

〈Special Issue〉

## 최근 3년간(1999-2001) 대한간호학회지의 통계기법 활용에 대한 분석

강 현 철\*

### 1. 서 론

통계적 기법의 활용은 경영, 경제, 이공, 의약 등 제 학문분야에 걸쳐 확산되고 있으며 사용되는 기법들의 종류도 다양해지고 있다. 또한 적용되는 통계적 자료분석 방법론의 수준도 점차 높아지고 있는 추세이다. 통계적 자료분석 방법론을 사용함에 있어 주로 문제가 되는 것은 전문적 지식의 결여에서 오는 적절하지 못한 기법의 선택이나 내용상의 잘못된 응용이다. 이러한 그릇된 응용은 학위논문이나 대학의 학술지에 게재된 논문들에서 뿐만 아니라 엄격한 심사를 거쳤다고 보여지는 전문학술지에 게재된 논문들에서도 흔히 발생하고 있는 문제이다. 따라서 통계적 기법을 사용하는 연구자는 올바른 기법의 선택과 적용 여부가 바로 연구의 성패를 가름한다는 사실을 명심해야 하며, 통계적 오류를 피하는 동시에 통계적 기법을 효율적으로 이용할 수 있는 방안을 적극적으로 탐색해야 한다. 이러한 문제는 기법의 제공자인 통계학의 측면에서만 아니라 기법의 이용자인 해당 학문분야의 발전에 중요한 저해요인으로 볼 수 있으므로 문제의 범위와 그 심각성에 대한 올바른 평가는 매우 중요하다고 할 수 있다.

통계적 기법을 활용한 연구논문에 대한 타당성 문제는 오래 전부터 여러 학문분야에서 제기되고 있는 문제이다. White(1979)는 British Journal of Psychiatry에 게재된 168편의 논문을 분석한 바 있고, Schor &

Karten(1966)은 의학분야의 10개 주요 학술지를 대상으로 타당성 평가를 시도한 바 있다. 또한 Fleming(1981)과 Kukuk & Baty(1979)는 인자점수를 이용한 다중회귀분석의 문제를 다루었고, Dawkins(1983)는 다중비교(multiple comparison)의 문제를 그리고 Delucchi(1983)는 카이제곱검정의 문제를 다룬 바 있다. 국내에서도 김병수 등(1987), 최중후 & 이재창(1990), 성삼경 등(1993)의 연구에서 여러 분야의 논문들에 대한 타당성 평가를 시도한 바 있다. 이러한 논문들은 통계자료의 오용, 분석상의 오류, 그릇된 결론도출을 공통적으로 지적하고 있다.

간호학 분야에서도 기초통계분석에서 고급의 통계적 기법에 이르기까지 통계적 방법론이 폭넓게 이용되고 있는 것이 현실이며, 이에 따른 적용상의 문제점도 증가하고 있다. 이은옥(1999)은 간호학 실험연구에서 흔히 발견되는 문제점을 몇 가지 측면에서 지적한 바 있으며, 최경숙 등(2000)은 창간호(1970년)에서부터 1999년까지 대한간호학회지에 게재된 총 959편의 논문들을 분석함으로써 국내 간호 연구의 동향을 살펴본 바 있다. 또한 이은현·김진선(2000)은 사회적지지의 효과모델에 관한 국내간호논문들에 대해 통계분석방법을 분석한 바 있고, 임난영 등(2001)은 구조방정식모형을 사용한 국내 간호연구의 동향에 대해서 살펴본 바 있다. 이러한 비교연구의 연장에서 본 고에서는 간호학 전문학술지로서 가장 포괄적이고 오랜 전통을 가지고 있는 대한간호

\* 호서대학교 정보통계학과 전임강사

학회지에 게재된 논문들에 대해 최근 3년간(1999-2001)의 논문들을 중심으로 통계적 측면에서 포괄적으로 살펴보고자 한다.

## 2. 통계적 기법 사용의 현황 및 전반적 검토

1970년에 창간된 대한간호학회지는 2002년 현재 제 32권에 이르고 있으며 년 6회(2, 4, 6, 8, 10, 12월) 발간되고 있다. 창간호(제1권 1호)부터 1999년 12월(29권 6호)까지 총 959편의 논문이 게재되었고, 최경숙 등(2000)은 이들 논문들에 대해 여러 가지 측면에서 국내 간호연구의 동향을 분석한 바 있다.

최근 3년간(1999~2001)의 게재논문은 총 329편으로 이 중 233편(70.8%)의 논문이 기초통계분석 이상의 통계적 기법을 사용하고 있었으며(표 1), 약 10% 내외의 논문이 단순빈도나 평균 등을 제시하고 있었다(이 안에는 일부 질적 연구도 포함되어 있다). 본 고에서는 이들 논문 중 기초통계분석 이상의 통계적 기법을 사용한 233편의 논문에 대해 여러 가지 통계적 측면에서 살펴보고자 한다.

최종후·김항규(1994)는 통계적 방법론을 활용하고 있는 논문들에서 주로 나타나는 문제점으로 “연구설계과정”에 있어 (1) 연구의 대상에 대한 문제, (2) 연구설계의 문제, (3) 자료탐색의 문제, (4) 실험연구에 있어 확률화, 반복화, 블록화의 문제를, “통계적 분석”에 있어 (1) 연구과제의 통계적 형식화, (2) 적정성의 문제, (3) 소프트웨어 활용의 문제, (4) 유의수준의 적용문제를, 그리고 “결론의 기술”에 있어 (1) 통계적 유의성의 의미해석 문제, (2) 유의성의 결여에 대한 문제, (3) 유의성에 대한 지나친 집착의 문제, (4) 연구가설에 대한 통계적 결론의 서술 등을 들고 있다. 이러한 문제점은 최근 3년 간의 대한간호학회지에 게재된 많은 논문에서도 공통적으로 발견되고 있는 사항이다.

먼저 ‘연구설계과정’에 있어 연구의 대상인 목표모집단(target population)과 연구자가 취급한 표본모집단(sampled population)이 구분되어 서술되어 있지 않

거나(소위 대표성의 문제는 주로 목표모집단과 표본모집단의 사이의 불일치에서 발생한다) 표본추출방법, 측정, 표집된 자료의 기술, 대표성의 고려, 표본의 크기, 그룹 선택의 기준 등에 대한 언급이 소홀히 다루어지고 있다. 또한 자료에 대한 탐색 및 기초통계에 대한 서술은 통계적 기법들이 가지고 있는 이론적 가정의 검토뿐만 아니라 표본의 구조와 특징을 파악하기 위해서 중요하게 다루어져야 하는 문제이나 많은 논문에서 이 문제를 크게 고려하지 않고 있으며, 실험연구에 있어 실험계획에서 지켜야 할 세 가지 큰 원칙인 확률화(randomization), 반복화(replication), 블록화(blocking)가 무시되고 있는 경우가 많았다. 간호학 분야의 실험연구에 대한 확률화(무작위 배치) 문제는 이은옥(1999)에 의해서도 지적된 바 있으며, 특히 분산분석을 적용함에 있어 블록에 대한 고려가 거의 이루어지고 있지 않았다. 연구설계과정에 대한 세부적인 언급은 이후 후속연구의 방향설정에 큰 도움을 준다는 점에서도 이는 적극적으로 고려되어야 할 것이다.

‘통계적 분석’ 과정에 있어서는 연구과제의 통계적 형식화나 적정성의 문제가 나타나는 논문이 다소 있었다(이러한 사항에 대해서는 다음 절의 기법별 평가에서 자세히 다룬다). 통계적 기법을 활용하는 연구자가 통계학에 대한 이해의 부족으로 인한 통계적 오류를 줄이는 동시에 통계적 방법론을 보다 효율적으로 활용할 수 있도록 여러 연구자들에 의해 일종의 점검표가 개발되어 있는데(White, 1979; Karten, 1966, Tabachnick & Fidel, 1989; Ford & Tortora, 1978; 김병수 등, 1987; 성삼경 등, 1993; 최종후·김항규, 1994) 연구자는 이러한 점검표를 참조함으로써 통계적 오류를 줄이고 적합한 통계분석 과정을 수행하는데 큰 도움을 얻을 수 있을 것이다.

한편 소프트웨어는 활용 시 이를 활용하는 연구자의 자세나 소프트웨어의 성능도 문제가 될 수 있다. 서로 다른 소프트웨어가 상이한 결과를 출력할 수도 있고, 같은 소프트웨어에 의해서도 옵션 사용 등의 처리절차에 따라 다른 결과를 얻을 수도 있다. 소프트웨어는 통계적

〈표 1〉 최근 3년간 대한간호학회지 논문게재 현황

| 년도   | 단순빈도 정도만<br>제시한 조사연구 |       | 기초통계분석 이상의<br>정량적 연구 |       | 게제<br>논문 수 |
|------|----------------------|-------|----------------------|-------|------------|
| 1999 | 15                   | 12.4% | 83                   | 68.6% | 121        |
| 2000 | 11                   | 8.9%  | 89                   | 72.4% | 123        |
| 2001 | 9                    | 10.6% | 61                   | 71.8% | 85         |
| 계    | 35                   | 10.6% | 233                  | 70.8% | 329        |

〈표 2〉 통계적 기법 활용의 기법별 점검표의 예 (t-검정)

| 점 검 항 목 | 대 책  |
|---------|--|
| 나. 적정성  |  |
| (1) 가정  | <ul style="list-style-type: none"> <li>• 정규성               <ul style="list-style-type: none"> <li>• P-P PLOT 또는 ROOTOGRAM SHAPIRO-WILK 검정으로 검토 (정규성 가정 훼손시 비모수적 검정)</li> </ul> </li> <li>• 독립성</li> <li>• 등분산성               <ul style="list-style-type: none"> <li>• BOX-PLOT 또는 F-검정으로 검토 (이분산의 경우 Behrens-Fisher 검정)</li> </ul> </li> </ul> |
| (2) 변환  |  |
| (3) 기타  | <ul style="list-style-type: none"> <li>• 짝진 표본의 필요성</li> <li>• 짝진 t-검정</li> </ul>  |

최종후·김향규(1994)에서 인용하였음.

자료분석을 위한 하나의 도구이므로, 이를 활용하는 연구자는 먼저 관련 통계적 지식을 충분히 습득한 후에 각 소프트웨어의 기능을 정확히 파악하고 사용해야 한다. 또한 통계 소프트웨어의 전문화 및 다양화를 고려할 때 활용한 소프트웨어와 처리절차를 자세히 명시하는 것이 바람직하다.

〈표 3〉 자료분석에 사용된 소프트웨어

| 소프트웨어   | 논문 수 | 퍼센트   |
|---------|------|-------|
| SPSS    | 120  | 51.5% |
| SAS     | 69   | 29.6% |
| LISREL  | 10   | 4.3%  |
| QUANAL  | 8    | 3.4%  |
| PCQ     | 3    | 1.3%  |
| MD-PCS  | 1    | 0.4%  |
| 명시되지 않음 | 33   | 14.2% |

“결론의 기술”에 있어서는 특히 통계적 유의성에 대한 해석 및 기술에서 적절하지 못한 서술이 발견되었다. 유

의수준(significant level,  $\alpha$ )의 선택은 가설에 대한 현실적인 경중과 관련이 있으므로 완전히 객관적일 수는 없으나, 연구자가 특정한 유의수준에 기초하여 결론을 도출하는 경우 이를 명시하는 것이 바람직하다(최근 3년 간의 대한간호학회지에서는 10편 정도에서만 가설검정에 대한 연구자의 유의수준을 명시하고 있다). 또한 유의확률(significant probability, p-값)은 귀무가설에 반하는 경험적 증거의 강약을 나타내므로, 연구결과에 이용자를 위해 가능한 이를 연구결과에 포함시키는 것이 좋다. 한편 통계적 유의성은 실제적 유의성과 더불어 검토될 문제이므로 통계적 유의성만의 지나친 집착과 확대해석은 옳지 않으며, 유의성이 검출되지 않은 연구결과도 때로는 그 자체로 훌륭한 의미를 지닐 수 있다는 점과 추후의 불필요한 연구를 막을 수 있다는 점에서 유의성 검출의 실패와 같은 연구결과도 발표되어야 할 것이다(유의성을 갖는 결과를 얻기 위해 여러 가지 검증을 동시에 적용하여 유의한 검증결과만을 보고하는 것도 바람직하지 않다).

〈표 4〉 한국통계학회 및 대한간호학회지의 용어 (t-검정 및 카이제곱검정)

| 한국통계학회                                   | 대한간호학회지  |
|--|--|
| 독립표본 t 검정<br>(independent sample t-test) | t 검정[검증], t-test<br>unpaired t-test  |
| 이표본 t 검정 (two sample t-test)             | student t-test   |
| t 검정[검증] (t-test)                        | two sample t-test<br>compared Independent sample t-test  |
| 쌍체 t 검정 (paired t-test)                  | paired t-test  |
| 쌍체표본 t 검정 (paired sample t-test)         | 짝비교 t-test[검정]   |
| 대응표본 t 검정, 대응 t 검정                       | 짝짓기 t-test   |
| 짝(지은) t 검정                               |  |
| 카이제곱검정 (Chi-square test)                 | 카이제곱검정, Chi-square test<br>X <sup>2</sup> -test[검정], X <sup>2</sup> -test<br>Chisquare test<br>이변량분석 |

용어의 부적절한 사용은 학문의 토착화와 학문간 교류의 측면에서 볼 때 바람직하지 않으며 연구결과를 잘 못 이해하게 하는 원인이 되기도 한다. 따라서 연구자는 단순히 선행 연구자의 잘못된 통계용어를 답습하는 자세를 지양하고 적절한 통계용어의 사용에 대한 노력을 기울여야 한다. <표 4>는 한국통계학회의 통계용어사전(1991), 통계학용어집(1997), 통계학용어대조표(<http://www.kss.or.kr> 참조)의 통계용어와 대한간호학회지에서 나타난 용어를 t-검정 및 카이제곱검정에 대해 대비한 것이다.

### 3. 세부 기법별 논의

<표 5>에서 볼 수 있듯이 t-검정은 가장 많이 사용되는 통계적 분석기법 중 하나이며(50% 이상의 논문이 t-검정을 사용하고 있다), 독립표본 t-검정은 주로 실험군과 대조군의 사전 동질성 비교나 평균차이 검정을 위해 그리고 쌍체표본 t-검정은 실험 전후의 평균차이 검정을 위해 사용되고 있었다. 그러나 두 방법 중 어느 방법을 사용했는지가 명시되지 않은 경우가 소수 있었으며, 쌍체표본(대응표본) t-검정을 사용해야 하는 상황에서 독립표본 t-검정을 사용한 것으로 보이는 경우도 있었다. 또한 분산분석을 수행해야 하는 문제를 t-검정을 반복적으로 사용한 경우도 있었다. 소수 몇 편의 논문에서는 평균 및 표준편차 등이 제시되지 않은 경우가 있었는데, 불가피한 경우(지면의 제한 등)가 아니라면 후속 연구를 위해 평균 및 표준편차를 모두 제시하는 것이 바람직하다.

한편 표본크기가 작은 경우 정규성 및 등분산성 등을 검토한 후 비모수적 검정을 실시할 필요가 있는데, 소수 몇 편의 논문에서 가정에 대한 아무런 검토 없이 t-검정을 사용하고 있다. 또한 t-검정은 표본크기가 큰 경우 가정에 대한 다소의 이탈에는 민감하지 않은 로버스트성(robustness)을 가지고 있는데, 표본크기가 충분히 큰 경우에도 아무런 검토 없이 부호순위검정이나 순위합검정을 수행한 논문이 다소 있었으며(비모수적 검정을 사용한 19편의 논문 중 표본크기가 50 이하인 경우가 14편, 51~100인 경우가 2편, 101~200인 경우가 2편, 301~400인 경우가 1편 이었다), 이는 검정의 검정력(power)을 떨어뜨리는 원인이 될 수 있으므로 바람직하지 않다.

카이제곱검정(Chi-square test)은 t-검정과 더불어 주로 여러 연구집단의 사전 동질성을 검토하기 위해 사용되고 있었다. 카이제곱검정을 수행함에 있어서 고려해야 할 사항 중 하나는 분할표(contingency table) 상에서 각 칸(cell)의 기대도수(expected frequency)가 충분히 큰가(대략 5이상)의 문제이다. 각 칸의 기대도수가 너무 작은 경우에도 도수에 대한 특별한 검토 없이 카이제곱검정에 기초하여 추론을 실시하는 경우가 있었으며, 기대도수가 5 이하인 경우 해당 변수를 분석에서 아예 제외한 경우도 있었다. 분할표 상에서 각 칸의 기대도수가 너무 작은 경우에는 주변의 칸과 통합하여 기대도수를 크게 하거나 Fisher의 정확검정(Fisher's exact test)을 사용하는 것도 한 방법이 될 수 있다(대략 5편 정도의 논문에서 Fisher의 정확검정을 사용하였다).

상관분석은 주로 피어슨의 상관계수(Pearson's correlation coefficient)에 기초하여 이루어지고 있었

<표 5> 주요 통계기법의 활용 현황

| 기법      | 세부기법      | 1999년     | 2000년     | 2001년     | 합계         |
|---------|-----------|-----------|-----------|-----------|------------|
| t-검정    | 쌍체표본 t-검정 | 11(13.3%) | 10(11.2%) | 6( 9.8%)  | 27(11.6%)  |
|         | 독립표본 t-검정 | 43(51.8%) | 44(49.4%) | 31(50.8%) | 118(50.6%) |
| 비모수적 검정 |           | 6( 7.2%)  | 9(10.1%)  | 4( 6.6%)  | 19( 8.2%)  |
| 카이제곱검정  |           | 26(31.3%) | 30(33.7%) | 25(41.0%) | 81(34.8%)  |
| 상관분석    |           | 27(32.1%) | 28(31.1%) | 22(36.1%) | 77(32.8%)  |
| 회귀분석    | 단순 회귀분석   | 2( 2.4%)  |           |           | 2( 0.9%)   |
|         | 다중 회귀분석   | 8( 9.6%)  | 11(12.4%) | 12(19.7%) | 31(13.3%)  |
|         | 일원 분산분석   | 35(42.2%) | 31(34.8%) | 20(32.8%) | 86(36.9%)  |
| 분산분석    | 공분산분석     | 4( 4.8%)  | 9(10.1%)  | 1( 1.6%)  | 14( 6.0%)  |
|         | 반복측정 분산분석 | 10(12.0%) | 9(10.1%)  | 9(14.8%)  | 28(12.0%)  |
|         | 다변량 분산분석  |           | 2( 2.2%)  |           | 2( 0.9%)   |
| 인자분석    |           | 10(12.0%) | 10(11.2%) | 6( 9.8%)  | 26(11.2%)  |
| 신뢰도 분석  |           | 38(45.8%) | 32(36.0%) | 31(50.8%) | 101(43.3%) |
| 구조방정식모형 |           | 4( 4.8%)  | 4( 4.5%)  | 5( 8.2%)  | 13( 5.6%)  |

다(77편의 논문 중 4편의 논문에서는 스피어만의 순위 상관계수(Spearman's rank correlation coefficient)를 사용하였다). 피어슨의 상관계수는 선형성(linearity)의 강도를 재는 척도이므로 두 변수의 관계가 곡선 또는 비선형적 형태를 가지고 있을 경우 이를 제대로 검출하지 못할 수 있다. 따라서 산점도(scatter plot) 등을 이용하여 자료의 형태를 탐색적으로 살펴보는 것도 필요한데, 대부분의 논문에서 이에 대한 언급이 다루어지고 있지 않다.

선형회귀분석(linear regression analysis)에서 요구되는 가정은 선형성(linearity), 정규성(normality), 독립성(independency), 등분산성(equal variance) 등이며, 이는 산점도 및 잔차그림(residual plot) 등을 통한 회귀진단으로 검토될 수 있는데 회귀분석을 사용한 대부분의 논문에서 가정에 대한 검토나 다중공선성 등 회귀진단 문제가 언급되지 않고 있었다. 회귀분석에서 발견되는 문제점 중 하나는 모형 설명력의 척도로 흔히 사용되는  $R^2$ (R-square)이 매우 작은 경우 이에 대한 특별한 언급이 없다는 점이다. F-검정에 의한 통계적 유의성이 인정된다 하더라도  $R^2$ 이 매우 작으면 실제적 유의성은 의심되므로 이를 언급하는 것이 바람직하다. 또한 다중회귀분석에서 각 회귀계수의 의미는 다른 모든 독립변수(설명변수)가 같도록 고정된(통제된) 경우 해당 독립변수가 1단위 변할 때 종속변수(반응변수)가 변하는 양이므로, 다중회귀분석에서의 회귀계수에 대한 해석은 단순회귀분석 또는 단순상관분석에서와는 구별되어 서술되어야 한다.

한편 전진 선택법(forward selection), 후진 제거법(backward elimination), 단계적 방법(stepwise method) 등의 변수선택방법을 사용하는 경우(대부분의 논문에서 단계적 방법을 사용하고 있음), 변수선택이나 제거시의 기준을 어떻게 설정하였느냐에 따라 결과가 달라질 수 있으므로 변수선택시의 기준이나 유의수준 등도 명시하는 것이 바람직하다(거의 대부분의 논문에서 이에 대한 언급이 없었다). 또한 변수선택방법을 사용한 대부분의 논문에서 회귀분석의 결과를 제시함에 있어 '변수 선택 요약'을 제시하고 있는데, 위계적(hierarchical) 회귀분석을 수행하고자 하는 것이 아니라면 회귀분석의 최종 결과를 제시하고 이에 기초하여 해석을 하는 것이 바람직한 것이다.

분산분석(analysis of variance)에서도 t-검정과 마찬가지로 정규성, 독립성, 등분산성 등이 요구된다. 특히 표본크기가 작은 경우 정규성 및 등분산성의 가정에서 크

게 이탈하였을 때는 비모수적 검정(예를 들면, Kruskal-Wallis 검정이나 Friedman 검정)을 사용하는 것이 바람직하다. 분산분석은 흔히 실험연구(experimental research)에서 많이 수행되는데, 적절하고 효율적인 실험은 연구의 성패를 결정짓는 중요한 요인이 된다.

일반적으로 확률화(randomization), 반복화(replication), 블록화(blocking)를 실험계획에서 지켜야 할 세 가지 큰 원칙으로 들 수 있다. 그러나 많은 실험연구 논문에서 실험설계상의 확률화 등에 대한 언급이 다루어지고 있지 않았다. 또한 블록화는 실험단위들이 서로 이질적인 경우 교락(confounding)되어 있는 요인들을 통제하여 실험의 정도(precision)를 향상시키고자 하는 원리이며, 동질적인 실험단위들을 하나의 블록으로 묶어 실험을 수행한 후 이를 고려하여 분석을 수행함으로써 실험 및 분석의 정도를 높일 수 있게 된다. 관측연구(observational study)의 경우에도 추정의 편(bias)을 줄이기 위해 블록(또는 공변량)을 고려하는 것이 바람직하다.

또한 다원 분산분석에서 교호작용(상호작용, interaction)이 유의적이면 주효과(main effect)에 대한 해석에 주의를 기울여야 한다. 분산분석을 사용한 거의 모든 논문에서 블록(block)이나 다른 요인(factor)을 고려한 다원 분산분석을 고려하지 않고 있었다.

공분산분석(analysis of covariance)은 공변량(covariate)을 일종의 블록으로 간주하여 실험 및 분석의 정도(precision)를 향상시키고자 하는 방법이다. 공분산분석에서는 공변량에 대한 기술기가 각 집단에서 동일하다는 가정이 필요한데(만약 기술기의 동일성이 인정되지 않는다면 일반선형모형으로 분석해야 한다), 거의 모든 논문에서 이에 대한 언급이 없었다.

반복측정(repeated measurement)에 의한 실험에서는 동일한 실험단위로부터 얻은 반응값들 간에 서로 상관이 생기게 되므로 이러한 상관성을 고려하여 분석이 수행되어야 하는데, 다변량(multivariate) 분산분석이나 구형성(또는 복합대칭성)을 고려한 일변량적 분석이 사용될 수 있다. 대부분의 논문에서 구형성에 대한 언급이 없이 일변량분석을 사용하고 있었으며, 다변량 또는 일변량 분산분석에서 각 F-검정의 의미를 잘 못 해석하고 있는 경우도 있었다.

다중비교(multiple comparison)는 모두 33편의 논문에서 수행되었는데(Duncan: 12, Scheffe: 10, Bonferroni: 7, Tukey: 3 등), 수행된 다중비교방법을 명시하지 않은 경우도 있었으며 다중비교를 실시해야 하

는 문제에서 t-검정을 반복적으로 사용하는 오류도 있었다. 특히 반복측정 분산분석에서는 소수의 논문에서만 다중비교가 실시되었으며(3편 정도의 논문에서 Bonferroni 방법을 사용함), 쌍체표본 t-검정을 여러 번 수행하는 경우도 있었다.

인자분석(factor analysis)은 가장 널리 사용되는 다변량 자료분석 기법 중 하나로써, 특히 측정도구의 타당성 검토와 관련하여 주로 사용되고 있었으며 Q-인자 분석을 수행한 논문도 11편이었다. 인자분석을 사용한 거의 모든 논문에서 공분산행렬과 상관계수행렬 중 어느 것에 기초하고 있는지에 대한 언급이 없었으며, 주성분 분석과 인자분석을 혼용(또는 혼동)하여 서술하는 경우도 있었다. 또한 인자회전을 수행하여 그 결과를 제시하는 경우가 대부분이었는데, 많은 논문에서 상관계수행렬(또는 공분산행렬)의 고유값(eigenvalue)과 인자의 설명분산(explained variance)을 혼동하여 서술하고 있었다.

한편 신뢰도(reliability)는 구조화된 측정도구를 사용하고 있는 대부분의 논문에서 제시되고 있었으며, 주로 사용된 신뢰도는 크로바흐의 알파(Cronbach's alpha)이었다.

구조방정식모형(structural equation model)은 특히 사회과학의 여러 분야에서 사용빈도가 증가하고 있다. 특별한 경우가 아니라면 구조방정식모형은 공분산행렬에 기초하여 분석을 수행하는 것이 바람직한데, 대부분의 논문에서 공분산행렬과 상관계수행렬 중 어느 것에 기초하고 있는지에 대한 언급이 없었다. 또한 많은 논문에서 상관계수 및 표준편차를 제시하지 않고 있는데, 이는 구조방정식모형의 기초자료이므로 후속연구를 위해 제시하는 것이 바람직하다. 한편 경로분석(4편 정도)에 있어서는 회귀분석을 여러 번 사용하여 그 결과를 경로 분석의 결과로 제시한 경우도 있었으며, 어떤 소프트웨어나 분석과정을 거쳐 분석결과를 얻었는지에 대한 언급이 없는 경우도 있었다.

기타 분석방법으로는 정준판별분석(3편), 로지스틱 회귀분석(3편), 다차원척도(1편), 시계열분석(1편), 메타분석(4편) 등이 사용되고 있었다. 특히 정준판별분석의 경우 공분산행렬의 집단 간 동일성 등 타당성에 대한 검토가 이루어지지 않았고, 판별함수의 유의성이나 판별변수의 공헌도 등에 대한 언급이 부족하였다.

#### 4. 결론 및 토의

본 고에서는 최근 3년간(1999-2001) 대한간호학회지에 게재된 논문들을 중심으로 통계적 자료분석 측면을 포괄적으로 살펴보았다. 본 고를 작성함에 있어 기존의 선행연구 및 주요 통계학 서적을 참조하여 객관적 서술이 되도록 노력하였으나, 이와 관련된 방대한 지식 및 시간의 부족과 무엇보다 간호학에 대한 무지로 인해 적절하지 못한 검토가 이루어진 측면도 있을 것이라 생각된다. 따라서 연구자는 다른 많은 연구논문 및 통계학 서적 등을 통해 활용하고자 하는 통계적 방법론에 대해서 충분히 검토하고 숙지할 필요가 있을 것이다.

한편 최중후·김항규(1994)는 통계적 기법 활용을 개선하기 위해 다음과 같은 방안을 제시하고 있다: (1) 통계학 교육의 확충, (2) 통계상담실(통계전문가)의 활용, (3) 레프리(referee) 제도의 활용, (4) 외국 학술지의 비판적 수용, (5) 통계적 기법을 활용하는 연구자의 올바른 자세. 이러한 개선 방안은 간호학뿐만 아니라 통계적 기법을 활용하는 모든 학문분야에도 적용될 수 있으리라 생각된다. 과학적 탐구에 있어서 무엇보다도 중요한 것은 올바른 연구문제의 정립이며, 그릇된 문제정립은 통계적 기법의 활용을 오도할 수 있으므로 특히 사회과학 분야의 연구에서 중시되어야 할 문제이다.

## References

- 김병수, 안윤기, 윤기중 (1987). 통계의 오류와 효율적 이용에 관한 연구, 산업과 경영, 24, 3-37.
- 성삼경, 최중후, 이재창 (1993). 경영학 연구논문의 통계적 오류, 경영학연구, 22, 163-198.
- 이은옥 (1999). 간호학 실험연구에서 흔히 발견되는 문제점, 대한간호학회지, 29, 1469-1473.
- 이은현, 김진선 (2000). 사회적지지 효과 모델 및 통계 분석방법에 관한 국내간호논문 분석, 대한간호학회지, 30, 1503-1520.
- 임난영, 강현숙, 이성은, 서연옥, 권영은 (2001). 공변량 구조분석을 이용한 국내 간호연구의 동향, 대한간호학회지, 31, 221-231.
- 최경숙, 송미순, 황애란, 김경희, 정명실, 신성례, 김남초 (2000). 대한간호학회지 논문 분석을 통한 국내 간호연구의 동향 - 창간호에서 1999년도까지 -, 대한간호학회지, 30, 1207-1217.
- 최중후, 김항규 (1994). 과학적 방법론과 통계기법의 활용, 자유아카데미.
- Dawkins, H. C. (1983). Multiple Comparisons

- Misused : Why so Frequently in Response-Curve Studies?, *Biometrics*, 39, 789-790.
- Delucchi, K. L. (1983). The Use and Misuse of Chi-Square : Lewis Burke Revisited, *Psychological Bulletin*, 94, 166-176.
- Fleming, J. S. (1981). The Use and Misuse of Factor Scores in Multiple Regression Analysis, *Educational and Psychological Measurement*, 41, 1107-1125.
- Ford, B. L. and Tortora, R. D. (1978). The Consultant's Forum - A Consulting Aid to Sample Design, *Biometrics*, 34, 299-304.
- Kukuk, C. R. and Baty, C. F. (1979). The Use of Multiple Regression With Composite Scale Obtained from Factor Scores, *Educational and Psychological Measurement*, 39, 277-290.
- Schor, G. W. and Karten, I. (1966). Statistical Evaluation of Medical Journal Manuscripts, *Journal of the American Medical Association*, 195, 1123-1128.
- Tabachnick, B. G. and Fidell, L. S. (1989). *Using Multivariate Statistics, 2nd Ed*, New York : Harper & Row, Publishers, Inc.
- White, S. J. (1979). Statistical Errors in papers in the British Journal of Psychiatry, *British Journal of Psychiatry*, 135, 336-342.