|  |
| --- |
| **실험 설계에서 일반화 선형 혼합효과 모형의 수행** |
|  |
| 일반화 선형 혼합효과 모형(generalized linear mixed-effects model: GLMM)은 여러 맥락에서 적용 가능한 모형으로서 최근에는 실험 설계 연구로부터 수집된 데이터 분석을 위해서도 사용이 증가하는 추세이다. GLMM의 실험데이터 분석을 위한 논의가 있었으나 그 설정이 매우 제한적이었고 연구자가 가정하는 상황에서 필요한 참가자와 문항 수를 결정하기 위한 정보를 제공하지는 않았다. 따라서 본 연구에서 GLMM의 사용을 고려하는 실험연구자들에게 필요한 정보를 제공하고 모형을 적절히 사용하기 위한 방법을 제시하기 위해 체계적인 몬테카를로 시뮬레이션 연구를 수행하였다.  본 연구에서는 두 가지 연구문제를 조사하였다. 첫 번째 연구문제는 연결함수(link function)에 따른 GLMM의 수행에 대한 평가이다. GLMM의 장점 가운데 하나는 연결함수를 사용함으로써 종속변인의 분포적 가정이 정규분포에 국한되지 않는다는 것이다. 종속변인이 연속형인 경우 사용할 수 있는 선형 혼합효과 모형(linear mixed-effects model: LMM)은 연결함수가 항등함수(identity function)인 GLMM의 일종이라고 볼 수 있다. LMM의 경우에 비해 실험 연구를 위한 GLMM의 성능을 평가한 논문의 수는 적다. Jaeger(2008)은 이산형 종속변인을 가진 실험데이터의 분석을 위해 변량분석 대신 GLMM을 써야하는 이유를 제시되었으나 체계적인 시뮬레이션 연구 결과는 제시하지 않았다.  두 번째 연구문제는 고정효과를 탐지하기 위한 모형선택 접근방법이다. 전통적으로 실험데이터를 분석하기 위해서 실험설계에 기반한 최대모형(maximal model)이 관행적으로 사용되었고, Barr 등(2013)은 LMM을 사용할 때 역시 최대모형의 사용을 제안하였다. 반면, 비실험 연구에서 회귀분석을 사용할 때 연구자의 관심 변인의 효과를 탐지하기 위해 모형 비교 후 선택된 최종모형에 기반하여 추론하는 과정이 일반적이다. Matuschek 등(2017)은 실험데이터 분석을 위해 LMM을 사용할 때도 모형 비교과정을 거칠 것을 제안하였다.  본 연구에서 참가자 수와 문항 수, 고정효과의 크기, 문항효과의 분산을 실험조건으로 조작하였다. 실험조건의 각 수준에서 1000개의 데이터셋을 생성하고 모형의 성능을 평가하였다. 모형의 성능은 추정값의 bias와 RMSE에 의해 평가되었고, 고정효과의 추론에 대한 1종 오류 비율과 검증력을 조사하였다. |

\*신청: kpa.me2021@gmail.com 로 보내주시기 바랍니다.