

연구방법론				번호: I - B - 7	
제 목	국문	경시적 연구에서 무시할 수 없는 중도탈락 자료의 분석을 위한 다단계 일반화 선형모형 적합			
	영문	Fitting Hierarchical Generalized Linear Model for Nonignorable Drop-out Data in Longitudinal Study			
저 자 및 소 속	국문	노맹석 ¹⁾ , 이무송 ²⁾ , 강위창 ³⁾ , 하일도 ⁴⁾ , 이영조 ¹⁾ 1) 서울대학교 통계학과, 2) 울산대학교 의과대학 예방의학교실, 3) 대전대학교 정보통계학과, 4) 경산대학교 통계학과			
	영문	Maengseok Noh ¹⁾ , Moo-Song Lee ²⁾ , Weechang Kang ³⁾ , Il-Do Ha ⁴⁾ , Youngjo Lee ¹⁾ 1) Department of Seoul National University, 2) Department of Preventive Medicine, University of Ulsan College of Medicine, 3) Department of Information and Statistics, Daejeon University, 4) Department of Statistics, Kyungsan University			
분 야	역 학 연구방법론	발 표 자	노맹석	발표형식	구 연
			일반회원		
진행상황	연구완료				
<p>1. 연구목적</p> <p>의학, 보건 분야의 많은 연구에서는 관측대상들을 여러 관측시점에 따라 반복측정된 경시적 자료(longitudinal data) 혹은 반복측정 자료(repeated measured data) 를 수집하여 분석을 수행하게 된다. 이때, 첫번째 시점에서는 연구대상에 포함되었으나 각 개인의 특성 및 기타요인으로 인하여 중간에 중도탈락하는 관측대상이 발생하게 되는 결측치 자료가 빈번하게 발생된다. 본 연구에서는 이러한 중도탈락 결측치 자료(dropout missing data)를 분석하는데 있어서 무시할 수 없는 결측치가 있을 경우 기존의 경시적 자료에서 적용되는 일반적인 통계분석 방법을 적용하게 되었을 때 편의된 추정 결과를 제시하게 되므로 반응변수에 대한 모형화와 함께 결측치 값에 대한 모형화를 함께 고려하는 새로운 HGLM 모형 및 적합방법을 제시한다. 그리고, 시뮬레이션 결과를 통하여 본 연구에서 제시되는 새로운 방법이 편의된 추정량을 제시하는 기존 방법을 보정하는 결과를 제시함을 보인다.</p> <p>2. 연구방법</p> <p>결측치의 종류는 완전 랜덤 결측(MCAR; Missing Completely At Random), 랜덤 결측(MAR; Missing At Random), 무시할 수 없는 결측(Nonignorable Missing)과 같이 세가지의 종류로 나누어진다. MCAR 인 경우에는 결측치가 관측된 값과 결측된 값에 영향을 받지 않는 경우이며, 결측된 자료를 제외한 통계적 추론이 가능하다. MAR 인 경우에는 결측치가 결측된 값에는 영향을 받지 않지만, 관측된 값에는 영향을 받는 경우이며, 결측치를 제외할 때 반응변수의 주변분포(marginal distribution)에 대해서는 잘못된 결과를 제시하지만,</p>					

설명변수가 고정된 경우의 통계적 추론은 가능하다. 무시할 수 없는 결측치는 결측치가 결측된 값에 영향을 받는 경우로 결측치에 대한 정보를 제외하면 편의된 통계적 추론을 줄 수 있다. 경시적 자료에서 중도탈락의 경우가 무시할 수 없는 결측치인 경우가 많이 있다. 예를 들어, 특정 의료기관을 계속 방문하여도 환자 자신에게 도움이 되지 않을 경우에 중도에 다른 의료기관으로 옮기는 경우는 그 환자가 가지는 특성에 영향을 받은 결측치이므로 무시할 수 없는 결측치의 경우이다. 이러한 각 관찰대상 고유의 특성치를 반응변수와 함께 결측치 정보에 동시에 고려하여 중도탈락이 있는 경시적 자료분석을 가능하게 하는 모형을 제시하고 이것은 두 개의 서로 다른 반응변수를 하나의 랜덤효과로 연결한 동시모형(joint model)로 설명이 가능하다.

3. 연구결과

시뮬레이션 연구를 통하여 실제 중도탈락이 있는 경시적 자료일 경우에 결측치를 무시하는 기존의 방법은 편의된 추정치를 제시하고 있음을 나타내고 있다. 특히 시간에 따라 변화하는 경향을 나타내는 모수 추정치의 편의는 시간에 상관없이 일정한 값을 가지는 설명변수에 해당되는 모수의 편의보다 훨씬 크게 된다. 이러한 경우에 동시모형을 통하여 분석을 하게 되면, 그 편의를 줄여주게 되며 실제 모든 관측값이 있다고 가정하는 이상적인 자료를 분석하여 얻게 되는 결과와 유사함을 보였다.

4. 고찰

본 연구에서는 중도탈락이 있는 경시적 자료분석에 적용되는 새로운 분석방법을 제시하고 기존에 결측치를 무시하고 분석하는 방법의 적용하였을 때 잘못된 결과를 보정하는 것을 시뮬레이션 연구를 통하여 보였다. 현재 개발된 방법은 관측되는 변수가 연속형 자료일 경우 즉 정규분포 자료일 경우에는 가능하다. 관측변수가 이산자료일 경우에는 본 연구에서 제시하는 방법을 확장하면 가능하며 향후 추가적인 연구가 필요하다.