

실험연구에 기반한 통계자료 분석: 비동등성 대조군 전후 설계

김 서 영*

I. 서 론

(1) 연구배경

최근 거의 모든 학문분야에 걸쳐 통계적 방법을 이용한 연구 결과가 속출하고 있다. 특히 사회과학, 보건복지 및 의약학 분야의 실험조사를 통한 연구 결과에 대해 정확한 통계적 방법에 의한 과학적 검증이 요구되고 있다. 이들 연구 분야는 분류기준에 따라 크게 실험유무에 따른 분류와 시간의 흐름에 따른 분류를 하고 있다. 이때 연구유형에 따라 조사설계 및 연구설계 방법이 다르고 이에 따라 적용되는 통계적 분석방법도 다르다. 각 연구설계 방법에 적합한 통계 분석 방법을 적용하기 위해서는 우선적으로 각 연구방법들의 특성을 명확하게 이해할 필요가 있다.

1) 실험유무에 따른 분류

실험유무에 따라 실험 연구(experimental study)와 비실험 연구(non-experimental study)로 나눌 수 있다. 실험연구란 실험실에서 시행되는 물리적 실험에 의한 연구만이 아니라, 사회과학 분야에서 널리 사용되는 인과론적 견지에서 증명하고자 하는 사전 인자와 사후 예측되는 결과 이외의 가능한 여러 간섭변수를 통제함으로써 얻고자 하는 사실만을 관찰하는 연구방법을 말한다.

비실험 연구는 예상되는 제3의 변수를 통제하지 않은 상태에서 인과적 관련성에 관한 논리적 근거를 찾아내는 방법으로, 일반 인구집단을 대상으로 하는 역학적 연구들이 이 분야에 속하는

*통계청 통계개발원 통계사무관

방법들이다. 비실험 연구는 두 개 이상의 집단으로부터 어떤 문제의 원인 또는 이들과 연관된 근거들을 밝힘으로써 문제를 해결하고자 하는 방법으로, 문제집단과 그렇지 않은 집단과의 비교연구(comparative study)를 시행하는 경우가 많다. 한편 비교연구는 연구자가 중재를 행하지 않고 관찰·조사하는 연구가 되기 때문에 관찰연구(observational study)로 분류되기도 한다(전우택, 1999). 연구자가 시행하는 많은 조사연구에 있어서 집단의 특성을 설명하는 것과 함께 비교에 중점을 두는 경우가 많다.

2) 시간적 순서에 따른 분류

시간적 순서에 따라 횡단면적 연구(cross-sectional study)와 추적연구(longitudinal study)로 분류할 수 있다. 횡단면적 연구는 연구대상에 대한 측정을 한번만 시행하는 경우로, 각 대상으로부터 자료 수집기간이 달라서 연구가 장기화 되는 경우라도 자료수집이 한번만 행해졌다면 횡단면적 연구에 속한다. 많은 연구들이 여기에 속하며 수행하기가 쉽기 때문에 질병의 원인을 파악하기 위한 목적으로 수행되고 있다. 그러나 한 번의 시행에 의한 단면적 연구로 인과관계를 추론하기에는 제약이 따르기 때문에, 자료를 해석하는 과정에서 이점에 대한 충분한 고려가 있어야 한다.

추적연구는 대상으로부터 자료 수집을 두 번 이상 시행함으로써 일정한 시간에 걸쳐 연구대상을 추적하는 연구를 말한다. 횡단면적 연구에 비해 시간과 노력이 많이 소요된다는 특성을 갖는다. 추적연구는 다시 시간경과에 따라 자료수집 시점 혹은 과거에 발생했던 것만을 다루는 후향적 연구(retrospective study)와 연구대상을 확정된 후 사건이 발생하기까지 추적한 후 이를 파악하는 전향적 연구(prospective study)로 나눌 수 있다.

3) 실험연구의 분류

실험연구는 연구를 제어할 수 있는 상황에 따라 실험 연구와 유사실험 연구(quasi-experimental study)로 분류될 수 있다. 실험연구는 연구자가 결과에 영향을 미칠 수 있는 여러 요인들을 직접 통제하고 관리할 수 있다는 측면에서 가장 정확한 방법이라 할 수 있다. 실험연구를 시행할 때는 실험의 대상이 되는 실험군(experimental group)과 비교하기 위한 대조군(control group)이 필요하고 대조군이 없는 실험을 하게 되면 연구자의 편견에 따라 왜곡된 결론을 내리기 쉽다.

한편, 실험연구의 경우 실험군과 대조군을 랜덤하게 할당하기 어려운 경우가 종종 발생한다. 이러한 대상자 할당 방법에 의한 실험군과 대조군에 대해서 실험군에는 중재를 하고 대조군에는 중재를 하지 않은 집단 간의 비교방식을 유사실험 중재연구(Intervention study)라 한다. 이러한 실험방법은 조사표 개발을 위한 연구라든가 제어하기 힘든 집단 간 비교를 목적으로

하는 연구 등에서 자주 사용된다.

(2) 연구의 목적

조사연구는 연구에 앞서 기본적으로 조사의 목적에서부터 결과해석에 이르기까지 조사에 관련된 모든 사항을 조사 설계 단계에서 구체적으로 명시하고, 이 절차에 따라 체계적인 수행이 이루어질 때 성공적인 연구결과를 얻을 수 있다고 보여진다. 조사연구의 경우, 경제적·시간적 비용을 투자가 많기 때문에 한 번의 계획된 조사에 의해 유익한 정보를 도출할 수 있어야 하고, 이를 위해서 자료에 적합한 분석방법의 선택은 매우 중요하다 할 것이다. 그러나 조사설계 단계에서 통계 분석 방법을 사전에 정하지 못하는 경우가 종종 발생하고, 이는 연구결과에 대해 오류를 범하게 하는 원인이 될 수 있다는 것을 간과해서는 안된다.

본 연구에서는 조사표 설계를 위한 사전연구 및 중재효과를 파악하고자 하는 연구에서 가장 일반적으로 사용되고 있는 유사실험 방법을 살펴봄으로써 올바른 통계분석방법의 필요성에 대해 살펴보고자 한다. 본 자료에서는 유사실험 방법 중 가장 자주 사용되는 비동등성 대조군 전후설계(nonequivalent control group pretest-posttest design)를 중심으로 이에 적합한 통계적 분석방법을 설명함으로써 사회과학 연구자들로 하여금 자료분석 방법에 대한 이해를 돕고자 한다.

2. 유사실험 설계

유사실험 설계 방법은 비동등성 대조군 전후설계, 비동등성 대조군 전후 시차설계(nonequivalent control group non-synchronized design)와 비동등성 대조군 사후 설계(nonequivalent control group posttest only design)등이 자주 사용되고 있다. 그 중에서는 본 절에서 비동등성 대조군 전후 설계 방법에 대해서 설명하기로 한다. 유사실험 설계는 실험설계 만큼 이상적이지는 않지만 실험설계를 실행할 없을 때 대안으로 적절한 방법이다. 사회조사에서는 두 군 간에 랜덤 할당이 불가능한 경우가 많기 때문에 이와 같은 유사실험 방법이 널리 사용되고 있다.

(1) 비동등성 대조군 전후 설계

비동등성 대조군 전후 설계는 대표적인 유사실험 방법으로 실험군과 대조군 설정에 있어서

두 집단을 랜덤하게 할당하지 못했기 때문에 사전조사를 통하여 두 집단의 특성 즉, 종속변수 상의 차이뿐만 아니라 이에 영향을 미치는 외생변수 상의 차이를 탐색하여 통제하는 방법이다. 이때 대조군은 사전조사의 결과가 실험군과 비슷하고 인구학적 특성과 같은 다른 측면에서도 비슷한 조건을 갖추고 있다는 가정을 만족해야 한다. 이때 대상자 선택에 대한 편중이나 제3의 변수의 개입 여부를 확인할 필요가 있다. 실험설계는 다음과 같다.

비동등성 대조군 전후 설계

| | 사전 | 중재 | 사후 |
|-----|----|----|----|
| 대조군 | 관측 | 없음 | 관측 |
| 실험군 | 관측 | 있음 | 관측 |

(2) 예제

비동등성 대조군 전후설계에 대한 간단한 예를 들어보자. 이 자료는 간호교육의 효과를 검증하기 위해 환자들을 대상으로 교육과 상담을 통한 지지간호를 실시하여, 질병에 대한 지식정도가 증가하는지를 살펴보고자 하였다(최옥자, 2005). 입원한 동맥 질환자 중 실험군 40명, 대조군 40명 총 80명을 대상으로, 연구자는 사전에 두 군 각각에게 질환에 대한 지식정도를 측정하였고, 중재(처리)로는 실험군에만 질환에 대한 교육과 상담을 실시하였다. 사후조사는 교육을 하고 일정시간이 지난 후 두 군에게 사전조사와 동일한 방법으로 질환에 대한 지식정도를 측정하였다. 연구가설은 '중재 후 실험군과 대조군 간에 질병관련 4개 요인에 대한 지식정도에 차이가 있을 것인가'를 규명하는 것이다. 여기서 4개 요인이란 질병특성, 위험요인, 투약방법 및 일상생활 변수를 말한다. 참고로 본 연구에서 사용된 통계기법 및 분석결과는 실제자료(최옥자, 2005)의 결과와는 상관없음을 미리 밝혀둔다.

(3) 통계분석방법

예제 자료에 적용 가능한 통계분석 기법을 살펴보자. 연구자들은 사전에 실험군과 대조군 간에 지식정도에 차이가 없다(그룹간 동일성 검정)는 것을 보이기 위해 독립 t-검정 (independent t-test)을 수행한다. 검정결과 통계적으로 그룹간 유의한 차이가 없다면, ① 사후 실험군과 대조군 간에 지식정도의 차이를 독립 t-검정을 통해 검정한다. 사전에 차이가 있다면 사후 결과에 대한 ② 공분산분석 (ANCOVA : Analysis of covariance)을 수행하거나, ③ 사전-사후에 대해서 독립 t-검정을 수행하기도 한다. 여기서 문제를 살펴보자.

첫째, 사전에 두 집단 간 지식정도가 동일하다는 결과가 사후에 대한 영향력 문제까지 뒷받침할 수 있을 것인가 하는 것이다. 이는 공분산분석의 목적에 대한 잘못된 이해로부터 파생된

문제라 여겨진다. 사전에 환자들의 지식정도가 환자 개개인에 따라 동일하지 않기 때문에 이 사전 지식정도는 사후 지식정도에 어느 정도의 영향을 줄 수 있을 것이다. 따라서 사전에 처리 그룹에 따라 지식정도가 동일하다는 결과와 사전의 지식이 사후의 지식정도에 영향을 줄 수 있는지의 문제는 별개의 문제로 보아야 한다.

둘째, 연구가설을 검정하기 위해 어떤 연구자들은 사전과 사후의 지식정도 차이(사전-사후)에 대해 독립 t-검정 기법을 적용하기도 한다. 이 방법은 공분산분석의 특수한 경우로 사전의 지식정도에 대해 사후 지식정도가 정비례적인 관계에 있을 경우에만 타당하다. 즉, 사전 지식정도가 1 증가하면 사후 지식정도도 정확히 측정단위의 1배 만큼 증가하거나 감소하는 경우에만 해당된다. 따라서 두 처리 그룹 간 전후 설계 자료에 대해서는 사전의 영향을 적절하게 제어해 줄 수 있는 분석 방법을 적용하는 것이 바람직하다 하겠다. 공분산분석에 대한 자세한 내용은 참고문헌을 참조할 수 있다(박동권, 1995).

3. 분석결과 비교

2절 예제에서 종속변수는 교육 후 대상자들의 질병관련 지식점수이고, 공변량(covariate)은 교육을 받기 전 지식점수, 처리변수(treatment)는 2개의 집단을 갖는 그룹변수이다. 자료분석의 목적은 중재(교육) 전후 지식정도의 처리 그룹에 따른 유의성을 검정하는 것이기 때문에 분석결과를 해석하기 위한 유의확률 p만 제시하고 나머지 기초통계량들은 생략한다.

(1) 전후 지식점수에 대한 독립 t-검정

<표 1>은 전후 지식정도의 차이에 대한 하위영역별로 분석 결과를 나타낸다. <표 1>에서 [사전 p]는 사전 지식정도, [사후 p]는 사후 지식정도, 마지막으로 [차이 p]는 사후와 사전의 지식 차이점수에 대해 독립 t-검정을 수행한 유의확률 값을 나타낸다.

먼저, 두 처리 그룹 간 사전 지식정도의 동일성 검정을 위해 독립 t-검정을 수행한 결과(사전 p), 유의수준 0.05하에서 투약방법을 제외한 나머지 영역은 두 집단 간 사전 지식정도에는 차이가 없었다. 이럴 경우 많은 연구자는 사전의 두 집단 간에 지식에 관련된 투약방법을 제외한 모든 하부 영역이 동일하다는 가정 하에 사후 결과에 대해서 독립 t-검정을 수행한다. 그 결과(사후 p) 지식전체(p=0.027) 및 일상생활(p=0.028)영역에서 유의수준 0.05하에서 유의하였다. 그러면, 우리는 사후의 검정결과를 신뢰할 수 있을 것인가? 하는 것이다. 즉, 사전에 두 집단 간 지식에 대한 차이가 없다는 결과가 사전과 사후 지식에 대한 영향력 관계까지 검정할 수 있는가 하는 문제이다. 이에 대한 대안으로 연구자들은 사후와 사전점수의 차이에 대해서 독립 t-검정

을 수행하기도 한다. 그 결과(차이 p), 유의수준 0.05하에서 지식전체 (p=0.009), 투약방법 (p=0.001)과 위험요인 (p=0.048)에서 유의한 차이가 있었다. 이 결과는 사후 지식에 대한 t-검정 (사후 p)와 다른 결과를 보여준다. 그러나 이미 서론에서 언급한 바와 같이 사후와 사전의 차이를 검정하는 것은 사전의 영향력을 제어할 수 있는 특수한 경우에 해당 될 뿐, 일반적으로 적용할 수 있는 방법은 아니다.

<표 1> 독립 t-검정에 의한 유의확률

| 변수 | 처리그룹 | 사전 p | 사후 p | 차이 p |
|------|------|--------------|--------------|--------------|
| 지식전체 | 실험군 | 0.227 | 0.027 | 0.009 |
| | 대조군 | | | |
| 질병특성 | 실험군 | 0.515 | 0.279 | 0.189 |
| | 대조군 | | | |
| 위험요인 | 실험군 | 0.390 | 0.202 | 0.048 |
| | 대조군 | | | |
| 투약방법 | 실험군 | 0.017 | 0.224 | 0.001 |
| | 대조군 | | | |
| 일상생활 | 실험군 | 0.630 | 0.028 | 0.112 |
| | 대조군 | | | |

(2) 상관분석

전후 지식정도에 대한 상호 관련성을 살펴보기 위해 상관분석을 수행하였다. 각 지식정도에 대한 피어슨 (pearson)의 이변량 상관계수 및 유의확률은 <표 2>와 같다.

<표 2> 전후 지식에 대한 피어슨의 상관계수 및 유의확률

| | 지식전체 | 질병특성 | 위험요인 | 투약방법 | 일상생활 |
|-------------|------------------|------------------|------------------|------------------|------------------|
| 전후 상관계수 (p) | 0.474 (0.001) | 0.314 (0.004) | 0.542 (0.001) | 0.379 (0.001) | 0.198 (0.076) |

일상생활을 제외한 모든 지식 영역은 상관계수와 유의수준 0.05 기준에서 전후 지식정도 간에 유의한 선형 상관을 갖는다는 것을 알 수 있다. 이 결과로부터 사전은 사후 지식정도에 어느 정도는 영향력을 끼친다는 것을 알 수 있고, 사전 지식정도를 적절하게 제어할 필요가 있다. 따라서 사전 지식을 공변량으로 하는 공분산분석을 수행하는 것이 바람직하다 하겠다.

(3) 공분산분석

공분산분석은 실험오차를 줄이고 정밀성을 높이기 위한 자료분석기법 중 하나이다. 이때 연구자는 처리변수 외에 결과변수에 영향을 주는 제3의 혼란변수 (confounding variable)를 실험에서 통제할 필요가 있고, 이 혼란변수를 공변량이라 한다. 공변량은 실험자에 의해 물리적으로 제어될 수 없는 변수로서, 공분산분석의 핵심은 결과변수에 미치는 공변량의 영향을 보정하는데 있다. 공분산분석의 첫 단계는 공변량이 결과변수에 미치는 영향력 정도를 파악하는 것으로 기울기에 대한 동일성 검정을 수행하는 것이다. 공변량과 처리변수와의 교호작용을 모형에 추가하여 기울기의 동일성을 검정한다. 기울기가 동일하지 않은 경우에 기울기 조정을 통한 분석을 수행해야 한다. 그러나 대부분의 통계패키지(SAS, SPSS)는 기울기가 동일하다는 가정 하에 분석이 이루어지고 있기 때문에 기울기가 동일하지 않은 경우에는 잘못된 결과를 도출할 수 있음을 명심해야 한다. 자세한 내용은 참고문헌을 참고할 수 있다(박동권, 1995).

1) 공분산분석결과

<표 3>은 지식정도가 실험군과 대조군 간에 차이가 있는지에 대한 공분산분석 결과를 나타낸다.

<표 3> 공분산분석 결과

| 종속변수 | 원인 | 유의확률(p) |
|--------|--------|---------|
| 사후지식전체 | 사전지식전체 | 0.001 |
| | 처리변수 | 0.001 |
| 사후질병특성 | 사전질병특성 | 0.003 |
| | 처리변수 | 0.007 |
| 사후위험요인 | 사전위험요인 | 0.001 |
| | 처리변수 | 0.037 |
| 사후투약방법 | 사전투약방법 | 0.001 |
| | 처리변수 | 0.015 |
| 사후일상생활 | 사전일상생활 | 0.054 |
| | 처리변수 | 0.020 |

<표 3>에서 사후투약방법의 경우, 처리변수에 대한 유의확률 $p=0.015$ 는 기울기 조정후의 사후 투약방법에 대해서 두 군간 t-검정을 수행한 결과이다. 또한 사전투약방법의 $p=0.001$ 은 사후와 사전 투약방법과의 선형회귀분석 후에 얻어진 결과를 나타낸다. 나머지 4개 영역은 통계패키지에서 사전지식을 공변량, 사후지식을 종속변수로 하여 공분산분석한 결과이다. 모든 지식관

런 요인들은 실험군과 대조군 간에 유의수준 0.05하에서 유의한 차이를 보이고, 또한 사후일상 생활을 제외한 나머지영역에서는 사전 지식정도가 사후지식정도에 유의한 영향을 미치고 있음을 알 수 있다. <표 3>의 결과는 <표 1>과 다른 결과를 제시하고 있다. <표 1>의 [사후 p]값은 사전의 지식정도가 실험군과 대조군 간에 차이가 없다는 것을 근거로 분석하였지만, 이 결과만으로는 사후에 대한 영향력을 반영할 수 없다는 것을 단편적으로 보여주는 결과이다. 사후와 사전의 차이만을 가지고 독립 t-검정한 결과인 [차이 p]값 역시 사후에 대한 사전의 영향력을 충분히 반영할 수 없음을 보여준다. 따라서 비동등성 대조군 전후 설계에 의한 자료에서처럼 사전의 영향력을 적절하게 제어함으로써 처리효과(중재)에 대한 정확한 검증을 위해서는 전후 영향력 관계를 정확히 파악한 후에 자료에 적합한 공분산분석 기법을 적용하는 것이 바람직할 것이다.

4. 결론 및 제언

지금까지 조사표 설계를 위한 선행연구라든가 다양한 연구들의 중재효과를 파악하기 위해 사용되는 비동등성 대조군 전후설계에 의한 유사실험 연구의 통계분석방법에 대해서 살펴보았다. 이로부터 통계기법의 올바른 적용은 연구 결과의 정확한 해석을 위해 필수적이라는 것을 알 수 있다. 본 연구에서 제시한 방법 외에도 잘못 적용되고 있는 사례는 많다. 이론을 중시한 개괄적인 소개 도 중요하지만, 실제로 사용자가 각 연구 분야에서 쉽고 정확하게 사용할 수 있도록 다양한 응용 사례를 통해서 통계적 방법론을 제공하는 것도 중요한 과제가 아닐까 한다. 이에 실제 응용분야에 대한 통계 이용자들의 적극적인 이해와 관심이 필요할 것으로 사료된다.

참고문헌

- 김혜자 (2002). 갱년기여성의 운동프로그램이 갱년증상, 칼슘, 지질, 심폐기능변화에 미치는 효과, 학위논문.
- 남기성 (2004). 사회과학에서 통계분석 방법의 선택에 관한 연구, *Journal of the Korean Data Analysis Society*, 6, 1255-1266.
- 박동권 (1995). *실험계획법*, 자유아카데미, 서울.
- 전우택 (1998). *사회과학 연구방법론*, 연세대학교 출판부.
- 최옥자 (2005). 관상동맥 조영술 후 지지간호가 관상동맥 질환자의 건강행위 이행에 미치는 효과, 학위논문.
- Rebecca, K. S. and Teresa, L. (2003). Evaluation of Individually Tailored Interventions on Exercise Adherence, *Western Journal of Nursing Research*, 25, 623-640.