들 중 임계치  $u$ 를 초과하는 관측치들을  $x'_1, x'_2, \dots, x'_{n_u}$  ( $n_u = \sum_{i=1}^n I_{\{x_i > u\}}$ )라 하고, 임계치초과값들(threshold excesses)을  $y_i = x'_i - u, i = 1, \dots, n_u$ , 라고 표기하자. 이때 임계치  $u$ 가  $x_F$ 에 충분히 가까운 값이면 Balkema-de Haan-Pickands 정리에 의해 임계치초과 데이터  $y_1, y_2, \dots, y_{n_u}$ 는 근사적으로 GPD모형을 따르는 데이터라고 볼 수 있다.

일반적으로 GEV모형이나 GPD모형에 내포된 모수들을 추정하는 방법으로는 최우추정법, 확률가중추정법(method of probability weighted moments) 그리고 비모수적 추정법 등이 있다. 그런데 Hosking과 Wallis(1987)에 의하면 꼬리지수가  $\xi > -0.5$ 이면 최우추정법에 의해 구한 모수들의 추정량은 점근적으로 정규분포를 따른다고 알려져 있다. 따라서 분포가 두터운 꼬리모양을 가질 경우에 모수들에 대해 최우추정법으로 구한 추정치들이 다른 방법으로 구한 추정치들보다 효율적이라고 볼 수 있다. 본 논문에서는 GPD모형에 내포된 모수들의 추정을 위해 최우추정법을 사용하기로 한다.

그런데 GPD모형에 내포된 모수  $\xi, \sigma_u$ 에 대한 최우추정치들은 임계치  $u$ 를 어디서 잡느냐에 따라 분산과 편의의 상반관계가 나타날 수 있다. 즉, 임계치를 너무 낮게 잡으면 극단치 데이터의 수가 많아져서 추정치의 분산은 작아지지만 분포의 점근성이 떨어져서 편의

가 커지게 되고 반면에 임계치를 너무 높게 잡으면 그 반대의 현상이 나타날 수 있다.

일반적으로 임계치를 잡는 방법에는 여러 가지 방법들이 있으나 본 논문에서는 Yeo(2006)에서 논의한 바와 같이 ME플롯(mean excess plot) 방식과 Hill플롯 방식을 사용하기로 한다. 먼저 ME플롯 방식은 평균초과함수(mean excess function)의 그래프를 이용하는 방법이다. 평균초과함수는 다음과 같이 정의된다.

$$e_X(u) = E(X - u | X > u) = \int_0^{x_F - u} x dF_u(x), \quad u < x_F. \quad (3.9)$$

그런데  $F_u(x)$ 가 GPD함수  $G_{\xi, \sigma_u}(x)$ 이면 평균초과함수는  $\xi < 1$  일 때 다음과 같이 주어진다.

$$e_X(u) = \frac{\sigma_u}{1 - \xi} = \frac{\sigma + \xi u}{1 - \xi}, \quad \sigma + \xi u > 0. \quad (3.10)$$

위의 식 (3.10)으로부터 평균초과함수의 그래프는  $u$ 에 관해서 선형임을 알 수 있다. 따라서 ME플롯 방식은  $\{(u, \hat{e}_X(u)) : x_{1:n} \leq u \leq x_{n:n}\}$ 으로 이루어진 플롯에서  $u$ 의 값을 변화시키기에 따라 양의 기울기를 가진 직선형태의 영역이 시작되는 점을 임계값으로 정한다. 여기서  $\hat{e}_X(u)$ 는 평균초과함수의 표본추정치로서 다음과 같이 주어진다.

$$\hat{e}_X(u) = \frac{1}{n_u} \sum_{i=1}^{n_u} y_i. \quad (3.11)$$

다음으로 Hill플롯 방식은  $\{(k, H_{k,n}) : k = 2, \dots, n\}$ 으로 이루어진 플롯에서  $k$ 의 값에 따라 변화시켜갈 때  $H_{k,n}$ 가 안정화되기 시작하는 값을 임계치  $u$ 로 선택한다. 여기서  $H_{k,n}$ 은 Hill(1975)에 의해 제시된 형태모수  $\xi$ 에 대한 추정치로서 다음과 같이 주어진다.

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^k \ln X_{i,n} - \ln X_{k,n}. \quad (3.12)$$

여기서  $X_{i,n}$ 은  $n$ 개의 관찰치들을 크기에 따라 내림차순으로 배열하였을 때  $i$ 번째에 위치한 관찰치를 의미한다. 즉,  $X_{1,n} \geq X_{2,n} \geq \dots \geq X_{n,n}$  이다.

이제 임계치  $u$ 와 GPD함수의 모수추정값들을 바탕으로 분포함수  $F$ 의 분위값을 추정하기로 한다. Balkema-de Haan- Pickands 정리로부터 임계치  $u$ 가  $x_F$ 에 충분히 가까우면  $F_u \sim G_{\xi, \sigma_u}(y)$ 가 된다. 이 결과를 식 (3.7)에 대입하고  $x = u + y$ 라 두면  $x > u$  일 때 자산 수익률분포의 꼬리부분은 근사적으로 다음과 같이 주어진다.

$$F(x) \approx (1 - F(u))G_{\xi, \sigma_u}(y) + F(u) \quad (3.13)$$

여기서 GPD함수에 내포된 모수  $\hat{\xi}, \hat{\sigma}_u$ 에 대해 최우추정치  $\xi, \sigma_u$ 를 대입하고  $F(u)$ 의 추정치로 경험적분포함수  $\hat{F}(u) = 1 - n_u/n$ 를 적용하면 식 (3.13)에 주어진  $F(x)$ 는 다음과 같이 추정된다.

$$\hat{F}(x) = 1 - \frac{n_u}{n} \left( 1 + \hat{\xi} \frac{x - u}{\hat{\sigma}_u} \right). \quad (3.14)$$



표 4.1: 화재손실 데이터의 기초 통계량(단위: 100만 크로나)

관측수	2167
평균	3.385
분산	72.376
표준편차	8.507
최소값	1.000
최대값	263.250
왜도	18.763
첨도	483.764

## 4. 실증분석

### 4.1. 기초자료분석

실증분석을 위한 데이터로 국내 데이터를 이용하려고 하였으나 관련 자료를 구하는데 많은 어려움이 있어 부득이 외국의 데이터를 사용하였는데, 1980년 1월부터 1990년 12월까지의 기간 동안 100만 크로나가 넘는 2,167개의 덴마크 화재손실 데이터로 구성되어 있다. 화재손실의 원인으로는 실화, 방화 등이었으며, 부동산과 동산의 총 피해액을 포함하고 있다.

표 4.1은 덴마크 화재손실 데이터에 대한 기초통계량을 보여준다. 표 4.1을 살펴보면 왜도가 18.763이므로 오른쪽으로 긴 꼬리를 갖는 분포가 되며, 첨도는 483.764의 큰 값을 갖고 있으므로 정규분포를 벗어나고 있다. 따라서 화재 손실 데이터에 정규분포 근사를 이용하는 것은 적절하지 못함을 알 수 있다.

극단치 이론을 응용하기 전에 극단치 분포 모형 사용의 정당성을 검증하기 위해서 몇

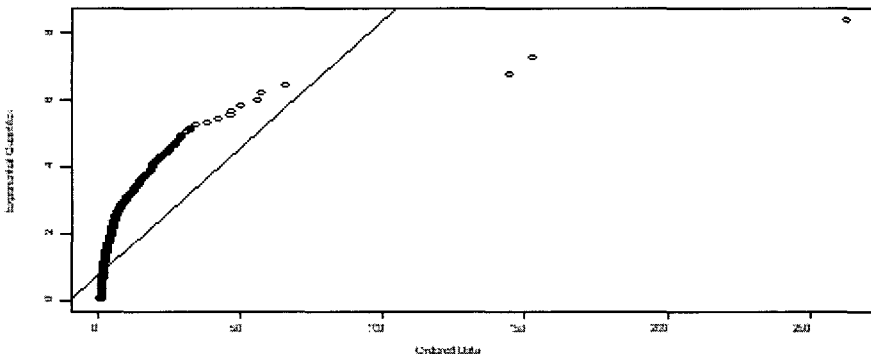


그림 4.1: 화재손실 데이터의 Q-Q 플롯

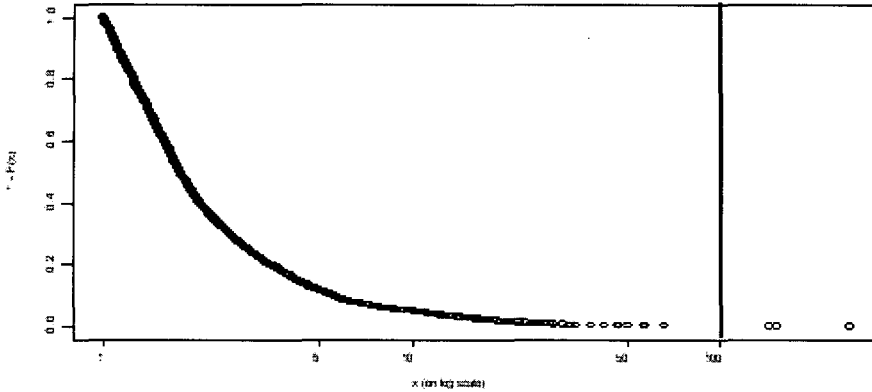


그림 4.2: 화재손실 데이터의 실증생존함수

가지 사전분석을 통해서 데이터가 두꺼운 꼬리를 가지고 있는지 여부를 확인할 필요가 있다. 이러한 기초 단계에서 보통 Q-Q플롯을 통해서 데이터에 대한 시각적인 점검을 실시한다. 그림 4.1은 지수분포에 대응하는 Q-Q플롯을 보여주고 있다. 지수분포를 기준분포로 그린 그래프에서 만약 데이터의 본래 분포가 지수분포이면 그림에서 나타난 점들은 양의 기울기를 갖는 직선 위에 있을 것이다. 만약 점들이 오목한 형태로 나타나면 실증분포가 두꺼운 꼬리분포를 의미한다. 그러나 볼록한 출발은 얇은 꼬리를 가진 분포라는 것을 의미한다. 그림 4.1을 살펴보면 오목한 형태로 나타나고 있음을 알 수 있고, 이는 데이터의 분포가 두꺼운 꼬리를 갖는다는 것을 의미한다.

그림 4.2는 1980년 1월부터 1990년 12월까지 관찰된 화재 손실 데이터의 실증생존함수(empirical survival function)를 나타낸다. 높은 초과 층을 분석하기 위해 그림 4.2에서 100을 초과하는 층의 손실을 추정해야 한다고 가정할 때, 이 구간에서는 단지 3개의 손실만을 가짐을 보게 된다. 만약 손해심도분포의 모든 데이터를 적합 시킨다면, 데이터의 수가 매우 적은 이러한 꼬리 부분에는 잘 적합되지 못할 것이다. 따라서 손해심도분포의 꼬리 부분에 대한 좋은 추정치를 얻을 수 있는 방법이 필요하다. 그러나 꼬리 부분에 대한 좋은 추정치를 얻었다고 해서 미래에 발생할 손실을 정확하게 예측할 수는 없다. 극단치 이론은 미래의 발생할 손실을 정확하게 예측할 수는 없지만, 과거의 극단적인 사건들을 설명할 수 있고, 이를 통해 미래의 손실을 대비할 수 있게 해줄 수 있다

#### 4.2. GPD모형의 적합성

GPD모형의 적합성을 알아보는 손쉬운 방법은 GPD모형의 분포함수와 화재손실 데이터에 대한 실증분포함수를 동시에 나타내는 그림을 통해 데이터가 GPD모형에 적합한 정도를 확인하는 것이다. 그림 4.3은 GPD모형의 분포함수와 화재손실 데이터에 대한 실증분포함수를 동시에 보여주고 있다. 이 그림에서 나타난 실선부분은 GPD모형의 분포함수를

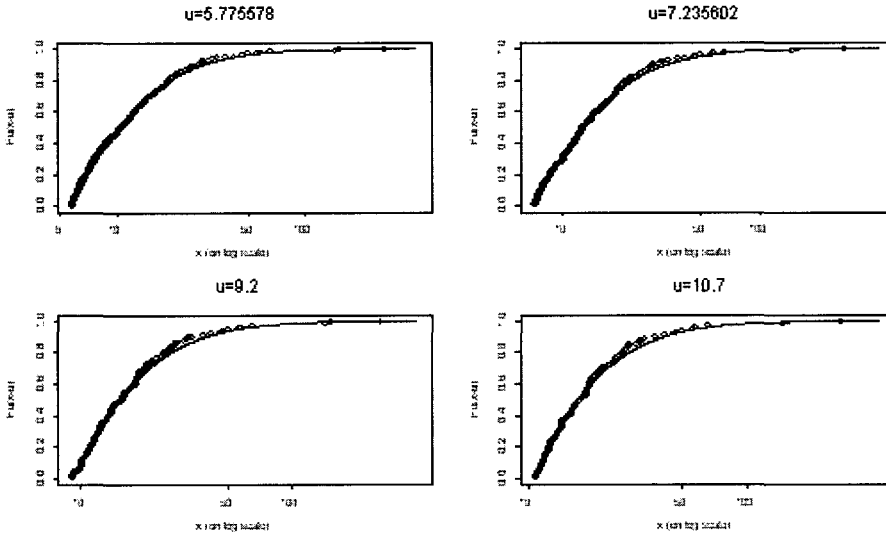


그림 4.3: 화재손실 데이터의 실증분포함수와 GDP분포함수

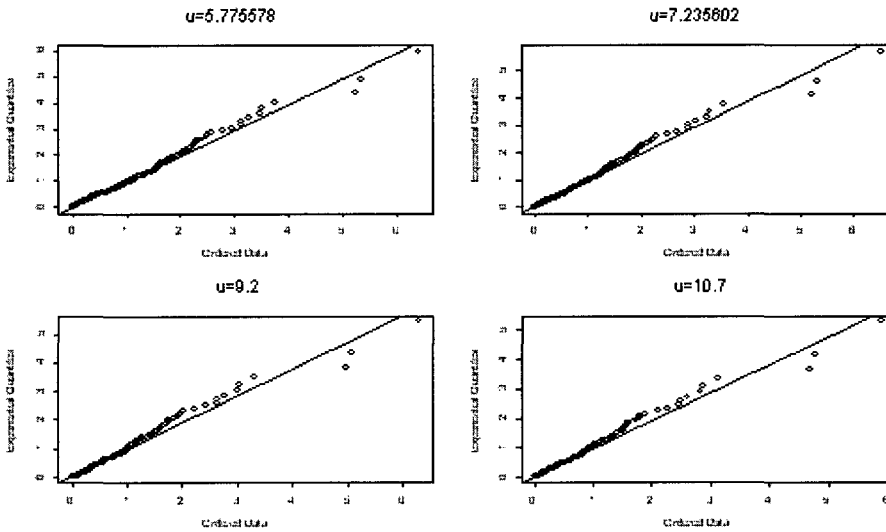


그림 4.4: 화재손실 데이터의 Q-Q 플롯

나타내고, 점선부분은 경험적 분포함수를 나타낸다. 그림 4.3에 나와 있는 네 개의 그림에서 점선부분은 거의 실선부분의 형태에 따라 분포되어 있다는 것을 확인할 수 있는데, 이는 실증자료가 GDP모형에 잘 부합하고 있다는 것을 의미한다. 이와 같은 결론은 Q-Q플롯을 통해서도 확인할 수 있다.

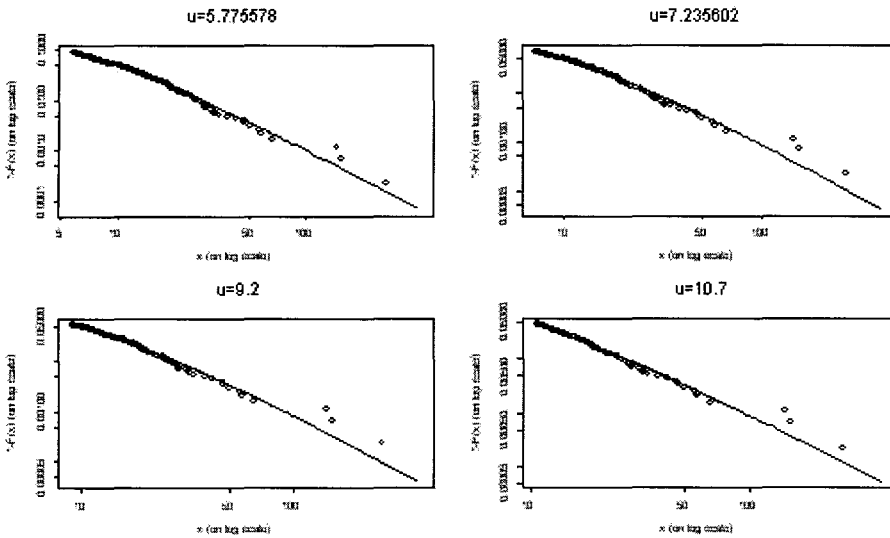


그림 4.5: 화재손실 데이터의 꼬리부분에 대한 GPD 모형의 적합도

표 4.2: 선택된 임계치에 상응하는 분위수와 초과치의 개수

임계치	분위수(%)	초과치의 개수
5.775578	9.13706	198
7.235602	6.92201	150
9.2	5.30688	115
10.7	4.42238	98

그림 4.4는 GPD를 기준분포로 그려진 임계치 초과값의 Q-Q플롯을 보여주고 있다. 네 개의 Q-Q플롯은 직선에 가까운 모양을 보여줌으로 데이터가 GPD모형에 잘 적합됨을 볼 수 있다. 그리고 그림 4.5는 화재손실 데이터에서 임계치를 초과하는 꼬리부분에 대한 실증생존함수와 GPD모형의 생존함수의 그래프를 보여주고 있다. 이 그림으로부터 화재손실 데이터의 임계치를 초과하는 꼬리부분이 GPD모형에 잘 적합됨을 알 수 있다.

### 4.3. GPD모형의 임계치와 모수추정

데이터가 GPD모형에 적합함을 살펴본 후, GPD모형의 모수를 추정하기 위해서는 먼저 임계치를 선택하는 것이 필요하다. 앞의 제 3절에서 언급한 바와 같이 임계치를 결정하기 위해 ME플롯 방식과 Hill플롯 방식을 적용하기로 한다.

그림 4.6은 화재손실 데이터에 대한 ME 플롯을 나타낸다. 이 그래프를 잘 살펴보면 전체적으로는 양의 기울기이지만 부분적으로 음의 기울기를 가지는 구간이 존재함을 보게 된다. 그런데 임계치  $u$ 값을 30이후로 잡을 경우 초과값의 개수가 12개 뿐이고, 이처럼 너무 적

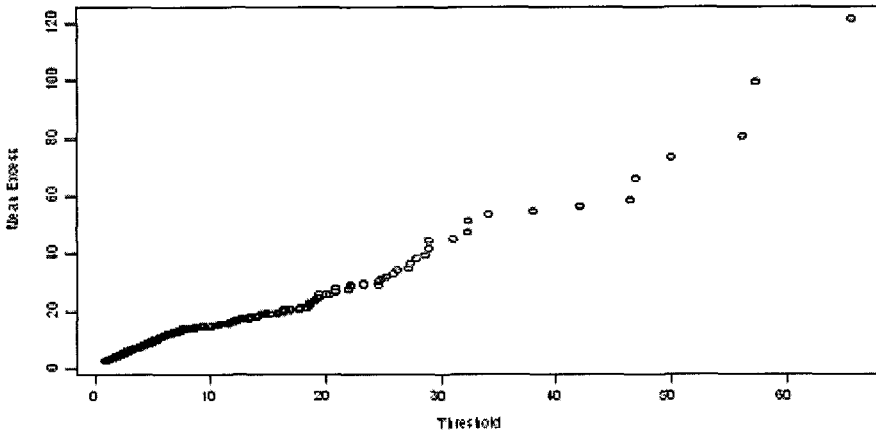


그림 4.6: 화재손실 데이터의 ME 플롯

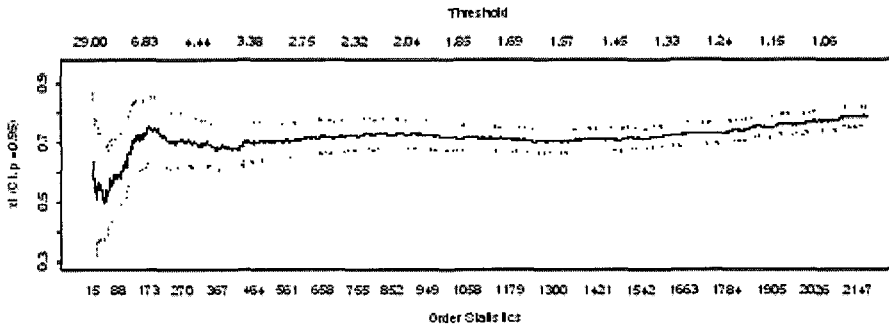


그림 4.7: 화재손실 데이터의 Hill 플롯

은 갯수의 임계치 초과값은 모수 추정의 오차를 초래할 수 있다. 이러한 점을 고려하여 먼저 임계치  $u$ 가 10.7 이후에서 지속적으로 양의 기울기를 보임을 알 수 있다. 따라서 10.7을 임계치  $u$ 의 값으로 선택한다. 동일한 방식을 적용하여 추가적으로 5.775578, 7.235602, 9.2를 임계치  $u$ 값으로 고려할 수 있다. 이와 같은 방법으로 선택된 각 임계치, 상응하는 표본의 분위수, 임계치 초과값의 개수가 표 4.2에 요약되어 있다.

한편, 그림 4.7은 화재손실 데이터에 대한 Hill 플롯을 보여주고 있다.  $u$ 값이 대략 5~15사이에 있을 때  $\xi$ 값이 안정적인 모습을 보이고 있으며 ME플롯 방식에 의해 선택된 임계치와 대체로 부합되는 모습을 보이고 있다. 즉, ME플롯 방식에 의해 선택된  $u$ 가 9.2와 10.7 근처에서  $\xi$ 값이 대체로 안정된 모습을 보이고 있다. 또한  $u$ 값으로 5.775578과 7.235602인 경우 초과치의 개수가 많아져서 모수 추정시 편향의 추정량이 될 가능성이 높아지므로 최종 임계점의 선택에서 배제하였다. 따라서 적절한 임계치로 10.7과 9.2의 두개만을 선택하기로 한다.

표 4.3: GPD모형의 모수 추정값

모형	임계치	모수추정값			
		$\hat{\xi}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\phi}^*$
GPD	9.2	0.4367526 (0.1210687)	130.4807 (40.30902)	60.5965 (31.25218)	7.626844 (1.123927)
GPD	10.7	0.4814762 (0.1379705)	137.6036 (47.04496)	68.64865 (38.92351)	7.547576 (1.238174)

$$*\hat{\phi} = \hat{\sigma} + \hat{\xi}(u - \hat{\mu})$$

이제 임계치  $u$ 값이 결정되면 다음 단계는 임계치 초과값 자료를 이용하여 GPD모형의 모수들을 추정하는 것이다. GPD의 모수를 추정하기위한 방법으로 본 논문에서는 최우추정법을 사용하였다. 임계치의 선택에 따라 GPD모형 하에서 모수추정치가 표 4.3에 제시되어 있다. 표 4.3을 보면 추정된 모든  $\hat{\xi}$ 값이  $-0.5$ 보다 큰 값으로 나타나므로 데이터의 분포가 두꺼운 꼬리를 가진다는 것을 의미하며,  $\hat{\xi}$ 값이 모두 양으로서 파레토타입 분포를 따른다는 것을 알 수 있다. 각 모수 추정치의 표준오차는 임계치의 값이 클수록 커지는데 이는 관찰치가 작아진 데 기인한다.

#### 4.4. GPD모형 하에서 분위수 추정

그림 4.8은 각각 99%, 99.5%, 99.9%, 99.93%의 분위수 추정치가 임계치의 선택에 의존한다는 사실을 보여주고 있다. 즉, 99.93% 분위수의 경우에 모형이 좋은 경우라면 100만 크로나를 초과하는 손실 중 매 10,000개마다 7개의 손실이 이 분위수를 초과한다고 예상할 수 있다. 그러한 손실들은 그 수가 매우 드물지만 보험회사 입장에서는 충분히 위협적일 수 있는 것이다.

표 4.4:  $\hat{\xi}$ 와 99.93% 분위수 추정치(단위: 100만 크로나)

모형	임계치	초과값 개수	$\hat{\xi}$	$\hat{\xi}$ 표준오차	분위수 추정치
GPD	9.20000	115	0.4367526	0.1210687	107.37286
GPD	10.7000	98	0.4814762	0.1379705	111.6602

표 4.5:  $\hat{\xi}$ 와 99.93% 분위수 추정치(최대치 제외)

모형	임계치	초과값 개수	$\hat{\xi}$	$\hat{\xi}$ 표준오차	분위수 추정치
GPD	9.20000	114	0.3340362	0.1149355	85.864057
GPD	10.7000	97	0.3695699	0.1306236	87.44088

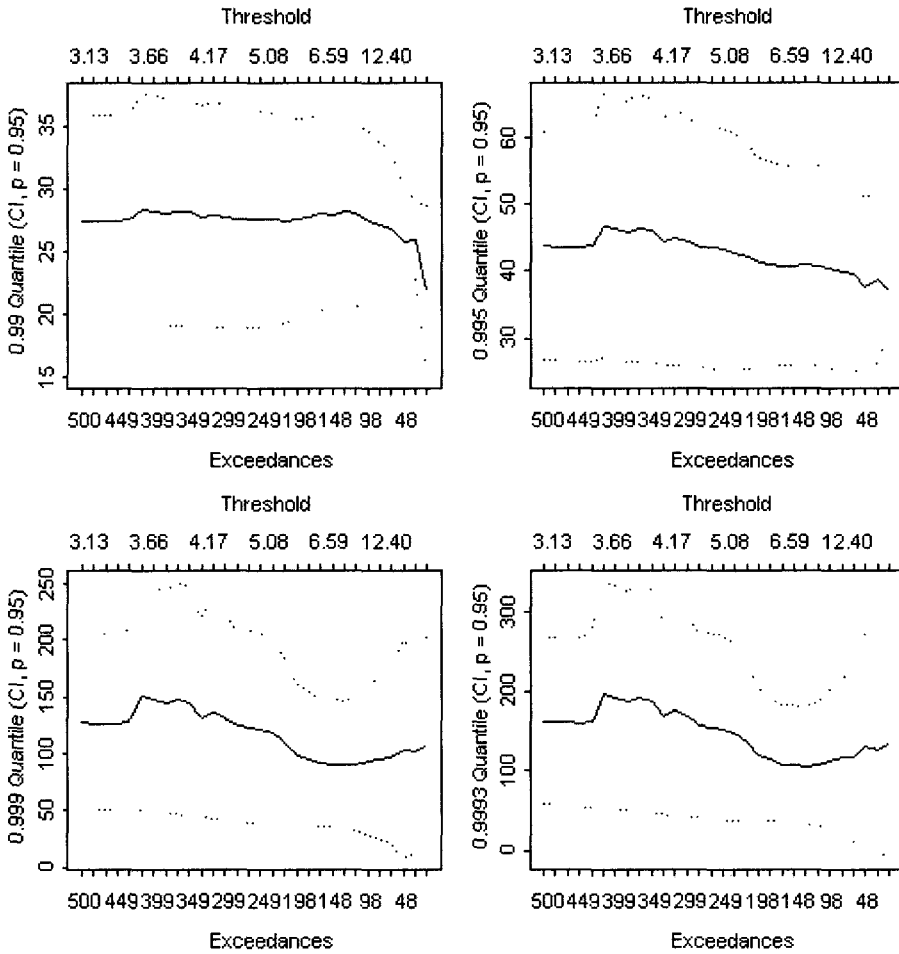


그림 4.8: 화재손실 데이터의 분위수 추정치

한편, 표 4.4는 99.93% 분위수에서의 초과값 개수, 형태모수의 추정치에 따른 분위수 추정치에 대한 결과를 보여주고 있다. 이러한 꼬리 부분에 대한 추정은 임계치의 선택에 매우 민감하다는 것을 알 수 있고, 특히 극단적인 손실의 경우에 얼마나 민감한지를 알아보기 위해 원래의 데이터에서 가장 큰 손실 한 개를 제외한 분석을 실시하였고 그 결과는 표 4.5에 나타나 있다. 표 4.4와 표 4.5를 비교하면 그 차이를 확실하게 알 수 있다. 원래 모형의 경우 임계치 10.7에서 초과값의 개수는 98개, 분위수 추정치는 111.6602으로 나타났으나, 가장 큰 손실 하나를 제외한 모형의 경우 임계치 10.7에서 초과값의 개수는 97개, 분위수 추정치는 87.44088로 나타나 크게 감소했음을 알 수 있다. 형태모수  $\hat{\xi}$ 의 추정값 또한 0.4814762에서 0.3695699로 크게 감소한다. 두 번째로 큰 손실까지 제외하고 분석하면 더욱 더 큰 감소를 보이게 된다.

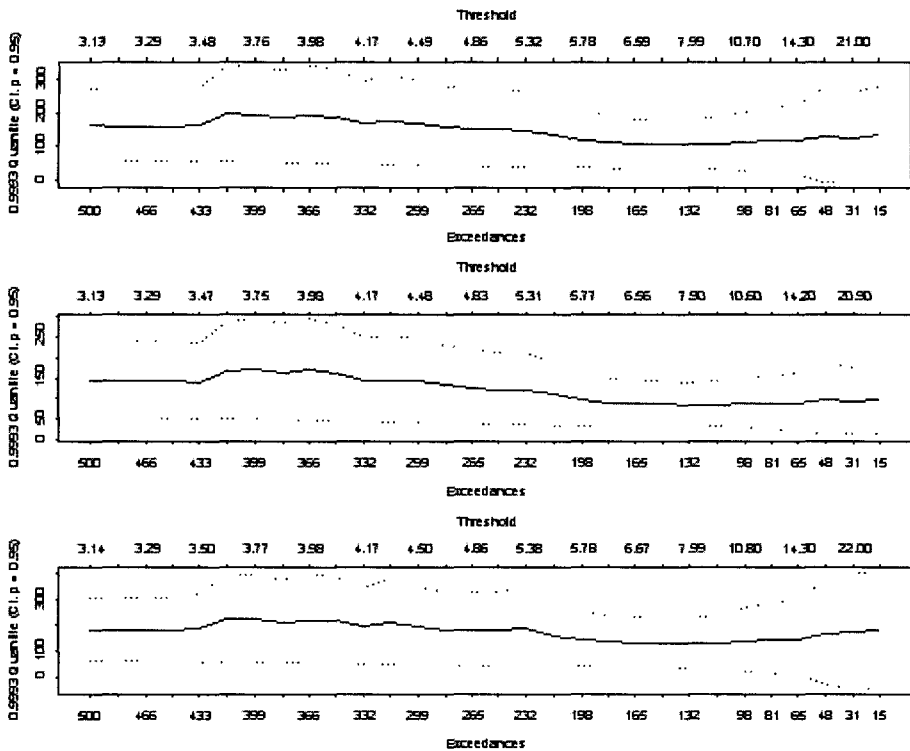


그림 4.9: 원 모형, 최대치 제거, 임의의 최대치 추가

표 4.6:  $\hat{\xi}$ 와 99.93% 분위수 추정치(최대치 추가)

모형	임계치	초과값 개수	$\hat{\xi}$	$\hat{\xi}$ 표준오차	분위수 추정치
GPD	9.20000	116	0.5181857	0.127278313	130.879188
GPD	10.7000	99	0.5701883	0.145316413	139.67928

한편, 원래 가장 큰 손실인 263.2504보다 더 큰 임의의 새로운 손실 270을 추가할 경우 임계치 10.7에서 초과값의 개수는 99개, 분위수 추정치는 139.67928로 크게 증가하며, 형태모수  $\hat{\xi}$ 의 값 또한 원래의 0.4814762에서 0.5701883으로 크게 증가함을 알 수 있으며, 표 4.6에서 나머지 임계치의 경우에도 동일한 사실을 확인할 수 있다. 그림 4.9는 각각의 경우에 대한 분위수 추정에 대한 그래프이다. 그래프의 비교를 통해서도 극단치가 분위수 추정에 어떠한 영향을 미치는지 확실히 알 수 있다. 이러한 변화는 데이터내의 최고 극단치들을 변화시킴에 따라 발생함을 알 수 있고, 추가 또는 감소되는 손실의 양이 적을수록 그 효과 또한 작아짐을 알 수 있다.



## 4.5. 위험자본의 추정

보험회사는 계약에 따른 보상책임에 대하여 그 모두를 부담하고 대량의 동질적인 리스크를 인수할 수 있어야 하겠지만, 보유능력을 초과하는 거대한 리스크나 위험분산이 어려운 소수의 독특한 리스크를 지닌 위험에 대하여는 당해 보험회사만으로는 그 모두를 부담할 수 없으므로 위험의 일부 또는 전부를 재보험으로 전가하는 것이 일반적이다. 재보험 또한 여러 가지 방식이 있으나, 본 논문에서는 2000만 크로나까지는 보험회사가 보유하고, 초과손실의 발생시 1억 크로나를 한도로 재보험회사가 보상을 하는 초과손해액 재보험(excess of loss reinsurance)을 가정하고 위험자본을 구하고자 한다.

표 4.7: 2000만 크로나 이상의 손실빈도 횟수

연도(1980-1990)	80	81	82	83	84	85	86	87	88	89	90
초과 손실의 횟수 ( $N_i$ )	3	4	5	0	0	3	1	4	8	5	3

표 4.7은 각 연도별로 2000만 크로나를 초과하는 손실의 횟수를 나타내고 있다. 표 4.7로부터 2000만 크로나 초과값들의 평균과 분산의 표본추정치는 각각 다음과 같이 얻어진다.

$$\hat{E}(N) = \frac{36}{11} = 3.27, \quad \hat{Var}(N) = \frac{\sum_{i=1}^{11} [N_i - E(N)]^2}{10} = 5.1074.$$

따라서 2000만 크로나 이상의 초과손실 횟수에 대한 분포를 포아송분포로 가정하고, 최우 추정치  $\hat{\lambda}$  을 구하면  $\hat{\lambda}=3.27$ 이 된다.

표 4.8: 포아송 분포에서의 손실 기대빈도

$N$	0	1	2	3	4	$\geq 5$
관찰빈도	2	1	0	3	2	3
확률값	0.038	0.124	0.203	0.221	0.181	0.232
기대빈도	0.418	1.364	2.233	2.431	1.991	2.552

표 4.8은 포아송분포 하에서의 관찰 빈도수, 추정확률, 기대빈도를 보여준다. 기대빈도는 포아송분포로부터 구한 각각의 확률값과  $N$ 의 곱으로 구해진다.

또한, 관찰값들이 포아송 분포를 따르는지에 대해 알아보기 위해 표 4.8에 나타난 결과를 바탕으로  $\chi^2$ 검정을 실시하면 다음과 같다.

$$\chi^2 = \sum \frac{(\text{관찰빈도} - \text{기대빈도})^2}{\text{기대빈도}} \sim \chi^2(k - p - 1).$$

여기서  $k$ 는 관찰그룹의 수이고  $p$ 는 추정된 모수값이다. 따라서  $\chi^2 = 8.528 < 9.488 = \chi_{0.05}^2(4)$

로서 유의수준 5% 하에서 관찰값들이 포아송분포를 따른다는 귀무가설을 기각할 수 없다.

선택된 최적 임계점에 따라 위험자본의 수준에 대해 알아보기 위해서  $r=2$ 천만 크로나에서  $R=1$ 억 크로나까지의 간격을 두면, 개별손실에 대한 기대가치는 식 (4.1)과 같이 주어진다.

$$E(Y_i|X_i > \delta) = \int_r^R (x-r)f_{X^\delta}(x)dx + (R-r)(1-F_{X^\delta}(R)). \quad (4.1)$$

식 (4.1)에서  $\int_r^R (x-r)f_{X^\delta}(x)dx$  부분은 손실발생액 중 초과손해액 재보험을 설정한 구간  $[r, R]$ 에 해당하는 손실을 의미하고,  $(R-r)(1-F_{X^\delta}(R))$  부분은 초과손해액 재보험의 최대 담보액  $R$ 을 초과한 손실을 의미한다. 또한  $F_{X^\delta}(x)$ 와  $f_{X^\delta}(x) = \frac{d}{dx}F_{X^\delta}(x)$ 는 각각  $\delta$ 에서 절단된 손실에 대한 분포함수와 확률밀도함수를 나타낸다. 따라서 2천만 크로나를 공제한 조건부 초과값을  $Y^{(20)} = (Y_i^\delta - 20|Y_i^\delta > 20)$ 으로 나타내면, 재보험을 고려한 개별 손실의 크기는 확률변수  $Y_i = \min(Y_i^{(20)}, 80)$ 로 나타내어지고, 개별 손실에 대한 기대가치는 다음과 같이 계산된다.

$$E(Y_i) = \int_{20}^{100} (y-20)f_{Y^\delta > 20}(y)dy + 80(1-F_{Y^\delta > 20}(100)).$$

따라서 총 손실의 기대가치는  $E(N)E(Y_i)$ 로서, 그 계산결과는 표 4.9에 나타나 있다.

표 4.9: 총 손실의 기대가치

임계치	$E(N)$	$E(Y_i)$	$E(N)E(Y_i)$
9.2	3.27	18.5783	60.751041
10.7	3.27	18.2935	59.819745

위험자본은 리스크 측면에서 기대가치의 변동성으로 정의된다. 즉, 예상평균 기대가치와 최악의 상황 발생 시의 기대가치와의 차액을 기준으로 위험자본을 산출한다. 이 때 기대가치란 회계상의 가치에 평가시점 현재 발생할 손익 중 B/S에 반영되지 않은 부분을 합산하여 산출한 기대가치를 말한다. 총 손실의 기대가치는  $E_{loss} = E(N)E(Y_i)$ 로서, S&P의 A등급에 해당하는 99.93% 분위수인  $Q_{99.93th}$ 를 최악의 상황 발생 시의 기대가치로 정의하면, 위험자본은  $ERC = Q_{99.93th} - E_{loss}$ 에 의해 측정할 수 있다.

표 4.10은 최적 임계점으로 선택한  $u=9.2, 10.7$ 에서 추정된 분위수와 손실의 기대가치 그리고 위험자본의 추정결과를 나타낸다. 표 4.10에서 볼 수 있듯이 임계치 9.2와 10.7에서 손실의 기대가치는 거의 유사하다. 그러나 위험 자본의 추정 결과 500백만 크로나의 차이를 보이는데, 이는 분위수 차이에 의한 결과로 볼 수 있다. 분위수간의 차이가 생긴 이유는 선택된 임계치 간의 초과값 개수가 차이를 보이고 있고, 이 초과값의 개수가 분위수 추정에 영향을 미치기 때문이라고 볼 수 있다.

## 5. 결론

IMF 금융위기에서도 알 수 있듯이 전 세계적으로 금융시장에서는 예측할 수 없는 대형 사고들이 지속적으로 일어나고 있으며, 특히 보험시장의 경우에는 대재해성(catastrophe)손

표 4.10: 최적임계점에서의 위험자본의 추정치

임계치	모수	$E_{loss} = E(N)E(Y_i)$	$Q_{99.93th}$	$ERC = Q_{99.93th} - E_{loss}$
	$\hat{\xi}=0.4367526$			
9.2	$\hat{\mu}=130.4807$ $\hat{\phi}=7.626844$	58.9581	107.37286	48.41476
	$\hat{\xi}=0.4814762$			
10.7	$\hat{\mu}=137.6036$ $\hat{\phi}=7.547576$	58.5715	111.6602	53.0887

실 등을 포함한 극단적 사고에 대한 예측이 날이 갈수록 어려워지고 있다. 많은 경우, 손실의 분포에 있어 정규분포보다 꼬리부분이 훨씬 두텁고 좌우 비대칭적인 형태를 보이고 있다. 따라서 정규분포에 근거하여 산출되는 위험수준에 대한 각종 지표들은 적절하다고 할 수 없다.

EVT모형은 극단적 사건들을 잘 반영할 뿐만 아니라 왼쪽 꼬리와 오른쪽 꼬리의 비대칭성을 고려할 수 있는 모형으로서 새로운 위험 측정방법으로 각광받고 있다. 본 연구에서는 EVT모형 중에서 특히 GPD모형을 사용하여 손실의 기대가치와 손해심도분포의 분위수를 추정하고 이를 바탕으로 적정위험자본의 추정치를 산출하고자 하였다.

본 논문의 실증분석 결과 화재손실 데이터의 분포는 비대칭적이면서 두터운 꼬리모양을 갖는 형태로서 GPD모형 중에서  $\xi > 0$ 인 파레토타입 분포를 따르는 것으로 알 수 있었다. 그런데 GPD모형을 적용하는 경우에 임계값의 정확한 선택이 중요한 점인데, 본 논문에서는 이를 해결하기 위해 ME플롯방식과 Hill플롯방식을 병행 사용하여 최적 임계치를 선택하고자 하였다.

끝으로 본 논문의 한계로 지적할 수 있는 것은 실증분석을 위해 국내 데이터를 구하지 못해 외국의 데이터를 사용한 점이라 할 수 있다. 또한 비상위험준비금과 같이 보험회사에서 적립하는 위험준비금과의 관계를 고려하지 않고 단순히 발생 손실액만을 대상으로 하여 위험자본을 산출했다는 점이라 할 수 있다. 따라서 추후 연구과제로 국내 데이터를 구해서 실증분석을 수행하면서, 위험준비금과의 관계를 연계시켜서 시나리오 분석, Stress Testing 등과 같은 위험분석기법을 사용하여 위험자본을 측정하고, 또한 리스크와 수익성을 동시에 고려하는 RAPM (risk adjusted performance measures)방식의 수익성 평가에 관한 연구도 필요할 것으로 판단된다.

## 참고문헌

- Balkema, A. A. and de Haan, L. (1974). Residual life time at great age, *The Annals of Probability*, **2**, 792-804.
- Coles, S. G. (2001). *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, London.

- Corradin, S. (2003). Economic risk capital and reinsurance: an extreme value theory's application to fire claims on an insurance company, *Astin Bulletin*.
- Embrechts, P., Klüppelberg, C. and Mikosh, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, Berlin.
- Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proceedings of the Cambridge Philosophical Society*, **24**, 180–190.
- Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution, *Annals of Statistics*, **3**, 1163–1173.
- Hogg, R. V. and Klugman, S. A. (1984). *Loss Distributions*, Wiley.
- Hosking, J. R. M. and Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, **29**, 339–349.
- Hürlimann, W. (2004). On the economic risk capital of portfolio insurance, *International Journal of Mathematics and Mathematical Sciences*, **41**, 2209–2218.
- Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological events. *Quarterly Journal of the Royal Meteorological Society*, **81**, 158–172.
- McNeil, A. J. (1997). Estimating the Tails of Loss Severity Distribution using Extreme Value Theory, *Astin Bulletin*, **27**, 117–137.
- Pickands, J. (1975). Statistical inference using extreme order statistics, *Annals of Statistics*, **3**, 119–131.
- Reiss, R. D. and Thomas, M. (2001). *Statistical Analysis of Extreme Values*, 2nd ed., Birkhauser Verlag, Basel.
- von Mises, R. (1936). La distribution de la plus grande de  $n$  valeurs. *Rev. Math. Union Interbalcanique*, **1**, 141–160. Reproduced in Selected Papers of Richard von Mises. *American Mathematical Society (1964)*, **2**, 271–294.
- Yeo, S. C. (2006). Performance analysis of VaR and ES based on extreme value theory, *The Korean Communications in Statistics*, **13**, 389–407.

[ 2006년 11월 접수, 2007년 2월 채택 ]

## Estimation of Economic Risk Capital of Insurance Company using the Extreme Value Theory

Sung Chil Yeo<sup>1)</sup> Dong Han Chang<sup>2)</sup> Byung Mo Lee<sup>3)</sup>

### ABSTRACT

With a series of unexpected huge losses in the financial markets around the world recently, especially in the insurance market with extreme loss cases such as catastrophes, there is an increasing demand for risk management for extreme loss exposures due to high unpredictability of those risks. For extreme risk management, to make a maximum use of the information concerning the tail part of a loss distribution, EVT(Extreme Value Theory) modelling may be the best to analyze extreme values. The Extreme Value Theory is widely used in practice and, especially in financial markets, EVT modelling is getting popular to analyze the effects of extreme risks. This study is to review the significance of the Extreme Value Theory in risk management and, focusing on analyzing insurer's risk capital, extreme risk is measured using the real fire loss data and insurer's specific amount of risk capital is figured out to buffer the extreme risk.

*Keywords:* Risk management, extreme value theory, GPD model, risk capital.

---

1) Professor, Department of Applied Statistics, Konkuk University, Seoul 143-701, Korea  
E-mail: scyeo@konkuk.ac.kr

2) Professor, Department of International Trade, Konkuk University, Seoul 143-701, Korea  
E-mail: dhchang@konkuk.ac.kr

3) Member, Tong Yang Financial Services Corp., Seoul 135-929, Korea  
E-mail: cutyboy77@nate.com