한국통계학회논문집 제4권 1호
1997년 4월 pp. 65-74

15대 국회의원 소속별 의석예측에 대한 대수선형모형

이 재 창[1], 전 명 식[2], 정 형 철[3]

요 약

1996년 15대 국회의원 선거의 소속별 예측의석수와 실제의석수에 대한 정방형 분할표를 구하여 적절한 대수선형모형을 적합시키고 모형의 해석을 시도하였다.

1. 서 론

본 연구에 사용된 자료는 1996년 4월 11일 실시된 15대 국회의원 선거결과를 토대로 작성한 것이다. 당시 국내 5개 여론조사기관과 TV 3사의 공동작업으로 253개 지역구에 대한 투표자조사를 실시하여 개표를 시작하기 전에 예측당선자와 각 정당별 의석수를 예측 발표하였다. 이 예측결과는, 선거법상 투표소 500미터 내에서 출구조사를 금지하여 전화조사에만 의존한 관계로 대면문화를 중요시하는 우리나라 현실상 오차가 존재하리라고 예상되었다. 소속정당별 예측의석수와 실제의석수는 다음 <표 1.1>에 주어져있다.

<표 1.1> 정당별 예측의석수와 실제의석수

예측 \ 실제	신한국당	국민회의	민주당	자민련	무소속	예측의석수
신한국당	119	9	6	16	5	155
국민회의	2	56	0	0	0	58
민주당	0	1	3	0	0	4
자민련	0	0	0	25	0	25
무소속	0	0	0	0	11	11
실제의석수	121	66	9	41	16	253

1) (136-701) 서울시 성북구 안암동 고려대학교 통계학과 교수
2) (136-701) 서울시 성북구 안암동 고려대학교 통계학과 교수
3) (136-701) 서울시 성북구 안암동 고려대학교 통계학과 박사과정
* 본 연구는 고려대학교 통계연구소의 일부지원을 받았음.

위의 분할표의 특징은 행변수의 범주와 열변수의 범주가 신한국당, 국민회의, 민주당, 자민련, 무소속의 5개로 같으며 범주의 순서가 짝지어진 따라서 분할표가 5×5 정방형이라는 점이다. 여기서 행(X)은 예측의석수를, 열(Y)은 실제의석수를 나타내며 각각 예측범주, 실제범주라고 하자. 신한국당의 경우 예측의석수는 155석이었으나 실제 투표결과에 의해 155 지역구중 119 지역구에서만 당선자가 정확히 예측되었다. 나머지 36개의 예측이 틀린 지역구에서의 실제 당선자는 국민회의 9석, 민주당 6석, 자민련 16석, 그리고 무소속 5석의 결과가 나왔다. 이때 예측이 맞은 곳은 대각원소에 그리고 틀린 곳은 비대각원소에 위치하게 된다. 이 표를 근거로 예측오차율을 계산하면 39/253 = 0.154로 15.4%의 오차가 있었음을 보여준다.

그런데, 범주형 자료에 근거한 분할표자료가 주어지면 이를 잘 설명하는 모형을 찾음으로써 범주형 변수간의 연관성까지 아울러 고려할 수 있다. 여기서는 개표전 당선자 예측결과와 개표후 실제결과의 관계를 대수선형모형(log-linear model)을 활용하여 자료의 적합성 여부(goodness of fit)와 모수의 해석 측면에서 자료를 분석하고자 한다(Agresti, 1990).

2. 분 석 내 용

두 범주형 변수 (X, Y)의 결합분포로부터 구한 X의 i번째, Y의 j번째 범주에 해당하는 확률을 $\{\pi_{ij} : i=1,\ldots,I,\ j=1,\ldots,J\}$ 라고 하자. 이제, $N=I\times J$개의 칸(cell)에 대해 크기가 n인 다항표본이 있을 때 기대도수는 $\{m_{ij}=n\pi_{ij}\}$로 표현되며, 대수선형모형은 $\log m_{ij}$에 대한 분산분석과 유사한 선형모형을 가정하게 된다. 이제, 짝지어진 범주형 변수에 대해 주어진 $I\times I$ 정방형 분할표에 대한 몇 가지 대수선형모형을 고려하고 앞절에서 소개된 자료를 잘 설명하는 모형을 찾아보기로 한다. 분석에 필요한 프로그램은 SAS를 사용하였다(허명회, 1989).

2.1 독립모형

먼저 예측의석수와 실제의석수가 독립이라는 대수선형모형

$$\log m_{ij} = \mu+\lambda_i^x+\lambda_j^y \ , \quad i,j = 1,\ldots,5 \tag{2.1}$$

의 적합을 시도하여 보자.

그런데, 주어진 분할표는 비대각요소에 관측값이 0인 경우가 많으므로, 0의 값을 갖고 있는 모든 칸에 0.05를 더하여 준 후 자료분석을 하였다(Goodman, 1968). 이 결과 전체빈도수는 253에서 253.7로 큰 증가를 보이지는 않는다. 식 (2.1)의 독립모형에 대한 적합결과는 다음 <표 2.1>과 같다.

<표 2.1> 실제값과 독립모형하에서 기대값 (기대값은 괄호 안에 있음)

예측 \ 실제	신한국당	국민회의	민주당	자민련	무소속	예측의석수
신한국당	119 (74.018)	9 (40.384)	6 (5.590)	16 (25.141)	5 (9.867)	155.00
국민회의	2 (27.769)	56 (15.151)	0.05 (2.097)	0.05 (9.431)	0.05 (3.702)	58.15
민주당	0.05 (1.982)	1 (1.081)	3 (0.149)	0.05 (0.673)	0.05 (0.264)	4.15
자민련	0.05 (12.034)	0.05 (6.565)	0.05 (0.908)	25 (4.087)	0.05 (1.604)	25.20
무소속	0.05 (5.348)	0.05 (2.918)	0.05 (0.403)	0.05 (1.816)	11 (0.713)	11.20
실제의석수	121.15	66.10	9.15	41.15	16.15	253.70

<표 2.1>에서 보듯이 실제값과 기대값은 상당한 차이를 나타내게 된다. 이러한 차이에 대한 우도비검정통계량은 G^2 = 364.810 이 얻어지며 이는 자유도가 16인 카이제곱분포를 고려할 때 매우 큰 값으로 독립모형은 기각된다. 또한 모형의 연관도를 나타내는 람다를 측정하면 예측정보가 주어졌을 때 실제득표정보의 비대칭 람다값은 0.700으로 행범주와 열범주가 상당히 밀접한 관련이 있음을 나타내준다. 즉 위의 자료는 예측결과와 실제결과가 상당히 관련이 있음을 나타내고 있으며 이는 매우 상식적인 결과라 할 수 있다.

2.2 대칭모형

각 범주의 주변확률을 $\pi_{i+}=\sum_j \pi_{ij}, \pi_{+j}=\sum_i \pi_{ij}$ 라고 할 때, $I\times I$ 분할표에서 각 범주에 대해 $\pi_{i+}=\pi_{+i},\ i=1,\ldots,I$의 관계가 성립하면 π_{ij}에 대해 주변확률이 동질하다고 정의하며, $\pi_{ij}=\pi_{ji},\ i\neq j$ 이면 π_{ij}의 확률이 대칭분포라고 말한다. 대칭성을 고려할 수 있는 대수선형모형으로 다음과 같은 대칭모형(symmetry model)을 들 수 있다.

$$\log m_{ij} = \mu+\lambda_i+\lambda_j+\lambda_{ij} \qquad i,j = 1,\ldots,5 \tag{2.2}$$
$$\sum_i \lambda_i=\sum_j \lambda_j=\sum_i \lambda_{ij}=\sum_j \lambda_{ij}=0$$

이때, 자유도는 $I(I-1)/2$로 주어지게 된다. 이 모형은 대각선을 제외한 각 범주의 기대값이 $m_{ij}=m_{ji}$ 인 것을 의미하며 두 범주 모두 단일인자 모수 λ_i 를 갖는다.

이에 대한 모형적합결과는 다음 <표 2.2>와 같다.

<표 2.2> 0의 칸에 0.05씩 더한 결과와 대칭모형적합에서의 기대값(기대값은 괄호 안에 있음)

실제 / 예측	신한국당	국민회의	민주당	자민련	무소속	예측의석수
신한국당	119 (119)	9 (5.500)	6 (3.025)	16 (8.025)	5 (2.525)	155 (138.075)
국민회의	2 (5.500)	56 (56)	0.05 (0.525)	0.05 (0.05)	0.05 (0.05)	58.15 (62.125)
민주당	0.05 (3.025)	1 (0.525)	3 (3)	0.05 (0.05)	0.05 (0.05)	4.15 (6.650)
자민련	0.05 (8.025)	0.05 (0.05)	0.05 (0.05)	25 (25)	0.05 (0.05)	25.20 (33.175)
무소속	0.05 (2.525)	0.05 (0.05)	0.05 (0.05)	0.05 (0.05)	11 (11)	11.20 (13.675)
실제의석수 (기대값)	121.15 (138.075)	66.10 (62.125)	9.15 (6.650)	41.15 (33.175)	16.15 (13.675)	253.70 (253.70)

우도비검정통계량은 G^2 = 41.69이며 자유도 10인 카이제곱분포에 근거한 유의확률은 0.00001로 적합이 매우 좋지않다. 그 주된 이유는 $m_{ij}=m_{ji}$라는 제약조건에 있는 것으로 생각된다. 일반적으로, 대칭모형은 매우 구조적이어서 적합이 만족스러운 경우는 드물다. 가령, 각 범주의 주변분포가 이질적인 경우에도 대칭모형은 부적절한데, 주어진 자료의 경우에는 신한국당에 대한 예측의석이 과대 추정되었으므로 두 범주의 주변분포가 동일하다고 보기 어렵다. 이러한 주변확률분포의 동질성 여부를 통계적으로 검증하여보자. 각 범주의 표본 주변확률을 p_{i+}, p_{+i} , $i=1,\ldots,I$라 할 때, 주변확률의 차이는 $d_i = p_{i+} - p_{+i}$ 로 나타낼 수 있다. $D' = (d_1,\ldots,d_{I-1})$ 로 표기하면 주변확률이 동질하다는 가정하에 $E(D)=0$ 이며 $\sqrt{n}D$의 표본분산공분산 행렬을 $\hat{V}$ 라고 할 때 $\hat{V}$ 의 각 요소는 다음과 같다.

$$\hat{v}_{ij} = -(p_{ij}+p_{ji}) - (p_{i+}-p_{+i})(p_{j+}-p_{+j}) \;, \quad i \neq j$$

$$\hat{v}_{ii} = p_{i+} + p_{+i} - 2p_{ii} - (p_{i+}-p_{+i})^2$$

다음의 Wald 통계량 $W = n D' \hat{V}^{-1} D$의 분포는 n이 증가함에따라 자유도 $I-1$인 χ^2분포에 수렴하게 된다. Bhapkar (1966)는 W 를 주변확률분포의 동질성을 평가하는 통계량으로 삼았

다. 자료에서 W = 34.74로 자유도 4에서 주변확률분포가 동질하다는 가설을 기각하게 된다. 즉, 국회의원 의석 예측결과 각 행범주와 열범주의 특성은 같은 성질이지만 각 정당별 예측결과와 실제결과와는 상당히 다르다고 판단할 수 있으며 이를 토대를 대칭분포의 적용이 무리스럽다는 것을 알 수 있다. 이제 이러한 결과를 토대로 주변확률의 이질성을 고려한 비대각요소가 대칭인 준대칭모형을 적용해보자.

2.3 준대칭모형

앞의 대칭모형의 한계를 넓히기 위하여, 모형 (2.2)의 단일인자 모수 λ_i의 동일성을 배제한 다음과 같은 준대칭모형(quasi symmetry model)을 생각할 수 있다.

$$\log m_{ij} = \mu+\lambda_i^x+\lambda_j^y+\lambda_{ij}^{xy}, \quad \lambda_{ij}^{xy}=\lambda_{ji}^{xy}, \quad i\neq j \tag{2.3}$$

앞의 식 (2.2)의 대칭모형은 식 (2.3)의 준대칭모형에서 $\lambda_i^x=\lambda_i^y$, $i=1,\ldots,I$로 놓은 특수한 경우에 해당됨을 볼 수 있다. 식 (2.3)의 준대칭모형의 자유도는 $(I-1)(I-2)/2$ 로 주어진다.

이 경우 각 칸의 추정값은 우도함수에 의해 다음과 같다.

$$\widehat{m}_{i+}=n_{i+} \quad i=1,\ldots,I\,;\ \widehat{m}_{+j}=n_{+j} \quad j=1,\ldots,J\,;\ \widehat{m}_{ij}+\widehat{m}_{ji} = n_{ij}+n_{ji},\ i\neq j \tag{2.4}$$

위 식 (2.4)에서 보듯 m_{ii}는 직접적인 추정이 가능하나 m_{ij}의 추정은 뉴턴방법에 의한 반복비율추정에 의해 가능하다. 이에 대한 모형적합의 결과는 <표 2.3>에 주어져있다.

<표 2.3> 0의 칸에 0.05씩 더한 결과와 준대칭모형적합에서의 기대값 (기대값은 괄호안에 있음)

예측 \ 실제	신한국당	국민회의	민주당	자민련	무소속	예측의석수
신한국당	119 (119)	9 (9.647)	6 (5.507)	16 (15.897)	5 (4.948)	155.00
국민회의	2 (1.352)	56 (56)	0.05 (0.616)	0.05 (0.093)	0.05 (0.087)	58.15
민주당	0.05 (0.542)	1 (0.433)	3 (3)	0.05 (0.091)	0.05 (0.082)	4.15
자민련	0.05 (0.152)	0.05 (0.006)	0.05 (0.008)	25 (25)	0.05 (0.031)	25.20
무소속	0.05 (0.101)	0.05 (0.012)	0.05 (0.017)	0.05 (0.068)	11 (11)	11.20
실제의석수	121.15	66.10	9.15	41.15	16.15	253.70

준대칭모형에서 G^2 = 3.06 이며 자유도 6인 카이제곱분포에 근거한 유의확률은 0.80128로 적합이 뛰어나다. 모형적합의 관점에서 준대칭모형의 적합이 훌륭하며 모수의 추정값은 다음 <표 2.4>와 같다.

<표 2.4> 준대칭모형의 모수 추정값

주 효 과	신한국당	국민회의	민 주 당	자 민 련	무 소 속
λ_i^x (예측범주)	3.261	0.445	-0.409	-1.737	-1.560
λ_j^y (실제범주)	0.699	-0.152	-0.653	0.344	-0.238
λ_{ij}^{xy}(교호작용)	신한국당	국민회의	민 주 당	자 민 련	무 소 속
신한국당	1.456	-0.205	-0.263	-0.202	-0.786
국민회의	-0.205	4.371	0.363	-2.521	-2.008
민 주 당	-0.263	0.363	2.799	-1.693	-1.207
자 민 련	-0.202	-2.521	-1.693	5.249	-0.833
무 소 속	-0.786	-2.008	-1.207	-0.833	4.834

이제 모수의 추정값을 살펴보자. 예측범주모수 λ_i^x의 추정값의 크기가 해당예측의석수 크기의 순서와 약간 다르며 이는 교호작용에 해당하는 모수의 추정값 λ_{ij}^{xy}에 의해 어느 정도 설명되고 있다. 반면, 실제범주모수 λ_j^y의 추정값의 크기는 해당되는 실제예측의석수의 크기와 순서가 상당히 다르다. 가령, 자민련의 경우 예측의석수는 25석이었으나 실제의석수는 41석이며 λ_4^x = -1.737 로 예측의석수의 순위에 비해 작은 값을 가지나 실제범주모수의 추정값은 λ_4^y = 0.344 로 상당히 크다. 한편, 교호작용에 해당하는 모수의 추정값이 (국민회의-민주당)의 경우를 제외하곤 모두 음인 것은 매우 당연해보인다. 이는 상삼각부분의 크기가 하삼각부분의 크기보다 (국민회의 - 민주당)의 경우를 제외하곤 모두 크거나 같다는 점과 더불어 어떤 당의 예측이 증가하면 다른 당의 실제 의석수는 감소하게 된다는 점을 나타낸다고 하겠다. 그런데, 자민련의 경우 λ_{13}^{xy} = -0.202로 음의 교호작용 중 가장 큰 값을 가졌다는 점에서 예측의석수와 실제의석수의 차이인 16석이 모두 신한국당 당선예측지역으로부터 온 사실을 일부분 반영하는 것으로 여겨진다. 그러나, 준대칭모형은

교호작용의 상삼각부분과 하삼각부분이 서로 같다는 제약조건 ($\lambda_{ij}^{xy}=\lambda_{ji}^{xy}$, $i\neq j$)을 두었으며 이에 어느정도의 한계를 가지는 것이 사실이다.

2.4 준독립모형

$I\times I$ 분할표에서 대각요소의 범주의 관측값이 비대각요소의 관측값보다 상당히 크고 독립모형이 잘 적합되지 않는 경우 주어진 비대각원소에 대한 단순한 구조를 고려하는 다음과 같은 준독립모형(quasi independence model)을 고려할 수 있다.

$$\log m_{ij} = \mu+\lambda_i^x+\lambda_j^y+\delta_i\, I(i=j) \tag{2.5}$$

여기서 $I(i=j)$는 지시함수이다. 위의 식 (2.5)를 보면 $\delta_i\, I(i=j)$가 생략된다면 완전한 독립모형임을 알 수 있다. 이 경우 자유도는 $(I-1)^2-I$로 주어지며, 이러한 준독립모형은 준대칭모형의 특수한 형태이다. 소속별 국회의원의석에 대한 분할표는 독립모형이 부적절했으며 대각요소에 큰 관측값들이 있음을 알 수 있다. 이제, 준대칭모형의 특수한 형태이며 자유도가 더 큰 준독립모형을 적합하여 보자. 이 경우 각 칸 최대우도추정값은 다음과 같다.

$$\widehat{m}_{i+} = n_{i+}\quad i=1,\ldots,I\ ;\ \widehat{m}_{i+} = n_{i+}\quad i=1,\ldots,I\ ;\ \widehat{m_{ii}} = n_{ii}\quad i=1,\ldots,I \tag{2.6}$$

그리고 m_{ij}의 추정은 뉴턴방법에 의한다. 이때 준독립모형은 대각요소에 대해서는 완전히 적합되고있음을 알 수 있다. 모형적합의 결과는 <표 2.5>에 주어져있다.

<표 2.5> 실제값에 0.5씩 더한 결과와 준독립모형적합에서의 기대값 (기대값은 괄호 안에 있음)

예 측 \ 실 제	신한국당	국민회의	민주당	자민련	무소속	예측의석수
신한국당	119 (119)	9 (9.870)	6 (5.913)	16 (15.329)	5 (4.886)	155.00
국민회의	2 (1.270)	56 (56)	0.05 (0.198)	0.05 (0.515)	0.05 (0.164)	58.15
민주당	0.05 (0.640)	1 (0.167)	3 (3)	0.05 (0.259)	0.05 (0.082)	4.15
자민련	0.05 (0.129)	0.05 (0.033)	0.05 (0.020)	25 (25)	0.05 (0.016)	25.20
무소속	0.05 (0.109)	0.05 (0.028)	0.05 (0.017)	0.05 (0.044)	11 (11)	11.20
실제의석수	121.15	66.10	9.15	41.15	16.15	253.70

준독립모형에서 우도비검정통계량 G^2는 4.78이며 자유도 11인 카이제곱분포에 근거한 유의확률은 0.9413로 적합이 매우 뛰어나다. 적합이 뛰어난 이유로 대각요소에 대해 완벽한 추정이 이루어지고 있고, 나머지 비대칭요소들은 대부분 0이기 때문이라고 생각된다. 준독립모형의 경우 준대칭모형보다 축소모형이면서 자료적합이 매우 훌륭하므로 모형적합 측면에선 적절한 모형으로 사료된다. 이제 준독립모형에서의 모수의 추정값들을 살펴보자.

<표 2.6> 준독립모형 모수의 추정값

모 형	신한국당	국민회의	민주당	자민련	무소속
λ_i^x (예측범주)	3.798	0.406	-0.279	-1.879	-2.044
λ_j^y (실제범주)	1.229	-0.113	-0.626	0.327	-0.817
δ_i (대칭범주)	1.147	5.128	3.399	6.166	6.653

예측의석수 크기와 예측범주모수 λ_i^x의 추정값의 크기는 순서가 약간 다르나, 실제의석수와 실제범주모수 λ_j^y의 추정값의 크기의 순서는 자민련의 경우 크게 다름을 볼 수 있다. 이는 앞서 다룬 준대칭모형에서와 비슷한 결과이다. 또한 대칭범주모수 δ_i의 값은 국민회의 경우 가장 크며 신한국당이 가장 작은 값을 가지는데 이는 신한국당에 비교하여 국민회의의 경우 예측이 실제결과와 잘 맞았다는 점을 반영한 결과라 여겨진다.

이상에서 준대칭모형이나 준독립모형이 주어진 자료를 잘 설명한다고 볼 수 있다. 그런데 이 두 모형의 적합도를 비교하여보면 우도비검정통계량은 각각 G^2 = 3.06, G^2 = 4.78로 큰 차이가 없으나, 준독립모형이 더 큰 자유도를 가지며 따라서 모형의 간결성(parsimoniousness)에 비추어 선호된다고 여겨진다.

3. 결 론

본 연구에서는 최근에 많은 관심을 모았던 $I \times I$ 정방형으로 주어진 1996년 15대 국회의원 선거의 소속별 의석수 예측결과에 대해 적절한 대수선형모형을 적합시키고 모형의 해석을 시도하였다. 분석결과 준대칭모형이나 준독립모형의 적합이 뛰어난 것으로 드러났다. 한편, 15대 총선의석예측은 신한국당에 대한 과대추정이 이루어졌으며 이에 신한국당과 다른당과의 연관을 고려한 준대칭모형의 변형도 시도할 수 있겠다.

부 록

자료분석에 사용된 SAS 프로그램은 다음과 같다. 모수추정의 경우 SAS의 결과를 모형에 맞추어서 사용할 수 있다.

```
DATA CAT; INPUT X Y COUNT @@; CARDS;
1 1  119.00  1 2   9.00  1 3  6.00  1 4  16.00  1 5   5.00
2 1    2.00  2 2  56.00  2 3  0.05  2 4   0.05  2 5   0.05
3 1    0.05  3 2   1.00  3 3  3.00  3 4   0.05  3 5   0.05
4 1    0.05  4 2   0.05  4 3  0.05  4 4  25.00  4 5   0.05
5 1    0.05  5 2   0.05  5 3  0.05  5 4   0.05  5 5  11.00
;
DATA CAT_QS;INPUT X Y Z COUNT @@; CARDS;
1 1 1 119.00  1 2 1  9.00  1 3 1 6.00  1 4 1 16.00  1 5 1  5.00
2 1 1   2.00  2 2 1 56.00  2 3 1 0.05  2 4 1  0.05  2 5 1  0.05
3 1 1   0.05  3 2 1  1.00  3 3 1 3.00  3 4 1  0.05  3 5 1  0.05
4 1 1   0.05  4 2 1  0.05  4 3 1 0.05  4 4 1 25.00  4 5 1  0.05
5 1 1   0.05  5 2 1  0.05  5 3 1 0.05  5 4 1  0.05  5 5 1 11.00
1 1 2 119.00  2 1 2  9.00  3 1 2 6.00  4 1 2 16.00  5 1 2  5.00
1 2 2   2.00  2 2 2 56.00  3 2 2 0.05  4 2 2  0.05  5 2 2  0.05
1 3 2   0.05  2 3 2  1.00  3 3 2 3.00  4 3 2  0.05  5 3 2  0.05
1 4 2   0.05  2 4 2  0.05  3 4 2 0.05  4 4 2 25.00  5 4 2  0.05
1 5 2   0.05  2 5 2  0.05  3 5 2 0.05  4 5 2  0.05  5 5 2 11.00
;
DATA CAT_IND;SET CAT; IF (X = Y) THEN DELETE;
TITLE 'SYMMTRY  MODEL';
PROC CATMOD DATA=CAT_QS; WEIGHT COUNT;
    MODEL X*Y*Z = _RESPONSE_ / ML NOGLS PRED=FREQ ;
LOGLIN X Y X*Y ;
TITLE 'QUASI SYMMTRY MODEL';
PROC CATMOD DATA=CAT_QS; WEIGHT COUNT;
    MODEL X*Y*Z = _RESPONSE_ / ML NOGLS PRED=FREQ ;
LOGLIN X|Y X|Z Y|Z;
TITLE 'QUASI INDEPENDENCE MODEL';
PROC CATMOD DATA=CATIND; WEIGHT COUNT;
    MODEL X*Y = _RESPONSE_ /ML NOGLS PRED=FREQ ;
LOGLIN X Y ;
RUN;
```

참고문헌

[1] 허명회 (1989). 「SAS 범주형 자료분석」, 자유아카데미, 서울.

[2] Agresti,A. (1990). *Categorical Data Analysis* John Wiley & Sons, New York.

[3] Bhapkar,V.P. (1966). A note on the equivalence of two test criteria for hypotheses in categorical data, *Journal of American Statistical Association,* Vol. 61, 228-235.

[4] Goodman,L.A. (1968). The analysis of cross-classified data:independence,quasi-independence and interactions in contingency tables with or without missing entries, *Journal of American Statistical Association,* Vol. 63, 1091-1131.

[5] Goodman,L.A. (1979). Simple model for the analysis of association in cross-classifications having ordered categories, *Journal of American Statistical Association,* Vol. 74, 537-552.