

연구방법론				번호: I - B - 10	
제 목	국문	폐암 사망수 분석시 과분산을 보정한 Poisson regression tree 적용			
	영문	Poisson Regression Trees for Analysis of Lung Cancer Mortality with Extra Variation			
저 자 및 소 속	국문	최윤희, 안홍식 ¹⁾ , 유근영 서울대학교 의과대학 예방의학교실, 뉴욕주립대 응용수학과 ¹⁾			
	영문	Yunhee Choi, Hongshik Ahn ¹⁾ , Keun-Young Yoo Department of Preventive medicine, Seoul National University College of medicine, Department of Applied math and statistics, SUNY at Stony Brook ¹⁾			
분 야	역 학 학	발 표 자	최윤희 일반회원	발표형식	구 연
진행상황	연구완료				
<p>1. 연구 목적</p> <p>특정기간 중 암의 발생 수나 실험조건상의 바이러스 균집 수와 같은 계측치는 실제 Poisson 모델에서 가정되는 분산보다 더 큰 분산을 가지는 경우가 많고 또한 독립변수와 종속변수와의 관계가 선형이 아닌 경우도 많아 이러한 데이터의 모델링 효과를 적합화 할 필요가 있다. 선형이 아닌 경우 모델링에서의 복잡한 결과 해석의 단점을 보완하면서 미국 Missouri주에서의 폐암 사망수에 대한 지역적 인구학적효과를 tree-structure을 이용하여 분석하였다.</p> <p>2. 연구 방법</p> <p>Tree-structure를 사용하여 데이터를 통계적으로 비슷한 점들끼리 나눈 후에 국소적으로 모델링을 하였다. 각 계측치의 분획에서 과분산(over-dispersion)을 고려하여 Quasi-likelihood을 사용하였으며 또한 과분산 모수에 대해 수치적인 방법이 필요 없는 불편추정치를 사용하여 모델적합 절차를 간편화 했다. 각 모델에서 공변량의 consistent한 검정을 위해 sandwich variance estimator (by Liang and Zeger)을 사용하였다. Tree-Structure에서의 불필요한 split을 5%내로 통제하기 위해 각 node에서 bootstrap resampling을 통한 검정을 시행하였다. Tsutakawa가 1985, 1988년에 수집한 폐암 사망수를 나이와 성별 그리고 인구수에 따라 이 방법으로 분석하였다. 과분산에 대한 tree-structure을 적용하기 전에 과분산에 대한 검정을 하였다.</p> <p>2. 연구 결과</p> <p>전체 폐암사망자 수의 과분산은0.043으로 추정 되었고 실제 과분산 존재에 대한 검정에서 거의 0에 가까운 P값을 얻었다. 이것은 일반적인 Poisson 모델보다 이 방법에 의해 모델이 만들어지는 것이 더 타당함을 보여준다. 폐암 사망률은 나이 65세를 기준으로 다른 패턴을 보여주었다. 65세 미만에서는 나이가 매우 중요한 요소였다. 75세 이상의 집단에서는 인구수가 남성의 폐암 사망률에 중요한 요인이었으며, 인구수는 다른 나이 집단에서도 중요한 요소였다. 전체적으로 남성이 여성보다 매우 높은</p>					

폐암 사망률을 보여 주었다. (자세한 결과는 tree그림참조)

4. 고찰

계측치를 분석하기 위해 일반 선형모델이 자주 사용되어진다. 하지만 데이터 전체를 한 선형모델로 해석하기에는 종종 어려움이 따르며 특히 독립변수의 수가 많거나 그들 중의 일부가 서로 상관되어 있을 때는 더욱 그러하다. 이러한 경우, 계측치를 제시된 방법으로 분석함으로 종속변수와 독립변수간의 관계를 좀 더 쉽게 해석 할 수 있으며 시각적인 결과보고 또한 이를 도와준다.