

# 재무비율의 극단치에 대한 통계적 분석

## Statistical Analysis of Extreme Values of Financial Ratios

주지환 (Jihwan Joo) KAIST 경영대학<sup>1)</sup>

### 〈 국문초록 〉

투자자들은 기업가치를 평가하기 위하여 재무비율을 활용하는데 특히 PER과 PBR은 적정 기업가치를 판단하는데 중요한 역할을 하는 대표적인 수치로 알려져 있다. 금융자료는 꼬리가 매우 두터운 형태의 분포를 따르는 경우가 많은데, PER과 PBR은 첨도가 매우 높으며 해당 재무비율의 극단치들은 기업의 다양한 이해관계자들의 의사결정 시 중요한 역할을 한다. 본 논문에서는 통계학의 극단치이론에서 주로 활용되는 GPD와 최근 새롭게 제안된 분포인 exGPD를 도입하고, 두 분포 간의 성능을 비교하기 위해 시뮬레이션을 수행하여 적합도를 살펴본 후 우측 꼬리에 속하는 90, 95, 99% 퍼센타일 값을 추정하여 실제 값과 비교한다. 다음으로 국내 증권시장에 상장된 정보기술군(IT) 기업들의 PER, PBR 자료에 근거하여 실증분석을 수행한다. 분석 결과 특히 PBR에서 exGPD가 GPD에 비해 자료의 우측 꼬리 영역을 보다 효과적으로 설명함을 확인하였다. 따라서, 재무비율에 기반한 기업가치평가 또는 위험관리 시 극단치의 특성을 효과적으로 반영할 수 있는 exGPD와 같은 분포를 활용한다면 꼬리 영역에 담긴 정보를 보다 정확하게 파악할 수 있다. 이는 기업 내부 위험관리자의 효과적인 지식경영을 돕고, 투자자를 비롯하여 다양한 외부 이해관계자들에게 유용한 지식을 제공할 수 있다.

주제어: 지식경영, 주가수익비율, 주가장부가치비율, 극단치이론

1) 제1저자 및 교신저자, jh1012.joo@kaist.ac.kr

## 1. 서론

### 1.1. 연구의 배경과 목적

재무비율(Financial Ratio)은 재무제표상의 한 항목을 다른 항목의 값으로 나눈 상대적인 비율이다. 이는 기업의 재무상태 및 경영상태를 나타내 주는 간결한 지표로, 비율분석을 통해 기업 간의 상대적 척도를 비교, 평가하는데 도움이 되는 지식을 제공하기 때문에 다양한 이해관계자들의 의사결정에 널리 활용된다. 단지 재무비율에만 의존하여 기업가치를 평가하는 것에는 한계가 있지만, 이는 복잡한 정보들을 하나의 수치로 간결하게 요약해 준다는 점에서 활용도가 매우 높아 기업의 경영자들 또한 해당 지표를 관리하는데 많은 관심을 가지고 있다.

기업의 재무상태를 요약하는 재무비율의 종류는 다양한데 투자 의사결정에서는 시장가치비율(Market Value Ratios)이 중요하게 여겨진다. 투자자들은 주가 수익비율(Price-to-Earnings Ratio, PER)이나 주가장부 가치비율(Price-to-Book Ratio, PBR)을 가장 핵심적인 지표로 여기는데, 이는 재무제표의 핵심정보를 요약한 값이며 기본적으로 해당 수치가 낮으면 기업가치가 저평가된 것으로 보아 투자 시 초과수익률을 기대할 수 있다. 반대로 위 재무비율이 높으면 기업가치가 과대평가됨을 의미하여 이에 투자하면 손실위험이 증가하는 것으로 본다. 즉, 위의 재무비율들은 적정 기업가치를 평가하는데 중요한 단서가 되는 지식을 제공한다.

금융자료를 통계적으로 분석할 때 해당 자료가 정규분포를 따른다고 가정하는 경우가 종종 있으나 첨도가 높은 경우에는 극단치에 담긴 정보를 제대로 반영하지 못하게 되어 투자 또는 위험관리 등과 관련된 의사결정에 다양한 문제를 일으킬 수 있다. 이 때문에

금융자료를 보다 잘 적합할 수 있는 방법론들이 개발되어 왔는데 극단치이론(Extreme Value Theory, EVT)을 활용하는 방법이 대표적으로 존재한다(강민정 등, 2013). 특히 VaR(Value at Risk) 등의 위험지표를 개선하기 위해 이를 적용한 연구가 많은데, 주로 기업 내부의 위험관리 책임자가 관련되는 업무를 원활하게 수행하는데 도움을 주기 위한 목적으로 수행되었다. 예를 들어 가장 활발하게 연구되었던 주식의 로그수익률에 대한 VaR의 계산에서는 투자수익률이 매우 낮은 왼쪽 꼬리의 1% 또는 5%에서의 백분위수 값을 추정하게 된다. 즉, 일정한 조건 하에서 발생할 수 있는 최대손실금액을 산출함으로써 투자포트폴리오 위험을 측정하였다. 기존의 실증분석에서는 주로 주식이나 환율의 로그수익률에 대한 일별 시계열자료를 활용하는데, 이는 정규분포보다 꼬리가 두터운 형태를 가진다. 두터운 꼬리를 효과적으로 반영하기 위해 일정한 임계값을 넘어서는 극단값들에 대하여 기존의 분포가 아닌 일반화파레토분포(Generalized Pareto Distribution, GPD)를 새롭게 적용하려는 시도를 하였다(Kellezi and Gilli, 2000; 오세경, 2005). 재무비율 중 PER과 PBR 값은 유동성비율이나 부채비율 등과 같이 직접적으로 재무적 위험을 나타내는 지표는 아니지만, 주식회사의 입장에서 주가의 변동은 기업가치나 경영자 보상 등에 많은 영향을 미치므로 해당 재무비율에 대해 면밀히 분석하고 지속적으로 모니터링하는 것은 필수적이다. 특히 PER과 PBR이 동종 산업군 내 다른 기업들보다 매우 큰 값을 나타내는 경우, 투자자들은 보유한 주식을 매도하고자 하므로 향후 기업가치가 하락할 가능성이 높아진다. 따라서, 본 논문에서는 극단치이론에서 전통적으로 활용되어 온 GPD와 최근 Lee and Kim(2019)에서 소개된 지수화된 일반화파레토분포(Exponentiated GPD, exGPD)에 근거하여 이러한 분포들이 재무비율의 극단치에 담긴 정보를

얼마나 효과적으로 설명하는지 비교 및 분석하였다.

본 논문의 구체적인 연구 동기는 다음과 같다. 첫째, 기존의 연구에서는 꼬리의 두께를 나타내는 첨도가 3인 정규분포보다 크다는 이유로 주식의 로그수익률을 이용하였으나 이는 첨도(Kurtosis)가 10을 넘어가지는 않으며 대부분 4에서 8 사이의 값을 나타낸다. 반면 PER이나 PBR와 같은 재무비율이 갖는 첨도는 수백에서 수천의 값을 가지며 꼬리가 두터운 자료의 특성이 보다 강하게 나타나기 때문에 해당 이론을 적용하는데 보다 적절한 금융자료가 된다. 둘째, 극단치 이론에서 전형적으로 사용되는 일반화파레토분포가 아닌 이를 개선할 수 있는 다른 분포를 도입하고 성능을 비교하여 두터운 꼬리 영역의 자료를 더욱 면밀하게 관찰하고 금융자료에 대한 적합도(Goodness of Fit)를 향상시킬 수 있음을 보여준다. 본 연구에서는 기존 GPD 뿐만 아니라 이를 로그변환하여 도출되는 exGPD를 활용해 비교 분석한다. 셋째, 경영학 선행연구들은 재무비율에 대하여 투자지표로서의 유용성을 확인하는 것에 초점을 맞추고 있는데, 재무비율의 극단치에 담긴 정보에는 관심이 상대적으로 적다. 그러나 재무비율의 극단치는 일반적으로 재무적 위험관리의 측면에서 핵심적인 지식을 포함하기 때문에, 극단치이론을 통하여 해당 기업의 재무적 위험 수준을 보다 정확하게 파악하고 궁극적으로는 효율적인 지식경영을 수행하는데 도움을 줄 수 있다. 넷째, 지식 집적 산업인 정보기술(IT)군 기업들을 대상으로 실증분석을 수행함으로써, 지식 사회의 핵심 산업군으로 주목 받는 정보기술군 기업들의 재무적 위험관리 및 자본시장 효율성 향상에 기여하고자 한다(김병수 등, 2010).

본 연구의 시사점은 다음과 같이 크게 세 가지로 정리할 수 있다.

우선, 재무적 위험관리의 측면에서 기업의 재무상

태에의 위험 수준을 보다 정밀하게 모형화하고 정확하게 파악함으로써 기업 내부의 위험관리 책임자가 관련된 업무를 원활하게 수행하는데 도움이 되는 지식을 제공하는 역할을 할 수 있다. 또한 해당 재무비율 자료는 투자자들의 투자 의사결정에 주로 활용하는 핵심 지표인 만큼 주식 투자자를 비롯하여 애널리스트나 신용평가기관 등의 다양한 외부 이해관계자들이 해당 기업에 대한 재무적 위험 수준을 보다 정확하게 파악하는데 유용한 지식을 제공하며, 궁극적으로는 자본시장의 효율성을 향상시키는데 도움이 된다. 본 논문에서 분석 대상으로 선정한 PER과 PBR 이외에도 첨도가 매우 높으며 재무적 상황을 평가하는데 활용될 만한 다른 여러 재무비율들로 쉽게 확장할 수 있다.

다음으로, 재무비율 정보는 기업 신용평가 및 부실예측을 위한 핵심정보로 활용될 수 있는데, 이와 관련된 연구에서는 부실기업과 건전기업 간 재무비율의 분포가 체계적으로 다르다는 가정을 전제로 한다. 일반적으로 위의 두 그룹 간 분포의 형태는 꼬리 영역에서 서로 맞물리게 되고, 이진분류 예측을 수행하는 여러 통계적 모형들은 제1종 오류와 제2종 오류를 최소화할 수 있는 특정 임계값을 정해야 하는 문제에 직면한다. 여기에서 exGPD를 활용한 모형은 오분류율을 감소시키는 역할을 할 수 있으므로, 부실예측 및 신용분석 등의 연구에서 재무비율이 가지는 정보를 더 효과적으로 반영하는데 도움을 준다. 실제로 많은 재무비율들은 꼬리가 두터운 분포를 가지므로, PER이나 PBR 이외에 다양한 종류의 재무비율을 활용하는 신용평가나 부도예측과 관련된 여러 모형들의 성능을 개선시킬 수 있다. 예를 들어, Altman(1968)은 부실예측을 위하여 Z-Score(판별점수)를 산출하는 모형을 개발하였는데, 총 5가지 재무비율 각각에 가중치를 부여하고 Z값을 산출하여 해당 값이 특정 점수(2.99)를 초과하면 건전기업으로, 반대로 또 다른 임계치(1.81)보

다 작으면 부실기업으로 판별하였다. Altman은 Z값을 산출하는데 쓰이는 재무비율들이 다변량 정규분포를 따른다고 가정하였는데, 해당 재무비율들이 정규분포 가정 때문에 꼬리 부분에 대한 정보를 제대로 반영하지 못하게 만들어 예측모형의 오분류율을 높일 수 있다. 이처럼 고위험 기업을 재무비율에 기초해 판단하는 모형에서 분포의 꼬리 영역이 exGPD와 같은 분포를 따르는 것으로 가정한다면 판별 능력을 보다 향상시킬 수 있다.

마지막으로, 기존의 위험관리 지표를 산출하는 과정에서 exGPD와 같은 분포를 새롭게 활용할 수 있다. 예를 들어, 금융기관의 위험관리에 가장 널리 활용되는 기법인 VaR(Value at Risk)은 하향위험(Downside Risk)의 한 척도로, 특정 허용수준 하에서 발생 가능한 최대 손실금액이나 수익률을 계산하는 지표다. 실무에서는 VaR을 계산하기 위하여 정규분포를 주로 가정하는데, 가정한 분포로부터 95% 또는 99% 신뢰수준 값을 구해 이를 기반으로 VaR을 측정하게 된다. 여기서 수익률과 같은 자료는 꼬리가 두터운 분포를 적용해야 더욱 효과적으로 측정할 수 있으며, 정규분포 외에 일반화파레토분포(GPD) 등 여러 분포에 대한 논의가 있어왔다. 최근 Lee and Kim(2019)의 연구에서 exGPD의 이론적 기반을 확립함으로써, exGPD를 VaR 등의 다양한 위험관리 측정모형에 활용할 수 있게 되었다. 이러한 상황에서 극단치이론 연구에서 이론적으로 정립된 새로운 확률분포인 exGPD의 우수성을 확인하여 지식경영 분야에 유용한 의미를 제공한다.

## 1.2. 논문의 구성

본 논문은 다음과 같이 구성되었다. 1장은 연구의 배경과 목표를 설명한 서론이다. 2장에서는 분석에 활용한 주요 재무비율의 기본적 특성과 일반화파레토분

포의 이론적 배경을 살펴본 뒤 exGPD로의 수리적 변환 과정을 소개한다. 다음으로 GPD와 exGPD의 분포적 특성을 그래프로 비교하여 각 확률분포의 기본적인 특성을 비교하여 exGPD에서 개선된 사항들을 살펴본다. 다음으로 3장에서는 시뮬레이션을 통하여 두 분포 간 성능을 비교 및 분석한다. 꼬리가 매우 두터운 가상의 자료를 생성한 후, 우측 극단치에 대해 각 분포별 적합도를 확인한다. 또한 위험관리를 목적으로 산출한 백분위수(Percentile) 추정값이 얼마나 정확한지 비교한다. 4장에서는 재무비율 데이터에 근거한 실증분석을 3장과 동일한 절차로 실시한 뒤, 특히 exGPD의 우수성을 실제 자료에 근거해 확인한다. 추가적인 분석으로 최근 6개년 자료에 근거하여 최종 선택된 분포로부터 재무비율 값을 추정하고 추정된 값들을 연도별로 비교해 본다. 마지막으로 5장에서는 연구결과를 요약하고 결론을 내린다. 또한 본 연구의 한계점 및 향후 과제를 의논한다.

## 2. 이론적 배경과 선행연구의 고찰

### 2.1. PER과 PBR의 개념 및 유용성

먼저 PER(Price-to-Earnings Ratio, 주가수익비율)은 개별 주식의 시장가격을 그 기업의 1주당 순이익(Earnings Per Share, EPS)으로 나눈 것을 말하며 단위는 배이다. 즉, PER은 주당순이익의 몇 배가 주식가격으로 형성되는지를 뜻한다. 예를 들어 PER이 10인 회사에 투자하면 투자원금을 회수하는데 10년이 소요된다고 해석할 수 있다. 주당순이익은 해당 사업연도의 순이익에서 우선주가 있는 경우 우선주배당금을 차감하고 난 후의 보통주(당기)순이익을 보통주 평균발행주식수로 나누어 계산된다.

PER은 직관적으로 이해하기 쉽다는 장점이 있고 기업가치를 간단한 방법으로 추정할 수 있기 때문에 투자 의사결정에 많이 활용된다(이영한과 박주영, 2016). 실제 PER이 동종 산업 PER의 평균 등을 이용해 도출된 정상적인 PER에 비해 상대적으로 높다면, 시장가격이 지나치게 비싸 주식이 과대평가되었다는 의미다. 반대로 실제 PER이 정상적인 PER보다 낮으면 기업의 이익에 비해 주가가 상대적으로 낮다는 것이므로 주식이 과소평가되었음을 뜻한다. 현실에서 PER이 낮은 주식이 반드시 저평가되었다고 볼 수는 없으나, 이는 초과수익률을 얻는데 도움을 주는 유용한 투자지표로 인정받고 있다.

여러 선행연구들은 PER을 기반으로 기업가치평가를 수행했을 때 정확도가 우수함을 확인하였다(Leclair, 1990; Kaplan and Ruback, 1995). 또한 PER을 사용하여 기업가치를 평가하고 투자의사결정을 할 때 동종 산업군 내에서 이루어져야 바람직함을 다양한 선행연구들에서 검증하였다. Alford(1992)는 산업군별 분석을 수행하여 PER을 사용했을 때 해당 재무비율이 내포하고 있는 재무적 위험 및 이익성장률을 효과적으로 파악할 수 있다고 주장하였다. 이후 주가배수를 이용한 관련 선행연구들은 PBR을 추가적으로 고려함으로써 PER과 PBR을 함께 사용하면 분석 정확도가 향상됨을 확인하였다(Beatty et al., 1999; Cheng and McNamara, 2000).

다음으로 PBR(Price-to-Book Ratio, 주가장부가치비율)은 개별기업 주식의 시장가격을 해당 기업의 1주당 순자산 또는 자기자본의 1주당 부가가치로 나눈 비율로, 자산가치가 반영된 상대적 주가수준을 측정하는 지표다. 다시 말해서 주가가 순자산에 비해서 1주당 몇 배로 거래되고 있는가를 나타낸다. 예를 들어 주당 순자산가치가 만원인 기업의 주가가 5천원일 경우 PBR은 0.5가 되는데, 이 경우에는 주가가 청산가치

에도 못 미치는 셈이다.

일반적으로 높은 수익성을 가지면서 내부유보가 많은 기업일수록 PBR은 낮다. PBR이 1이면 특정시점의 주가와 기업의 1주당 순자산이 동일하다는 의미다. PBR이 낮을수록 해당 기업의 자산가치가 증시에서 저평가되었음을 의미한다. 주당순자산가치(장부가치)가 실질가치를 정확히 반영할 경우 PBR은 1이 되어야 하나 현실에서 대부분의 기업들은 그렇지 않다. 기업마다 경영여건의 제반 요인들에 따라 미래의 현금흐름이 다르고 주당순자산의 질적 차이도 반영되기 때문이다. PER과 마찬가지로 PBR을 이용하면 미래의 이론적인 주가를 추정할 수 있다. 종합적으로 볼 때, PER은 전통적으로 기업가치를 추정하는데 가장 유용한 재무비율이고 PBR을 함께 활용하면 보다 정확한 정보를 얻을 수 있게 되며 동종 산업군 내 기업들 간에 재무비율 분석이 이루어져야 효과적임을 보여주었다(이영한과 박주영, 2016).

다만 위와 같은 재무비율 분석에서 주의해야 할 점은 기업 간 비율의 상대적인 높고 낮음에 대한 가치판단은 비교 가능한 동종 또는 유사 기업집단 내에서만 가능하다는 것이다. 재무비율이 갖는 편리함의 이면에는 정보의 손실 등 한계점도 공존한다. 궁극적으로 투자의 성과를 극대화하기 위해서는 재무비율을 포함한 재무제표만을 이용해 투자하는 것보다는 경제 및 산업분석을 통해 거시적인 상황을 우선 파악한 후, 재무제표 정보를 활용하는 것이 바람직하다.

## 2.2. 극단치이론 및 확률분포

### 2.2.1. 극단치이론 및 일반화파레토분포(Generalized Pareto Distribution, GPD)

극단치이론(Extreme Value Theory, EVT)은 임의의 분포를 가지는 확률변수가 취할 수 있는 여러 가능한

값들 중에서 분포의 중앙에서 크게 벗어난 꼬리부분의 자료에 대한 분포이론이다(Hill, 1975; Dekkers et al., 1989; Coles, 2001; De Hann and Ferreira, 2006; Magnou, 2017). 즉, 전체자료가 아닌 꼬리부분에만 집중하여 특정 임계값을 초과하는 극단치들에 분포를 적용시킨다. 이를 이용하면 극단적인 자료를 더욱 정확하게 관측할 수 있다. 극단치를 정의하기 위한 방법은 Block Maxima(BM) 방법과 Peaks Over Threshold(POT) 방법이 있는데, 전자는 극단치들을 여러 그룹으로 나눈 다음 각 그룹의 최대값을 골라 극단값으로 정의하고 후자는 특정 값을 임계값으로 결정한 후 이를 초과하는 값들을 극단치로 정의한다. 여기에서 임계값보다 큰 자료들의 초과분포는 임계값이 무한대로 증가할 때 일반화파레토분포(GPD)로 수렴함이 이론적으로 증명되었다. POT 방법을 사용하는 경우에 분포로는 GPD를 적용한다. 둘 중에서 POT 방법은 자료를 더욱 효율적으로 사용한다. 따라서 본 논문에서는 이를 이용하였다. 또한 최적 임계값의 결정은 매우 중요한 문제로 다양한 방법이 존재한다. R의 함수에 의해 관심대상인 PER와 PBR 자료의 임계값을 직접 추정해 본 결과 90% 백분위수 근방에서 결정됨을 확인하였다. 따라서 90% 백분위수를 임계값으로 결정하고 이후의 분석을 진행하였다.

POT 방법은 전체 데이터를 오름차순으로 배열하였을 때 특정 임계값을 넘는 관측치들의 극한분포로 정의된다. 확률변수의 수열  $X_1, X_2, \dots, X_n$ 이 *i.i.d.* (independent and identically distributed)이고 알려지지 않은 한 분포  $F$ 에서 추출되었다고 하자. 특정 임계값 ( $u$ )을 초과하는 값들에 대하여 분포함수  $F_u$ 를 추정하는 것이 목표이다. 분포함수  $F_u$ 는 임계치 초과분포함수라 하며 이는 다음과 같이 정의된다.

$$F_u(y) = P(X-u \leq y \mid X > u) = \frac{F(u+y) - F(u)}{1 - F(u)}; 0 \leq y \leq x_F - u \quad (1)$$

단,  $x$ 는 확률변수,  $u$ 는 주어진 임계값,  $y = x - u$ 는 일정 손실크기를 넘는 초과손실,  $x_F \leq \infty$ 은  $F$ 의 오른쪽 끝점(right endpoint)이다. Balkema and De Haan(1974) 및 Pickands(1975)의 정리에 의하면 위의 초과분포함수  $F_u$ 는 임계값( $u$ )이 양의 무한대로 증가할 때 GPD로 수렴한다. 2개의 모수를 갖는 GPD는 다음과 같은 방법으로 정의된다. 따라서, GPD의 누적분포함수는 다음과 같다.

$$G_X(x) = 1 - (1 + \frac{\xi x}{\sigma})^{-\frac{1}{\xi}}; \xi \neq 0 \quad (2)$$

단,  $x$ 의 범위는  $\xi > 0$ 일 때  $x \geq 0$ 이고,  $\xi < 0$ 일 때  $0 \leq x \leq \frac{-1}{\xi}$ 이다.

$$G_X(x) = 1 - e^{-\frac{x}{\sigma}}; \xi = 0, x \geq 0. \quad (3)$$

따라서, GPD의 확률밀도함수는 다음과 같다.

$$g_X(x) = \begin{cases} \frac{1}{\sigma} (1 + \frac{\xi x}{\sigma})^{-\frac{1}{\xi}-1}; \xi \neq 0 \\ \frac{1}{\sigma} e^{-\frac{x}{\sigma}}; \xi = 0 \end{cases} \quad (4)$$

단,  $\sigma$ 는 척도(scale)모수이고  $\xi$ 는 형태(shape)모수이다.

이 분포는 일반화된 형태이므로  $\xi$ 의 범위에 따라 세 가지의 분포로 나누어진다.  $\xi > 0$ 이면 두터운 꼬리를 갖는 Type I 파레토분포나 t분포 등이 해당되는데 이는 금융자료의 형태와 유사하다. 또한  $\xi < 0$ 이면 Type II 파레토분포,  $\xi = 0$ 이면 지수분포에 속한다 (Grimshaw, 1993; Castillo and Hadi, 1997; Zhang, 2007; De Zea Bermudez and Kotz, 2010).

2.2.2. 지수화된 일반화파레토분포(Exponentiated GPD, exGPD)

Lee and Kim(2019)에 의하면, exGPD의 함수는 GPD의 분포함수 대하여 로그변환을 취함으로써 얻을 수 있다.  $X \sim GPD$ 이고  $Y = \log X$ 라 하자.

$$F_Y(y) = P(Y \leq y) = P(\log X \leq y) = P(X \leq e^y) = G_X(e^y) = 1 - (1 + \frac{\xi e^y}{\sigma})^{-\frac{1}{\xi}}; \xi \neq 0 \tag{5}$$

단,  $\xi > 0$ 일 때  $-\infty < y < \infty$ 이고,  $\xi < 0$ 일 때  $-\infty < y \leq \log(\frac{-1}{\xi})$ 이다.

$$F_Y(y) = 1 - e^{-\frac{y}{\sigma}}; \xi = 0, -\infty < y < \infty. \tag{6}$$

따라서, exGPD의 확률밀도함수는 다음과 같다.

$$f_Y(y) = \begin{cases} \frac{e^y}{\sigma} (1 + \frac{\xi e^y}{\sigma})^{-\frac{1}{\xi} - 1}; \xi \neq 0 \\ \frac{1}{\sigma} e^{-\frac{y}{\sigma}}; \xi = 0. \end{cases} \tag{7}$$

위 함수에서는 모든  $\xi$ 의 범위에 대하여 정의하였으나, 본 논문에서는 주로 양수 값을 갖는 경우에 대하여 논의하고자 한다.

확률변수의 적률(moment)은 분포의 형태를 결정하는데 매우 중요한 역할을 한다. 1차와 2차 적률을 통하여 exGPD 분포의 특징을 살펴보고자 한다. 먼저 exGPD의 1차 적률은 다음과 같다.

$$E(Y) = \log(\frac{\sigma}{\xi}) + \xi - \sum_{k=1}^{\infty} (\frac{1}{k(\xi k + 1)}); \sigma, \xi > 0. \tag{8}$$

GPD와 달리 exGPD의 경우에는 척도모수( $\sigma$ )가 위치모수의 성격 또한 가진다. 왜냐하면 1차 적률이  $\sigma$ 에 대한 증가함수로 표현되기 때문이다. 또한, GPD의 1차 적률이 존재하는 범위가 제한적( $\xi < 1$ )인데 반하여, exGPD는 1차 적률의 존재조건이 완화됨을 확인할 수 있다.

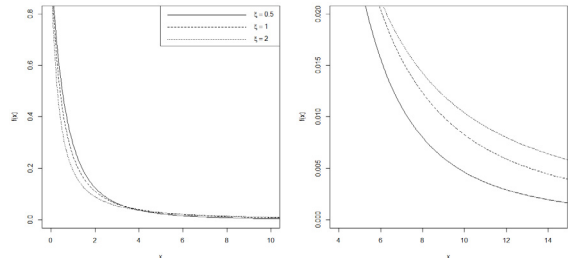
다음으로 2차 적률은 아래와 같이 표현할 수 있다.

$$E[Y^2] = 2\xi^2 + 2\xi \log(\frac{\sigma}{\xi}) + (\log(\frac{\sigma}{\xi}))^2 - 2\log(\frac{\sigma}{\xi}) (\sum_{k=1}^{\infty} \frac{1}{k(\xi k + 1)}) - \sum_{k=1}^{\infty} (\frac{2\xi}{k(\xi k + 1)^2}) + \sum_{k=1}^{\infty} H_{k-1} (\frac{2}{k(\xi k + 1)}); \sigma, \xi > 0. \tag{9}$$

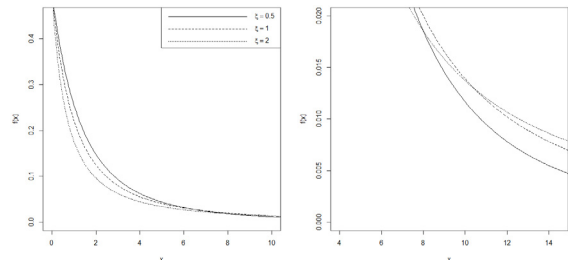
exGPD의 2차 적률은  $\sigma$ 와  $\xi$ 가 양수일 때 유한한 값을 가지며 잘 정의된다. 반면에 GPD의 경우에는  $\xi < \frac{1}{2}$ 라는 비교적 엄격한 조건을 가지고 있다.

2.2.3. GPD와 exGPD 간 특성 비교

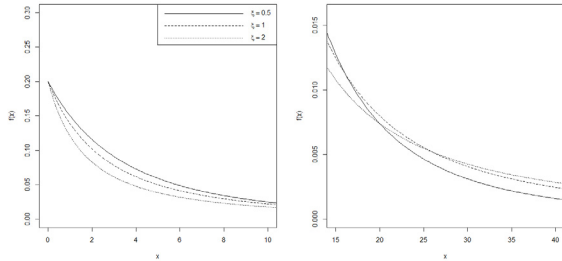
두 분포의 모수 변화에 따른 확률밀도함수의 형태를 시각적으로 비교하고자 한다. 이를 위하여 지정된 모수의 조합마다 그래프를 제시하였다. 특히 우측꼬리의 두께를 면밀히 살펴보기 위하여 각 그림마다 이를 확대해 보았다.



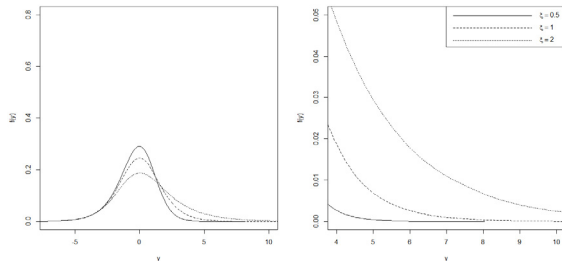
<그림 1>  $\sigma=1$ 일 때 GPD의 확률밀도함수 그래프



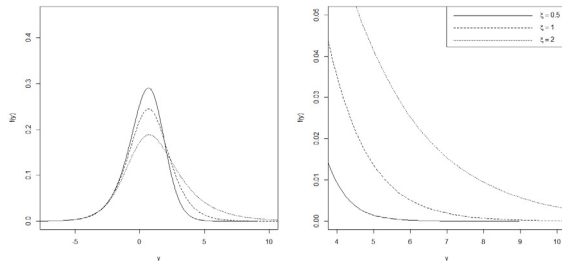
<그림 2>  $\sigma=2$ 일 때 GPD의 확률밀도함수 그래프



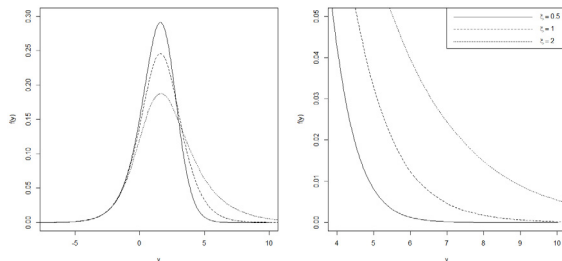
〈그림 3〉  $\sigma=5$ 일 때 GPD의 확률밀도함수 그래프



〈그림 4〉  $\sigma=1$ 일 때 exGPD의 확률밀도함수 그래프



〈그림 5〉  $\sigma=2$ 일 때 exGPD의 확률밀도함수 그래프



〈그림 6〉  $\sigma=5$ 일 때 exGPD의 확률밀도함수 그래프

위의 그림들은 총 9가지의 모수의 조합  $\{(\sigma, \xi) | \sigma=1, 2, 5; \xi=0.5, 1, 2\}$ 마다 GPD와 exGPD의 확률밀도함수 그래프이다. 전반적으로 볼 때, 두 분포간의 차이점은

0보다 큰 데이터에 대하여 로그변환을 실시하면 음수를 포함한 모든 실수로 범위가 확장되는 점이다. 또한, 기존의 우측 꼬리부분에 해당하는 자료들은 로그변환을 함으로써 상당히 작은 값으로 이동함과 동시에 조밀해진다. 재무비율과 같은 자료의 경우 첨도가 매우 크고 우측 꼬리가 과도하게 길어 비정상적인 소수의 극단값들이 존재하므로 기존의 GPD를 통해서 꼬리부분을 설명하기에 한계가 있다. 반면 exGPD에서는 비현실적인 극단값들이 로그함수의 특성에 의해 분포의 중심에서 크게 벗어나지 않아 통제하기 쉬운 범위에서 꼬리를 형성하므로 더욱 손쉽게 관찰할 수 있다.

〈그림 1〉은 척도모수( $\sigma$ )가 1일 때 형태모수( $\xi$ )를 증가시키면서 GPD의 확률밀도함수의 변화를 살펴본 것이다. 기본적으로 형태모수가 양수인 경우 두터운 꼬리를 반영하고 있으며, 그 값이 증가할수록 더욱 두터운 꼬리를 가짐을 오른쪽 확대그림을 통하여 확인할 수 있다. 즉, GPD의 그래프에서 형태모수가 클수록 분포의 꼬리가 더욱 두터워짐을 확인할 수 있다. 〈그림 4〉의 exGPD에서도 형태모수가 클수록 분포의 꼬리가 두터워지는 성질은 동일하게 유지된다. 특히 exGPD에서 우측꼬리의 두터운 수준을 훨씬 유연하게 조절할 수 있다는 장점이 잘 드러나고 있다. GPD는 꼬리부분의 두께가 형태모수가 증가하면서 두터워지는 하나 그 차이가 크지 않은 반면, exGPD는 형태모수에 따라 꼬리의 모양이 크게 변화한다. 그러므로 첨도가 매우 큰 자료에 대해서 exGPD를 활용하면 두터운 꼬리를 더욱 효과적으로 반영할 수 있게 된다. 나머지 그림들에서도 위에서 설명한 것과 동일한 특징들을 찾을 수 있다. 또 다른 특성으로는 exGPD의 경우, 척도모수가 증가하면서 그래프의 중심이 오른쪽으로 이동하는 것을 확인할 수 있는데 이는 1차 적률이 척도모수에 대하여 증가함수라는 점을 반영하는 것이다.



### 2.2.4. exGPD에 대한 선행연구

exGPD에 대한 선행연구로는 Lee and Kim(2019)이 존재한다. 이 논문에서는 GPD의 로그변환에 의한 exGPD 이론을 수리적으로 유도하고 기본적인 특징을 살펴보았다. 또한 exGPD의 MLE를 추정하기 위하여 새로운 초기값 알고리즘을 고안하고, 이 알고리즘이 수치적 실험에서 기존 GPD에서 사용되었던 MM(Method of Moments) 방법보다 우수함을 확인하였다. 이를 근거로 exGPD 모형이 GPD의 대안으로 사용될 수 있음을 주장하였다.

## 3. 시뮬레이션

### 3.1. 시뮬레이션의 목적

매우 두터운 꼬리를 갖는 재무비율 자료에 대한 실증분석에 앞서, 이와 유사한 가상의 자료를 반복 생성하여 극단치이론을 적용하고자 한다. 관심의 대상이 되는 재무비율 자료들은 주로 로그정규분포를 따르므로 극단치를 제외한 영역에서는 이를 적용한다. 동시에 극단치 부분에는 GPD 또는 exGPD를 적용한다. 따라서 전체데이터에 대한 모형의 적합도를 살펴보고, 오른쪽 꼬리에 대한 백분위수를 얼마나 정확하게 추정하는지 알아보하고자 한다. 본 모의실험의 목적은 기존의 GPD와 새로 제안한 exGPD 중 어느 분포가 극단치들을 더 효과적으로 설명하는지 탐색하는 것이다.

### 3.2. 시뮬레이션 수행 및 결과

2장에서 소개한 여러 선행연구들에 따르면 PBR은 투자자에게 매우 유용한 재무비율임을 강조하였다. 본 논문에서 소개한 PBR 자료(음수는 제외한 뒤 일괄적으

로 1을 더함)로부터 로그정규분포, 임계값, 그리고 GPD(exGPD) 등 총 5가지의 모수들을 R의 함수 `flognormgpd()`에 의해 MLE 방법으로 먼저 추정하였다. 그 다음에는 추정한 모수들을 간소화하여 가상자료 생성을 위한 모수적 분포를 결정하였다. 즉, 로그정규분포의 두 모수( $\mu=0.9, \sigma=0.3$ (GPD);  $\mu=0.0, \sigma=0.5$ (exGPD)), 임계값( $u=2.8$ (GPD);  $u=1.9$ (exGPD)); 약 90% 백분위수), 그리고 극단치분포에 필요한 두 모수( $\sigma=1.5, \xi=0.6$ (GPD);  $\sigma=0.7, \xi=-0.1$ (exGPD))와 같이 모수들을 결정하였다. 위의 지정된 모수집합에 대하여 1000개의 랜덤넘버를 생성하였다. 즉, 실증분석에서와 동일하게 Mixture model을 사용하여 임계값 이하에서는 로그정규분포를, 임계값 초과하는 영역에서는 GPD를 적용하였다. 다음으로, RMSPE와 MAPE를 계산하였으며 이와 함께 90%, 95%, 그리고 99% 백분위수를 구하였다. 이러한 절차를 5000번 반복하여 얻어지는 출력결과들은 5000개의 원소를 갖는 각 벡터들에 저장하였다. 여기에서 활용하는 적합도 통계량은 척도에 의존하지 않는 RMSPE(Root Mean Square Percentage Error)와 MAPE(Mean Absolute Percentage Error)에 기반하여 계산하였는데, 이들은 다음과 같이 정의된다.

$$RMSPE(\%) = \sqrt{\frac{1}{n} \sum \left[ \frac{Actual - Forecast}{Actual} \right]^2} \times 100 \quad (10)$$

$$MAPE(\%) = \frac{1}{n} \sum \frac{|Actual - Forecast|}{|Actual|} \times 100 \quad (11)$$

먼저 GPD를 이용한 전체 모형의 적합도는 RMSPE(19.80%), MAPE(14.81%)로 나타났다. 단, RMSPE와 MAPE는 반복된 과정에 의해 계산된 5000가지 결과들의 평균값이다. 90% 백분위수를 넘어서는 꼬리부분에 대한 모형의 적합도만 보면 RMSPE(0.30%), MAPE(0.24%)로 계산된다.

이제 exGPD를 적용하기 위한 가상의 데이터를 생성하자. 앞에서 지정한 모수집합의 분포로부터 1000개의 랜덤넘버를 추출하였다. 다음으로, 변환된 랜덤넘버에 대하여 RMSPE, MAPE를 계산함과 함께 90%, 95%, 그리고 99% 백분위수를 구하였다. 이 과정을 5000번 반복함으로써 5000개의 원소를 갖는 각 벡터에 출력결과들을 저장하였다. exGPD를 이용한 전체 모형의 적합도는 RMSPE(11.96%), MAPE(8.32%)로 나타났다. 단, RMSPE와 MAPE는 각각 5000개의 평균값이다. 90% 백분위수를 넘어서는 꼬리부분에 대한 모형의 적합도만 계산하면 RMSPE(0.28%), MAPE(0.22%)와 같다.

위의 두 분포에 대한 시뮬레이션으로부터 도출된 결과는 다음과 같이 정리할 수 있다. 우선 전체 데이터에 대한 적합도를 살펴볼 경우에는 exGPD가 GPD보다 우수할 뿐만 아니라, 전체가 아닌 임계값(90% 백분위수)을 넘어서는 극단치 자료만을 대상으로 계산한 적합도의 비교에서도 exGPD의 성능이 조금 더 앞선 것이 확인되었다. 즉, 모형의 적합도 측면에서는 exGPD 모형이 GPD 모형보다 더 자료의 형태를 정확하게 확률분포로 나타냄을 확인할 수 있다.

## 4. 실증분석

### 4.1. 자료의 구성 및 분석방법

본 연구는 KISVALUE에서 2015년 말을 기준으로 한국 유가증권시장에 상장된 정보기술(Information Technology)군에 속하는 493개 기업의 PER 및 PBR을 분석대상으로 하였다. 산업분류는 주식시장에서 실무적으로 가장 많이 활용되는 GICS(Global Industry Classification Standard)에 기반하였다. 이는 1999년

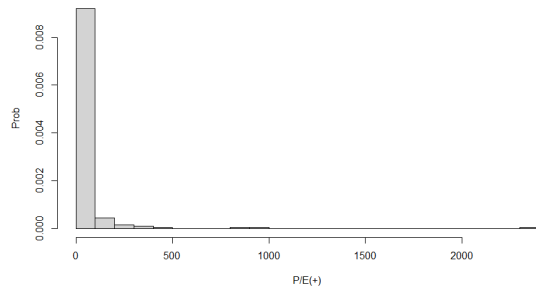
S&P와 MSCI가 개발한 업종 분류법으로 총 4단계의 계층으로 구성된다. 이는 조금씩 변경될 수 있는데 10 Sectors(대업종), 24 Industry Groups(중업종), 67 Industries(소업종)와 156 Sub-industries(세부업종)으로 분류하고 있다(2015년 기준). 재무비율 수치는 동종 또는 유사한 기업집단 내에서 비교해야 하는데 세분화된 분류를 따를수록 비교가능성은 더욱 높아진다. 정보기술군은 대분류에 속하는데 이보다 하위분류를 사용할 경우 표본기업의 수가 매우 적어 극단치이론을 적용하기 어려워지므로 대분류 기준을 적용하였다. 단, 정보기술군 이외의 다른 대분류 산업군들은 기업 수가 작은 경우가 많아 극단치 자료에 대한 분포 적합을 수행하기 곤란하여 분석 대상에서 제외하였다.

2015년 말 정보기술군 기업들의 PER 값에 대한 기초통계량은 평균(16.38), 중위수(10.75), 최소값(-823.8), 최대값(2385), 왜도(8.11) 및 첨도(136.28)와 같이 나타난다. 정보기술군 내 기업은 PER 값이 약 16이면 평균 수준이라고 볼 수 있다. 첨도는 백 단위에 달하므로 꼬리가 매우 두터운 자료에 해당한다.

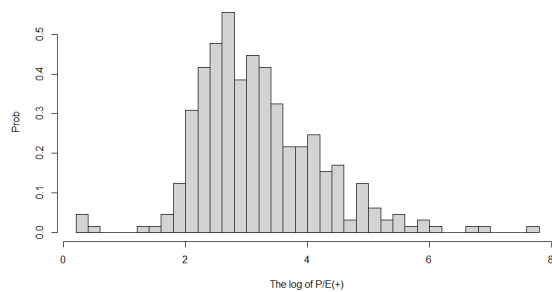
PER이 0보다 작은 기업은 169곳에 이른다. 기본적으로 PER은 작을수록 좋지만 음수인 경우에는 문제가 된다. 기업이 이익을 내지 못하여 당기결산에서 적자를 나타낸 것이기 때문이다. PER이 음수라는 것은 투자 측면에서 적신호이다. 그러나 음수를 나타내는 모든 기업들이 실제로 경영성과가 매우 악화되었다고 단정 지을 수 없다. 왜냐하면 주당순이익은 영업이익이 아니라 당기순이익이기 때문이다. 재무제표의 특성상 영업외손실 때문에 당기순이익이 일시적으로 음수가 될 수 있다. 예를 들어 대규모 설비투자를 위한 차입금으로 금융 부담이 급격히 증가했거나, 예기치 못한 환차손의 발생, 또는 지분 보유기업의 사고나 영업손실 등의 우발적 사고 때문일 수 있다. 이러한 경우에는 일시적으로 PER이 음수 값을 보이는 것일 뿐이다.

그럼에도 불구하고 음수의 PER 수치는 명백히 부정적인 신호임에 틀림없다. PER은 오로지 직전 회계연도의 당기손익에 근거해 계산된 값이므로, 이후의 분기 또는 월별 당기손익의 등을 참고하여 종합적으로 기업가치를 판단해야 한다. 또한 PER이 음수인 기업은 N/A로 공시되는 경우가 많으므로 이후의 실증분석에서도 제외하였다. PER이 음수인 기업들은 극단치이론을 통하여 통계적으로 분석하기보다는 재무제표에 대한 전반적인 검토가 필요하다.

PER가 양수인 기업 324곳은 실증분석 대상이 되는데, 추후 로그변환을 수행하는 과정에서 양수의 PER 값을 유지하기 위하여 일괄적으로 1을 더하여 살펴보고자 한다. 기본적으로 정보기술군 기업들의 해당 재무비율 값의 분포는 아래 <그림7>와 같다. 기초통계량을 살펴보면 평균(51.13), 중위수(20.89), 최소값(1.30), 최대값(2386), 왜도(11.46) 및 첨도(157.40)로 나타난다. 특히 첨도가 매우 크기 때문에 이 경우 극단치이론을 적용하는 것이 바람직하다. 아래의 <그림7>에서 주가에 비해 당기손익이 적은 기업은 PER이 큰 값을 가지므로 주로 분포의 오른쪽 꼬리부분에 위치하게 된다. <그림7>에 해당하는 자료에 대하여 관심 분포 중에서 exGPD를 적용하기 위해서는 로그변환이 필요하다. <그림8>은 기존 데이터를 로그변환한 결과를 시각화한 것이다. 로그변환 후 자료의 기초통계량은 평균(3.19), 중위수(3.04), 최소값(0.26), 최대값(7.78), 왜도(0.80) 및 첨도(2.01)로 나타났다. 자료의 범위가 약 7에 불과하고 첨도가 매우 낮아졌는데 아래의 <그림8>에서 우측 꼬리부분의 자료들이 조밀해지고 과도한 극단값이 존재하지 않게 되었다.



<그림 7> 정보기술군 기업의 PER 자료  
(2015년 말 기준, 음수는 제외, 일괄 +1)



<그림 8> 정보기술군 기업의 로그변환된 PER 자료  
(2015년 말 기준, 음수는 제외, 일괄 +1)

2015년 말 정보기술군 기업들의 PBR은 아래 <그림 9>와 같이 분포되어 있다. 여기에서는 PER와 달리 PBR이 음수인 기업은 7곳에 불과하므로 이를 제외한 양수인 기업 486곳이 실증분석 대상이 된다. 다만, 이후 로그변환을 하는 과정에서 양수의 PBR 값을 유지하기 위하여 일괄적으로 1을 더한 후 살펴보고자 한다. 따라서, 기초통계량을 살펴보면 평균(4.12), 중위수(2.69), 최소값(1.25), 최대값(84.88), 왜도(8.05) 및 첨도(88.11)로 나타났다. 정보기술군 기업은 PBR 값이 약 3이면 평균 수준임을 알 수 있다. 첨도는 PER보다는 작은 값이나 약 90에 달하여 여전히 크다. 아래의 <그림 9>를 보면 우측 꼬리가 과도하게 긴 것을 확인할 수 있다.

한편 PBR이 0보다 작은 기업들의 재무적 특성을 정리하면 아래의 <표 1>과 같다.

<표 1> PBR이 음수인 기업들의 주요 재무정보

(기말주가, 주당순자산, 주당순이익의 단위는 원)

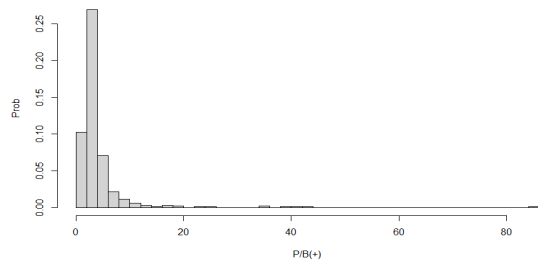
기업	기말주가	주당순자산	주당순이익	PBR	PER
A	19850	-24.7	-418	-803.6	-47.5
B	6450	-50.8	8	-126.9	806.3
C	4645	-92.1	196	-50.4	23.7
D	1870	-72.1	-746	-25.9	-2.5
E	970	-58.6	-295	-16.6	-3.3
F	3900	-375.6	-1154	-10.4	-3.4
G	2800	-1285.2	-19554	-2.2	-0.1

해당 기업들은 구름컴퍼니(A), 와이디온라인(B), 비나텍(C), 아이카이스트랩(D), 옐로페이(E), 유비온(F), 코아로직(G) 등 7곳이다. 위 기업들은 재무적 위험이 수치상으로 높게 나타났다. 단, 기업 C는 PER이 비교적 양호하게 나타났다. 이러한 경우가 존재할 수 있기 때문에 PBR에만 의존하여 기업가치를 평가하는 것에는 한계가 있다. PBR와 직접적으로 관계되는 주당순자산이 모두 음수일 뿐만 아니라 기말주가 대부분 낮고, 주당순이익은 매우 적거나 음수값을 보였다. 위 5가지에 대한 평균값들을 살펴보면 기말주가(5783.57), 주당순자산(-279.87), 주당순이익(-3137.57), PBR(-147.99), PER(110.44)와 같다. PBR와 PER의 구성요소인 기말주가, 주당순자산, 그리고 주당순이익은 모두 PBR이 양수인 어떠한 집단보다도 낮았다.

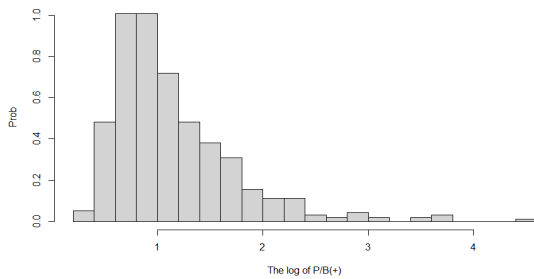
직관적으로 보더라도 PBR이 음수인 기업은 위험군에 속할 것으로 기대되며, 뿐만 아니라 PBR이 지나치게 큰 양수값을 갖는 기업의 경우에도 재무적으로 심각한 문제가 있을 수 있다. PBR이 상위 1%인 기업 5곳에 대해 살펴보면 평균적으로 기말주가(6423.8), 주당순자산(145.13), 주당순이익(-628.4), PBR(47.5), PER(-23.51)으로 나타났다. 기말주가는 비교적 낮고 주당순자산은 매우 적으며 주당순이익이 음수를 나타내므로 경영성과가 안 좋을 가능성이 높다. 또한 PBR

가 매우 크고 PER는 음수이므로 수치상으로 해당 기업들의 재무적 상황이 나쁘지만, PBR이 음수인 기업보다는 조금이나마 양호한 수치를 보여주고 있다. 즉, 기말주가, 주당순자산, 그리고 주당순이익 모두 PBR이 음수인 기업보다는 높게 나타났다.

<그림9>에 해당하는 자료에 대하여 관심 분포 중에서 exGPD를 적용하기 위해서는 자료의 로그변환을 필요로 한다. <그림10>은 기존 데이터를 로그변환한 결과를 시각화한 것이다. 로그변환 후 자료의 기초 통계량은 평균(1.15), 중위수(0.99), 최소값(0.22), 최대값(4.44), 왜도(1.82) 및 첨도(4.79)로 나타났다. 자료의 범위는 약 4에 불과하며 첨도가 크게 감소하였는데 아래의 <그림10>에서 우측 꼬리부분의 자료들이 조밀해지고 과도한 극단값이 존재하지 않게 되었다.



<그림 9> 정보기술군 기업의 PBR 자료  
(2015년 말 기준, 음수는 제외, 일괄 +1)



〈그림 10〉 정보기술권 기업의 로그변환된 PBR 자료  
(2015년 말 기준, 음수는 제외, 일괄 +1)

이제 분석대상이 되는 PBR이 양수인 기업의 각종 재무적 정보들을 비교해 보면 다음과 같다. 먼저 PBR이 0보다 큰 기업들의 자료를 오름차순으로 정렬한 후 20% 백분위수마다 다섯 개의 집단으로 나누어 보았다. 즉,  $PBR \in [1.25, 1.99)$  (0~20%, 98곳),  $PBR \in [1.99, 2.42)$  (20~40%, 97곳),  $PBR \in [2.42, 3.10)$  (40~60%, 96곳),  $PBR \in [3.10, 4.55)$  (60~80%, 98곳),  $PBR \in [4.55, 84.88)$  (80~100%, 97곳) 등의 다섯 가지 그룹으로 표현할 수 있다. 양측 극단(0~20%[그룹1], 80~100%[그룹2]) 그룹의 경우 각각 주가가 과소, 과대평가되었다고 볼 수 있으므로 투자자들이 PBR에 근거하여 투자 포트폴리오를 구성할 경우, 전자에 속한 기업은 포함시키는 반면 후자에 해당되는 기업은 배제하려는 성향이 있다. 따라서, 위 두 집단의 다른 재무적 특성을 정리해 보면 PBR이 낮은 기업들([그룹1])의 경우 평균적으로 기말주가(7512.67), 주당순자산(9814.89), 주당순이익(530.63), PBR(0.77, 1을 더하지 않은 원자료 기준), PER(-1.56, 1을 더하지 않은 원자료 기준)를 갖는 것으로 조사되었다. 반면에 PBR이 높은 기업들([그룹2])의 경우 평균적으로 기말주가(23151.14), 주당순자산(3706.29), 주당순이익(372.76), PBR(9.07, 1을 더하지 않은 원자료 기준), PER(23.28, 1을 더하지 않은 원자료 기준)으로 나타났다. 해석하자면 [그룹1]은 PBR이 양수인 기업 전체에 비해 주당순자산이 평균

적으로 1788.03원 높았으나 [그룹2]의 경우에는 평균적으로 4320.574원 낮았다. 특히 [그룹2]는 다섯 개 집단 중에서 평균 주당순자산 및 주당순이익이 가장 낮은 것으로 드러났다. 그룹의 분류기준과는 무관한 PER에서는 위와 같은 추세를 발견할 수 없었다. 다섯 개 집단에 대한 평균 주당순이익을 살펴보면 순서대로 530.63, 394.04, 1254, 478.86, 372.76으로 나타났다. [그룹1]은 두 번째로 높았으나 [그룹2]는 가장 낮았고 두 집단 간의 차이는 157.87원이었다. 따라서 [그룹1]은 PER의 측면에서도 양호하므로 투자할 유인이 크지만, 반대로 [그룹2]는 투자대상에서 제외하는 것이 좋다.

#### 4.2. 실증분석 및 결과

위의 자료에 대해 다음과 같이 극단치이론을 적용한다. 모형은 Extreme Value Mixture Model을 기반으로 하는데, 이는 임계값을 하나의 모수로 취급하여 추정하고 결정된 임계값의 왼쪽 영역은 Bulk 분포, 오른쪽에는 Tail 분포를 동시에 결합하는 방법이다. 보통 Bulk 분포는 정규분포, 감마분포, 와이블분포, 또는 로그정규분포를, 그리고 Tail 분포는 GPD를 사용한다.

먼저 PER 자료의 대부분을 설명하는 Bulk 분포를 결정해야 하는데 대상 자료의 형태를 가장 잘 반영하는 모수적 분포를 선택하면 된다. 즉, 분포별로 적합도 통계량(즉, RMSPE, MAPE)을 산출 및 비교하여 값이 가장 작은 분포를 찾는다.

Bulk 분포의 후보로서 위에서 언급한 네 가지의 성능을 각각 살펴본 결과, RMSPE와 MAPE는 정규분포(RMSPE=850.22%, MAPE=186.38%), 감마분포(RMSPE=157.29%, MAPE=56.43%), 와이블분포(RMSPE=196.94%, MAPE=67.32%), 로그정규분포(RMSPE=60.57%, MAPE=25.45%)로 나타났다. 두 통계량은 공통적으로 로그정

규분포일 때 적합도가 가장 우수하므로 이를 Bulk 분포로 선택하였다.

또한 Tail 분포는 두터운 꼬리를 갖는 자료에 대해 효과적인 GPD 또는 exGPD를 적용하여 Mixture Model의 전반적인 성능을 살펴보면 된다. Mixture Model에서의 임계값은 R의 함수 `flognormgpd()`에 의해 MLE 방법으로 추정하였는데, 원자료의 경우 87.65% 백분위수, 로그변환된 자료에서는 88.89% 백분위수가 임계값으로 추정되었다. 이를 근거로 하여 90% 백분위수를 실증분석 시의 임계값으로 선정하였다.

이제 로그정규분포(Bulk)에 대하여 GPD(Tail)를 혼합한 전체분포를 [분포1], 동일한 Bulk에 대하여 exGPD(Tail)를 혼합한 전체분포를 [분포2]라 하자. [분포1]의 적합도 통계량은 RMSPE=27.08%, MAPE=13.08%이고 [분포2]는 RMSPE=46.25%, MAPE=23.65%와 같이 나타났다. 즉, PER 자료는 GPD(즉, [분포1])의 경우에서 분포의 전반적인 적합도가 좋았다. 그러나 이를 바탕으로 GPD가 exGPD보다 주어진 자료를 더욱 잘 설명한다고 단언하기는 어렵다. 왜냐하면 [분포1]과 [분포2] 간에는 Tail 분포의 차이뿐만 아니라, 로그변환에 의해 자료의 형태가 변형되었다는 점 역시 전체 적합도의 차이에 영향을 미칠 수 있기 때문이다.

본 연구에서 모수적 분포를 결정하려는 궁극적인 목적은 우측 꼬리부분의 백분위수를 추정함으로써 관심기업의 상대적 위험수준을 가늠하고 궁극적으로는 투자자의 의사결정에 도움을 주기 위함이다. 따라서 [분포1]과 [분포2] 각각에 대해 90%, 95%, 및 99% 백분위수를 추정하여 이를 실제 자료의 백분위수와 비교하고자 한다.

원자료의 실제 백분위수는 89.86(90%), 144.88(95%), 및 434.09(99%)이다. [분포1]에 대하여 추정된 백분위수는 68.97(90%), 121.79(95%), 420.78(99%)이다. 다음으로, [분포2]에 대하여 추정된 백분위수는 84.65(90%),

138.87(95%), 461.89(99%)이다. 각각의 백분위수마다 (추정된 백분위수-실제 백분위수)의 절대값을 계산하여 성능 척도로 활용해 보면 다음과 같이 계산된다. [분포1]에서는 20.89(90%), 23.09(95%), 13.31(99%)이고, [분포2]에서는 5.21(90%), 6.01(95%), 27.8(99%)으로 나타난다. 위 값은 작을수록 좋은데, 90%와 95% 백분위수에서는 exGPD가, 99% 백분위수에서는 GPD가 우수한 성능을 보여준다.

이제 두 번째 자료인 PBR에 대해서도 분석해 보면 다음과 같이 나타난다. 우선 Bulk 분포를 결정해야 한다. Bulk 분포의 후보로는 이전의 네 가지와 동일하고 RMSPE와 MAPE를 계산해 본 결과 정규분포(RMSPE=871.45%, MAPE=169.55%), 감마분포(RMSPE=368.01%, MAPE=81.91%), 와이블분포(RMSPE=603.89%, MAPE=124.90%), 및 로그정규분포(RMSPE=192.90%, MAPE=48.69%)로 나타났다. 결과적으로 로그정규분포의 성능이 가장 우수한 것으로 나타나 이를 Bulk 분포로 결정하였다.

또한 Tail 분포는 두터운 꼬리를 갖는 자료에 적절한 GPD(exGPD)를 선택해 Mixture Model의 전반적인 성능을 알아보려고 한다. Mixture Model에서의 임계값은 원자료의 경우 52.47% 백분위수, 로그변환된 자료에서는 90.53% 백분위수로 추정되었다. 이를 근거로 각각 50%, 90% 백분위수를 임계값으로 결정하였다.

로그정규분포(Bulk)에 대하여 GPD(Tail)를 혼합한 전체분포를 [분포3], 동일한 Bulk에 대하여 exGPD(Tail)를 혼합한 전체분포를 [분포4]라 하자. [분포3]의 적합도 통계량은 RMSPE=26.63%, MAPE=12.97%이고 [분포4]는 RMSPE=11.30%, MAPE=6.08%로 나타났다. 즉, PBR는 exGPD(즉, [분포4])를 이용하면 자료 전체를 보다 잘 설명할 수 있다. PER의 경우와 마찬가지로 [분포3]과 [분포4] 간에는 Tail 분포의 차이뿐만 아니라 로그변환에 의한 분포의 형태변화 또한 적합도의

차이에 영향을 줄 수 있으므로, 각 분포별로 90%, 95%, 및 99% 백분위수를 추정해 실제 백분위수와 비교해 보겠다.

원자료의 실제 백분위수는 6.38(90%), 9.68(95%), 및 34.39(99%)이다. [분포3]에서 추정된 백분위수는 2.81(90%), 4.11(95%), 10.24(99%)이다. [분포4]에서 추정된 백분위수는 6.57(90%), 10.43(95%), 27.59(99%)이다. 각각의 백분위수마다 (추정된 백분위수-실제 백분위수)의 절대값을 계산하여 성능 척도로 활용해 보면 다음과 같이 계산된다. [분포3]에서는 3.57(90%), 5.57(95%), 24.15(99%)이고, [분포4]에서는 0.19(90%), 0.75(95%), 6.80(99%)으로 나타난다. 위 값은 작을수록 좋는데, 모든 백분위수에서 exGPD가 GPD보다 일관되게 우수한 성능을 보여준다.

위의 실증분석 결과들은 아래의 <표 2>와 같이 정리할 수 있다. 즉, PER에 대한 분포의 적합도 평가에서는 GPD가 효과적으로 나타났고, 그 이외의 경우들에서는 대체로 exGPD가 보다 우수한 결과를 보여주었다. 우측 꼬리에 해당하는 영역, 즉 90% 백분위수 이상의 극단적인 값들에 관심이 있는 경우에는 PER에서 99% 백분위수를 추정하는 경우만 제외하면 exGPD 모형을 채택하는 것이 종합적으로 바람직하다고 결론지을 수 있다.

특히 재무비율 중에서 PBR을 대상으로 분석할 경우에는, 언제나 exGPD가 일관되게 뛰어난 성능을 보여줌으로써 재무비율의 통계적 분석 시 exGPD가 중요한 분포로써 활용될 수 있음을 시사한다.

<표 2> 실증분석 결과 성능이 우수한 분포

재무비율		PER	PBR
평가기준	RMSPE	GPD	exGPD
	MAPE	GPD	exGPD
백분위수	90%	exGPD	exGPD
	95%	exGPD	exGPD
	99%	GPD	exGPD

추가적으로 아래의 <표 3>와 <표 4>은 최근 6개년 자료에 대하여 각각 PER 및 PBR를 위의 <표 2>에서 선택된 분포로부터 추정한 값을 정리한 것이다. 꼬리 영역에 대한 분포로서는 PER의 99% 백분위수를 추정하는 경우만 제외하고 모두 exGPD 모형을 일괄 적용하였다.

기존의 실증분석에 활용되지 않았던 2016년부터 2020년까지의 자료를 추가로 살펴보고 비교함으로써 안정적인 백분위수 추정값을 얻는다면, 투자 의사결

<표 3> 최종 선택된 모형에 의해 추정된 PER의 백분위수

(2015~2020년, 음수는 제외)

		PER		
모형		exGPD		GPD
백분위수		90%	95%	99%
연도	백분위수			
2020		98.03	160.32	656.43
2019		111.28	169.16	529.16
2018		75.97	126.60	246.26
2017		110.20	197.50	1017.19
2016		81.67	141.25	689.59
2015		83.65	137.87	419.78
평균		93.47	155.45	593.07
표준편차		15.23	25.78	263.67

<표 4> 최종 선택된 모형에 의해 추정된 PBR의 백분위수

(2015~2020년, 음수는 제외)

		PBR		
모형		exGPD		
연도	백분위수	90%	95%	99%
2020		6.22	9.65	28.36
2019		4.58	5.86	12.53
2018		3.91	5.78	12.96
2017		3.46	4.99	11.29
2016		4.81	6.46	16.26
2015		5.57	9.43	26.59
	평균	4.76	7.03	18.00
	표준편차	1.02	2.00	7.54

정 및 위험관리에 활용 가능한 가이드라인을 각 재무 비율의 특정 임계값을 기준으로 제한할 수 있을 것으로 기대된다. 우선 <표 3>에서 최근 연도별로 추정된 PER 값들을 살펴보면, 평균적으로 93.47을 넘어서면 정보기술군 기업들 중에서 90% 백분위수 이상에 속한다고 볼 수 있고, 155.45는 95% 백분위수, 그리고 593.07은 99% 백분위수에 대응되는 PER 값으로 볼 수 있다. 마찬가지로 <표 4>에서 최근 연도별로 추정된 PBR 값들을 살펴보면, 평균적으로 4.76을 넘어서면 해당 산업군 내 기업들 중에서 90% 백분위수 이상에 속하는 것으로 판단할 수 있고, 7.03은 95% 백분위수, 그리고 18.00은 99% 백분위수에 위치하는 PBR 값으로 볼 수 있다. PBR의 경우 연도별로 추정된 재무비율 값의 변동이 낮은 편으로 PER에 비해 보다 안정적인 가이드라인으로 활용할 수 있다는 점에서 의의가 있다. 따라서, 이후 연도에도 표본기업의 수가 일정하고 여러 경제적 상황이 유사하다면 해당 분포에 근거하여 특정 산업군 내에서 재무비율 값이 가지는 상대적인 수준을 예측하는데 활용할 수 있다.

다만, PER와 PBR의 분석결과를 종합하여 투자 의사결정을 내릴 때 두 재무비율이 일관된 가이드라인을 제시하지 못할 수도 있다는 점을 유의해야 한다. 이전의

실증분석에서 PBR은 음수임에도 불구하고 PER 측면에서는 양호한 값을 나타내는 기업들을 일부 찾을 수 있다. 이처럼 PER과 PBR 간에 기업가치가 다르게 나타나는 사례들을 구체적으로 정리하면 다음과 같다. PBR 값이 0과 1 사이에 속하는 기업 100곳에 대하여, PER이 부정적인 신호를 나타내는 경우들이 존재한다. PER 값이 양수인 기업들의 평균값(50.13)보다도 높게 나타나는 기업은 총 6곳이다. 위 기업들의 PER은 평균(140.41), 최소값(51.94), 및 최대값(450)을 가진다. 또한 PER가 음수인 기업들은 총 35곳으로, 평균(-48.60), 최소값(-432.86), 및 최대값(-1)으로 나타난다. 이를 종합해보면 전체 100곳 중에서 PER이 평균을 넘는 기업은 6%, PER이 음수인 기업은 35%를 차지한다. 다시 말해서, PBR 상으로 우수한 그룹에 속하는 기업들 중 41%는 PER 측면에서 살펴볼 때는 재무적 상황이 좋지 않다고 판단된다는 점이다. 이처럼 두 재무비율이 일관되지 못한 경우가 존재하기 때문에 단 하나의 재무비율에만 의존하여 기업가치를 평가하고 투자 의사결정에 활용하는 것은 바람직하지 못하다.

또한 위와 정반대인 사례를 살펴보면 다음과 같다. PER 측면에서는 우수하지만, PBR 지표가 위험군인 대표적인 기업으로는 (주)넥슬론을 꼽을 수 있다. 이 기



업의 PER은 0.49에 불과하지만, PBR은 9.50으로 나타났다. 이러한 상황에서 PER에만 의존하여 가치판단을 수행하기에는 위험성과 한계점이 존재하는 것이다.

해당 재무비율이 양수인 기업에 대하여 PER와 PBR 간 상관관계수 값은 0에 가까운 것으로 나타났다. 반면, 자료에 대한 여러 상관관계 중에서 주가와 순이익 사이의 상관관계수는 0.88이고 주가와 순자산 사이의 상관관계수는 0.90로 매우 큰 값을 보였다. 즉 기업의 순자산과 순이익은 주가를 결정하는 핵심적인 재무정보이며 PER 및 PBR는 투자 측면에서 유용성이 높다는 것이다. 결론적으로 PER 및 PBR는 어느 하나가 나머지의 정보를 대체하거나 대표할 수 없으며 상호보완의 관계를 가지므로, 두 재무비율을 종합적으로 분석해야 할 것이다.

## 5. 결론

### 5.1. 연구결과의 요약 및 시사점

이전 4장에서 Extreme Value Mixture Model을 활용하여 모수적 분포를 결정하고 해당 재무비율의 우측 꼬리 영역에 대하여 GPD 또는 exGPD를 적합한 결과는 아래와 같이 정리할 수 있다. 우선 PER 자료의 RMSPE, MAPE, 99% 백분위수에 대해서는 GPD가 exGPD보다 더 나은 성능을 보여주었고, 또한 PER 자료의 90%, 95% 백분위수와 PBR 자료의 경우에는 모든 평가기준에 대하여 exGPD가 GPD보다 일관되게 우수한 결과를 나타내었다. 따라서 위 실증분석 결과들을 종합해 보면, 재무비율을 분석하는데 있어 극단치이론에서 전통적으로 활용되어 온 GPD뿐만 아니라 exGPD 또한 모수적 분포를 결정하고 보다 정확하게 재무비율을 분석하는데 유용하게 활용됨을 확인하였다.

본 연구의 시사점은 다음과 같이 요약할 수 있다. 첫째, 최근 통계학의 극단치이론 연구분야에서 새롭게 제안된 확률분포인 exGPD의 유용성을 실제 재무비율 자료 중에서 침도가 매우 크고 중요한 정보를 포함하는 PER과 PBR에 기초하여 실증적으로 살펴보았다. 둘째, 위의 두 재무비율 이외에도 기업의 신용분석, 부실예측, 위험관리 등을 수행하기 위해 종종 활용되는 유동성비율이나 부채비율 등의 다양한 재무비율들을 면밀히 분석하는 경우 exGPD를 확률분포에 대한 가정으로 사용하는 것이 보다 효과적인 대안이 될 수 있다.

### 5.2. 연구의 한계점 및 향후 과제

본 연구의 한계점은 크게 세 가지로 정리할 수 있다. 우선, 연도별로 백분위수를 추정할 값은 특히 PER에서 일관되지 못하기 때문에, 이전 5.1절에서 보여주고 있는 결과들은 안정적이면서 객관적인 위험지표로 활용되기에는 불완전할 수 있다. 다음으로, 동종 산업 내 기업들의 수가 충분히 커야 극단치이론을 적용하기에 유리하지만, 한국의 경우 그 수가 해외에 비하여 많이 부족하여 분석하고 활용하는데 한계가 존재한다. 마지막으로, 기업의 가치를 정확하게 파악하려면 관심대상이 되는 산업군이 공통적으로 영향을 받는 외부적인 상황 등을 다각적으로 고려해야 한다. 그러나 본 논문에서는 PER 및 PBR의 수치에만 근거하여 기업의 재무적 위험수준을 판단하였다. 이러한 접근법이 간편하고 직관적이기는 하나 중대한 의사결정을 내릴 경우에는 바람직하지 않을 수 있다. 따라서, PER과 PBR 이외에 투자자들이 활용할 수 있는 다양한 재무정보들을 살펴보는 것이 중요하고, 또한 해당 산업에 대한 전반적인 검토와 거시경제적 상황에 대해 이 해 역시 동반되어야 할 것이다.

exGPD 이론을 수리적으로 살펴본 연구는 존재하나, 이를 활용한 실증분석은 거의 존재하지 않는다. 특히 기업에 대한 신용평가, 부실예측, 또는 위험관리를 수행하는 경우에는, 모형의 분포에 대한 가정이 예측 성능을 결정하는데 매우 중요한 역할을 한다. 이러한 연구에서 사용되는 금융자료가 가지는 특성을 더욱 잘 반영하고, 분포의 꼬리 영역의 자료를 보다 효과적으로 관찰하기 위해 exGPD는 효과적인 대안으로 활용될 수 있다. 따라서, exGPD를 이용하여 금융자료를 보다 정확하게 설명하고 다양한 통계적 모형에 이를 응용함으로써 모형의 예측 능력이 향상됨을 구체적으로 확인하는 후속 연구가 이루어지길 희망한다.

## 〈참고문헌〉

### [국내 문헌]

1. 강민정, 김지연, 송종우, 송성주 (2013). Peaks Over Threshold를 이용한 Value at Risk: 모수추정 방법론의 비교. **응용통계연구**, 26(3), 483-494.
2. 김병수, 허용석, 한인구, 이희석 (2010). 지식 경영 활동의 혁신 역량으로의 연계: IT 서비스 산업 중심으로. **지식경영연구**, 11(1), 97-113.
3. 오세경 (2005). 극단치이론(Extreme Value Theory)과 Value-at-Risk: GPD 모형을 중심으로. **금융연구**, 19(1), 69-112.
4. 이영한, 박주영 (2016). 국제회계기준(IFRS) 도입이 주가이익 배수(PER)와 이익의 질과의 관련성에 미치는 영향. **경영학연구**, 45(3), 1003-1039.

### [국외 문헌]

5. Alford, A. W. (1992). The Effect of the Set of Comparable Firms on the Accuracy of the Price-Earnings Valuation Method. **Journal of Accounting Research**, 30(1), 94-108.
6. Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. **The Journal of Finance**, 23(4), 589-609.
7. Balkema, A. A., & De Haan, L. (1974). Residual Life Time at Great Age. **The Annals of Probability**, 2(5), 792-804.
8. Beatty, R. P., Riffe, S. M., & Thompson, R. (1999). The Method of Comparables and Tax Court Valuations of Private Firms: An Empirical Investigation. **Accounting Horizons**, 13(3), 177-199.
9. Castillo, E., & Hadi, A. S. (1997). Fitting the Generalized Pareto Distribution to Data. **Journal of the American Statistical Association**, 92(440), 1609-1620.
10. Cheng, C. S., & McNamara, R. (2000). The Valuation Accuracy of the Price-Earnings and Price-Book Benchmark Valuation Methods. **Review of Quantitative Finance and Accounting**, 15, 349-370.
11. Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*(1st Ed.), Springer Series in Statistics, London, England.
12. De Haan, L., & Ferreira, A. (2006). *Extreme Value Theory: An Introduction*(1st Ed.), Springer Science & Business Media, Berlin, Germany.
13. De Zea Bermudez, P., & Kotz, S. (2010). Parameter Estimation of the Generalized Pareto Distribution Part I. **Journal of Statistical Planning and Inference**, 140(6), 1353-1373.
14. Dekkers, A. L., Einmahl, J. H., & De Haan, L. (1989). A Moment Estimator for the Index of an Extreme-Value Distribution. **The Annals of Statistics**, 1833-1855.
15. Grimshaw, S. D. (1993). Computing Maximum Likelihood Estimates for the Generalized Pareto Distribution. **Technometrics**, 35(2), 185-191.
16. Hill, B. M. (1975). A Simple General Approach to Inference about the Tail of a Distribution. **The Annals of Statistics**, 3(5), 1163-1174.
17. Kaplan, S., & Ruback, S. (1995). The Valuation of Cash Flow Forecasts: An Empirical Analysis. **The Journal of Finance**, 50(4), 1059-1093.
18. Kellezi, E., & Gilli, M. (2000). Extreme Value Theory for Tail-Related Risk Measures. **The Computational Finance 2000 Conference**, London Business School.
19. Leclair, M. S. (1990). Valuing the Closely-held Corporation: The Validity and Performance of Established Valuation Procedures. **Accounting Horizon**, 4(3), 31-42.
20. Lee, S. Y., & Kim, J. H. T. (2019). Exponentiated Generalized Pareto Distribution: Properties and Applications towards Extreme Value Theory. **Communications in Statistics-Theory and Methods**, 48(8), 2014-2038.
21. Magnou, G. (2017). An Application of Extreme Value Theory for Measuring Financial Risk in the Uruguayan Pension Fund. **Compendium**, 4(7), 1-19.
22. Pickands III, J. (1975). Statistical Inference using Extreme Order Statistics. **The Annals of Statistics**, 3(1), 119-131.
23. Wade, W. R. (2009). *An Introduction to Analysis*(4th Ed.), Pearson, New Jersey, U.S.

24. Zhang, J. (2007). Likelihood Moment Estimation for the Generalized Pareto Distribution. *Australian & New Zealand Journal of Statistics*, 49(1), 69-77.

---

● 저 자 소 개 ●

---



**주 지 환 (Jihwan Joo)**

현재 KAIST 경영공학부 박사과정에 재학 중이다. 중앙대학교 경영학부 및 수학과 학사, 고려대학교 통계학과 석사학위를 받았다. 주요 관심분야는 기업공시자료(텍스트)분석, 기업가치평가, 재무제표분석 등이다.

〈 Abstract 〉

# Statistical Analysis of Extreme Values of Financial Ratios

Jihwan Joo\*

Investors mainly use PER and PBR among financial ratios for valuation and investment decision-making. I conduct an analysis of two basic financial ratios from a statistical perspective. Financial ratios contain key accounting numbers which reflect firm fundamentals and are useful for valuation or risk analysis such as enterprise credit evaluation and default prediction. The distribution of financial data tends to be extremely heavy-tailed, and PER and PBR show exceedingly high level of kurtosis and their extreme cases often contain significant information on financial risk. In this respect, Extreme Value Theory is required to fit its right tail more precisely. I introduce not only GPD but exGPD. GPD is conventionally preferred model in Extreme Value Theory and exGPD is log-transformed distribution of GPD. exGPD has recently proposed as an alternative of GPD(Lee and Kim, 2019). First, I conduct a simulation for comparing performances of the two distributions using the goodness of fit measures and the estimation of 90-99% percentiles. I also conduct an empirical analysis of Information Technology firms in Korea. Finally, exGPD shows better performance especially for PBR, suggesting that exGPD could be an alternative for GPD for the analysis of financial ratios.

Key Words: Knowledge Management, Price-to-Earnings Ratio(PER), Price-to-Book Ratio(PBR), Extreme Value Theory(EVT)

---

\* College of Business, KAIST