

최근 10년간(2000-2009) 사상체질의학회지 게재논문의 통계분석방법에 대한 연구

박기현 · 김상혁 · 김홍기* · 이시우

한국한의학연구원
*충남대학교 통계학과

Abstract

An Evaluation of the Statistical Techniques Used in Papers of the Journal of Sasang Constitutional Medicine (2000-2009)

Ki Hyun Park · Sang-Hyuk Kim · Hong Gie Kim* · Si Woo Lee

Korea Institute of Oriental Medicine

*Chungnam National University. Statistics Dept.

1. Objectives

The purpose of this study is to assess the statistical methods used in the Journal of Sasang Constitutional Medicine and to identify the types of errors in the statistical analysis.

2. Methods

We have reviewed quantitative articles published in the Journal of Sasang Constitutional Medicine from 2000 to 2009. Except for the case report, a total of 505 articles was reviewed by year. We focused on the methods of statistics in these articles. We have analysed the number of thesis used the statistical methods, the number of the statistical methods, the types of the statistical methods, and the errors of the statistical methods with the object of 202 thesis used the statistical method.

3. Results

1. Of the 370 articles, 202 used inferential statistics.
2. The most dominant was the thesis with one type of the statistical method forming 61.9% (125 thesis) of the total.
3. The ANOVA was the statistic of choice overall (40.1%), followed by the t-test (35.6%).
4. By examining the errors of the statistical methods, there were 49 thesis with errors among 81 thesis using ANOVA, 25 thesis among 72 thesis using t-test, and 18 over 62 thesis using chi-squared test.

4. Conclusions

We have found various mistakes or misuses in the applications of statistical methodologies in the articles published in the Journal of Sasang Constitutional Medicine. It is necessary to present the cases of inappropriate statistical methods in order to improve the quality of academic researches. So long as Sasang constitutional medicine uses the statistical methods, researchers should fully examine and be aware of the statistical methods which they wish to use through other research articles, and statistic publications.

Key Words : Application of statistical method, ANOVA, t-test, chi-squared test

• 접수일 2010년 11월 01일; 심사일 2010년 11월 02일;
승인일 2010년 11월 12일
• 교신저자 : 이시우
대전시 유성구 전민동 461-24
한국한의학연구원 체질의학임상연구센터

Tel : +82-42-868-0555 Fax : +82-42-868-9480
E-mail : bfree@kiom.re.kr

* 이 논문은 2010년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 20100020618).

I. 緒 論

연구의 통계분석이란 관찰이나 실험 또는 조사를 통하여 얻어진 결과들이 기존에 알려진 사실들과 비교하여 어떠한 차이를 보이는지 객관적으로 증명하기 위한 과정이다. 이러한 통계분석은 많은 분야에서 연구 결과를 살펴볼 때 사용하고 있으며, 이때 쓰이는 기법들의 종류도 다양해지고 내용 또한 그 수준이 높아지고 있는 추세이다. 그러나 연구자들의 통계에 대해 충분한 교육은 이루어지지 못하고 있는 실정이며, 이로 인해 부적절한 통계기법의 적용, 결과해석에서의 오류를 범하기 쉽다. 예를 들어 김태일(1997)은 1992년에서 1996년까지의 행정 분야 주요학회지에 게재된 논문에 대한 통계기법 오류를 살펴보았으며¹, 이경미 등(2003)은 1998년에서 2002년 8월까지 발간된 디자인학연구에 게재된 논문에 대해 각 연구 단계에 따른 통계기법 타당성 검정을 실시하였다.²

의학 분야에서도 기초통계분석에서 고급의 통계적 기법에 이르기까지 통계적 방법론이 폭넓게 이용되고 있으며, 이에 따른 적용상의 문제점도 증가하고 있다. 이형기 등(1991)은 1980년부터 1989년까지 대한의학 협회지에 게재된 총 297편의 논문들을 분석함으로써 의학 분야에서 흔히 발견되는 문제점을 몇 가지 측면에서 지적한 바 있다³. 한의학에서도 이정열 등(1998)은 한의학 연구에서 사용된 통계기법의 오류에 대해 살펴본 바 있다⁴. 본고에서는 이러한 연구의 연장으로 사상체질의학 전문학술지로서 가장 포괄적이고 오랜 전통을 가지고 있는 사상체질의학회지에 게재된 최근 10년간(1989-2010)의 논문들을 중심으로 통계적 측면에서 포괄적으로 살펴보고자 한다.

II. 研究對象 및 方法

1. 연구대상

1989년에 창간된 사상체질의학회지는 연 4회(3, 6, 9, 12월-2010년 기준) 발간되고 있으며, 창간호(제1권 1호)부터 2009년 12월(21권 3호)까지 총 813편의 논문이 게재되었다. 본고에서는 최근 10년간(2000~2009) 게재논문 총 505편 중 증례를 제외한 논문 370편을 분석대상으로 하였다(Table 1).

2. 연구방법

연구대상 370편의 논문 중 통계기법을 사용하지 않았거나 단순 빈도분석만을 사용한 논문은 제외하였으며, 한 개 이상의 통계기법을 사용한 논문을 대상으로 사용된 통계기법에 대하여 전반적으로 살펴보고, 가장 많이 사용된 통계기법을 중심으로 오류를 살펴 보았다.

III. 結 果

1. 통계기법을 적용한 논문 수

전체 370편의 논문을 연도별로 살펴보면 2000년에 35편, 2001년에 38편, 2002년에 32편, 2003년에 36편, 2004년에 32편, 2005년에 31편, 2006년에 40편, 2007년에 47편, 2008년에 34편, 2009년에 45편이었으며, 그중 통계기법을 사용한 논문은 202편(54.6%)으로 2000년에 19편(54.3%), 2001년에 21편(55.3%), 2002년에 15편(46.9%), 2003년에 17편(47.2%), 2004년에 18편(56.3%), 2005년에 14편(45.2%), 2006년에 25편

Table 1. Number of Articles

(unit : number)

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	Total
Case Report	6	12	18	14	17	19	7	15	14	13	135
	14.6%	24.0%	36.0%	28.0%	34.7%	38.0%	14.9%	24.2%	29.2%	22.4%	26.7%
Original Article	35	38	32	36	32	31	40	47	34	45	370
	85.4%	76.0%	64.0%	72.0%	65.3%	62.0%	85.1%	75.8%	70.8%	77.6%	73.3%
Total number of articles	41	50	50	50	49	50	47	62	48	58	505

Table 2. Number of Articles Using Statistics (unit : number)

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	Total
Number of articles not using statistics	16	17	17	19	14	17	15	15	14	24	168
	45.7%	44.7%	53.1%	52.8%	43.7%	54.8%	37.5%	31.9%	41.2%	53.3%	45.4%
Number of articles using statistics	19	21	15	17	18	14	25	32	20	21	202
	54.3%	55.3%	46.9%	47.2%	56.3%	45.2%	62.5%	68.1%	58.8%	46.7%	54.6%
Total number of articles	35	38	32	36	32	31	40	47	34	45	370

Table 3. Number of Statistical Techniques Applied (unit : number, (%))

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	Total
1	13	14	11	8	8	10	18	19	14	10	125(61.9)
2	5	5	2	5	9	2	4	10	3	7	52(25.7)
3	1	2	2	1	1	2	2	2	2	1	16(7.9)
4				3				1	1	3	8(4.0)
5							1				1(0.5)
Total	19	21	15	17	18	14	25	32	20	21	202

(62.5%), 2007년에 32편(68.1%), 2008년에 20편(58.8%), 2009년에 21편(46.7%)이었다(Table 2).

적용된 통계기법 수를 살펴보면 총 202편 중 1가지 방법만 사용한 경우가 125편(61.9%)으로 가장 많았고, 2가지 52편(25.7%), 3가지 16편(7.9%), 4가지 8편(4.0%), 5가지 1편(0.5%)이었다(Table 3).

2. 통계적용 기법

통계기법을 적용한 논문 총 202편을 모수적인 통계기법과 비모수적 통계기법으로 나눠 보면, 모수적인 통계기법만 적용한 논문이 114편(56.4%)으로 가장 많았고, 비모수적 통계기법만 적용한 논문은 42편(20.8%)이었으며, 모수와 비모수적 방법을 함께 사용한 논문은 46편(22.8%)이었다(Table 4).

적용된 통계기법수를 살펴보면 총 202편의 논문에

서 285가지의 통계기법이 사용되었다. 세부별로 살펴보면 분산분석이 81번으로 가장 많이 사용되었고, T검정이 72번, 카이제곱검정이 62번, 신뢰도분석이 10번, 회귀분석이 9번, 판별분석이 9번, 상관분석이 6번으로 많이 사용되었으며, 카이제곱검정을 제외한 비모수검정은 31번 사용되었다. 연도별로 적용된 통계기법을 살펴보면 2000년에 19편의 논문에서 25가지의 통계기법이 사용되었고, 2001년에 21편의 논문에서 30가지, 2002년에 15편의 논문에서 21가지, 2003년에 17편의 논문에서 32가지, 2004년에 18편에서 27가지, 2005년에 14편에서 20가지, 2006년에 25편에서 32가지, 2007년에 32편에서 46가지, 2008년에 20편에서 29가지, 2009년에 21편에서 39가지의 통계기법이 사용되었다. 편당 사용된 통계기법수를 살펴보면 2003년과 2009년에 각각 1.88가지와 1.86가지로 가장 많았

Table 4. Number of Parametric and Non-parametric Methods Used (unit : number, (%))

Method	Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	Total
Parametric		14	11	7	7	11	11	16	15	15	7	114(56.4)
Non-parametric		3	7	4	3	3	1	5	10	1	5	42(20.8)
Both		2	3	4	7	4	2	4	7	4	9	46(22.8)
Total		19	21	15	17	18	14	25	32	20	21	202

Table 5. Number of Statistical Techniques Used According to Data Form

Techniques	Year											Total
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009		
ANOVA	6	5	7	7	5	5	9	14	12	11	81	
t-test	10	10	5	7	8	9	3	6	6	8	72	
Chi-square test	4	6	6	8	5	3	5	11	3	11	62	
Non-parametric	1	5	1	2	3	2	4	8	3	2	31	
Correlation analysis	0	1	0	1	2	0	1	0	0	1	6	
Regression analysis	0	0	0	1	1	1	1	0	2	3	9	
Reliability analysis	0	1	1	3	1	0	1	1	1	1	10	
Discriminant analysis	0	1	0	2	0	0	2	3	1	0	9	
The others	4	1	1	1	2	0	6	3	1	2	21	
Total	25	30	21	32	27	20	32	46	29	39	285	
Total number of articles	19	21	15	17	18	14	25	32	20	21	202	

Table 6. Check List

Statistical technique	Check list
ANOVA	1. Validity
	2. Assumption
	3. Multiple comparison
	4. Conclusion
t-test	1. Validity
	2. Assumption
	3. Conclusion
Chi-square test	1. Validity
	2. Sample size
	3. Conclusion

고, 2000년과 2006년에 각각 1.32가지와 1.28가지로 가장 적었으며, 전체로 살펴보면 평균 1.41가지의 통계기법이 사용되었다(Table 5).

3. 적용된 통계기법 오류

최중후 · 김항규(1994)⁵가 제시한 기준을 바탕으로 만들어진 점검표(Table 6)로 가장 많이 사용된 통계기법인 분산분석, T검정, 카이제곱 검정의 타당성에 대해 살펴보았다.

1) 분산분석

분산분석을 사용한 논문의 81편중에서 오류가 발견된 논문은 49편(60.5%)이었다. 연도별로 살펴보면 2000년 3편(50.0%), 2001년 5편(100%), 2002년 3편(42.9%), 2003년 3편(42.9%), 2004년 4편(80.0%), 2005

년 2편(40.0%), 2006년 6편(66.7%), 2007년 8편(57.1%), 2008년 8편(66.7%), 2009년 7편(63.6%)으로 매년 적지 않은 오류가 발견되었다.

발견된 오류로는 표본의 크기가 매우 작음에도 불구하고 정규성 및 등분산성의 가정에 대한 아무런 언급이 없이 분산분석을 시행한 경우가 38편(46.9%)으로 가장 많았고, 분산분석 시행 후 유의한 결과에 대한 사후검정을 실행하지 않은 경우가 28편(34.6%), 시작부터 끝까지 유의수준을 언급하지 않았으나 유의하다는 결론을 낸 경우가 6편(7.4%)이었다. 연도별로 살펴보면 2000년 2편, 2001년 4편, 2002년 3편, 2003년 2편, 2004년 3편, 2005년 2편, 2006년 3편, 2007년 6편, 2008년 8편, 2009년 3편의 논문에서 표본 크기가 작음에도 아무런 가정 언급 없이 분산분석을 사용하였으며, 사후검정을 실행하지 않은 논문은 2000년 1편, 2001년

2편, 2003년 2편, 2004년 1편, 2006년 3편, 2007년 6편, 2008년 6편, 2009년 7편이 발견되었다. 유의수준을 제시하지 않은 논문은 2000년 1편, 2001년 1편, 2002년 1편, 2006년 2편, 2007년 1편이었다. 그 밖에도 분산분석 후 사후검정에서 다중비교를 T검정의 반복으로 대체 실행한 경우도 있었다.

2) T검정

T검정을 사용한 논문은 72편이었으며, 그 중에서 오류가 발견된 논문은 25편(34.7%)이었다. 연도별로 살펴보면 2000년 2편(20.0%), 2001년 2편(20.0%), 2002년 1편(20.0%), 2003년 5편(71.4%), 2004년 5편(62.5%), 2005년 5편(55.6%), 2007년 3편(50.0%), 2009년 2편(25.0%)으로 2006년과 2008년을 제외한 논문에서 꾸준히 발견되었다.

발견된 오류를 살펴보면 분산분석에서 나타난 오류와 같이 표본크기가 매우 작음에도 연구에서 정규성, 등분산성에 대한 언급 없이 T검정을 사용한 경우가 24편(37.5%)으로 가장 많았다. 연도별로 살펴보면 2000년 2편, 2001년 2편, 2002년 1편, 2003년 5편, 2004년 5편, 2005년 5편, 2007년 3편, 2009년 1편이었다. 그 밖에 나타난 오류로는 대응표본 T검정을 독립표본 T검정으로 분석한 경우가 있고, 오류라고 할 수는 없겠으나 T검정을 분산분석으로 분석한 경우가 있었다. 그리고 T검정을 실행한다는 언급만하고 다른 분석만 실행한 경우도 있었다.

3) 카이제곱검정

카이제곱검정을 사용한 논문은 62편이었으며 그

중에서 오류가 발견된 논문은 18편(29.0%)이었다. 연도별로 살펴보면 2001년 3편(50.0%), 2002년 2편(33.3%), 2003년 3편(42.9%), 2006년 3편(60.0%), 2007년 2편(20.0%), 2009년 5편(45.5%)으로 2000년, 2004년, 2005년, 2008년을 제외한 논문에서 오류가 발견되었다.

카이제곱검정에서 발견된 오류의 대부분은 각 칸의 기대도수가 너무 작은 경우에도 도수에 대한 특별한 검토 없이 카이제곱검정을 실행한 경우로 모두 17편(27.4%)이었다. 연도별로 살펴보면 2001년 3편, 2002년 2편, 2003년 3편, 2006년 3편, 2007년 2편, 2009년 4편에서 도수에 대한 검토 없이 카이제곱검정을 시행했다. 그밖에 오류로는 결론에 대한 언급에서 유의확률을 제시하지 않은 경우와 카이제곱검정을 실행한다는 언급만하고 다른 분석 결과만 제시한 경우가 있었다.

이상의 내용을 그래프로 나타내면 Figure. 1.과 같다.

IV. 考察 및 結論

연구자들은 여러 분야의 연구에서 통계기법을 사용하고 있으며, 그 쓰임새 또한 다양해지고 있다. 하지만 적절한 통계기법을 사용하지 못한다면 연구결과에 대한 신뢰성을 떨어뜨리고 잘못 분석된 연구 결과를 토대로 임상 진료가 이루어진다면 더 큰 문제가 생길 것이다. 그래서 본고에서는 기존의 선행연구 및 주요 통계학 서적을 참조하여 최근 10년간 (2000-2009) 사상체질의학회지에 게재된 논문에서 증례를 제외한

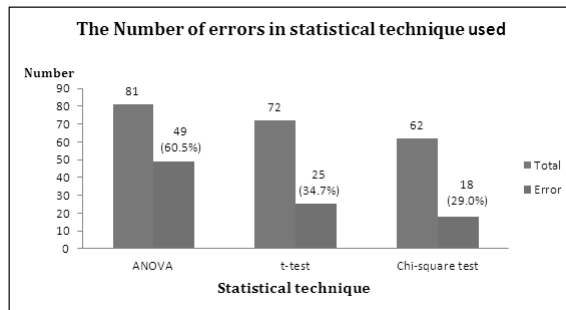


Figure. 1. The number of errors in statistical technique used

370편의 논문 중 통계기법이 적용된 202편의 논문을 대상으로, 통계기법을 적용한 논문편수, 적용된 통계기법 수, 적용된 통계기법의 종류, 그리고 통계기법의 오류 등을 분석하였다. 그 내용을 정리하면 다음과 같다.

통계기법을 적용한 논문 편수는 분석대상으로 삼은 370편의 논문 중에서 202편(54.6%)에서 통계기법을 사용하였고, 2007년에 32편(68.1%)으로 가장 많이 사용되었다. 적용된 통계기법 수는 1가지 방법을 사용한 논문이 125편(61.9%)으로 가장 많았으며, 2가지 방법을 사용한 논문은 52편(25.7%) 있었고, 최대 5가지 방법을 사용한 논문도 1편이 있었다. 적용된 통계기법으로는 분산분석을 사용한 논문이 81편으로 가장 많았고, T검정이 72편으로 두 번째 많았으며, 카이제곱검정이 62편으로 세 번째 많이 사용되었다. 그 밖에도 신뢰도분석, 회귀분석, 판별분석, 상관분석 등이 사용되었으며, 비모수적 검정도 31편에서 사용되었다. 통계기법을 가장 많이 사용한 연도로는 2007년 32편의 논문에서 46가지가 사용되었고, 편당 사용수가 가장 많은 연도는 2003년이었으며 17편의 논문에서 32가지로 편당 1.88가지의 통계기법을 사용하였다. 통계기법의 오류는 분산분석을 사용한 논문 81편 중 49편(60.5%), T검정은 72편 중 25편(34.7%), 카이제곱검정은 62편 중 18(27.4%)편이 있었다. 가장 많이 발견된 오류로는 분산분석과 T검정 모두에서 크기가 매우 작음에도 불구하고 정규성 및 등분산성 가정에 대한 아무런 언급이 없이 해당 분석을 실행한 경우로 한 집단의 표본수가 3인 경우도 있었다. 카이제곱검정에서는 각 칸의 기대도수가 너무 작은 경우에 도수에 대한 특별한 검토 없이 카이제곱검정을 실행한 경우가 많이 있었다.

분산분석의 경우 표본의 크기가 너무 작으면 유의한 분석결과가 도출되어도 그 결과에 대한 신뢰성을 떨어뜨리므로 올바른 분석이라고 할 수 없다. 이때는 비모수적 검정, 예를들면 Kruskal-Wallis 검정이나 Friedman 검정을 사용하는 것이 바람직하다. 그리고 분산분석에서 집단 간 차이가 유의한 결과에 대해 사후검정을 실행하지 않는 것은 연구 결과에 대한 더

많은 설명을 포기하는 것과 같다. 유의한 값을 통해 집단 간 차이가 있음을 보였다면 사후 검정을 통해 어떤 집단 간 차이가 있었는지 더 자세한 분석결과를 보여주는 것이 옳바르다. 또 유의수준의 선택은 귀무가설에 대한 새로운 대립가설의 채택과 기각의 기준이 되므로 연구자가 특정한 유의수준에 기초하여 결론을 도출하는 경우 이를 명시하는 것이 바람직하다. 그러므로 유의수준을 미리 정하고 그에 따른 결론을 기술하는 것은 중요한 일이다. 마지막으로 분산분석 후 사후검정에서 다중비교를 T검정으로 두 집단을 반복 비교하면, 전체 검정의 유의수준이 커져 잘못된 검정이 된다.

T검정에서도 표본의 크기가 너무 작으면 결과에 대한 신뢰성을 떨어뜨리므로 올바른 분석이 아니다. 그러므로 이때는 T검정에 대응하는 비모수검정, 예를 들면 Mann-Whitney 순위합검정이나 Wilcoxon 순위검정을 사용하는 것이 바람직하다. 그리고 일반적으로 대응표본 T검정에서 짝지어 대응이 있는 실험을 하는 목적은 두 가지 처리 결과에 대한 비교의 정도를 높이기 위한 것이므로 독립표본 T검정을 사용하면 안된다.

카이제곱검정은 다항분포에서 각 범주에 속할 확률이 (p_1, p_2, \dots, p_k) 인지를 검정하고자 하는 것이다. 이때 i 번째 범주에 해당하는 관측도수 n_i 과 이의 기대도수 np_i 와의 차이 $n_i - np_i, i = 1, 2, \dots, k$ 에 의존하는 검정통계량을 생각하게 되고 검정통계량의 값이 크면 귀무가설 즉, (p_1, p_2, \dots, p_k) 의 확률모형을 따른다는 가설이 기각될 것이다. 이 검정통계량은 표본크기 n 이 큰 경우 근사적으로 χ^2 분포를 따르게 된다. 표본의 크기가 크고 기대도수 np_i 가 5보다 같거나 클 때 이 근사적 접근이 무리가 없다고 알려져 있다. 따라서 분할표상에서 각 범주의 기대도수가 5미만의 값을 가지면 주변의 칸과 통합하여 기대도수가 5 이상이 되도록 수정하거나 Fisher의 정확검정(Fisher's exact test)을 사용해야한다⁵. 그리고 유의확률을 제시하지 않은 경우는 귀무가설에 반하는 경험적 증거의 강약을 나타내는 것이 유의확률이므로 이를 연구결과에 포함시켜야 한다.

본고는 통계기법의 오류를 살펴보는데 있어 기존의 선행연구와 통계학서적을 참조하여 객관적 서술이 되도록 노력하였으나 논문에 기술된 내용만으로 자료의 정보를 살피는데 한계가 있었고, 통계학적 지식부족과 사상체질의학에 대한 이해부족으로 적절하지 못한 검토가 이루어진 측면도 있을 것이다. 하지만 여러 학문 분야에서와 같이 사상체질의학에서도 통계적 기법을 사용하고 있는 이상 연구자는 다른 연구 논문 및 통계학 서적 등을 통해 활용하고자 하는 통계적 방법론에 대해서 충분히 검토하고 숙지한 후 연구에 임해야 할 것이다.

최종후·김항규(1994)⁵는 통계적 기법 활용을 개선하기 위해 방안으로 통계학 교육의 확충, 통계상담실(통계전문가)의 활용, 레프리(referee) 제도의 활용, 외국학술지의 비판적 수용, 통계적 기법을 활용하는 연구자의 올바른 자세를 주문했다. 사상체질의학 분야에서도 이러한 개선방안을 통해 통계적 기법을 더욱 자유자재로 활용한다면 향후 연구 결과에 대한 객관성과 신뢰성을 더욱 높일 수 있을 것이다.

V. 感謝의 글

이 논문은 2010년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 20100020618)

VI. 參考文獻

1. Kim TI. A Review of Statistical Methods in Public Policy Studies. Public Administration Review. 1997;31(3):19-35. (Korean)
2. Lee KM, Paik JK, Yoo YS. A Study on the Adequacy of the Statistical Method Using in the Design Papers. Journal of Korean Society of Design Science. 2003; 16(2):5-436. (Korean)
3. Lee HK, Ahn YO, Huh BY. An Assessment of Methodological and Ststistical Validity of Medical Articles Published in Korea, From 1980 to 1989. Korean Journal of Family Medicine. 1991;12(6):46-67. (Korean)
4. Lee JY, Lee SD. A Study Trend on the Research Designs and Statistical Data Analysis in Korean Oriental Medicine Literatures. Korean Journal of Oriental Medicine. 1998;1(1):105-118. (Korean)
5. 최종후, 김항규. 과학적 방법론과 통계기법의 활용. 파주:자유아카데미. 1994.