

반복측정된 실험자료 분석에 관한 고찰

하일도¹, 노규정², 고정환³

요약 본 논문은 의뢰인의 Pilot Study를 상담한 것으로서 당뇨병 및 암 환자에게 효능이 있는 약으로 밝혀진 Steroid계통의 Methyl Prednisolone이 척수손상 환자에게 효능이 있는지를 알아보기 위해, 토끼를 실험대상으로 하여 얻은 반복측정자료를 분석하였다.

1. 서론

반복측정자료(Repeated measures data)란 동일 실험개체(Subject)에 대해 여러가지의 다른 시간이나 실험조건으로부터 반복측정을 하여 얻어진 자료를 말한다. SAS, SPSS, GLIM, GENSTAT 등 여러 패키지의 개발로 인해, 최근 의약학 및 행동과학 분야 등의 연구에서 실험을 통해 얻어진 반복측정자료의 분석을 흔히 다루고 있다. 대부분의 경우 반복측정실험을 하는 중요한 이유는 실험개체의 수가 적어서 충분한 자료를 얻기가 힘들기 때문이다. 예를 들어, 실험의 대상이 어떤 특수한 질병을 가진 환자군 또는 실험개체의 관리에 주의를 많이 기울여야 하는 경우 등에서는 반복측정실험을 실시하면 충분한 자료를 얻을 수 있어서 효과적인 분석을 할 수 있다.

한편, 반복측정자료는 반응값의 형태(연속형, 이산형)에 따라 다양한 통계모형을 사용하여 추정과 검정을 할 수 있다. 대표적인 모형으로 연속형인 경우에는 정규분포에 기초한 선형모형을 사용한다. 이산형인 경우에는 명목형 자료(Nominal data)와 순서형 자료(Ordinal Data)로 나눌 수 있는데, 명목형 자료에는 일반화로지트모형(Generalized logit model)을 순서형 자료에는 비례오즈모형(Proportional odds model)이 각각 사용된다.

본 논문은 의뢰인의 의학 Pilot Study를 상담한 것으로서 다음과 같이 반복측정된 실험자료를 분석하여, 당뇨병 및 암 환자에게 효능이 있는 약으로 밝혀진 Steroid계통의 Methyl Prednisolone이 척수 손상 환자에게도 효능이 있는지를 고찰해 보고자 한다. 우선 토끼의 척수를 손상시킨 후 Methyl Prednisolone을 투약하여 각 시점별로 토끼의 운동신경 반응을 측정함으로써 이 약의 효능 여부를 조사 분석한다. 부차적으로 척수 손상 전 토끼의 포도당과 척수 손상 후 토끼의 운동 신경 반응과의 영향 관계를 알아보려고 한다.

1 서울대학교 자연과학대학원 계산통계학과 통계학전공 박사과정
2 서울대학교 의과대학원 의학과 마취과전공 박사과정
3 안동대학교 자연과학대학 통계학과

2. 실험방법 및 자료의 구조

연구의 목적을 달성하기 위한 실험 절차 및 자료 수집 과정은 다음과 같다.

절차 1, 시장에서 토끼 22마리를 구입한 후 임의로 대조집단(Control Group)에 12마리, 실험집단(Experimental group)에 10마리를 각각 할당한 후 모든 토끼의 포도당을 측정한다.

절차 2, 두 집단 22마리 모두 똑같은 조건으로 20분간 척수 혈관을 막아서 척수 손상 후, 실험집단만 Steroid계통의 Methyl Prednisolone을 투여하고 대조집단은 이 약을 투여하지 않는다.

절차 3, 주어진 신경학적 검사 점수 채점표에 근거하여, 2시간, 1일, 2일, 3일 후 각각 토끼의 운동신경 반응의 점수를 측정한 후 자료를 얻는다.

이러한 실험 결과, 관측된 자료는 순서형 반복측정 자료(Repeated ordered categorical response data)이다. 그 자료는 표 1과 같다. 표 1의 두드러진 특징 및 의문점을 살펴보면, 대조집단에서는 죽은 토끼(4, 9, 11번째)가 3마리이고 3일간 계속 점수가 0(7, 8번째)과 4(10, 12번째)인 토끼가 각각 2마리이다. 그리고 실험집단에서는 3일간 계속 점수가 0인 토끼가 많은 편이다. 한편 포도당 측정에서 4개의 결측치(*)가 있다.

3. 의뢰인과의 토론

표 1에 나타난 두드러진 특징과 의문점에 대해서는 의뢰인과 다음과 같은 충분한 협의의를 거친 다음 자료를 분석했다.

첫째, 대조집단에서 죽은 자료는 결측치가 아니고 점수 채점표에도 없으므로 분석에서 제외하기로 했다.

둘째, 대조집단에서 3일간 계속 점수가 0과 4인 자료는 척수의 손상 정도에 의문점이 생길 것 같지만 똑같은 실험 조건이었고 점수의 등급 측면에서의 자료이기 때문에 분석에 포함시켰다.

셋째, 실험집단에서 3일간 계속 점수가 0인 자료는 투여한 약의 효과일 가능성도 있기 때문에 분석에 포함시켰다.

4. 통계적 분석

4.1 자료의 정리

표 1의 자료를 정리하면 다음 표 2와 같다.

표 2를 상대 도수 측면에서 보면, 각 시점에서 각각 대조집단은 점수 4, 실험집단은 점수 0의 비중이 높다. 또한 모든 시점에서 평균(Mscore)을 살펴보면 대조집단이 실험집단보다 높다는 것을 알 수 있다. 결국 실험에 사용된 약의 효능은 없음을 어느 정도 추측할 수 있다.

표 2의 각 집단에서의 시점간 변화 양상을 보다 쉽게 파악하기 위해, 꺾은선 그래프로 나타내면 다음 그림 1과 같다.

표 1 척수 손상 유발된 토끼의 자료

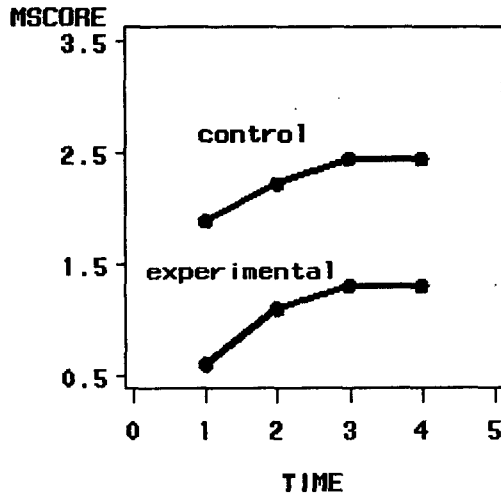
| 집단 | 토끼 | 척수 손상전 포도당(mg/ml) | 점 수 | | | |
|----|----|----------------------|--------|--------|----|----|
| | | | 2시간 | 1일 | 2일 | 3일 |
| 대조 | 1 | 142 | 3 | 3 | 1 | 1 |
| | 2 | 168 | 1 | 3 | 4 | 4 |
| | 3 | 142 | 1 | 1 | 4 | 4 |
| | 4 | 220 | expire | | | |
| | 5 | 185 | 0 | 3 | 4 | 4 |
| | 6 | 133 | 4 | 2 | 1 | 1 |
| | 7 | 134 | 0 | 0 | 0 | 0 |
| | 8 | 184 | 0 | 0 | 0 | 0 |
| | 9 | 129 | 0 | expire | | |
| | 10 | 132 | 4 | 4 | 4 | 4 |
| | 11 | * | 0 | expire | | |
| | 12 | * | 4 | 4 | 4 | 4 |
| 실험 | 1 | * | 1 | 1 | 1 | 1 |
| | 2 | 154 | 0 | 3 | 4 | 4 |
| | 3 | 156 | 3 | 4 | 4 | 4 |
| | 4 | 163 | 0 | 0 | 0 | 0 |
| | 5 | * | 2 | 3 | 4 | 4 |
| | 6 | 257 | 0 | 0 | 0 | 0 |
| | 7 | 210 | 0 | 0 | 0 | 0 |
| | 8 | 255 | 0 | 0 | 0 | 0 |
| | 9 | 132 | 0 | 0 | 0 | 0 |
| | 10 | 189 | 0 | 0 | 0 | 0 |

신경학적검사 점수 채점표

- 점 수 0 : 하반신 마비와 함께 하지의 운동 기능이 전혀 없다.
 1 : 하반신의 움직임이 약하다. 약간 움직이거나 누르면 약간 움직임을 보인다.
 2 : 약간의 하지 움직임이 있고 누르면 지탱하는 힘이 있으나 다리를 몸쪽으로 끌거나 뗄 수는 없다.
 3 : 다리를 몸쪽으로 끌 수 있고 뗄 수도 있으나 정상적이지 못하다.
 4 : 정상적인 운동 기능이 있다.

표 2 자료의 정리

| 도수 (열%) | 대조 집단 | | | | 실험 집단 | | | |
|------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | 2시간 | 1일 | 2일 | 3일 | 2시간 | 1일 | 2일 | 3일 |
| 점수 0 | 2 (22.22) | 2 (33.22) | 2 (22.22) | 2 (22.22) | 7 (70.00) | 6 (60.00) | 6 (60.00) | 6 (60.00) |
| 1 | 2 (22.22) | 1 (11.11) | 2 (22.22) | 2 (22.22) | 1 (10.00) | 1 (10.00) | 1 (10.00) | 1 (10.00) |
| 2 | 0 (0.00) | 1 (11.11) | 0 (0.00) | 0 (0.00) | 0 (0.00) | 1 (10.00) | 0 (0.00) | 0 (0.00) |
| 3 | 1 (11.11) | 3 (33.33) | 0 (0.00) | 0 (0.00) | 1 (10.00) | 2 (20.00) | 0 (0.00) | 0 (0.00) |
| 4 | 3 (33.33) | 2 (22.22) | 5 (55.56) | 5 (55.56) | 0 (0.00) | 1 (10.00) | 3 (30.00) | 3 (30.00) |
| 도수합 | 9 | 9 | 9 | 9 | 10 | 10 | 10 | 10 |
| 평균 | 1.89 | 2.22 | 2.44 | 2.44 | 0.60 | 1.10 | 1.30 | 1.30 |
| 표준편차 | 1.83 | 1.56 | 1.88 | 1.88 | 1.07 | 1.60 | 1.89 | 1.89 |



GROUP: c=control group
 e=experimental group
 TIME: 1=2 hours
 2=1 day
 3=2 days
 4=3 days
 GROUP ●—●—● c ●—●—● e

그림 1 각 시점별 평균의 꺾은선 그래프

그림 1의 특징을 살펴보면 다음과 같다.

첫째, 대조집단이 실험집단보다 모든 시점에서 평균이 더 크게 나타나고, 두 집단에서의 그래프가 평행에 가깝다.

둘째, 집단간에는 차이가 있지만 시점간에는 차이가 거의 없는 것 같다.

셋째, 시간과 집단간의 교호작용이 없다.

다음 단계로 이러한 특징들의 통계적 유의성 여부를 구체적으로 분석하였다.

4.2 각 시점에서 두 집단간 차이의 검정

어느 시점에서 약의 효과가 있는지 알아보기 위해 각 시점에서 대조집단과 실험집단의 차이 검정에 Wilcoxon의 순위합 검정과 T-검정을 사용하였고, 분석한 결과는 다음 표 3과 같다.

표 3 각 시점에서 두 집단간 차의 비교

| 시 간 | P-값(Wilcoxon검정) | P-값(T-검정) |
|------|-----------------|-----------|
| 2 시간 | 0.084 | 0.089 |
| 1 일 | 0.144 | 0.140 |
| 2 일 | 0.156 | 0.204 |
| 3 일 | 0.156 | 0.204 |

표 3에 의하면 두 집단간 차이가 있는 시점이 없다고 할 수 있다. 즉, 약의 효능이 있는 시점이 없음을 알 수 있다.

4.3 분산 분석

순서형 반복측정자료에 대한 분석법이 거의 알려져 있지 않는 실정이어서 다소 정보의 손실은 수반되었지만 주어진 자료를 연속형으로 간주하고 분석을 실시하였다. 이 자료를 분석하기 위하여 다음과 같은 분할법 모형을 적용하였다.

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + \delta_{ik} + \varepsilon_{ijk},$$

$$i = 1, 2; \quad j = 1, 2, \dots, n_i (n_1 = 12, n_2 = 10); \quad k = 1, 2, 3, 4$$

여기서 μ : 전체 모평균,

α_i : i 번째 집단의 효과 ($i=1$ =대조, $i=2$ =실험),

$\beta_{j(i)}$: i, j 번째의 오차 $\sim iid N(0, \sigma_\beta^2)$

γ_k : k 번째 시점의 효과,

δ_{ik} : i 번째 집단과 k 번째 시점간의 교호작용,

ε_{ijk} : i, j, k 번째의 오차 $\sim iid N(0, \sigma_\varepsilon^2)$

이 모형에 근거한 분산분석 결과는 다음 표 4와 같다.

표 4 분산분석

| Source | F 값 | P-값 |
|--------|-------|-------|
| 모형 | 8.850 | 0.001 |
| 시간 | 1.890 | 0.144 |
| 집단 | 2.860 | 0.109 |
| 집단*시간 | 0.030 | 0.992 |

표 4에 의하면 모형에 관한 전체적인 적합성은 유의하지만 시간에 따른 점수의 차이, 집단간 점수의 차이, 집단과 시간간의 교호작용 모두 유의하지 않다. 따라서 연구의 주목적인 집단간의 차이가 없으므로 실험된 약의 효능은 없음을 알 수 있다.

4.4 상관 분석

여기서는 연구의 부차 목적을 분석하기로 하겠다. 즉, 척수 손상전의 토끼의 포도당과 척수 손상 후 이 약을 투여한 토끼의 운동신경 반응(각 토끼의 점수의 합)과의 상관관계를 분석한 결과는 다음 표 5와 같다.

표 5 포도당과 점수간의 상관분석

| 구분 \ 상관계수 | Kendall | Pearson |
|-----------|----------------|----------------|
| 대조집단 | -0.144 (0.733) | -0.038 (0.899) |
| 실험집단 | -0.411 (0.274) | -0.367 (0.250) |

표 5에 의하면 대조집단보다 실험집단이 약간 높은 음의 상관관계 즉, 포도당이 높을수록 운동신경 반응이 나빠지는 경향이 척수 손상 유발후 약이 투여된 실험집단이 다소 강하다. 하지만 통계적 유의성은 없음을 알 수 있다.

5. 결론 및 연구과제

본 논문은 Steroid 계통의 Methyl Prednisolone이 척수 손상 환자에게 효능이 있는지 알아보기 위해 반복측정된 실험자료를 분석결과 다음과 같은 결과를 얻었다.

척수 손상 후 실험에 사용된 약의 효능 그리고 척수 손상전 토끼의 포도당과 척수 손상 후 토끼의 운동 신경 반응과의 상관관계 두 가지 모두 통계적 유의성이 없음을 표 4 와 표 5로부터 각각 알 수 있었다.

다음으로 본 논문에서의 문제점 및 연구 과제를 살펴보면 다음과 같다.

첫째, 처리 배치에서 확률화(Randomization)의 고려가 결여된 점이다. 의뢰인이 토끼 22마리를 대조집단과 실험집단에 각각 11마리씩 균등한 확률로 할당하여 실험함으로써 생각지도 못한 교락요인을 제거할 수 있겠다. 또한 토끼의 유전적 차이(genetic difference)로 인해 반응의 변이가 크기 때문에 한 배(Litter)에서 태어난 토끼들을 대상

으로 하여, 짝짓기 기법을 사용하여 두 집단을 나누어 실험하는 방법도 생각해 볼 수 있겠다.

둘째, 죽은 자료를 처리 분석하는 방법상의 어려운 점이다. 죽은 자료도 하나의 정보이기 때문에 분석에서 제외한다는 것이 최선은 아닌 것 같다. 죽은 자료가 포함된 점수 채점표 작성을 의뢰인에게 건의했지만 주어진 점수 채점표가 아주 중요하기 때문에 이 채점표만으로 분석할 것을 요구했다.

셋째, 순서형 반복측정자료를 분석하기가 쉽지 않다는 점이다. 일반적으로 측정된 반응 자료가 순서형 반복측정자료일 때 적용하는 비례오즈모형을 근거로 하여 주어진 자료를 분석하기 위해서, $5^4 = 625$ 개의 cell를 고려해야 하지만 625개의 cell에 대한 해당 도수가 0이 너무 많아서 의미도 없다는 점과 알고리즘 구현도 힘들다는 점이다. 이를 위해 점수의 그룹화도 고려할 수 있지만 앞에서 언급한 바와 같이 주어진 점수 채점표만을 사용해야 한다는 점 때문에 측정된 자료를 연속형으로 간주하고 분산분석을 실시하였으므로 다소 정보의 손실이 있었을 것으로 추측된다.

끝으로 만약 위의 3가지 문제점이 적절하게 해결되었을 때, 실험에 사용된 약의 효능에 대한 통계적 유의성이 없더라도 의학적 의미는 있을 수 있다. 왜냐하면 척수 신경 손상 환자에게 이 약을 사용하지 않으면 될 것이고 차후 이 약의 연구에 대한 방향을 제시해 주기 때문이다.

참 고 문 헌

박용구, 송혜양 (1991), 반복측정 자료의 분산분석법, 자유아카데미.

허명희 (1993), 선형모형 방법론, 자유아카데미.

Alan Agresti (1990), *Categorical Data Analysis*, John Wiley & Sons, New York.

Stram, D. O., Wei, L. J. and Ware, J. H. (1988), *Analysis of Repeated Ordered Categorical Outcomes with possibly missing observations and time-dependent covariates*, Journal of American Statistical Analysis, Vol. 83, pp. 631-637.

Stuart, R. L. and Kyungmann Kim (1994), *Analysis of Repeated Categorical Data using generalized estimating equations*, Statistics in Medicine, Vol. 13, pp. 1149-1163.