<table>
<tr><td colspan="3">역학 III(분석방법)</td><td colspan="3">번호: III - H - 6</td></tr>
<tr><td rowspan="2">제 목</td><td>국문</td><td colspan="4">SAS의 GENMOD와 MULTTEST procedure들을 이용한 r*c 분할표의 다중비교 방법 소개 - 두개내 뇌동맥 및 두개외 경동맥 동맥경화증 발현여부와 관상동맥 동맥경화증 정도간의 관련성 분석 -</td></tr>
<tr><td>영문</td><td colspan="4">Introducing a multiple comparison test for a r*c contingency table using the GENMOD and MULTTEST procedures in SAS.</td></tr>
<tr><td rowspan="2">저 자
및
소 속</td><td>국문</td><td colspan="4">이준영[1], 배희준[2]
1) 고려대학교 의과대학 예방의학교실,
2) 을지의과대학교 을지병원 신경과</td></tr>
<tr><td>영문</td><td colspan="4">Juneyoung Lee[1], Hee-Joon Bae[2]
1) Department of Preventive Medicine,
College of Medicine, Korea University,
2) Eulji General Hospital, Eulji University School of Medicine</td></tr>
<tr><td>분 야</td><td>역 학
기타</td><td>발 표 자</td><td>이준영
일반회원</td><td>발표형식</td><td>구 연</td></tr>
<tr><td>진행상황</td><td colspan="5">연구완료</td></tr>
</table>

1. 연구 목적:

다중비교(multiple comparison, 또는 post-hoc test)란 다수의 집단들이, 통계적으로 모두 동일한 특성을 가지지는 않는 것으로 판단되는 경우, 두 집단씩 짝을 지어 통계적인 검정을 실시하여 어느 집단끼리 유의한 차이를 보이는지 알아보는 방법이다. 하지만 모든 가능한 짝에 대해 일정한 유의수준(예, 0.05)하에서 통계적인 검정을 실시하게 되면 전반적인 유의수준(overall level of significance)은 우리가 실제로 원하는 유의수준 0.05를 훨씬 초과하게 된다. 따라서 각 짝 진 비교에서 얻어지는, 관찰된 유의수준(observed level of significance, 즉, p-value)들을 보정(adjust)해 주어야 하는데, 이것이 바로 다중비교 내용의 핵심이다.

분산분석(ANOVA)의 경우 다중비교 방법론이 많이 적용되고 있지만, r*c 분할표로 나타나는 범주형 자료의 경우, 전체 rc개 칸들을 대상으로 한 다중비교 방법은 아직 널리 사용되고 있지 않은 실정이다. 이는 연구자들이 이 다중비교 방법을 알지 못하거나 이 방법의 유용성을 아직 인식하지 못하고 있는 데에도 일부 기인하겠지만, 또 한편으로는, 범주형 자료의 경우, SAS나 SPSS 등과 같은 기존의 통계분석용 프로그램들이 다중비교 방법을 손쉽게 적용할 수 있도록 준비되어있지 않은 데에도 그 원인이 있다고 할 수 있겠다.

따라서 본 연구에서는 분할표에서 얻어지는 각 칸들에 대해 다중비교를 실시하는 방법을 소개하고, 이 방법을 적용해 두개내 주요 뇌동맥 동맥경화증(Intracranial Cerebral Atherosclerosis)과 관상동맥 동맥경화증(Coronary Atherosclerosis) severity 정도간의 관련성 및 두개외 경동맥 동맥경화증(Extracranial Carotid Atherosclerosis)과 관상동맥 동맥경화증 severity 정도간의 관련성 여부를 보다 심도 있게 분석하고자 한다.

2. 연구 방법:

SAS의 MULTTEST procedure는 여러 개의 통계적 검정을 동시에 실시할 때 발생하는 전체 유의수준의 증가문제를 조절하기 위해 각 검정결과들의 p-value들을 보정해주는 프로그램이다.

한편, r*c 분할표의 경우, 일반 카이제곱 검정을 사용해 두 변수간의 관련성 여부를 파악할 수도 있지만, 각 칸의 값들은 전체 표본의 수에서 포아송 분포에 근거해 얻어진다는 점을 이용한 포아송 회귀모형(Poisson regression model)을 사용하면 좀더 심도 있는 분석을 할 수 있다. (포아송 회귀모형은, 반응값이 두 범주인 경우에 사용되는 로지스틱 회귀모형에 대응하는, r*c 분할표에 대한 통계모형이라 할 수 있다) 이 모형은 SAS의 GENMOD procedure를 사용해 분석될 수 있다.

따라서 SAS의 GENMOD procedure를 통해, r*c 분할표의 칸들간 평균 차이에 관한 짝 진 비교를 실시하고, 여기서 얻어진 p-value 들에 대해 SAS의 MULTTEST procedure를 이용한 보정을 실시하면 r*c 분할표 자료에 대한 다중비교를 할 수 있게 된다. 참고로 이때 사용되는 각 칸들의 평균값으로는, 산술평균이 아닌, 회귀모형의 다른 효과들을 적절히 보정시킨 최소제곱평균(least squares means)을 이용하여야 한다.

3. 연구 결과:

동일한 연구대상에 대해 관상동맥의 동맥경화증(CA)과 두개내 뇌동맥의 동맥경화증(ICAS) 및 두개외 경동맥의 동맥경화증(ECAS)의 상관관계를 조사한 연구는 아직 알려진 바 없다. 이를 파악하기 위해, coronary artery bypass graft(CABG) surgery를 실시한 246명의 환자들을 대상으로 ICAS 및 ECAS의 존재유무(Absent/Present), 그리고 관상동맥조영술에 근거한 CA 정도(Mild/Moderate/Severe)를 조사하였다.

카이제곱 분석결과, ICAS와 CA 정도간에는 관련이 없는 것으로 나타나, 이들간의 상호작용으로 고려하지 않은 포아송 회귀분석을 실시하였다. 각각의 수준에 대한 다중비교를 실시한 결과, CA 정도 효과를 보정시킨 상태에서 연구대상자의 ICAS 여부에는 차이가 없는 것으로 나타났으나, ICAS 여부에 관한 효과를 보정시킨 상태에서 CA 정도는 mild한 군과 moderate한 군, 그리고 moderate한 군과 severe한 군간에 통계적으로 유의한 차이가 있는 것으로 나타났다.

한편 ECAS와 CA 정도간에는 상호 관련이 있는 것으로 나타났으며, 이에 따라 이들간의 상호작용을 고려한 포아송 회귀분석을 실시하였다. 다중비교 결과, CA 정도가 mild한 군 및 moderate한 군에서는, ECAS가 발현되지 않은 경우가 발현된 경우보다 통계적으로 유의하게 높은 사례를 보이는 것으로 나타났지만, CA 정도가 severe한 군에서는 ECAS 발현여부에 차이가 없는 것으로 나타났다. 따라서 CA는 적어도 그 정도가 severe 하기 전 까지는 ECAS 발현여부에 관한 예후인자로서의 역할 가능성이 있는것으로 보인다.

4. 고찰:

일반적인 분할표 분석에서 두 변수간 관련성이 있는 것으로 보이는 경우, 사후분석으로 각 변수의 수준 조합 별 비교가 필요하게 된다. 본 연구에서는 이와 같은 분석이 포아송 회귀모형을 이용한 다중비교 방법을 통해 이루어질 수 있음을 소개하였다.

짝을 지어 비교하는 다중비교 결과의 해석은 해당 짝에 대해서만 이루어져야지 짝들간의 결과에 대해 삼단논법 형태의 추론방식을 적용해서는 안 된다. 이에 덧붙여, 다중비교는 가능한 모든 짝들을 기계적으로 비교하는 것이기 때문에, 임상적인 해석이 곤란한, 유의한 짝들에 대해서까지 다중비교 결과의 의미를 해석하려는 시도는 적절하지 않다고 본다.