

번호: OP-C-005					
제 목	자기기입식 설문자료에서 이상치 판별과 인구사회학적 요인과의 연관성 The Outliers in Questionnaire and their Associated Factors				
저 자 및 소 속	최성용 1), 김기랑 1), 조수희 1), 박태성2), 김미경1), 최보율1) 1)한양대학교 예방의학교실, 2) 서울대학교 통계학과 Sungyong Choi 1), Kirang Kim 1), SuHee Jo 1), Taesung Park 2), Mi Kyung Kim 1), Boyoul Choi 1) 1) Hanyang University, School of Medicine, Department of Preventive Medicine, 2) Seoul National University, Department of Statistics				
분 야	역 학 [의료정보 및 통계]	발 표 자	최성용 일반회원	발 표 형 식	구 연
<p><b>목적:</b> 역학연구에 있어 자료를 수집하는 방법으로는 설문조사방법과 기계를 이용한 측정방법이 있다. 표준화된 기계를 이용한 측정방법의 경우 이상치는 거의 존재하지 않는다. 하지만, 설문조사방법의 경우 면접원이나 응답자에 의해 많은 영향을 받는다. 특히 자기기입식 설문조사방법의 경우 응답자의 심리적 요인과 오류, 조사 도구의 문제, 랜덤 오류 등의 원인으로 인해 이상치가 발생하며, 이상치 응답을 판별방법으로는 평균±3표준편차, 다변량 기법, 통계적 모형, 신경망구조등의 방법이 있다. 본 연구에서는 (i) Grubbs(1950)가 제안한 방법을 응용하여 이상치 응답을 판별하고, (ii) 이상치 응답자들의 인구사회학적 연관성을 분석하고자 한다.</p> <p><b>방법:</b> 이 연구의 대상 설문은 자기 기입 방법으로 조사하는 PWI-SF이었다. 2001년부터 2005년 1월까지 양평군에서 시행하였던 양평코호트연구의 기반조사에 참여하였던 주민 2,683명 중 65세 이상이며, 독해가 불가능하여 자기 기입 방법으로 설문조사를 시행할 수 없었던 사람을 제외한 1,968명을 분석 대상으로 하였다. 이 연구에서는 Grubbs(1950)가 제안한 방법을 이용하여 이상치 응답을 판별하였다. 또한 제안한 방법에 의해 이상치 응답을 구분 후, 이상치 응답자들의 인구사회학적 연관성을 판별분석을 통해 분석하였다. Grubbs가 제안한 방법은 문항별 보편적인 응답으로부터 많이 벗어난 응답일수록 이상치 응답으로 간주하여 판단하는 방법이다. Grubbs에 의한 방법은 설문지와 같이 응답의 범위가 제한적인 경우에 적용하기가 적당하며, 이상치 판별을 위한 계산방법이 비교적 간단하다는 장점이 있다. 이 연구에서 보편적인 응답을 벗어난 정도를 측정하기 위한 도구로 아래의 측도를 사용하였다. 여기서, <math>m_j</math>와 <math>s_j</math>는 j번째 문항에 대한 전체 응답자들의 평균과 표준편차이다. 이 연구에서는 <math>Z_i</math>값의 상위 5%를 이상치로 간주하였다.</p> $Z_i = \sum_{j=1}^k \left  \frac{x_{ij} - m_j}{s_j} \right $ <p><b>결과:</b> Grubbs 제안한 방법에 의해 판별된 이상치 응답자는 98명(4.9%), 정상응답은 1,870명(95.1%)이었다. 이상치 응답과 성별과의 연관성에서 성별 차이가 존재하였으며(<math>p &lt; 0.01</math>), 여자가 남자에 비해 약 2.25배 만큼 높았다. 교육수준과 이상치의 응답의 연관성에서 교육수준에 따른 차이는 존재하였으며(<math>p &lt; 0.01</math>), 학력수준이 낮을수록 이상치의 응답비율이 증가하였다. 결혼상태에 따른 차이는 존재하여 별거/이혼/사별인 집단의 이상치 응답율이 높은 것으로 나타났다(<math>p &lt; 0.01</math>). 직업과 이상치 응답의 연관성에서 직업에 따른 연관성이 존재하며(<math>p &lt; 0.05</math>), 무직인 사람의 이상치 응답비율이 높았다. 또한, 연령과 가계수입 수준에 따라서도 이상치 분포가 다른 것을 발견하였다. 현재, 다른 이상치 응답 판별방법을 이용하여 추가 분석을 시행하고 있으며, 이들 방법들 간의 장단점을 비교 정리하여 제시할 예정이다.</p>					