**<Bootstrapping for Fuzzy Mediation, Moderated-Mediation Analysis>**

**Abstract**

매개 분석은 독립변수X가 종속변수 Y에 어떻게 영향을 미치는지에 관해 가설을 검정하는 통계 방법이며, 매개변수를 경유하여 전달되는 매개효과를 분석하는데 중점을 둔다. 매개분석은 심리학, 교육학, 경영학 등 다양한 분야에서 활발하게 이용되고 있다. 매개효과를 검정하는 방법은 Baron and Kenny의 단계적 검정, Sobel 방법 등 전통적인 방법들을 거쳐 최근 부트스트랩을 이용한 연구가 많이 진행되었다. 그러나 선행연구 대부분의 매개분석은 언어적 의미는 전달이 되지만 정확한 수치적 데이터로는 다루기 어려운 모호한 데이터조차 정확한 숫자 데이터로 이루어져있다. 이런 모호한 데이터는 crisp 숫자 보단 퍼지 숫자와 같은 soft 숫자로 표현하는 것이 적절하다. Yoon에 의해 퍼지 매개 분석에 대한 연구가 진행되었지만, 아쉽게도 전통적인 방법인 Sobel 방법을 사용하였다. 즉, 표본 분포가 정규분포를 이룬다는 가정을 전제로 분석이 진행되어, 편향적인 분포를 보이는 표본분포에 대해선 정확한 분석이 이루어지지 않았음을 알 수 있다. 따라서 우리는 표본분포가 알려지지 않은 상태에서, 표본자료를 이용하여 경험적인 분포를 형성하여 이를 토대로 표본분포를 추정하는 통계적 기법인 부트스트랩을 사용하여 다양한 모델에 대해 퍼지 분석을 제안한다. R에 대한 PROCESS MACRO를 이용하여 각 모델의 회귀 계수를 부트스트랩을 이용하여 추정하였으며, 간접효과에 대한 신뢰구간을 기존 방법인 Sobel 방법과 비교하여 다름을 확인하였다.

**1장**

  사회과학자들 및 행동과학자들은 인간의 행동을 설명하는 데 있어 어떤 자극에 노출이 되면 바로 반응하는 것이 아닌, 유기체 내에서 작동하는 메커니즘을 통해 반응한다고 보고 한 현상이 다른 현상에 어떤 영향을 미치는지에 대한 메커니즘에 관심을 가져왔다. 그들은 독립변수와 종속변수 간의 인과관계를 파악하고 이들의 관계성을 더 명확하게 파악하기 위해 제3의 변수를 추가하여 연구를 시도하였는데 이 변수는 매개변수와 조절변수로 나뉜다.

매개변수는 독립변수와 종속변수의 인과관계 중간에 논리적으로 개입되는 변수로서 어떤 이유로 또는 어떻게를 설명하기 위해 필요한 변수이다. 예를 들면, 어느 기업의 제품 만족은 고객의 만족으로 이어질 것이다. 다시 말해, 제품에 만족하는 고객들은 고객만족도 또한 높을 것이며, 반대로 제품에 불만족하는 고객은 고객만족도가 낮은 경향을 보일 것이다. 이 경우, 제품의 만족도와 고객만족도는 정적인 상관관계가 나타나는데, 이 관계가 어떻게 영향을 미치는데 설명해주는 변수가 매개변수이다. 이 관계에서는 제품의 회사에 대한 믿음, 즉 신뢰감이 형성되었을 시 제품 만족도와 고객 만족도의 양의 상관관계는 더 견고해진다고 가정할 수 있다. 이런 두 변수 사이에서 둘 간의 관계를 중간에 더 잘 설명할 수 있는 변수를 찾는 것을 목적으로 유의미한 효과가 있는지를 검정하는 것이 매개효과분석이다. 조절변수는 독립변수와 종속변수 간의 관계의 크기와 방향에 영향을 주는 변수이다. 조절효과분석은 조절변수에 의해 독립변수와 종속변수 간의 관계의 강도나 방향이 달라지는 분석하는 것인데, 두 변수 간의 관계를 설명하는 데 어떤 조건에서, 언제 또는 누구에게서 더 약한지 강한지를 연구하는데 목적이 있다. 이러한 매개효과와 조절효과는 여러 분야의 저자들에 의해서 연구되어왔다. [ ] 이 뿐만 아니라 변수간의 질적인 이해를 위해 매개효과와 조절효과를 결합하여 유의성을 검증하는 연구도 과거부터 진행되어왔다. [ ] 특히, 본 논문에서 소개 할 조절된 매개효과는 이것의 한 예이다. 조절된 매개효과는 James & Bret에 의해 1984년에 처음으로 소개되었으며[ ], 매개효과의 강도나 특정변수에 의해 조절된다는 것을 의미한다. 즉, 조절변수의 값이 높아질수록 매개효과가 강화되거나 약화되는 것을 의미한다(Jame & Bret, 1984). 매개효과와 조절효과 그리고 조절된 매개효과의 간단한 모델은 Fig. 1에 나타내었다.

 회귀분석에 기반을 둔 sobel(1982)검정법과 Baron and Kenny(1986)의 방법 그리고 Aroian, Goodman 검정법들은 지난 기간동안 많은 논문에서 사용되어 온 매개분석 방법이다. 그러나 Baron and Kenny(1986)를 이용한 매개효과분석은 매개효과의 유무를 판단할 뿐, 매개효과의 통계적 유의성을 판단하지 못하며, 그 외의 sobel, Aroian, Goodman 방법들은 매개효과의 통계적 유의성의 경우 분석순서와 판단절차가 단순하지 않다는 점이 실상이다. 또한 이 방법들은 통계적 검정력이 약할 뿐만 아니라 연구모형에서 측정오차를 반영하지 못하고 있으며, 정확하지도 않다는 비판과 함께 매개모형을 검증하는데 한계를 갖고 있는 것으로 보고, 이를 극복하기 위한 방법으로 부트스트랩을 이용한 방법이 최근 많은 논문에서 사용되고 있다.

(기존 선행연구 (부트스트랩 이용))

 그 동안은 대개 이런 매개모형들을 분석할 때 ‘정확한 숫자’로 분석이 진행되었다. 그러나 현실에선 정확한 숫자로 나타내기 어려운 애매한 표현을 하는 데이터들도 존재한다. 예를 들어 ‘적당한’, ‘몇몇’과 같이 언어적 의미는 전달이 되지만 정확한 수치적 데이터로는 다루기는 어렵다. 특히 심리학을 다루는 사회과학 분야에서는 이런 모호한 데이터들을 자주 만나게 되는데, 이런 데이터를 정확한 숫자로 표현화하는 과정에서 정보의 손실 뿐 만 아니라 표현하는 것 자체에 어려움을 겪을 것이다. 예를 들어, 사람의 스트레스의 정도를 변수로서 측정한다고 해보자. 정확한 숫자로는 사람의 정신적인 부분을 완전히 표현할 수 없다는 것은 사실이며, 이를 수치적으로 표현한다고 해도 사람마다 평가 척도는 다를 수밖에 없어 같은 데이터 수치라고 해도 실상은 다른 수치일 수 있으므로 이를 그대로 코딩하면 정보의 손실은 필연적일 것이다. 따라서 이를 Zadeh에 의해 처음 도입된 퍼지수와 같은 소프트한 수로 표현하는 것이 합리적이다.

 퍼지이론을 사용한 매개분석은 2020년도에 Yoon에 의해서 진행되어왔다[ ]. 그러나 fuzzy mediation과 fuzzy moderated - mediation 이용한 부트스트랩 논문은 연구가 전혀 진행되지 않았다. 부트스트랩은 기존의 매개효과 검증 방법으로 많이 사용되었지만 통계적인 추론 오류와 정확하지 않은 통계방법인 Baron & Kenny 방법과 표본분포가 정규분포를 이룬다는 가정을 전제로 편향적인 분포를 보이는 표본분포는 통계적으로 유의한 매개효과를 검증할 수 없다고 비판을 받는Sobel방법의 대안으로서 변수의 분포와 표본분포에 대해 어떤 가정도 하지 않는다는 점에서 강점을 가지고 있는 방법이다. 특히 수천 번의 재표본추출 작업을 필요로 하는 부트스트랩의 방법은 컴퓨터의 성능 향상 및 R, python 등의 통계 패키지를 통해 접근이 용이해짐에 따라 최근 많이 쓰이고 있어, 본 논문에서는 부트스랩핑 방법을 이용하여 fuzzy mediation model과 fuzzy moderated - mediation model을 ( )을 이용하여 분석하는 것을 제안하다.

This paper is structured as follows: Sect. 2 통계적 검정 방법을 기존 검정 방법과 본 논문에서 사용할 부트스트랩 방법에 대해 소개하고, Sect. 3 제안된 퍼지 매개 및 조절된 매개 모델을 제안하며, Sect. 4에선 제 2절에서 제안된 다양한 모델을 부트스트랩을 이용하여 추정하는 방법을 제공하고 있다. Sect.5 에선 앞에서 설명한 모델 및 추정 방법을 이용하여 데이터 분석을 데이터 세트와 함께 제공하며 제 4절에서 결과를 결론짓는다.

(This paper is structured as follows: Sect. 2 introduces the existing statistical test method and the bootstrap method to be used in this paper, Sect. 3 proposed fuzzy mediation and moderated - mediation models, and Sect. 4 provides an estimation method using bootstrap using the various models proposed in Sect. 2. In Sect. 5, data analysis is provided along with the data set using the model and estimation method described above, and the results are concluded in Sect. 4.)

**2. 통계적 검정 방법**

**2.1 기존의 매개효과 검정 방법**

**2.1.1 Baron & Kenny**

Baron & Kenny의 연구는 매개변수와 조절변수를 명확하게 정의하고, 매개효과 검증의 논리를 직관적으로 이해하기 쉽게 제시했으며, 매개효과가 실제로 어떻게 검증할 수 있는지를 자세히 규명하여 매개효과 검정법으로 논문에 가장 많이 인용되고 있는 방법이다.

그러나 최근에는 이 방법이 여러 문제점이 있다고 보고 비판을 받고 있다. 매개효과의 크기를 산출할 때 그 크기가 유의한지 통계적 추론을 통한 검증이 아닌 다른 수치들을 산출하여 차례로 검증함으로써 간접적으로 매개효과에 대한 결론을 내린다. 어떤 가설을 검증할 때 오류를 저지를 가능성은 항상 존재한다. 검증해야 하는 가설의 수가 많을수록 오류가 나타날 확률은 높아질 수밖에 없는데, 다수 가설의 순차적 검정으로 인한 오류 과다로 검정력이 약하는 사실이 밝혀졌다(e.g., Fritz & MacKinnon, 2007; Hayes & Schaarkow, 2013). 또한 Baron & Kenny 검정법은 독립변수가 종속변수에 미치는 영향이 통계적으로 유의해야 한다는 가정을 기반으로 매개효과를 분석하는데 이는 사실이 아니다. 이는 매개효과 검증 방법이 통계적으로 엄밀하지 않은 것이 아닌 정확하지 않은 통계 방법이라는 비판을 받고 있다.

**2.1.2 Sobel Test**

Baron & Kenny 검증방법의 핵심적인 문제점이 매개 효과를 간접적으로 검증함으로써 발생하는 것이라면, Sobel(1982) 방법은 직접적으로 그 효과의 크기를 산출하여 검증하는 점에서 Baron & Kenny의 방법보다 진일보한 방법이라고 볼 수 있다. 비교적 간단하게 매개효과를 검증할 수 있다는 점에서 Sobel test는 연구자들에 의해 자주 사용되어왔다. 그러나 Sobel 검정 방법 또한 결함을 가지고 있는 것으로 밝혀졌다. Sobel 검정법으로 매개효과의 유의도를 검증할 때 그 값의 표본분포가 정규분포를 이룬다는 가정을 전제로 하고 있는데, 다수의 연구자들에 의해 매개효과 검증 시 사용되는 표본 분포가 정규분포가 아닌 대개 편향적인 분포를 보이는 것으로 밝혀졌다(Bollen & Stein, 1990; Shrout & Bolger, 2002). 따라서 Sobel 방법 또한 통계적으로 유의한 매개효과를 제대로 판단하지 못한다는 한계가 있다(Fritz & MacKinnon, 2007; Hayes & Scharkow, 2013).

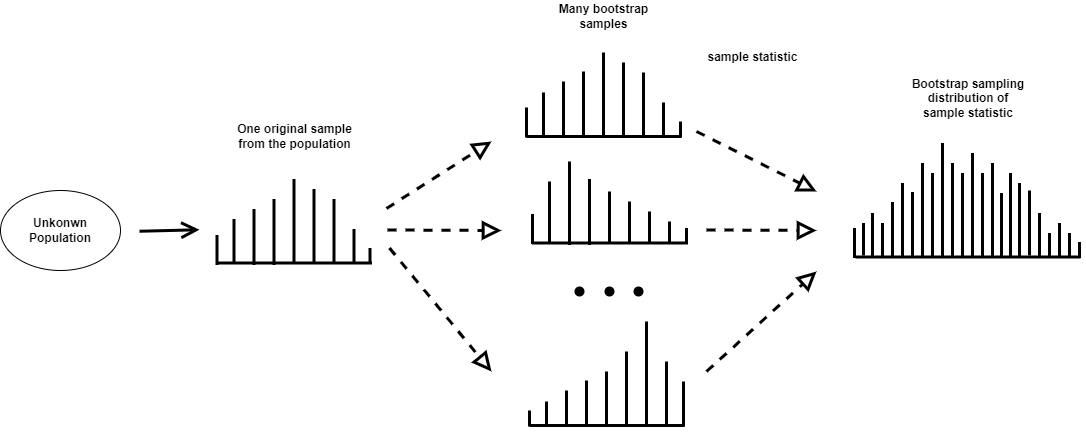
**2.2 Bootstrapping**

일반적으로 매개효과의 신뢰구간을 계산하는 방법은 표본추출분포가 정규분포 또는 t분포를 따른다고 가정한다. 그러나 추출된 표본들이 정규분포 또는 t분포를 따르지 않을 경우가 존재한다. 표본 추출의 분포가 대칭적이 않을 수 있는데 이 경우 정규분포 또는 t분포로 가정하여 신뢰구간을 계산하게 된다면, 올바른 신뢰구간의 근사치를 제공해 줄 수 있지만, 항상 우수한 근사치를 제공하지는 못할 것이다. 이에 대한 대안으로 제시되고 있는 bootstapping 방법은 최근에 들어서야 연구자들에게 점차 보편화되었다.

부트스트래핑 방법은 표본분포가 알려지지 않은 상태에서, 표본자료를 이용하여 경험적인 분포를 형성하여 이를 토대로 표본분포를 추정하는 통계적 기법이다. 즉, 변수의 분포나 표본분포에 대한 어떤 가정도 하지 않고 표본자료에서 표본크기가 동일한 수의 표본들을 반복적으로 무작위 복원 추출한 후, 추출된 표본들을 통해 반복 추정하여 추정 표본분포의 근사적 표준오차, 신뢰구간과 유의확률을 계산하게 된다는 점에서 매개효과를 검증하는데 강점을 가지고 있다. (Fig() 참조)

부트스트래핑 방법으로 매개효과를 검증하는 절차로 두 가지 방법이 제시되고 있다. 먼저 재추출한 표본분포의 신뢰구간에 0이 포함되는지의 여부로 판단하는 방법과 검정 매개변수 효과 분해를 통해 총간접효과의 유의확률로 판단하는 방법이 있다. 신뢰구간을 계산하는 기법으로는 Percentile의 방법과 Bias-corrected 방법이 있다.

사실 부트스트래핑 방법을 이용하여 매개효과를 검증하는 방법은 1990년대부터 일부 학자들에 의해 소개되어 왔다(Bolen & Stein, 1990). 이러한 강점에도 불구하고 부트스트래핑 방법이 보편화되지 못했던 이유는 컴퓨터를 이용하지 않고서는 엄청난 계산량을 실행하기가 어려웠으며, 복잡한 프로그래밍으로 인해 적용하는데 제약이 있었다. 그러나 최근 컴퓨터의 발달로 다양한 통계패키지를 통해 부트스트래핑을 사용하는 절차가 간소화됨으로써 여러 학문분야에서 그것의 활용도가 증가하고 있다. 따라서 본 논문에서는 부트스트랩 사용하여 퍼지 매개모델의 통계적 유의성을 설명하였다.



**2.2.1 Percentile 방법 (백분위 부트스트랩)**

독립변수가 매개변수에 미치는 영향의 크기를 a, 독립변수의 영향을 통제시킨 상태에서 매개변수가 종속변수에 미치는 영향의 크기를 b라 할 때, 간접효과를 ab로 정의한다. 이 간접효과의 크기인 ab의 표본분포의 대한 가정이 필요 없는 추론방법 중 대표적으로 부트스트랩 신뢰구간을 이용하여 간접효과를 검정하는 방법이 있다. 신뢰구간을 설정하는 방법 중 하나인 백분위 부트스트랩 방법에 의해 신뢰구간(95%) 설정 방법의 절차는 다음과 같다(Shrout & Bolger 2002). 모든 단계는 Hayes에 의해 개발된 컴퓨터 프로그램인 PROCESS macro에서 자동으로 실행된다.

1. 모집단에서 추출된 표본크기가 N개인 원표본에서 복원 추출하여 원표본과 동일한 N크기인 부트스트랩 표본을 추출한다.
2. 1단계에서 얻은 부트스트랩 표본을 이용하여 재표본에서 간접효과의 통계치를 추정한다.
3. 1과 2단계의 과정을 k번 반복하여 k개의 부트스트랩 표본을 생성하고 이를 이용하여 k개의 간접효과를 추정하여 저장한다.
4. k개의 간접효과 추정치들을 가장 낮은 값에서부터 가장 높은 값으로 분류한다.
5. 95%신뢰구간을 사용하는 경우 하한선을 앞에서 구한 통계치의 분포에서 0.5(100-95)번째 백분위에 해당하는 통계치로 정의하고 상한선은 오름차순으로 정리된 k개의 통계치 분포에서 [100-0.5(100-95)]번째 백분위에 해당하는 통계치로 정의한다. 하한선 값과 상한선 값이 95% 신뢰구간의 양 끝점으로 결정된다.

이런 95% 신뢰구간에 0이 포함되지 않으면 간접효과가 통계적으로 유의하다고 한다.

**2.2.2 Bias-corrected 방법 (편향조정 부트스트랩)**

백분위 부트스트랩 신뢰구간에 존재하는 잠재적 편향을 보완하는 편향조정 부트스트랩은 Efron과 Tibshirani(1986)에 의해 제안되었다. Bias-corrected 접근법은 백분위 신뢰구간과 기본적으로 동일하지만 k개의 부트스트랩 표본에서 계산된 k개의 간접효과 추정치들 중 원표본의 간접효과의 점추정치보다 작은 개수의 비율을 이용해 편향 상수(bias constant)를 계산하고 이를 이용하여 백분위 부트스트랩 신뢰구간의 양 끝점의 오차율이 같아지도록 수정한 신뢰구간이다. 이는 부트스트랩 추정치 분포의 비대칭성을 더욱 엄밀하게 반영해서 신뢰구간의 상한과 하한을 결정한다. 따라서 추정치의 표본분포가 비대칭적일 때는 bias-corrected의 방법이 더 정확한 결과를 얻을 수 있다. 그러나 편향조정 부트스트랩이 검정력은 높을지라도 1종오류를 발생시키기 때문에 적절한 검정 방법이 아닐 수 있다는 사실이 최근 보고되고 있다(Biesanz et al, 2010, Hayes & Scharkow, 2013, Falk & Biesanz, 2015, Tofighi & Kelly, 2020).

**3. Fuzzy Mediation and Moderated-Mediation Analysis**

In this section, referring to the basic concepts in [1], we introduce the definition of fuzzy numbers by Zadeh [2], and simple fuzzy mediation models with mediators and fuzzy moderated-mediation model introduced by Yoon [3,4].

**3.1 Fuzzy number**

퍼지 숫자는 실수 R에서 정의되는 퍼지 집합으로서 정규화되고 볼록할 때를 의미한다. 퍼지집합은 Membership function이라고 불리는 함수에 의해 0과 1사이의 실수 값을 소속척도로 취하는 원소들로 구성된다. Membership function의 형태는 객관적이거나 주관적인 가능성을 고려하여 정의할 수 있어 일반적인 규칙이 존재하지 않는다. 따라서 특정한 경우로 LR-퍼지 숫자라고 하는 퍼지 숫자의 parametric class가 사용된다. 퍼지 숫자A가 다음과 같은 조건을 만족하면 LR 퍼지숫자라 한다.

where L and R are reference functions called left and right shape functions of X and have the following properties : L,R :R→[0,1] are left-continuous and decreasing function with R(0) = L(0) = 1, R(1) = L(1) = 0. And ‘m’ means the mode of the LR-fuzzy number A. ‘l’ and ‘r’ are greater than 0 and mean the width of the left and right sides. We abbreviate the LR-fuzzy number as . And LR-fuzzy number, one of the triangular numbers, has the following two operations.

= (,

.

**3.2 Simple Fuzzy Mediation Model**

Baron and Kenny의 Simple mediation model의 매개분석방법은 단순 회귀분석을 통해 인과관계를 단계별로 분석한다. 매개분석은 인과선행변수(X)가 결과변수(Y)에 어떻게 영향을 미치는지에 관한 가설을 검정하는 통계방법이다. Baron and Kenny의 회귀식은 3개가 도출되는데 독립변수와 종속변수 간의 regression, 독립변수와 mediator 사이의 regression 그리고 mediator과 종속변수 사이의 regression 이며 다음과 같이 제안된다[5].

이 모델에서 와 은 중요한 회귀상수이다. 여기서 은 X가 Y에 영향을 미치는 ‘직접효과’를 추정한 것이며, 은 X가 M을 경유하여 Y에 영향을 미치는 ‘간접효과’를 와 의 적항으로 수치화한 것이다. 또한 은 ‘총효과’이며 X의 직접효과와 간접효과의 합계인 와 같다. 다시 말해 직접효과(은 총효과(보다 작다는 사실을 보여준다.

‘많다’, ‘적다’, ‘행복하다’와 같이 수치화 하기 애매할 때, 이를 변수로 하는 경우 crisp 숫자보단 퍼지 숫자를 사용하여 표현하는 것이 더 합리적이다. Fuzzy Mediation Model은 다음과 같이 제안된다.

In the model as above, is the total effect, is the indirect effect, and is the direct effect. Note that it is easily checked that

**텍스트, 시계, 손목시계이(가) 표시된 사진

자동 생성된 설명**

**3.3 Fuzzy Mediation Model for Multiple Mediators**

현실세계는 훨씬 더 복잡하고 다양한 원인에 의한 인과관계로 인해 종종 simple fuzzy mediation model 대신 매개변수가 여러 개인 simple fuzzy mediation for multiple mediators를 이용하는게 현명할 때가 있다.

매개변수가 k(k>1)인 simple fuzzy mediation model은 다음과 같이 제안된다.

where

In the model as above, is the total effect, is the indirect effect through and on , and is the direct effect. In other words, there are k indirect effects.



**3.4 Fuzzy Mediated-Moderation Model**

**3.4.1 Mediatied-Moderation Model**

인과관계에서 네 번째 변수인 조절변수(W)로 인해 독립변수(X)가 매개변수(M)을 통해 종속변수(Y)로 가는 간접효의 영향력이 조절될 수 있는 메커니즘을 moderated-mediation이라 한다. 조절된 매개효과는 통계적으로 조건부 간접효과(conditional indirect effect)를 의미하고 현재는 동일한 의미로 혼용하여 사용한다(Preacher, Rucker & Hayes, 2007)

조절분석과 매개분석을 통합한 조건부과정분석은 한 변수가 다른 변수에게 영향을 전달하는 메커니즘의 조건부 성격과 이런 조건부효과에 관한 가설검정을 이해하는 것을 목적으로 한다. 따라서 매개분석의 관심사인 독립변수(X)의 직접효과와 간접효과를 추정하고 조절분석에서 독립변수(X)가 매개변수(M)에 미치는 효과가 조절변수(W)에 의해 조절되는 경우 X가 M에 미치는 영향은 단일 수치가 아닌 W의 함수로 나타난다.

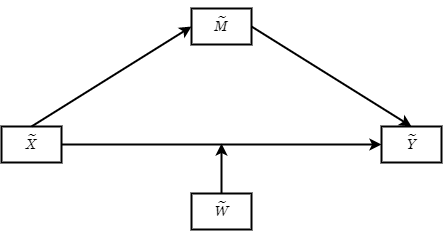
**3.4.2 Fuzzy Moderated-Mediation Model**

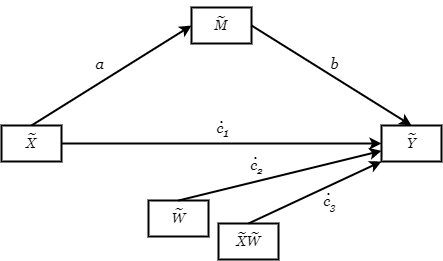
모호한 변수를 사용한 인과관계에서 를 퍼지 예측변수라하고 를 퍼지 반응변수 과 을 각각 fuzzy mediator, fuzzy moderator이라 하자. 아래 모형 in the Fig은 에서 의 경로가 에 의해 조절되고 다른 경로는 조절되지 않는 가장 단순한 조절된 매개모형이다.

의 직접효과는 에 의해 조절되기 때문에 에 조건적이다. 다음 모델 in Fig을 표현하면:

=

여기서 은 에서 의 ‘퍼지 조건부 효과’이며, = 의 함수로 정의된다. 이 모형에서의 간접효과는 단순 퍼지 매개모형과 마찬가지로 a와 b의 적항으로 정의되며, 또는 경로가 조절되지 않으므로 간접효과는 비조건적이다. 따라서 이 모형에서는 가 을 통해 간접적으로 에 영향을 미치며 또한 는 에 의존하면서 에 직접 영향을 미치기도 한다.

****

****

지금까지 Yoon[3,4]에 의해 다양한 퍼지 매개 모델들이 제안되었다. 그러나 아직까지 이런 퍼지모델을 부트스트래핑을 이용하여 검정한 논문은 존재하지 않는다. 다음 4장에서는 부트스트래핑을 통한 단순 퍼지 매개 분석 및 조절된 매개 분석을 먼저 제안하고, 5장에선 기후 데이터 및 다양한 데이터를 부트스트래핑을 이용하여 분석한 것을 제안한다.

**4. Bootstrapping for Fuzzy Mediation and Moderated-Mediation Analysis**

**4.1 Estimation for Fuzzy Mediation Analysis**

When using least squares estimation with fuzzy data, it is necessary to have a suitable metric in the fuzzy set spaces. A helpful type of metric can be established through the use of support functions. The support function of any compact, convex set can be represented as which is determined by the following formula for all :

where is the (d-1)-dimensional unit sphere in and represents the scalar product in . It should be noted that for compact and convex sets the support function is uniquely defined. A metric in a fuzzy number set is established through the use of the *-* metric in the space of Lebesgue integrable functions, represented as:

This leads to the definition of an *-* metric for fuzzy numbers as:

A fuzzy regression model was previously introduced in the author's studies [24,25] and is expressed as follows:

.

The variables are represented by and for It is assumed that are the fuzzy random errors that account for the fuzziness. It is worth mentioning that all cases can be covered by defining and as follows:

where represent the left and right spreads of respectively.

The estimators are obtained by minimizing the following objective function:

where *q* is the number of the regression model in this fuzzy mediation analysis and *k=1,2,…,q*,. The objective function is based on the *-*metric, and the *-* distance can be calculated as:

To minimize the above equation, we obtain the normal equation applying

The normal equation has as its solution, and for each value of , the normal equation can be written as follows:

To determine the solution vector, we introduce a *triangular fuzzy matrix* *(t.f.m.)* which is expressed as

,

and abbreviated as , where is a triangular fuzzy number forandAdditionally, we define a triangular fuzzy vector

.

To minimize the objective function mentioned above, we apply the fuzzy operations, fuzzy numbers and estimators defined in our previous studies [26-29]. The fuzzy operations are as follows:

*.*

The following operations are defined for two triangular fuzzy matrices, , , and a crisp matrix :

,

*,*

*,* .

.

The solutions to the normal equation fuzzy estimators are derived for each by using the above operations and algebraic properties, with

where

and , for Note that the solution (16) exists only if .

**4.2 Statistical inferences of Fuzzy Mediation Model and Moderated-Mediation Analysis**

**4.2.1 기존 방법을 이용한 구간 추정(Interval Estimation)**

Fuzzy least squares estimators ( )가 점근적으로 정규분포를 따른다는 가정은 이전 연구[ ]를 통해 알려진 바이다. 모집단의 분산(variance)를 알 수 없는 경우는 t-분포를 적용하고, 데이터의 크기가 큰 경우는 정규분포를 따른다고 가정하여 z값을 통해 총효과와 직접효과에 대한 (1-)100% 신뢰구간을 나타낼 수 있다.

*CI for the total effect :*

*CI for the direct effect*

직접효과와 총효과의 표준 오차(standard error) se는 다음과 같이 정의한다.

표본 평균에 경우 다음과 같은 성질을 적용할 수 있다.[ ]

*,* where is a fuzzy random variable.

In addition, hypothesis tests

,

,

간접효과의 경우 “Sobel test” 또는 “delta method” 또는 “product of coefficients method”를 통해 신뢰구간을 추론할 수 있다[ ]. 간접효과 ab는 표본을 기준으로 한 의 추정값이다. 간접효과의 유의 수준은 아래와 같이 매개효과 추정치를 추정치의 표준오차 로 나눈 값을 검정통계량으로 하여 정규분포(Z)를 통해 판정한다.

where .( second order standard error estimator)

는 a와 b의 표준 오차를 의미한다.

*CI for the indirect effect*

where .

Based on above assumptions, the hypothesis test

,

Sobel 검정의 경우 간접효과의 표본 분포의 정규성을 가정하고 표준 정규 분포를 사용하여 도출한다. 이러한 가정은 표본이 큰 경우에는 합리적이지만, 작은 표본에서는 그렇지 않다. 일반적으로 상황에 따라 달라지는 가정은 낮은 검정력을 산출한다. 아직까지 많은 분야의 전문가들이 sobel 검정을 여전히 사용하고 있다.

**4.2.2 부트스트랩을 통한 구간 추정(Interval Estimation) or Bootstrap confidence interval**

본 절에서는 3절에서 제안한 모델에 대한 부트스트랩을 이용한 신뢰구간을 통한 통계적 추론을 제안한다.

정규성을 가정하는 매개효과 검정 방법들이 가지는 한계를 극복하기 위해 부트스트랩은 Efron(1979)에 의해 처음 고안되었다. 부트스트랩은 매개효과 추정치의 표집분포에 대한 가정이 필요 없는 비모수적이고 경험적인 재표집 방법이다. 위에서는 신뢰 구간을 구하는 고전적인 접근 방식을 검토했다. 다양한 부트스트랩 방법 중 부트스트랩 분포의 표본 백분위수로 신뢰 구간을 추정하는 방식을 제안한다.

원래의 표본에서 크기가 n인 재표본을 추출한 후 재표본에 대한 추정치를 산정하고 이 두 단계를 여러 번 반복하여 재표본 추정치의 산포를 파악한다. 주어진 확률분포 , *, ,,*로부터 복원 추출 방법(sampling with replacement)으로 부트스트랩 표본 , , , 을 얻는다. 이 부트스트랩 표본에 대해 구하고자 하는 통계량을 얻을 수 있으며 이를 통해 추정 통계량 의 부트스트랩 분포를 추정하게 된다. 정확한 신뢰구간 추정값을 계산하기 위해 최소 1000개의 부트스트랩 표본이 적절하며 B의 값이 클수록 좋다.

백분위수 부트스트랩 방법(percentile bootstrap method ‘ PB) B개의 부트스트랩 표본을 통한 추정 통계량 {, , }를 크기순 ≤ ≤ · · · ≤ 으로 나열하여 구한 에 대한 (1−α) × 100% PB 신뢰구간을 다음과 같다.

[ , ]

여기서 BL(Bootstrap Lower bound)은 의 percentile과 다음 수의 평균값이고, BU(Bootstrap Upper bound)은 의 percentile과 이전 수의 평균값이다.

총효과와 직접효과에 대한 (1-)100% 신뢰구간은 다음과 같이 표현한다.

*CI for the total effect* [ *,*

*CI for the direct effect* [ *,*

직접효과와 총효과의 표준 오차(standard error) se는 다음과 같이 정의한다.

표본 평균은 다음과 같은 성질을 적용한다.

*,* where is a fuzzy random variable.

In addition, hypothesis tests

,

,

간접효과 ab의 신뢰구간을 부트스트랩을 통해 정의하면,

*CI for the indirect effect*

이며, 간접효과의 표준 오차(standard error) se는 다음과 같이 정의한다.

이때 에서 로 가는 경로의 영향력인 a의 표준 오차는 다음과 같으며,

에서 로 가는 경로의 영향력인 b의 표준 오차는 다음과 같다.

우리는 B=5000번의 반복을 통해 95% 부트스트랩 신뢰구간을 제안한다. 따라서 α=0.05이며, BL은 와 백분위수의 평균값이므로 125번째와 126번째의 추정 통계량의 평균값이 되며, BU은 4,874번째와 4,875번째의 추정 통계량의 평균값이 된다.

Based on above assumptions, the hypothesis test

,

1. Hayes, A.F. ‘‘Introduction to mediation, moderation, and conditional process analysis’’, Guilford (1e, 2013/2e, 2018)

2. Zadeh, L.A.: Fuzzy sets. Information and control. 8, 338-353 (1965)

3. Yoon, J.-H.: Fuzzy mediation analysis. Int. J. Fuzzy Syst.22(1), 338-349 (2020)

5. Baron R.M., Kenny D.A.: The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 51(6), 1173-82 (1986)