

생물통계의 현황(1981-1990)

신한풍* 송혜향** 김병수*** 이증협**** 한원식*****

요약

생물통계학(biometrics)의 여러 분야중 의학, 보건학, 농학 및 독성학 등 각 분야별로 지난 10년간(1981-1990) 학술전문지에 실린 논문을 통해 사용된 통계 방법론의 빈도를 연도별로 혹은 자료유형별로 조사 분석하였다. 주요 보건통계의 현황과 농업연구 분야에서의 통계적 기법 사용상의 문제점을 살펴보고 개선방향을 제시하였으며, 의과대학 교과과정중 통계학 과목의 현황과 방법론으로써 전반적인 활용을 살펴 보았다. 또한 독성학 분야에서 일반적으로 다루어지는 통계적 문제를 소개하고 우리나라 독성학 분야 논문에서 가장 많이 다루는 것으로 나타난 용량-반응 분석기법에 대한 고찰을 하고 있다.

본 연구를 통하여 우리나라의 생물통계의 현황을 파악하고 장래의 방향을 제시하고자 한다.

* 고려대학교 통계학과

** 가톨릭 의과대학 통계학과

*** 연세대학교 응용통계학과

**** 한국보건사회연구원

***** 농촌진흥청

생물통계학(biometrics)이 국내에 소개된 지는 분야별로 다양한 역사를 가지고 있지만 관련분야에서 그 활용현황이 체계적으로 조사된 것은 거의 없었다. 따라서 본 저자들은 지난 10년간(1981-1990)을 중심으로 생물통계학의 여러분야중 의학, 보건학, 독성학 및 농학분야에서 통계적 방법론이 어떻게 활용되고 있는지, 그 활용상의 문제점은 무엇인지 그리고 각 분야별 통계 및 자료분석현황이 어떠한지를 종합적으로 평가하여 생물통계 분야에서 장래 통계인들의 역할에 필요한 정보를 제공해 주고자 한다. 본 논문에서 조사한 내용을 각 분야별로 살펴보면 다음과 같다.

의학분야

우리나라의 의학통계는 1980년대에 들어서면서 그 발전의 토대가 될 수 있는 여러 변화에 접하게 되었다. 첫째로는 의료보험 시대의 시작으로써 이는 1960년대의 공무원 연금법에서 그 시초를 볼 수 있으며, 현재 국민 모두가 큰 어려움없이 의료기관을 이용할 수 있다. 따라서 병원을 찾는 빈도가 증가되었고 입원환자의 수도 함께 증가하였다(한국인구보건 연구원, 1987). 이러한 사실은 다른 한편으로 역학·의학연구의 활성화를 의미하기도 한다. 둘째로는 우리나라 사회 전반에 걸친 전산화 시대로의 변화를 들 수 있다. 병원도 예외없이 병원 전산화를 시작하게 되었으며 큰 대학 병원들이 지난 5년사이에 전산화의 장기계획을 출범시켜 현재 이 계획들을 진행중에 있다(개인적 문의). 이와 같은 변화에 접하여 의학통계의 발전면에서 문제시되는 점을 본 논문에서 생각해 보고자 한다. 또한 의학분야의 연구와 통계적 방법론의 활용이 활발해지려면 통계학 교육의 질적인 향상이 요구되는 바로써 이를 뒷받침하고 있는 우리나라 의과대학 교과목에서의 통계적 방법론의 교육현황을 알아 보고자 한다. 마지막으로 지난 10년간 의학연구에서 응용되고 있는 통계적 방법론을 연구전문지를 통하여서 조사해 보고자 한다.

보건학분야

“국민의 건강상태 및 보건의료자원이나 서어비스의 실태와 그 이용에 관한 통계”라고 정의할 수 있는 보건통계의 현황을 파악하고자 통계법에 의거하여 작성승인된 보건분야의 주요통계를 작성기관, 작성주기 및 조사내용을 중심으로 조사하였다. 총 344종의 통계중 6.4%만이 보건분야에 관련된 것으로 나타났다. 1980년대 이후 도시의 산업화, 인구의 노령화, 식생활 개선 등 환경의 변화로 국민의 의료수요가 급증하여 국민보건에 대한 중요성이 날로 증대되고 있다. 이러한 상황에서 통계인들이 늘어나고 있는 보건수요에 어떻게 대응해야 하는지를 알아보기 위해 지난 80년대 4개의 학술전문지 즉, 「한국보건통계학회지」, 「보건교육학회지」, 「인구보건논집」, 「보건학논집」을 연도별 각 한권을 임의로 선택하여 논문에 사용된 통계적 방법론을 연도별로 조사하였다. 총 299편의 논문중 83.6%의 논문이 다양한 통계기법

을 사용하고 있는 것으로 나타났다.

독성학분야

독성학은 대부분의 사람들에게 생소한 학문분야이고 우리나라에서는 1985년에 한국 독성학회가 설립되었다. 본 연구에서는 우리나라의 독성학 분야에 대한 간단한 소개와 일반적으로 독성학 분야에서 다루어지는 통계적 문제들을 살펴본 후 국내 독성학 분야 연구 논문중 「한국 독성학회지」(85-90)에 실린 논문을 조사하여 제시된 통계자료 및 통계분석의 유형을 살펴보고자 한다.

「한국 독성학회지」(85-90)의 총 84편의 발표논문 중에 60편(71.4%)의 논문이 통계적 분석을 시도하고 있다. 제시된 통계자료는 69건으로 크게 7가지로 분류될 수 있으며, 그 중 용량-반응 분석에 관련된 통계자료가 24건(34.8%)으로 가장 많았다. 제시된 통계자료의 유형과 사용된 통계분석의 기법은 차이를 나타내고 있으며, 이러한 자료의 유형과 분석 기법의 차이, 즉 적절한 분석기법의 사용은 이미 국내의 의학 분야에서도 지적된 바 있다.

본 연구에서는 조사된 논문에서 발견된 통계적 오류의 구체적인 형태를 밝히고 있으며, 용량-반응 분석에서 많이 다루고 있는 에임즈 쉘모넬라 변이원성 시험자료에 대한 분석과 최기형성자료의 동복자 효과에 대한 통계적 분석방법을 소개하고 있다.

농학분야

실제 농사시험 연구과정에서 통계적 방법이 사용된 것은 1962년 농촌진흥청 산하의 시험장, 연구소가 확대 설치되면서 부터이다. 그러나 1960년대에는 컴퓨터 보급이 원활치 못해 통계적 방법의 적용에 한계가 있었으며 개발된 많은 분석방법이 보편화되지 않았다. 그러던 중 1970년도에 농촌진흥청에 FACOM 230/10이 설치된 것을 기점으로 통계적 방법의 응용과 시험연구 방법에 있어 많은 발전을 거듭해 오다가 1978년 전산시스템이 확장되고, 연구소, 시험장 및 도 농촌진흥청에 단말기가 설치됨에 따라 사용되는 통계적 방법도 다양해지게 되었다. 따라서 현재까지 농업실험연구에서 사용되고 있는 통계적 방법론의 현황을 살펴보는 것은 이 분야 발전에 가치있는 일일 것이다. 농촌진흥청에서 개발하여 사용중인 “한국농업연구문헌정보” 데이터베이스를 이용하여 1960년부터 1991년까지 통계적 방법이나 응용에 관한 색인어를 가지고 있는 문헌을 학회지별로 검색하였다. 그 결과 품종개량을 위한 육종, 작물재배, 축산 등에 통계적 방법등이 많이 이용되고 있으며 보다 상세하게 파악하기 위하여 5년 간격으로 통계적 방법의 빈도와 사용된 통계적 방법의 개략적 내용을 주제별로 요약하였으며 이에 대한 사용상의 문제점과 개선 방안을 제시하였다.

의학, 보건학, 독성학 그리고 농학 분야에서의 통계현황 및 통계적 방법론의 활용현황을 각 분야별로 상세히 살펴보기로 하자.

1. 의학

가. 의학통계 자료의 전산화에 따른 문제점

각 대학병원이 독립적으로 전산화 작업을 실행시킴은 의학통계의 발전면에서 문제를 제기하게 된다. 각 병원이 의무기록을 독립적으로 전산화할 때 병원과 병원간의 통일성이 유지될 수 없으며, 장래 병원자료의 비교 및 교환의 가능성이 희박해 진다. 이러한 사실은 큰 규모, 또는 전국 내지는 국제적인 규모의 역학 및 의학연구의 가능성을 배제한다고 하겠다. 외국의 경우에 중요한 역학·의학연구의 다수가 병원자료에 근거하였음을 우리는 이미 알고 있다. 특히, 여러 질병간의 관련성에 관한 연구와 암연구에 병원자료가 중요한 역할을 하였다. 그러므로 현실적으로 공동의 개발이 불가능하다 하더라도 의무기록, 환자통계 및 기타 병원관리 통계들의 수집과 분석을 위한, 통일된 관리 체계의 확립이 요구된다. 여러 병원 통계 관리자들이 함께 모여서 최소의 자료를 규정함이 필요하며, 또한 공통된 용어의 정의, 일치된 질병 및 치료의 분류방법과 자료의 일관성있는 수집방법이 추구되어야 한다.

여러 병원으로부터의 자료의 종합은 이미 등록부(registry)를 통하여 이루어지고 있다. 저자중 1인이 참석한 우리나라 투석치료 및 신이식환자 등록부(Korea Dialysis and Transplant Registry)가 그 한 예이다(송혜향 등, 1988). 외국의 경우에는 이미 오래전에 등록부가 확립되었으며, 세계적으로 널리 알려진 영국 암등록부(England and Wales National Cancer Registry)는 1945년에 시작되었다. 이 영국 암등록부는 현재 전세계적으로 암에 관련된 가장 큰 자료를 확보하고 있으며, 매해 20만명이란 놀라운 숫자의 자료가 등록되고 있다고 한다. 사회 전반의 전산화와 더불어 특수성을 유지하는 여러 질병 등록부가 우리나라에서 시작되었으며, 의미있는 통계적 분석을 위해서는 자료의 완결성과 정확성이 유지되는 등록부 자료가 확보되어야 한다. 우리나라 투석치료 및 신이식환자 등록부의 경우에 자료수집과정에서의 문제점과 불성실한 자료의 결과와 앞으로의 대응책을 제시하였다(송혜향 등, 1988).

나. 의학분야에서 통계적 방법론의 교육현황과 전망

의학통계를 정의하게 되면 통계학에서 다루고 있는 통계 또는 통계적 방법론과 의학분야에서만 특이하게 적용되고 있는 통계 또는 통계적 방법론 전부를 일컫는다. 여기서는 이 둘을 구분치 않을 것이며 우리나라 의과대학교과정중에서의 통계학의 위치를 파악함으로써 의학통계의 현황을 알아 보

고자 한다.

1991년 현재 우리나라에는 총 34개의 의과대학 또는 의학부가 있다. 의과대학 교육과정에서의 의학통계의 현황을 파악하기 위해 1991년에 실시되고 있는 교과목중에서 통계방법론에 관련된 과목에 대해 조사하게 되었다. 1971년에 발족한 한국의과대학장 협의회에서 격년제로 발간하고있는 「의과대학교육현황」(1990-1991년도, 제6집)에 실린 자료가 이 조사의 근거가 되었으며 통계학과 관련이 깊은 전산과목도 함께 조사하였다. 「의과대학교육현황」에 실린 31개의 대학이 조사대상이 되었고 이들 대학이 제공하는 통계과목 및 전산과목의 현황은 [표 1-1]과 같다.

통계학 과목이 있는 의과대학은 총 29개 대학이 되며 그러나 통계학이 다수의 선택과목으로 함께 묶여서 보고된 경우에는 조사의 대상이 될 수 없었다. 의과대학에서의 통계학 과목은 매우 다양한 명칭으로 불리우고 있다. 대다수인 16개 대학에서는 통계학으로 불리우고 있고, 8개 대학에서는 의학통계학, 2개 대학에서는 생물통계학, 각각 1개 대학에서는 보건통계학과 통계수학으로 불리우고 있다. 통계학 과목은 1-4학점이며 이는 일반적으로 의예과 과정에 배정되어 있다. 통계학과목의 강의시간은 대학에 따라 매우 큰 차이가 있으며 16-72시간이 된다. (경상대학교 의과대학의 경우는 잘못 제시된 경우로 판단되어 여기서 제외되었는데 학점이 3학점인데도 불구하고 총 강의시간이 3시간으로 되어 있다.) 통계학 과목의 실습이 실시되고 있는 학교는 2개 대학에 불과하여 과연 통계학의 내용이 의학도들에게 충분히 전달되고 있는지에 대해서는 의아스럽다 하겠다. 2개 대학에서는 통계학 과목이 의예과 과정에서보다도 의학부 과정에서 배우도록 되어 있다.

[표 1-1]에서 알 수 있듯이 대다수의 의과대학 교과목에 통계학이 포함되어 있으며 또한 이들의 학점수가 총학점수에 비하여 작은 비중이 아닌 것으로 미루어 보아서 의과대학 교육과정에서 통계학의 위치가 일단 성공적으로 확립되어 있다고 할 수 있다.

대학에서도 전산화 시대의 반영을 볼 수 있는데 전산학 과목이 있는 의과대학은 총 23개의 대학이 되며, 이 과목은 고려대학(6학점)을 제외하고는 1-4학점이며 전부가 의예과 과정에 배정되어 있다. 전산학 과목의 강의시간은 16-112시간으로 대학에 따라 매우 큰 차이를 보이고 있으며 이는 통계학 과목에 비하여 매우 많은 시간이 되는데 그 이유로는 8개의 대학에서 전산학 과목이 실습을 실시하고 있기 때문이다. (경상대학교 의과대학의 경우는 잘못 제시된 경우로 판단되어 여기서 제외되었는데 학점이 3학점인데도 불구하고 총 강의시간이 3시간으로 되어 있다.) 통계학과 전산학을 종합한 학점은 2-8학점이다.

의학분야에서 통계적 방법론의 활용과 연구가 활발해지려면 우선 의과대학에서의 통계교육이 그 기본이라 할 수 있다. 그러나 현재 과연 적절하고

[표 1-1] 의과대학 교과목중에서 통계학 및 전산학 과목의 현황

대학	통계학 과목					전산학 과목					통계학 전산학 총학점 (의예과)				
	교과명	강의 시간	강의 학점	실습 시간	실습 학점	계 시간	계 학점	교과명	강의 시간	강의 학점		실습 시간	실습 학점	계 시간	계 학점
가톨릭대학	통계학	72	4	36	108	4	정보과학	36	2	36	72	2	6(91)		
건국대학교	통계학	48	3		48	3	전자계산	48	3		48	3	6(80)		
경북대학교	통계학	48	3		48	3							3(86)		
경상대학교	의 학	3	3		3	3	컴퓨터 프로그램	3	3		3	3	6(80)		
경희대학교	통계학 의 학	32	2		32	2							2(83)		
계명대학교	통계학 의 학	32	2		32	2	전자계산	32	2		32	2	4(88)		
고려대학교	통계학 이 해	48	2		48	2	컴퓨터	56	3	56	3	112	6	8(74)	
고신대학	통계학	48	2		48	2	전산학	64	2		64	2	4(83)		
단국대학교	의 학	16	2		16	2	의 학	16	1		16	1	3(82)		
동국대학교	통계학 의 학	32	2		32	2	컴퓨터 전자계산			64	2	64	2	4(86)	
동아대학교	통계학	48	3		48	3	전자계산	32	2		32	2	5(114)		
부산대학교	생 물 통계학	48	3		48	3							3(90)		
서울대학교															
순천향대학	통계학	48	3		48	3	전산학 개 본 프로그램	64	4		64	2	7(92)		
아주대학교															
연세대학교			3			3									
연세 (원주)	통계학*	36	2		36	2	전산학 개 본	32	2	32	64	2	4(89)		
영남대학교	통계학	48	3		48	3	전산개본	32	2		32	2	5(76)		
울산대학교	통계학	48	2		48	2	전 산 프로그램	64	2		64	2	4(87)		
원광대학교	보 건 통계학	32	1		32	1	전산이론	32	2		32	2	3(87)		
이화여대															
인제대학	통계학 생 물	32	1		32	1									
인하대학교	통계학	32	2		32	2									
인하대학교	통계학	48	3		48	3	컴퓨터 언 어	48	3		48	3	6(80)		
전남대학교	통계학*	48	3		48	3	전산학*	48	3		48	3	6(80)		
전북대학교	통계학	32	2		32	2	전산학 개 본*	48	3		48	3	5(84)		
조선대학교	의 학	32	2			2							2(81)		
중앙대학교	통계학 의 학	32	2		32	2	컴퓨터	32	2		32	2	4(97)		
충남대학교	통계학 개 본*	32	2		32	2	전산입문	48	3		48	3			
충남대학교	의 학	16	1		16	1							6(81)		
충북대학교	통계학	48	3		48	3	전산학	32	2	32	1	64	3	6(82)	
한림대학	의 학	32	2	32	1	64	3	전산개본	32	2	32	1	64	3	6(95)
한양대학교	통 계 수 학	32	2		32	2	컴퓨터 언 어			64	2	64	2	4(87)	

* 선택과목임.

* 의학과 과정의 과목임.

** 통계학이 두 과목으로 구성되어 있음.

주) 서울대학교, 이화여대는 통계학이나 전산과목이 없었음.

연세대학교는 1990-1991년도 자료에 의예과 과목이 제시되지 않았기에 1988-1989년도 자료에 근거하였음.

유용한 통계교육이 이루어지고 있는지에 대해서는 확신할 수 없는데, 그 이유로는 첫째로 의학인들에게 적절한 의학분야의 예제로 이루어진 통계교과서가 몇몇에 불과하다는 점이며, 둘째로 이들 통계학 과목이 의학통계 또는 통계 전문교수에 의해 강의되기보다는 통계전공이 아닌 의학전공의 교수들에 의해 강의되고 있는 실정이라는 점이다. 따라서 의과대학에서의 통계교육의 질적인 향상이 앞으로의 과제로써 남아있다 하겠다. 또한 장래의 통계 발전에는 통계적 방법론의 이용이 필수적인 의학연구자와 통계인과의 밀접한 협조체제가 확립되어야 하는데, 이를 위해서는 의학통계를 전문으로 하는 통계인의 수적인 증가와 이들을 수용할 수 있는 교실이나 연구기관이 확보됨이 우선적이며, 이의 빠른 실현을 장차 기대하는 바라 하겠다.

다. 의학분야에서 통계적 방법론의 활용방안

이제 의학연구 분야에서의 통계적 방법론의 전반적인 활용에 대해 알아보려고 한다. 질병의 예방이나 감소에 심혈을 기울이는 예방의학의 연구자나 역학자는 질병 원인이 될 수 있는 요인의 발견에 있어 통계적 방법론에 깊이 의존하고 있으며, 다른 어떤 분야보다도 이들 전문지에 실린 논문들에서 다양한 통계적 방법론을 많이 발견할 수 있다. 그러므로 이러한 배경을 가진 예방의학회지는 통계적 방법론의 전반적인 활용에 대해 알아볼 수 있는 좋은 검토자료가 되며, 또한 장래 의학인들을 교육하거나 장래 상담자를 접하는데 있어서 통계인들에게 중요한 정보를 제공하게 된다.

대한 예방의학회에서 발간하고 있는 「예방의학회지(Korean Journal of Preventive Medicine)」중 1981년부터 1990년까지의 각 권에서 제1호만을 선정하여 원저논문에서 사용된 통계적인 방법론에 대해 조사하였다. 매해 발간되는 권수가 다르므로 일관성있게 제1호만을 선정하였다. 통계적 방법론의 분류는 Hokanson외 3인(1987)의 논문을 참조하였으며 그러나 예방의학회지에 실린 논문에서 사용된 통계방법론에 맞추어 적절하게 수정하였다.

통계적 방법론이 전혀 사용되지 않은 경우와 통계가 사용되었어도 기술 통계치(descriptive statistics), 즉 평균, 중앙치, 최빈치, 백분율만이 언급된 경우를 따로 구분하였다. 또한 기술 통계치(descriptive statistics), 산포(dispersion)와 더불어 더욱 복잡한 통계 방법론이 사용된 경우에는 각각을 누계토록 하였다. 한 논문에 예를 들어서 t-test가 여러번 사용되고 언급되었다 하더라도 한번으로 누계되었다. 유의확률 p가 제시된 것으로 보

아 분명 검정법이 사용되었으나 구체적인 검정법이 논문 내용이나 표에 전혀 언급되지 않은 경우는 따로 구분하여 누적하였다. [표 1-2]에 제시된 23가지 분류에 속하지 않을 때에는 기타에 포함시켰으며 사용된 방법론을 명시하였다.

총 191편의 논문중에서 통계적인 방법론이 전혀 언급되지 않은 논문이 11편(5.8%)이며 기술 통계치(descriptive statistics), 즉 평균, 중앙치, 최빈치, 백분율의 보고에 그치고 있는 논문이 27편(14.1%)이며 이들의 연도별 분포는 [표 1-2]와 같다. 기술 통계치를 위시하여 통계적인 방법론이 보고된 180편(94.2%)의 논문중에서 총 465건의 통계적 방법론이 사용되어 한편의 논문에서 여러 방법론이 사용되었음을 시사하고 있다.

10년간의 논문을 종합하여 이에 사용된 통계적 방법론의 빈도를 살펴보면 기술 통계치와 산포(dispersion)를 제외하고는 첫째로 카이제곱검정이 가장 많이 사용되었고(50건), 둘째로 t-test(32건)와 단순회귀와 중회귀를 총괄한 회귀(32건)가 동일하게 많이 사용되었다. 다음으로 상관계수(29건), 역학적인 통계(24건), ANOVA(24건)의 순서이다. 이외에도 비모수 검정법이 8건, 로지스틱 회귀 5건, 생존분석 4건, 판별분석 4건, 인자분석 4건, 만텔-헨젤 방법과 로그 리니어 모형이 4건이었다. 특히 생존분석, 판별분석, 인자분석, 로지스틱 회귀, 만텔-헨젤 방법과 로그 리니어 모형과 같이 컴퓨터 사용이 불가피한 통계적인 방법론은 일반적으로 최근에 와서야 사용되었음을 볼 수 있다. 통계적인 검정 방법론의 사용 후 논문에 언급함에 있어 부족한 경우로써 통계 검정법의 명시가 전혀 없이 p값만이 보고된 경우라 하겠다. 이는 모두 18건으로 매우 많은 수라 하겠다.

기술통계치, 산포를 비롯하여 t-test, 회귀, ANOVA, 카이제곱, 상관계수는 모든 기초 통계교과서에서 취급하고 있는 방법론이라 하겠다. 그 외의 방법론들, 특히 다변량 분석법은 의학 교과과정에서 배우지 못하는 방법론이며, 더욱 바람직하고 효율적인 의학연구를 위해서 이러한 고차원적인 통계 방법론들이 사용된 빈도수를 참조하여 장래 통계교육자가 가르쳐야하고 주력해야 하는 방법론들의 우선 순위를 세울 수 있다([표 1-2]).

[표 1-2] 예방의학회지(81-90) 논문에서 사용된 통계 방법론의 빈도수

No.	Category	Description	'81	'82	'83	'84	'85	'86	'87	'88	'89	'90	Total
1.	Descriptive statistics	Mean, median, mode, percentage	7	12	17	25	14	12	13	13	10	9	132
2.	Dispersion	SD, SEM, range coefficient of variation	1	8	7	19	10	7	5	8	6	7	78
3.	t-test	One-sample, matched-pair, two-sample	2	4	2	8	2	2	2	4	5	1	32
4.	Analysis of Variance	ANOVA, repeated measures	4	3	2	4	3			3	3	2	24

5. χ^2 -test	Contingency tables, Yate's correction, McNemar's test	2	6	6	10	6	3	5	6	5	1	50
6. Pearson Correlation	Bivariate product - moment correlation	1	3	6	8	2	1	3	2	1	2	29
7. Simple Linear regression	least - squares regression with one predictor and response variables	1	3	2	5		1	3	1	2		18
8. Multiple linear regression	Stepwise regression polynomial regression	1	1	2	1	1		2	3	2	1	14
9. Nonparametric Tests	Wilcoxon signed - rank, Mann - Whitney U	1	1					1	2	2	1	8
10. P-value not otherwise specified	Use of statistical test of significance with no explicit identification of methodology	1	3	3		3	1	2	2	2	1	18
11. Epidemiologic statistics	Relative risk, odds ratio, log odds, incidence, prevalence, measures of association, sensitivity, specificity	1	1	7	2	1	4	3	3	2		24
12. Survival analysis	Actuarial life tables, Kaplan - Meier, Breslow's									3	1	4
13. Multiple comparisons	Brekson - Gage, Cox Multiple inferences on same data sets: Scheffe's contrasts, Duncan multiple range procedure, Newman - Keuls procedure	1										1
14. Mathematical models	Model using Polya - Eggenberger distribution, simulation			1							1	2
15. Transformation	Logarithm	1			1							2
16. Discriminant function analysis	Mahalanobis distance, Rao Wilk's statistics				1			1		2		4
17. Factor analysis	Principal components: varimax, quartimax or orthogonal rotation					1		1	1		1	4
18. Logistic regression	Logistic regression	1						1	1	1	1	5
19. Multiway tables	Mantel - Haenszel procedure, log - linear model	2								1	1	4
20. Nonparametric correlation	Spearman's ρ	1										1
21. k statistics	Coefficient of agreement	1										1
22. Reliability	Reliability assessment	1						1			1	3
23. Nonlinear regression	Simple or multiple nonlinear regression			1					1			2
24. Other	LD50, Kolmogorov - smimov test, time - series analysis, Bayesian probability	1	1				1			1	1	5
Number of articles reviewed		14	24	23	40	15	13	15	17	19	11	191
Number of articles with no statistics		2	3	2	0	0	1	0	1	2	0	11
Number of articles with descriptive statistics only		6	6	5	4	1	1	1	3	0	0	27
Total number of procedures		18	56	52	88	45	33	44	50	49	34	465

2. 보건학

가. 보건통계의 현황

과거 고도성장 주도의 경제사회발전계획 수행과정중에서 발생한 환경의 오염, 도시화, 산업화 및 인구의 노령화 등에 대처할 수 있는 국가적 차원의 대응책은 국민적 욕구를 충족시킬 수 있는 보건의료 정책이라 하겠다. 따라서 보건관련 분야의 기초통계의 현황을 살펴봄으로써 정부의 보건정책의 효율성을 파악하고 보건통계의 문제점과 개선책에 대해 알아보하고자 한다.

보건통계라 함은 “국민의 건강상태 및 보건의료 자원이나 서어비스의 실태와 그 이용에 관한 통계”라 정의할 수 있는데 1989년 11월 1일 현재 통계법 제3조의 규정에 의거 작성 승인된 보건분야의 주요 통계는 [표 2-1]에 제시된 총 22종이 있는데 이는 다음과 같이 크게 세 부문으로 분류할 수 있다.

국민건강상태를 측정하는 기초통계 ————— 15종(국민영양 조사 등)
 의료자원에 관한 통계 ————— 3종(의료기관실태 보고 등)
 의약품에 관한 통계 ————— 4종(방역약품생산 및 지출상황표 등)

이중 보건사회부에서 작성되는 보건분야(사회복지부분 제외)의 통계는 20종에 불과하고, 2종은 한국보건사회연구원에서 조사작성되고 있는데 이는 총 344종의 승인통계중 6.4%에 불과한 것으로 국민의 보건의료에 대한 수요를 충족시키기 위한 통계자료가 부족함을 시사하고 있다. 또한 통계자료의 내용이 단편적이고 표준분류의 작성미비로 보건전반에 관한 종합적인 상호연계분석이 어려운 상태이다. 전문인력의 부족으로 통계의 정확성이 결여된 것 같으며 통계자료를 이용하는 데 용이하지 않았던 것도 사실이다.

따라서 통계전문가를 통한 보건통계교육 및 보건대학원에 위탁교육을 실시하여 보건통계 전문인력을 양성하여야 하며, 보건사회부의 통계담당관실을 확충하여 보건통계 작성기준의 표준화 작업 및 자료수집과 집계를 위한 전산화 작업을 추진하는 것이 바람직할 것이다. 사회복지, 성인병, 노인실태 및 약물중독 등과 같은 새로운 정책개발을 위한 새로운 통계의 개발에 적극적으로 참여하여 국민보건수요에 대응해야 할 것이다.

나. 보건학 분야에서 통계적 방법론의 활용현황

지난 10년간 생활환경의 변화로 새로운 질환들이 발생되고 있으며, 국민의 의료수요가 급증하여 국민보건에 대한 중요성이 날로 증대되고 있다. 따라서 어느 국가나 지역사회의 보건상태를 파악하거나, 보건학 관련분야의 여러 변량들의 행태를 연구하는데 있어 통계적 방법론이 다양하게 사용되고 있다. 본 논문에서는 보건학 계통의 학술전문지인 「한국보건통계학회지」(81-89), 「보건교육학회지」(83-90), 「인구보건논집」(81-90), 「보건학논집」

[표 2-1] 보건분야의 주요통계 현황

통계명칭	작성기관	작성주기	조사내용
국민영양조사	보건사회부	년	식품섭취, 식생활, 건강상태
국민건강조사	한국보건사회연구원	3년	가구에 관한 사항, 병태개별조사
전국결핵실태조사	보건사회부	5년	BCG 접종율, 결핵유병율
가족보건실태조사	한국보건사회연구원	3년	가족계획실태, 모자보건, 임신 및 출산, 인공임신중절실태
법정전염병환자월보	보건사회부	년	법정전염병 종류별 발생 및 사망수
결핵관리사업실적	보건사회부	월	결핵예방접종, 검진, 환자발견
만성병관리사업실적	보건사회부	분기	기생충관리 및 구충실적, 성병관리 및 치료실적
가족보건사업실적보고	보건사회부	월	가족보건사업의 대상인구, 등록, 분만, 사망 및 건강상담
전국기생충감염 실태조사	보건사회부	5년	지역별, 연령별 및 학업별 기생충감염현황
환자조사	보건사회부	2년	기관조사표, 외래환자조사표 퇴원환자조사표
식품산업기초조사	보건사회부	년	사업체의 일반사항 및 투자액, 생산비
구강보건사업현황	보건사회부	분기	치과이동분류 및 불소실적
국제검역상황	보건사회부	분기	입항선박 및 상륙인원, 검역 구역내 전염병 발생 및 사망
의료기관실태보고	보건사회부	반기	병원 및 병실 수, 환자 및 퇴원자수
보건소 및 보건지소 진료소현황	보건사회부	분기	인력현황 및 활동현황
조산원교육병원현황	보건사회부	년	교육생현황, 분만실적
방역약품생산 및 지출상황표	보건사회부	월	예방약품현황, 진단액, 진단용 혈청
의약품 등 생산 실적보고	보건사회부	반기	업소별, 품목별, 효능별 생산량, 생산금액
식품환경관계 실태 현황	보건사회부	월	유해업소 및 식품업소 현황
불소함유량 수질 검사 보고	보건사회부	월	불소함유량, PH 온도, 대장균, 배수량
부정의료업자 단속현황	보건사회부	분기	적발 및 처리 현황(고발, 폐쇄)
약사감시상황	보건사회부	분기	의약품, 향정신성의약품, 위생용품

[표 2-2] 보건학분야의 전문잡지·논문에서 사용된 통계방법론의 빈도수

No.	Category	Description	'81	'82	'83	'84	'85	'86	'87	'88	'89	'90	Total	
1.	Descriptive statistics	Mean, median, mode, percentage	10	12	15	27	21	16	15	22	22	12	172	
2.	Dispersion	SD, SEM, range, CV			5	4	11	5	6	1	4	1	37	
3.	t - test	One - sample, matched - pair, two - sample t tests			1	2	1	2		3	1	4	14	
4.	Analysis of Variance	ANOVA, ANCOVA, MANOVA repeated measures	1	2			5	1	1	2		3	15	
5.	χ^2 - test	Contingency tables, Yate's correction McNemar's test			1	5	3	3	2	3	5	4	7	33
6.	Correlation	Pearson correlation	1	4	4	5	3	4	1		3	3	28	
7.	Linear regression	Simple, multiple, polynomial, stepwise regression	3	3	2	4	7	5		1	5	4	34	
8.	Logistic regression	Logistic regression	2			2		1	1		1		7	
9.	Path analysis	Path analysis	2			1	2	1					6	
10.	P-value not otherwise specified	Use of statistical test of significance without explicit identification of methodology			2	3	3	4	2	1	2	1	18	
11.	Multiple comparison	Multiple inferences on same data sets: Scheffe's contrasts, Duncan multiple range procedure				1				1	2		4	
12.	Epidemiologic statistics	Relative risk, odds ratio, log odds, incidence, measures of association, sensitivity						1	2	1		1	5	
13.	Survival analysis	Life tables, Kaplan-Meier, Cox proportional hazard		2	2	3		2	5	1		2	17	
14.	Discriminant analysis	Mahalanobis distance, Rao Wilk's statistics		1				1					2	
15.	Factor analysis	Principal components: varimax, quartimax or orthogonal rotation		1			2					1	4	
16.	Multiway tables	Mantel - Haenszel procedure, log - linear model							1	1	1		3	
17.	Sampling	Simple random sampling, PPS	1		1	1		2	1		1		7	
18.	Other	Anything not fitting above headings: includes screening test, matching method, tobit analysis, various index, etc	4	5	4	1	8	1	1	6	4		34	
Number of articles reviewed			20	24	32	39	35	30	37	31	32	19	299	
Number of articles with no statistics			5	2	7	1	3	8	14	4	3	2	49	
Number of articles with descriptive statistics and/or dispersion			3	9	5	20	12	7	10	12	16	7	101	
Total number of procedures			24	37	42	62	62	48	33	47	48	37	440	

(81-89) 등 4개 학술전문잡지에 실린 논문을 중심으로 활용되고 있는 통계적 방법론을 살펴봄으로써 보건수요에 통계인들이 대응할 수 있는 정보를 제공하고자 한다.

4개의 전문잡지 각 권에서 임의로 한 호를 선정하여 실린 논문에 사용된 통계적 방법론을 조사하였다. 통계적 방법론의 분류는 Hokanson의 3인 (1987)의 논문을 참고하였으나 일부 수정보완하였다. 통계적 방법론이 전혀 사용되지 않은 경우와 사용되었다라도 기술통계(descriptive statistics)와 산포(dispersion)만이 사용된 경우를 따로 구분하였다. 또한 한 논문에서 여러가지 통계기법이 사용되었을 경우 이를 중복해서 누계하였으며, 표본설계를 통해서 얻은 조사자료를 분석한 경우에는 분석방법으로써 표본기법(sampling technique)에 포함시키지 않았다.

총 299편의 논문중 통계적 방법론이 전혀 언급되지 않은 논문이 49편 (16.4%)이며, 기술통계치나 산포만이 사용된 논문이 101편(33.8%)이다. 이들의 연도별 분포는 [표 2-2]와 같다. 통계적 방법론의 빈도를 살펴보면, 기술통계치와 산포를 제외하고는 가장 많이 사용된 방법론이 회귀 34건, 카이제곱검정 33건, 상관분석 28건 순으로 이들의 사용빈도가 거의 비슷하였다. 그 다음으로는 유의확률 p가 제시되었지만 구체적 검정법이 논문내용이나 표에 언급되지 않은 경우가 18건 이었고, 생존분석 17건, t-검정 14건, 분산분석(ANCOVA, MANOVA 각 1건 제외) 13건 순으로 사용되었다. 이외에도 THI(Todai Health Index)와 같은 다양한 지수들이 사용되고 있거나 임상검사에서 주로 사용되는 Screening test의 여러 측도와 같이 뚜렷하게 분류할 수 없는 경우가 기타항목으로써 34건에 이르고 있다.

이상에서 살펴본 바와 같이 33.8%의 논문이 기초통계(기술통계치와 산포)를 이용하여 통계분석을 하고 있는 것으로 보아 현상 및 문제 파악에 중점을 두고 있는 것 같고, 앞으로 심층분석이 추가될 개연성이 내포되어 있다고 사료된다. 기술통계치, 산포를 비롯하여 회귀, 카이제곱 검정, 상관계수, 분산분석(MANOVA 제외), t-검정 등은 전부 보건통계학 교과과정에서 취급되고 있는 방법론들이다. 그 이외의 방법들 특히, MANOVA, 생존분석, 요인분석과 같은 통계방법론들은 교과서에 잘 소개되어 있지 않은 실정이다. 따라서 통계교육자들은 보건관계 종사자들과 긴밀한 협조체계를 구축하여 장래 보건학 연구를 위해 필요한 방법론 선정에 [표 2-2]의 결과를 참조할 수 있을 것으로 생각된다.

3. 독성학*

가. 우리나라에서 독성학의 개관

* 본 원고의 초고를 읽고 도움 말씀을 주신 연세대학교 의과대학 예방의학 교실의 신동천 교수께 감사를 드린다.

한국독성학회는 1985년에 설립되었으며 1985년 12월에 「한국독성학회지 (Korean Journal of Toxicology)」 1권을 출간하면서 현재 6권까지 발간되어 있다. 이와 유사한 성격의 학회인 한국 환경성 돌연변이원·발암원 학회는 1981년 설립되어 현재 한국 환경성 돌연변이원·발암원 학회지를 10권까지 발간하고 있으며, 두 학회는 하나로 통합되어 단일 기구를 구성하고 있다.

한국독성학회의 주 구성원은 독성학, 약리학, 의학, 생물학, 미생물학, 수의학, 환경보건학 관계 학자 및 관련 제약업계이고 주요 관련 연구소로서는 화학 연구소내 Toxicology Center, 한국 과학 기술원내 Doping Control Center, 국립보건안전연구원, 국립환경보건연구원 등이 있다.

독성시험은 신물질 개발, 혹은 신약개발 과정에서 필수적으로 거쳐야 하는 시험과정이다. 신약개발과 독성시험에서 통계학의 역할에 대하여는 백운봉(1991)에서도 다루고 있다. 구미선진국에서는 새로운 화학물질을 등록할 때 독성학의 기초시험이라고 할 수 있는 에임즈 켈모넬라 돌연변이 시험 (Ames Salmonella mutagenicity test)를 포함한 단기시험(short term tests)의 결과를 함께 보고하도록 법제화하고 있다(Tennant et al, 1987). 또한 장기동물시험(long term rodent test)을 통하여 발암성등 만성적인 효과를 규명하는 시험도 독성시험의 범주에 속한다. 현재까지 확인된 화학물질의 발암원여부는 담배나 석면 등을 제외하고는 거의 모두 이 장기동물시험을 통하여 알려졌다.

우리나라에서는 아직 신물질 개발에 대한 시험이 본격적으로 이루어지고 있지 않고 있으며, 일부 시도하려는 실험실에서도 아직까지는 실험실 설치 단계라 할 수 있다. 그리고, 아직 보건 분야의 투자 규모가 작으므로 비용이 많이 드는 장기동물시험은 일부 실험실에서만 시행되고 있다. 그러나 신물질 개발이 국책사업으로 이루어지고 있는 만큼 향후 10년 동안 이 분야에 많은 투자가 이루어 질 것이며, 따라서 독성시험에 대한 수요는 급증하리라 기대된다.

나. 독성학 분야에서 통계적 분석 모형의 고찰

1991년 5월 일본의 동경에서 「독성학 분야에서 보건 통계학 국제회의 (International Biostatistics Conference in the Study of Toxicology)」가 열렸고, 1990년 9월 독일의 라이프치히에서 개최되었던 Eurotox'90 학회에서는 「독성학에서 통계적 방법론(Statistical Methodology in Toxicology)」 워크숍(Hothorn, 1991)이 열렸다. 이 두 학회의 주제 및 성격은 아주 흡사하며, 이 학회의 개최는 유럽과 일본에서 이 분야에 대한 관심을 나타내 준다고 말할 수 있다. 일본에서 개최된 「독성학 분야에서 보건 통계학 국제회의」의 주제는 크게 3가지로 분류되는데 이는 독성학 분야에서 다루어지는 통계적 분석 내용을 잘 나타내 주고 있다.

(1) 장기동물시험에서 생존시간, 종양 발생기간 (time-to-tumor), multiplicity, combining data에 관한 연구

이는 발암원을 규명하는데 사용되는 장기동물시험에 관한 내용으로서 이에 관련된 통계학 문제들은 Krewski and Brown(1981), McKnight and Crowley(1984), Haseman(1984), Portier(1986), Portier and Dinse(1987) 등에서 찾아볼 수 있으며 특히 다음과 같은 문제들을 다루고 있다.

- ① 실험동물에게 사용하는 고용량과는 달리 인간이 실제로 접하는 저용량에 대한 위해를 추정하는 기전적 모형을 구성하고 이를 식별하고 추정하는 연구
- ② 고용량의 독성은 수명에 영향을 주기 때문에 시간을 조정된 분석(time adjusted analysis)이 필요한 점
- ③ 종양발생기간이 직접 관찰되지 않고 또한 종양의 종류에 따라서 사망율이 각각 다른 점에 대한 고려, 죽음과 연관되지 않은 종양(incidental tumor)과 급치사 종양(rapidly lethal tumor)의 분리 고려
- ④ 3가지 상태 모형(three state model)의 tumor-free상태에서 tumor상태로 가는 전이율 추정
- ⑤ 과거대조군 자료(historical control data)의 활용

(2) 단기시험에서 시험 설계, 분석 및 성과

유전독성시험인 단기시험은 장기동물시험용 화학물질을 선별하는 과정으로 사용되고 있고, 궁극적으로는 비용이 많이 소요되는 장기동물시험의 대체(surrogate)가 되기를 희망하는 시험이다. 단기시험에는 가장 대표적인 에임즈 켈모넬라 돌연변이 시험(SAL)외에 Chinese Hamster Ovary(CHO) cell에서 Chromosome Aberration Assay(ABS)와 Sister Chromatid Exchange Induction(SCE), 그리고 Mouse Lymphoma L5178Y cell mutagenesis assay (MLY) 등이 있고 이러한 단기시험들과 장기동물시험과의 관계는 Tennant et al(1987)과 Haseman et al(1990)에서 볼 수 있다. Tennant et al(1987)의 주요 결과는 ① 위의 4가지 단기시험들은 발암원성과의 질적인 일치도에 있어서 큰 차이가 없고, 대개 60%의 일치도를 보이고 있다. ② 발암원성의 예측에 관한 한 SAL Assay이외에 4가지 단기시험들의 어떠한 조합도 발암원성의 예측도를 개선시키지는 못한다. ③ 단기시험(STT)의 음성반응은 비발암원성(noncarcinogenicity)을 예측하여 주지는 않지만 단기시험의 양성반응은 발암원성을 다소 예측하여 준다는 것이었다. Tennant et al(1987)의 이러한 결과에 대한 유전독성학계의 처음 반응은 다소 엇갈렸다. 위 시험이 혹시 잘못된 것은 아니었느냐 하는 반응도 있었고 이에 대한 응답으로서 추가적인 실험이 시행되었는데 추가실험의 결과는 애초의 Tennant et al(1987)의 결과와 다를 바가 없다는 보고였다 (Haseman et al, 1990).

또한 Margolin, Kim & Risko(1989)는 에임즈 켈모넬라 돌연변이 시험의

의사결정원칙으로서 통계적 의사결정원칙과 독성학자들이 관례적으로 사용하여 온 대조군 2배수 원칙(two fold background rule)을 비교하여 통계적 분석모형의 타당성을 보여주었다.

(3) 위해성 평가 (risk assessment)

생체외(in vitro)에서 실시되는 단기시험의 결과를 생체내(in vivo)로 외삽시키는 방법, 장기동물시험에서 동물에게 주어지는 고용량의 효과를 인간이 실제로 접하는 저용량의 수준으로 외삽시키는 방법, 또한 랫(rat)과 마우스(mouse)를 사용한 장기동물시험의 결과를 인간에게 적용시키는 절차와 방법 등을 다루고 있다. 최근의 Statistical Science(1988)는 제3집에서 이 문제를 특집으로 다루고 있다.

그리고 위에서 언급한 3가지 주제이외에도 동복자 효과(litter effect)를 주로 다루는 최기형(teratology)시험은 발달독성학(developmental toxicology)의 범주에 속한다. 최기형시험은 화학물질을 임신한 모태에 투여한후 그 태자(fetus)들을 통하여 화학물질의 효과를 관찰하고자 하는 시험이다. 태자들을 통하여 관찰하려는 반응변수가 연속형이고 정규분포로 근사할 수 있을 경우 동복자 효과는 분산분석에서 내포된 임의 효과모형(nested random effect model)을 구성한다. 또한 그 반응변수가 기형의 발생이나 죽음의 경우처럼 이분형일 경우 발생빈도는 이항분포가 아닌 이항분포의 혼합분포를 이루게 된다. 동복자 효과에 대한 자세한 언급은 Haseman and Soares(1976), Haseman and Kupper(1979)와 Kupper, Portier, Hogan and Yamamoto(1986), Yamamoto and Yanagimoto(1988)에서 찾아볼 수 있다. 그리고 Piegorsch and Haseman(1990)은 최기형 시험 자료의 분석에 있어서 동복자(litter)가 시험단위(experimental unit)로서 사용되어야 하는 점을 특히 강조하고 있다.

다. 우리나라의 독성학 분야 연구 논문에서 통계적 분석에 대한 고찰

(1) 현황 및 문제점

우리나라의 독성학 분야 연구논문중 「한국독성학회지」를 1권(1985)부터 6권(1990)까지 모든 논문을 조사하여 통계적 분석이 시도된 논문의 수와 통계자료의 유형을 각각 다음 [표 3-1]과 [표 3-2]와 같이 얻었다. 본문에서 통계적 분석이란 통계적 검정이나 추정이 이루어진 절차를 의미하며, 단순한 기술통계량의 나열은 통계적 분석으로 간주하지 않았다.

[표 3-2]에서 두개 이상의 유형에 분류되는 논문은 모두 8편이었다. 통계적 분석이 실시된 60편의 논문에서 저자들이 사용한 통계적 분석 기법은 모두 10개 정도로서 구체적으로는 다음 [표 3-3]과 같다.

[표 3-1] 한국독성학회지의 통계적 분석 논문 편수

년도	권	발표 논문수	통계적 분석이 실시된 논문수	통계적 분석논문 /발표논문수
1985	1	5	1	20.0 %
1986	2	10	8	80.0 %
1987	3	15	13	86.7 %
1988	4	15	12	80.0 %
1989	5	19	14	73.7 %
1990	6	20	12	60.0 %
		84	60	71.4 %

[표3-2] 제시된 통계자료의 유형

유형 (분석기법)	회수
1. 두그룹간의 비교	10
1.1 연속형 자료	6
1.2 범주형 자료	3
1.3 포아송 자료	1
2. 일원배열자료(1-AOV, MANOVA 혹은 Kruskal-Wallis 검정)	21
3. 반복관찰치가 있는 일원배열 자료 (1-AOV with repeated measure, Friedman Test, Quade Test)	9
4. 이원 배열 자료	2
5. 용량-반응 분석	24
5.1 에임즈 켈모넬라 돌연변이 시험자료 분석	3
5.2 용량에 따른 추세검정 및 기울기의 추정	9
5.3 두개 이상의 용량-반응 곡선의 비교	1
5.4 동복자 효과(litter effect)가 있는 기형학 자료분석	8
5.5 용량-반응 곡선의 추정이나 ED50 추정	2
5.6 두 약제의 상승효과나 반대효과 검정	1
6. 회귀분석	2
7. 기타	1
69	

[표 3-3] 사용된 통계적 분석 기법

내 용	회수
1. 아무런 언급없이 대조군과 비교한 P 값만 보고	11
2. 이 표본 t검정 (unpaired t test 로 불리워짐)	34
3. Mann-Whitney 검정 (Wilcoxon Rank Sum Test)	4
4. 일원배열 분산분석 (One Way ANOVA)	8
5. Kruskal-Wallis 검정	1
6. 반복관찰이 있는 일원배열 분산분석	0
7. 이원배열 분산분석	1
8. χ^2 검정 혹은 피셔의 정확한 검정	5
9. 목측으로 판단	3
10. Mean \pm SE , Mean \pm SD 혹은 수치만 보고	2
11. ED50추정이나 모수를 사용한 용량-반응 곡선 추정	2
12. 회귀계수의 추정	2
	73

[표 3-3]의 분류에서 두개이상의 통계적 분석기법을 사용한 논문은 모두 11편이었다. 반복관찰이 있는 일원배열 분산분석의 빈도가 0인 것은 어느 논문도 분석기법을 “반복관찰이 있는 일원배열 분산분석”으로 명시하지 않았다. 그러나 저자에 따라서는 일원배열 분산분석에 포함시켰을 수도 있으리라 생각되지만 확인을 할 수 없었다.

[표 3-2]와 [표 3-3]에서 볼 수 있는 자료의 유형과 사용된 분석기법의 분포상의 불일치는 이미 여러 곳에서 지적된 바 있다 (White(1979), 김병수, 안윤기, 윤기중(1988), 송혜향(1989)). 특히 송혜향(1989)교수가 지적한 5년전 의학논문의 통계적 오류중 다음은 본 저자가 지적하고자 하는 오류와 일치한다.

- ① 논문에 사용된 검정법이 명시되지 않았다. 60편의 논문중 20편의 논문만 통계적 분석기법을 명시하였다.
- ② 세 군 이상에서 처리군과 대조군의 구별없이 두군의 평균을 비교할 때는 이 표본 t검정을 사용하지 말고 일원배열 분산분석이나 크루스칼-월리스 검정을 사용하여야 한다. 또한 세 군이상에서 처리군 각각과 대조군의 평균을 비교할 때는 이 표본 t검정보다는 더넷검정(Dunnet, 1964)을 사용하는 것이 바람직하다.
- ③ 각 검정에서 요구되는 가정의 검토가 이루어지지 않았다. 정규분포의 검정이나 등분산검정이 이루어진 후에 t검정이나 일원배열 분산분석이 이

루어질 수 있다.

- ④ 자료요약에 Mean±SE나 Mean±SD가 논문간에 일관성이 없이 사용되고 있다.

그 이외에도 본 연구에서는 다음과 같은 문제를 추가적으로 지적하고자 한다.

- ⑤ 추정된 회귀식을 기술하는데 회귀계수의 표준오차, R² 등이 빠져있다.
 ⑥ 짝을 이룬 t검정(paired t-test)과 이 표본 t검정(two sample t-test, unpaired t-test)은 구별이 명시되어 있는데 일원배열 분산분석과 반복 관찰치가 있는 일원배열 분산분석은 구별이 되어 있지 않다. 실제로 반복 관찰치가 있는 일원배열 분산분석 자료유형은 9건이 있는데 그러한 기법이 명시된 논문은 한편도 없었다.
 ⑦ p값에 대한 기술이 잘못되어 있다. 몇 가지 잘못된 기술을 인용하면 다음과 같다.

“Probability was performed at 95% (p<0.05) and 99% (p<0.01).”

“The difference between treated groups and control groups were estimated at the levels of 95% (p<0.05) and 99% (p<0.01).”

“F-ratio간의 유의성을 p<0.05 및 p<0.01 에서 검정하였다.”

- ⑧ 검정력을 낮게 해줄 수 밖에 없는 작은 표본 규모 : 랫(rat)의 경우 그룹당 5마리~10마리 정도를 사용하고 있다.

(2) 용량-반응 분석의 통계적 분석 모형

의학논문에서는 임상실험결과를 분석하는 생존분석(survival analysis)이 주요한 분석도구라 할 수 있겠다. 독성학 관련 논문에서는 동물이나 미생물을 실험대상으로 사용하므로 각각의 용량마다 독립된 표본을 구성하여 사용할 수 있고, 또한 필요에 따라 부검도 실시할 수 있다. 그러므로 독성학 관련 자료분석에서는 용량에 따른 독립적인 반응을 분석하는 이른바 용량-반응 분석이 중요한 분석도구라 할 수 있겠다. [표 3-2]에서 알 수 있듯이 통계적 분석이 시도된 논문 60편중 24편이 용량-반응 분석을 다루고 있다.

한국독성학회지에서 논의된 용량-반응 분석은 크게 4가지 유형으로 나누어 볼 수 있는데 각각의 유형에서 통계적 분석모형을 요약하면 다음과 같다.

① 에임즈 셸모넬라 돌연변이 시험자료 분석

에임즈(Ames)셸모넬라 돌연변이(Ames et al, 1975)자료에 대한 통계적 분석모형으로는 크게 Margolin, Kaplan and Zeiger(1981)와 Bernstein et al(1982)이 있다. 이 두 모형은 모두 용량-반응 곡선에서 원점 근처의 기울기를 추정하여 그 기울기에 대한 검정을 실시하고 그 기울기가 0보다 유의적으로 클 경우 시험대상물은 돌연변이원으로 판명이 된다. 또한 그 기울기

의 역수를 역가(potency)로 사용하고 있다. 최근에 Margolin, Kim & Risko (1989)는 독성학자들이 흔히 사용하는 대조군 2배수 원칙을 통계적 모형과 비교하여 통계적 분석모형의 타당성을 보여 주었다.

② 용량에 따른 추세 검정 및 기울기의 추정

용량-반응 곡선에서 반응변수가 이분형일 경우 용량에 따른 추세 검정 (trend test)은 카크란-아미티지(Cochran-Armitage) 추세검정(Snedecor and Cochran, 1980)을 사용할 수 있고, 또한 피셔의 정확한 검정(Fisher's exact test)도 사용할 수 있다. 그리고 반응변수가 연속형일 경우 주어진 생물학적 이론을 배경으로하여 로지스틱 곡선과 같은 모수적 용량-반응 곡선(parametric dose-response curve)을 모형으로 사용하고, 추정된 모수의 유의성 검정을 통하여 용량과 관련된 추세(dose-related trend)를 밝혀낼 수 있다.

③ 두개 이상의 용량-반응 곡선의 비교

두개 이상의 용량-반응 곡선을 비교할 때는 모수적 용량-반응 곡선을 모형화한 후, 각 처리군에서 생성된 자료로써 용량-반응 곡선의 모수를 추정하여 그 추정치를 비교하는 것이 바람직한 방법이라 할 수 있겠다.

④ 동복자 효과가 있는 최기형(teratology)자료

동복자 효과란 같은 모태에 있는 태자들은 다른 모태에 있는 태자들보다는 서로 비슷한 반응을 보이는 추세를 말한다. 최기형 시험에서는 임신한 동물에게 관심대상의 화학물질을 투여하고 출산직전에 그 동물을 희생시켜 개개의 태자에 대하여 반응을 관찰한다. 이 경우 반응변수는 태자의 무게와 같이 연속형이 될 수도 있고, 어떤 기형의 발생여부처럼 이분형이 될 수도 있다. 반응변수가 연속형이고 정규분포에 근사할 경우는 분산분석의 내포된 임의효과모형(Searle, 1971)을 구성할 수 있다. 그러나 반응변수가 이분형일 경우 어떤 분석방법이 다른 여러가지 가능한 분석방법보다 더 우수한가에 대하여는 아직 알려지지 않고 있다. 단지 χ^2 동질성 검정이나 피셔의 정확한 검정(Fisher's exact test)은 적절한 방법이 아니라고 지적되고 있다(Kupper and Haseman, 1979). 이런 경우에 사용할 수 있는 분석방법으로는 0과 1사이의 값을 취하는 표본비율을 전환하여 분산분석을 실시할 수 있는 Arc Sine 전환, Freeman-Tukey 전환이 있을 수 있고, Mann-Whitney U test (MWU)나 Kruskal-Wallis 검정도 가능하다. 또한 용량의 증가에 따른 추세 검정으로는 비모수 추세검정법인 Jonckheere 검정도 적용이 가능하다.

라. 우리나라에서 독성학 분야와 통계학의 학제간 연구의 전망 :

향후 10년

3장의 가에서 지적된 문제점들은 독성학 관련논문에서만 지적되는 문제는 아니다. 송혜향(1989)교수의 관찰에 준거하여 보아도 이러한 문제는 향후 5년 후에는 많이 없어지리라 기대되고 동시에 독성학 관련분야에서도 기

초통계학에 대한 인식을 새롭게 하리라 기대된다.

그러나, 3장의 나에서 논의한 용량-반응분석의 문제는 주로 동물시험이나 미생물시험에서 제기되는 분석법이고, 독성학의 고유한 통계적 분석방법이라고 말할 수 있다. 이 분야는 통계학자의 전문적인 연구나 상담이 없이는 독성학자들이 쉽게 적용하기 어려운 분야이므로 앞으로 통계학자들의 꾸준한 참여가 요구되고 있다. 앞으로 신물질 개발이 활발히 이루어지면서 독성학분야의 발전이 이루어지면 이 분야에서 실험의 설계는 독성학자와 통계학자가 공동으로 하고 실험자료의 생성은 독성학자가 그리고 자료의 분석은 통계학자가 하는 이른바 학제간 연구가 바람직한 연구형태로 기대된다.

4. 농학

가. 서론

우리나라에서 근대적인 농업연구가 시작된 것은 현재 농촌진흥청의 전신인 권업모범장이 설립된 1906년부터라고 할 수 있다. 그러나 1960년대 이전에는 농사시험연구에서 체계적인 실험계획법과 결과에 대한 통계적 분석이 많이 이루어지지는 않은 듯하다. 그간의 문헌을 살펴보면 1930-1940년대에도 작물육종이나 재배법 개선 등의 실험에서 간단한 실험계획법이나 상관계수 분석 등의 이용 사례가 있으나 근대적인 실험계획법이 사용된 예는 찾아보기 힘들다.

농사시험 연구에서 보다 체계적인 기술개발이 이루어지고 이를 위한 시험연구 과정에서 통계적 방법이 사용된 것은 1962년 현재의 농촌진흥청과 산하에 시험장, 연구소가 확대 설치되면서부터라고 볼 수 있다. 특히 농사시험연구의 주종을 이루어 왔던 농작물의 품종개량에서는 1960년대 이후 완전확률화계획법, 난괴법 등의 실험과 분할구계획법등의 요인실험이 상당히 많이 사용되었고 특히 품종비교실험과 같은 처리가 많은 시험에서는 당시에 격자계획법을 사용한 예도 많다.

그러나 실제로 시험연구 결과를 분석하는 과정에서는 성능이 우수한 계산기를 필요로 하게 되는데 수동이나 전동계산기가 전부였던 60년대에는 계산과정상의 제약으로 인하여 통계적 방법의 적용에 한계가 있었으며 개발된 많은 실험계획법이나 분석방법이 보편화되지는 못하였다.

그러던중 1970년도에 농촌진흥청에 소형컴퓨터(FACOM 230/10)가 설치되었는데 이를 기점으로 우리나라 농사시험연구에 있어서 통계적 방법의 응용과 시험연구방법에 있어서 많은 발전을 가져오게 되었다.

1978년 전산시스템이 확장되고 연구소, 시험장 및 도 농촌진흥원에 단말기가 설치되면서 데이터 분석이 용이하여짐에 따라 사용되는 통계적 방법도 매우 다양하여지게 되고 이용자수도 급격히 증가하게 되었으며 이로 인한 농업기술개발과정에서의 효과는 양적으로 표현할 수 없으나 대단히 크다고

할 것이다.

나. 농업실험연구에서 통계적방법의 사용

농업연구를 크게 분류하면 농작물이나 가축의 품종개발, 재배법 및 사양 기술연구, 농작물이나 가축에 대한 환경연구 등으로 나눌 수 있는데 최근에는 기술 및 생산관리 체계의 변화로 각종 자동화 및 경영관리 기법에 대한 연구도 많은 비중을 차지하고 있다.

농업연구에서 많이 사용되는 통계적 방법을 파악하기 위하여 현재 농촌진흥청에서 개발 사용중인 「한국농업연구문헌정보(Korean Agriculture Research Information System : KARIS)」를 통하여 통계적 방법이나 응용에 관한 색인 용어를 갖고 있는 문헌을 검색한 결과 1960년부터 1991년까지 총 617건의 문헌이 검색되었는데 그 내용을 학회지 별로 보면 다음의 [표 4-1]과 같다.

[표 4-1] 학회지별 검색문헌 색인 건수

학회지명	건 수	학회지명	건 수
한국작물학회지	51	한국농공학회지	24
한국원예학회지	17	한국식물학회지	7
한국육종학회지	130	한국연초학회지	21
농사시험연구보고	57	한국임학회지	63
한국농업경제학회지	12	한국농업교육학회지	5
한국축산학회지	65	기 타	165

이 결과를 보아 농업연구에서 통계적 방법을 가장 많이 사용하는 것은 품종개량을 위한 육종, 작물재배, 축산 등임을 개략적으로 파악할 수 있는데 이는 품종개량에서 작물이나 가축의 유전적 연구를 위한 통계유전학 모델이 많이 사용되기 때문인 것으로 보인다.

농업연구에서 사용되는 통계적 기법을 보다 상세하게 파악하기 위하여 검색된 논문을 대상으로 사용된 통계적 기법의 빈도와 개략적인 내용을 조사하였는데 그 결과는 [표 4-2]와 같다.

이 결과에서도 보여주듯이 농업연구에서 가장 통계적 방법이 많이 보이는 것은 역시 유전통계분야라고 볼 수 있고 특히 80년대 후반에 와서는 컴퓨터의 보급과 함께 계산이 복잡한 다변량 해석이나 수리적 모형 또는 시뮬레이션 기법등의 응용도 점진적으로 증가하는 것을 볼 수 있다.

여기에서 조사된 논문들의 통계적인 기법의 개략적인 내용은 다음과 같다.

[표 4-2] Frequency of the statistical method in agriculture research

Statistical Method	61-65	66-70	71-75	76-80	81-85	86-90	Total
Descriptive Statistics	6	6	36	38	25	34	145
χ^2 tset	0	0	0	6	6	4	16
T-test	2	0	2	4	1	2	11
Experimental Design & ANOVA (balanced unbalanced)	3	6	18	25	25	16	93
Simple Regression & Correlation	3	4	13	12	11	17	56
Multiple Regression	0	2	13	12	22	25	74
Diallel cross (Jinks model)	0	0	3	10	21	14	48
Diallel cross (combining ability)	1	1	7	17	40	53	122
Genetic correlation & heritability	2	6	15	9	48	75	154
Path analysis	1	0	0	6	8	4	19
Selection index	1	2	0	1	3	20	27
Principle Component analysis	0	2	0	1	13	7	23
Cluster analysis	0	0	0	4	2	2	8
Mathematical Model	0	0	3	0	4	7	14
Simulation	0	1	2	3	4	15	25
others	2	2	13	3	4	12	36

(1) 기술통계 및 기초통계량의 계산

농업연구에서는 다른 분야에 비하여 데이터의 양이 비교적 많고 각종 관여인자들이 복잡하여 거의 모든 논문에서 간단하나마 통계적 방법을 적용한다. 기술통계와 관련하여 주로 사용되는 통계적 방법은 다음과 같다.

- 조사된 변수의 평균, 분산, 표준편차, 도수분포표 및 그래프작성
- 유전연구에서 질적 형질의 멘델식 유전분리비 검정을 위한 χ^2 검정

- 두가지 이상 질적 특성의 분리비에 관한 χ^2 독립성 검정
- 두개의 처리 또는 표본간의 차이 검정을 위한 t검정

(2) 실험계획법과 분산분석

실험계획법의 많은 부분이 농사시험에서 발견된 것과 같이 여러 종류의 실험계획법이 가장 다양하게 사용되는 것이 바로 농업분야인 것 같다. 농업연구에서 사용되는 실험계획법의 종류는 농업분야외에서도 대상재료와 내용에 따라 상당히 다양하며 거의 모든 연구의 기초가 되고 있다.

본 조사에서는 실험계획법의 사용빈도가 93건으로 조사되었으나 실제로 실험계획법은 모든 농업연구의 기본적 사항이므로 아마도 실제 실험계획법의 사용빈도는 이보다 훨씬 클 것으로 볼 수 있다. 추측컨대 농업연구의 각 분야에서 통계적 실험계획법이 사용되는 연구 프로젝트는 70% 이상이 될 것으로 생각되며 주로 사용되는 실험계획법은 다음과 같다.

- 작물품종육성 초기세대에서 사용되는 Augmented design(Federer 1961)
- 완전확률화계획법 : 각종 개체비교 및 특히 실내실험, 포트실험 등에서 사용
- 난괴법 : 품종비교, 지역적응시험, 각종 처리비교를 위한 포장시험
- 분할구계획법 : 2요인이상의 포장시험
- 세세구 계획법 : 3요인이상의 요인 포장실험
- 격자계획법 : 처리수가 많은 품종비교시험
- Hierarchical classification 계획법 : 축산의 교배실험에서 사용

이상의 여러가지 실험계획법의 빈도를 보기 위하여 농사시험연구중 주로 작물연구를 위한 포장실험에서 사용되는 실험계획법의 빈도를 한 등이 (1980) 조사한 것을 보면 다음의 [표 4-3]과 같다.

이 결과를 보면 포장실험에서 주로 사용되는 실험계획법은 난괴법으로 70.5%를 차지하고 있다. 또한 분할구계획법이 21.2%로 상당한 부분을 차지하고 있는데 이는 농업실험에서 포장배치나 처리의 관리상 편리한 분할구계획법의 장점때문인 것으로 생각된다.

그러나 실험수행과정에서 교락법이나 부분실시법의 사용에는 거의 없는 것으로 보인다. 이는 이러한 방법들이 처리나 인자가 많을 경우 상대적으로 실험의 효율성을 높일 수 있는 장점이 있음에도 불구하고 실험설계나 분석과정에서의 난점이 있기 때문인 것으로 생각된다.

한편 축산, 잠업 등의 동물을 대상으로 하는 연구에서는 표본수가 불균형인 예가 많다. 특히 가축의 육종이나 사양시험에서는 조사 개체수를 균형있게 맞추기가 어려워져 대부분의 경우 관측치가 불균형인 자료를 구성하게 되는데 이 경우 최소제곱에 의한 해를 구하는 방식으로 각 모수를 추정하고 있으며 이를 위하여는 현재 농촌진흥청의 통계분석 패키지(AGRISP)를 사용하고 있다.

[표 4-3] Type of experimental design of crop experiment (Hahn, 1980)

CROPS	C. R. D	R. C. B	latin square	split plot	S. S. plot	simple lattice	Triple lattice	others	계
Rice	17	882	15	333	62	19	18	9	1355
Wheat	1	108	-	20	2	1	1	3	136
Barley	1	197	2	32	3	1	-	2	238
Soybean	4	94	-	28	3	7	-	-	136
Sweet potato	-	33	-	4	3	4	-	-	44
Corn	-	60	-	6	2	-	-	-	68
Sesame	1	22	2	6	-	2	-	-	33
Peanut	-	19	-	3	1	2	2	-	27
Stevia	-	25	-	6	-	-	-	-	31
Rape	-	13	-	6	2	-	-	1	22
Vegetables	4	139	-	40	2	3	-	-	189
Fruit trees	2	150	-	34	-	-	-	-	185
Floriculture	3	15	-	15	-	-	-	-	33
others	-	57	-	12	3	-	4	-	76
Total	32	1814	19	545	83	39	25	15	2573
Percent	1.3	70.5	0.7	21.2	3.2	1.5	1.0	0.6	100

(3) 회귀상관 분석법

다른 분석에서와 마찬가지로 회귀 및 상관분석법이 가장 보편적으로 사용되는 통계적기법으로서 특히 컴퓨터의 보급과 함께 중형회귀 사용이 많아지고 있다.

단순회귀가 응용된 주요 예로는 작물의 생장곡선연구, 작물의 수량을 구성하는 각 형질들간의 관계, 작물의 환경요인과 수량과의 관계 등의 분석 등에서 사용된 예가 많다.

중형회귀분석은 작물연구에서 수량구성요소들과 수량과의 관계분석, 비료수준(N, P, K)과 작물의 수량과의 관계분석, 기상, 토양 등 환경요인과 작물의 수량과의 관계분석 등이 많고 임업연구에서 나무의 특성치에 의한 산림체적의 추정등의 예가 있다.

(4) 유전통계분석

농업연구분야에서 가장 많이 사용되는 통계적 기법중의 하나는 작물이나 가축의 품종개량을 위하여 기본적으로 필요한 유용한 형질들에 관한 유전적 특성과 이들 상호간에 관계분석을 위하여 사용되는 유전통계 방법이다. 특히 polygene의 영향을 받는 것으로 추정치는 대부분의 유전 양적 형질들은 계량적인 방법을 많이 쓰고 있으며 이러한 방법의 이론적인 근거와 응용의 기초를 제시한 사람이 Fisher라고 할 수 있다.

그 동안 유전통계분석을 위하여 많이 사용되고 있는 통계적 기법을 들면 다음과 같다.

- 세대평균분석법(Generation Mean Analysis)
- 완전확률화계획법이나 난괴법 배치에 의한 유전력 및 유전상관분석
- 이면교잡에 의한 조합능력검정(Diallel cross analysis ; Griffing, Gardner 모델)
- 이면교잡에 의한 유전자 분석(Diallel cross analysis ; Mather & Jinks 모델)
- 선별지수의 계산(Selection Index)
- 작물의 안정성 검정 분석(Stability parameters ; Eberhart 모델)
- 경로계수분석(Path analysis)

한편 축산연구에서는 유전실험계획법에서 cross factor design보다는 nested factor design을 사용하는 실험계획법을 통하여 유전모수를 추정하는 경우가 많다.

(5) 다변량해석법

우리나라 농사실험연구에서 다변량해석법이 사용된 것은 1970년대 컴퓨터가 도입되면서부터이며 주로 주성분 분석과 집락분석(cluster analysis) 등이 사용되었다.

주성분분석이 사용된 예로는 여러가지 조사전 특성을 기호로 한 품종의 분류 환경변수를 이용한 지역의 분류 등이 많으며 작물의 성장변화를 시계열로 사용하여 주성분분석을 한 예도 있다.

개체의 분류를 목적으로 하는 집락분석도 작물의 품종군 분류, 환경변수를 기호로한 환경 및 지역의 분류 등에 많이 사용되는데 특히 최근에는 전기 영동법을 사용한 식물체의 염기구조를 이용하여 작물의 분류를 한 예도 있다.

(6) Mathematical Model 및 Simulation

최근에 컴퓨터의 확장보급과 더불어 Mathematical Model과 Simulation을 이용하는 연구가 급증하고 있는 것 같다. 이러한 기법이 주로 응용되는 분야는 농업토목 및 농업기계 분야로서 주요 응용내용으로는 운동시물레이션,

농업기상에 있어서 강수량, 증발산량 등의 추정모델, 농용관계수의 저수지에 있어서 유입량과 방출량의 연구등이 있다.

다. 통계적 기법 사용상의 문제점

농업연구분야에서의 통계적 기법의 사용상의 문제점과 개선방안을 들어 보면 다음과 같다

(1) 현재 개선된 모형과 실제 응용 수준의 차이가 심하다

최근의 여러가지 통계적 모형들이 연구 개발되고 있음에도 불구하고 이들이 직접 연구과정에서 사용되는 예는 상당히 적다. 이는 실제로 개발된 모델을 연구원들이 직접 이해하여 응용한다는 자체가 대단히 어렵기 때문이다. 따라서 이를 위하여 통계 연구분야에서는 보다 실용적인 응용연구에 중점을 둘 필요가 있으며 연구된 내용에 대하여 사례연구를 각 분야의 연구원들과 공동으로 수행할 필요가 있다. 또 실제응용분야의 연구원들은 보다 적극적인 자세로 새로 개발된 모형의 이용법을 이해하고 사용할 수 있도록 노력해야 할 것이다.

(2) 연구 논문에서의 실험의 배경 및 통계적 모델의 명시가 미흡하다

연구논문에서 실험의 환경과 통계적 모델을 제시하지 않은 논문이 많다. 따라서 연구자는 자기가 사용하는 모델에 대한 정확한 개념을 갖지 못하는 경우가 있으며 이로 인하여 그 모델에 있어서의 가정의 검토를 무시한 채 분석하려는 경우가 많아 결과의 해석에 있어어도 오류를 범할 수가 있다. 이러한 점은 통계전문가가 연구계획의 수립이나 결과의 해석 및 평가시 참여하여 자문을 하여주고 문제점을 지적하여 개선하여 주는 방법으로 시정을 유도해야 할 것이다.

(3) 조사항목에 비하여 분석 항목이 미흡하다

실제로 연구계획서를 작성할 때는 많은 조사항목을 포함시켜 많은 노력을 들여서 연구를 하고 조사를 하였음에도 불구하고 분석내용은 극히 일부에 불과한 경우가 많다. 따라서 조사항목에 대한 최적의 내용을 계획있게 포함시키고 또 조사된 내용은 철저하게 분석할 수 있도록 하여야 할 것이다. 이를 위하여서는 보다 쉬운 통계분석 소프트웨어와 함께 컴퓨터 보급이 수반되어야 할 것이며 특히 연구원에 대한 통계교육이 통계학회를 중심으로 많이 이루어져 연구원들의 통계연수 참여 기회를 넓혀 주어야 할 것이다.

(4) 농업분야 응용통계연구조직 및 인원 미흡

현재 농업분야에서는 많은 통계적 방법이 사용되고 있음에도 불구하고 새로운 모델의 개발이나 개발된 모델의 응용을 연구하는 기구와 조직이 매우 미흡하다. 현재 농촌진흥청 산하연구기관과 농과계 대학에서 통계학의

응용을 연구하는 연구원이 소수이다. 이 정도로는 미흡하다. 따라서 통계학회내의 생물통계 연수회를 발전시켜 농업분야의 연구원이 보다 많이 참여할 수 있도록 하고 통계학회에서도 보다 많은 지원을 하여야 할 것이다. 이를 위하여 연구기관과 통계학과간의 공동연구 프로젝트 설정도 가능한 하나의 과제이며 통계학자들의 농업 및 생물통계에 대한 관심도가 높아져야 할 것이다.

참 고 문 헌

1. 김병수, 안운기, 윤기중(1987). 통계의 오용과 효율적 이용에 관한 연구, 산업과 경영, 24권, 3-37.
2. 경제기획원 조사통계국(1989). 한국통계조사현황, 경제기획원.
3. 백운봉(1991). 이제는 통계학의 실천에 눈을 돌릴 때, 한국통계학회 춘계 학술논문 발표회 발표논문.
4. 송혜향(1989). 병원통계자문의 변천과정과 바람직한 방향, 한국통계학회 추계 학술논문 발표회 발표논문.
5. 송혜향, 박용규(1988). 우리나라 투석치료 및 신이식환자 등록부의 시계열적 자료현황, 한국보건통계학회지 제13권, 1-17.
6. 한국의과대학장 협의회(1990). 의과대학교육현황(1990-1991년도), 제6집, 계축문화사.
7. 한국인구보건 연구원(1987). 제5장, 의료보 제도와 지불보상제도, 의료자원과 관리체계에 관한 조사, 보건사회부에 제출한 최종보고서, 445-536.
8. Ames, B.N., McCann, J. and Yamasaki, E.(1975). Methods for detecting carcinogens and mutagens with the Salmonella/microsome mutagenicity test, *Mutation Research*, Vol.31, 347-364.
9. Bernstein, L., Kaldor, J., McCann, J.M. and Pike, M.C.(1982). An empirical approach to the statistical analysis of mutagenesis data from the Salmonella test, *Mutation Research*, Vol.97, 267-281.
10. Dunnet, C.W.(1964). New tables for multiple comparisons with a control, *Biometrics*, Vol.20, 482-491.
11. Hokanson, J.A., Stiernberg, C.M., McCracken, M.S. and Quinn, F.B. (1987). The reporting of statistical techniques in otolaryngology, *Journals. Arch. Otolaryngol. Head Neck Surg*, Vol.113, 45-50.
12. Haseman, J.(1984). Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies, *Environmental Health Perspectives*, Vol.58, 385-392.
13. Haseman, J.K. and Kupper, L.L.(1979). Analysis of dichotomous

- response data from certain toxicological experiments, *Biometrics*, Vol. 35, 281-293.
14. Haseman, J.K., Zeiger, E., Shelby, M.K., Margolin, B.H. and Tennant, R.W. (1990). Predicting rodent carcinogenicity from four in vitro genetic toxicity assays: An evaluation of 114 chemicals studied by the National Toxicology Program, *Journal of the American Statistical Association*, Vol. 85, 964-971.
 15. Hothorn, L. (ed.) (1991). *Statistical Methodology in Toxicology*, Lecture Notes in Medical Informations, Springer Publications, Vol. 43, Heidelberg.
 16. Krewski, D. and Brown, C. (1981). Carcinogenic risk assessment: A guide to the literature, *Biometrics*, Vol. 37, 353-366.
 17. Kupper, L.L., Portier, C., Hogan, M. and Yamamoto, E. (1986). The impact of litter effect on dose-response modeling in teratology, *Biometrics*, Vol. 42, 85-98.
 18. Margolin, B.H., Kaplan, N. and Zeiger, E. (1981). Statistical analysis of the Ames Salmonella/microsome test, *Proceedings of the National Academy of Science*, Vol. 78, 3779-3783.
 19. Margolin, B.H., Kim B.S. and Risko, K.J. (1989). The Ames Salmonella/microsome mutagenicity assay: Issues of inference and validation, *Journal of the American Statistical Association*, Vol. 84, 651-661.
 20. McKnight, B., and Crowley, J. (1984). Tests for differences tumor incidence based on animal carcinogenesis experiments, *Journal of the American Statistical Association*, Vol. 79, 639-648.
 21. Piegorsch, W.W. and Haseman, J.K. (1991). Per-litter analyses for studies of developmental toxicity, *Statistical Methods in Toxicology* (L. Hothorn, Ed.), Lecture Note in Medical Informatics, Springer Publication, Heidelberg, Vol. 43, 86-95.
 22. Portier, C. (1986). Estimating the tumor onset distribution in animal carcinogenesis experiments, *Biometrika*, Vol. 73, 371-378.
 23. Portier, C.J. and Dinse, G.E. (1987). Semiparametric analysis of tumor incidence rates in survival/sacrifice experiments, *Biometrics*, Vol. 43, 107-114.
 24. Searle, S.R. (1971). *Linear Models*, Wiley, N.Y.
 25. Snedecor, G.W. and Cochran, W.G. *Statistical Methods*, 7th Ed., Iowa State University Press, Iowa.

26. Soper, K.A. and Clark, R.L.(1991). Data-based scores in exact trend tests for fetal survival, *Proceedings of the International Biostatistics Conference in the Study of Toxicology*, Tokyo, 17-22.
27. Articles from the symposium on statistics and the environment, *Statistical Science*, Vol.3, No.3.
28. Tennant, R. W., Margolin, B.H., Shelby, M.D., Zeiger, E., Haseman, J.K., Spalding, J., Caspary, W., Resnick, M., Stasiewicz, S., Anderson, B. and Minor, R.(1987). Prediction of chemical carcinogenicity in rodents from in vitro genetic toxicity assays, *Science*, Vol.236, 933-941.
29. White, S.(1979). Statistical errors in papers in the British Journal of Psychiatry, *British Journal of Psychiatry*, Vol.135, 336-342.
30. Yamamoto, E. and Yanagimoto, T.(1988). Litter effects to dose-response curve estimation, *Journal of Japan Statistical Society*, Vol.18, 97-106.