대한이식학회지 : 제 4 권 제 1 호
Vol. 4, No. 1, November, 1990

퍼스널 컴퓨터를 이용한 의학 통계(I)

연세대학교 의과대학 외과학교실

이우정 · 김순일 · 김유선 · 박기일

= **Abstract** =

Basic Statistics For Medical Researches Using Personal Computer

—Part I: Special Emphasis on the T-test, Chi-square test, ANOVA and other nonparametric tests—

Woo Jung Lee, M.D., **Soon Il Kim**, M.D., **Yu Seun Kim**, M.D. and **Kiil, Park**, M.D.

Department of Surgery, Yonsei University College of Medicine

Basic understanding and computerized processing for T-test, Chi-square test, ANOVA and other nonparametric tests are described in this part I for basic and clinical investigators in the field of Medical Researches which have been frequently used in our department of Surgery since 1985.

서 론

대학에서 처음 시작한 통계학 공부는 이론위주이었으며, 실제적으로 임상의학 연구에 적용을 시도할 경우 계산의 복잡성과 난해한 용어 등으로 어려움이 많았고, 이의 해결을 위해 컴퓨터의 도움을 받으려고 대학교나 연구소에 있는 대형컴퓨터에 대해 관심을 갖는 경우도 많으나 이 경우도 실제로 대형컴퓨터 사용방법의 복잡성과 컴퓨터언어에 대한 무지 및 절차상의 문제점 등으로 쉽게 포기하는 경우가 왕왕 있어 왔다. 저자는(W.J.L) 1982년부터 애플컴퓨터를 이용하여 의학통계처리를 시도하여 왔으며 통계에 대해서도 관심을 가지게 되었다. 애플컴퓨터에서 가능한 프로그램(ABSTAT)은 초보단계이기는 하여도 간단한 T-test, Chi-square test, ANOVA, 그리고 간단한 Regression 처리까지는 가능하였으나 점차 이 프로그램의 한계를 알게되었고 좀 더 낳은 프로그램의 필요성을 절감하게 되었다. 최근들어 16 bit 컴퓨터의 보급이 대중화되고, 퍼스널 컴퓨터에서 이용가능한 통계 package인 SPSS[1], SAS[2], BMDP[3] 등이 소개되면서 임상의학 연구에 꼭 필요한 Survival Analysis 및 그 검정과, Cox Regression의 처리도 가능하게 되었다.

의학연구 분야도 이제는 어떤 결론을 유도할 때 적절한 통계적 검증을 거치고 않고는 의의있는 결과를 얻을 수 없게 되었으므로 의학 연구자에게도 필수적으로 통계의 이해가 필요하다고 생각되어 저자는 연세대학교 의과대학 외과학교실에서 주로 사용하는 통계기법 및 컴퓨터 사용기법에 대해 간단히 소개하고자 한다.

이 글은 지면 관계상 I, II, III 세편으로 나누어지며 본 I편에서는 우선 통계에 대해 전혀 모르는 사람을 대상으로 간단한 통계용어를 먼저 설명하여 이해를 돕고, 흔히 사용되는 T-test, Chi-square test, ANOVA(분산분석), 기타 비모수 검정에 대해서 설명을 하겠다. II편에서는 Survival Analysis(생존분석) 중 생명표와 생존함수에 대해서 즉 의학에서 흔히 사용되는 5년생존율이나 평균생존기간 그리고 각기 다른 집단간의 생존율의 차이의 유의성 검정에 대해서 설명하겠다. III편에서는 상관분석(correlation analysis), 회귀분석(regression analysis), 다변량분석(multivariate analysis) 및 Cox Regression에 대해서 설명하고, 실제로 퍼스널 컴퓨터의 이용은 각기 통계처리 예제에서 조금씩 설명하겠다.

저자는 통계전문가가 아닌 의학연구자이므로 본 논문의 내용을 가능한한 실제적으로 흔히 의학연구에서 사용되는 범위에 국한시켰으며 좀 더 자세한 분야는 통계학 전문가에게 상의하는 것이 마땅할 것으로 사료된다.

본 론

1. 용어 설명 및 통계방법의 선택

일반적으로 통계는 다음과 같이 분류될 수 있다. 이 분류는 약간은 무리가 있으나 앞으로 설명할 통계의 이해에 도움을 주기위해 다음의 표로 정리 하면,

표 1. 통계의 분류

- 기술통계학 (descriptive statistics)
- 추측통계학 (inferential statistics)
  - 모수치 추정
  - 가설검정
    - 비모수 (nonparametric) 검정
    - 모수 (parametric) 검정

기술통계란 어떤 모집단 또는 표본으로부터 수집한 수량적 혹은 분류된 자료들을 요약, 정리하는 통계적 방법을 말한다. 여기에는 대표값, 분산 등이 있으며 컴퓨터 package (예 : SPSS)에서는 단일변량분석을 사용해서 결과를 얻을 수 있다. 그러나 표본에서 얻은 자료를 분석해서 얻은 결과를 모집단의 그것으로 일반화하는 데에는 항상 문제점이 따르게 된다. 왜냐하면 표본으로부터 얻어진 자료에는 항상 오차가 게재되기 때문이다. 따라서 표본에서 얻어진 수치를 가지고 전체집단(모집단)의 특성을 추리하는 통계가 추측통계이다. 다시말하면, 어떤 실험이나 환자집단에서 나온 자료를 정리하면 그것이 기술통계이고 그 값을 기초로 어떤 결론(가설)을 추측하거나 차이나 관련성을 추정하는 방법이 추측통계이다. 우리가 실제로 의학에서 가장 많이 사용하는 통계가 바로 추측통계에서의 가설검정이다 (표 1).

여러가지 복잡한 통계용어가 있으나 초보자에게는 "변수"란 용어의 정확히 개념을 잡기가 어려우며 "변수"에 대해서 정확히 모르고는 앞으로 전개할 통계에 어려움이 많이 따르므로, 이 글에서는 먼저 "변수"에 대해서 설명하고, 그 다음 통계처리에 대해서는 간단한 도표 하나를 들어 어떤 경우에 어떤 통계를 이용해야 되는지를 설명하겠다.

"변수"란 영어로 Variables, 한자로는 **變數**라고 하는 것처럼 같은 차원의 것으로 두가지 이상의 값을 갖는 척도의 기준이다. 어렵게 생각할 것 없이 모든 것이 변수가 될 수 있다. 하지만 통계에서는 다음 2가지를 확실히 구별할 수가 있어야 하며 그 구별을 근거로 어떤 통계를 사용할 것인가도 알 수 있게 된다. 2가지 구별은 다음과 같다.

1) 독립변수와 종속변수 : 독립변수를 independent variable, 종속변수를 dependent variable이라고 하는데 우리말로 독립변수를 '설명변수' 또는 '원인변수'라

표 2. 측정수준에 따른 변수의 분류

- 명목변수 (nominal variable) : 조사대상의 고유한 특성을 숫자나 기호로 표시한 것으로서 예를 들면, 성별, 인종, original disease 종류, ABO 적합성 유무, ATN 유무, 간염항원 보균 유무, 수혈유무, 투석요법의 종류, 면역억제제의 종류 등
- 순위변수 (ordinal variable) : 많고 적음에 따라 인위적으로 순서를 정하는 것으로 예를 들면, 교육수준, 생활수준, HLA mathing 정도, 거부반응회수, 수혈양의 정도등이며 상기명목변수와 순위변수는 가감승제를 할수없다.
- 간격변수 (interval variable) : 비교서열간에 일정한 간격을 두고 수치값을 부여하는 것을 말하며 예를 들면, 체온, 나이, 투석요법의 기간, 면역억제제 투여량등이 해당된다.
- 비율변수 (ratio variable) : 일정한 간격을 가지고 절대 영점이 있을 때를 말하며 예를 들면, 체중, 신장, RBC수, WBC 수 등이 있다.

고, 종속변수를 결과변수로 하면 무난하다. 예를 들면, 신장이식 환자에서 HLA matching의 정도에 따른 조기거부반응의 빈도를 설명하고자 하면 matching의 정도는 독립변수(설명변수)이고 조기거부반응의 빈도는 종속변수(결과변수)가 된다.

2) 측정수준에 의한 분류 : 변수는 측정수준에 따라 다음 4가지로 분류할 수 있다 (표 2).

3) 통계방법의 선택(모수통계와 비모수통계) : 통계방법의 선택은 얻어진 자료(변수)가 측정수준에 의한 분류로 어디에 속하느냐에 달려있다고 하겠다. 그러나 이에 앞서 먼저 고려하여야 할 것이 있는데, 그것은 처리하고자 하는 자료의 모집단이 다음 세가지의 조건에 합당한가 하는 문제이다. 즉 모집단이 정규분포이어야 하고, 집단내의 분산이 같아야 하고, 변수는 최소한 등간척도로 되어야 하는 것이다. 이러한 조건이 구비되면 모수(parametric) 통계로 처리할 수 있다. 그러나 이와는 달리 비모수검정방법은 모집단의 정규 분포를 전제하지 않으므로 분포를 알 수 없을 때(특히 표본의 크기가 적을 때 : <30)와 명목변수, 순위변수에 의하여 측정되었을 때 할 수 있다.

이를 요약하면 얻어진 자료가 명목변수이면 Chi-square 검정으로 제한적인 것이 보통이며, 얻어진 자료가 순위변수이면 일반적으로 비모수검정으로 제한된다. 그러나 자료가 간격이나, 비율변수이면 모수와 비모수검정을 다 택할 수 있는데 선택의 기준은 위에서 설명한 것과 같이 자료의 크기와 정규분포의 여부에 달려 있다[5].

다음 표 3과 표 4에 비모수통계와 모수통계를 간단히 정리하여 보았다.

물론 위 표로 설명되지 않는 것도 있다. 특히 Survival Analysis는 환자의 추적관찰이 100%이라면 모르겠으나 실제 임상에서는 환자가 다른 질환으로 사망하거나 또는 추적이 어떤 사정에 의해 중단되는 경우가 있기 때문에 보통의 통계방법이 아닌 이러한 예외의 자료처리가 가능한 방법들이 개발되어 왔다. 이에 대해서는 다음의 II, III편 글에서 설명하도록 하겠다 (표 3, 4, 5).

2. 통계처리 실제

1) 가설 : 원래 영가설(null hypothesis, Ho)은 통계

표 3. 비모수검정으로 실행 가능한 대표적 방법들

자료의 종류	명목변수	순위변수
단일표본	Chi-square test	Kolmogrov-Smirnov
2짝표본(pair)	McNemar test	Wilcoxon signed rank
K짝표본(pair)	Cochran's Q test	Friedman, Kendall
2독립표본	Chi-square test Fisher exact test	Mann-Whitney (Wilcoxon rank sum test) Kolmogorov-Smirnov
K독립표본	Chi-square test	Kruskal-Wallis
상관관계		Spearman's rho

표 4. 모수검정으로 실행가능한 대표적인 방법들

종속변수 / 독립변수	간격 및 비율변수 (continuous)	명목 및 순위변수 (categorical)
간격 및 비율변수 (continuous)	Correlation (상관분석) Regression (회귀분석) Factor Analysis (인자분석)	Discriminant analysis (판별분석) Logistic analysis (로지스틱분석)
명목 및 순위변수 (categorical)	T-test ANOVA ANCOVA	Chi-square test

표 5. 상응하는 모수검정 및 비모수검정

자료의 종류	모수 검정	비모수 검정
2 independent samples	Unpaired T-test	Mann-Whitney test
2 related samples	Paired T-test	Wilcoxon signed rank test
K-related samples (독립변수 1개)	One way ANOVA	Kruskal-Wallis test
K-independent samples (독립변수 2개)	Two way ANOVA	Friedman test

적으로 나타난 차이가 관계가 우연의 법칙으로 인해 생긴 것이라는 뜻으로 수학적으로는 $u_1 = u_2$로 표시된다. 그러나 많은 의학연구에서는 표본들의 차이가 우연히 발생하는 것이 아니라, 그 두 개의 표본이 대표하는 모집단 평균치 사이에는 분명한 차이가 있다는 것의 진실여부를 알아보는데 있다. 즉 연구의 목적은 영가설을 부정하고 이론적인 대안을 받아들이는 데 있는 것이다. 이렇게 밝히고자 하는 대안을 대립가설(alternative hypothesis, H_1) 또는 연구가설이라 하고, 영가설은 귀무가설이라 한다. 이러한 이야기가 잘 이해가 않되는 사람을 위해 다음과 같이 다시 한번 이야기 하면, 연구를 하는 사람이 증명하고저 하는 의도가 바로 대립가설 또는 연구가설이고 이것을 증명하기 위해 이론적으로 만들어 부정을 함으로써 대립가설을 받아들이고저 하는데 이 이론적인 가설이 바로 영가설 또는 귀무가설이라고 할 수 있다.

2) 유의수준 : 흔히 유의수준은 확률(probability)을 나타내는 알파벳 소문자 p로 표시하는데 이는 연구자가 영가설을 부정하기 위한 기준으로 설정한 **확률수준**으로 정의내려지는데, 이는 영가설을 받아들일 확률이 이러한 유의수준 이하로 떨어질 때 영가설은 부정되고 연구가설이 받아들여지는 것이다. 통계계산상 p값이 0.05로 나왔다는 것은 두 표본이 같은 모집단의 일부일 가능성(확률)이 0.05 즉 5%라는 이야기이다.

위에 이야기 한 것처럼 유의수준을 설정하여 어떤 가설을 검정하고저 할 때 염두에 두어야 할 것이 있는데 이는 제 1 종 오차의 제 2 종 오차이다. 예를 들면 혈액형이 같은 경우와 혈액형이 다른 경우의 신장이식을 시행한 경우 거부반응의 횟수를 가지고 검증을 해서 만일 p값이 0.04로 나왔다고 하면 이는 p값이 설정한 유의수준인 0.05보다 적으므로 영가설을 부정하고 대립가설(혈액형이 같은 경우의 신장이식이 거부반응이 적다)을 받아들이게 되는데 여기에는 다음과 같은 가능성이 있다. 즉 실제로는 두 집단간에 차이가 없는데 우리가 유의수준을 0.05로 설정하므로써 차이가 있다고 결론을 내릴 가능성인 것이다. 위와 같이 실제로는 같은 모집단인데 다르다고 하는 오류를 제 1 종 오차라고 한다. 이와는 반대로 실제로는 다른 모집단인데 같다고 결론을 내릴 가능성이 제 2 종 오차인 것이다.

보통 유의 수준이 0.05나 0.01인 경우에 가장 제 1 종 오차의 제 2 종 오차가 적은 것으로 알려져 있다. 그렇다고 모든 통계검정을 이 유의수준만 가지고 하는 것은 아니다. 이러한 유의수준의 결정은 연구자가 어떠한 오류를 감수할 것이냐에 달려 있다. 예를 들어 현대의학의 큰 과제인 암에 대한 새로운 치료법을 실험하고자 할 때는 효과가 없다는 영가설의 유의수준을 0.2 혹은 0.1로 하여 연구가설을 받아들인다 하여도 그것은 의미가 있을 것이다. 반대로 이미 많이 나와있는 비슷비슷한 감기약의 효능을 검사하는 연구에서는 유의수준을 0.001 혹은 0.0001로 정할 수도 있을 것이다.

3) T-test : 위의 표에서 보듯이 T-test는 독립변수가 Categorical variable이고 종속변수는 Continous variable 경우에 두 집단의 차이를 검정하는 데 사용되는 모수통계방법으로 Chi-square test와 더불어 의학통계에서 가장 많이 이용되고 있다[4]. 비교적 우리가 흔히 쓰는 통계방법이지만 짚고 넘어갈 것이 있다. 독립적(independent or unpaired) T-test와 짝비교(paired) t-test는 둘 다 T-test이지만 표본의 성격상 개념이 다르다. 예를 들면 같은 사람에게서 어떤 약제의 투여 전과 투여 후의 맥박수를 각각 측정하였다면, 이는 실제로 맥박수가 변수가 되는 것이 아니고 투여후와 투여전의 맥박수의 차이가 변수가 되는데 이런 경우처럼 측정치가 서로 쌍(pair)를 이루거나 대응을 이루는 경우에 쓰는 통계방법이 짝비교 T-test (paired T-test)이고, 그 외에 표본수가 다르거나 혹 같은 표본 수라도 표본이 서로 다른 집단인 경우에는 독립적 T-test를 사용하면 된다.

독립적 T-test에 상응하는 비모수 검정법은 중앙값검정법(median test) 또는 Mann-Whitney (Wilcoxon Rank Sum test) 검정이고, 짝비교 T-test에 상응하는 비모수 검정법은 Wilcoxon singed rank 검정으로 이들 모두는 SPSS/PC+에서 가능하다.

□ SPSS/PC+ 이용한 실제 □

a) 독립적 검정(independent T-test) **예**

신장이식후 CsA를 사용한 A group과 기존의 면역억제제를 사용한 B group 사이에 신장이식후 3개월 지나서 평균동맥혈압을 측정하여 약제의 부작용을 비교해 보고자 할 경우

A group : 77, 91, 101, 81, 76, 68, 72, 82 mmHg

B group : 72, 62, 51, 81, 47, 74, 52, 65 mmHg

<SPSS/PC+ 입력>

```
DATA LIST FREE/GR DRUG.
BEGIN DATA.
1 77 2 72 1 91 2 62
1 101 2 51 1 81 2 81
1 76 2 47 1 68 2 74
1 72 2 52 1 87 2 65
END DATA.
T-TEST GROUPS=(1,2)/VAR=DRUG.
```

<결 과>

SPSS/PC+

Independent samples of GR

Group 1: GR EQ 1.00 Group 2: GR EQ 2.00

t-test for DRUG

	Number of Cases	Mean	Standard Deviation	Standard Error
Group 1	8	81.0000	10.637	3.761
Group 2	8	63.0000	12.259	4.334

	Pooled Variance Estimate				Separate Variance Estimate		
F Value	2-Tail Prob.	t value	Degrees of Freedom	2-Tail Prob.	t Value	Degrees of Freedom	2-Tail Prob.
1.33	.717	3.14	14	.007	3.14	13.73	.007

<결과 해석>

F 값은 두 모분산의 가설을 검정한 것으로 같다는 것을 알 수 있다. pooled variance of t-test는 두 그룹의 모분산이 같고 풀링하여 추정할 때 쓴다. 일반으로 separate의 값을 보면 무난하다. 단측검정에 대한 명령은 없다. p값이 0.05미만이므로 평균값의 유의한 차이가 있다고 볼 수 있다. 즉 기존의 면역억제제를 사용하는 경우보다 CsA를 사용하는 경우 혈압이 높다고 말할 수 있다. 위의 예제는 사실상 샘플의 수가 적기 때문에 비모수 검정을 해야 하나, 예제이기 때문에 모수 검정법을 사용했다. 위의 예를 가지고 비모수 검정(Mann-Whitney test)을 시행해 보면 다음과 같다.

<입 력>

```
NPAR TESTS M-W=DRUG BY GR(1,2).
```

<결 과>

SPSS/PC+

Mann-Whitney U - Wilcoxon Rank Sum W Test

DRUG

by GR

Mean Rank	Cases	
11.50	8	GR = 1.00
5.50	8	GR = 2.00
	16	Total

		EXACT	Corrected for Ties	
U	W	2-tailed P	T	2-tailed P
8.0	92.0	.0104	-2.5242	.0116

<결과 해석>

P=0.0104(p<0.05)로 평균값의 유의한 차이가 있다고 할 수 있다.

b) 짝비교 검정(paired T-test) **예**

신장 이식후 6명의 환자에게 이뇨제(Lasix)를 주기전과 주고나서의 일일 소변량의 비교를 하고자 한다.

주기전 : 4.3, 5.6, 3.9, 5.7, 4.8, 5.2 liter/day

준 후 : 5.7, 5.4, 4.6, 7.3, 6.0, 5.3 liter/day

<SPSS/PC+ 입력>

```
T-TEST PAIRS=BEFORE, AFTER
```

<결 과>

SPSS/PC+

Paired samples t-test: BEFORE

AFTER

Variable	Number of Cases	Mean	Standard Deviation	Standard Error
BEFORE	6	4.9167	.719	.294
AFTER	6	5.7167	.906	.370

(Difference) Mean	Standard Deviation	Standard Error	Corr.	2-Tail Prob.	Degrees of Freedom	2-Tail Prob.
-.8000	.729	.298	.619	-2.69	5	.043

<결과 해석>

p=0.043(<0.05)으로 평균값이 유의한 차이가 있다고 볼 수 있다.

위의 예를 가지고 비모수검정(Wilcoxon paired-sapmple test)를 해보면 다음과 같다.

<SPSS/PC+ 입력>

```
NPAR TEST    WILCOXON=BEFORE WITH AFTER.
```

<결 과>

```
      SPSS/PC+

------Wilcoxon Matched-pairs Signed-ranks Test

     BEFORE
with  AFTER

Mean Rank     Cases

   2.00         1    - Ranks  (AFTER LT BEFORE)
   3.80         5    + Ranks  (AFTER GT BEFORE)
                0      Ties .  (AFTER EQ BEFORE)
               ---
                6      Total

    Z =  -1.7821      2-tailed P=  .0747
```

<결과 해석>

p=0.0747(>0.05)로 유의한 차이가 없는 것으로 나타났다. 일반적으로 비모수 검정은 검출력이 떨어진다.

4) ANOVA (Analysis of Variance)

T-test는 위에서 보듯이 두 집단의 평균간의 차이를 검정하는데 사용하지만 ANOVA는 세개 이상의 모집단 평균간의 차이를 검정하는데 사용되는 통계기법이다. 즉 독립변수가 명목변수이면서 그 명목(표본)이 3개 이상인 경우에 사용하는 모수통계방법이다. 만일 독립변수가 1개이면 단일변량분석(Oneway ANOVA)을 하게되고 독립변수가 2개 이상인 경우에는 다원변량분석(Multiple ANOVA)을 하게되는 것이다. 물론 짝비교 t-test(paired t-test)처럼 반복측정 분산분석도 가능하다. 본 글에서는 단일 변량분석만 이야기 하겠다. 한가지 더 이야기할 것은 이러한 ANOVA검사가 나타내는 것은 세집단(만일 세 집단인 경우)의 차이만을 나타내지 각각 2집단의 차이가 유의한가는 알 수 없다는 것이다. 그러나 연구자의 관심에 따라 이런 각각 2집단의 차이가 있는가를 알아볼 수도 있는데, 이러한 경우에 사용되는 통계방법을 '추후검증' 또는 '사후검증'이라 한다. 이러한 검증에 자주 사용되는 통계기법은 Tukey의 다중비교법, Scheffe의 검증법, Duncan의 다중범주비교법, Newman-Keuls(or Hartly)의 검증법 등인데 여기서는 자세히 설명하지 않고 밑의 SPSS/PC+ 이용실제에서 Newman-Keuls의 경우만 예를 들겠다.

단일변량분석(ANOVA)에 상응하는 비모수검정법은 Kruskal-Wallis 검정이고 이원변량분석(Two way ANOVA)에 상응하는 비모수검정법은 Friedman 검정으로 이들 모두는 SPSS/PC+에서 가능하다.

□ SPSS/PC+ 이용한 One way ANOVA 실제 □

세집단에서 약물분포 용적의 비교를 하고자 한다.

집단 1 : 465, 293, 371, 405, 451 mg/liter
집단 2 : 291, 225, 287, 302, 210 mg/liter
집단 3 : 192, 212, 270, 251, 220 mg/liter

<SPSS/PC+ 입력>

```
DATA LIST FREE/GR VD.
BEGIN DATA.
1 465 2 291 3 192
1 293 2 225 3 212
1 371 2 287 3 270
1 405 2 302 3 251
1 415 2 210 3 200
END DATA.
ONEWAY VAR=VD BY GROUP(1,3)
/RANGES=SNK
/STAT=ALL
```

<결 과>

```
              SPSS/PC+
      --------------------O N E W A Y ---------------
   Variable  VD
 By Variable  Group

                          Analysis of Variance
                          Sum of           Mean        F       F
           Source   D.F.    Squares         Squares   Ratio    Prob.
 Between Groups       2    70120.0000     35060.0000  12.8292  .0010
 Within  Groups      12    32794.0000      2732.8333
 Total               14   102914.0000
```

```
              SPSS/PC+
      --------------------O N E W A Y ---------------
   Variable  VD
 (Continued)
                         G  G  G
                         r  r  r
                         p  p  p

      Mean      Group    3  2  1
    243.0000    Grp 3
    263.0000    Grp 2
    397.0000    Grp 1    *  *
```

<해 석>

세집단의 비교를 위해 oneway ANOVA를 시행하여 여러 결과 중 일부를 실었다. 결과 표중 위의 것을 보면 F 값이 0.001로 유의수준 0.05로 설정하면($p<0.05$) 영가설을 기각할 수 있다. 즉 세 집단간의 약물분포 용적에는 차이가 있다고 할 수 있다. 그리고 추후검정으로 집단 상호간의 비교를 위한 다중비교는 student-Newman-Keulis test를 시행하였으며 집단간의 차이가 있으면 보기 좋게 *표로 표시되는데, 이 결과로 집단 1 집단 2, 그리고 집단 1과 집단 3간에는 차이가 있는 것으로 나타났다. 즉 집단 2와 집단 3간에는 유의한 차이가 없음을 알 수 있다. 이 경우도 샘플의 수가 적은 데도 설명을 하기 위하여 모수 검정법을 사용했음을 이해하기 바란다.

위의 예를 가지고 비모수 검정(Kruskal-Wallis test)을 시행해 보면 다음과 같다.

<SPSS/PC+ 입력>

```
NPAR TESTS  K-W=VD BY GROUP(1,3).
```

<결 과>

```
   ------Kruskal-Wallis  1-way ANOVA

    VD
 by GROUP

   Mean Rank     Cases
     12.80          5      GROUP  =  1
      6.60          5      GROUP  =  2
      4.60          5      GROUP  =  3
                   ---
                   15      Total
                                           Corrected  for Ties
   CASES   Chi-Square   Significance   Chi-Square    Significance
     15     9.1400          .0104        9.1400          .0104
```

<결과 해석>

p=0.0104(<0.05)로 세집단간에서 유의한 차이가 있음을 보여주고 있다.

5) Chi-square test

변수의 측정값이 어떤 특성이 있는가 또는 없는가로 측정되는 통계자료나 또는 측정 값이 어떤 범위에 들어가는지 않들어가는지로 구분되는 통계자료로부터 모집단값에 대한 가설검정은 Chi-square test를 이용할 수 있다[5]. Chi-square 검정법은 독립변수와 종속변수가 다 Categorical variable인 경우에 사용하는 검정으로, 모수검정은 물론 비모수 검정법으로도 많이 쓰인다. 가령 성별이라는 변수는 남녀 두 집단으로 나누어 질 수 있고, 이러한 집단들에 빈도가 주어졌을 때 그 빈도 간에 유의적 차이가 있나 없나를 알아보고자 하는 경우에 쓸 수 있다. Chi-square 검정은 이와같이 기대치와 관찰치의 차이를 검증하는 방법으로 적합도검정(Goodness of fit test)이라고도 한다.

Chi-square 검정법에는 특히 독립적으로 추출된 두 표본의 2×2 검정인 경우 이산분포에 근접시키는 방법으로 Yate's Correction법을 사용할 수 있고, 샘플의 수가 20<n<40이면서 기대값이 5이하일 때는 Fisher exact probability test[5]가 더 유용하다. SPSS에서는 Crosstab이라는 명령어로 같이 처리가 되는데, Corrected Chi-square 값을 함께 계산하고, 샘플의 수가 30 이하인 경우 Chi-square 값에 Fisher's exact test 값을 계산하여 나타낸다.

Chi-square 검정법중, 한 집단내의 부속집단간의 빈도의 차이를 검정하고자 할 때 쓰이는 것이 '단일표본의 Chi-square검정'이고, 변수의 두 개 혹은 그 이상의 집단간의 독립성을 검정하는데 쓰이는 것이 '분할표의 Chi-square검정'이다. 실제 의학연구에서 가장 많이 사용되는 검정법이 바로 이 '분할표(contingency table)의 Chi-square검정'이다.

예를 들어 신장이식후 사용하는 면역억제제의 효과를 알아보기 위해 환자를 A group(CsA 사용군과)과 B group(AZA 사용군)으로 나누어 생존율을 구할 경우, 각 group간의 HLA matching 정도가 다르지 않다는 것을 먼저 검정한 후 이식신 생존율의 차이를 구해야 한다. 각 group에서 HLA matching 정도에 따른 빈도를 아래의 표에 정리하면 다음과 같다.

HLA matching	Identical	Haplo-identical	Full-mismatch	Total
A group	3	3	4	10
B group	2	5	8	15
Total	5	8	12	25

위의 표에서 chi-square값을 구하면 1.076이 나오고 유의수준 0.05로 설정했을 때의 값 5.99보다 작으므로 영가설은 부정할 수 없게 된다. 따라서 이 두 group간에는 HLA matching의 정도에는 차이가 없다고 결론 지울 수 있다.

Chi square 검정은 위의 표처럼 자료가 이미 집계되어 있는 경우나, 아니면 처음상태의 자료로 입력되어 있더라도 컴퓨터로 처리가 가능하다.

□ SPSS/PC$^+$ 이용한 실제 □

신장이식전에 수혈을 받은 집단과 수혈을 받지 않은 집단간의 급성거부반응의 발생빈도를 비교하고자 한다.

수혈 받은 집단 : 급성거부반응 생긴 환자수 8명(1), 생기지 않은 환자수 22명(2)

비수혈 집단 : 급성거부반응 생긴 환자수 16명(3), 생기지 않은 환자수 9명(4)

<SPSS/PC$^+$ 입력>

```
DATA LIST  FREE/BLOOD  REJECTION. (참고 Blood 1은 수혈, 2는 비수혈이고
                                   REJECTION 1은 급성거부 발생, 2는 미발생)

BEGIN DATA.
1111111111
111111
1212121212
 .
 .  (생략)
 .
2 4 2 4 2 4 2 4
END DATA.
CROSSTAB  TABLES=BLOOD BY REJECTION.
/STAT=ALL.
```

<결 과>

```
          SPSS/PC+

Crosstabulation :     BLOOD
               By  REJECTION

REJECTION Count                    Row
                  1.00    2.00    Total
     BLOOD ------------------------------
        1.00        8      22      30
        ------------------------------
        2.00       16       9      25
        ------------------------------
        Column     24      31      55
        ------------------------------
        Total     43.6    56.4    100.0
Chi-Square      D.F.     Significance    Min E.F.   Cell with E.F. < 5
----------      ----     ------------    --------   ------------------
 6.28422         1          .0122         10.909          None
 7.72760         1          .0054        (Before Yates Correction)
```

<결과 해석>

p=0.0122(<0.05)로 수혈집단이 급성거부반응이 적다고 할 수 있다.

결 론

본편에서는 기초통계용어 및 간단한 T-test, Chi-square test, ANOVA 및 기타 비모수 검정법의 적용 및 실제에 대하여 간단히 살펴보았으며, 다음 II편에서

는 Survival analysis를 설명하고 III편에서는 Regression과 Cox proportional hazard model에 대하여 쓰도록 하겠다.

참 고 문 헌

1) Nie NH, Hull CH, Jenkins JG, Steinbrunner K, Bent DH SPSS: *statistical package for the social sciences. 2nd ed, New York, McGraw-hill, 1975*
2) SAS users' guide: *statistics-1982 edition. Cary, N. C.: SAS Institute, 1982*
3) Dixon WJ, Brown MB, Engelman L, et al: *BMDP statistical software 1983 Berkeley, Calif.: University of California Press, 1983*
4) Godfrey K: *Statistics in practics. N Engl J Med 313: 1450-1456, 1985*
5) 이동우 : 보건통계학 방법. 신광출판사, *1986*