The Korean Communications
in Statistics Vol. 6, No. 2, 1999
pp. 375-382

Sensitivity Analysis for Ordered Categorical Data[1)]

Il Hyun Cho[2)], Taesung Park[3)]

Abstract

Linear-by-linear association models are commonly used to analyze ordered categorical data. To fit these models, appropriate scores need to be chosen. In this paper, we perform sensitivity analyses in two-way contingency tables to investigate the effect of scores on goodness-of-fits and on tests of significance. In addition, we show that the best score, which yields the best fit of data, can be selected based on the sensitivity analysis results.

1. 서 론

범주형 변수들 중 순서척도를 가진 변수를 순서 범주형 자료라고 한다. <표 1>의 자료는 십이지장 환자들에 대한 처방과 그 처방에 의해 나타날 수 있는 구토 고통정도를 조사한 자료로 위절제 처방(변수 1)과 구토 고통정도(변수 2)를 나타내는 두 변수에 의해 분류된 4×3 분할표이다. 처방은 위의 절제 정도를 나타내는 변수로 각 범주들은 처방의 정도에 따라 순서를 갖고 있다. 처방 A는 배액치료와 위신경 절제, 처방 B는 25%절제와 위신경 절제, 처방 C는 50%절제와 위신경 절제, 처방 D는 75%절제와 위신경절제를 나타낸다. 변수 2는 구토 고통정도를 나타내며 처방에 의해 생길 수 있는 부작용을 나타내고 있다. 역시 이 변수의 범주값도 순서를 갖고 있으며 부작용이 없는 범주(N, none), 약한 증상을 나타내는 범주(S, slight), 고통을 나타내는 범주(M, moderate)로 구분되어 있다.

보통의 로그 선형모형들은 변수들의 순서가 없는 명목형(nominal)으로 처리하기 때문에 순서에 대한 정보를 무시하는 문제점을 갖고 있다(남궁평, 홍종선,1991). 로그 선형모형 중에서 범주들 간의 순서를 고려하는 모형으로 선형-대-선형 연관성 모형(linear-by-linear association model)이 있다(홍종선, 1995; Agresti, 1984, 1990). 이 모형은 순서형 변수의 각 범주에 적당한 점수(score)를 사용하여 순서에 대한 정보를 분석에 이용하게 된다. 자주 사용되는 점수에는 등간격(equal interval) 점수와 단위간격 점수(unit spaced scores) 등이 있다. 합한 점수가 영이 되도록 단위간격점수를 변형시킨 점수가 사용되기도 하며 또한 표준화된 점수들도 사용된다.

그러나 순서형 범주에 적당한 점수를 선택하는 것은 그리 간단하지가 않다. 비록 범주들간에 순서가 존재하지만 질적인 형태의 값을 수치로 표현 가능한 양적인 점수 형태로 변환을 시켜줘야 하기 때문이다. 또한 선택한 점수에 따라 모형의 결과가 상당히 차이가 날 수도 있다.

---

1) This research was supported by S.N.U. Research Fund.
2) Graduate, Department of Statistics, Hankuk University of Foreign Studies, Kyungki-Do, 449-791, Korea
3) Associate Profesoor, Department of Statistics, Seoul National University, Seoul, 151-742

본 논문에서는 <표 1>에 있는 분할표와 같은 이원 분할표에서 두 변수가 모두 순서형으로 주어진 경우에 선형-대-선형 연관성 모형을 사용할 때에 점수변화에 대한 민감도 분석을 실시할 수 있는 방법을 제안하였다. 특별히 점수의 변화에 의한 모수의 유의성에 어떤 영향을 미치는지 살펴보았고 또한 모형의 적합도통계량에 미치는 영향도 분석해보았다.

2절에서는 순서 범주형 자료의 분석에 사용되는 선형-대-선형 연관성 모형에 대하여 정리해 보고 모형의 적합도 검정에 대하여 소개하였다. 3절에서는 <표 1>의 예제를 이용하여 선형-대-선형 연관성 모형에 대한 검정결과를 보여주고 또한 점수 선택이 모수에 대한 유의성검정에 어떤 영향을 주는가에 대한 민감도 분석을 실시하였다.

<표 1> 위 절제 처방과 구토고통정도 자료

처방 \ 구토고통	N	S	M	합계
A	61	28	7	96
B	68	23	13	104
C	58	40	13	110
D	53	38	16	107

(출처 : Grizzle 등 (1969))

2. 선형-대-선형 연관성 모형

일반적인 이차원 $I \times J$ 분할표에서 모형의 행과 열 범주에 각각 $u_1 \le u_2 \le \ldots \le u_I$ 와 $v_1 \le v_2 \le \ldots \le v_J$ 와 같은 순서 점수를 부여해보자. 선형-대-선형 연관성 모형은

$$\log m_{ij} = \mu + \lambda_i^X + \lambda_j^Y + \beta(u_i - \bar{u})(v_j - \bar{v}) \tag{1}$$

과 같이 정의된다(홍종선, 1995; Agresti, 1984, 1995). 여기서 μ는 전체 평균을 나타내고 λ_i^X는 행의 i번째 범주의 효과를 나타내며 λ_j^Y는 열의 j번째 범주의 효과를 나타낸다. 모형에 대한 제약조건은 $\sum_i \lambda_i^X = \sum_j \lambda_j^Y = 0$ 이다. 이 모형은 포화모형과 독립성모형의 중간에 위치하는 모형이다. 즉, 이 모형은 포화모형의 교호작용 항인 λ_{ij}^{XY} 항을 $\beta(u_i - \bar{u})(v_j - \bar{v})$으로 대체한 모형으로 포화모형의 특별한 경우로 간주될 수 있다. 또한 u_i 와 v_j 가 고정된 점수이므로 독립성모형에서 모수 β 만이 추가된 모형임을 알 수 있다. 이 모형에서 $\beta = 0$이면 독립성모형이 된다.

모형의 적합도 검정을 위한 통계량의 자유도는

$$\begin{aligned} \text{자유도} &= (\text{독립성모형의 자유도}) - 1 \\ &= IJ - I - J \end{aligned}$$

이다.

점수 u_i 와 v_j 의 정의에 따라 연관성의 강도를 나타내는 β 에 대한 해석이 달라질 수 있다. 자주 사용되는 u_i 와 v_j 의 점수 형태는 등간격 점수와 단위간격 점수가 있다. 등간격을 갖는 점수는 다음을 만족하는 점수이며

$$u_2 - u_1 = u_3 - u_2 = \cdots = u_I - u_{I-1}, \quad v_2 - v_1 = v_3 - v_2 = \cdots = v_J - v_{J-1}$$

해석하기가 용이하다는 장점을 갖고 있다. 단위 간격 점수는

$$u_i = i, \quad v_i = j$$

와 같이 정의된다. 또한 위의 점수들을 변형시켜 합이 영이 되도록 변환된 점수들과 표준화된 점수들도 널리 사용된다. 더 자세한 설명은 홍종선(1995)을 참조하기 바란다.

선형-대-선형 연관성 모형의 모수 β 는 다른 로그 선형모형의 모수와 마찬가지로 오즈비(odds ratio)를 이용하여 해석할 수 있다. 임의의 $h<i$ 인 두 행과 $j<k$ 인 두 열에 대해 모형 (1)로 부터, 로그 변환된 오즈비는

$$\log\left(\frac{m_{hj} \cdot m_{ik}}{m_{hk} \cdot m_{ij}}\right) = \beta(u_i - u_h)(v_k - v_j)$$

와 같이 표현될 수 있다. 행과 열점수간의 차이가 크면 로그 변환된 오즈비의 절대값은 커지게 된다. 이러한 사실로부터 선형-대-선형 연관성 모형의 β 값에 대하여 오즈비를 이용하여 해석할 수 있다. 행과 열점수의 간격이 $u_i - u_h = v_k - v_j = 1$ 인 경우에 모든 오즈비는 e^{β} 로 일정하게 나타난다. 따라서 순위변수를 갖는 모형의 특징을 설명하기 위해 다음과 같이 서로 인접한 4개의 칸에 대하여 오즈비를 정의할 수 있다.

$$\theta_{ij} = \frac{m_{ij} \cdot m_{i+1\,j+1}}{m_{i\,j+1} \cdot m_{i+1\,j}}$$

이러한 오즈비를 국소오즈비(local odds ratio)라고 부른다. 선형-대-선형 연관성 모형에서 로그 변환된 국소 오즈비는

$$\log \theta_{ij} = \beta(u_{i+1} - u_i)(v_{j+1} - v_j)$$

이며 만약 등간격 점수를 사용한 경우에는 모든 국소 오즈비의 값이 e^{β} 로 동일하게 된다. 이런 경우에 행과 열간에 균일연관성(uniform association)이 존재한다고 하며 이 모형을 균일연관성모형(uniform association model)이라고 부른다 (Agresti, 1984).

선형-대-선형 연관성 모형의 모수 추정량은 보통의 로그 선형모형에서 사용되는 최대우도추정

법을 사용하여 구할 수 있다. 추정량을 구하기 위해서는 반복비율적합(iterative proportionate fitting, IPF) 알고리즘이나 Newton-Raphson 알고리즘과 같은 반복적인 알고리즘을 이용하여 추정량을 구할 수 있다(홍종선, 1995; Agresti, 1984).

3. 민감도 분석

이제 <표 1>의 자료에 대하여 여러 형태의 점수를 사용하여 선형-대-선형 연관성 모형을 적합시켜 보자. 먼저 점수간의 간격이 1인 균일연관성모형을 적합시켜 보자. <표 2>에는 독립성모형과 균일 연관 모형을 적합시킨 결과가 정리되어 있다. 이 표에서 보는 것처럼 독립성모형보다는 균일 연관 모형이 대체로 더 적합이 잘 되었음을 알 수 있다. 특히 (1,3) 칸과 (4,1) 칸에서는 이러한 경향이 두드러지게 나타난다.

<표 2> 위 절제 처방과 구토고통정도 자료의 기대값 비교

처방 \ 구토고통	N	S	M	합계
A	61 (55.21) [63.54]	28 (29.70) [29.17]	7 (11.06) [7.29]	96
B	68 (59.86) [65.38]	23 (32.17) [22.12]	13 (11.97) [12.50]	104
C	58 (63.31) [52.73]	40 (34.03) [36.36]	13 (12.66) [10.91]	111
D	53 (61.58) [49.53]	38 (33.10) [35.51]	16 (12.32) [14.95]	107

( ) : 독립성모형의 기대도수 추정값

[ ] : 균일연관성모형의 기대도수 추정값

편의상 선형-대-선형 연관성 모형을 $L\times L$로 표시하고 독립성모형을 I로 표시하자. <표 3>에는 독립성모형의 적합도검정통계량인 $G^2(I)$ 값과 선형-대-선형 연관성 모형의 적합도검정통계량인 $G^2(L\times L)$ 값이 정리되어 있다. 여기서 간격이 1인 단위 간격 점수를 사용하였기 때문에 $L\times L$모형은 균일연관성모형이 된다. G^2는 로그 우도비 검정통계량(log-likelihood ratio statistic)으로 모형의 적합도를 나타내는 통계량이다. $G^2(L\times L)$ 값으로부터 모형 $L\times L$이 모형 I 보다도 1개의 모수를 더 갖고 있지만 더 적합이 잘 되었음을 알 수 있다. 두 모형의 G^2 값의 차이로부터 모수 β 에 대한 유의성 검정을 할 수 있다. 이 차이를 $G^2(I \mid L\times L)$ 라고 표시하자. 이 값은 모형 $L\times L$이 옳다는 가정 하에서 β 의 유의성을 검정하기 위한 조건부 검정통계량이 된다. 이 값이 6.29이고 유의확률이 0.012이므로 유의수준 0.05에서 귀무가설을 기각할 수 있다. 따라서 독립성모형보다 선형-대-선형 연관성 모형이 더 적합하다고 할 수 있다. 이 모형을 적합

하기 위한 SAS 프로그램은 홍종선(1995)을 참조하기 바란다.

<표 3> 독립성 모형과 선형-대-선형 연관성 모형의 적합도 검정

모 형	G^2	df	Prob
$G^2(I)$	10.88	6	0.0922
$G^2(L\times L)$	4.59	5	0.4680
$G^2(I \mid L\times L)$	6.29	1	0.012

<표 2>의 선형-대-선형 연관성 모형의 적합 결과는 간격이 1인 등간격 점수를 사용했을 때에 얻어진 결과였다. 만약 다른 형태의 점수를 사용한다면 어떤 결과가 얻어질까? 이제 여러 다른 점수를 사용하여 점수가 모형의 적합에 미치는 영향이 어떠한 지에 대해 민감도 분석을 실시하기로 하자.

점수를 선정하는 방법은 먼저 첫 번째 범주와 마지막 범주의 점수값을 선정한 다음에 중간 범주의 무게 중심을 앞쪽 범주에서부터 뒤쪽 범주로 조금씩 옮기는 방법을 사용하였다. 점수를 선정하기 위해 우선 행점수 u_i를 먼저 선택하기로 하자. 행이 처방을 나타내고 4개의 범주를 갖고 있으므로 다음과 같은 4개의 점수를 고려해보자.

$$(u_1, u_2, u_3, u_4) = (1,\ 2,\ 3,\ 4),\ (1,\ 1.5,\ 2,\ 4),\ (1,\ 2.5,\ 3.5,\ 4),\ (1,\ 3,\ 3.5,\ 4)$$

모든 첫 번째 범주의 점수값은 1로 마지막 범주의 점수값은 4로 고정시켰다. 첫 번째 점수는 등간격 점수이고 두 번째 점수는 앞쪽의 범주에 무게를 두었고 세 번째 점수는 가운데 범주에 무게를 두었고 네 번째 점수는 뒤쪽의 범주에 무게를 두었다. 즉 가운데 범주를 앞쪽의 범주에 가까운 것처럼 취급하는 점수와 가운데 위치하는 것처럼 취급하는 점수와 뒤쪽의 범주에 가까운 것처럼 취급하는 점수로 분류할 수 있다.

다음으로 선택된 행점수 (u_1, u_2, u_3, u_4)에 대하여 열점수 (v_1, v_2, v_3)를 선택하였다. 점수 v_j를 선정하는 방법도 역시 u_i와 마찬가지로 첫 번째 범주 점수 1과 세 번째 범주 점수 3을 고정한 후에 가운데 범주값의 점수를 다음과 같이 변화시켰다

$$v_2(w) = 1\cdot(1-w) + 3\cdot w = 1 + 2w,\ \ 0 \le w \le 1.$$

여기에서 $v_2(w)$는 점수 1과 3사이의 가중평균값으로 정의되었다. w값은 0과 1 사이에서 정의된다. 여기서는 0.1부터 0.9까지 변화시키면서 $v_2(w)$값을 구해 보았다.

이제 이 선정된 여러 형태의 점수들이 선형-대-선형 연관성 모형의 통계량 β 의 유의성 검정에 어떠한 영향을 주는가에 대하여 살펴보도록 하자. 또한 $v_2(w)$를 통하여 모수 β 의 유의성 검정에 어떠한 영향을 미치는 지 또한 모형의 적합도검정통계량인 G^2의 값에 어떤 영향을 미치는 지를 살펴보도록 하자.

<표 4>는 4개의 행점수와 8개의 열점수에 대하여 선형-대-선형 연관성 모형을 적합시킨 후에

얻어진 모형의 적합도통계량 G^2값과 p-값을 구해 정리해 놓았다. 이 결과들은 SAS 프로그램의 CATMOD를 사용하여 구했다(허명회, 1995).

<표 4> 점수변환에 의한 G^2 $(L \times L)$와 p-값의 변화표

행점수 (u_1, u_2, u_3, u_4)	통계량 (p-값)	열점수 $v_2(w)$ 1.2	1.4	1.6	1.8	2.2	2.4	2.6	2.8
(1, 2, 3, 4)	G^2	7.68	6.68	5.78	5.07	4.32	4.22	4.23	4.31
	(p-값)	(0.1751)	(0.2454)	(0.3283)	(0.4074)	(0.5042)	(0.5186)	(0.5169)	(0.5051)
(1, 1.5, 2, 4)	G^2	8.01	7.19	6.45	5.85	5.20	5.12	5.15	5.25
	(p-값)	(0.1557)	(0.2066)	(0.2652)	(0.3216)	(0.3917)	(0.4011)	(0.3979)	(0.3867)
(1, 2.5, 3.5, 4)	G^2	7.66	6.76	5.99	5.43	4.93	4.91	4.98	5.10
	(p-값)	(0.1759)	(0.2390)	(0.3073)	(0.3659)	(0.4241)	(0.4266)	(0.4183)	(0.4038)
(1, 3, 3.5, 4)	G^2	7.46	6.80	6.32	6.05	6.02	6.16	6.33	6.53
	(p-값)	(0.1886)	(0.2361)	(0.2765)	(0.3013)	(0.3041)	(0.2914)	(0.2752)	(0.2582)

<표 5> β 의 추정치와 β 에 대한 Wald 통계량과 우도비 검정통계량의 변화표

행점수 (u_1, u_2, u_3, u_4)	통 계 량 (p-값)	열점수 $v_2(w)$ 1.2	1.4	1.6	1.8	2.2	2.4	2.6	2.8
(1, 2, 3, 4)	β	0.1293	0.1488	0.1610	0.1652	0.1554	0.1458	0.1352	0.1246
	Wald 통계량	3.12	4.08	4.95	5.66	6.43	6.55	6.55	6.47
	(p-값)	(0.0772)	(0.0435)	(0.0260)	(0.0174)	(0.0112)	(0.0105)	(0.0105)	(0.0110)
	LRT 통계량	3.2	4.2	5.1	5.81	6.56	6.66	6.65	6.57
	(p-값)	(0.0736)	(0.0404)	(0.0239)	(0.0159)	(0.0104)	(0.0098)	(0.0099)	(0.0104)
(1, 1.5, 2, 4)	β	0.1127	0.1284	0.1390	0.1437	0.1375	0.1296	0.1204	0.1110
	Wald 통계량	2.95	3.78	4.52	5.10	5.69	5.75	5.71	5.61
	(p-값)	(0.0859)	(0.0519)	(0.0335)	(0.0239)	(0.0171)	(0.0165)	(0.0169)	(0.0179)
	LRT 통계량	2.87	3.69	4.43	5.03	5.68	5.76	5.73	5.63
	(p-값)	(0.0902)	(0.0547)	(0.0353)	(0.0249)	(0.0172)	(0.0164)	(0.0167)	(0.0177)
(1, 2.5, 3.5, 4)	β	0.1322	0.1503	0.1600	0.1612	0.1473	0.1368	0.1259	0.1155
	Wald 통계량	3.01	3.84	4.59	5.17	5.72	5.81	5.76	5.66
	(p-값)	(0.0827)	(0.0500)	(0.0321)	(0.0229)	(0.0164)	(0.0160)	(0.0164)	(0.0174)
	LRT 통계량	3.22	4.12	4.89	5.45	5.95	5.97	5.9	5.78
	(p-값)	(0.0727)	(0.0424)	(0.0270)	(0.0196)	(0.0147)	(0.0145)	(0.0151)	(0.0162)
(1, 3, 3.5, 4)	β	0.1423	0.1557	0.1597	0.1558	0.1355	0.1236	0.1122	0.1016
	Wald 통계량	3.09	3.69	4.19	4.51	4.67	4.58	4.43	4.25
	(p-값)	(0.0789)	(0.0546)	(0.0407)	(0.0336)	(0.0306)	(0.0323)	(0.0353)	(0.0392)
	LRT 통계량	3.42	4.08	4.56	4.83	4.86	4.72	4.55	4.35
	(p-값)	(0.0644)	(0.0434)	(0.0327)	(0.0279)	(0.0274)	(0.0298)	(0.0329)	(0.0370)

<표 4>에서 나타난 결과를 보면 모형의 적합도통계량의 값이 행점수의 변화나 열점수의 변화에 따라 등간격인 모형(균일연관성모형)과 어느 정도 차이를 보이고 있음을 알 수 있다. 그러나 이 차이는 $L \times L$ 모형의 적합도에는 영향을 미치지 않는 것을 알 수 있다. 왜냐하면, <표 4>의 모든 p-값이 0.05보다 크기 때문이다.

적합도통계량의 값은 열점수의 변화에 크게 영향을 받지 않는 것으로 나타난다. 예를 들어 w가 0.1일 때 즉 $v_2(w)=1.2$ 일 때 $G^2(L \times L)$의 값은 행점수가 변함에 따라 7.68, 8.01, 7.66, 7.46으로 나타나므로, 이 값들은 u_i의 변화에 따라 별로 민감하게 변하고 있지 않음을 알 수 있다. w가 0.9일 때 $v_2(w)=2.8$ 일 때도 마찬가지로 $G^2(L \times L)$의 값이 4.31, 5.25, 5.10, 6.53으로 많은 차이를 보이지 않는다. 그러나 열점수의 변화에 따른 적합도통계량은 w가 0.1에서 0.9로 변함에 따라 조금씩 줄어들다가 나중에 다시 증가하는 경향을 보이고 있다. 예를 들어 행점수가 (1, 2, 3, 4)일 때 열점수가 변함에 따라 $G^2(L \times L)$의 값이 7.68에서 4.22까지 줄어들었다가 다시 증가하면서 4.31의 값을 가짐을 알 수 있다. 따라서 열점수가 행점수보다는 $G^2(L \times L)$의 값에 좀 더 큰 영향을 미치고 있음을 알 수 있다. 아마 그 이유는 행의 네 범주의 관찰 도수들은 비슷한 값을 갖지만 열의 세 범주의 관찰 도수는 범주 N과 S에서 크고 M에서 작기 때문인 것 같다. 이들 점수 중에서 가장 좋은 적합 결과를 제공하는 점수는 $(u_1, u_2, u_3, u_4)=(1, 2, 3, 4)$, $(v_1, v_2, v_3)=(1, 2.4, 3)$임을 알 수 있다. 이상을 종합해보건대 행점수와 열점수가 변하더라도 $G^2(L \times L)$ 통계량은 대략적으로 4와 8사이 값을 가지므로 $L \times L$ 모형이 적합이 잘되었음을 나타내고 점수변화의 효과가 크지 않음을 나타낸다.

<표 5>는 4개의 행점수와 8개의 열점수에 대하여 선형-대-선형 연관성 모형을 적합시켜 각 변환된 점수에 따라 나타나는 선형-대-선형 연관성 모형의 모수 β 값과 그에 해당되는 Wald 통계량과 우도비 검정통계량의 변화를 나타낸 표이다. <표 5>에서 Wald 통계량의 값은 SAS 프로그램의 CATMOD로부터 직접 출력된 결과이며, 우도비 검정통계량 $G^2(I \mid L \times L)$ 은 독립성 모형의 적합도검정통계량의 $G^2(I)$ 값에서 선형-대-선형 연관성 모형의 적합도검정통계량 $G^2(L \times L)$ 값의 차로 계산된 결과이다. p-값은 모수 β 의 유의성 검정에 대한 유의확률값을 나타낸다.

<표 5>에서 보는 것처럼 행점수의 변화에 따른 β 의 추정값의 변화는 그리 크지 않음을 알 수 있다. 예를 들어 w가 0.1일 때 즉 $v_2(w)=1.2$ 일 때 행점수가 변함에 따라 β 의 추정값들이 각각 0.1293, 0.1127, 0.1322, 0.1423으로 변하게 되나 이 값들의 변화는 그리 크지 않음을 알 수 있다. 또한 모수 β 의 유의성을 검정하기 위한 Wald 통계량이나 우도비 검정통계량의 값도 역시 행점수 u_i의 변화에는 별로 민감하게 변하고 있지 않음을 알 수 있다. $v_2(w)=1.2$ 일 때 Wald 통계량은 3.12, 2.95, 3.01, 3.09로 변하며 우도비 검정통계량은 3.2, 2.87, 3.22, 3.42로 크게 차이가 나지 않음을 볼 수 있다. 다른 w의 값에 대해서도 같은 경향을 볼 수 있다.

그러나 열점수는 행점수보다 β 의 추정값와 β 의 유의성에 좀 더 큰 영향을 미치고 있다. 예를 들어 행점수가 (1, 2, 3, 4)일 때 열점수가 변함에 따라, 즉 w가 0.1에서 0.9로 변함에 따라 β

의 추정값이 0.1293에서 0.1652까지 증가했다가 다시 감소하면서 0.1246까지 줄어드는 것을 볼 수 있다. 또한 Wald통계량과 우도비 통계량도 증가하다가 감소하는 경향을 보이고 있다. 이런 경향은 역시 p-값에서도 볼 수 있다. 고정된 행점수에 대하여 w가 0.1과 0.2일 때에는 유의수준 0.1에서 β가 유의하며 β가 0.3이상인 경우에는 유의수준 0.05에서 항상 유의한 검정결과를 얻을 수 있다.

요약하면 유의수준 0.1에서는 어떠한 점수를 선택하던지 상관없이 항상 유의한 결과를 얻을 수 있다. 그러나 유의수준 0.05에서는 β의 유의성검정결과가 점수의 선택에 따라 상반된 결과를 얻게 된다. 이들 점수 중에서 가장 큰 검정통계량의 값, 즉 가장 낮은 p-값을 제공하는 점수는 <표 4>에서 얻은 것과 마찬가지로 $(u_1, u_2, u_3, u_4) = (1, 2, 3, 4)$, $(v_1, v_2, v_3) = (1, 2.4, 3)$임을 알 수 있다.

4. 결 론

본 논문에서는 이차원 분할표에서 두 변수가 모두 순서형으로 주어진 경우에 선형-대-선형 연관성 모형을 사용하여 점수변화에 대한 민감도 분석을 실시하였다. 민감도 분석을 통해서 행점수와 열점수가 모형의 적합도통계량이나 순서척도의 의미를 지닌 모수의 유의성에 영향을 미치는 지를 알아 볼 수 있었다. 일반적으로 순서형 범주에 적당한 점수를 선택하는 것은 그리 간단하지가 않다. 범주들간에 존재하는 질적인 형태의 순서를 수치로 표현 가능한 양적인 점수 형태로 변환시켜야하기 때문이다. 또한 선택한 점수에 따라 모형의 결과가 상당히 차이가 날 수도 있다. 본 논문에서 실시한 민감도 분석을 통해서 점수 선택이 적합도검정통계량과 모수의 유의성에 어떠한 영향을 미치는 지를 구체적이고 체계적으로 살펴볼 수 있었다. 이런 민감도 분석은 유사한 형태의 이차원 분할표의 자료에 쉽게 응용될 수 있으리라 생각된다. 또한 민감도 분석을 통해서 모형의 적합도를 가장 좋게 하고 모수의 유의성 검정력을 가장 높게 만드는 최적의 점수를 선택할 수 있게 된다.

참 고 문 헌

[1] 남궁평, 홍종선 (1991). 범주형 자료의 통계분석, 자유 아카데미
[2] 허명회 (1995). SAS 범주형 데이타 분석, 자유 아카데미
[3] 홍종선 (1995). 대수 선형 모형, 자유 아카데미
[4] Agresti, A. (1990). Categorical Data Analysis. New York: John Wiley.
[5] Agresti, A. (1984). Analysis of Ordinal Categorical Data. New York: John Wiley.
[6] Grizzle, J. E., C. F. Starmer and G. G. Koch. (1969). Analysis of categorical data by linear models. *Biometrics,* 25, 458-504.