

제 목	국 문	생태학적 자료에서 분류오류의 영향과 분석모형의 선택		
	영 문	A study on the influence of the misclassification in ecological data and the selection of ecological model		
저 자 및 소 속	국 문	남정모, 강형곤 연세대학교 예방의학교실		
	영 문	Chung Mo Nam, Hyung Gon Kang <i>Dept. of Preventive Medicine and Public Health, Yonsei Univ. College of Medicine</i>		
분 야	역 학	발 표 자	남 정 모	
발표 형식	구 연	발표 시간	15 분	
진행 상황	연구완료 (), 연구중 (0) → 완료 예정 시기 : 96 년 10 월			
<p>1. 연구 목적</p> <p>이 연구는 개인단위에서 이분형으로 측정된 위험요인과 혼란변수가 모두 분류 오류가 있으며 이들이 질병위험도와 지수모형을 따르는 경우 몇가지 생태학적 모형을 제시하고 각각의 생태학적 모형에서 위험요인과 질병위험도의 추정치에 대한 생태학적 오류와 분류오류로 야기된 치우침을 알아보고자 하였으며 또한 그 치우침을 수정할 수 있는 방법을 수리적으로 알아보고자 하였다. 그리고 이러한 생태학적 오류의 크기와 분류오류로 야기된 치우침의 크기를 비교하여 생태학적 모형의 선택기준을 알아보았다.</p> <p>2. 연구 방법</p> <p>지역별 위험요인과 혼란변수에 각각 폭로된 사람의 비율과 평균 질병위험도의 관계를 생태학적 모형 1은 지수모형, 모형 2는 가법모형, 그리고 모형 3은 모형 2에 추가적으로 교호작용을 포함하여 각각 설정하였다. 이러한 3가지 생태학적 모형에서 각각 추정된 회귀계수를 이용하여 개인에서의 비교위험도를 외삽(extrapolation)하는 경우 발생하는 생태학적 오류를 시뮬레이션을 통해 조사하였다. 또한 3가지 생태학적 모형에서 위험요인과 혼란변수의 분류오류로 야기된 외삽된 비교위험도의 치우침을 수학적으로 알아보았고 또한 그 치우침을 제거할 수 있는 방법을 제안하였으며 시뮬레이션을 통해 치우침의 정도와 수정정도를 알아보았다. 시뮬레이션은 FORTRAN과 SAS의 NLIN 프로시저를 이용하였다.</p>				

3. 연구결과

1) 위험요인과 혼란변수 모두 분류오류가 있는 경우 개인수준으로 외삽된 위험요인의 비교위험도를 수학적으로 살펴보았을 때 생태학적 모형 1은 위험요인의 분류오류와 생태학적 오류로 인한 치우침이 존재하였고, 모형 2는 위험요인의 분류오류, 혼란변수의 분류오류, 그리고 생태학적 오류로 인한 치우침이 존재하였다. 그러나 모형 4는 단지 위험요인의 분류오류로 인한 영향만 존재하였다.

2) 시뮬레이션 결과 생태학적 모형 2는 분류오류가 없는 경우에도 모형의 부적절성으로 인한 생태학적 오류가 아주 크게 존재하였고 또한 위험요인과 혼란변수의 분류오류로 인한 영향도 매우 크게 받아 모형 2와 모형 3에 비해 상대적으로 매우 좋지 않은 모형이었다.

3) 위험요인에 분류오류가 없는 경우 모형 1은 모형의 부적절성으로 인한 추정치의 치우침이 존재하였고 참 위험요인의 비교위험도가 커질수록 그 정도가 점점 커졌으나 모형 3은 상대적으로 그 치우침의 정도가 작았다. 그러나 모형 3은 위험요인의 분류오류가 커질수록 추정치의 치우침은 점점 커졌으며 모형 1은 큰 변화가 없었다. 한편 분류오류의 영향을 수학적으로 수정한 비교위험도를 보면 모형 3은 참 비교위험도에 매우 근접하였으나 모형 1은 모형의 부적절성으로 인한 추정치의 치우침이 여전히 존재하였다.

4. 고찰

위험요인과 질병과의 관계를 알아보고자 하는 역학적 연구에서 생태학적 연구가 가지는 많은 단점에도 불구하고 자료접근의 용이성으로 아직까지 이러한 연구를 수행하는 경우가 있다. 이 연구는 개인수준에서의 위험요인과 질병과의 관계가 지수모형인 경우 만약 위험요인과 기타 혼란변수에 분류오류가 없거나 또는 분류오류의 크기를 정확하게 알수 있는 경우에는 생태학적 모형 3을 선택하는 것이 좋다는 것을 제시한다. 그러나 분류오류가 존재하고 그 크기를 정확히 알 수 없는 경우에는 모형 2를 선택하는 것이 좋다는 것을 아울러 제시하여 준다. 한편 일반적으로 생태학적 연구에서 많이 사용되는 모형 2는 모형의 부적절성으로 인한 오류와 분류오류의 영향으로 인해 생태학적 모형으로는 부적절하다는 것을 또한 제시하여 준다.