

2018년 봄호

연구방법논총

Journal of Research Methodology

【연구논문】

- 한국 언론의 보도행태의 다양성 : 최순실 사건 보도를 중심으로
이성우 · 이호준 · 이병근 / 1
- 효과적인 남북한 재난대응 협력을 위한 법제도적 문제 고찰
김진원 · 정한범 / 25
- 스마트건강증진서비스 효과성 분석 연구 : 서울시 사례
오수경 · 신창환 / 57

【연구경향】

- Varimax Rotation and Thereafter : Tutorial on PCA Using Linear Algebra, Visualization, and Python Programming for R and Q analysis
Byung Lee / 79



차 례

【연구논문】

- 한국 언론의 보도행태의 다양성: 최순실 사건 보도를 중심으로
.....이성우·이호준·이병근 / 1
- 효과적인 남북한 재난대응 협력을 위한 법제도적 문제 고찰
..... 김진원·정한범 / 25
- 스마트건강증진서비스 효과성 분석 연구: 서울시 사례
..... 오수경·신창환 / 57

【연구경향】

- Varimax Rotation and Thereafter: Tutorial on PCA Using Linear Algebra,
Visualization, and Python Programming for R and Q analysis
.....Byung Lee / 79
- 『연구방법논총』 원고작성의 일반적 요령 131
- 『연구방법논총』 저술 윤리강령 138
- 『연구방법논총』 편집 및 심사 규정 140
- 『연구방법논총』 편집위원명단 144

[DOI] <http://dx.doi.org/10.21487/jrm.2018.5.3.1.1>

【연구논문】

한국 언론의 보도행태의 다양성: 최순실 사건 보도를 중심으로

이 성 우*·이 호 준**·이 병 근***

논문요약

본 논문은 텍스트 마이닝 기법과 전문가 설문조사를 활용하여 한국언론의 최순실 사건에 대한 보도행태의 다양성과 그 원인을 분석하였다. 한국정치에서 언론은 객관적 정보전달자라기 보다 정치세력과 함께 연관된 하나의 행위자로 인식되는 경우가 많다. 또한 언론은 이익을 창출해야하는 기업으로서 이익과 연관된 기업의 영향력으로부터 자유롭지 못하다. 분석결과, 분석대상 언론 중 가장 보수적이며 새누리당과 연계된 편향성을 가진 언론은 SBS 로 나타났다. 가장 진보적이며 민주당 편향성을 갖는 언론은 한겨레와 JTBC이다. 그리고 각 사건에 대응하는 보수언론의 패턴은 예상보다 다양하게 나타나고 있으며, 대중의 여론 흐름을 쫓아서 보도편향을 바꾼 언론도 등장하였다. 특히 가장 보수적인 언론으로 인식되어온 조선일보, 중앙일보, 동아일보 각기 다른 보도행태를 보여주었다. 그리고 각 언론사들이 정치적 연관성과 기업과의 관계를 고려하여 보도행태를 정한다는 것을 일정하게 확인하였다. 분석결과에 기반하여, 본 논문은 언론의 보도행태와 정치세력과 기업과의 관계에 대한 추가적인 연구의 필요성을 제기한다.

주제어 : 언론, 정치적 스캔들, 텍스트 마이닝, 보도행태

* 고려대학교 평화와 민주주의연구소 연구교수, 교신저자 woo9505@korea.ac.kr

** 고려대학교 컴퓨터학과 학부생

*** 고려대학교 컴퓨터학과 학부생

I. 서론

본 논문은 최순실 사건에 대한 한국의 주요 언론의 보도 행태와 그 변화를 분석하는 것을 목적으로 한다. 주요 9개 중앙언론사들의 기사를 대상으로 자동화된 텍스트 기법을 이용하여 그 보도의 편향성과 사건의 진행과정에 따라 그 보도편향의 변화를 분석한다. 그리고 언론인을 대상으로 한 설문조사를 실시하여 그 원인에 대한 추정을 진행하였다.

뉴스는 어떤 언론사가 보도하느냐에 따라 달라진다. 언론사가 처한 환경, 조직 구조, 주요 구성원들의 정치적 믿음, 보도관행 등의 많은 요인들의 작용에 의해 만들어진다. 그래서 언론사마다 뉴스가 다른 논조와 관점에서 생산되는 것은 당연한 것이다. 언론사마다 각자의 조건 속에서 해당 언론만의 프레임(frame)이 있다. 이 프레임이 지속되기도 하고, 이슈와 시기에 따라 변화가 되기도 한다. 그 변화는 여론의 흐름을 반영한 것이기도 하고, 자신들의 이익의 변화를 보여주는 것이기도 하다.

실제 언론현실에서 한국의 언론은 저마다 공정한 보도를 이야기하지만, 정파적인 구분이 명확한 언론들이 여론을 주도한다. 그리고 광고매출을 중심으로 한 수익구조로 인하여 언론은 자본의 영향력에서도 자유롭지 못하다.

최순실 사건은 한국 정치에 매우 큰 충격을 준 사건이었다. 최순실 사건으로 인해 한국의 정치적 지형은 보수우위에서 진보우위로 변화되었다. 대다수의 언론들은 국민 여론에 맞추어 최순실 사건과 박근혜 정부에 대한 비판적 입장에서 보도를 하였다. 하지만, 실제 보도를 분석한 결과 한국언론들이 매우 다양한 보도행태를 보여주고 있다는 것을 확인하였다. 본 논문은 텍스트 마이닝(Text Mining) 기법을 활용하여 기사를 수집하고 분석을 진행하였다. 그리고 보조적 지표로서 관련 전문가들에 대한 설문조사를 진행하였다. 이 두가지 분석방법을 활용하여 본 논문은 한국언론의 최순실 사건에 대한 보도행태와 그 원인에 대한 분석을 진행하였다. 그리고 본 논문에서 진행한 텍스트 마이닝 기법은 최근 각광받는 빅데이터 분석방법 중 하나로서 향후 정치학 분야에서 텍스트 마이닝 기법을 통한 연구의 활성화의 필요성과 그 활용 가능성을 보여주고자 한다.

II. 기존연구검토와 가설

1. 기존연구검토

언론이 가지는 태도의 다양성은 그 자체가 민주주의의 상징이기도 하다. 해당 사회가 얼마나 가치와 의견의 다양성을 존중하고 다양한 의견을 얼마나 성숙하게 받아들여야 할지 보여주는 하나의 척도이기도 하다.

언론들이 특정이슈나 사안에 대해서 각기 다른 입장을 가진다. 언론은 저마다 뉴스 가치에 대한 전문적 인식, 조직과 기술적 틀, 수용자 혹은 정기적인 독자에 대한 감각 등이 다르다. 이러한 차이가 합쳐져 신문의 매우 상이한 ‘사회적 성격(Social personalities)’을 만들어낸다(Hall et al 1978, 60).

언론의 다양한 보도 행태의 원인에 대해서는 여러 가지 입장이 존재한다. 우선, 정치권력과 연관성에 대한 입장이 있다. 언론이란 본질적으로 체제의 산물이며 권력 의지의 이익에 봉사하는 도구적 존재일 수밖에 없다는 것이다(Altschull 1984). 이와 연관되어 언론이 자신들의 이익을 위해 정치권력과 연계를 가진다는 설명도 있다. 언론은 정치권력과 직업적인 연계를 가진다. 뉴스의 주 정보원으로 정치권력을 이용한다. 일반적으로, 많은 국가들에서 언론과 정치권력과의 연계는 오래 지속되어왔고, 정치적 격동기를 거치면서도 크게 흔들리지 않는다(Hallin 1994). 윤영철(2000)은 언론의 산물은 권력엘리트와 언론의 상호관계 속에서 나타나며, 언론은 지배권력의 이해를 반영하는 이념적 시각을 전파함으로써 권력관계를 재생산한다고 설명한다. 언론의 위력은 언론 그 자체의 힘보다 그것을 이용하려는 정치권력이 밀접하게 연결이 되면 그 힘은 배가된다. 이 시각은 언론의 소유여부보다 언론이 그 사회의 권력엘리트들과 어떤 관계를 맺느냐에 주목한다.

언론은 사회의 이데올로기를 재생산하는 역할을 하기도 한다(Hall 1979). 언론의 이데올로기적 기능에 대한 논의는 오래전부터 있어왔다. 언론은 타 계급과 집단의 생활을 인지하게 하는 토대를 제공하고, 분열에서 합의로 가는 과정을 더욱 수월하게 한다. 그리고 사회의 다원화되고 복잡한 면들을 하나의 집합적 표현으로 내보내는 역할을 한다. 언론은 이러한 기능을 통해 사회에 기여를 하기도 하고, 자신들의 이익을 스스로 창출하기도 한다. 매체경제력 지배, 국가권력에 대한 영향력, 매체인력에 대한

통제, 이 세 가지를 기반으로 성립되는 언론권력은 수구 이데올로기를 재생산한다(김승수 2003, 44).

다른 한편에서는 경제적 측면을 더 강조하여, 자본주의 체제 하에서 언론의 소유구조에 주목한다. 자본주의 사회에서 언론은 그 공적기능이 있음에도 하나의 산업이다. 언론의 소유구조가 언론의 생산물 즉 뉴스의 방향, 이슈, 주제 등에 영향을 미치는 가장 중요한 요인이라는 것이다. 자본주의 사회에서 언론은 소수의 사람들이 소유하는 사유 기업의 형태를 띠며, 이들 언론조직이 생산해내는 산물은 소유주의 이익과 가치, 그리고 이데올로기를 재생산 할 수 밖에 없다(Curran 1990, 139; McNair 1994, 40). 결국 소유한 자본의 형태, 구조가 뉴스의 편집 방향과 의도 등을 결정하는 구조적 요인이 된다는 입장이다. 언론권력의 원천은 소유독점과 여론지배이다. 언론권력은 시장지배적인 매체를 소유하는 것은 물론 부동산업, 건설업, 유통업, 제조업, 문화사업 등 다양한 곳에 많은 이권을 갖고 있는데 이들은 자신의 이익을 위해 정부 또는 경쟁자를 회유하거나 위협하기도 한다(김승수 2003, 40-41).

결국 언론은 정치권력과 연계성, 자본과의 관계성에 따라 자신들의 생산물의 내용과 형태를 결정한다. 이 둘이 따로 떨어져 있는 것은 아니다. 정치권력과 자본은 자신들의 이익을 위해 서로를 이용한다. 언론을 소유한 기업은 자신들의 이익을 위해 언론을 활용하고, 언론을 통해 정치권력내의 자신들에 우호적인 사람은 적극 지지하고 그렇지 않으면 가혹할 정도로 비판한 후 그래도 마음에 안 들면 쫓아버린다(김승수 2003, 51).

오랜 권위주의 체제하에서 한국의 보수언론은 정당성 없는 독재 권력에 기생하여 살아왔고, 그 결과 보수언론들은 큰 영향력을 형성하였다. 민주화 이후 이 영향력을 활용하여 언론사 스스로의 이익과 사주의 이익을 보장받았다. 90년대 중앙 신문사 중 조선일보, 동아일보, 중앙일보는 신문시장의 70%를 과점하였다. 그 결과, 정치인, 관료, 기업인들 모두 언론의 눈치를 보는 또 다른 권력이 되기도 하였다. 하지만, 시대의 변화와 기술의 발전으로 인하여 2000년대 이후 다양한 인터넷 언론이 등장하게 되었다. 그리고 시민들 스스로 뉴스를 생산하는 시대가 되었다. 이러한 시대의 변화로 인하여 현재 언론은 과거와 같은 영광은 누릴 수 없는 구조가 되었다.

자본주의 경제체제에서 언론기업은 수익을 창출하여 기업으로 생존하여야 한다.

언론기업은 이윤을 지속적으로 확보할 수 있어야 한다. 시장경제에 불확실성과, 인터넷과 뉴미디어의 발달로 인해 독점 언론의 영향력이 감소하였다. 이 환경 속에서 독점 언론은 두 가지 선택을 할 수 있다. 첫째, 소유구조와 소유권을 더욱 집중하여 통제력을 강화하고, 다른 종류의 대중매체를 소유하는 커뮤니케이션 복합기업화(Cross-media Ownership) 하거나, 둘째, 대중매체와 관련이 없는 대기업이 대중매체를 소유하는 일반복합기업화 하는 것이다. 2009년 7월22일 신문법(신문 등의 자유와 기능보장에 관한 법률)과 방송법, 인터넷TV(IPTV)법안들(이하 미디어법)이 국회에서 통과되었고, 이를 근거로 조선일보, 중앙일보, 동아일보는 종합편성(중편) 채널을 확보했다. 한국 보수 신문사들이 종합편성채널 신설과 경영을 위해 노력했던 것은 커뮤니케이션 복합기업화 현상이라고 볼 수 있다. SBS의 경우 일반 복합기업화로 판단할 수 있다. 기업의 규모는 어떤 형식으로도 더 커졌고, 시장의 규모는 한정되어있는 속에서 언론들은 생존을 위해 노력하게 된다. 과거에는 정치권력의 비호아래 큰 노력 없이 이익을 향유하였다면, 다양한 미디어가 발달된 현재에는 보수언론 내에서도 여러 가지 분화현상이 등장하게 된다. 이 측면에서 최순실 사건 보도가 TV조선에서 시작되었다는 것은 시사 하는 바가 크다. 2014년 최순실과 갈등을 빚게 된 고영태로부터 비선실세 최순실 관련 제보를 받은 TV조선은 1년여 동안의 준비과정을 거쳤고, 관련 내용을 특종 보도하였다.

최근의 정치적 사건으로서 세월호와 최순실 사건에 대한 언론의 태도 분화는 매우 주목받을만한 현상이다. 최진봉(2016)의 연구를 보면 한국의 주류 방송들이 정파적 입장에 매여 재난보도의 기본원칙도 지키지 않는다는 것을 보여준다. 연구결과, 우리나라 주류 방송사인 지상파 방송사들은 재난보도가 지켜야 하는 보도의 정확성과 객관성에 미흡함을 보였으며, 언론이 재난보도 과정에서 지양해야 하는 선정성과 오보·양산 등 재난보도의 문제점에서 벗어나지 못하고 있는 것으로 나타났다. 그러나 JTBC의 경우 주류 방송사들의 보도와 달리 재난보도 과정에서 사건 발생의 원인규명과 재발 방지를 위한 대안을 제시하고, 지속적인 사회적 이슈로 논의할 수 있는 장을 마련한 것으로 나타났다.

박근혜대통령 탄핵 보도에 대한 김병진·이상은·양종훈의 연구에서는 진보와 보수 언론을 대표로 MBC와 JTBC의 보도행태를 비교하였다. 탄핵이 인용된 날부터 일주일

간 양 방송사의 저녁 종합뉴스의 아이템 수 변화, 주제의 편향성, 뉴스 프레임 등 3가지 측면을 분석하였다. 연구결과, JTBC와 MBC의 관련 보도는 양적으로 현저한 차이를 보였고, JTBC는 촛불집회 쪽에, MBC는 태극기집회 쪽 관련 주제를 집중적으로 보도했다. 뉴스 프레임 측면에서 MBC는 탄핵인용 관련 보도에서 박근혜 대통령과 이를 옹호하는 태극기집회 참여자들 쪽에서, JTBC는 반대로 반 박근혜 대통령과 탄핵의 정당성을 주장하는 촛불집회 참여자 입장에서 보도한 것으로 나타났다(김병진·이상은·양중훈 2017).

이 같은 보도행태의 분화는 곧 정치적 이슈의 미디어 효과를 극명히 보여주는 것이기도 하다. 언론사 마다 보여주는 정치적 편향이 분명하기 때문에 이를 감안하여 시민들은 자신들이 좋아하는 언론의 보도를 더 보게 된다. 결국 언론들은 자신들의 독자를 위해서라도 더욱 정치적 편향성이 드러나는 보도를 하게 된다. 실제 이소영의 연구에서는 한국의 언론들이 채널별로 상당히 일관된 정파성을 가지고 유권자의 정치적 태도에 영향을 미치고 있는 것을 보여준다. 특히 JTBC, 보수신문, 진보신문은 모든 이슈에 대한 태도에 강한 영향을 미치고 있으며, 그 중에서도 대북정책·안보 이슈와 원자력발전 중단 이슈에 대한 태도 변화에 큰 효과가 있는 것으로 조사되었다. 특정 미디어에 선택적으로 빈번히 노출되었을 경우를 가정한 시뮬레이션 결과, 보수적인 미디어 중에서는 보수신문이, 진보적인 미디어 중에서는 진보신문이 유권자의 태도를 정파적으로 극화시키는 데 가장 큰 역할을 하는 미디어 채널로 나타났으며, 종합편성채널과 JTBC 또한 강한 태도변화 효과를 가지는 것으로 나타났다(이소영 2017).

기존 연구에서 언론의 보도 편향에 미치는 여러 가지 요인과 실제로 언론의 보도편향의 다양성이 생겨나게 되는 조건들을 살펴보았다. 향후 한국의 미디어 정치의 영향과 효과의 다양한 변화가 예상되는 속에서, 본 논문은 최순실 사건 보도에 대한 주요 언론들의 보도행태와 편향을 분석함으로써 이를 실증적으로 보여주고자 한다.

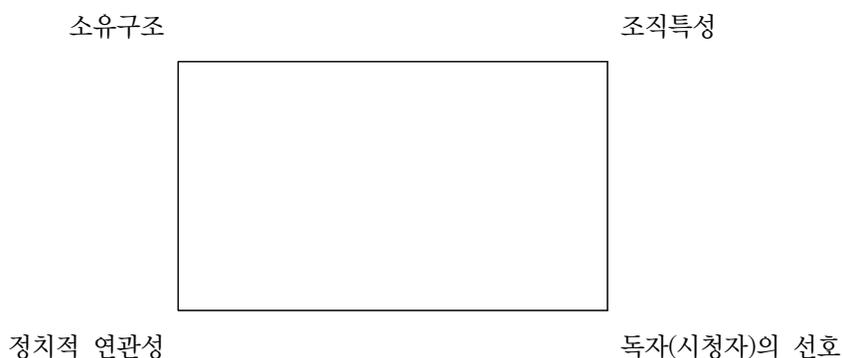
2. 분석의 틀과 가설

본 논문은 언론은 스스로 이익을 설정하고, 이익을 위해 보도행태를 변화시킨다는 기본 가정을 가진다.

기존 연구에서 살펴본바와 같이, 언론의 정파적 보도행태에 영향을 미치는 중요

요인으로는 정치적 연관성, 소유구조에 따른 경제적 이익 추구가 있다는 것을 확인할 수 있다. 그리고 추가적으로 영향을 미치는 요인으로는 오랜 기간 동안의 보도행태로 인하여 형성된 개별 언론별 조직 특성을 꼽을 수 있고, 기존 언론의 소비자들의 선호를 영향요인으로 인식할 수 있다. 이 요인들은 서로 연계되어있고, 언론사마다 그 중요도는 다를 수 있다. 다만, 언론사별로 소유구조에 따라 그 이익의 설정자가 다를 것이다.

[그림 1] 언론의 보도방향 설정의 요인



일반적인 경우에는 이 요인들의 작용에 의한 보도행태의 변화는 잘 느낄 수 없다. 특이한 경우가 아니고서는 자율적으로 보도행태가 정해질 것이다. 하지만, 큰 정치적 또는 경제적 변동이 발생한다면 보도행태는 변화된다. 이 네가지 요인이 변화되지 않는다고 하더라도 이 중 하나의 요인만으로도 보도행태가 변동될 수 있다. 이 요인들 중 하나의 요인이 경직적으로 변동되지 않는다고 하더라도, 다른 요인의 변동으로 인하여 보도행태가 변화될 수 있는 것이다.

민주주의 체제하에서 언론은 자신들의 이익을 위해서 그 보도행태를 변경할 수 있다. 시대 변동에 따라 경제적 이익을 우선할 수도 있고, 정치적 연관성이 큰 정치세력을 보호하는데 중점을 둘 수도 있다. 때로는 광고를 많이 주는 기업을 위해 자신과의 정치적 연관성이 깊은 정치세력과의 갈등을 벌일 수 있다. 그 갈등의 수준이 고조되지 않는 선에서 갈등을 봉합하기도 한다.

언론의 보도행태는 고정적 위치가 어느 정도 정해져 있지만, 변동되지 않는 것은 아니다. 궁극적으로 시대의 변화에 따라 자신들의 이익을 위해 변화를 스스로 추구할 수 있다.

그렇다면, 언론은 자신의 보도행태를 결정하는데 깊이 연관되는 정치적 연관성과 경제적 이익을 위한 파트너들이 위기상황에 빠졌을 때 어떤 보도행태를 보이는가?

반대로, 자신들의 정파성과 반대의 정치세력과 경제세력이 위기상황에 처했을 때 어떤 보도행태를 보일 것인가?

첫째, 정치적 연계와 기업과의 연계로 인한 정치적 편향을 가진 언론들은 자신들의 편향에 기반하여 이슈에 대한 방어와 공격의 보도를 진행할 것이다. 따라서 정파적 언론은 이슈에 대한 보도양의 차이가 클 확률이 높다. 공격하는 언론은 많은 보도를 생산할 것이고, 방어하는 언론은 우선 보도를 하지 않을 것이다.

둘째, 앞서 언급한 정치적 연계, 경제적 이익, 독자의 선호 변화, 이슈에 대한 조직내부의 이해와 결정에 따라 언론들은 기존의 보도 행태와 다른 행태가 나타날 가능성이 크고, 기존의 보수 언론들 사이에서는 보도행태의 차이가 분명하게 나타날 가능성이 크다. 매우 큰 정치적 사건이며, 그 파급효과에 따라 언론 자신들의 이익관계에도 큰 영향을 미칠 것이기 때문에 그 판단과 보도행태는 다양하게 나타날 가능성이 있다.

셋째, 언론들은 매우 중요한 정치적 사건에 대처하는 과정에서 자신들이 이익을 위하여 필요에 따라 자신들과 연계된 정당과 기업들과의 연계를 단절할 수 있고, 연계를 더욱 강화할 수도 있다.

위와 같은 예상되는 기본적인 언론의 행태를 확인하기 위하여 본 논문은 텍스트 마이닝 기법과 설문조사 결과를 활용하여 언론의 보도행태를 실증적으로 분석할 것이다.

Ⅲ. 데이터와 연구방법

1. 데이터

본 논문은 두 가지 분석방법을 사용한다. 텍스트 마이닝 기법과 인터넷 설문조사 방법을 사용하였다.

첫째, 본 논문은 텍스트 마이닝(Text Mining) 기법을 활용하여 최순실 사건과 연관된 기사를 수집하여 분석하였다. 텍스트 마이닝(Text Mining)은 비정형 데이터(Unstructured Data)중에서 문자, 문서의 형태로 저장된 데이터들을 처리하기 위하여 개발되었다. 자연어처리(Natural Language Process)에 기반하고 있으며, 최근 컴퓨터 성능의 발전으로 인하여 더욱 고도화되고 있다. 앞으로 정치학뿐만 아니라 사회과학 전반에서 많이 사용될 것으로 생각된다.

분석대상 언론기사들은 2016년 3월 10일부터 탄핵일인 2017년 3월 10일까지 1년여 동안의 박근혜를 키워드로 검색한 기사들로서 주요 언론과 방송사의 기사를 수집하였다. 기사는 파이썬(python)을 이용한 웹 크롤링(web crawling)으로 수집하였다. 최순실 게이트는 8월 이후부터 본격적으로 보도되기 시작하였는데 1년 동안의 일반적인 경향과 8월 이후의 경향성을 살펴보고 위기상황에서의 보도형태 변화를 살펴보고자 하였다.

대상 언론은 총 10개로서 조선일보, 중앙일보, 동아일보, 매일경제, 한국일보, 한겨레, MBC, KBS, JTBC, SBS 이다. 대표적인 보수, 중도, 진보 언론으로 선정하였다. 본 논문은 언론의 정파성을 측정하기 위한 방식으로 박종희(2016)의 방식을 차용하였다. 최순실 사건에 대한 정당 간 태도는 극명하게 갈릴 것으로 보고, 더불어민주당과 새누리당의 공식 논평과 브리핑 자료를 수집하여, 두 당의 논평에서 사용하는 어휘를 기준으로 언론의 정파성을 측정하였다.

둘째, 언론의 보도행태의 차이를 만드는 원인을 추정하기 위하여 언론계 종사자들을 대상으로 전문가 설문조사를 진행하였다. 텍스트 마이닝 기법을 통해서 기사를 모으고 언론 보도의 정치적 편향을 보여줄 수는 있지만, 왜 최순실 사건에 대한 언론의 정치적 편향이 기존의 보도행태와 다르고 같은지를 설명할 수 없다. 따라서 그 원인을 추정하고자 전문가 설문조사를 진행하였다. 설문조사 진행기간 2018년 3월 23일부터

3월 31일까지이며 현재 국회 출입하는 정치부 기자와 최순실 관련 사건의 기사를 작성한 사회부 기자를 주 대상으로 하였고, 언론사 간부와 평론가 일부도 포함되었다. 전체 700명 이상의 기자들을 대상으로 이메일 설문을 진행하였으나 응답자는 65명이다. 언론사 형태별로는 일간지 종사자 35명, 지상파 방송국 6명, 종합편성 방송국 8명, 인터넷 언론사 11명, 주간지 4명, 월간지 1명이다. 취재경력으로는 5년 이상이 46명으로 가장 많고, 3~5년 사이가 7명, 3년 미만인 12명이다.

2. 연구방법

본 논문은 앞서 언급한바와 같이 정치적 이슈 특히 중요한 정치적 스캔들과 관련된 기사는 언론사별로 각기 정치적 연관성이 더 깊은 정당과 사용하는 어휘가 유사할 것이라고 가정한다. 분석 방법으로 박종희(2016)의 방식을 차용하였다. 첫째, 여론 흐름의 동적인 변화를 고려하여 한 주 단위로 정당과 언론사들의 정치적 경도를 새로 측정하여 해당 주제 각 언론사들이 기사에서 사용한 어휘가 더불어민주당, 새누리당이 논평에서 사용한 어휘와 얼마나 유사한지를 측정하였다. 둘째, 각 언론사들이 기사에서 사용한 어휘 중 1어절을 제외하고 더불어민주당과 겹치는 2어절, 3어절의 어휘와 새누리당(자유한국당)과 겹치는 2어절, 3어절의 어휘의 사용 빈도를 측정하였다. 이때 분석 대상의 어휘는 개별 정당의 정파적 입장을 나타낼 수 있는 어휘로서 특정 정당에서만 사용되는 어휘들 중 한 주 동안 2번 이상 언급된 어휘로 설정하였다.

분석에 사용하는 어휘들은 한국어 정보처리 패키지인 KONLPY의 TWITTER 라이브리리를 이용해 형태소(morpheme)로 정제하여 사용한다. 형태소를 사용한 이유는 어휘를 그대로 사용할 경우 같은 의미의 단어 묶음들이 다르게 수집되기 때문이다. 예를 들어 '박근혜 대통령과', '박근혜 대통령이'는 같은 의미의 2어절 단어지만 형태소로 정제하지 않으면 다르게 수집된다. 형태소 분석 뒤 조사, 어미, 접미사, 관형사, 특수문자 등을 제거하여 분석에 사용한다.

앞서 분석한 두 정당의 주차별 2, 3어절 어휘 묶음과 각 언론사의 어휘 묶음을 비교할 때 특정 정당에서 사용하는 어휘 묶음의 개수가 더 클 경우 겹치는 어휘도 더 많을 것으로 보았다. 따라서 겹치는 어휘 사용 빈도를 해당 정당의 어휘 묶음 개수로 나눠 주었다.

$$\text{언론 정치적 연관성 관측값} = \log \frac{(\text{자유한국당과 겹치는 어휘의 사용 빈도} / \text{자유한국당의 말뭉치 크기} + K)}{(\text{더민주와 겹치는 어휘의 사용 빈도} / \text{더민주의 말뭉치 크기} + K)}$$

log는 단조 증가 함수로서, 값이 +에 가까울수록 자유 한국과 유사한 어휘 사용이고, 값이 -에 가까울수록 더 민주와 유사한 어휘를 사용한 것이다. 중간에 특정한 예외 값이나 이상치를 보정하기 위해서 K를 더해줌으로서 평활화(smoothing) 처리하였다.

IV. 분석

1. 언론별 정당별 이슈 대응 빈도 분석

최순실 사건을 통한 박근혜 대통령의 정치적 스캔들은 시간이 갈수록 진정되지 않고, 결국 대통령 탄핵까지 이어졌다. 언론은 독자들의 관심, 소유구조, 정치적 연관성, 자신의 조직특성을 고려하여 보도태도를 정한다. 자신들이 불리한 정치적 위기상황에서 언론이 처음 반응하는 패턴은 옹호기사 생산과 기사를 생산하지 않는 것 두가지로 추정해볼 수 있다. 또한, 반대편 정파의 위기 상황에서 언론은 더 적극적으로 기사를 생산하고, 여론을 주도하려고 할 것이다. 정당들도 비슷한 패턴을 보이겠지만, 언론은 더욱 다양하고 더 많은 내용의 기사를 생산하려고 할 것이다. 따라서 최순실 사건에 대한 언론의 기사 수는 정당과의 연계성에 따라 크게 차이가 날것으로 추정할 수 있다.

우선 정당별로는 새누리당은 1140개, 더불어 민주당에서는 2098개로 2배에 달하는 논평과 브리핑 개수의 차이가 있다. 또한, ‘박근혜’를 키워드로 검색해서 나온 더불어민주당의 공식논평과 브리핑 수는 1243개이고, 새누리당의 논평 수는 273개이다. 더불어 민주당은 59.2%에 달하는 논평이 ‘박근혜’와 관련 있었던 반면, 새누리당은 논평 중 23.9%만이 ‘박근혜’와 관련이 있었다.

또한, 언론사별 기사의 수에서도 큰 차이가 있다. 언론사의 기사 차이는 각 언론사의 역량의 차이가 존재하기 때문에 기사건수의 차이로 해당 언론의 의도와 정치적 연계성

을 단정 지을 수는 없다. 하지만, 비슷한 역량을 갖춘 언론사들 끼리 기사의 건수가 크게 차이가 나고 있다. 가장 많은 기사를 만들어낸 곳은 15,916건의 기사가 작성된 JTBC 이다. 그 다음이 12,903건의 SBS 이며, KBS가 10,646 건이다. 신문사 중에는 유일하게 매일경제가 10,000건이 넘어서 10,162건의 기사를 생산했다.

방송국 중에 MBC는 5,405건으로 의도적으로 뉴스생산을 하지 않은 것으로 추정된다. 조선일보, 중앙일보, 동아일보도 기사가 많이 생산되지는 않았다. 언론사 역량으로 보면 비교가 안되는 한겨레가 8,140건의 기사를 보도했는데 그보다 더 큰 역량을 갖춘 조선일보, 동아일보, 중앙일보가의 보도건수가 8000건이 되지 않는 것은 의도적으로 기사를 많이 다루지 않은 것으로 유추할 수 있다.

〈표 1〉 언론사 별 박근혜 관련 기사수

	기사수
조선일보	7,901
중앙일보	6,221
동아일보	7,649
매일경제	10,162
한겨레	8,140
한국일보	5,917
MBC	5,405
KBS	10,646
SBS	12,903
JTBC	15,916

기사수를 시기별로 나누어서 보면, 거의 모든 시기 가장 많은 보도를 한 언론은 JTBC로, JTBC가 사실상 최순실, 박근혜 국정농단 사건에 대한 여론을 주도해온 것을 알 수 있다.

전문가 여론조사에서는 “박근혜 대통령관련 기사가 아래와 같이 언론사별 기사건수가 큰 차이를 보이고 있습니다. 방송과 신문의 특성을 고려하더라도 정치적 연관성이 강한 언론사별 보도건수의 차이가 커 보이는데 의도된 것으로 생각하십니까?”라고 물었다.

그 결과, ‘그럴 것이다.’라고 응답한 인원이 28명으로 전체 43.1%이며, ‘대체로 그럴 것이다.’로 응답한 인원이 23명으로 35.4%로 나타났다. 이는 전체 인원 중 78.5%가 기사건수의 차이가 언론사별 정치적 연관성과 연관된 의도된 것으로 판단하고 있다는 것을 의미한다.

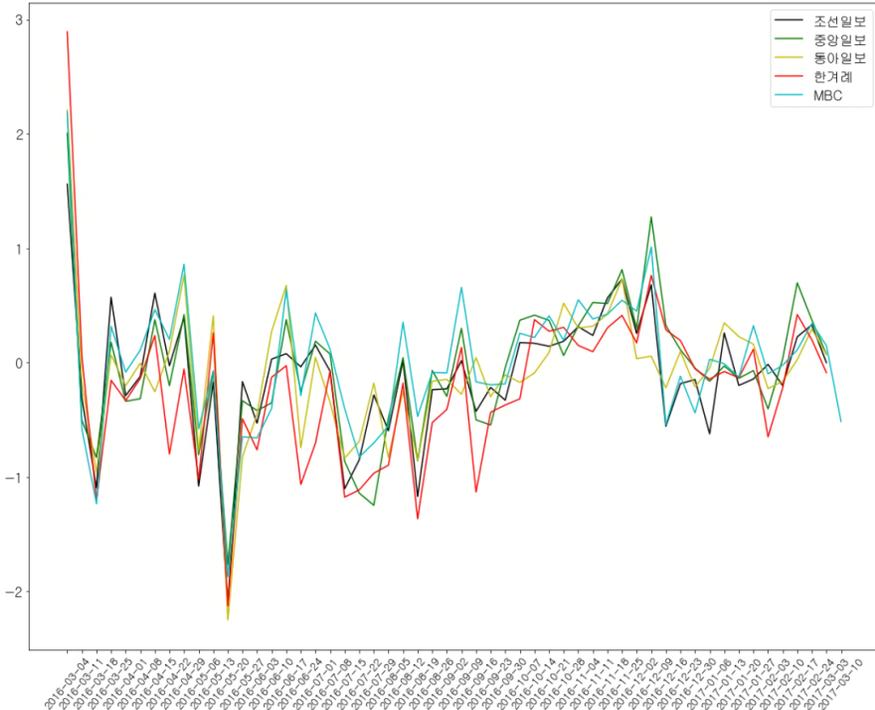
2. 시기별 언론의 정파적 연관성 변화

다음으로 시기별로 언론이 정파적 연관성의 변화가 발생하는지 살펴보았다. 최순실 사건은 꽤 오랜 기간 동안 여러 가지 많은 사건들이 이어지면서 발생하였다. 따라서 각 사건별로 언론의 정파적 경도 값의 차이가 발생하게 된다. 특정 사건에 대해서 각 언론사들의 판단이 각 정당과의 연관성이 더 있을 수도 있고 없을 수도 있다. 이는 언론사의 개별적인 판단이라고 이해될 수 있다. 따라서 각 정당의 판단과 언론의 판단이 동일할 수도 있고 다를 수도 있지만, 서로 상호작용한다고 볼 수 있다. 하지만, 이는 기본적으로 여론의 흐름이 반영된 것으로 이해된다.

본 논문은 여론을 조사한 각 언론매체들의 측정된 정치적 연관성 값들의 평균으로 정의하였다. 언론매체별 정치적 연관성은 언론매체별 정치적 연관성 관측 값에서 여론을 뺀 값으로 정의하였다(언론매체별 정치적 연관성 정도 = 언론매체별 정치적 연관성 관측 값 - 여론). 시기별 언론의 정치적 연관성 관측 값의 패턴을 살펴봄으로써 여론의 흐름과 함께 정파적 대응의 패턴도 일정하게 이해할 수 있다.

[그림 2] 시기별 주요언론 정치적 연관성의 정도 변화

언론사별 정파성 관측 값



가독성을 위해 일부 언론사들만 차트위에 표기(다른 언론사들의 흐름도 몹시 유사)

시기별로 언론의 정치적 경도의 변화를 살펴본 결과 몇 가지 특징적인 것을 발견하였다. 우선, 시기별로 편차가 있지만, 새누리당과 가장 가까운 보도행태를 보인 것은 MBC와 중앙일보이다. 2016년 6월 정운호 게이트가 롯데그룹 비자금 조성 의혹 사건으로 발전되어가던 시기에 이를 방어하는 새누리당과 가장 가까운 보도를 한 것이 MBC와 중앙일보였다. 그리고 2016년 8월 16일 MBC에서는 우병우 수석을 조사하고 있는 이석수 특별감찰관과 조선일보 기자의 감찰내용 누설 의혹에 대한 폭로 보도가 나왔다. 이시기 새누리당과 가장 가까운 정파적 보도를 한 것도 MBC였다. 마지막으로, 정치적 경도가 새누리당과 가장 가까웠던 시기는 2016년 12월 9일 박근혜 대통령 탄핵 직전으로서 이를 막기 위한 시도로 이해할 수 있다.

두 번째, 조선일보는 2016년 4월 이후 새누리당과의 정치적 연관성이 사라진 보도를 지속해왔다. 이는 청와대와의 불화가 4월부터 본격화 되었던 것으로 이해된다. 새누리당과의 연관성이 가장 높은 보도를 한 것이 2016년 4월이 마지막이다. 탄핵시기를 제외하고는 조선일보는 2016년 4월부터는 새누리당과 정치적 연관성이 거의 없었다. 2016년 7월 18일 조선일보는 보도로 우병우 민정수석의 비리의혹을 처음 제기하였고, 7월 26일 TV조선을 통해 미르재단 관련 폭로를 보도하였다.

세 번째, 각 언론의 정치적 연관성의 변화는 결국 여론의 흐름이 반영된 것으로 이해할 수 있다. 진보, 보수 언론 모두 정치적 연관성의 정도가 각기 시기별로 변화되는데 그 변화의 양상이 비슷하게 흘러간다. 이는 각 시기별 여론에 기반하여 언론의 보도방향에 반영한 것으로 볼 수 있다. 각각 대표적인 진보, 보수 언론인 MBC와 한겨레는 자주 각 정당과의 연관성이 높은 보도를 하였지만, 전반적인 여론 흐름에서 벗어난 것은 아닌 것으로 이해할 수 있다.

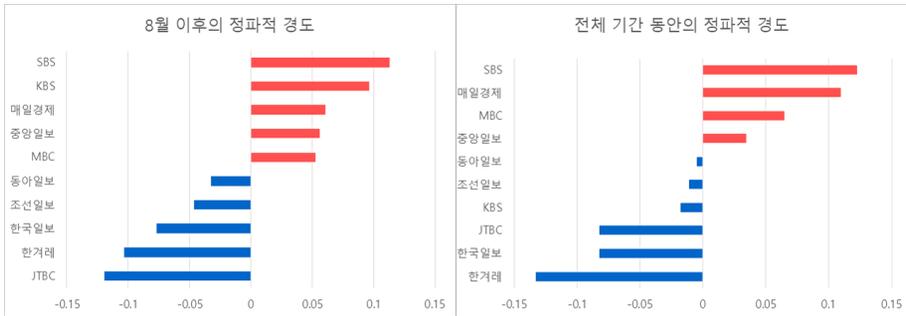
전문가 설문조사에서는 “한국의 언론사들이 여론의 흐름을 따라간 것으로 생각하십니까?”라고 물었다. 그 응답 중 22명(33.8%)이 ‘그럴 것이다.’, 38명(58.5%)가 ‘대체로 그럴 것이다.’라고 응답하였다. 정치적 연관성에 따른 편차는 존재하지만, 전반적으로 전체적인 여론의 흐름은 완전히 거부하는 형태의 보도는 이루어지지 않고, 여론의 흐름이 진보, 보수 언론 모두에 반영된 보도가 진행되었다라고 이해할 수 있다.

3. 언론사별 정치적 연관성

다음으로 언론사별 정치적 연관성의 변화를 측정하였다. 전체 1년간의 정치적 연관성과 최순실 사건이 본격적으로 보도되기 시작한 2016년 8월 이후 보도에서의 정치적 연관성의 정도를 각기 따로 보여준다. 박근혜-최순실 게이트 관련 보도는 8월 이후부터 본격적으로 보도되기 시작했다. 2016년 7월 26일, 조선일보는 TV조선을 통해 미르재단 관련 폭로를 최초로 보도하였고, 8월 2일, K스포츠재단의 비리와 미르와 K스포츠의 유사성에 대한 기사가 처음 보도되었다. 분석 기간 동안 수집된 90,860개의 기사들 중 72433개인 80%에 달하는 기사들이 보도되었다. 따라서 8월 이후의 정파적 연관성 정도가 박근혜-최순실 게이트 관련 정파적 연관성 정도를 가장 잘 표현할 수 있다고 볼 수 있다.

[그림 3]에서 보는바와 같이, 전체기간동안과 8월 이후 모두 가장 보수적인 보도연관성을 보인 언론사는 SBS이다. 전체기간동안 보수적인 연관성을 보인 언론사들은 SBS, 매일경제, MBC, 중앙일보이다. 진보적 연관성을 보인 언론은 한겨레, 한국일보, JTBC, KBS, 조선일보, 동아일보였다. 조선일보, 동아일보, KBS는 기본적으로 보수언론으로서 연관성 정도의 값이 매우 낮은 것으로 보여 이는 민주당 편향이라기 보다 여론의 흐름에 따라 보도편향을 변경한 것으로 이해할 수 있다.

[그림 3] 언론사별 정파적 경도



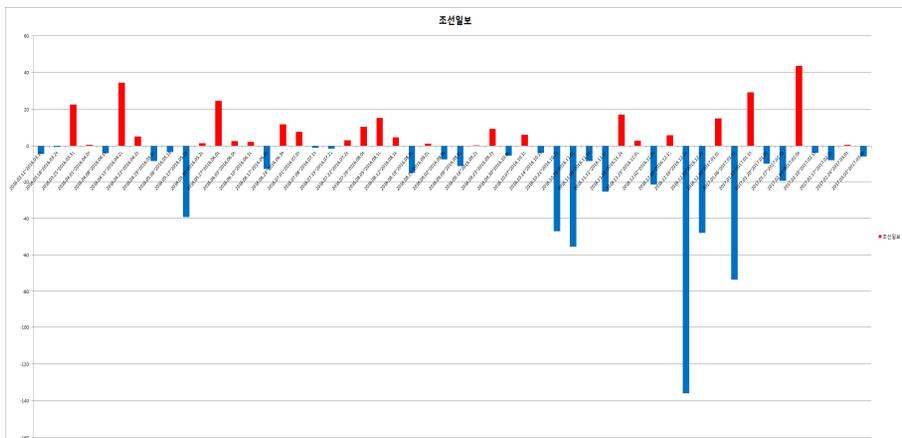
8월 이후 정치적 연관성 정도분석에서 보수적 연관성을 보인 언론은 SBS, KBS, 매일경제, 중앙일보 MBC 순으로 나타났다. 이 언론들은 적극적으로 박근혜 대통령을 보호하기 위한 보도 경향을 보였다고 볼 수 있다. 진보적 연관성을 보인 언론은 JTBC, 한겨레, 한국일보, 조선일보, 동아일보 순으로 나타났다. JTBC와 한겨레는 최순실 사건에 대한 추적보도를 지속했고, 결정적인 보도를 함으로써 박근혜 대통령 탄핵까지 이어지는 전개에 가장 큰 공헌을 한 언론사이다. 이 역할이 전체 보도편향 분석결과에 그대로 드러난 것이다. 한국일보는 원래 중도적 언론사로 알려져 있다. 본 논문에서도 중도적 언론 중의 하나로 선정하여 분석하였다. 한국일보의 보도 편향도 여론의 흐름에 따라 민주당 편향이 드러난 보도 흐름을 보여준다. 조선일보와 동아일보는 보수언론이지만, 대중의 여론 변화에 맞추어서 그 편향을 변경하였다고 이해할 수 있다. 이는 전체기간 분석결과에 비추어 봤을 때, 좀 더 적극적으로 민주당 편향의 보도를 진행하였다고 이해할 수 있다.

전문가 설문조사 결과에서도 언론사들의 정치적 연관성에 기반 한 보도행태가 언론사별 정파적 연관성을 반영한 결과라는 인식을 확인할 수 있다. “언론사별 정파적 연관성을 측정한 결과, 전체기간동안 새누리당과 가장 연관이 있는 보도행태를 보인 것은 SBS, 매일경제, MBC 이며, 가장 민주당과 연관성이 있는 매체는 한겨레, 한국일보, JTBC로 파악이 됩니다. 이 분석결과가 해당 언론의 정치적 연관성을 반영한 것으로 생각하십니까?” 라는 질문에 ‘반영한 것이다.’라고 응답한 인원이 22명(33.8%)이고, ‘대체로 반영한 것이다.’라고 답한 인원 32명(49.2%)이었다. 전체적으로 65명 중 54명(83%)가 각 언론사들의 정치적 연관성이 반영된 보도행태라고 응답하였다.

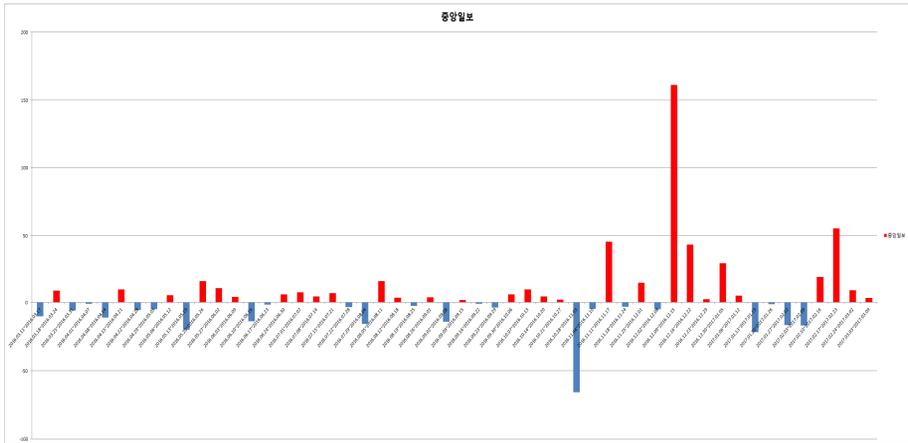
4. 보수 언론별 정파적 경도의 변화

앞선 분석에서 조선일보와 동아일보의 보도편향이 기대와 달리 민주당쪽으로 편향이 드러난 것은 예상외의 결과였다. 이를 통해 그동안 보수언론의 대표 주자였던 조선일보, 중앙일보, 동아일보가 각기 다른 보도편향을 드러내고 있음을 확인하였다. 이를 좀 더 구체적으로 확인하기 위하여 주별로 조선일보, 중앙일보, 동아일보의 주별 보도편향을 별도로 비교하였다. [그림 4], [그림 5], [그림 6]은 각각 조선일보, 중앙일보, 동아일보의 주별 보도편향을 보여준다. 이는 각 언론사의 정치적 연관성 정도에 기사수를 곱한 값으로서 주별로 보도의 편향을 보여준다.

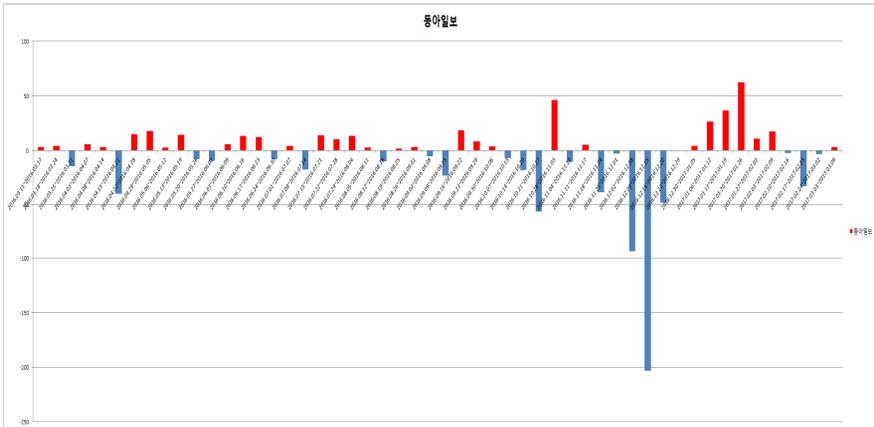
[그림 4] 조선일보의 주별 보도편향



[그림 5] 중앙일보의 주별 보도편향



[그림 6] 동아일보의 주별 보도편향



조선일보는 중앙일보와 동아일보에 비해 기사 건수가 상대적으로 더 많고, 주별로 구분하여 분석한 결과 새누리당 편향의 보도편향을 보이는 경우도 상당히 나타났다. 보수 편향이 분명하게 드러나는 주가 17주정도 되며, 민주당 편향의 보도를 한 경우는 24주로 나타났다. 조선일보는 2016년 7월 26일 TV조선을 통한 미르재단 500억 모금 지원 기사를 보도한 이후 더욱 분명하게 정권에 대한 비판적 태도를 보였고, 더불어

민주당과 유사한 어휘 사용을 보여주었다. 이런 흐름은 박근혜 정부와의 불화와 여론의 흐름에 따른 결과로 이해된다.

반면에, 중앙일보는 매우 분명하게 보수편향의 태도를 유지하고 있는 것으로 보인다. 그래서인지 앞서 살펴본 바와 같이 기사수도 가장 적다. 민주당 편향의 보도가 있긴 하였지만, 최순실 사건이 분명해지고 박근혜 대통령의 탄핵안이 가결됨과 동시에 보수적 색채가 더욱 강해지며, 기사수도 크게 증가한다.

앞서 살펴본 [그림 2]에서 확인할 수 있듯이, 중앙일보는 최순실 사건에 대한 비판적 여론이 거세지는 동안 9~10번째로 적은 보도를 하다 탄핵안이 가결됨과 동시에 5~8번째의 보도 숫자로 변화한다. 보수세력의 위기가 분명해진 속에서 조선일보와 달리 보수성향을 더욱 분명하게 한 것으로 이해될 수 있다. 그리고 최순실 게이트에서 삼성의 말 지원 등 모기업이라고 볼 수 있는 삼성의 연관성으로 인해 무조건적인 비판을 할 수 없는 상황도 반영된 것으로 이해된다. 삼성과의 관계는 단순히 소유주의 혈연관계뿐만 아니라 광고로 연결된 경제적 이해관계의 반영이기도 하다. 더군다나 JTBC의 태블릿 PC 특종 보도 이후 삼성과의 관계는 좋지 않은 상황에서 삼성과 연관된 최순실 사건에 대한 비판에 집중할 수 없었을 것으로 추정된다.

동아일보는 전반적으로 보도의 방향이 새누리당 편향인 주가 많지만, 기사수가 많지 않다. 반면에 탄핵 이후 민주당 편향의 기사수가 급증하였다. 중앙일보가 새누리당 편향의 보도를 많이 생산한 것과 정반대로 민주당 편향의 기사수가 탄핵이후 급증하였다. 이는 대중의 여론 흐름을 쫓아간 결과로 이해된다. 몇 주 후 다시 새누리당 편향의 보도가 주를 이룬 것과 잘 대비되는 결과이기도 하다.

이상에서 살펴본바와 같이, 같은 보수 언론이라도 각기 다른 양태로 여론의 흐름을 반영하기도 하고, 지속적으로 원래의 정치적 편향을 그대로 유지하기도 하는 등 다양한 선택을 하고 있다는 것을 알 수 있다.

전문가 설문조사에서는 이와 관련하여 두가지 설문항을 언론인들에게 물었다. 첫째, “대표적인 보수언론인 조선일보, 중앙일보, 동아일보의 최순실 사건에 대한 보도 행태를 측정한 결과 각 언론들이 모두 다른 패턴을 보였습니다. 보수언론의 보도패턴의 다양성이 발생한 원인이 무엇이라고 생각하십니까?” 라는 설문에는 26명(40%)은 ‘정치적 연관성의 유지와 단절에 대한 다른 판단.’이라고 응답했고, 두 번째로 많은

수인 13명(20%)이 ‘독자들의 선호에 대한 다른 판단’이라고 답했다. 세 번째는 10명(15.4%)이 ‘전체적인 국민여론의 흐름을 따른 결과’라고 답했다. 소유주의 판단의 결과라는 응답도 9명(13.8%)이 있었고, ‘광고와 매출에 영향을 미치는 기업과의 관계 고려’도 7명(10.8%)이 응답했다. 설문결과, 정치적 연관성이 가장 중요한 판단 근거가 된다는 의견이고 그 다음이 독자들의 선호를 고려한 것이었다. 소유주의 판단과 기업과의 관계도 무시하기는 힘든 수준의 응답수가 나왔다. 언론의 보도행태가 정해지는데 있어서 다양한 원인들이 복합적으로 고려된다고 이해될 수 있다. 두 번째로는 정치와 경제적 연관성에 중심을 두고 문항을 작성하였다. “중앙일보, 매일경제, SBS 세 언론사가 가장 보수적인 보도경향을 보였습니다. 그 원인이 정당과 기업과의 연관 중 무엇 때문이라고 생각되십니까?”라고 물었다. 가장 많은 응답은 ‘정당(정부)과 기업과의 관계를 모두 고려한 것이다.’였고 전체 39명(60%)이 응답했다. 그 다음으로는 12명(18.5%)이 ‘기업과의 연관성 때문이다.’라고 응답했다. 이처럼 언론인들의 의견에서도 보도경향을 결정하는데 있어서 정치적 연관성과 기업과의 관계가 많이 고려된다는 것을 알 수 있다. 대표적인 보수언론인 중앙일보, 매일경제, SBS, 조선일보, 동아일보 등이 최순실 사건에 대한 보도의 방향을 설정할 때 국민의 여론도 고려하였지만, 정치적 연관성과 기업과의 관계도 함께 고려하여 보도행태가 정해진 것으로 이해할 수 있다.

V. 결론

본 논문은 텍스트 마이닝 기법과 전문가 설문조사를 활용하여 한국언론의 최순실 사건에 대한 보도행태의 다양성과 그 원인을 분석하였다. 한국정치에서 언론은 객관적 정보전달자라기 보다 정치세력과 함께 연관된 하나의 행위자로 인식되는 경우가 많다. 이를 실증적으로 확인하기 위해서 최근 가장 큰 정치적 스캔들이었던 최순실 사건에 대한 보도에서의 언론사별 보도행태의 차이와 그 다양성의 원인을 분석하고자 하였다. 각기 다른 정치적 편향성을 측정하기 위하여 새누리당과 민주당의 공식 브리핑과 논평에서 사용된 어휘들을 기준으로 사용하였다.

그동안 한국 언론은 정치적 편향에 많이 노출되어 정권의 성격에 따라 많은 변화를 겪어왔다. 따라서 한국의 언론은 정치적, 경제적, 사회적 이익을 위하여 스스로 정치적 행위자가 되기도 하고, 권력으로부터의 압력으로 인해 정치적 행위자가 되기도 하였다. 최근 10년 동안 보수정권이 집권하면서 한국 언론의 보수 편향성을 더욱 증가되었다. 하지만, 최순실 사건으로 인해 박근혜 대통령 까지 탄핵이 되고 기소되어 재판까지 받는 상황이 되면서 한국의 정치적 보수는 위기를 맞았고, 보수적인 언론도 이 위기에 대응해야했다. 이 대응에서 각 언론들은 각기 다른 패턴을 보인다.

분석결과, 분석대상 언론 중 가장 보수적이며 새누리당과 연계된 편향성을 가진 언론은 SBS 로 나타났다. 가장 진보적이며 민주당 편향성을 갖는 언론은 한겨레와 JTBC이다. 그리고 각 사건에 대응하는 보수언론의 패턴은 예상보다 다양하게 나타나고 있으며, 대중의 여론 흐름을 쫓아서 보도편향을 바꾼 언론도 등장하였다. 특히 가장 보수적인 언론으로 인식되어온 조선일보와 동아일보의 보도편향의 변화는 시사하는 바가 크다.

향후 후속 연구에서는 진보언론의 대응과 비교하는 연구를 진행할 필요가 있을 것이다. 최근 한국의 정치적 진보가 위기상황에 있었던 노무현 정부 말기의 한겨레와 경향신문 등 진보적 편향성을 갖고 있다고 인식된 언론의 보도편향과 보수언론의 보도편향을 비교분석해봄으로서 좀 더 언론의 정치적 편향성에 대한 종합적인 분석을 시도할 수 있다.

또한, 앞으로 텍스트 마이닝 기법을 활용한 정치분석 시도는 향후 계속 증가할 것이다. 언론과 정당들의 정치적 어휘 분석, 정치적 어휘의 실질적인 의미와 그 시대별 변화 등 많은 연구들이 진행될 수 있을 것이다.

참고문헌

- 고영신. 2007. “정권의 성격변화와 언론보도.” 『커뮤니케이션 이론』 제3권 1호: 156-196.
- 김병진·이상은·양종훈. 2017. “박근혜 대통령 탄핵 보도 편향성에 관한 연구.” 『한국콘텐츠학회논문지』 제17권 11호: 554-566.
- 김승수. 2003. “언론권력의 정치경제학: 베네주엘라, 이탈리아, 영국의 사례분석.” 『한국언론학보』 22호: 39-75.
- 박종희. 2016. “세월호 참사 1년 동안의 언론보도를 통해 드러난 언론매체의 정치적 경도.” 『한국정치학회보』 제50권 제1호: 239-269.
- 윤영철. 2000. “권력 이동과 신문의 대북 정책 보도: 신문과 정당의 병행 관계를 중심으로.” 『언론과 사회』 제 27호: 48-81.
- 이소영. 2017. “한국 유권자의 정치적 태도에 대한 미디어 효과.” 『21세기정치학회보』 제27권 4호: 1-29.
- 최진봉. 2016. “방송 뉴스의 재난보도 콘텐츠에 대한 분석.” 『한국콘텐츠학회논문지』 제16권 12호: 539-550.
- Altschull, Herbert J. 1984. *Agents of Power: The Role of the News Media in Human Affairs*. New York: Longman.
- Curran, J. 1990. “The new revisionism in mass communication research: A reappraisal.” *European Journal of Communication* 5: 135-164.
- Hall, S., Critcher, C., Jefferson, T., Clarke, J., and Roberts, B. 1978. *The social production of news*. na.
- Hall, S. 1979. “Culture, The Media and Ideological Effect.” in J. Curran et al. eds. *Mass Communication and Society*, Haverly Hills: Sage. 315-348.
- Hallin, Daniel. 1994. “From Vietnam to the Salvador : Hegemony and Ideological Change.” *We keep America on Top of the World*. London: Routledge Press, 58-86.

- McNair, B. 1994. *News and Journalism in the UK: A textbook*. London: Routledge.
- Park, Eunjeong and Sungzoon Cho. 2014. “KoNLPy: Korean natural language processing in Python.” Proceedings of the 26th Annual Conference on Human & Cognitive Language Technology.

The Varieties of Newsgathering and Processing Activities of Korean Press: Focusing on Choi Soon-Sil Scandal

Sungwoo Lee* · Hojoon Lee** · Byungkun Lee***

ABSTRACT

This article aims to analyze the newsgathering and processing activities of Korean Press about Choi, Soon-Sil Scandal, a 2016 South Korean political scandal, using Text mining technique and specialist survey for journalist's opinions. Many Korean people have realized that the press in Korean Politics is important actors linked to particular political groups rather than objective reporter, and that the press in Korean Society linked to conglomerates because of their profits as a company. As a result of an analysis, this article finds that SBS, private broadcasting, is the most conservative press linked to the Saenuri Party (conservative party), whereas two progressive presses are Hangyoreh shinmun, JTBC linked to the Democratic Party (progressive party). We find very various newsgathering and processing activities of conservative presses. Also, many presses changed their positions about the political issues according to major public opinions. Especially, this article finds that the Chosun Ilbo, the JoongAng Ilbo, Dong-A Ilbo, three major conservative presses, took a very different positions about Choi, Soon-Sil Scandal. And, we find that many presses make different positions about political economic issues based on the political connections and relationship with big business. Based on the result, this article shows the need of researches about relationship between the newsgathering and processing activities of Korean Press and Political Parties and Big businesses.

Keywords : The Press, Political Scandal, Text Mining, Report Behavior of Press

투고일: 2018.04.06. 심사일: 2018.04.10. 게재확정일: 2018.04.20.

* Korea University

** Korea University

*** Korea University

[DOI] <http://dx.doi.org/10.21487/jrm.2018.5.3.1.57>

【연구논문】

스마트건강증진서비스 효과성 분석 연구: 서울시 사례*

오 수 경**·신 창 환***

논문요약

본 연구에서는 서울시 7개 자치구에서 지역사회서비스투자사업의 스마트건강증진서비스를 1년간 참여한 만 19세 이상의 성인 282명을 연구대상으로 하였다. 서비스 효과성 측정은 서비스를 실시하기 전·후의 대사증후군 변화, 연령대별 대사증후군 변화를 살펴보고, 또한 대사증후군 측정 항목별 진단 기준에 따른 유소견자 현황을 살펴보고, 대사증후군 유소견자의 변화를 살펴보고, 대사증후군은 체질량, 허리둘레, 수축기혈압, 이완기혈압, 공복혈당, 중성지방, HDL-콜레스테롤 등 7가지를 측정하였다. 본 연구의 주요 결과는 다음과 같다. 첫째, 대사증후군 측정 항목별 변화에서 서비스 이용 전보다 이용 후 체질량, 허리둘레, 수축기혈압, 이완기혈압, 공복혈당, 중성지방이 감소하였고, HDL-콜레스테롤이 증가하여 대사증후군의 긍정적인 변화를 보였다. 둘째, 연령대별 대사증후군 변화의 차이가 있는 것으로 나타났고, 40~60세의 연령대에서 서비스 효과가 가장 많이 나타난 것으로 분석되었다. 셋째, 대사증후군 유소견자 감소율은 최소 1.5% ~ 최대 83.3%로 나타났고, 이러한 감소율은 연령대별 집단에 따라 차이가 있었다. 넷째, 대사증후군 유소견자만을 대상으로 분석한 결과, 7가지 측정 항목 모두에서 긍정적인 변화를 보였다. 본 연구결과에 따라 서비스 이용 대상의 확대, 인구학적 특성에 따른 차별적 서비스 지원방안, 유소견자에 대한 추가서비스 지원, 서울시 종합적 대사증후군관리 차원의 정책으로 보건소 사업과의 연계 또는 조정 등을 제안하였다.

주제어 : 사회서비스, 바우처, 효과성, 대사증후군, U-Health 시스템

* 이 논문은 서울시 지역사회서비스지원단에서 수행한 2016년 서비스 효과성 측정 연구사업의 내용을 수정·보완한 것임.

** 주저자, 이화여자대학교 사회복지대학원 석사, (전)서울복지재단 지역사회서비스지원단

*** 교신저자, 경북대학교 사회복지학부, 부교수

I. 서론

통계청 발표에 따르면 암, 심장 질환, 뇌혈관 질환, 당뇨병, 고혈압의 만성질환과 관련된 질환들이 한국인의 10대 사망원인으로 조사되었다(통계청, 2016). 이러한 뇌혈관 질환, 당뇨, 고혈압과 같은 만성질환은 건강한 생활습관 실천 및 관리로 예방이 가능하지만 관련 유병률을 지속적으로 증가하고 있는 추세이다. 한국의 대사증후군 유병률을 살펴보면, 2001년 기준 20세 이상 성인의 22%, 2005년 30.4%, 2007 31.3%로 조사되었고, 가장 최근인 2014년에는 한국인 30대 이상 성인 4명 중 1명이 대사증후군으로 진단된 것으로 조사되었다(대한비만학회, 2012; 국민건강보험공단, 2014). 특히 연령이 증가할수록 그 심각성은 더욱 심해지는 것으로 조사되고 있어 연령대별 관리의 필요성 또한 주목되고 있다(Ford et al., 2004; 강상경·전해숙, 2013). 이처럼 대사증후군 질환의 체계적 관리가 필요함에도 불구하고 적극적 관리가 이루어지지 않고 있는 것을 알 수 있다.

이러한 시대적 배경에 따라 보건복지부에서는 2010년 지역사회서비스투자사업의 일환으로 사전·예방적 건강관리서비스를 통해 건강증진, 질병예방 및 의료비 절감을 도모하며, 궁극적으로는 삶의 질 향상을 위해 지역별 U-Health(Ubiquitous-Health)를 활용한 대사증후군 관리서비스를 추진하였다(송태민, 2011). 이러한 서비스는 국가적으로 만성질환이나 건강관리에 대한 관심도가 증가되고, 보건의료 패러다임이 치료 중심에서 예방 중심으로의 변화를 반영한 것으로 정책적 차원에서 U-Health를 활용한 대사증후군 관리서비스를 도입한 것으로 볼 수 있다.

이에 따라 서울시에서는 지역사회서비스투자사업의 일환으로 스마트건강증진서비스(구, U-Health 시스템을 활용한 대사증후군 서비스)를 2013년부터 도입하여 서울시민들에게 제공하고 있다. 기존의 의료·건강서비스는 건강관리 비용 부담, 직접 내방해야 하는 시공간적 요인, 거동이 불편하거나 건강이 좋지 못할 때 건강 요인, 건강관련 정보의 미충족 등 이용시 등 장애요인이 많았다면, 스마트건강증진서비스는 의료·건강관리에 정보기술을 접목 및 유무선 통신망을 활용하여 서비스 제공자-이용자간 연결을 통해 시공간 제약 없이 이용자의 건강상태를 파악하고 진단할 수 있다(박정선, 2012). 서울시 스마트건강증진서비스 대상자 선정기준은 기준 중위소득 120% 이하의

만 19세 이상 성인으로서 허리둘레, 체질량 지수, 혈압, 중성지방, HDL-콜레스테롤, 공복혈당 등 대사증후군 측정 항목 중 1개 이상 유소견자를 대상으로 한다. 서비스 제공 내용을 살펴보면, 건강 측정할 수 있는 기기와 태블릿 PC를 이용자에게 지급하여 이용자가 언제 어디서든 자가검진할 수 있고, 건강기록 홈페이지를 통해 상시 모니터링 할 수 있다. 또한 이용자의 지속적 건강관리실천을 위해 건강측정기록을 전문가가 실시간 모니터링하며, 매월 최소 1회 대면 전문상담 진행, 문자메시지를 통해 독려하고 이메일로 맞춤형 건강관리 정보를 제공한다.(서울시지역사회서비스지원단, 2016).

2015년 국민건강통계에 따르면 최근 1년간 병의원에 가지 못한 19세 이상 성인은 남자 10.1%, 여자 15.4%로 총 25.5%에 달하며, 연령대별로 살펴보면, 30대 15.5%, 40대 10.1%, 50대 11.7%, 60대 11.1%, 70대는 13.8%로 조사되었다. 또한 병의원 미충족의료 이유로 '시간이 없어서(4.8%)'가 1위를 차지했다.(보건복지부 질병관리본부, 2016). 이러한 차원에서 서울시의 스마트건강증진서비스는 시공간 제약 없이 직접 찾아가는 서비스 제공, 전문상담, 모니터링 및 건강관련 맞춤형 정보 제공이 가능하기 때문에 바쁜 직장인, 복지사각지대, 독거노인을 중심으로 중요한 역할을 할 수 있을 것으로 판단된다.

따라서 본 연구에서는 서울시 스마트건강증진서비스에 대한 효과성을 분석하고자 한다. 서비스 효과성은 이용자들의 서비스 이용 전과 후의 대사증후군 측정항목별 변화를 살펴볼 것이며, 구체적인 연구목적은 다음과 같다. 첫째, 대사증후군 측정항목별 이용자의 변화를 살펴보고, 둘째, 연령별 집단을 구분하여 각 집단의 대사증후군 측정항목별 변화를 살펴보고자 한다. 셋째, 대사증후군 진단 기준에 따라 서비스 이용 전과 후의 유소견자 수 감소율을 살펴보고, 넷째, 유소견자 수 감소율을 살펴보고, 넷째, 유소견자의 서비스 이용 전과 후의 변화를 살펴볼 것이다.

이러한 효과성 측정 연구는 실질적으로 이용자에게 서비스가 효과적으로 작용하였는지, 이용자의 실질적인 건강상태가 긍정적으로 변화하였는지 등 사업의 성과를 확인하고, 궁극적으로는 정부 정책의 당위성 확보에 필요한 기초자료로 활용될 것이다.

II. 연구방법

1. 연구대상 및 자료수집

본 연구의 모집단은 2015년 3월부터 2016년 5월까지 서울시 스마트건강증진서비스 이용자이며 이 중 이용정지, 이사, 연락두절 등 중도탈락자를 제외하고 사전·사후 검사가 모두 완료된 이용자 282명을 연구대상으로 하였다. 자료수집은 1차적으로 제공기관과 간담회를 통해 본 연구의 목적과 방향에 대해 설명한 후, 연구의 가능성을 타진하고 분석내용 및 자료수집 방법에 대해 논의하였다. 이후 제공기관의 협조를 받아 이용자의 일반적 특성 및 각 자치구별 보건소에서 측정한 대사증후군 사전·사후 검사결과를 약 1개월간 2차례에 걸쳐 이메일로 회신 받아 분석하였다. 분석에 불필요한 자료는 즉시 폐기하였으며 이용자의 개인정보가 절대 노출되지 않도록 하였다.

2. 자료분석

본 연구는 대사증후군 관련 맞춤형 건강관리 서비스를 제공하여 대사증후군의 측정 항목별 변화 정도를 사전·사후로 비교한 단일집단 사전사후 설계(One-group pretest-posttest design)방법을 사용하였다. 수집된 자료는 에디팅, 코딩·편칭, 신뢰도를 높이기 위한 데이터 클리닝 과정을 거친 후 SPSS 21.0을 이용하여 분석하였다. 연구대상자의 일반적 특성 및 생활습관, 대사증후군 측정 항목별 유소견자 수를 알아보기 위해 기술통계 및 빈도분석을 실시하였으며, 서비스 이용 전과 서비스 이용 후 생활습관 변화, 전체집단의 대사증후군 측정항목별 변화, 연령별 변화, 유소견자 변화 등을 확인하기 위해 대응표본 t검증(Paired t-test)을 실시하였다.

3. 측정도구

1) 대사증후군 측정 항목별 건강측정지

서울시에서 실시하고 있는 진단 기준을 살펴보면¹⁾, 허리둘레: 남성 90cm 이상, 여성 85cm 이상 또는 체질량 지수 25kg/m² 이상, HDL-콜레스테롤: 남성은 40mg/dl

미만, 여성 50mg/dl 미만, 중성지방: 150mg/dl 이상, 공복혈당: 100mg/dl 이상, 혈압: 수축기 130mmHg 이상 또는 이완기 85mmHg 이상인 경우 유소견자로 진단하고, 5가지 기준 가운데 3가지 이상 해당되는 경우를 대사증후군으로 보고 있다. 하지만 스마트건강증진서비스는 조기개입을 통한 예방 및 만성질환 위험요인 감소를 위한 목적으로 1개 이상 유소견자를 서비스 대상자로 선정하고 있다.

2) 생활습관 설문지

생활습관은 약물복용, 음주, 흡연, 운동, 영양 등 5가지로 습관을 조사하였다. 먼저 음주는 최근 1개월 동안의 음주 횟수를 조사하였고, 흡연은 흡연자/비흡연자로 구분하여 흡연 여부를 조사하였다. 약물복용은 당뇨, 고지혈증, 고혈압 등 질환의 치료제 복용 여부를 조사하였다. 운동습관은 운동 규칙성, 운동 빈도를 조사하였다. 운동 규칙성은 숨이 찰 정도의 운동을 지난 3개월간 규칙적으로 실행하였는지, 운동 빈도는 땀이 날 정도의 운동을 1주일에 몇 번 하는지 응답하도록 하였다. 영양습관은 한 끼 식사 시 걸리는 식사시간을 조사하였다.

III. 분석결과

1. 이용자의 일반적 특성

스마트건강증진서비스 이용자의 일반적 특성을 살펴보면, 남성은 106명(37.6%), 여성 176명(62.4%)으로 남성보다 여성이 많았다. 연령은 60대가 85명(30.1%), 40대가 57명(20.2%), 70대 이상이 50명(17.7%), 30대 이하가 43명(15.3%)의 순으로 조사되어 60대가 가장 많았으며, 평균연령은 56.3세로 나타나 주로 4~60대 이용자가 많은 것을 알 수 있다. 학력은 전문학사·학사 78명(52%), 고졸 62명(41.3%), 석사 이상 10명(6.7%) 순으로 조사되어 주로 고졸, 전문학사·학사 수준의 이용자가 많이 이용하는 것을 알 수 있다. 결혼 유무를 살펴보면 기혼 129명(86%), 미혼 21명(14%)

1) 서울시 대사증후군관리사업지원단, <http://www.5check.or.kr/index.html>

으로 조사되어 기혼이 미혼보다 압도적으로 많았다. 가구원수는 3인 84명(29.8%), 4인 58명(20.6%), 5인 45명(16.0%), 2인 44명(15.6%), 1인 31명(11%) 순으로 나타나 3인 가구가 가장 많은 것으로 나타났다. 월 평균 소득은 100만 원 이상~200만 원 미만 81명(28.7%), 200만 원 이상~300만 원 미만 54명(19.1%), 100만 원 미만 49명(17.4%)으로 나타나 주로 100만 원 미만에서 300만 원 미만 사이의 소득이 있는 이용자가 많은 것으로 나타났다. 특히, 0원인 경우에는 건강보험료 면제 대상자로서 기초생활수급자가 29명(10.3%)으로 분석되었다.

〈표 1〉 이용자의 일반적 특성(n=282)

구분		빈도(명)	비율(%)
성별	남	106	37.6
	여	176	62.4
연령 (M=56.3)	30대 이하	43	15.3
	40대	57	20.2
	50대	47	16.7
	60대	85	30.1
	70대 이상	50	17.7
학력	고졸이하	62	41.3
	전문학사·학사	78	52.0
	석사 이상	10	6.7
결혼 유무	기혼	129	86.0
	미혼	21	14.0
가구원수 ²⁾	1인	31	11.0
	2인	44	15.6
	3인	84	29.8
	4인	58	20.6
	5인	45	16.0
	6인 이상	20	7.0
월평균 소득	0원*	29	10.3
	100만 원 미만	49	17.4
	100만 원 이상 ~ 200만 원 미만	81	28.7
	200만 원 이상 ~ 300만 원 미만	54	19.1
	300만 원 이상 ~ 400만 원 미만	46	16.3
	400만 원 이상 ~ 500만 원 미만	18	6.4
	500만 원 이상	5	1.8
		282	100

*기초생활수급자

2. 이용자의 생활습관 변화

이용자의 생활습관 변화를 살펴보기 위해 월 평균 음주 정도, 흡연 여부, 약물복용 수, 운동습관, 영양습관 등 총 5가지를 분석하였다.

먼저 월 평균 음주 정도를 살펴보면, 사전에는 68.1%의 이용자가 금주하였다면, 사후에는 69.1%의 이용자가 금주하는 것으로 나타났고, 월2회 이상 음주하는 경우에는 서비스 이용 전보다 이용 후 감소하였다. 약물복용 수를 살펴보면, 복용하지 않는 이용자가 55.7%인 것으로 나타났고, 1가지 복용하는 이용자는 이용 전보다 이용 후 감소하였으나 3가지 복용하는 이용자는 증가한 것으로 나타났다. 이러한 약물복용 이용자 수의 증가는 서비스 제공을 통해 자신의 건강상태를 인지하고 적절한 약물 투여를 통한 적극적 자기주도 건강관리가 잘 이루어지고 있는 것으로 해석할 수 있다. 운동습관은 서비스 이용 전보다 이용 후 더욱 규칙적으로 운동하는 것으로 나타났고, 운동 빈도는 주3회 미만은 감소한 반면 주 3회 이상 운동하는 이용자 수는 증가한 것으로 나타났다. 영양습관을 살펴보면, 서비스 이용 전 보다 이용 후 10분 이내로 식사하는 이용자 수는 감소한 반면 10분 이상 식사하는 이용자 수는 증가한 것으로 나타났다.

따라서 이용자의 생활습관이 서비스 이용 전보다 이용 후에 긍정적인 변화를 보인 것으로 분석되었다.

2) 월 평균 소득은 서울시의 협조를 받아 직장가입자 기준 건강보험료로 가정하고 추정 한 소득이다. 계산식은 총 건강보험료=보수월액*건강보험료율(2016년 기준, 6.12%)에서 건강보험료의 개인부담 50%를 적용하여 건강보험료(개인)=보수월액*3.06%로 계산 하여 월 평균 소득을 추정하였다.

〈표 2〉 이용자의 생활습관 변화(n=282)

구분		사전 N(%)	사후 N(%)	
월 평균 음주 정도*	금주	192(68.1)	195(69.1)	
	월 1회 이하	43(15.2)	50(17.7)	
	월 2~4회	38(13.5)	34(12.1)	
	월 5회 이상	9(3.2)	3(1.1)	
흡연 여부	예	26(9.2)	23(8.2)	
	아니오	256(90.8)	259(91.8)	
약물복용 수 (고혈압, 당뇨, 고지혈증)	없음	157(55.7)	157(55.7)	
	1가지	73(25.9)	72(25.5)	
	2가지	43(15.2)	43(15.2)	
	3가지	9(3.2)	10(3.5)	
운동습관***	운동 규칙성	예	134(47.5)	172(61.0)
		아니오	148(52.5)	110(39.0)
	운동 빈도	주3회 미만	177(62.8)	126(44.7)
		주3~4회	63(22.3)	91(32.3)
주5회 이상		42(14.9)	65(23.0)	
영양습관***	식사 시간	10분 이내	90(31.9)	59(20.9)
		10~20분	144(51.1)	160(56.7)
		20분 이상	48(17.0)	63(22.3)
			282(100)	282(100)

***p<.001, **p<.01, *p<.05

3. 대사증후군의 항목별 변화

1) 전체 이용자의 변화

스마트건강증진서비스를 받은 이용자는 대사증후군 각 측정 항목별 서비스 이용 전과 이용 후, 유의미한 차이가 있는지 살펴보기 위해 대사증후군 관련 항목의 전후 평균차이를 Paired t-test로 분석하였다.

대사증후군의 7가지 측정 항목 모두 통계적으로 유의미한 것으로 분석되었으며,

서비스 이용 전보다 이용 후 체질량, 허리둘레, 수축기혈압, 이완기혈압, 공복혈당, 중성지방은 감소하였고, HDL-콜레스테롤은 증가하여 대사증후군의 긍정적인 변화를 보였다. 따라서 전반적으로 대사증후군관리서비스의 효과성이 나타난 것으로 확인되었다.

〈표 3〉 서비스 제공 전·후 대사증후군 변화(n=282)

측정 항목	전체				
	Pre-test M(SD)	Post-test M(SD)	Difference M(SD)	t	p
체질량(kg/m ²)	25.10(3.21)	24.87(3.27)	.23(1.26)	2.973	.003**
허리둘레(cm)	86.50(8.50)	85.32(8.52)	1.18(5.10)	3.125	.002**
수축기혈압(mmHg)	129.95(17.02)	124.74(13.81)	5.20(15.01)	3.887	.000***
이완기혈압(mmHg)	80.60(10.59)	78.25(9.42)	2.34(10.97)	5.826	.000***
공복혈당(mg/dl)	103.95(21.76)	100.62(15.19)	3.32(18.76)	3.586	.000***
중성지방(mg/dl)	150.59(92.60)	137.23(69.67)	13.36(87.58)	-6.434	.000***
HDL-C(mg/dl)	48.33(15.06)	53.30(12.93)	-4.97(12.98)	2.562	.011*

*p<.05, **p<.01, ***p<.001

2) 연령대에 따른 변화

스마트건강증진서비스의 효과는 연령대별 차이가 있는지 확인하기 위해 실험집단을 39세 이하, 40~60세, 61세 이상 등 세 집단으로 구분하여 대사증후군 측정 항목별 서비스 이용 전과 후의 차이를 Paired t-test를 통해 분석하였다.

먼저 39세 이하의 경우 허리둘레, 수축기혈압, 이완기혈압, HDL-콜레스테롤 등 4가지 항목에서 통계적으로 유의미한 변화가 있었고, 40~60세의 경우, 체질량, 허리둘레, 수축기혈압, 이완기혈압, 중성지방, HDL-콜레스테롤 등 6가지 항목에서 유의미한 변화가 나타났다. 61세 이상의 경우 체질량, 수축기혈압, 공복혈당 등 3가지 항목에서 유의미한 변화가 나타났다.

결과적으로 39세 이하의 경우, 허리둘레, 수축기혈압, 이완기혈압이 감소하였고,

HDL-콜레스테롤이 증가하여 대사증후군의 긍정적인 변화가 있었다. 40~60세의 경우 체질량, 허리둘레, 수축기혈압, 이완기혈압, 중성지방이 감소하였고, HDL-콜레스테롤이 증가하여 긍정적인 변화를 보였다. 61세 이상의 경우 체질량, 수축기혈압, 공복혈당이 감소하여 마찬가지로 긍정적인 변화를 보였다. 따라서 연령대별로 통계적으로 유의미한 항목에 차이가 있는 것으로 나타났고, 40~60세 연령대가 39세 이하 및 61세 이상의 연령대보다 서비스 효과성이 가장 크게 작용한 것으로 나타났다.

〈표 4〉 서비스 제공 전·후 연령별(39세 이하) 대사증후군 변화(n=49)

측정 항목	전체				
	Pre-test M(SD)	Post-test M(SD)	Difference M(SD)	t	p
체질량(kg/m ²)	25.23(3.92)	24.94(4.25)	.28(1.64)	1.225	.226
허리둘레(cm)	87.58(9.54)	85.08(10.66)	2.49(5.16)	3.385	.001**
수축기혈압(mmHg)	121.95(15.38)	116.85(12.88)	5.10(14.96)	2.387	.021*
이완기혈압(mmHg)	79.40(9.72)	73.77(10.50)	5.63(12.48)	3.157	.003**
공복혈당(mg/dl)	99.28(16.70)	99.08(15.24)	.20(17.94)	.080	.937
중성지방(mg/dl)	141.67(88.22)	132.32(44.72)	9.34(82.54)	.793	.432
HDL-C(mg/dl)	45.16(14.14)	54.08(12.02)	-8.91(14.45)	-4.319	.000***

*p<.05, **p<.01, ***p<.001

〈표 5〉 서비스 제공 전·후 연령별(40~60세) 대사증후군 변화(n=103)

측정 항목	전체				
	Pre-test M(SD)	Post-test M(SD)	Difference M(SD)	t	p
체질량(kg/m ²)	25.50(3.43)	25.23(3.35)	.21(1.06)	2.062	.042*
허리둘레(cm)	86.75(8.60)	85.51(8.37)	1.24(4.64)	2.709	.008**
수축기혈압(mmHg)	129.67(17.57)	123.95(14.89)	5.72(11.76)	4.940	.000***
이완기혈압(mmHg)	82.86(11.11)	79.16(9.02)	3.69(9.49)	3.952	.000***
공복혈당(mg/dl)	104.22(21.19)	101.99(17.56)	2.23(16.09)	1.408	.162
중성지방(mg/dl)	171.43(113.39)	147.77(73.05)	23.66(95.91)	2.503	.014*
HDL-C(mg/dl)	44.77(14.51)	51.81(12.89)	-7.03(12.57)	-5.680	.000***

*p<.05, **p<.01, ***p<.001

〈표 6〉 서비스 제공 전·후 연령별(61세 이상) 대사증후군 변화(n=130)

측정 항목	전체				
	Pre-test M(SD)	Post-test M(SD)	Difference M(SD)	t	p
체질량(kg/m ²)	24.74(2.68)	24.51(2.73)	0.23(1.26)	2.094	.038*
허리둘레(cm)	85.91(8.01)	85.26(7.77)	0.64(5.37)	1.359	.176
수축기혈압(mmHg)	133.18(16.24)	128.34(11.89)	4.83(17.26)	3.196	.002**
이완기혈압(mmHg)	79.26(10.25)	79.23(8.87)	0.03(11.05)	.032	.975
공복혈당(mg/dl)	105.48(23.70)	100.13(13.03)	5.36(20.82)	2.935	.004**
중성지방(mg/dl)	137.45(71.17)	130.73(73.95)	6.72(82.25)	.932	.353
HDL-C(mg/dl)	52.34(14.94)	54.20(13.28)	-1.85(12.05)	-1.753	.082

*p<.05, **p<.01, ***p<.001

3) 대사증후군 진단 기준에 따른 변화

결과적으로 서비스 이용을 통해 대사증후군 유소견자가 서비스 이용 전보다 서비스 이용 후 얼마나 감소했는지 검증하기 위해 서비스 이용 전·후의 유소견자 수를 분석하였고, 그 감소율을 함께 분석하였다.

대사증후군 유소견자의 측정 항목별 진단 변화를 살펴보면, 전체 집단의 경우 최소 1.5% ~ 최대 45.0%로 나타났다. 연령대별로는 39세 이하의 경우 최소 8.3% ~ 최대 83.3%의 감소율을 보였고, 40~60세는 최소 8.9% ~ 최대 41.9%, 61세 이상은 최소 7.1% ~ 최대 42.3%의 감소율을 보였다. 연령대별로 비교해보면, 체질량, 허리둘레, 이완기혈압, HDL-콜레스테롤은 39세 이하에서 유소견자 감소율이 높은 반면 수축기혈압, 중성지방은 40-60세인 경우에 감소율이 높았고, 공복혈당은 61세 이상에서 감소율이 높은 것으로 나타났다.

〈표 7〉 서비스 제공 전·후 대사증후군 변화

(n=282, 39세 이하 49명, 40~60세 103명, 61세 이상 130명)

측정항목 구분	체질량		허리둘레		수축기혈압		이완기혈압		공복혈당		중성지방		HDL-C		
	사전	사후(%)*	사전	사후(%)*	사전	사후(%)*	사전	사후(%)*	사전	사후(%)*	사전	사후(%)*	사전	사후(%)*	
전체	132	130(1.5)	135	119(11.9)	141	101(28.4)	100	55(45.0)	139	122(12.2)	103	92(10.7)	131	77(41.2)	
연령대	39세 이하	24	22(8.3)	24	20(16.6)	11	8(27.3)	18	3(83.3)	16	16(0.0)	18	16(11.1)	26	14(46.1)
	40~60세	54	54(0.0)	56	48(14.3)	51	32(37.2)	43	25(41.9)	56	51(8.9)	43	37(13.9)	53	33(37.7)
	61세 이상	54	54(0.0)	55	51(7.3)	79	61(22.8)	39	27(30.8)	67	55(17.9)	42	39(7.1)	52	30(42.3)

*유소견자 감소율

4) 대사증후군 유소견자의 변화

대사증후군 측정 항목별로 유소견자만 추출하여 서비스 이용 전과 이용 후에 유의미한 차이가 있는지 살펴보기 위해 Paired t-test로 분석하였다.

분석결과, 체질량, 허리둘레, 수축기혈압, 이완기혈압, 공복혈당, 중성지방, HDL-콜레스테롤 등 측정 항목 모두에서 통계적으로 유의미한 것으로 분석되었다. 즉, 유소견자는 서비스 이용 전보다 이용 후 체질량, 허리둘레, 수축기혈압, 이완기혈압, 공복혈당, 중성지방은 감소하였고, HDL-콜레스테롤은 증가하여 대사증후군의 긍정적인 변화를 보였다. 따라서 전반적으로 스마트건강증진서비스의 효과성이 나타난 것으로 확인되었다.

〈표 8〉 서비스 제공 전·후 유소견자의 대사증후군 변화(n=282)

측정 항목	전체				
	Pre-test M(SD)	Post-test M(SD)	Difference M(SD)	t	p
체질량(kg/m ²)	27.77(2.41)	27.37(2.51)	.40(1.13)	4.057	.000***
허리둘레(cm)	89.56(9.24)	87.34(9.42)	2.22(5.14)	5.011	.000***
수축기혈압(mmHg)	142.99(12.38)	130.55(12.69)	12.43(14.59)	10.122	.000***
이완기혈압(mmHg)	91.63(6.13)	82.00(10.11)	9.63(10.53)	9.143	.000***
공복혈당(mg/dl)	116.67(24.64)	105.85(18.07)	10.82(22.03)	5.789	.000***
중성지방(mg/dl)	238.38(101.19)	175.31(83.25)	63.07(114.38)	5.597	.000***
HDL-C(mg/dl)	42.96(14.68)	51.40(13.96)	-8.43(13.35)	-7.228	.000***

*p<.05, **p<.01, ***p<.001

IV. 결론 및 제언

본 연구는 최근 뇌혈관 질환, 당뇨, 고혈압 등 대사증후군 관련 질환의 적극적 건강관리 및 사전 예방적 필요성이 대두됨에 따라 서울시 지역사회서비스투자사업에서 수행되고 있는 U-Health 시스템을 활용한 스마트건강증진서비스 효과성을 분석하였다.

분석결과, 연구대상인 282명을 대상으로 서비스 이용 전과 후 체질량, 허리둘레, 수축기/이완기혈압, 공복혈당, 중성지방은 통계적으로 유의미하게 감소하였고, HDL-콜레스테롤은 증가하는 것으로 나타났다. 이를 연령대별로 살펴보면, 39세 이하의 경우 허리둘레, 수축기혈압, 이완기혈압은 감소하였고 HDL-콜레스테롤은 증가하여 총 4가지 대사증후군 측정 항목에서 통계적으로 유의미하였다. 40~60세의 경우 체질량, 허리둘레, 수축기/이완기혈압, 중성지방은 감소하였고, HDL-콜레스테롤은 증가하여 총 6가지 항목에서 유의미하였다. 61세 이상인 경우 체질량, 수축기혈압, 공복혈

당이 유의미하게 감소하였다. 즉, 전반적으로 스마트건강증진서비스는 효과성이 나타났고, 특히 40~60세 이용자에게 서비스 효과가 가장 크게 작용한 것을 확인할 수 있다.

또한 대사증후군 진단 기준에 따라 유소전자 수와 감소율을 살펴보면, 측정항목별로 유소전자 수의 감소는 39세 이하에서 가장 많았고, 다음으로 40~60세, 61세 이상 순으로 분석되었다. 유소전자만을 대상으로 대사증후군 변화를 살펴보면, 대사증후군 7가지 항목 모두에서 통계적으로 유의미하였는데, 체질량, 허리둘레, 수축기/이완기 혈압, 공복혈당, 중성지방은 감소하였고, HDL-콜레스테롤은 증가한 것으로 나타났다. 즉, 유소전자를 대상으로 한 분석에서도 서비스 효과성이 나타났음을 확인할 수 있다.

이러한 분석결과를 통해 향후 긍정적 건강생활습관 실천의 스마트건강증진서비스를 위한 정책적 제언을 하면 다음과 같다.

첫째, 스마트건강증진서비스 이용 대상의 확대가 필요하다. 본 연구결과에 따르면 대사증후군 전체 이용자를 대상으로 한 분석에서도, 유소전자만을 대상으로한 분석에서도 통계적으로 유의미하였으며, 긍정적인 변화를 보여 서비스 효과성이 나타난 것으로 확인되었다. 이에 따라 서울시민이 스마트건강증진서비스를 보다 많이, 적극적으로 이용할 수 있도록 예산 규모 및 소득기준 확대 등 방안의 검토가 필요할 것으로 판단된다.

둘째, 서비스 이용자의 인구학적 특성에 따른 서비스 지원방안이 요구된다. 본 연구의 연령대별 대사증후군 분석결과 40~60세의 중장년층에서 서비스 효과성이 가장 높게 나타났다. 2012년 건강검진통계연보에 따르면, 30대는 다른 연령대에 비해 흡연율 및 비만율이 높고, 가장 신체활동이 적은 연령대인 것으로 조사되었다(국민건강보험공단, 2013). 이는 30대의 경우 고된 직장생활로 인해 운동할 여유를 갖지 못하고, 육류위주 섭취, 잦은 음주 등 부정적 식습관을 갖고 있기 때문에 비만이나 고혈압, 당뇨 등 만성질환에 대한 심각성을 느끼지 못하는 것으로 판단된다. 또한 61세 이상 노인층에서는 신체적 노화로 인해 신체회복율이 낮기 때문에 서비스 효과가 가장 낮게 나타난 것으로 예측할 수 있다. 반면 40~60대의 경우 신체활동 실천율이 가장 높고, 중년기 진입에 따라 건강 관심이 증가되면서 건강검진 이용빈도 및 종류가 증가된다는 점에서 서비스 효과가 가장 크게 나타난 것으로 예측할 수 있다. 따라서

효과적인 대사증후군 건강관리서비스를 위해 인구학적 특성에 따른 맞춤형 건강관리를 위한 지원이 필요할 것으로 판단된다. 예컨대 지리적 접근이 어려운 농촌 지역주민, 보건소나 병원 방문이 어려운 바쁜 직장인, 거동이 불편한 노인, 장애인 등 대사증후군 관리 사각지대에 놓인 이용자를 대상으로 맞춤형 건강상담 및 교육, 관리 등 서비스 접근방법 및 내용의 차별성을 통해 서비스 효과를 극대화할 수 있을 것이다. 이는 향후 발생할 수 있는 사회적 비용, 의료비 및 건강보험 비용 등의 절감효과를 기대할 수 있을 것으로 판단된다.

셋째, 스마트 기기를 활용한 대사증후군관리가 필요하다. 앞서 살펴봄과 같이 스마트건강증진서비스 효과성은 40~60대의 경제활동인구의 연령대에서 집중적으로 긍정적인 변화가 나타났다. 대부분의 직장인들은 바쁜 업무로 인해 병원에 내원할 시간적 여유가 없고, 개인별 영양상담 및 교육 제공이 어려울 뿐만 아니라 지속적인 건강관리가 어려워 대사증후군 등 만성질환 위험에 노출되어 있다(박세윤, 양윤정, 김유리, 2011). 이러한 측면에서 스마트건강증진서비스는 인터넷 기반의 개인 맞춤형 건강관리서비스로서 언제, 어디서든 건강측정이 가능하고, 찾아가는 서비스를 통해 서비스 접근이 용이하며 대면, 인터넷, SMS 등 다양한 방법을 활용하여 맞춤형 건강상담 및 정보가 제공되기 때문에 바쁜 직장인들을 대상으로 적절히 활용한다면 건강관리의 효율성을 높일 수 있을 것으로 판단된다.

넷째, 대사증후군 유소견자의 긍정적 변화에 대한 적극적 지원이 필요하다. 대사증후군 진단 기준에 따른 유소견자 변화를 살펴보면, 서비스 이용 전보다 이용 후 대사증후군 7가지 항목 모두에서 긍정적으로 변화한 것으로 나타났다. 따라서 이들에 대한 적극적인 지원을 통해 장기적인 건강생활을 실천할 수 있도록 하는 장치 마련이 필요하다. 예컨대 국민건강보험공단은 금연치료 지원사업을 병의원·약국과 연계하여 금연 치료를 희망하는 국민에게 상담, 치료서비스, 금연보조제 등 치료비용의 상당 부분을 지원하고 있으며, 프로그램 이수자에게는 인센티브를 지원하고 있다(금연길라잡이, 2017). 이처럼 대사증후군 기존의 유소견자에서 무소견자로 달성 시 소정의 인센티브를 지원하는 방안을 고려해볼 수 있다.(보건복지부·한국건강증진개발원 국가금연지원센터, 2017).

다섯째, 서울시의 종합적 대사증후군관리서비스 정책이 필요하다. 서울시의 경우

2009년부터 서울시민들의 심뇌혈관, 당뇨, 암 등 질병들로부터 경제적 부담 및 고통에 대해 적극적 대처하고자 질병 예방사업으로 대사증후군관리사업을 보건소에서 시행하고 있다(대사증후군 오락프로젝트, 2017). 보건소에서 시행 중인 대사증후군관리사업은 보건소에 직접 내소하는 시민을 대상으로 1년의 서비스 기간 동안 무료로 서비스를 제공하고 있으며, 건강 변화에 따라 이용기간의 연장이 가능하다. 반면 지역사회서비스투자사업의 스마트건강증진서비스는 소득 및 욕구기준을 충족한 시민을 대상으로 서비스 비용의 10~20%는 이용자가 부담하여 최대 1년간 서비스를 이용할 수 있으며, 건강측정기기 제공을 통해 시공간 제약 없이 자가측정이 가능하고 평생 이용이 가능하다. 또한 이용자의 건강측정 결과를 상시 제공인력이 모니터링 하며, 찾아가는 전문상담 및 맞춤형 건강정보가 제공된다. 단, 서비스 기간이 종료되면 관리대상자에서 제외된다. 따라서 두 사업 간의 연계 및 조정을 통한 서울시 사업의 효율성 강화가 모색되어야 한다. 스마트건강증진서비스는 이용기간이 종료됨에 따라 관리대상자에서도 제외되기 때문에 건강예방의 연속성을 위해 스마트건강증진서비스 종료시 보건소에서 시행 중인 대사증후군관리사업과 연계하여 대사증후군 유소견 이용자의 추적관리와 서울시민의 건강개선을 위해 노력해야 할 것이다.

본 연구는 서울시 지역사회서비스투자사업 중 대사증후군관리서비스 이용자의 일반적 특성, 생활습관 현황, U-health 서비스를 통한 효과성과 연령을 구분한 서비스 효과성을 분석하였다. 이를 통해 2011년 시범사업 결과에 따른 서비스 효과성 측정 연구(송태민, 2011) 이외 사업이 본격화 된 후 대사증후군관리서비스의 효과성을 측정하여 긍정적인 연구결과를 확인하였다는 점에서 큰 의의가 있다. 그럼에도 불구하고 본 연구는 다음과 같은 연구의 제한점을 갖는다. 서비스 이용 전과 이용 후의 차이가 실험처치(experimental treatments)의 작용으로 인한 차이인지, 이용자의 개인적 환경이나 약물복용으로 인한 생리적 위험요인의 감소 따른 차이인지에 대한 내적타당도를 확보하지 못하였다. 또한 사업의 효과성을 측정하기 위해서는 서비스 기준정보에 따른 필수 서비스 제공 이외 자가 측정을 포함한 서비스 이용 빈도 및 시간과 추가 상담빈도 수 등 이용자의 자발적 노력에 의한 서비스 이용과 관련된 정보가 필수적이거나, 자료의 한계로 이를 연구에 반영하지 못하였다. 따라서 제한점을 해결하기 위해 이용자 인터뷰 및 U-health 기기 사용 빈도 조사를 통해 연구의 보완이 필요할 것이

다. 마지막으로 서비스 종결 이후 서비스 효과의 지속성 및 중단연구를 통해 보다 명확한 효과성 검증이 필요하리라 판단된다.

참고문헌

- 강상경·전해숙. 2013 “연령대별 건강궤적과 예측요인에 대한 탐색적 연구.” 『사회과학연구』 제24권 제3호: 61-84.
- 강창현·문순영·김기창. 2012. “바우처 제도의 현황과 효과성 분석.” 국회예산정책처. 국민건강보험공단. 2013. 『2012년 건강검진통계연보』 .
- 국민건강보험공단. 2017. 『2016년 건강검진통계연보』 .
- 김선명. 2009. “조직 성격의 차이에 따른 지역사회서비스사업의 비교 분석: S시·K시의 ‘장애아동의 교육 및 치료지원 사업’을 중심으로.” 『한국조직학회보』 제6권 제3호: 115-142.
- 김선애. 2004. “DEA를 이용한 대학도서관의 효율성 평가: 프랑스의 대학도서관을 대상으로.” 『한국문헌정보학회 학술발표논문집』 제17집: 7-32.
- 김신미·김순이. 2008. “노인의 노화불안과 삶의 질.” 『정신간호학회지』 제17권 제3호: 281-291.
- 김용득. 2008. “사회서비스 정책의 동향과 대안-시장 기제와 반-시장 기제의 통합.” 『사회복지연구』 36호: 5-28.
- 김은정. 2008. “사회서비스 재정지원방식의 변화와 품질관리를 위한 정책 과제.” 『사회복지정책』 35호: 141-168.
- 김은정·정은아. 2012. “지역사회서비스투자사업 이용자의 서비스 효과성 인식과 관련 요인: 대구시 문제아동조기개입서비스를 중심으로.” 『사회서비스연구』 제3권 제1호: 67-94.
- 김현제·조성한. 2006. “DEA를 이용한 서울시 자치구별 교육서비스의 효율성 측정.” 『경영교육논총』 제42집: 189-205.
- 김희선·윤석준·이건세·김형수·오상우·유호신·추진아·이계희·유병욱·이동욱·박동림. 2011. “지역사회 대상의 대사증후군 자가관리 프로그램 중재 효과 - 서울특별시 대사증후군 관리 시범사업.” 『보건교육·건강증진학회지』 제28권 제2호: 51-62.
- 김희선·윤석준·이건세·김형수·오상우·유호신·추진아·이계희·유병욱·이동욱·박동림. 201

1. “지역사회 대상의 대사증후군 자가관리 프로그램 중재 효과 - 서울특별시 대사증후군 관리 시범사업.” 『보건교육·건강증진학회지』 제28권 제2호: 51-62.
- 남궁근·하혜수. 2004. “DEA에 의한 공공서비스 투입자원 재배분방안 연구.” 『한국조직학회보』 제1권 제1호: 1-24.
- 대한비만학회. 2012. 『비만치료 지침』. 서울: 청운.
- 박경효. 1992. “공공서비스의 민간화에 대한 평가.” 『한국행정학보』 제25권 제4호: 459-479.
- 박대권. 2009. “미국의 학교선택제.” 『교육정책포럼』 197(통권): 16-21.
- 박정선. 2012. “u-Health 정책현황과 향후 추진 방향.” 『과학기술정책』 188: 21-29.
- 보건복지부 질병관리본부. 2016. “2015 국민건강통계 I: 국민건강영양조사 제6기 3차년도.”
- 보건복지부. 2016. 2016년 지역사회서비스투자사업안내.
- 보건복지부. 2017. 2017년 지역사회서비스투자사업안내.
- 보건복지부·한국건강증진개발원 국가금연지원센터. 『2017 지역금연 민간보조사업 안내』.
- 서울시지역사회서비스지원단. 2016. 서울시 지역사회서비스투자사업 표준안.
- 서재룡·배상수. 2011. “일개 보건소 대사증후군 관리프로그램의 효과분석.” 『농촌의학·지역보건』 제36권 제4호: 264-279.
- 송태민. 2011. “u-Health 서비스 효과 분석: 대사증후군 관리서비스를 중심으로.” 『보건·복지 Issue&Focus』 79권: 1-8.
- 아시아투데이. 100세 시대 위협하는 대사증후군… 남자 ‘혈압’·여자 ‘콜레스테롤’ 관리(2016년 6월 16일).
- 양기용. 2013. “사회서비스 공급체계변화와 공공성.” 『한국공공관리학보』 27(1): 89-114.
- 이애경·이상아·강임옥·정백근·한준태·박일수. 2007. “건강검진 사후관리 서비스의 건강증진 효과 - 건강주의자를 대상으로.” 『보건교육·건강증진학회지』 24(1): 127-138.

- 이영익. 2013. “사회서비스 바우처 프로그램의 효과성 및 이용만족에 대한 영향요인: 스포츠바우처를 중심으로.” 『한국체육정책학회지』 11(1): 31-43.
- 이은희·김혜경·이윤희·문순영·권은주·지선하. 2007. “생활습관 중재프로그램이 성인의 건강생활습관실천과 대사증후군 개선에 미치는 효과 분석.” 『보건교육·건강증진학회지』 24(9): 1-19.
- 이재원. 2011. “사회서비스 산업 육성을 위한 법적 기반 확충 과제: 사회서비스 산업 진흥법 제정의 필요성과 과제.” 국회사회서비스포럼·한국사회서비스학회 공동 정책 세미나 자료집. 15-43.
- 정소연·김은정. 2009. “노인돌보미 바우처서비스 효과성과 영향 요인: 노인부양가족 생활상의 긍정적 변화를 중심으로.” 『가족과 문화』 21(1): 117-149.
- 지재환·성지동·최윤희·정재훈·민용기·이명식·김광원·이문규. 2006. “한국인에서 다양한 진단기준에 따른 대사증후군의 유병률과 인슐린 저항성 반영의 차이.” 『한국 지질·동맥경화학회지』 16(2): 299-308.
- 통계청. 2016. “2015년 사망원인통계.”
- 현순화·이규영. 2012. “서울시 일개 지역 중년여성 대상 ‘여성건강리더’교육프로그램의 효과.” 『보건사회연구』 32(4): 551-576.
- 홍현주. 2004. 『조기유학 알고 보내자』. 서울: 넥서스BOOKS.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ford, E. S. 2004. “Prevalence of the metabolic syndrome in US populations.” *Endocrinology and metabolism clinics of North America* 33(2): 333-350.
- Ford, E. S., Giles, W. H., & Mokdad, A. H. 2004. “Increasing prevalence of the metabolic syndrome among US adults.” *Diabetes Care* 27(10): 2444-2449.
- Witte, J. F. 2016. *Evaluating Voucher Programs: The Milwaukee Parental Choice Program*. Society for Research on Educational Effectiveness.

A Study on Measuring Effectiveness of metabolic syndrome management using U-Health system : Focusing on Seoul city

Su Kyung Oh* · Chang Hwan Shin**

ABSTRACT

In this study, 282 adults aged 19 and over who participated in the smart health promotion service of community service investment project in 7 boroughs in Seoul were studied. Service effectiveness was measured by the changes in the metabolic syndrome before and after the service, and the changes in the metabolic syndrome by age. In addition, we investigated the present condition of the patients with metabolic syndrome according to the diagnostic criteria and examined the changes of metabolic syndrome. Metabolic syndrome was measured by body mass, waist circumference, systolic blood pressure, diastolic blood pressure, fasting blood glucose, triglyceride, and HDL-cholesterol. The main results of this study are as follows. First, the change in metabolic syndrome was more significant than before service, and the body mass, waist circumference, systolic blood pressure, diastolic blood pressure, fasting blood glucose, and triglyceride were decreased, and the metabolic syndrome was positively changed by increasing HDL-cholesterol. Second, there was a difference in the metabolic syndrome by age group, and it was analyzed that the service effect was the highest in the age group of 40-60 years old. Third, the rate of decrease of metabolic syndrome was 1.5%-83.3%. The rate of decrease was different according to the total population and age group. Fourth, only the metabolic syndrome cases showed positive results in all 7 items. Based on the results of this study, we propose extension of service utilization, provision of discriminative service according to demographic characteristics, support for additional services for the suspected persons, and linkage or duplication adjustment with the public health center business with the policy of managing comprehensive metabolic syndrome in Seoul.

Keywords : Social Service, Voucher, Measuring Effectiveness, Metabolic Syndrome, U-Health system

투고일: 2018.03.07. **심사일:** 2018.04.10. **게재확정일:** 2018.04.23.

* Kyungpook National University

** Ewha Womans University, Graduate School of Social Welfare, Master of Social Welfare

[DOI] <http://dx.doi.org/10.21487/jrm.2018.5.3.1.79>

【연구경향】

Varimax Rotation and Thereafter: Tutorial on PCA Using Linear Algebra, Visualization, and Python Programming for R and Q analysis

Byung Lee*

ABSTRACT

Principal component analysis (PCA) has been heavily used for both academic and practical purposes. This tutorial would help individuals who want to better utilize PCA as well as R scholars interested in Q analysis. Many statistical commercial packages can handle PCA, but one may reach a deeper understanding by running PCA code for oneself. Cranking out results using PCA features on one of the commercial statistical packages is easy, just importing data and clicking a few buttons. For a deep understanding of PCA, however, an individual needs to see what happens when running code. Many scholars may have a basic knowledge of a popular programming language like Python. But PCA code in Python is not neatly compiled in one place. A user must gather Python code that is scattered around the Internet, tweak it for compatibility, and fill any remaining gaps. Also, commercial statistical packages, which have been geared toward R analysis, cannot be used for rotation for theoretical or exploratory analysis in Q studies beyond varimax. This tutorial offered all Python code needed for PCA while comparing its results with Statistical Package for the Social Science (SPSS) output. This tutorial also covered the theoretical or exploratory rotation of factor axes, which is a must for Q analysis.

Keywords : PCA, Principal Component Analysis, linear algebra, graphs, Python code, varimax rotation, and R and Q analysis

* Associate Professor, Elon University(byunglee@elon.edu)

I. Introduction to PCA

A) What is PCA?

"A principal component analysis can be considered as a rotation of the axes of the original variable coordinate system to new orthogonal axes, called principal axes, such that the new axes coincide with directions of maximum variation of the original observations" (Campbell & Atchley, 1981, pp. 268).

Items, whether human beings or objects, can be observed by measuring their characteristics or attributes, so-called variables. These variables constitute data dimensions. When graphically displayed, they represent coordinate axes.

If one chooses to have the same number of principal components in PCA as original variables, for example, 10, one can account for all variations that the original ones can. PCA finds principal components in descending order of variations explained. The front components account for more variations than the later ones. The 1st principal component accounts for the maximum amount of variations possible in data, and the 2nd principal component extracts the maximum possible variations in data after excluding what was explained by the 1st component. Extractions can be done until all the variations are accounted for by the last principal component.

Original variables vary together in some pattern. PCA takes advantage of this covariation and accounts for more variations than the original ones in the case of the front-end principal components. If one keeps all principal components, there is no benefit. The benefit arises when it is used for data reduction by discarding the backend principal components, which can be interpreted as noise or small enough to be ignored without causing much loss to the original data. Also, the principal components of PCA are statistically independent of each other.

To understand the central factors that influence political behavior, for

example, a study (Markaki, Chadjipandelis, & Tomaras, 2014) using PCA extracted eight factors out of the original 64, explaining 34% of the total variation. Through PCA, the original dataset was reduced to a new dataset, which is one-eighth of the original amount.

PCA is "a data-reduction technique that transforms a larger number of correlated variables into a much smaller set of uncorrelated variables called principal components" (Kabacoff 2015, p. 319).

B) The purpose of this manuscript

To understand PCA and use it correctly, it requires more than punching numbers into software incorporating the PCA feature and cranking out results. It requires a linear algebra background to understand its theoretical foundation. Abstract concepts in linear algebra can be more easily followed by graphical geometrical displays. The operation of PCA beyond three dimensions can be better understood by actually running computer code on data.

Not much literature exists to simultaneously cover these three areas--linear algebra, visualization, and computer code--at a level that social scientists without much mathematical background are able to understand. Many books on factor analysis have described PCA as part of factor analysis. Since it involves linear algebra, some factor analysis books even include an introduction to matrix operations (Horst 1965; Cureton & D'Agostino, 1983; Gorsuch 1983). Linear algebra books always include a chapter on eigenvalues and eigenvectors, but only recently these books began to include PCA around at their end (Strang 2016; Lay 2005). Popular computer packages like SPSS, SAS, and Stata have PCA features, but how these programs operate under the hood is hidden. Many social scientists are familiar with Python, a general-purpose scripting language, but Python codes for PCA are often scattered around the Internet without much annotation. Python code has been

developed mainly for R analysis (traditional factor analysis in which variables are correlated), not for Q analysis, in which research participants are correlated, and exploratory or theoretical rotations of factor axes beyond varimax rotation are adopted. The commercial statistical programs are not different. The author found that literature rarely cover all three approaches together¹⁾: numerical analysis based on linear algebra, geometrical displays, and Python programming code, with which researchers can dirt their hands with data.

II. Graphical Geometric Explanation

This article will start with a visual description of PCA, the easiest among the three approaches. For a graphical illustration of PCA process, the author created a sample dataset measuring two variables with 30 observations in this example: people's favor for Chocolate and their favor for Candy, as shown below. This dataset is posted at "<http://bestelon.com/pca/pca.csv>".

	A	B	C
1	name	x_value	y_value
2	A	2.7	4.6
3	B	15	15
4	C	23.9	9.5
5	D	14.3	9.3
27	Z	20.4	16.6
28	A1	32.6	17.4
29	B1	28.6	12.7
30	C1	18.1	12.3
31	D1	46.1	7.9

1) When factor analysis began to be developed, some book authors introduced Fortran code along with linear algebra and geometrical illustrations.

A) Import the original dataset

```

1. import pandas as pd
2. filepath = "http://bestelon.com/pca/pca.csv"
3. df = pd.read_csv(filepath)

```

B) Draw a graph based on the dataset

```

1. import numpy as np
2. import matplotlib.pyplot as plt
3. import random
4. labels = df["name"]
5.
6. random.seed( 5 )
7. randColor = np.random.random(30)
8.
9. # to set the figure size explicitly
10. plt.figure(figsize=(6, 6))
11.
12. plt.scatter( df.iloc[:, 1], df.iloc[:, 2], marker='o',
               c=randColor , s = 100, cmap=plt.get_cmap('Spectral') )
13.
14. #c for color or sequence of color; s for size
15. # marker for marker style
16. #spectral for the black-purple-blue-green-yellow-red-
    white spectrum
17. # refer to https://matplotlib.org/api/\_as\_gen/matplotlib.pyplot.scatter.html
18.
19. for label, x, y in zip(df.iloc[:, 0], df.iloc[:, 1],
                       df.iloc[:, 2]):
20.     plt.annotate(
21.         label,
22.         xy=(x, y), xytext=(5, 5),
23.         textcoords='offset points', ha='right', va='bottom')
24.
25. plt.xlim(0,60)
26. plt.ylim(0, 60)
27. plt.xlabel('Favor for Chocolate')
28. plt.ylabel('Favor for Candy')
29. plt.title("Figure 1. People's favor for Chocolate vs.
    Candy",fontsize = 16)
30. plt.show()

```

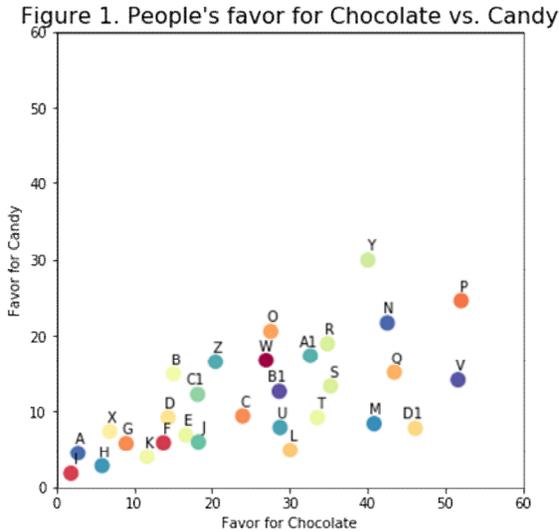
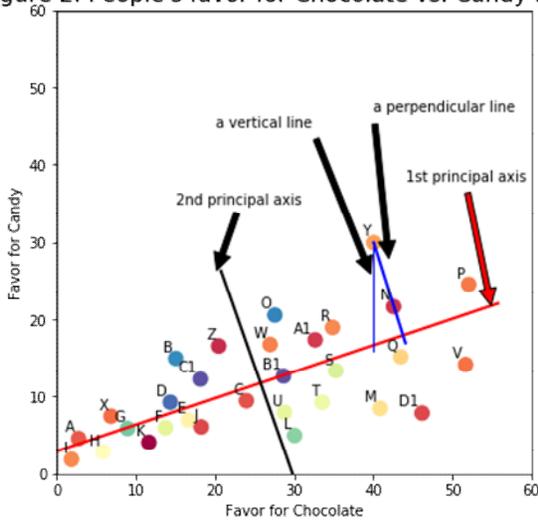


Figure 1 shows 30 people in a coordinate system who favor chocolate on the x-axis and who favor candy on the y-axis. Here one can see a positive correlation between people's favor for chocolate and favor for candy. Rather than measuring the people's favor twice to understand their taste, a researcher can measure favor for sweets at once with a straight line that goes through the middle of all dots.

The straight line should be placed so that the distance between each dot and the line is the shortest, more specifically when the sum of the squared distance for each dot is the smallest. The distance from the line here indicates the perpendicular distance, not the vertical distance (refer to Figure 2).

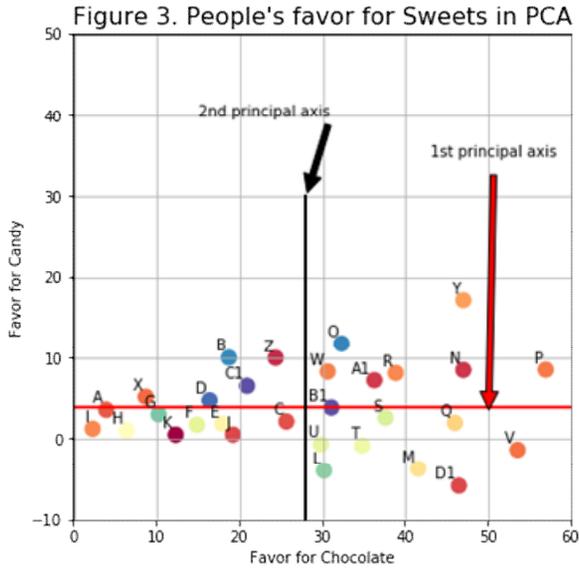
Then the line becomes a new axis, called the 1st principal component axis, and another line that is perpendicular to the 1st component axis, the shortest horizontally off all dots when measured from the 2nd component axis itself, becomes the 2nd principal component.

Figure 2. People's favor for Chocolate vs. Candy in PCA



C) Perspective change

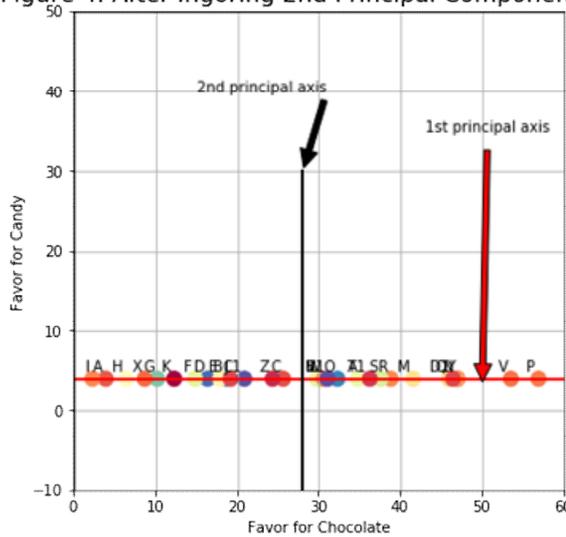
To better understand what the axis change means, let's look at the configuration of dots from a new perspective, meaning under new coordinates. To do that, one can imagine oneself hovering over the graph. Let's rotate oneself by about 15 degrees counterclockwise to align the 1st principal axis horizontally and the 2nd principal axis line vertically. They will be a new x-axis and y-axis respectively. Or you can just rotate the graph by 15 degrees clockwise so that the 1st and 2nd principal component axes will become x- and y-axes (refer to Figure 3). PCA allows one to see the data from a new frame of reference.



D) Discard back-end principal components

If one discards the 2nd principal component for data reduction, it means ignoring the data that is accounted for by this principal component. This disregard results in the effect that each data dot in Figure 3 moves to a new position as shown in Figure 4.

Figure 4. After Ingoing 2nd Principal Component Axis



III. Numerical Analysis: Comparison Between Eigendecomposition and SVD

The second approach is based on numerical analysis using linear algebra. Among many methods to find eigenvalues and eigenvectors, this article covered eigendecomposition and SVD.

A) Formula for calculating variance/covariance or correlation

In PCA, one needs to find eigenvalues and eigenvectors. One can use eigendecomposition, which uses the formula of $Ax = \lambda x$. Here A is an $n \times n$ matrix, x is a vector, and λ is a scalar. When a vector, x , is premultiplied by A , it can be a scalar multiple of x . If the x is not a zero vector, then the x becomes an eigenvector and the λ an eigenvalue.

Eigenvectors match the direction of principal component axes. Eigenvalues

and eigenvectors are paired. If an eigenvector pairs with a higher eigenvalue, it matches with a more front-end principal component. For example, if an eigenvector pairs with the highest eigenvalue, it matches the 1st principal component axis.

Since eigendecomposition requires a symmetric matrix, either the covariance or correlation of the dataset is used instead of the original dataset in rectangular format. If variables are measured using the same or similar measures, one can use a variance/covariance matrix, as shown in Equation 1. Covariance and variance can be calculated with equations 2, 3 and 4. To best estimate the variance or covariance of the population data, (n-1) is used with a sample dataset instead of n, the number of observations.

$$\begin{pmatrix} Var(x, x) & Cov(x, y) \\ Cov(y, x) & Var(y, y) \end{pmatrix} \quad \text{Equation 1}$$

$$Cov(x, y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad \text{Equation 2}$$

$$Var(x, x) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x}) \quad \text{Equation 3}$$

$$Var(y, y) = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})(y_i - \bar{y}) \quad \text{Equation 4}$$

If variables are calculated with different measures, one is recommended to use correlations since the measurement unit may influence PCA results. For example, the use of inches or feet to measure heights will make a difference. The formula for calculating the correlation between x and y variables is shown in Equation 5. The correlation between x and x or between y and y can be calculated in the same way by substituting x for y or y for x.

$$\text{Corr}(x, y) = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x)}\sqrt{\text{Var}(y)}} \quad \text{Equation 5}$$

B) Center or standardize data

The original dataset was converted below into dfO (the original 30 x 3 dataset matrix), dfM (a centered dataset by subtracting the mean value from each observation), and dfZ (a standardized dataset). dfO2, dfM2, and dfZ2 represent the numerical portion of dfO, dfM, and dfZ, these datasets' last two columns each.

```

1. dfO = df
2. dfO2 = df.iloc[:, 1:]
3. dfM2 = dfO2 - dfO2.mean()
4. dfZ2 = (dfO2 - dfO2.mean()) / dfO2.std()
5. dfM = dfM2.copy()
6. dfM.insert(loc = 0, column = "name", value = dfO.iloc[:, 0])
7. dfZ = dfZ2.copy()
8. dfZ.insert(loc = 0, column = "name", value = dfO.iloc[:, 0])

```

C) Eigendecomposition with variances/covariances

Variance and covariance can be calculated using either the original dataset or centered dataset.

```

1. dfO2Cov = dfO2.cov()
2. print(dfO2Cov)
3. dfM2Cov = dfM2.cov()
4. print(dfM2Cov)

```

The covariance was plugged into the np.linalg.eig function to produce eigenvalues and eigenvectors. The resulting eigenvalues and their corresponding eigenvectors are sorted in descending order below:

```

1. dfM2Cov = dfM2.cov()
2. eValM2Cov, eVecM2Cov = np.linalg.eig(dfM2Cov)
3.
4. #sort results based on eigenvalues
5. idxM2Cov = eValM2Cov.argsort()[::-1]
6. eValM2Cov, eVecM2Cov = eValM2Cov[idxM2Cov], eVecM2Cov[:,
   idxM2Cov]
7. print(eValM2Cov.round(3))
8. print(eVecM2Cov.round(3))

```

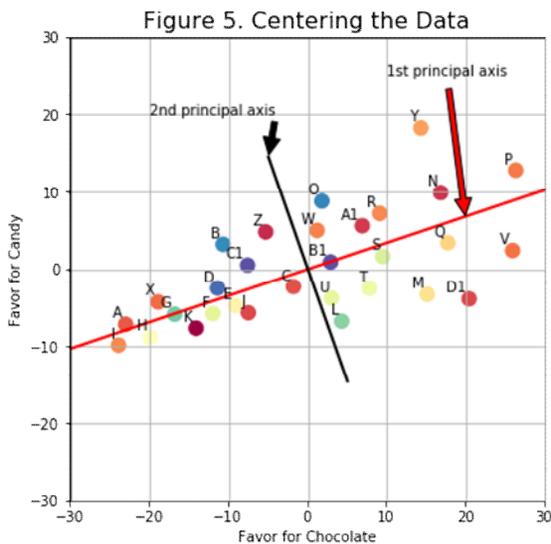
The first line in the following code indicates two eigenvalues, while lines 2 and 3 indicate two eigenvectors:

```

1. [ 231.35  25.482]
2. [[ 0.946 -0.325]
3. [ 0.325  0.946]]

```

If centered values are plotted and lines are drawn based on the eigenvectors above, one can see exactly the same graph, as seen in figures 1 and 2, except for a movement of the center point by the mean of x and y variables (refer to Figure 5).



D) Eigendecomposition with correlations

Correlations among variables can be calculated using any of the following four different ways: the covariance of the standardized dataset, or the correlation of the original dataset, the centered dataset, or the standardized dataset.

```
1. dfZ2Cov = dfZ2.cov()
2. dfO2Corr = dfO2.corr()
3. dfM2Corr = dfM2.corr()
4. dfZ2Corr = dfZ2.corr()
```

A matrix of correlations was plugged into the `np.linalg.eig` function to calculate eigendecomposition:

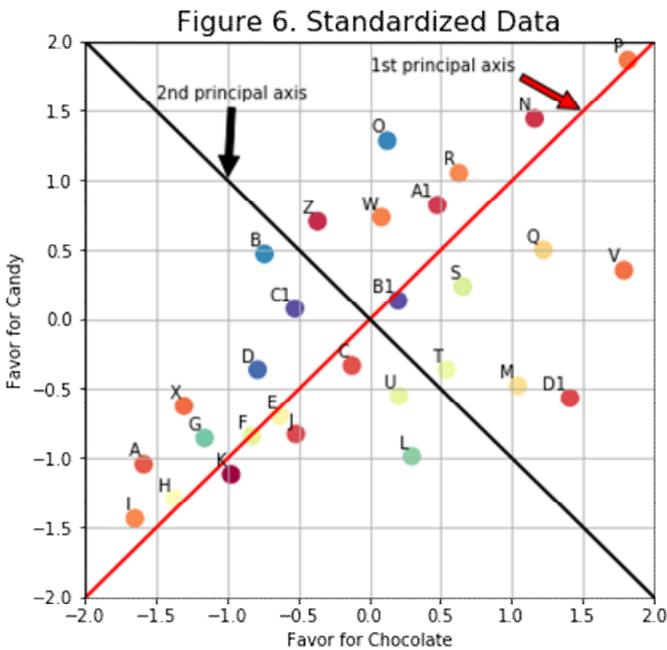
```
1. covZ = dfZ.corr()
2. eValZ1, eVecZ1 = np.linalg.eig(covZ)
3.
4. #sort results based on eigenvalues
5. idxZ = eValZ1.argsort()[::-1]
6. eValZ1, eVecZ1 = eValZ1[idxZ], eVecZ1[:, idxZ]
7. print(eValZ1.round(3))
8. print(eVecZ1.round(3))
```

The result below differs from what the eigendecomposition produced with variances/covariances earlier:

```
1. [ 1.636  0.364]
2. [[ 0.707 -0.707]
3. [ 0.707  0.707]]
```

The data standardization changes the variation of x and y values. Now x and y variables have the same amount of variance, 1. In Figure 6, one can

see the data dots are more widely scattered vertically than they were. It strengthens the influence of variables that were weak relative to stronger ones. Now the 1st principal component is only about 4.5 times as strong as the 2nd principal component. Before the standardization, it was 9.1 times.



E) Singular value decomposition (SVD) with a symmetric data matrix

Among the many ways to decompose a matrix, SVD is preferred by researchers because it always decomposes a matrix, while eigendecomposition or other methods may fail to do it. Eigendecomposition has to use a symmetric matrix while SVD can use any rectangular matrix including a square form. SVD produces eigenvalues and eigenvectors with a dataset in a symmetric

form, which is created by premultiplying a dataset with a transpose of itself like $M.T.dot(M)$. Here M indicates a data matrix, T represents a transpose of the matrix, and "dot" means an inner product of the two matrices. If one divides this product of the two matrices by $n-1$ (n representing the number of observation), it will be the same as a variance/covariance matrix shown in Equation 1 earlier.

```

1. n = 30
2. import numpy as np
3. covMat2 = (1/ (n-1)) *dfM2.T.dot(dfM2)
4. U, S, V = np.linalg.svd(covMat2)
5. print(S)
6. print(V)

```

This code above produced the same result as the eigendecomposition method in C) Eigendecomposition with variances/covariances above, except for the sign of eigenvectors. The first line below represents eigenvalues and the rest, eigenvectors.

```

[ 231.35017587  25.48200804]
[[-0.94582661 -0.32467218]
 [-0.32467218  0.94582661]]

```

The line 3 of the code above used a division of a matrix by $(n-1)$. The following code used the same code without the division:

```

1. import numpy as np
2. covMat2 = dfM2.T.dot(dfM2)
3. U, S, V = np.linalg.svd(covMat2)
4. print(S)
5. print(V)

```

The result shows the same eigenvectors, so the directions of eigenvectors remain the same. But the eigenvalues are (n-1) times as big as the earlier ones:

```
1. [ 6709.15510011  738.97823322]
2. [[-0.94582661 -0.32467218]
3. [-0.32467218  0.94582661]]
```

F) Singular value decomposition (SVD) with a rectangular data matrix

SVD can decompose the original centered data matrix without relying on the correlation or covariance of datasets:

```
1. import numpy as np
2. U, S, V = np.linalg.svd(dfM2)
3. print(S)
4. print(V)
```

The result of the code above is:

```
1. [ 81.9094323  27.18415408]
2. [[ 0.94582661  0.32467218]
3. [ 0.32467218 -0.94582661]]
```

The vectors in the second and third line are the same except for the sign difference. The numbers in the first line calculated with SVD using the non-symmetric matrix are called singular values. If each is multiplied by itself, however, it becomes the same as the eigenvalue of the symmetric matrix in E) Singular value decomposition (SVD) with a symmetric data matrix above.

$$81.9094323 * 81.9094323 = 6709.1511$$

$$27.18415408 * 27.18415408 = 738.978233$$

PCA can use any of these two methods, eigendecomposition or SVD, although the latter is more stable than the former since it always succeeds in producing eigenvalues and eigenvectors.

IV. Analysis of a Real Dataset

A) Lipset dataset

In real life, the data are more complex than the 2 dimensions the current author used as an example. If there are more than 3 variables, it is impossible to display the data graphically, so one has to deal with PCA by the numerical or computer coding approach.

The following data, so-called the Lipset data²⁾, has been often used as an example by Q community when it illustrates how to use PCA and factor analysis for Q analysis. Generally, Q analysis is understood in R matrix users as correlating people instead of variables in PCA or factor analysis. In fact, the Q researchers who subscribe to William Stephenson's ideas use Q analysis only with a dataset collected in a specific way. They try to understand people's subjective schema on any events, people, issues, etc. To do so, the Q researchers usually select questions based on stratified sampling and have research participants consider other question items while rank ordering each

2) The questions items for the Lipset data were made by William Stephenson, and the Lipset data themselves were collected by Stephen Brown. This dataset as a sample is also included in the most popular Q analysis package, PQMethod. The author uses this after getting permission from Brown. For python analysis, the dataset has been adapted and posted at <http://bestelon.com/pca/lipset.csv>.

question item on a card in a Gestalt way in the two-step sorting process (Brown 1980; Stephenson 1953).

	A	B	C	D	E	F	G	H	I	J
1	item	v1	v2	v3	v4	v5	v6	v7	v8	v9
2	p01	-1	-1	2	3	-4	1	2	-2	3
3	p02	0	0	-2	1	-1	-3	0	2	1
4	p03	-2	-1	-2	-3	3	0	-2	0	0
5	p04	0	-3	4	-1	-1	3	1	-3	1
						1	3	0	4	-4
28	p27	4	2	1	-1	3	0	0	-4	-1
29	p28	2	3	-1	2	0	4	4	3	-4
30	p29	2	-1	0	0	1	-1	-3	0	0
31	p30	-3	-4	-1	2	-2	2	1	-2	-1
32	p31	-2	-2	-1	1	1	1	-1	-2	3
33	p32	-2	-3	4	2	-2	2	3	-2	-1
34	p33	4	4	1	0	-1	-3	-2	2	4

B) Import dataset and calculate correlations among variables

SPSS is the statistical package that quantitative social scientists are familiar with. As a reference point, SPSS results are compared with the results from Python code. First, let's import the Lipset data with 33 items and 9 variables:

```

1. import pandas as pd
2. filepath = "http://bestelon.com/pca/lipset.csv"
3. df = pd.read_csv(filepath)
4. print(df)

```

Correlations among variables were calculated with the following:

```

1. # select only numbers
2. numCols = df.iloc[:, 1:10]
3. numCorr = numCols.corr()
4. print(numCorr)

```

The calculation produced the 9 x 9 correlation matrix below:

	v1	v2	v3	v4	v5	v6	v7	v8	v9
v1	1.00000	0.53750	0.20625	0.22500	0.10000	-0.22500	-0.31875	0.23750	0.05000
v2	0.53750	1.00000	-0.07500	0.08750	0.17500	-0.02500	-0.16250	0.38125	0.06875
v3	0.20625	-0.07500	1.00000	0.40000	-0.54375	0.08750	0.05000	-0.09375	0.11250
v4	0.22500	0.08750	0.40000	1.00000	-0.56250	0.27500	0.16875	0.05625	0.02500
v5	0.10000	0.17500	-0.54375	-0.56250	1.00000	-0.06250	-0.13125	0.01875	-0.03125
v6	-0.22500	-0.02500	0.08750	0.27500	-0.06250	1.00000	0.61875	-0.36875	-0.21250
v7	-0.31875	-0.16250	0.05000	0.16875	-0.13125	0.61875	1.00000	-0.28750	-0.03125
v8	0.23750	0.38125	-0.09375	0.05625	0.01875	-0.36875	-0.28750	1.00000	-0.21250
v9	0.05000	0.06875	0.11250	0.02500	-0.03125	-0.21250	-0.03125	-0.21250	1.00000

SPSS produced the exact same correlation matrix, except for the numbers being rounded to 3 decimal places:

Correlation Matrix

	v1	v2	v3	v4	v5	v6	v7	v8	v9
Correlation v1	1.000	.538	.206	.225	.100	-.225	-.319	.238	.050
v2	.538	1.000	-.075	.088	.175	-.025	-.163	.381	.069
v3	.206	-.075	1.000	.400	-.544	.088	.050	-.094	.113
v4	.225	.088	.400	1.000	-.563	.275	.169	.056	.025
v5	.100	.175	-.544	-.563	1.000	-.063	-.131	.019	-.031
v6	-.225	-.025	.088	.275	-.063	1.000	.619	-.369	-.213
v7	-.319	-.163	.050	.169	-.131	.619	1.000	-.288	-.031
v8	.238	.381	-.094	.056	.019	-.369	-.288	1.000	-.213
v9	.050	.069	.113	.025	-.031	-.213	-.031	-.213	1.000

C) Eigendecomposition using np.linalg.eig

The Python code below shows how to use the Python command of np.linalg.eig() for eigendecomposition:

```

1. import numpy as np
2. eVal_corr, eVec_corr = np.linalg.eig(numCorr)
3. idx_corr = eVal_corr.argsort()[::-1]
4. eVal_corr = eVal_corr[idx_corr]
5. eVec_corr = eVec_corr[:, idx_corr]
6. print(eVal_corr.round(3))
7. print(eVec_corr.round(3))

```

The code above produced nine eigenvalues and eigenvectors:

```

1. [ 2.384  2.015  1.341  1.127  0.732  0.512  0.401  0.32  0.168]
2. [[ 0.312  0.427 -0.197 -0.281 -0.424  0.08 -0.435 -0.266  0.396]
3. [ 0.324  0.276 -0.468 -0.317  0.197 -0.157  0.539 -0.21 -0.317]
4. [-0.251  0.47  0.229 -0.018 -0.378 -0.607  0.045  0.309 -0.233]
5. [-0.27  0.511 -0.148  0.041  0.179  0.599 -0.214  0.258 -0.374]
6. [ 0.34 -0.422 -0.25 -0.29 -0.219 -0.033 -0.32  0.521 -0.37 ]
7. [-0.447 -0.076 -0.511 -0.153 -0.115  0.033  0.281  0.392  0.513]
8. [-0.46 -0.128 -0.333 -0.155  0.304 -0.379 -0.494 -0.362 -0.158]
9. [ 0.366  0.225 -0.202  0.459  0.481 -0.309 -0.213  0.345  0.27 ]
10. [ 0.017  0.09  0.439 -0.69  0.471 -0.023 -0.043  0.218  0.225]]

```

D) Eigendecomposition using np.linalg.svd for SVD

SVD can produce eigenvalues and eigenvectors:

```

1. U, S, V = np.linalg.svd(numCorr)
2. print(S.round(3))
3. print(V.T.round(3))
4. [ 2.384  2.015  1.341  1.127  0.732  0.512  0.401  0.32  0.168]
5. [[-0.312  0.427  0.197 -0.281  0.424 -0.08  0.435 -0.266 -0.396]
6. [-0.324  0.276  0.468 -0.317 -0.197  0.157 -0.539 -0.21  0.317]
7. [ 0.251  0.47 -0.229 -0.018  0.378  0.607 -0.045  0.309  0.233]
8. [ 0.27  0.511  0.148  0.041 -0.179 -0.599  0.214  0.258  0.374]
9. [-0.34 -0.422  0.25 -0.29  0.219  0.033  0.32  0.521  0.37 ]
10. [ 0.447 -0.076  0.511 -0.153  0.115 -0.033 -0.281  0.392 -0.513]
11. [ 0.46 -0.128  0.333 -0.155 -0.304  0.379  0.494 -0.362  0.158]
12. [-0.366  0.225  0.202  0.459 -0.481  0.309  0.213  0.345 -0.27 ]
13. [-0.017  0.09 -0.439 -0.69 -0.471  0.023  0.043  0.218 -0.225]]

```

The two methods, eigendecomposition and SVD, show the same results except for the reversed signs of some values. This occurs because eigenvectors are not unique. Multiplying by any constant gives another valid eigenvector. This is clear given the formula describing an eigenvector: $A \cdot v = \lambda \cdot v$. If all values measuring a concept, for example, love, switch their number signs, they will mean all values for a concept of hate. To get consistent results, the author calculated the sum of eigenvectors for each column and flip the sign by multiplying each by -1 if their sum is less than 0.

```

1. def flip_vector_sign(eVec):
2.     for i in range( eVec.shape[1]): ## changed from 1 to 0
       inside shape
3.         if (eVec[:, i].sum() < 0):
4.             eVec[:, i] = -1 * eVec[:, i]
5.
6.     return eVec

```

The function above was called below:

```

1. # To print each row on one line
2. np.set_printoptions(linewidth = 100)
3. Vt = flip_vector_sign(V.T)
4. Vt.round(3)

```

A new 9 x 9 matrix was produced:

```

1. array([[ -0.312,  0.427,  0.197,  0.281, -0.424, -0.08 ,  0.435, -0.266, -0.396],
2.        [ -0.324,  0.276,  0.468,  0.317,  0.197,  0.157, -0.539, -0.21 ,  0.317],
3.        [  0.251,  0.47 , -0.229,  0.018, -0.378,  0.607, -0.045,  0.309,  0.233],
4.        [  0.27 ,  0.511,  0.148, -0.041,  0.179, -0.599,  0.214,  0.258,  0.374],
5.        [ -0.34 , -0.422,  0.25 ,  0.29 , -0.219,  0.033,  0.32 ,  0.521,  0.37 ],
6.        [  0.447, -0.076,  0.511,  0.153, -0.115, -0.033, -0.281,  0.392, -0.513],
7.        [  0.46 , -0.128,  0.333,  0.155,  0.304,  0.379,  0.494, -0.362,  0.158],
8.        [ -0.366,  0.225,  0.202, -0.459,  0.481,  0.309,  0.213,  0.345, -0.27 ],
9.        [ -0.017,  0.09 , -0.439,  0.69 ,  0.471,  0.023,  0.043,  0.218, -0.225]])

```

E) Select principal components

Many ways of deciding on the number of factors have been suggested (Brown 1980, pp. 222-224; Field 2005, pp. 632-634; Watts & Stenner, 2012, pp. 105-110). They include checking the inflection point of the scree plot as the cut-off point or counting principal components whose eigenvalues are larger than a specific number like 1 or 0.7. If these two methods are used, the following code would be beneficial since it organizes eigenvalues for individual principal components and those in cumulative form:

```

1. exp = eVal_corr *100 / np.sum(eVal_corr)
2. accSum = np.cumsum(exp)
3. pcNum = list(range(1, 10) )
4. data = np.array([pcNum, eVal_corr, exp, accSum])
5. eigenValues = pd.DataFrame(data.T, columns = ['PC #', 'Eigenvalue', "% of Variance Exp", "Cumulative %"])
6. #Keep original for later use
7. eigenNumbers = eigenValues.copy()
8. format_mapping={'PC #': '{:,.0f}', 'Eigenvalue': '{:,.3f}',
    , '% of Variance Exp': '{:.3f}%', 'Cumulative %': '{:.3f}%'}
9. for key, value in format_mapping.items():
10.     eigenValues[key] = eigenValues[key].apply(value.format)
11. eigenValues

```

	PC #	Eigenvalue	% of Variance Exp	Cumulative %
0	1	2.384	26.486%	26.486%
1	2	2.015	22.393%	48.879%
2	3	1.341	14.898%	63.777%
3	4	1.127	12.526%	76.303%
4	5	0.732	8.136%	84.439%
5	6	0.512	5.685%	90.125%
6	7	0.401	4.459%	94.584%
7	8	0.320	3.551%	98.135%
8	9	0.168	1.865%	100.000%

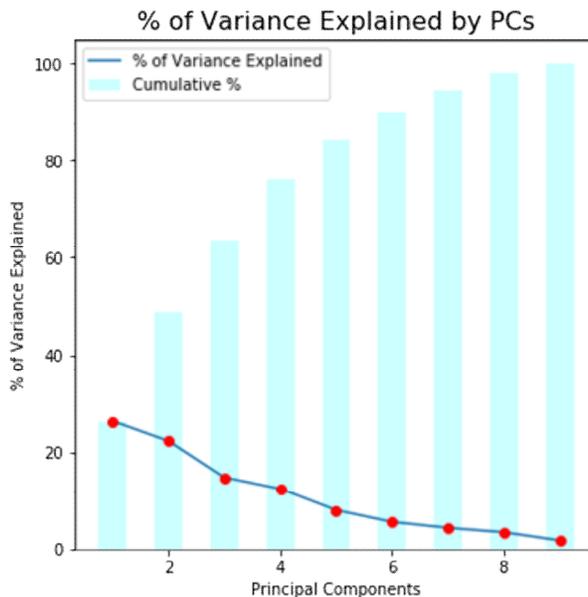
The following code was written to plot eigenvalues:

```

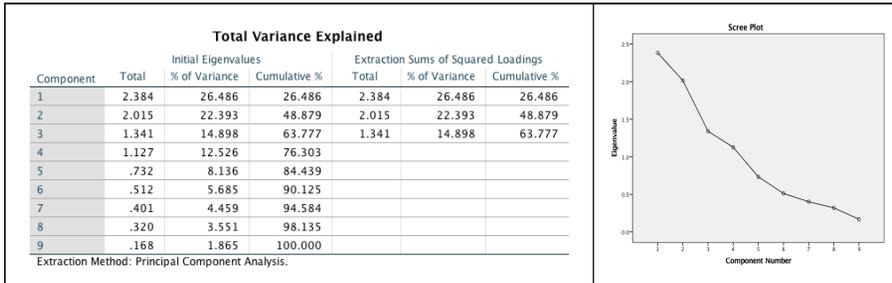
1. import matplotlib.pyplot as plt
2. plt.figure(figsize=(6, 6))
3. eachExp = eigenNumbers.iloc[:, 2]
4. plt.bar(pcNum, accSum, width = 0.5, color = 'cyan', alpha
   = 0.2, label = "Cumulative %")
5. plt.plot(pcNum, eachExp, label = "% of Variance Explained")
6. plt.plot(pcNum, eachExp, 'ro', label = '_nolegend_')
7. plt.xlabel("Principal Components")
8. plt.ylabel("% of Variance Explained")
9. plt.title("% of Variance Explained by PCs", fontsize = 16)
10. plt.legend(loc = 'upper left')
11. plt.show()

```

A scree plot and a bar chart for the cumulated percentage of variance were drawn in the same graph.



SPSS produced similar outcomes:



F) Loadings

Let's assume that three principal components were selected for analysis based on theory or an inflection point. The next analysis shows how to calculate factor loadings. Among the principal components (PCs), only the first three were selected:

```

1. eVec_corr3 = Vt[:, :3]
2. eVal_corr3 = eVal_corr[:3]
3. loadings3 = eVec_corr3 * np.sqrt(eVal_corr3)
4. print(loadings3.round(3))

```

Those three were transposed and titled with the following code:

```

1. loadings3T = loadings3.round(3).T
2. loadingsDF = pd.DataFrame(loadings3T, index = ["PC1", "PC2", "PC3"],
3.     columns = ["v1", "v2", "v3", "v4", "v5", "v6", "v7", "v8", "v9" ])
4. print(loadingsDF)

```

The loadings matrix in output shows the relationship between old variables with new principal components by calculating the coordinate of the old variables along the PC (principal component) axes:

1.	v1	v2	v3	v4	v5	v6	v7	v8	v9
2. PC1	-0.481	-0.501	0.388	0.416	-0.525	0.691	0.711	-0.565	-0.027
3. PC2	0.607	0.392	0.667	0.725	-0.600	-0.107	-0.182	0.320	0.128
4. PC3	0.228	0.542	-0.265	0.171	0.289	0.591	0.386	0.234	-0.509

The loadings can be interpreted as the Pearson correlation between a principal component and a variable. One can see v6 and v7 are high with PC1, meaning that these two variables contribute heavily to PC1. On the other hand, v9 is low, little contributing to the formation of PC1.

The loadings also can be interpreted as regression coefficients, as shown below (Field, 2005, p. 622).

For example, the three principal components are composed of the nine variables:

$$\begin{aligned}
 \text{PC1} &= -0.481 * v1 - 0.501 * v2 + \dots - 0.027 * v9, \\
 \text{PC2} &= 0.607 * v1 + 0.392 * v2 + \dots 0.128 * v9, \\
 \text{PC3} &= 0.228 * v1 + 0.542 * v2 + \dots - 0.509 * v9
 \end{aligned}
 \quad \text{Equations 6}$$

The nine loading numbers are used as regression coefficients for each of the three equations above.

SPSS created Component Matrix, the same as the loadings matrix above.

Component Matrix^a

	Component		
	1	2	3
v1	-.481	.607	.228
v2	-.501	.392	.542
v3	.388	.667	-.265
v4	.416	.725	.171
v5	-.525	-.600	.289
v6	.691	-.107	.591
v7	.711	-.182	.386
v8	-.565	.320	.234
v9	-.027	.128	-.509

Extraction Method: Principal Component Analysis.

a. 3 components extracted.

G) PC score coefficients and PC scores

Principal component scores for items can be calculated using Equations 6, which is known as a weighted average (Field, 2005, p. 625). However, more sophisticated ways were invented for calculating principal component scores. Among the most popular is the regression method. In this method, PC loadings are premultiplied by the inverse of the original matrix to calculate PC score coefficients (refer to Equation 7), which is to "take account of the initial correlations between variables...in doing so, differences in units of measurement and variable variances are stabilized." (Field, 2005, p. 626)

PC score coefficients = the inverse of the original matrix * PC loadings
Equation 7

```
1. import numpy as np
2. pcScoreCoef = np.linalg.inv(numCorr).dot(loadings3)
3. print( pcScoreCoef.round(3))
```

The following are PC score coefficients:

```
1. [[-0.202  0.301  0.17 ]
2.  [-0.21   0.194  0.404]
3.  [ 0.163  0.331 -0.198]
4.  [ 0.175  0.36   0.127]
5.  [-0.22  -0.298  0.216]
6.  [ 0.29  -0.053  0.441]
7.  [ 0.298 -0.09   0.288]
8.  [-0.237  0.159  0.174]
9.  [-0.011  0.063 -0.379]]
```

These are exactly the same as what SPSS produced as Component Score Coefficient Matrix.

Component Score Coefficient Matrix

	Component		
	1	2	3
v1	-.202	.301	.170
v2	-.210	.194	.404
v3	.163	.331	-.198
v4	.175	.360	.127
v5	-.220	-.298	.216
v6	.290	-.053	.441
v7	.298	-.090	.288
v8	-.237	.159	.174
v9	-.011	.063	-.379

Extraction Method: Principal Component Analysis.
Component Scores.

PC scores were calculated by multiplying the data matrix by the PC score coefficients, as shown below:

1. `zScore = (numCols - numCols.mean()) / numCols.std()`
2. `pcScore = zScore.dot(pcScoreCoef)`
3. `pcScore.round(5)`

The Python code produced the result (below left), and SPSS produced the same (below right):

	0	1	2		FAC1_1	FAC2_1	FAC3_1
0	1.55125	0.92806	-0.85854	3	1.55125	.92806	-.85854
1	-0.57445	0.23950	-0.46792	3	-.57445	.23950	-.46792
2	-0.66767	-1.45335	-0.29508	3	-.66767	-1.45335	-.29508
3	1.42814	0.00750	-0.73250	3	1.42814	.00750	-.73250
4	0.57579	-1.00634	1.29582	7	.57579	-1.00634	1.29582

V. Varimax Rotation

A) The purpose of varimax rotation and algorithm

PCA often needs rotation for easier interpretation. The current author used the most popular method, called Varimax rotation. Varimax orthogonal rotation tries to maximize variance of the squared loadings in each factor so that each factor has only a few variables with large loadings and many other variables with low loadings. This process helps to obtain Thurstone's simple structure (ttnphns, 2015). Kaiser suggested normalization before rotating factors (Kaiser, 1958) and offered suggestions for computer programming (Kaiser, 1959).

Varimax can proceed by rotating two principal components successively or all principal components simultaneously (Horst, 1965). Multiple versions of Python code for varimax rotation can be found on the Internet (Biggs, 2017; "Factor_rotation," n.d.). Biggs's varimax function used Kaiser's normalization and a simultaneous rotation. Another method of successive rotation for comparing each pair of principal components is also available online (Zaiontz, n.d.).

For this article, the varimax function in Biggs' factor analysis class was saved as a separate function, `varimaxFunction`, and imported for data analysis, as shown in the code below. Since the function requires a Pandas' DataFrame input, line 1 below is used to convert the data format from a numpy array to a Pandas DataFrame.

```
1. loadings3 = pd.DataFrame(loadings3)
2. from varimaxFunction import varimax
3. rotatedLoading, rotationMtx = varimax(loadings3)
4. print(rotationMtx.round(3))
5. print(rotatedLoading.round(3))
```

The function created a varimax transformation matrix (lines 1-3 below) and a rotated loading matrix (lines 4-13).

```

1. [[ 0.637  0.529  0.56 ]
2. [-0.557  0.818 -0.14 ]
3. [-0.533 -0.223  0.816]]
4.      0      1      2
5. 0 -0.766  0.191 -0.168
6. 1 -0.826 -0.066  0.107
7. 2  0.016  0.811 -0.093
8. 3 -0.230  0.776  0.271
9. 4 -0.154 -0.833  0.026
10. 5  0.184  0.146  0.885
11. 6  0.349  0.141  0.739
12. 7 -0.663 -0.089 -0.171
13. 8  0.183  0.204 -0.448
    
```

If the original loadings matrix of loadings3 is multiplied by the varimax transformation matrix, rotationMtx, as shown below, it will produce the same rotated loading matrix shown above.

```

1. check = loadings3.dot(rotationMtx)
2. print(check.round(3))
    
```

SPSS created the following two, which are the same as what the Python code calculated except for some structural differences explained below.

Component Transformation Matrix				Rotated Component Matrix^a			
Component	Component			1	2	3	
	1	2	3				
1	.530	-.637	.560	v1	.191	.766	-.168
2	.818	.557	-.140	v2	-.066	.826	.107
3	-.223	.533	.816	v3	.811	-.016	-.093
				v4	.776	.230	.271
				v5	-.833	.154	.026
				v6	.146	-.184	.885
				v7	.141	-.349	.739
				v8	-.089	.663	-.171
				v9	.204	-.183	-.448

Extraction Method: Principal Component Analysis.
 Rotation Method: Varimax with Kaiser Normalization.

Extraction Method: Principal Component Analysis.
 Rotation Method: Varimax with Kaiser Normalization.
 a. Rotation converged in 5 iterations.

B) A comparison of the SPSS and Python results

SPSS and Python produced the results, which showed differences in the signs and column order. The Python calculations need to be adjusted in two ways. First, the number sign was flipped with the earlier `flip_vector_sign` function, as shown below. The varimax loadings matrix can be directly modified without first modifying its corresponding transformation matrix before recalculating the varimax loadings matrix.

```
1. # flip_vector_sign requires a numpy matrix
2. rotLoadNPMatrix = rotatedLoading.as_matrix()
3. rotatedLoadingFlip = flip_vector_sign(rotLoadNPMatrix)
4. print(rotatedLoadingFlip.round(3))
```

Second, the principal component columns should be rearranged based on the amount of variance that the principal components account for. After varimax rotation, it is possible that the original 1st principal component may not explain the largest amount of variance. This is what happened with the data above. Therefore, the term factors is preferred over principal components after a varimax rotation is performed on principal component axes.

To check the amount of variance each principal component accounts for, the following `matrixDescReorder` function summed loadings for each column and rearranged all columns in descending order.

```
1. def matrixDescReorder(matrix):
2.     ### squared each number -- the portion of explanation
3.     rotatedPctExplained = matrix**2
4.     ### sum columnwise
5.     rotatedPctExplainedSum = np.sum(rotatedPctExplained,
6.     axis = 0)
7.     ### sort them in descending order
```

```

8. ### final loading matrix is rotatedLoadingFlipOrdered
9.     indexN = rotatedPctExplainedSum.argsort()[::-1]
10.    rotatedReordered = matrix[:, indexN]
11.    return rotatedReordered

```

The flipped rotated loading matrix was used as an argument for the `matrixDescReorder` function:

```

1. rotatedLoadingFlipReordered = matrixDescReorder(rotatedLoadingFlip)
2. print(rotatedLoadingFlipOrdered.round(3))

```

The result below produced the same as what SPSS did earlier. Columns were reordered and number signs were adjusted.

```

1. [[ 0.191  0.766 -0.168]
2. [-0.066  0.826  0.107]
3. [ 0.811 -0.016 -0.093]
4. [ 0.776  0.23  0.271]
5. [-0.833  0.154  0.026]
6. [ 0.146 -0.184  0.885]
7. [ 0.141 -0.349  0.739]
8. [-0.089  0.663 -0.171]
9. [ 0.204 -0.183 -0.448]]

```

C) Factor score coefficients

Factor score coefficient matrix can be calculated as the PC score coefficient matrix was calculated in Equation 7.

```

1. cscmRotated = np.linalg.inv(zScore.corr()).dot(rotatedLoadingFlipReordered)
2. print(cscmRotated.round(3))

```

The result (on the left below) is the same as that produced by SPSS (on the right).

<ol style="list-style-type: none"> 1. [[<u>0.102</u> <u>0.387</u> -0.016] 2. [-<u>0.042</u> <u>0.457</u> <u>0.185</u>] 3. [<u>0.401</u> -<u>0.024</u> -<u>0.117</u>] 4. [<u>0.358</u> <u>0.157</u> <u>0.151</u>] 5. [-<u>0.408</u> <u>0.089</u> <u>0.094</u>] 6. [<u>0.012</u> <u>0.021</u> <u>0.53</u>] 7. [<u>0.02</u> -<u>0.087</u> <u>0.415</u>] 8. [-<u>0.034</u> <u>0.332</u> -<u>0.013</u>] 9. [<u>0.131</u> -<u>0.16</u> -<u>0.325</u>]] 	<p>Component Score Coefficient Matrix</p> <table border="1"> <thead> <tr> <th rowspan="2"></th> <th colspan="3">Component</th> </tr> <tr> <th>1</th> <th>2</th> <th>3</th> </tr> </thead> <tbody> <tr><td>v1</td><td>.102</td><td>.387</td><td>-.016</td></tr> <tr><td>v2</td><td>-.042</td><td>.457</td><td>.185</td></tr> <tr><td>v3</td><td>.401</td><td>-.024</td><td>-.117</td></tr> <tr><td>v4</td><td>.358</td><td>.157</td><td>.151</td></tr> <tr><td>v5</td><td>-.408</td><td>.089</td><td>.094</td></tr> <tr><td>v6</td><td>.012</td><td>.021</td><td>.530</td></tr> <tr><td>v7</td><td>.020</td><td>-.087</td><td>.415</td></tr> <tr><td>v8</td><td>-.034</td><td>.332</td><td>-.013</td></tr> <tr><td>v9</td><td>.131</td><td>-.160</td><td>-.325</td></tr> </tbody> </table> <p>Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Component Scores.</p>		Component			1	2	3	v1	.102	.387	-.016	v2	-.042	.457	.185	v3	.401	-.024	-.117	v4	.358	.157	.151	v5	-.408	.089	.094	v6	.012	.021	.530	v7	.020	-.087	.415	v8	-.034	.332	-.013	v9	.131	-.160	-.325
	Component																																											
	1	2	3																																									
v1	.102	.387	-.016																																									
v2	-.042	.457	.185																																									
v3	.401	-.024	-.117																																									
v4	.358	.157	.151																																									
v5	-.408	.089	.094																																									
v6	.012	.021	.530																																									
v7	.020	-.087	.415																																									
v8	-.034	.332	-.013																																									
v9	.131	-.160	-.325																																									

Factor scores were calculated with the following code:

```

1. pcScoreRotated = zScore.dot(cscmRotated)
2. print(pcScoreRotated.round(5))
    
```

The Python code produced a little different result (on the left below) from what SPSS did (on the right). But the difference, which might occur due to the computer's rounding errors, was negligible. Only the first five rows are shown instead of the entire 33 rows to save space.

<ol style="list-style-type: none"> 1. 0 1 2 2. 0 1.77244 -0.92830 <u>0.03845</u> 3. 1 -0.00379 0.24997 -0.73743 4. 2 -1.47725 -0.54182 -0.41141 5. 3 0.92569 -1.29567 <u>0.20132</u> 6. 4 -0.80778 -0.23698 1.52144 	<table border="1"> <thead> <tr> <th></th> <th>FAC1_2</th> <th>FAC2_2</th> <th>FAC3_2</th> </tr> </thead> <tbody> <tr><td></td><td>1.77245</td><td>-.92827</td><td>.03843</td></tr> <tr><td></td><td>-.00380</td><td>.24998</td><td>-.73743</td></tr> <tr><td></td><td>-1.47725</td><td>-.54184</td><td>-.41140</td></tr> <tr><td></td><td>.92571</td><td>-1.29566</td><td>.20130</td></tr> <tr><td></td><td>-.80776</td><td>-.23700</td><td>1.52144</td></tr> </tbody> </table>		FAC1_2	FAC2_2	FAC3_2		1.77245	-.92827	.03843		-.00380	.24998	-.73743		-1.47725	-.54184	-.41140		.92571	-1.29566	.20130		-.80776	-.23700	1.52144
	FAC1_2	FAC2_2	FAC3_2																						
	1.77245	-.92827	.03843																						
	-.00380	.24998	-.73743																						
	-1.47725	-.54184	-.41140																						
	.92571	-1.29566	.20130																						
	-.80776	-.23700	1.52144																						

VI. Theoretical or Exploratory Rotation

In Q analysis, factor axes are often rotated for theoretical reasons (Brown, 1980, pp. 224-239). Factor axes may be rotated toward variables with same characteristics (in the case of Q analysis, people are treated as variables). Of course, factor axes also can be rotated for serendipity that may lead to unexpected insights. This type of rotations is unique to Q analysis that Q community has adopted.

In the case of R community, it just reverses observations and variables of any dataset to obtain a Q matrix, but does not rotate factor axes after varimax rotation. Thus, SPSS and other statistical packages, which are not focused on the Q community, have not incorporated the feature for theoretical or exploratory rotations.

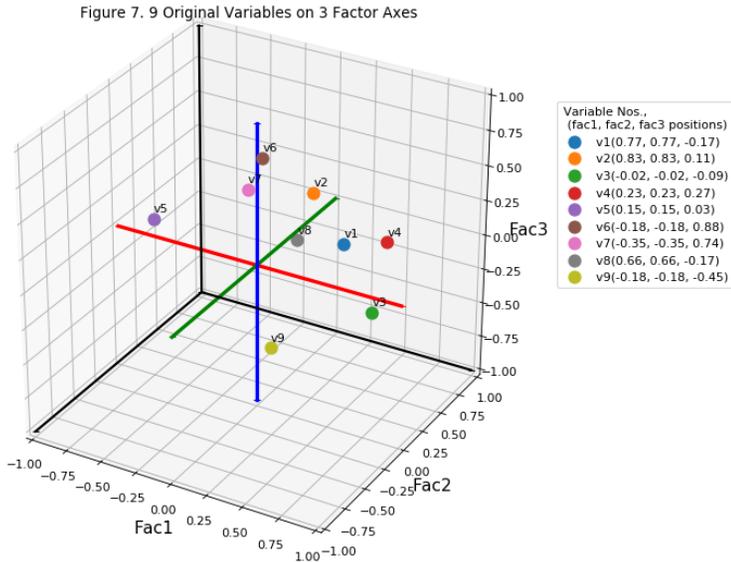
For rotation of axes, a user needs to see an arrangement of variables with factor axes. Either two or three axes can be displayed.

A) 3-dimension graph

A 3-dimensional graph can be drawn using `facsGraph3` in Appendix 1. This function is applied to an argument of the varimax factor loadings matrix, `loadingsRot`, which is the Pandas' DataFrame version of `rotatedLoadingFlipReordered`.

```
1. # pcsGraph3 requires DataFrame
2. loadingRot = pd.DataFrame(rotatedLoadingFlipReordered.copy())
3. facsGraph3("Figure 7", loadingRot)
```

The following graph displays the coordinate of variables along the three factor axes, and the legend shows the specific coordinates of each variable (refer to Figure 7).



B) Rotate two axes at a time while viewing a 3-dimensional graph

Although three axes can be changed theoretically, it will be daunting to human brains. Thus, the author wrote code, transMatrix3 in Appendix II, after incorporating the following three formulae ("Rotation Matrix," n.d.). Rotation theta degrees were multiplied by -1 since the computer code actually rotates variable dots on the graph, while people tend to think about the rotation of the axes. For example, rotating data dots by 5 degrees counterclockwise means rotating the axes by -5 degrees counterclockwise.

$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$	$R_y(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}$	$R_z(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$
--	--	--

The function above takes the number of degrees for a clockwise rotation and two axes as its arguments. One may consider rotating the x- and z-axis along the y-axis by some degrees, like -8 degrees, counterclockwise after viewing Figure 7. Then the following code will do the job:

```
1. import numpy as np
2. rotMatrix1 = transMatrix3(-8, 1, 3)
3. rotMatrix1.round(3)
```

In linear algebra, a rotation by some angles can be translated into a rotating matrix like the following:

```
1. array([[ 0.99 ,  0.   ,  0.139],
2.        [ 0.   ,  1.   ,  0.   ],
3.        [-0.139,  0.   ,  0.99 ]])
```

The rotMatrix1 above is used as a factor transformation matrix in the following code to rotate the varimax factor loadings matrix:

```
1. loadingRot1 = loadingRot.dot(rotMatrix1)
2. print(loadingRot1.round(3))
```

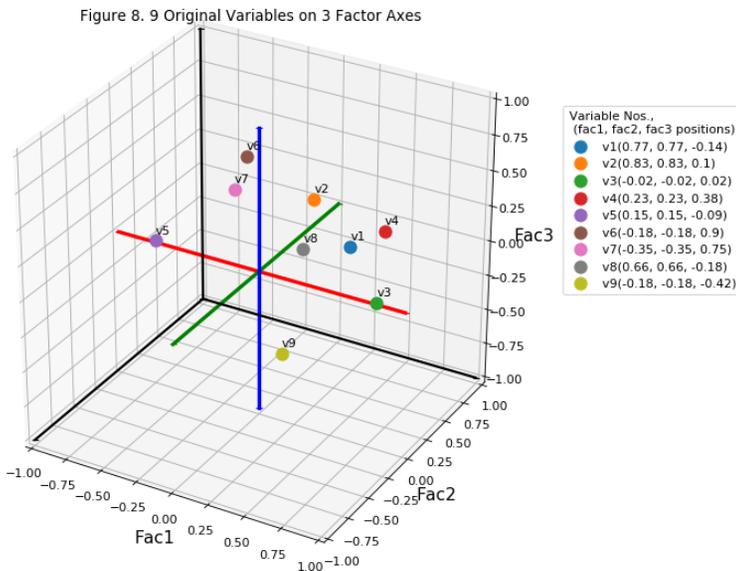
C) A new factor loadings matrix after a rotation

The newly emerged factor loadings changed the first and third columns while keeping constant the second column that represents the y-axis.

```
1.          0      1      2
2.  0  0.213  0.766 -0.140
3.  1 -0.080  0.826  0.097
4.  2  0.816 -0.016  0.021
5.  3  0.731  0.230  0.376
6.  4 -0.829  0.154 -0.090
7.  5  0.021 -0.184  0.897
```

8.	6	0.037	-0.349	0.751
9.	7	-0.064	0.663	-0.182
10.	8	0.264	-0.183	-0.415

When the new factor loadings matrix was plugged into the `fac3Graph3` function, a new 3-dimensional graph was created (refer to Figure 8). Notice the rotation of only the x- and z-axis in Figure 8 below.



D) Calculate a factor score coefficient matrix and factor scores

To create a factor score coefficient matrix, a loadings matrix is premultiplied by the inverse of the original correlation matrix as in Equation 7.

```
1. cscmRot1 = np.linalg.inv(zScore.corr()).dot(loadingsRot1)
2. print(cscmRot1.round(3))
```

The following is the factor score coefficient matrix:

1.	[0.103	0.387	-0.002]
2.	[-0.069	0.457	0.177]
3.	[0.414	-0.024	-0.06]
4.	[0.335	0.156	0.2]
5.	[-0.416	0.089	0.037]
6.	[-0.062	0.023	0.526]
7.	[-0.039	-0.088	0.413]
8.	[-0.032	0.333	-0.018]
9.	[0.175	-0.159	-0.304]]

Then, factor scores can be calculated by multiplying the data matrix by this factor score coefficient matrix. The first five rows of the result are shown below:

1.		0	1	2
2.	0	1.749	-0.929	0.284
3.	1	0.099	0.248	-0.732
4.	2	-1.405	-0.541	-0.613
5.	3	0.889	-1.293	0.329
6.	4	-1.014	-0.238	1.396

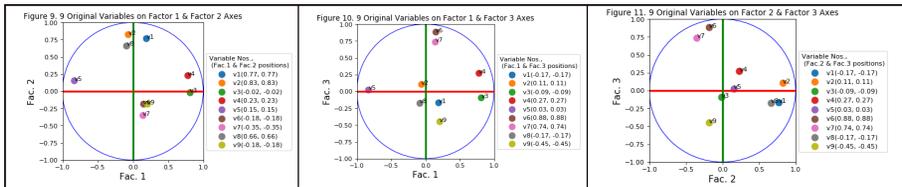
E) 2-dimensional graph and rotation

One may prefer to see loadings in 2-dimensional graphs instead to decide the first two axes to rotate. If one has three axes in total, like factor 1, factor 2 and factor 3, three graphs are possible: one with factors 1 and 2, and the second one with factors 1 and 3, and the last with factors 2 and 3. The four factors will lead to six 2-dimensional graphs.

The function `facfsGraph2` in Appendix III was written to draw a 2-dimensional graph with function arguments of a graph title, a loading matrix, the number representing two axes (1 for factor 1, 2 for factor 2, etc.), and the width and height of a graph in centimeters. The three lines of code created

three figures 9, 10 and 11.

1. `facsGraph2("Figure 9", loadingRot, 1, 2, 4, 4)`
2. `facsGraph2("Figure 10", loadingRot, 1, 3, 4, 4)`
3. `facsGraph2("Figure 11", loadingRot, 2, 3, 4, 4)`



F) Rotate axes in a 2-dimensional graph

First, a rotation of axes by theta degrees can be converted into a rotation matrix with the following `transMatrix2` function:

```

1. from math import sin, cos, radians
2. def trans2DMatrix(theta):
3.     theta = -theta
4.     a1 = cos(radians(theta))
5.     a2 = -sin(radians(theta))
6.     a3 = sin(radians(theta))
7.     a4 = cos(radians(theta))
8.     transMat = np.array([ [a1, a2], [a3, a4] ])
9.     return transMat

```

The two axes in Figure 11 may be rotated by 15 degrees counterclockwise. The `trans2DMatrix` is applied to the 15 degrees for a rotation transformation matrix:

```

1. trans2DMat1 = trans2DMatrix(15)
2. print(trans2DMat1.round(3))

```

The function call created the following matrix:

```

1. [[ 0.966 -0.259]
2. [ 0.259  0.966]]

```

This 2D transformation matrix needs to be premultiplied by the varimax rotation matrix, specifically only two columns of the matrix that are related to the factors to be rotated. The following rotate2Factors function needs four arguments--the current loading matrix, the 2D transformation matrix, and the factor numbers representing the two factors.

```

1. def rotate2Factors(matrix, trans2DMat, fNum1, fNum2):
2.     #select 2 columns
3.     twoFacs = matrix.iloc[:, [fNum1 - 1, fNum2 - 1] ]
4.     twoFacRot = twoFacs.dot(trans2DMat)
5.
6.     matrix.iloc[:, fNum1 - 1]= twoFacRot.iloc[:, 0]
7.     matrix.iloc[:, fNum2-1]= twoFacRot.iloc[:, 1]
8.     ## return the whole matrix
9.     return matrix

```

To prevent the same varimax rotation matrix from being used in mixed way in both 3d and 2d rotations, one needs to create a new copy from the varimax loadings matrix.

```

1. # newly create a copy of the loading matrix after varimax
2. loadingRot2D = pd.DataFrame(rotatedLoadingFlipReordered.copy())
3. # the earlier loading matrix, loading transmission matrix,
   two columns to be rotated
4. loadingRot2D1 = rotate2Factors(loadingRot2D, trans2DMat1,
   2, 3)
5. print(loadingRot2D1.round(3))

```

A new loading matrix was calculated:

```

1.      0      1      2
2. 0  0.191  0.696 -0.361
3. 1 -0.066  0.826 -0.110

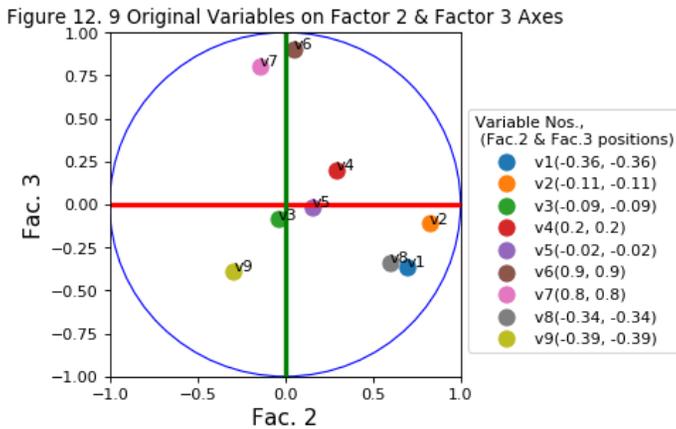
```

4.	2	0.811	-0.040	-0.086
5.	3	0.776	0.292	0.202
6.	4	-0.833	0.155	-0.015
7.	5	0.146	0.051	0.902
8.	6	0.141	-0.146	0.804
9.	7	-0.089	0.596	-0.337
10.	8	0.204	-0.293	-0.385

G) Check the graph for rotation

The new loadings matrix, loadingRot2D1, can be plugged into facsGraph2 for a new graph.

1. pcsGraph2("Figure 12", loadingRot, 1, 2, 4, 4)



One can notice the rotation based on the changes to the position of data dots. Of course, more rotations can be done as needed. Since rotation may change the number sign and the amount of variance each factor accounts for, one may need to flip the number sign and reorder the data columns if needed.

VII. A Summary of All Procedures

PCA goes through many different steps to produce the results a researcher wants. As the principal component analysis (PCA) process chart in Appendix IV shows, eigenvalues and eigenvectors have to be calculated for eigendecomposition in the beginning. This can be done by creating a correlation or covariance matrix from the data (step 1 in Appendix IV), which is further processed through either eigendecomposition (edc) or singular value decomposition (svd) (step 2). They also can be derived directly from the original data matrix through svd (step 3).

PC loadings, which is called Component Matrix in SPSS, can be calculated with the formula of eigenvectors $\ast \sqrt{\text{eigenvalues}}$ (step 4).

If rotation is not needed, the principal component (PC) score coefficient matrix can be calculated using one of many methods, like regression method, with the formula, the inner product of the inverse of the original correlation or covariance matrix and the loadings matrix (step 5). This is called Component Score Coefficient Matrix in SPSS. If the matrix of the original standardized or centered data is multiplied by PC score coefficient matrix, it will produce PC scores (step 6). In SPSS, a user has to select "Analyze" from the top menu, "Dimension Reduction" from the pulldown menu, "Scores" in the Factor Analysis dialog box, "Regression" in the Factor Analysis Factor Scores dialog box, and "Save as variables." Then principal factor scores will be appended to the end of the original dataset.

Varimax rotation will rearrange the amount of variance among principal components, so principal components lose their characteristics of extracting the maximum variance possible each time when an additional principal component is added. Thus, the term factors is preferred over principal components after varimax rotation. Rotation of principal components can be

performed on two principal components at a time or all principal components simultaneously. All the rotation angles are translated into the factor loading transformation matrix (step 7). This transformation matrix is called Component Transformation Matrix in SPSS. If the original PC loadings is multiplied by the transformation matrix, a new varimax factor loadings matrix is produced (step 8). SPSS calls it Rotated Component Matrix. One can calculate a new factor score coefficient matrix (step 9) and factors scores (step 10) using the same formula as in steps 5 and 6. SPSS appends new factor scores to the end of the dataset.

If one wants to continue to rotate factor axes after varimax rotation for further insights, rotation of factors axes needs to be expressed as a transformation matrix (step 11). This matrix can be used to produce a new factor loadings matrix (step 12), the same as in step 8. This process can be done until one is satisfied with the loadings of variables with factors. When one comes to a satisfactory loadings matrix, one can calculate its corresponding factor score coefficient matrix (step 13) and factor scores (step 14).

Q scholars who are interested in theoretical rotations may consult theory (step 15) and rotate factor axes as the theory suggests (step 16). Then, the factor score coefficient matrix and factor scores can be calculated in the same way as in step 13 and step 14 earlier.

VIII. Conclusion

Cranking out results using PCA analysis on one of the commercial statistical packages is easy, just importing data and clicking a few buttons. For a deep understanding of PCA, however, one needs to go inside the program and see

what happens inside the black box by running code. Many scholars may have a basic knowledge of a popular programming language like Python. But PCA code in Python is not compiled in one place. A user has to gather code scattered around the Internet, tweak it for compatibility and fill any gaps left.

Also, commercial statistical packages have been geared toward R analysis so they cannot be used for theoretical or exploratory analysis beyond varimax rotation, a unique Q approach. Q analysis emphasizes the importance of factor scores. R analysis, on the other hand, often ignores them.

This tutorial shows comparisons of Python and SPSS. Also, this tutorial shows all the steps needed for PCA along with Python code beyond varimax rotation. Therefore, this would help anyone who wants to run PCA at a deeper level and R scholars who want to learn Q analysis.

References

- Biggs, J. 2017. *Factor_analyzer* [Web page]. Retrieved from github.com: https://github.com/EducationalTestingService/factor_analyzer/blob/master/factor_analyzer/factor_analyzer.py
- Brown, S. R. 1980. *Political subjectivity: Applications of Q methodology in political science*. New Haven, CT: Yale University Press.
- Campbell, N. A., and Atchley, W. R. 1981. The geometry of canonical variate analysis. *Systematic Zoology* 30(3): 268-280. Retrieved from <http://www.jstor.org/stable/2413249>
- Cureton, E. E., and D'Agostino, R. B. 1983. *Factor analysis: An applied approach*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Markaki, E. N., Chadjipandelis, T., and Tomaras, P. 2014. *Procedia - Social and Behavioral Sciences* 147: 554-560.
- Factor_rotation. n.d. [Web page]. Retrieved from github.com: https://github.com/mvds314/factor_rotation/blob/master/README.md
- Field, A. 2005. *Discovering statistics using SPSS : (And sex, drugs and rock 'n' roll)* (2nd ed.). London: Sage.
- Gorsuch, R. L. 1983. *Factor analysis*. Hillsdale, N.J: L. Erlbaum Associates.
- Horst, P. 1965. *Factor analysis of data matrices*. New York: Holt, Rinehart and Winston Inc.
- Kabacoff, R. A. 2015. *R in action : Data analysis and graphics with R*. Shelter Island, N.Y.: Manning Publications. Retrieved from Amazon.
- Kaiser, H. 1959. Computer program for varimax rotation in factor analysis. *Education and Psychological Measurement* 19(3): 413-420.
- Kaiser, H. F. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23(3): 187-200.

- Lay, D. C. 2005. *Linear algebra and its applications* (3rd ed.). Boston: Addison-Wesley.
- Rotation Matrix. n.d. [Web page]. Retrieved from en.wikipedia.org: https://en.wikipedia.org/wiki/Rotation_matrix
- Stephenson, W. 1953. *The study of behavior: Q-technique and its methodology*. Chicago, IL: University of Chicago Press.
- Strang, G. 2016. *Introduction to linear algebra*. Wellesley, Massachusetts: Wellesley-Cambridge Press.
- ttnpnhs. 2015. *Factor rotation methods (varimax, oblimin, etc.) - what do the names mean and what do the methods do?* [Web page]. Retrieved from stats.stackexchange.com: <https://stats.stackexchange.com/questions/185216/factor-rotation-methods-varimax-oblimin-etc-what-do-the-names-mean-and-what-do-the-methods-do>
- Watts, S., and Stenner, P. 2012. *Doing Q methodological research: Theory, method and interpretation*. Los Angeles: Sage.
- Zaiontz, C. n.d. Rotation | *Real Statistics Using Excel*. Retrieved from www.real-statistics.com: <http://www.real-statistics.com/multivariate-statistics/factor-analysis/rotation/>

Appendix 1: facsGraph3

```

1. def facsGraph3(figNum, loadingMat):
2.     # import modules
3.     from mpl_toolkits.mplot3d import Axes3D
4.     #from matplotlib.collections import LineCollection
5.     import matplotlib.pyplot as plt
6.     import numpy as np
7.
8.     #basic frame
9.     fig=plt.figure(figsize=(9, 9), dpi= 80, facecolor='w', edgecolor='k')
10.
11.     # start to ceate a 3d graph
12.     #111 means 1 row, 1 column, and the first slot
13.     ax = fig.add_subplot(111, projection='3d')
14.
15.     #displaying dots
16.     for i in range(9):
17.         fac1 = loadingMat.iloc[i:i+1, 0 ]
18.         fac2 = loadingMat.iloc[i:i+1, 1 ]
19.         fac3 = loadingMat.iloc[i:i+1, 2 ]
20.         #ax.scatter(fac1, fac2, fac3, c = "r", marker = "o")
21.         scatter1 = ax.scatter(fac1, fac2, fac3, s = 100, cmap="Diverging")
22.
23.     arrX = loadingMat.iloc[:, 0 ]
24.     arrY = loadingMat.iloc[:, 1 ]
25.     arrZ = loadingMat.iloc[:, 2 ]
26.     # create labels starting with v1
27.     labels = []
28.     for i in range(loadingMat.shape[0]):
29.         labels.append("v" + str(i + 1) )
30.     labels = np.array(labels)
31.     #give the labels to each point
32.     for x, y, z, label in zip(arrX, arrY, arrZ, labels):
33.         ax.text(x, y, z+ 0.05, label )
34.
35.     # 3 principal component axes
36.     ax.plot([-1, 1], [0, 0], [0, 0], color='r', lw=3, marker="_", label='_
nolegend_')
37.     ax.plot([0, 0],[-1, 1], [0, 0], color='g', lw=3, marker="_", label=
'_nolegend_')
38.     ax.plot([0, 0], [0, 0], [-1, 1], color='b', lw=3, marker="_", label=
'_nolegend_')

```

```

39.
40.     # 3 axes on the border
41.     ax.plot([-1, 1], [1, 1], [-1, -1], color='k', lw=2, marker="_", label
='_nolegend_')
42.     ax.plot([-1, -1], [-1, 1], [-1, -1], color='k', lw=2, marker="_", lab
el='_nolegend_')
43.     ax.plot([-1, -1], [1, 1], [-1, 1], color='k', lw=2, marker="_", label
='_nolegend_')
44.
45.     arrX = arrY.round(2)
46.     arrY = arrY.round(2)
47.     arrZ = arrZ.round(2)
48.
49.     lgd = ax.legend( [str(label)+ "(" + str(x) + ", " + str(y) + ", " +
str(z) + ")" for x, y, z, label in zip(arrX, arrY, arrZ, labels) ], loc='u
pper left', bbox_to_anchor=(1, 0.8))
50.     lgd.set_title("Variable Nos. \n (fac1, fac2, fac3 positions)")
51.
52.     ax.set_xlabel('Fac1', fontsize = 14)
53.     ax.set_ylabel('Fac2', fontsize = 14)
54.     ax.set_zlabel('Fac3', fontsize = 14)
55.     ax.set_title(figNum + ". " + str(loadingMat.shape[0]) + " Original Vari
ables on " + str(loadingMat.shape[1]) + " Factor Axes")
56.
57.     ax.set_xlim3d(-1, 1)
58.     ax.set_ylim3d(-1, 1)
59.     ax.set_zlim3d(-1, 1)
60.     plt.show()
61.     return None

```

Appendix II. transMatrix3

```
1. def transMatrix3(theta, axis1, axis2):
2.     import math
3.     theta = -theta
4.     if(axis1 == 2 and axis2 == 3):
5.         ##transMatrixX(theta):
6.         transMat = np.array([ [1, 0, 0 ],
7.                               [0, cos(math.radians(theta)), -sin(math.radians
8. (theta)) ],
9.                               [ 0, sin(math.radians(theta)), cos(math.radians
10. (theta))]
11.                               ])
12.     elif(axis1 == 1 and axis2 == 3):
13.         #transMatrixY(theta):
14.         transMat = np.array([ [cos(math.radians(theta)), 0, sin(math.rad
15. ians(theta)) ],
16.                               [0, 1, 0 ],
17.                               [-sin(math.radians(theta)), 0, cos(math.radians
18. (theta))]
19.                               ])
20.     elif(axis1 == 1 and axis2 == 2):
21.         #transMatrixZ(theta):
22.         transMat = np.array([ [cos(math.radians(theta)), -sin(math.radian
23. s(theta)), 0 ],
24.                               [sin(math.radians(theta)), cos(math.radians(theta
25. )), 0],
26.                               [0, 0, 1 ],
27.                               ])
28.     else:
29.         pass;
30.     return transMat
```

Appendix III. facsGraph2

```

1. import matplotlib.pyplot as plt
2. import numpy as np
3.
4. def facsGraph2(figNum, loadingMat, f_input1 = 1, f_input2 = 2, figW =8, fig
   H = 8):
5.     # import modules
6.     f1 = f_input1 -1
7.     f2 = f_input2 -1
8.
9.     #basic frame
10.    fig=plt.figure(figsize=(figW, figH), dpi= 80, facecolor='w', edgecolor=
    'k')
11.
12.    # start to ceate a 3d graph
13.    #111 means 1 row, 1 column, and the first slot
14.    ax = fig.add_subplot(111 )
15.
16.    #displaying dots
17.    for i in range(loadingMat.shape[0]):
18.        fac1 = loadingMat.iloc[i:i+1, f1 ]
19.        fac2 = loadingMat.iloc[i:i+1, f2 ]
20.
21.        #ax.scatter(f1, f2, f3, c = "r", marker = "o")
22.        scatter1 = ax.scatter(fac1, fac2, s = 100, cmap="Diverging")
23.
24.    arrX = loadingMat.iloc[:, f1]
25.    arrY = loadingMat.iloc[:, f2]
26.
27.    # create labels starting with v1
28.    labels = []
29.    for i in range(loadingMat.shape[0]):
30.        labels.append("v" + str(i + 1) )
31.    labels = np.array(labels)
32.
33.    #give the labels to each point
34.    for x, y, label in zip(arrX, arrY, labels):
35.        ax.text(x, y, label)
36.
37.    # 3 principal component axes

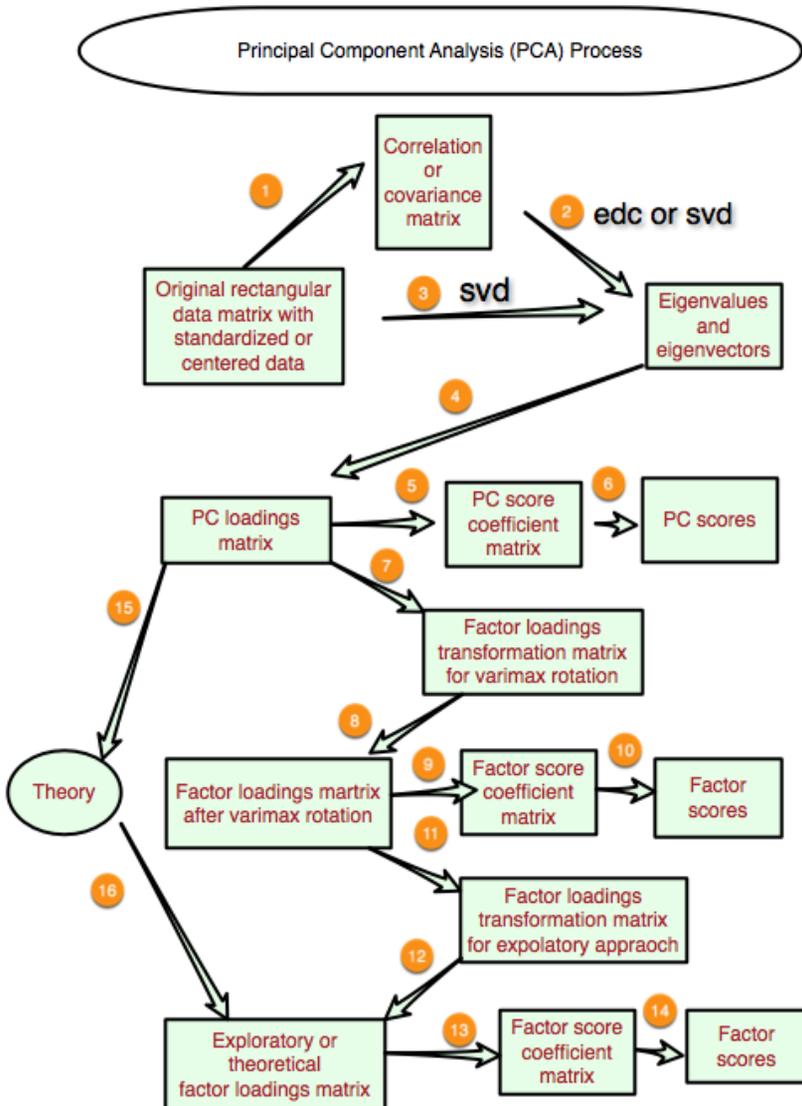
```

```

38.     ax.plot([-1, 1], [0, 0], color='r', lw=3, marker="_", label='_nolegend
    _')
39.     ax.plot([0, 0],[-1, 1], color='g', lw=3, marker="_", label='_nolegend
    _')
40.
41.     circle = plt.Circle((0, 0), 1, color='b', fill=False)
42.     ax.add_artist(circle)
43.
44.     arrX = arrY.round(2)
45.     arrY = arrY.round(2)
46.
47.     lgd = ax.legend( [str(label)+ "(" + str(x) + ", " + str(y) + ")" for
    x, y, label in zip(arrX, arrY, labels) ], loc='upper left', bbox_to_anchor
    =(1, 0.8))
48.     lgd.set_title("Variable Nos. ,\n (Fac." + str(f_input1) + " & Fac." +
    str(f_input2) + " positions)")
49.
50.     # [ 'Lag ' + str(lag) for lag in arrX]
51.     #xyz labels
52.     ax.set_xlabel('Fac. ' + str(f_input1), fontsize = 14)
53.     ax.set_ylabel('Fac. ' + str(f_input2), fontsize = 14)
54.     ax.set_title(figNum + ". " + str(loadingMat.shape[0]) + " Original Vari
    ables on Factor " + str(f_input1) + " & Factor " + str(f_input2) + " Axes
    ")
55.
56.     #xyz range
57.     ax.set_xlim(-1, 1)
58.     ax.set_ylim(-1, 1)
59.
60.     #display
61.
62.     plt.show()
63.
64.     return None

```

Appendix IV. Principal Component Analysis(PCA) Process



Notes: edc in step 2 indicates eigendecomposition, and svd in step 3, singular value decomposition.

Varimax 회전 및 그 이후: R 및 Q 분석을 위한 선형 대수, 시각화 및 Python 프로그래밍을 사용한 PCA에 대한 사용지침서

이 병 수*

논문요약

주성분 분석(PCA)은 학문 및 실제 목적을 위해 많이 사용되었습니다. 이 사용지침서는 Q 분석에 관심있는 R 학자 뿐만 아니라 PCA를 더 잘 활용하고자 하는 개인에게 도움이 될 것입니다. 많은 상용 통계 패키지가 PCA를 처리할 수 있지만 PCA 코드를 직접 실행하면 PCA에 대한 더 깊게 이해할 수 있습니다. 상업용 통계 패키지에서 PCA 기능을 사용하여 데이터를 입력하고 몇 개의 버튼을 누르기만 하면 통계 결과가 나옵니다. 하지만 PCA에 대한 깊은 이해를 위해서는 개인이 코드를 실행할 때 어떤 일이 일어나는지 알아야합니다. 많은 학자들이 파이썬과 같은 인기있는 프로그래밍 언어에 대한 기본 지식을 보유하고 있을 수 있습니다. 그러나 파이썬 코드는 한 곳에서 깔끔하게 정리되어 모아져 있지 않습니다. 사용자는 인터넷에 흩어져 있는 파이썬 코드를 수집하고 호환성을 위해 조정할 다음 부족한 부분이 있다면 채워야 합니다. 또한 R 분석을 대상으로 한 상업 통계 패키지는 배리맥스 방법을 넘어선, 즉 Q 연구를 위한 이론적 또는 탐색적 분석을 위한 팩터 회전에 사용할 수 없습니다. 이 사용지침서에서는 PCA에 필요한 모든 Python 코드를 제공하면서 결과를 SPSS (Statistical Package for the Social Science)출력과 비교했습니다. 이 사용지침서는 Q 분석을 위한 필수 요소인 팩터 축의 이론적 또는 탐색적 팩터 회전도 다루었습니다.

주제어 : PCA, 주성분 분석, 선형 대수, 그래프, 파이썬 코드, varimax 회전, R 및 Q 분석

투고일: 2018.04.15. 심사일: 2018.04.16. 게재확정일: 2018.04.23.

* 일란대학교, 부교수

[DOI] <http://dx.doi.org/10.21487/jrm.2018.5.3.1.25>

【연구논문】

효과적인 남북한 재난대응 협력을 위한 법제도적 문제 고찰

김진원* · 정한범**

논문요약

본 연구는 효과적인 남북한 재난대응 협력을 위한 법제도적 문제를 살펴보기 위한 연구이다. 연구에서는 과거 주요한 남북 재난대응 협력사례를 살펴보고, 북한지역에서의 재난 대응을 위한 현재 한국의 법적, 제도적 문제를 고찰하고 이를 개선하는 방안을 제시 할 것이다. 남북한 간의 정치적, 군사적 이슈로 인한 마찰과 교류의 단절은 북한지역에서의 재난 발생 시 효율적인 대응과 지원을 어렵게 한다. 하지만 재난대응과 인도주의적 지원은 시의적절하고 충분해야만 그 목적을 달성할 수 있는 만큼, 효율적이고도 신속한 개입이 필요하다. 이를 위하여 북한에서 발생한 재난의 종류에 따른 접경지역 지방자치단체와 중앙정부의 행동 규범이 필요하다. 또한 인도주의적 지원 시 지원물품 품목 및 지원수단 등이 적혀있는 매뉴얼이 필요하며 이 매뉴얼의 작동을 위한 법적제도의 정비도 필요할 것이다.

주제어 : 남북교류협력, 재난대응, 인도주의적 지원

I. 서론

1. 연구의 배경과 목적

2014년 4월 16일 발생한 ‘세월호 사건’으로 인해 한국 사회 내에서의 재난·재해의

* 국방대학교 안보정책학과, 대학원생

** 교신저자, 국방대학교 안보정책대학원 안보정책학과, 교수

대응과 관리에 대한 관심과 체계적인 재난관리 체계에 대한 요구가 크게 증가하였다. 재난과 재해에 철저히 대비하고 재난과 재해가 발생하였을 시 신속한 조치로 사건 발생과 사후 처리에 걸친 전 과정을 효과적으로 통제하기 위한 체계를 갖추는 것은 국민의 생명과 재산을 지키기 위해 존재하는 국가의 중차대한 과제일 것이다. 당시 국내에서는 국가안전처 등을 신설하였고, 현재에도 효과적인 재난 대응을 위한 법적, 제도적 개선 작업이 이루어지고 있어 그 추이가 주목된다.

하지만 헌법상 한국의 국민들이 거주하고 있는 한국의 영토, 휴전선 이북지역에서 발생하는 재난·재해의 대응과 관리는 분단의 현실 아래 한국의 영향력에서 벗어나 있는 것이 사실이다. 남북한 간의 관계는 양 체제의 규범체계상 국가와 국가 간의 관계라고 특정할 수 없으면서도 현실적으로는 어느 한 체제의 국내법과 정치, 군사, 경제, 사회적 영향력이 미치지 않는 특수한 관계이다. 또한 1963년 휴전 이후로 군사 분쟁과 대치를 지속하고 있는 정전체제를 유지하고 있다.

이러한 현실에서 북한체제 아래 실질적 통치를 받고 있는 북한지역의 재난·재해에 까지 관심을 가지는 것은 시기상조 혹은 어불성설이라는 의견이 있을 수도 있다. 그러나 헌법 제 3조가 대한민국의 영토를 한반도와 그 부속도서로 하고 있고

대법원은 1996년 판결에서 북한 주민 역시 한국 국민으로 인정하고 있는 만큼¹⁾ 휴전선 이북에서 일어나 북한 주민들에게 큰 영향을 끼치는 재해와 재난에 대해서도 한국 정부와 국민이 수수방관할 수는 없는 일이다. 이 외에도 인도주의적 차원의 문제와 접경지역에서의 재해와 재난은 휴전선 이남의 한국 영토와 국민에게 피해를 입힐 수도 있다는 점에서 한국의 북한지역 재해·재난에 대한 대비책 구비는 더욱 그 필요성이 강조되고 있다.

본 연구에서는 ① 선행연구 및 문헌조사를 통해 북한의 재난대응 개념과 법적,

1) 대판 1996.11.12, 96누1221에 따르면 북한과 남한 전 국가의 국민인 조선인을 부친으로 하여 출생한 자는 남조선과도정부법률 제11호, ‘국적에 관한 임시조례의 규정’에 따라 조선국적을 취득하였다가 제헌헌법의 공포와 동시에 대한민국 국적을 취득하였다고 할 것이라고 판시하였다. 이에 따라 대한제국시절 출생하여 북한 주민이 된 자에 대해 우리나라 국적을 인정하고 있으며 동 국민을 부모로 둔 북한 주민, 즉 일제강점기와 해방 직후, 또한 북한 정권아래에서 출생한 북한 주민일지라도 국적법 제2조에 따라 대한민국국민을 부모로 하여 출생한 자이기 때문에 대한민국 국민으로 인정한다.

제도적 준비사항에 대한 분석을 실시하여 북한의 재난대응 및 관리체계를 살펴보고

② 남북한 재난대응 협력사례를 참고하여 기존 재난대응 협력체계의 문제점과 법적 제약요소 및 미비점을 도출한 후, ③ 국제적 규범, 현행법, 남북한 간 조약 등을 기반으로 남북한 간의 특수 관계 하에서 한국정부가 북한지역의 재해와 재난에 대해 합법적이고도 효율적인 대응 및 지원을 실시하기 위한 방안을 모색하고자 한다.

2. 연구의 범위 및 방법

연구의 범위를 선정하기 위해 선행적으로 재난과 재해의 개념에 대한 정의를 내리고자 한다. 천재지변이나 인재로 인한 재난(災難)의 피해로 인해 사람이 겪는 피해를 재해(災害)라고 부르는 것이 사전상의 정확한 구분이기는 하나 재난대비 및 관리가 완화(예방)단계(Mitigation & Planning), 대비단계(Preparedness & Planning), 대응단계(Response), 복구단계(Recovery) 등으로 그 발생시점 이전부터 발생시점 이후까지 이루어지는 반면 재해는 재난의 발생으로 말미암아 발생하는 개념이기에 포괄적인 차원의 관리와 대응방안을 논하기에는 ‘재해’, 혹은 재해와 재난을 혼용하는 것 보다는 ‘재난’의 개념이 더욱 적절하다고 할 것이다.

재난의 사전적 의미는 국민의 생명·신체 및 재산과 국가에 피해를 주거나 줄 수 있는 것으로서 태풍·홍수·호우·폭풍·폭설·가뭄·지진·황사·전염병 등 자연현상으로 인하여 발생하는 재해와 화재·붕괴·폭발·교통사고·환경오염사고 등과 같은 사고로써 대통령령이 정하는 규모 이상의 피해를 뜻한다.²⁾ 상기한 대통령령 제25478호, ‘재난 및 안전관리 기본법 시행령’ 제2조³⁾에는 재난의 범위에 대해 국가 또는 지방자치단체 차원의 대처가 필요한 인명 또는 재산의 피해 혹은 이에 준하는 피해를 뜻한다고 명시하고 있다.

따라서 본 연구에서는 북한에서 발생한 위와 같은 규모의 재난과 당시 남북한 간

2) 경찰학사전, (2012. 11 .20, 법문사)

3) 「재난 및 안전관리 기본법」(이하 "법"이라 한다) 제3조제1호나목에서 "대통령령으로 정하는 규모 이상의 피해"란 다음 각 호의 어느 하나에 해당하는 것을 말한다.

1. 국가 또는 지방자치단체 차원의 대처가 필요한 인명 또는 재산의 피해
2. 그 밖에 제1호의 피해에 준하는 것으로서 안전행정부장관이 재난관리를 위하여 필요하다고 인정하는 피해

재난대응사례를 분석할 것이며 동 범위 내에서 한국정부가 휴전선 이북지역에서 발생한 자연재해, 대형사고, 전염병 등의 재난에 효과적으로 대응하고 재난관리에 일정부분 관여하며 인도주의적 지원이 신속하게 이루어 질 수 있는 방법을 모색해 보고자 한다.

남북한 간의 재난대응 및 관리방안에 대해 논하기 위해 먼저 북한의 재난대응 개념과 법적 제도 현황을 간략히 살펴본다. 북한의 재난대응 관련 규범과 북한 재난관리 체계의 의의를 살펴 본 후 이들 규범과 체계에 의해 조직되어지고 운영되는 북한 재난관리·대응 기구의 특징을 분석할 것이다.(II)

이후 북한 내의 재난 시 남북한 재난협력사례를 분석하기 위해 2004년의 룡천역 열차 폭발사고와 2006년의 대규모 수해, 2009년의 신종인플루엔자 확산사태에서의 남북 재난대응협력사례를 분석하여 의의와 문제점을 도출한다.(III)

도출된 문제점을 바탕으로 북한지역에서의 발생하는 재난의 재난관리 및 복구지원을 위한 법적 제도마련에 대해 국제법 및 규제규범화, 국내법적 근거 마련, 남북한 간 합의로 나누어 제안 할 것이다.(IV)

II. 북한의 재난대응 개념과 법적 제도 현황

1. 북한 재난대응 관련 규범

세계 대다수 국가는 재난대응과 관련하여 관련된 법이나 제도 등이 규범 안에 존재한다. 하지만 북한에는 재난관리에 관한 명백한 법이나 체계가 없다.(이금순, 2001, 43) 따라서 이 장에서는 북한의 규범상에서 드러나는 재난대응을 알아보고자 한다. 또한 이를 사전대비와 사후대응의 측면으로 나누어 북한의 재난대응의 체계가 어느 쪽으로 치우쳐있는지 알아보고자 한다.

1) 토지법

북한의 토지법에는 수해와 산불, 병해충 등의 재난재해와 관련된 내용이 명시되어

있지만 모두 사전대비의 성격을 지니고 있다. 전체 제6장80조로 구성된 토지법 중 제17조제4항과 제20조, 제21조, 제22조, 제23조, 제24조, 제28조, 제29조는 홍수 등의 수해를 대비하기 위한 조항이다. 제38조는 산불을 예방하기 위한 것이며 제39조와 제40조는 병풍해 등의 피해로부터 산림을 지키기 위한 법조항이다.

〈표 1〉 북한 토지법 각 조항의 성격 분석

법 조항	내 용	성 격
제17조제4항	국토건설총계획에는 다음과 같은 내용이 포함된다. 4. 강하천, 호소, 저수지의 건설 및 정리방향, 큰물피해를 막기 위한 시설물들의 배치 및 물의 종합적리용대책	사전대비
제20조	강하천정리사업은 큰물피해로부터 농경지를 비롯한 나라의 귀중한 재산을 보호하며 국토의 면모를 개변하는 중요한 사업이다. 국가는 강하천건설을 해당 지대의 자연지리적조건과 특성에 맞게 관개공사와 병행하여 진행하며 큰 강과 중소하천정리를 다같이 밀고나간다.	사전대비
제21조	국토관리기관과 농업지도기관은 강하천정리사업을 설계에 근거하여 계획적으로 전망성있게 조직진행하여야 한다. 큰 강과 중요강하천의 정리와 관리는 국토관리기관이 하며 중소하천의 정리는 해당 기관, 기업소, 협동농장이 한다. 강하천정리는 큰물피해가 심한 중요산업지구, 주민지구, 농경지보호면적이 많은 지대부터 먼저 집중적으로 하여야 한다.	사전대비
제22조	국토관리기관과 지방정권기관을 비롯한 해당 기관, 기업소, 협동농장은 해마다 국가가 정한 시기에 담당하고있는 강하천의 변동정형과 제방, 시설물의 상태 등을 구체적으로 조사등록하고 해당한 대책을 세워야 한다.	사전대비
제23조	국토관리기관은 강하천의 통일적인 보수관리체계를 세우고 강하천보수관리전문기업소의 역할을 높여 보수관리를 전문화, 과학화하도록 하여	사전대비

	야 한다. 사회안전기관은 강하천과 기술상태를 정상적으로 검열하고 해당 대책을 세워야 한다.	
제24조	국토관리기관과 해당 기관, 기업소, 협동농장은 무더기비에도 토지가 류실되지 않도록 강바닥파기와 강줄기바로잡기, 제방공사, 호안공사, 옹벽공사, 모래잡이언제공사를 질적으로 하여야 한다.	사전대비
제28조	농업지도기관과 해당 기업소는 논밭이 물에 잠길수 있는 지대에 고인물빼기시설을 완비하여 그에 대한 유지보수사업을 정상적으로 조직진행하여야 한다.	사전대비
제29조	국영 및 협동농장은 토지가 류실되지 않도록 강기슭의 논밭머리에 버들을 심거나 돌담을 쌓으며 산경사지의 밭머리에는 돌림물길을 만들어야 한다.	사전대비
제38조	국토관리기관은 산불에 대한 감시사업을 강화하며 필요한곳에 산불막이선을 치거나 인원과 설비의 동원체계를 세우는 등 산불로부터 산림을 보호하기 위한 사업을 조직진행하여야 한다.	사전대비
제39조	국토관리기관은 송충을 비롯한 병해충에 의한 산림의 피해를 막을수 있도록 산림을 제때에 소독하며 해로운 벌레를 잡아먹는 유익한 동물을 보호증식하는 등 필요한 사업을 조직 진행하여야 한다.	사전대비
제40조	국토관리기관과 농업지도기관을 비롯한 해당 기관, 기업소는 지대적특성에 맞게 방풍림, 사방림, 위생풍치림, 수원함양림 등 보호림을 조성하며 사방야계구조물을 설치하여 자연재해로부터 국토를 보호하며 나라의 풍치를 아름답게 하여야 한다.	사전대비

*출처 : “북한법령집” 국가정보원, 2017

2) 인민보건법

북한의 인민보건법은 전염병과 관련되어 있으며 토지법의 경우와 같이 사전대비의 성격을 지닌 법조항만을 가지고 있다. 제18조와 제19조, 제27조에서 드러나듯 북한은 인민들이 전염병에 걸리지 않게 노력하겠다고 말하지만 실제로는 전염병에 걸린 책임은 인민들에게 돌아간다. 이는 북한이 자연재해 중 인위적 재해로 분리시켜 그 책임을 인민들에게 돌리는 것과 무관하지 않다.

〈표 2〉 북한 인민보건법 각 조항의 성격 분석

법 조항	내 용	성 격
제18조	국가는 인민들이 병에 걸리지 않도록 미리 대책을 세우는것을 자기 활동의 중요한 임무로 여기며 예방에 선차적인 힘을 넣는다.	사전대비
제19조	보건기관을 비롯한 기관, 기업소, 단체는 위생선전사업과 교양사업을 강화하여 인민들자신이 위생문화사업에 자각적으로 참여하며 과학적으로 건강을 보호 관리하고 질병을 예방하도록 하여야 한다.	사전대비
제27조	국가는 전염병을 미리 막기 위한 방역대책을 철저히 세운다. 보건기관과 기관, 기업소, 단체는 전염병의 발생조건을 없애고 소독사업을 강화하며 주민들에 대한 면역대책을 철저히 세워야 한다. 해당 기관은 다른 나라에서 전염병이 들어오지 못하도록 검역사업을 강화하여야 한다.	사전대비

*출처 : “북한법령집” 국가정보원, 2017

3) 산림법

북한의 산림법은 산림에서 발생하는 화재와 병해충과 관련이 있는데 산림법의 제20조, 제22조, 제23조, 제24조, 제26조는 사전대비의 성격을 지닌다. 이 중 제22조 중 “산불이 일어났을 경우 해당 지역의 인민위원회와 인민보안기관은 조직과 지휘를

바로하여 산불을 제때에 꺼야 한다”라는 내용과 제23조 중 “병해충이 발생한 산림구역에서 나무와 풀, 흙 같은것을 내갈수 없다”라는 내용에서 드러나듯 상기한 토지법과 인민보건법과 달리 사후대응의 성격 또한 지니고 있다.

하지만 사후대응의 두 단계 역시 대응단계와 복구단계 중 산림법의 제22조와 제23조는 대응단계에 속해 복구단계와 관련된 내용은 미비한 것을 알 수 있다.

〈표 3〉 북한 산림법 각 조항의 성격 분석

법 조항	내 용	성 격
제20조	국가는 산불피해를 막기 위하여 산불방지기관을 정한다. 산불방지기관을 정하는 사업은 내각이 한다.	사전대비
제22조	국토환경보호기관, 림업기관과 해당기관, 기업소, 단체는 산불막이선을 치고 잘 관리하며 산불감시체계를 세워 산불이 일어나지 않도록 하여야 한다. 산불이 일어났을 경우 해당 지역의 인민위원회와 인민보안기관은 조직과 지휘를 바로하여 산불을 제때에 꺼야 한다. 산불이 일어난 지역의 기관, 기업소, 단체는 산불끄기에 필요한 인원과 설비, 수단을 의무적으로 동원하여야 한다.	사전대응 + 사후대응
제23조	국토환경보호기관, 림업기관과 기업소, 단체는 산림병해충예찰 및 검역체계를 세우며 발생한 산림병해충을 제때에 없애야 한다. 병해충이 발생한 산림구역에서 나무와 풀, 흙 같은것을 내갈수 없다.	사전대응 + 사후대응
제24조	산림과학연구기관과 해당 기관, 기업소, 단체는 산림병해충을 없애기 위한 효능 높은 농약과 생물학적방법, 현대적인 기술장비 같은 것을 연구 도입하여야 한다.	사전대비
제26조	국토환경보호기관, 림업기관과 해당 기관, 기업소, 단체는 무더기비, 사태 같은 자연피해로부터 산림을 보호할 수 있게 사방야계공사를 계획적으로 하고 그 시설물을 정상적으로 보수정비하여야 한다.	사전대비

*출처 : “북한법령집” 국가정보원, 2017

4) 적십자회법

북한의 적십자회법은 추후 조선적십자회로 개칭된 북조선적십자회가 창설되면서 만들어진 법률로 전술한 법률들과는 달리 사전대비의 성격보다 사후대응의 성격을 지닌 조항들로 이루어져 있다. 제8조의 제2항은 북한 주민들의 질병 예방을 위한 조항으로 사전대비의 성격을 가지고 있는데 적십자회법 제8조의 제1항, 제3항, 제9조는 사후대응의 성격을 지니고 있다.

위에서 살펴본 대로 북한의 재난재해 관련 법령은 대체로 사전대비 측면에 치우쳐 있다. 이는 북한의 재난재해 발생 시 그에 대한 적절한 대처를 할 수 있는 능력이 부족한 것이 큰 원인이다. 물론 사전대비에 따라 피해가 최소화되기 때문에 사후대응보다 사전대비가 중요한 것은 사실이다. 사후대응은 피해를 최소화시키는 작업이지만 사전대비는 피해가 발생하지 않기 때문이다.

〈표 4〉 북한 적십자회법 각 조항의 성격 분석

법 조항	내 용	성 격
제8조제1항	큰물, 태풍, 해일 같은 각종 재난에 대처하기 위한 준비사업을 하며 피해자의 생명을 구원하고 생활을 안정시키며 피해후과를 가시기 위한 활동을 한다.	사전대비 + 사후대응
제8조제2항	적십자회의 임무는 다음과 같다. 2. 인민들의 건강 복리를 증진시키고 전염병을 비롯한 각종 질병을 예방하기 위한 구급의료봉사활동, 보건위생선전사업을 진행하며 국가보건기관의 사업을 협조한다.	사전대비
제8조제3항	전시에는 상병자들에게 의료상방조를 제고하며 인민들에 대한 보호, 구호사업을 한다.	사후대응
제9조	적십자회는 큰 규모의 재난시에 국제적지원을 요청할수 있다. 공화국령역안에서 진행되는 국제적인 적십자협조사업은 적십자회를 통하여 한다.	사후대응

*출처 : “북한법령집” 국가정보원, 2017

2. 북한 재난관리 체계의 의의

재난관리 단계는 두 가지 대분류와 4가지 소분류로 나눌 수 있다. 크게는 사전대비와 사후대응의 단계로 나눌 수 있고, 좁게는 완화와 예방, 대비와 계획, 응급대응, 복구로 나뉜다.(이금순, 2001, 2) 완화와 예방, 대비와 계획 단계가 사전대비에 속하고 응급대응과 복구는 사후대응에 속한다.

북한의 재난관리 체계는 사전대비와 사후대응 중 전자의 특징을 가진다. 북한의 법 중 재난재해관리와 관련이 많지 않다는 점을 주목했을 때있는 재난과 관련이 있는 법으로는 토지법과 인민보건법, 산림법, 적십자회법이 있는데 모든 법에 사전대비 성격을 지니고 있는 조항이 포함되어 있지만 사후대응의 성격을 지닌 조항은 산림법⁴⁾과 적십자회법⁵⁾에서만 나타난다. 또한 산림법에서 드러나는 사후대응 조항은 재해가 발생했을 경우의 행동만 명시하고 있어 재난관리상 4단계 중 3단계인 대응단계까지만 해당된다. 비교적 할 일이 명확한 대응단계와는 달리 복구단계는 인과관계가 복잡하게 얽혀있고 그 해결이 어렵다.(최희천, 2010, 215) 따라서 북한의 재난재해 법령에서 드러나는 사후대응 단계의, 특히 복구단계의 미비는 북한의 재난재해 법령이 가지는 큰 문제점이다.

이 외에도 북한의 재난관리 체계의 특징은 추상적이고 선언적인 규정이 많다는 것이다.(이효원, 2004, 49) 북한의 재난관리 체계는 상기한 법에 따르면 크게 감독사업, 조직지도사업, 지도관리협조의 세 부분으로 나눌 수 있다.(이금순, 2001, 48) 감독사업 부문에서 문제를 해결하기 위한 계획이나 대책을 세우면 조직지도사업부문에서 세워진 계획 혹은 대책을 바탕으로 현장지휘를 한다. 지도관리협조 부문에서는 감독사업부문과 조직지도사업부문의 필요한 정보와 인력, 물자를 제공한다. 하지만 관련법에 이를 수행하는 조직에 대한 언급은 전무하다. 즉, 원론적인 이야기만을 하며 구체적으로 어떠한 조치를 취해야 하는지 명시되어 있지 않다는 것이다. 이는 북한의 관련 법령이 구체적 사안에 대한 적용을 고려하지 않았다는 것을 반증하며, 그들 스스로가 세세한 성문법화의 필요성을 느끼지 못한 결과(윤대규, 2003, 29)라는 분석이 존재한다.

4) 제22조와 제23조.

5) 제8조제1항, 제8조제3항, 제9조.

3. 북한 재난관리·대응 기구의 특징

1) 내각 산하 조직

북한 법에서 드러난 재난관리를 위한 내각 산하 조직은 농업성, 국토환경보호성, 국토관리기관 및 해당 기관이 있다. 내각 산하 조직은 재난대응에 있어서 감독사업, 조직지도사업, 지도관리협조를 각각 담당하고 있다.

북한 내각 산하의 중앙기구는 재난을 직접 담당하는 부서는 존재하지 않으며 각 성에서 해당 기구의 기능의 분야에 들어가는 재난을 총괄할 뿐이다.

농업성은 수해를 피하기 위한 대책을 수립하는 등 농촌지역 재해예방을 위한 사업을 추진한다. 구체적으로는 수해 가능성이 높은 지역의 보수작업과 농경지의 배수로 정비, 저수지나 하천의 정리 사업, 황사 대비 등의 사업을 펼치고 있다.(한국농촌경제연구원, 2006, 43-45)

국토환경보호성은 전국적인 감시체계를 설립하여 전반적인 재해 가능성을 확인하며 이를 해결하기 위한 계획을 세우는 일을 한다.(김정순, 2010, 100) 또한 세운 정책을 집행하는 과정에서 요구되는 설비와 자재를 제시간에 제공하기 위한 대책을 마련해야 하는 임무를 지니고 있다. 또한 국토환경보호성은 국토관리기관과 해당 기관을 조직하여 각각의 부문의 문제를 조사하여 계획을 수립한다.(박경석, 2013, 13-14) 예를 들어 산림 부문에서는 지속적으로 발생하는 수해를 막기 위해 토지법 제30조⁶⁾, 제39조, 제40조에 명시되어 있는 예방 사업을 해당 기관이 실시하도록 명령을 내린다.

2) 군 조직

북한의 행정부 산하 조직들이 재해를 예방하고 재난 피해를 최소화할 계획을 수립하여 명령을 내릴 때, 이를 현장에서 직접 수행하는 것은 군 조직이다. 특히 인민군이

6) 산림건설사업은 토지를 철저히 보호하며 나라의 부강발전과 후손만대의 번영을 위한 만년대계의 대자연개조사업이다. 국가는 토지류실을 방지하며 나라의 자연부원을 늘이기 위한 산림건설사업을 전망성있게 조직진행한다.

직접 움직이는 경우도 있지만 대체로 노동적위군이 이 일을 맡아서 진행한다.

노동적위군은 북한 전 주민 중 조선인민군이나 교도대에 편성되지 않은 17~60세의 남성과 17~30세의 미혼 여성으로 구성된, 북한의 예비군 조직 중 하나이다. 이들은 전시에는 정규군의 지원 및 해당 지역의 방어 임무를 수행하는데 평시에는 민방위대 역할과 주요시설 보수 및 재난복구 등의 일을 한다.(이관세, 2012, 78)

북한은 재난구조 및 복구 활동을 진행하기 위한 대규모 인력동원을 위해 조선인민군을 활용한다. 94년 조선인민군 최고사령관 명령 제0051호⁷⁾에서는 최초로 조선인민군에게 비군사적 임무를 부여한 사례였으며, 이후 조선인민군은 재난재해를 겪는 상황에서 인력이 필요한 모든 사업에 동원되었다.(NK조선, 01/08/13)

3) 조선적십자회 재난대처센터

조선적십자회의 원 설립목적은 '근로자들의 자원적이며 대중적인 보건위생 및 구호단체'로 규정되어 조선노동당의 외곽단체로서의 구실 또한 충실히 수행하고 있다. 과거의 북한에는 재난관리라는 개념 자체가 존재하지 않았다. 하지만 1990년대 중반부터 발생한 잦은 자연재해 이후로 임시적인 대책이 마련되기 시작했고 1999년 국제적십자연맹의 후원을 받아 재난대처센터가 개설되면서 재난관리·대응 체계가 재정비되었다.(명수정 외, 2008, 64) 이에 따라 조선적십자회는 북한 인민의 구호나 응급서비스의 증진 활동을 펼치는 등의 일을 수행하게 되었다.

Ⅲ. 남북 재난대응 협력사례

1. 2004년 룡천역 열차 폭발사고

7) 조선인민군 최고사령관 명령 제0051호, “평양시에 청류다리(2단계)와 금릉2동굴을 건설할데 대하여” 94.11.09

1) 사건 개요

2004년 4월 22일 오후 12시가 조금 넘은 시간 평안북도 룡천군 룡천역에서 폭발이 일어났다. 질산암모늄을 싣고 있던 열차가 부주의로 고압선에 접촉하면서 열차가 폭발했고(서동렬 외, 2004, 138) 역 주변에 있던 소학교와 문화회관이 완전히 무너져 내렸다. 또한 8,000여 채의 일반 주택이 반파되거나 완파되면서 룡천역을 중심으로 반경 4킬로미터 이내는 완전히 폐허가 되었다.(손종도, 2012, 70) 사건 초기 북한은 이 사고를 일절 보도하지 않았었다.(북한연구소, 2004, 170) 하지만 이틀 후인 24일 이를 공식발표하고 그 이후에는 조선중앙통신을 통해 사고가 어떻게 진행되고 있는지 보도하는 등 공식적으로 여러 나라와 국제사회의 인도적 지원을 바란다고 밝혔다.(손종도, 2012, 70)

2) 진행 과정

정부에서는 북한이 구호 공식 요청이 들어온다면 지원방안을 검토할 수 있을 것이라고 발표하였다. 북한이 공개적으로 여러 나라와 국제기구에 원조를 요청한 이후에 당시 여당이었던 열린우리당은 물론 당시 한나라당 또한 남북협력기금의 사용승인 요청이 들어오면 신속하게 처리하겠다고 룡천 지역을 지원해야 한다는 공통된 의견을 가지고 있었다.(장용훈, 2004, 40)

4월 24일 정식으로 지원 요청이 들어오면서 정부 중앙청사에서는 북한 룡천역 열차 폭발사고에 따른 재해대책 관계장관회의가 소집되었다.(장용훈, 2004, 39) 27일에는 개성에 있는 자남산 여관에서 재해복구 지원을 위한 ‘긴급구호 회담’을 개최하여 북한에 1차적으로 1백만 달러 상당의 구호물자를 전달하기로 하였다.(평화문제연구소, 2004, 4) 이 외에도 29일 국정현안 정책조정 회의에서 북측 요청 물량을 가급적 전량지원하기로 결정하였다. 또한 직접적인 지원 외에도 국제사회의 대북지원활동 참가와 관련하여 WHO(세계보건기구)에 20만 달러를 지원하고 민간차원에서 모인 구호물품 수송비용을 지원하는 등 정부는 약 700억 원 가량의 지원을 펼쳤다.

정부가 약 700억 원 가량의 지원을 펼친 결과(북한연구소, 2006, 215)로 같은

연도 9월 7일 피해지역에 1,650가구의 살림집이 건설되었고(북한연구소, 2004, 4) 사고 후 1년 만에 복구 공정을 80% 완료하는 등 빠른 재건이 이루어졌다. 이는 북한의 빠른 구호 요청과 정부의 신속하고 적절한 구호지원이 일궈낸 결과라고 할 수 있다.

〈표 5〉 룡천역 사고 당시 대북지원

구 분	지 원 품 목
1차 지원	40만 달러 상당의 생필품(누비이불, 속옷, 수건, 비누 등의 생활용품, 컵라면, 생수, 담요, 운동복 등)
	40만 달러 상당의 의약품
	20만 달러 상당의 WHO 지원비
NGO 지원	수송비 지원
북한 피해 입은 장비 복구	수송비 포함 약 310억 원 추정

3) 문제점

민간 차원에서도 역시 북한에 대한 지원에 열을 올렸다. 하지만 이런 긍정적인 지원 과정에도 문제점은 발견되었다. 첫째, 잦은 중복지원으로 인한 대북지원의 비효율성이다. 당시의 기준으로 현재 무엇이 필요한지, 그리고 무엇을 지원받았고 무엇을 지원받지 못했는지를 알 수 없었던 만큼 같은 물품을 중복 지원하게 되는 경우가 발생했다.

2004년 4월 29일 북한의 조선민족경제협력연합회 단동사무소 전성근 대표는 “담요는 이미 5만장 이상 지원됐다”, “룡천지역 재건과 재활에 초점을 맞춰 장기적인 지원사업을 계획해 달라”고 하며(세계일보, 04/05/03) 대북지원의 비효율성 문제를 지적하였다.

두 번째 문제점은 대북 지원 과정에서 법이 지켜지지 않았다는 점이다. 당시 남북교류협력에관한법률 제9조제3항⁸⁾과 남북교류협력에관한법률시행령 제19조⁹⁾에 따르

8) 남한의 주민이 북한의 주민등과 회합·통신 기타의 방법으로 접촉하고자 할 때에는 통일원장관의 승인을 얻어야 한다.

9) 법 제9조제3항의 규정에 의한 승인을 얻고자 하는 자는 접촉 15일전까지 다음 각호의

면 우리 국민이 북한 주민과 접촉하기 15일 전에 승인을 받아야 했다. 하지만 민간단체들이 경쟁적으로 사고수습을 위한 지원에 나서면서 접촉 며칠 전에 승인을 받거나 승인을 받지 않은 채로 중국 단둥을 거쳐 북측에 구호물자를 전달하는 민간단체들이 다수 나타났다. 당시 남북교류협력에관한법률 제27조제1항¹⁰⁾에 따르면 이들은 3년 이하의 징역 또는 1,000만 원 이하의 벌금형을 받아야 했다. 하지만 이들 중 처벌받은 사람은 없었다.(세계일보, 04/06/15) 따라서 효율적 집행을 위해 남북교류협력에관한법률 제9조제3항과 남북교류협력에관한법률시행령 제19조는 조건이 완화되어야 할 필요가 있다.

2. 2006년 대규모 수해

1) 사건 개요

2006년 7월 16일 북한의 조선중앙TV는 “대동강에 16년 만에 큰물이 났다”고 보도하였다. 당시 보도와 권태진(2006, 4)에 따르면, 10일부터 16일 사이에 평양(227mm), 양덕(485mm), 평강(382mm), 개성(366mm), 맹산(301mm), 원산(300mm), 북창(260mm) 등 북한에 폭우가 내렸는데 주로 14일과 15일 사이에 집중되었다. 또한 태풍 에위니아의 영향을 받아 더욱 커진 이 폭우는 대동강과 예성강 상류 지역을 중심으로 내렸기 때문에 이 하천 범람으로 인한 하천 주변 지역의 피해가 극심했다. 이로 인해 주변 농경지와 농경시설이 파괴되었고 따라서 농작물 손실 또한 약 2만 2천 톤으로 추정되었다. 뿐만 아니라 이번 피해로 감소될 것으로 예상되는 쌀 생산량이 약 50,000톤으로 추정되었고 추가피해가 발생할 가능성과 복구에 시간이 걸린다는 점에서 감소 규모가 10만 톤에 이를 수도 있었다.

서류를 통일부장관에게 제출하여야 한다...(후략)

10) 다음 각 호의 1에 해당하는 자는 3년이하의 징역 또는 1,000만원이하의 벌금에 처한다. 1. 제9조제1항의 규정에 의한 증명서를 발급받지 아니하고 남한과 북한을 왕래하거나 동조제3항의 규정에 의한 승인을 얻지 아니하고 회합·통신 기타의 방법으로 북한의 주민과 접촉한 자...(후략)

2) 진행 과정

같은 연도 8월 9일 6·15공동선언실천 북측 위원회에서 남측 위원회로 건설자재와 건설장비, 쌀, 담요, 의약품 등을 지원해달라는 요청이 들어왔다.(송영대, 2006) 14일 정부는 전화통지문을 통해 북한에 위로의 뜻을 전했고 남북적십자 실무접촉을 제안했다. 북한이 이를 수락하여 19일 금강산에서 남북적십자 실무접촉이 이루어졌다. 이 접촉에서 남북은 쌀 10만 톤과 복구자재 장비 등과 복구가 어느 정도 이루어진 뒤 현장 방문을 통해 모니터링을 실시하기로 합의를 하였다. 이후 우리나라는 쌀 8만 9500톤과 시멘트 2만 9585톤, 철근 3800톤, 덤프트럭 50대, 굴착기 50대, 페이로더 60대, 모포 2만 장, 긴급구호세트 1만 개, 의약품 등을 지원했다.(남북회담본부, 2006)

〈표 6〉 2006년 대규모 수해 당시 대북지원

품목	예정 지원량	실제 지원량
쌀	100,000t	89,500t
시멘트	100,000t	29,585t
철근	5,000t	3,800t
트럭	100대	50대
굴삭기	50대	50대
페이로더	60대	60대
모포	80,000장	20,000장
응급구호세트 및 의약품	10,000개	10,000개

3) 문제점

2006년 북한 수해 지원에서는 큰 문제점이 발견되었는데, 인도적 대북지원의 일관성이 지켜지지 않았다는 점이다. 정부는 7월 5일 북한이 미사일을 발사한 것을 이유로 북한에게 추가적인 쌀·비료 지원을 하지 않겠다고 발표했다. 정부는 북한에게 쌀과

비료를 남북협력기금을 사용하여 지원하고 있었는데 당시 남북협력기금법 제7조 1항에 따르면 남북협력기금은 통일부장관이 운용하고 관리하게 되어있었다. 북한의 미사일 발사 때문이라고는 하지만 인도적 지원을 아무런 절차 없이 끝냈던 것이다. 그런데 북한이 집중호우로 인한 홍수로 큰 피해를 입으면서 다시 지원을 하게 되었다.

이후 북한의 핵실험 이후 다시 지원은 취소되었다. 당시 남북협력기금은 남북협력기금법 제7조제4항¹¹⁾에 명시되어 있는 절차에 따라 운용되었을 텐데 정부의 입장에 따라 빈번하게 기금 운용에 변동이 생긴다는 것은 위의 절차만으로는 인도적 대북지원의 일관성을 지킬 수 없었다는 것이다. 최소한의 인도적 지원이 유지되었어야 한다. 물론 북한이 국제적 도발을 감행한 것은 사실이므로 이전같이 지원하는 것은 옳바르지 않다. 하지만 최소한의 인도적인 지원은 계속 되었어야 했다.

3. 2009년 신종인플루엔자 확산사태

1) 사건 개요

2009년 4월 29일 북한은 조선중앙통신을 통해서 ‘멕시코 돼지독감’, ‘신형 돌림감기’ 등의 용어를 사용하면서 신종 인플루엔자 피해소식을 보도했다. 또한 계속적으로 신종 인플루엔자 바이러스를 연구하고 있다는 점과 전염을 막기 위해 노력하고 있다는 점을 보도했다. 그렇게 5월 19일 조선중앙통신에 “H1N1 돌림감기에 대한 방역체계가 정연하게 수립되었다”라는 소식이 들렸다.(강석운, 2009, 167)

하지만 조선중앙통신의 보도가 무색하게 같은 연도 11월부터 평안북도 신의주를 포함한 북한 전 지역에 신종 인플루엔자가 퍼졌고, 12월 9일 북한은 9명의 감염자와 수십 명의 사망자가 발생했다고 발표했다.(황나미, 2010, 76)

2) 진행 과정

북한의 공식 발표 전인 8일 이명박 전 대통령은 “북한에 최근 신종플루가 발생했다

11) 기금의 운용·관리에 관한 다음의 사항에 대하여는 남북교류협력에관한법률 제4조의 규정에 의한 남북교류협력추진협의회 심의를 거쳐야 한다.

는 소식이 있는 만큼 사실관계를 확인해 도와줄 수 있는 방안을 모색하라”며 지시했다. 이후 9일 북한의 공식 발표 이후 사안이 시급하다고 판단한 정부는 당시 남북 관계가 상당히 경색되어 있었음에도 불구하고 신종 인플루엔자 치료제를 약 50만 명분(약 178억 원어치)을 지원하겠다고 발표했다. 10일 북한에서는 정부의 지원의사를 수용한다는 입장을 통보했다.(황나미, 2010, 76) 그리고 18일 판문점을 통해서 신종 인플루엔자 치료제 50만 명분이 북한으로 전달되었다. 또한 22일에는 신종 인플루엔자를 예방하기 위한 손 소독제도 북한으로 보내졌다.

〈표 7〉 2009 대북 신종 인플루엔자 치료제 지원 현황

구 분	지 원 양	가 격
신종 인플루엔자 치료제	50만 명분	약 168억 원
손 소독제	20만 ℓ	약 10억 원

3) 문제점

정부의 대북 신종 인플루엔자 치료제 지원은 속전속결로 이루어졌다. 하지만 신속하게 진행되어야 한다는 기조 아래에서 일을 조급하게 처리하면서 문제점이 발생했다. 바로 남북협력기금 운용의 투명성 문제이다. 정부는 북한에 지원할 치료제의 양을 50만 명분으로 결정했다. 하지만 어째서 50만 명분을 지원하기로 한 것인지 근거를 제시하지 못했다. 치료제 지원으로 소모된 178억 원은 남북협력기금에서 나온 금액으로, 세금을 사용하는 일이었던 만큼 그 사용에 있어서 합당한 근거가 필요했다. 하지만 정부는 신속한 지원만을 추구한 나머지 북한에 발병현황 자료 등을 요구하지 않았다.(연합뉴스, 09/12/17) 즉, 정부는 북한 주민 중 몇 명이 감염되었고 몇 명이 감염 징후를 보이는지 알 수 없었다는 것이다.

그런데도 정부는 신종 인플루엔자 치료제 50만 명분을 지원하기로 결정했다. 12월 17일 대북 지원계획을 밝힌 브리핑에서 이종주 당시 통일부 부대변인은 “(북한의) 신종플루 환자의 규모를 정확하게 얘기하기 어렵다”고 말했다. 이미 치료제의 규모를 밝힌 상황에서 17일 위와 같은 답변이 나왔다는 것은 북한의 신종 인플루엔자 감염추이도 모르는 상태에서 50만 명분이라는 어마어마한 양을 지원한 것이 된다. 치료제

지원의 규모가 합당한지 증명되지 않은 채 국민의 세금이 178억 원이 사용되었다. 위와 같은 사례를 막기 위해서는 남북협력기금 운용에 관한 심의가 보다 엄격하게 시행될 필요가 있었다. 따라서 남북교류협력에 관한 법률 제7조제2항¹²⁾의 심의통과 조건을 보다 엄격하게 강화하거나 새로운 법 조항을 신설하여 추가적인 심의 절차를 만드는 것이 필요했다. 물론 절차의 엄격성과 지원의 적시성은 따로 고려되어야 할 필요가 있다.

또한 이 문제점은 법적인 접근 외에 제도적인 접근법으로도 바라볼 수 있다. 만약 정부에서 북한의 신종 인플루엔자 감염 추이를 묻는 것을 잊어버렸다고 해도 정해진 제도에 필요한 자료들을 받아야만 진행이 되는 절차가 있었다면 정부가 북한 주민의 신종 인플루엔자 감염 추이나 감염자 수를 모르는 일은 없었을 것이다.

4. 소 결

룡천역 열차 폭발사고 지원과 2006년 수해 지원, 그리고 2009년 대북 신종 인플루엔자 치료제 지원을 살펴보면서 각각의 지원 과정과 문제점을 살펴보았다. 각 사례에서 발견된 문제점은 각각 법에 개선이 필요한 경우와 제도의 미비로 나눌 수 있다. 또한, 법에 개선이 필요한 경우는 또 법 조항의 조건을 완화하거나 강화하는 것으로 나눌 수 있다.

먼저, 법에 개선이 필요한 경우에는 룡천역 열차 폭발사고 지원 과정에서 법이 잘 지켜지지 않았던 것과 2006년 북한 수해 지원 과정에서 발견된 남북협력기금 운용의 잦은 변동, 마지막으로 2009년 대북 신종 인플루엔자 치료제 지원 사례의 근거 없는 남북협력기금 운용이 있었다.

이 중, 법 조항의 조건을 완화해야 하는 경우는 룡천역 열차 폭발사고 지원 과정에서 범법행위가 당연시되는 것으로 당시 남북교류협력에관한법률시행령 제19조의 ‘15일 조건’이 현실에 맞지 않는다는 것이 드러났기 때문에 그 조건을 느슨하게 하게 해야 했다.

법 조항의 조건을 강화해야 하는 경우에는 2006년과 2009년의 대북 지원 과정에서 드러난 남북협력기금의 운용문제이다. 2006년 북한 수해 지원의 사례에서 남북협력

12) 협의회의 회의는 재적위원 과반수의 출석과 출석위원 과반수의 찬성으로 의결한다.

기금의 운용이 남북교류협력에관한법률 제4조의 규정에 의한 남북교류협력추진협의회 심의를 거치고 있었음에도 불구하고 운용의 변동이 잦았기 때문에 심의 절차를 더욱 엄격하게 해야 했다. 또, 2009년 대북 신종 인플루엔자 치료제 지원의 경우엔 타당한 이유가 없음에도 남북교류협력기금이 운용되었으므로 기금의 운용을 심사하는 절차가 더욱 엄격하게 실시되어야 했다.

법 이외에 제도의 미비로 발생한 문제점도 있는데 첫째는 룡천역 열차폭발 사고 지원 과정에서 밝혀진 중복 지원으로 인한 대북 지원의 비효율성이다. 남북 간에 북한이 무엇을 필요로 하고 우리가 지원한 품목이 무엇인지, 그 품목을 얼마나 북측으로 보냈는지 수시로 확인할 수 있는 제도가 있었다면 중복 지원으로 인한 대북 지원의 비효율성은 크게 감소했을 것이다.

제도의 미비로 인한 문제점 둘째는 2009년 대북 신종 인플루엔자 치료제 지원 과정에서 찾을 수 있다. 바로 북한에 어떤 품목을 지원하기로 결정하기 전에 북한에 해당 품목이 필요한지 확인할 수 있는 제도가 없었다는 점이다. 실제로 이와 같은 제도가 존재했다라면 북한으로 보낸 50만 명분의 신종 인플루엔자 치료제가 합당한 이유로 설정된 것이라 국민들에게 설득할 수 있었을 것이다. 또한 남북적십자회담 등의 연구소로 바뀌준다.

상기한 세 가지 문제점 중 현재는 수정이 된 것이 있다. 바로 남북교류협력에관한법률시행령 제16조제1항은 “법 제9조의 제1항 본문에 따라 미리 신고하려는 남한의 주민은 접촉 7일 전까지 북한주민접촉 신고서에 다음 각 호의 서류를 첨부하여 통일부장관에게 제출해야 한다”로 개정되어 기존의 문제점이었던 ‘15일 조건’이 ‘7일 조건’으로 완화된 것을 알 수 있다. 또한 제도가 미비했던 것으로 분류된 문제점은 남북간 직통전화(Hot Line)을 사용하는 것으로 문제점을 해결할 수 있다.

하지만 남북협력기금법 제7조제4항은 “기금의 운용·관리에 관한 다음 각 호의 사항은 「남북교류협력에 관한 법률」 제4조에 따른 남북교류협력 추진협의회 심의를 거쳐야 한다”로 심의를 보다 엄격하게 진행하거나 추가적인 절차를 걸치는 등의 매뉴얼은 만들어지지 않았다. 이 문제가 해결되지 않는다면 차후 유사사례 발생 시에도 같은 문제가 반복될 수밖에 없다.

〈표 8〉 남북 재난대응 교류협력 사례에서 발견된 문제점과 개선 여부

구분		문제점	법 조항	개선 여부
개선 필요	완화	법의 구속력 약화	당시 남북교류 협력에관한 법률시행령 제19조	개선됨
	강화	남북협력기금 운용 기준의 잦은 변동	남북교류협력에관 한법률 제4조	개선되지 않음
		적절한 이유 없는 남북협력기금 운용	남북교류협력에관 한법률 제4조	개선되지 않음
제도미비		지원품목과 필요품목 알 수 없음	X	개선되지 않음

IV. 북한 재난관리 및 복구지원을 위한 법적 제도 마련

지금까지 북한의 재난관리·복구 체계를 살펴보며 문제점을 찾아보았고 그 중 개선된 것도 존재하지만 개선되지 않고 여전히 문제점을 가지고 있는 부분도 존재하고 있다는 것을 알게 되었다. 이 문제점이 해결되지 않는다면 미래에도 같은 문제가 동일하게 발생할 수 있기 때문에 북한의 재난관리 및 복구 체계를 향상시키기 위한 법적 제도 마련이 시급하다. 그렇기 때문에 필요한 법적 제도를 각각 국제법 및 국제기구, 국내법, 남북합의서의 측면에서 알아보는 것은 중요하다.

1. 국제기구와의 협력을 통한 국제규범화

북한은 심각한 경제적 난관과 식량난이 가중되면서 상당한 어려움을 겪고 있다. 이러한 어려움을 극복하기 위해서 북한은 대내적인 노력과 대외적인 노력을 병행하고 있다. 특히 국제사회의 원조를 받기 위해 대외적인 노력을 통해서 정부간국제기구와 국제비정부기구 등의 국제기구와의 접촉을 늘려가고 있다.(박재영, 1999, 91)

북한이 다른 지원주체보다 국제기구 및 개발원조 기구를 선호하고 교류를 지속하고 있는 이유는 지원의 규모는 비교적 작은 편이면서 보다 실질적인 도움을 줄 수 있기 때문이다. 이와 같은 개발원조 기구와의 교류는 북한 체제에 미치는 악영향이 거의 없다고 판단하기 때문이다.(박재영, 1999, 94)

북한과 잦은 교류를 가지는 개발원조 기구 중 북한의 재난관리와 복구지원을 지원하는 국제기구로는 국제연합아동기금(UNICEF), 세계보건기구(WHO), 국제연합식량농업기구(FAO), 유엔개발계획(UNDP), 국제적십자사연맹(IFRC), 국제적십자위원회(ICRC), 유엔세계식량계획(WFP), 유엔인구기금(UNFPA), 유엔환경계획(UNEP), 녹색기후기금(GCF) 등이 있다 이 중 북한의 질병예방에 도움을 주는 기구는 UNICEF, WHO, IFRC, ICRC, UNFPA이며 긴급상황에서 식량을 지원해주는 기구는 FAO와 WFP가 있다. 이 외에 FAO와 UNDP는 북한의 복구 건설에 힘을 쓰며 UNEP와 GCF는 북한의 자연재해 예방 프로그램을 진행하고 있다.

위의 국제기구들은 북한과 지속적으로 교류해오면서 북한에 대한 영향력을 높여왔다. 북한에서 원조국이나 국제기구에 모니터링을 허용하는 것은 이례적이라고 할 수 있다.(이원웅, 2009, 17) 그런데 WFP는 이미 현장에서 식량 분배 모니터링을 실시해왔고 다른 국제기구들도 2일 전에 신청을 하면 특별한 사유가 없는 이상 북한에서 모니터링을 허용했다.(이종운 외, 2012, 10-13) 이렇게 북한에 대한 국제기구의 영향력이 높아지고 있는 상황에서 대북 지원을 국제기구와의 협력을 통해 진행하는 것은 북한 재난관리 및 복구지원을 하는 데에 있어서 큰 도움이 될 수 있다.

〈표 9〉 북한과 교류하고 있는 주요 국제기구 목록

기구명	지원 내용
국제아동기금(UNICEF)	전염병 예방
세계보건기구(WHO)	
국제적십자사연맹(IFRC)	
국제적십자위원회(ICRC)	
유엔인구기금(UNFPA)	
유엔세계식량계획(WFP)	식량 지원
국제연합식량농업기구(FAO)	식량 지원 + 복구 건설
유엔개발계획(UNDP)	복구건설
유엔환경계획(UNEP)	자연재해 예방
녹색기후기금(GCF)	

실제로 정부가 북한과 교류하고 있는 국제기구를 통해서 지원을 한 사례가 존재한다. 1995년 정부는 세계기상기구(WMO)를 통해서 북한의 재난복구를 도와주었다.(최은석, 2010, 40) 당시 북한은 집중호우로 인해 기상예측장비가 망가져 기상예측장비를 쓸 수 없었다. 정부는 남북협력기금을 사용하여 5만 달러를 WMO에 지원했고, WMO는 이를 북한의 기상예측장비 복구에 활용하였다.

국제기구를 통해 북한을 지원하는 방법 외에도 국제인도법의 측면도 존재한다. 국제인도법은 전쟁 등의 국내외적 무력 충돌이 발생했을 때 인도주의 정신을 바탕으로 무력충돌 희생자를 보호하고 그 피해를 최소화하기 위해 성립된 국제법이다.(대한적십자사, 2009, 2) 위와 같은 국제기구와 국제법을 활용한 대북재난대응협력은 한국의 입장에서 지원의 법적근거 마련과 인도주의적 지원이라는 정당성 확보라는 측면에서 유의미하며, 북한의 입장에서도 한국으로부터 직접 지원을 받는다는 것에 대한 부담감을 감소시켜 지원과 협력의 과정에서 불필요한 갈등의 요소가 줄어드는 효과를 얻을 수 있을 것이라 판단된다.

2. 국내법적 제도 마련

하지만 실질적인 법적제도의 마련을 위해서는 국제법적 근거만이 아니라, 대북재난 대응 및 지원에 관한 국내법에서 드러나는 법적 모순 또는 미비점을 찾아 고치는 것도 필요하다. 2006년 북한 수해 지원의 경우와 2009년 대북 신종 인플루엔자 치료제 지원의 사례에서는 남북협력기금 운용의 심의절차가 부족하다는 공통된 문제점이 있었다. 남북교류협력에관한법률 제4조에 따르면 남북협력기금의 운용은 남북교류협력 추진위원회의 심의를 거쳐야 한다. 하지만 2006년에는 정부의 입장 발표에 따라 남북협력기금 운용에 잦은 변동이 있었다. 2009년에는 북한에게 신종 인플루엔자 치료제를 지원하기로 했는데 북한의 신종 인플루엔자 감염 추이 등을 인지하지도 못한 상황에서 신종 인플루엔자 치료제 50만 명분을 지원하기로 계획했고 심의를 통과하여 남북협력기금에서 약 178억 원이 불분명한 이유로 사용되었다.

따라서 심의과정을 명백하게 드러내고 심의절차를 엄격하게 진행하여 남북협력기금 운용의 방향이 자주 변경되는 경우가 없어야 하며 타당하지 않은 근거로는 남북협력기금의 운용을 할 수 없게 해야 한다.

〈표 10〉 남북교류협력에관한법률 제26조에서 준용되는 법령

법 조항	준용 법령
제1항	대외무역법
제2항	관세법
제3항	외국환거래법, 외국인투자 촉진법, 한국수출입은행법, 무역보험법, 대외경제협력기금법, 법인세법, 소득세법, 조세특례제한법, 수출용원재료에 대한 관세 등 환급에 관한 특례법

남북협력기금 운용의 심의 문제 외에도 국내법적으로 다른 문제가 존재한다. 남북교류협력에관한법률 제26조제1항과 제2항, 제3항에 따르면 법의 적용에 있어서 대외무역법과 외국환관리법, 외국인투자 촉진법, 한국수출입은행법, 무역보험법, 대외경제협력기금법, 수출용원재료에 대한 관세 등 환급에 관한 특례법을 준용하고 있다.

하지만 남북관계 발전에 관한 법률 제3조제1항¹³⁾과 제2항¹⁴⁾에 따르면 남북의 관계는 국가 간의 관계가 아닌 통일을 지향하는 과정에서 잠정적으로 형성되는 특수 관계이며 남북 간의 거래는 국가 간의 거래가 아닌 민족 내부의 거래이다. 그런데 남북교류 협력에관한법률 제26조에 따라 민족 내부의 거래에 대외무역에 적용되는 법률을 준용하는 것은 자칫 북한을 국가로 인정한다는 것으로 해석될 여지가 있다.(박건우, 2011, 47-48) 이를 해결하기 위해선 남북 간의 거래에 기존처럼 대외무역 관련법을 적용하지 않고 국내거래 관련법을 활용하는 것이 필요할 것이다.

3. 남북한 간 합의사항 이행 및 제도화

국제법적, 국내법적 근거 마련과 함께 남북간의 합의사항을 이행하는 것도 중요하다. 1992년 남북기본합의서는 남북한관계를 '쌍방 사이의 관계가 나라와 나라 사이의 관계가 아닌 통일을 지향하는 과정에서 잠정적으로 형성되는 특수관계'라고 밝히고 있다. 물론 남북기본합의서의 법적 성격과 효력에 대하여 논란이 있었으나, 국내법적으로는 헌법재판소와 대법원은 그 법적 효력을 인정하지 않고 단지 '정치적 공동성명' 내지 '신사협정'에 불과¹⁵⁾하다고 규정하였다. 따라서 남북기본합의서는 국가 간의 조약으로는 인정되지 않아 법제도적 요건을 구성하지는 못하였지만 남북한의 분단상황을 현실적으로 인정하는 바탕 위에 경제사회문화분야의 교류협력을 추진하는 기본 원칙과 내용이 담겨 있어(이효원, 2011, 4) 향후 재난대응 및 교류협력의 대원칙으로 삼을 수 있을 것이다.

또한 2000년대 개성공단사업과 금강산관광사업 등으로 확대된 남북간의 교류협력 과정에서 다양한 합의문과 관련법령을 제정하는 등 통일법제적 요소를 가진 법률을 구비하였다.(이효원, 2011, 3) 개성공단 관련 법제에는 열차운행, 투자보장, 왕래자의 휴대물품, 통신, 통관, 회계, 외화, 자동차, 부동산, 세관, 보험 등 다양한 부분에 대한 세밀한 법제화가 이루어졌고 개성공업지구와 금강산관광지구의 출입 및 체류에 관한 합의서(2007) 등의 합의도 이루어졌다. 현재 개성공단과 금강산관광지구의 폐쇄로

13) 남한과 북한의 관계는 국가간의 관계가 아닌 통일을 지향하는 과정에서 잠정적으로 형성되는 특수관계이다.

14) 남한과 북한간의 거래는 국가간의 거래가 아닌 민족내부의 거래로 본다.

15) 현재 92헌바6결정, 대법원 98두14525판결

인해 유명무실해지긴 했지만, 위와 같은 제도화 경험은 남북 재난대응 및 협력을 법제화 하는 데 유의미한 자산으로 활용될 수 있을 것이다.

특히 새로운 합의도 중요하지만, 이미 체결된 합의서에 대해서도 명확한 기준을 정하여 구체적인 분류를 한 후 법제화를 추진해야 할 것이다.(한명섭, 2011, 110) 또한 이러한 남북교류에 관한 국내법제화에 있어 다양한 남북합의의 법적 구속력 부여에 대한 문제는 남북 재난대응과 협력뿐만이 아니라, 향후 남북통일에 관한 남한의 법제 정비의 가장 핵심적인 과제라는 지적이 존재한다.(박건우, 2011, 62) 최근 남북한 간 정상회담으로 도출된 ‘판문점 선언’(18년 4월 27일)에 대한 국회인준 필요성에 대한 논의는 남북합의의 국내법제화에 대한 논의에 다름 아니며, 남북합의를 근거로 한 다양한 교류와 협력의 국내법적 제도화 노력으로 평가된다 할 수 있다.

IV. 결 론

만성적인 식량부족 및 경제난과 함께 국제제재를 오랫동안 받은 북한은 재난에 대응한 예비자원에 대한 준비가 매우 취약할 수밖에 없다. 또한 고지대의 산이 많고 정비가 미흡한 하천과 유역에 수자원이 집중되어있는 북한은 매년 수해에 노출되어 있으며 하천 준설과 유역관리에 대한 시도가 미흡하기에 예상치 못한 규모의 호우나 태풍에 큰 피해를 입을 가능성이 크다.

자연적 여건과 체제의 특성이 맞물려 어느 국가나 체제보다 재난에 취약하지만 개혁과 개방에 소극적인 북한과의 재난 대응 교류협력은 북한의 제도나 개방을 요하는 재난의 예방보다는 발생 후 관리와 복구 쪽이 비교적 가능성이 높다 할 것이다. 이후 관련 교류협력이 증대되고 남북한 간의 신뢰가 형성 될 시 임진강 유역 등 남북한이 분할하여 점유하는 강줄기에 대한 유역관리 및 수해방지, 북한지역에서의 상류 댐 방류 시 협조체제 구축, 군사분계선과 해상군사분계선(NLL)에서의 재난, 해난 발생 시 공동대책방안 수립 등 실질적인 남북한 재난관리 교류협력으로 발전 할 수 있을 것이라고 예상된다.

이를 위해서는 남북한 간 실질적인 재난대응협력과 제도화가 중요할 것이다. 그러

나 현재의 남북 간의 경제격차나 재난대응 역량으로 미루어 볼 때 북한과의 교류협력, 특히 북한지역에서 발생한 재난대응 및 관리에 관한 교류협력은 한국의 일방적인 지원에 머무를 수밖에 없다. 또한 최근 북한이 비핵화를 위한 조치를 행하고는 있다고 하나, 핵무장을 완성한 이러한 상황에서 교류협력은 시기상조, 혹은 어불성설이라는 의견이 존재할 수 있다.

하지만 재난의 관리와 인도주의적 지원은 시의적절하고 충분해야만 인명피해 등의 피해를 최소화하는 소기의 목적을 달성할 수 있는 만큼, 효율적이고도 신속한 개입이 요구된다. 더욱이 민족통일을 헌법에 명시하고 국가적 과업으로 삼는 대한민국은 북한 지역에서의 재난에 개입해야할 명분과 의무가 있다. 대규모 재난이 발생할 경우 잠재적인 대한민국 국민인 북한 주민들에 대한 동포애적·인류애적 지원은 필요할 것이다. 또한 접경지역에서의 대규모 재난은 한국에도 영향을 미칠 수 있기에 재난대응에 관한 협력은 중요하다 할 것이다.

본 연구에서는 효과적인 남북한 재난대응 협력을 위하여 현재 남북교류협력법을 비롯한 기본 법제도적 정비가 필요하다고 분석하였다. 현재의 남북 교류협력 관련 법령들은 북한지역의 재난 발생 시 적시적이고 효과적인 대응을 하는데 걸림돌이 되는 과한 제제를 가하고 있거나 혹은 지원물자나 품목에 대한 심의규정이나 절차가 정립되어있지 않아 과도한 지원이 발생하는 등 지원품목과 규모에 대한 행정기관의 자의적 판단이 가능하게 한 면이 과거 북한지역 재난관리 협력 사례에서 발견되었다.

이에 대한 개선책으로는 첫째, 북한이 체제에 대한 제제가 비교적 약하고 규모가 큰 지원이 가능하다는 점에서 재난 발생 시 국제기구의 도움을 직·간접적으로 요구하는 경우가 많으므로 북한지역에서의 재난발생 시 국제기구와의 협력을 통해 북한 지역의 재난에 공동개입하는 협력체제를 구축해야 할 것이다. 이는 국제기구라는 창구를 통해 비교적 쉽게 재난에 대한 관리 및 대응이 가능하며 한국 단독의 지원이 아닌 국제기구와의 공동지원이라는 점에서 위와 같은 사례가 지속되어 축적된다면 북한재난 발생 시 한국의 개입이 국제규범화 될 수 있다는 데 의의가 있다.

둘째, 국내법적 제도 및 근거를 정립하는데 심혈을 기울여야 할 것이다. 복잡하고 미묘한 남북한의 특수관계 아래에서 신속하고 정확한 개입 및 재난대응을 위해서는 남북교류협력법에서의 각종 제제범위와 제제방안 등에 대한 재정비가 필요할 것이라

판단되며 이와 같은 작업이 어려울 경우 북한지역의 재난대응 및 관리에 관한 법률 조항 또는 특별법을 제정하여 비단 북한 지역이 천재지변, 인재에 뿐만이 아니라 혹시 발생할지 모를 북한 내부의 급변사태로 인한 대규모 인명피해와 재산피해에 대비할 근거를 마련하는 것이 옳을 것이라 판단하였다.

셋째, 남북한 간의 기존 합의를 이행하는 것이 가장 상책일 것이나 남북한 간 합의는 정치·군사적 이슈에 민감한 바, 이로 인해 북한 주민들의 재난에서의 생존권이 침해 받는 것을 방지하기 위해 북한 재난 발생 시 우리가 제안할 수 있는 최소한의 인도주의적 지원방책을 사전에 마련해야 할 것이다.

이를 위하여 북한에서 발생한 재난의 종류에 따른 접경지역 지방자치단체와 중앙정부의 행동 규범과 인도주의적 지원 시 지원물품 품목 및 지원수단 등을 명시한 매뉴얼(manual)이 필요하며 이 매뉴얼이 작동하기 위한 법적제도의 정비가 필요하다고 판단하였다. 본 매뉴얼은 국내법적 정비와 맞물려 법령으로 구체화가 가능할 것이며 차후 남북관계의 진전에 따라 북한과의 합의를 통해 더욱 적극적인 재난대응·관리 분야에서의 교류협력을 진전시킬 수 있으며 전무하다시피 한 북한의 재난관리체계를 개선시키기 위해 우리의 노하우(know-how)를 전수할 수도 것이다.

매우 고조되었던 위기에서 18년 4월 남북정상회담까지 급격한 정세변화를 맞이하고 있는 한반도에서 남북이 교류와 협력을 통해서 평화를 정착시키고자 한다면, 과거 교류협력의 합의와 이행에서 도출된 문제를 진지하게 성찰하는 자세가 필요하다. 재난 대응과 같은 시급하고 중요한 분야에 대한 우선적인 합의와 이를 이행하고자 하는 강한 의지는 법제화를 통해 발현되고 실천될 수 있을 것이다.

참고문헌

- 강석윤. 2009. “지금 북한에서는.” 『북한』 451권: 170-179.
- 권태진. 2006. “북한의 홍수 피해 상황과 농업부문 영향.” 『한국농촌경제연구원논집』 제8권 제2호: 3-20.
- 김정순. 2010. “북한 환경보호법의 고찰.” 2010 남북법제연구보고서(II). 법제처. 남북회담본부. “대북수해복구 지원 관련 남북적십자 실무접촉.” 2006.08.19.
- 대한적십자사. 2009. 『국제인도법이란?』. 대한적십자사.
- 명수정·홍현정·최현일·정주철. 2008. “북한의 자연재해취약지 추정 및 남북협력방안 연구.” 한국환경정책평가연구원 보고서.
- 박건우. 2011. “남한의 통일법제.” 서울대학교 헌법통일센터 2011 학술대회.
- 박경석. 2013. “북한의 산림현황 및 산림정책.” 『KREI 북한농업동향』 제15권 제3호, 3-24.
- 박재영. 1999. “북한과 국제기구: 개발을 위한 북한의 대국제기구 정책의 특징.” 『경상대학교 민족통일론집』 제15권: 91-106.
- 북한농업팀. 2006. 『KREI 북한농업동향』 제8권 제3호.
- 북한연구소. 2004. “북한 용천역 열차 폭발.” 『북한』 제389호: 166-170.
- 북한연구소. 2004. “폭발사고로 폐허화 된 용천시가지 국제사회 도움으로 복구사업 완료.” 『북한』 제394호: 1-14.
- 북한연구소. 2006. “북한 수해 및 지원 현황.” 『북한』 제417호: 211-216.
- 서동렬·조영동. 2004. “용천역 폭발사고의 시사점.” 한국철도학회 학술발표대논문집. 한국철도학회, 137-152.
- 손종도. 2012. “용천역 폭발사고 지원 남북 긴급재난구호 모델 만들어.” 『통일한국』 제344호: 70-72.
- 윤대규. 2003. “북한사회에서의 법의 성격.” 『북한법연구』 제6호: 19-32.
- 이관세. 2012. 『북한지식사전 연구 개발』. 통일부.
- 이금순. 2001. 『남북한 재난관리 협력방안』. 통일연구원.

- 이원웅. 2009. “북한인권 개선 행위자들의 활동 평가: 국제기구, 미국정부, 한국정부, 국내 외 인권NGO를 중심으로.” 2009북한인권국제회의. (사)북한민주화네트워크.
- 이중운·홍이경. 2012. “북한의 식량난과 국제사회의 대북 지원 현황 분석.” 대외 경제 정책연구원.
- 이효원. 2004. “남북한 산림협력합의서 체결 방안.” 『통일과 법률』 19호: 41-74.
- 이효원. 2011. “통일법제 연구의 저변확대 방안과 과제.” 서울대학교 헌법통일센터 2011년 학술대회.
- 이효원. 2012. “한반도 통일과 DMZ의 효율적 관리를 위한 법제도화.” 『통일전략』 제12권 제1호: 45-85.
- 장용훈. 2004. “용천역 폭발사고 피해현황과 대북지원 - 한국의 적극 지원에 북한당국 ‘평가.’” 『통일한국』 제245호: 38-40.
- 최은석. 2010. “남북한 기상법제 비교 연구.” 법제처.
- 최희천. 2010. “재난관리 단계의 기존인식에 대한 비판적 고찰: 대응단계와 복구단계의 비교를 중심으로.” 『한국위기관리논집』 제6권 제1호: 201-218.
- 평화문제연구소. 2004. “북한 용천역 대형 폭발사고 - 한국과 국제사회의 신속한 복구지원.” 『통일한국』 제245호: 1-4.
- 한명관. 2010. “법치주의에 입각한 남북관계의 정립.” 『통일과 법률』 제4호: 1-10.
- 한명섭. 2011. “남북합의서의 법적 성격과 법제화 방안.” 『법조』 통권 제5호: 83-115.
- 현대경제연구원. 2004. “북 용천역 폭발참사 관련 정부 지원대책 현황(4.29).” 현대경제연구원.
- 황나미. 2010. “북한, 신종플루 발생 방역체계 현주소는?” 『통일한국』 제313호: 76-77.

A Study of Legal Issues for Effective ROK-DPRK Emergency Response Cooperation

Jinwon Kim* · Hanbeom Jeong**

ABSTRACT

The purpose of this study is to examine legal and institutional issues for effective cooperation between South and North Korea in disaster response.

In this study, we will examine the major cases of between South and North Korea disaster response cooperation in the past, and review the current legal and institutional issues for disaster response in North Korea and suggest ways to improve them. Disruption of friction and exchange caused by political and military issues between North and South Korea makes it difficult to effectively respond to and support North Korea in the event of a disaster. However, disaster response and humanitarian assistance need to be efficient and swift as long as they are timely and sufficient to achieve their objectives. To this end, it is necessary to adopt the code of conduct of the local governments and the central government according to the kinds of disasters in North Korea. In addition, manuals that describe the items to be supported and the means of support for humanitarian assistance are needed, and the legal system for the operation of this manual will need to be improved.

Keywords : South and North Exchange and Cooperation, correspond to disaster, humanitarian assistance

투고일: 2018.04.04. 심사일: 2018.04.10. 게재확정일: 2018.04.16.

* KNDU

** KNDU