

主題公園 利用者들의 選擇行動 推定에 關한 研究

- Nested Logit Model의 適用 -

洪性權

建國大學校 農科大學 園藝科學科

A Study on Choice Behavior of Theme Park Visitors

- Application of Nested Logit Model -

Hong, Sung-Kwon

Dept. of Horticultural Science, Kon-Kuk University

ABSTRACT

This study was carried out to identify users' choice behavior of theme parks. Everland, Lotte World, Seoul Land, Dreamland and Children's Grand Park were selected as study areas. Both multinomial logit model(MNL), nested logit model(NMNL) and joint logit model were test, using a choice-based sample collected on study areas.

Hausman-McFadden test showed that the MNL is not appropriate because the IIA assumption is violated. To avoid the problematic IIA assumption, the NMNL was tested. It splits similar alternatives into groups and nests separate decisions into hierarchical order to avoid the IIA assumption. Cluster analysis and discriminant analysis were conducted to find applicable nest structures. The inclusive value coefficient was 0.7788. It meant that sufficient condition of this model is met and users' choice behavior can be better understood by NMNL than MNL. The ρ^2 value and accuracy of prediction of this model were 0.402 and 46.33%, respectively.

Several comments were suggested to make the NMNL to be more reliable for future research on users' choice behavior of theme park.

I. 序論

主題公園은 首都圈 주민들에게 가장 중요한 레크레이션 시설이 되었으며, 외국의 경우를 참조하더라도 이용자 수의 증가는 당분간 지속될 것으로 보인다(嚴, 1991). 이에 따라, 국내에서도 주제공원의 연구가 활성화 되고 있다(洪, 1995^a; 洪, 1995^b; 嚴, 1994; 鄭, 1993). 民間部門의 주제공원은 市場優位 확보가 운영의 궁극적 목적이기 때문에, 選擇行動(choice behavior)은 레크레이션 분야에서 가장 활발히 연구되어 온 주제중의 하나이다(Fesenmaier, 1990). 왜냐하면, 특정 주제공원이 선택되는 과정(mechanism)과 영향 요인들을 이해해야만 관리자들은 이용자들이 추구하는 바를 알 수 있으며, 그것들을 실제로 반영할 수 있기 때문이다(Clark and Downing, 1985).

현재 여러 종류의 離散선택모델(discrete choice model)이 선택행동의 추정에 사용되고 있다. 대부분의 모델들은 개인 n이 대상지 i에 대해 갖는 效用(U_{in})을 개인특성(a_n)과 대상지특성(X_{in})으로 구성되는 독립변수(V_{in}) 및 誤差(ε_{in})로 설명하며(식 1), 개인은 효용이 최대인 대상지를 합리적으로 선정한다고 假定한다(식 2). 그러나, 독립변수에 포함되지 못한 변수들로 인해 발생하는 오차(ε_{jn}, ε_{in})의 분포를 어떻게 설정하는가에 따라 모델들의 특성과 용도는 달라진다(洪, 1994^b; McFadden, 1978).

$$\begin{aligned}
 &U_{in}=U(X_{in}, a_n, \epsilon_{in})= V_{in} + \epsilon_{in} \\
 &U_{jn}=U(X_{jn}, a_n, \epsilon_{jn})= V_{jn} + \epsilon_{jn} \text{ ----- (식 1)} \\
 &P_n(i)=Pr(U_{in} \geq U_{jn})=Pr(V_{in}+\epsilon_{in} \geq V_{jn}+\epsilon_{jn}) \\
 &=Pr[(\epsilon_{in}-\epsilon_{jn}) \geq \beta(X_{jn}-X_{in})] \text{ ----- (식 2)}
 \end{aligned}$$

線形確率모델(linear probability model)은 오차의 분포(ε=ε_{jn}-ε_{in})를 상수로, 二項 프로빗모델(binary probit model)은 오차의 분포를 N(0, σ²)로 가정한다. 二項 로짓모델(binary logit model)은 ε_{in}과 ε_{jn}이 서로 독립적이며(independent), 分散이 동일한

(identical) Gumbel(Weibull)분포로 가정하여 선택확률을 계산한다(식 3)(그림 1). (식 2)의 양변에 동일한 상수(μ:scale parameter)를 곱해도 선택확률은 변하지 않기 때문에, 二項 로짓모델은 이 조건을 만족시키는 모든 경우 중에서(μβ) 하나의 계수값만을 정산하기 위해 μ를 보통 1로 간주한다. 이로 인해 二項 로짓모델로 정산한 계수는 二項 프로빗모델 계수보다 π/√3배 크다. 프로빗모델과 로짓모델은 거의 유사한 결과로 정산되나, 프로빗모델은 多項(multinomial)일 경우 계산이 어려워 로짓모델을 흔히 사용한다(Ben-Akiva and Lerman, 1985; Aldrich and Nelson, 1984).

$$\begin{aligned}
 P_n(i) &= \frac{\text{Exp}(\mu \cdot V_{in})}{\text{Exp}(\mu \cdot V_{in})+\text{Exp}(\mu \cdot V_{jn})} \\
 &= \frac{\text{Exp}(V_{in})}{\text{Exp}(V_{in})+\text{Exp}(V_{jn})} \text{ ----- (식 3)}
 \end{aligned}$$

$$P_n(i) = \frac{\text{Exp}(V_{in})}{\sum_{j=1}^J \text{Exp}(V_{jn})} \text{ ----- (식 4)}$$

多項 로짓모델(MNL, multinomial logit model)은 二項 로짓모델의 一般型으로(그림 2), (식 4)로 선택확률을 계산한다. 多項 로짓모델의 가정은 二項 로짓모델과 동일하지만, 비교 대상지가 3개 이상이어서 ‘오차의 독립성’ 가정은 부적절 代案으로 부터의 獨立性(IIA, Independence from Irrelevant Alternatives)을 동시에 의미한다(Luce, 1977). IIA假定은 “2 대상지가 선택될 확률의 비율이 다른 대상지에 의해 전혀 영향을 받지 않는다”로, 모든 대상지들간의 代替性(substitution)이 동일할 때만 성립한다. 그러나, 연구에 포함될 대상지중 일부는 類似할 경우가 많기 때문에, IIA를 가정하는 MNL로써는 선택확률을 정확히 계산할 수 없다. 예를 들어, 승용차, 빨간색 버스의 선택확률이 각각 1/2, 1/2인 상황에서 파란색 버스가 도입될 경우, MNL은 승용차, 빨간색

버스, 파란색 버스의 선택확률이 각각 1/3, 1/3, 1/3이라는 불합리한 결론을 유도한다.

이러한 문제점을 제거하기 위해 McFadden (1978)은 MNL의 一般型인 네스티드 로짓모델 (NMNL, nested multinomial logit model)을 제시하였다. NMNL은 IIA가정으로 인한 MNL의 단점을 해결하였을 뿐만 아니라, MNL의 장점인 계산의 용이성을 동시에 갖고 있다. 외국의 경우, NMNL은 교통 및 주택분야 등에서 오래 전부터 사용되었으나(Börsch-Supan, 1987; Ortuzar, 1983), 레크레이션 분야는 Peterson et al.(1983)이 적용 가능성을 제시한 이래 Siderelis et al.(1995)과 Lin et al.(1988)이 NMNL을 사용해 MNL보다 훨씬 우수한 결과를 보고하였다. 국내 레크레이션 연구의 경우, 崔 外 4人(1995)과 李(1993)는 二項 로짓모델로써 근린여가시설들의 선택행동을, 洪(1994^b)은 MNL로써 서울시內와 近郊의 레크레이션 시설의 선택행동을 추정하였으나, NMNL이 적용된 예는 별로 없다. 이에 본 연구는,

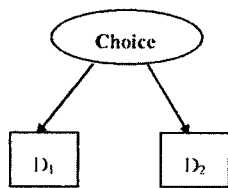
- (1) 서울시內와 近郊에 위치한 주제공원의 이용자들을 대상으로,
- (2) 多項 로짓모델의 一般型인 네스티드 로짓모델을 적용하여,
- (3) 특정 주제공원의 선택 과정과 이에 영향을 미치는 주요 변수들을 확인함으로써,
- (4) 주제공원들간의 類似性 有無에 관계없이, 일반적이면서도 체계적으로 선택행동을 연구하고자 한다.

Ⅱ. 네스티드 로짓모델 (Nested Multinomial Logit Model)

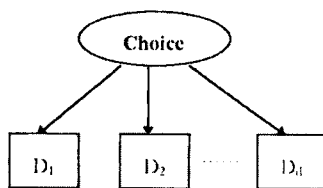
1. 네스티드 로짓모델의 필요성

MNL은 모든 대상지들간의 代替性이 동일하다고 가정한다. 그러나, 일부 대상지들이 類似할 경우에는 오차간의 相關性(correlation)이 존재하므로, '오차의 독립성' 가정은 지켜질 수 없어 Red Bus/Blue Bus문제가 발생한다. 또한 MNL은 대상지의 선택이 多次元的(multi-dimensional)이며 位階的(hierarchical)이라는 점을 고려하지 않는다. 주제공원의 선택시 이용자들은 방문목적(또는 추구효익)이 있는데, 이 목적은 특정 주제공원에서 뿐만 아니라 多數의 다른 주제공원에서도 이루어 질 수 있다. 이용자는 방문목적을 이룰 수 있는 주제공원집단을 먼저 선택한 후, 다른 요인들을 고려해 선택한 주제공원집단내 1곳을 최종 선택한다(그림 3). 즉, 선택가능한 모든 대상지중에서 (a)방문목적이나 욕구 같은 동기요인(motivational factor)을 적용해 유사한 대상지들의 集團(이하 上位代案이라고 칭함)을 high level에서 선택한 후, (b)비용, 거리등과 같은 중재요인들(intervening factors)로써 선택한 h^{th} 上位代案에 포함된 l^{th} 대상지(이하 下位代案이라고 칭함)를 low level에서 최종 선택한다(김 과 안, 1995; Clark and Downing, 1985; Schreyer et al., 1985; Styne and Peterson, 1984).

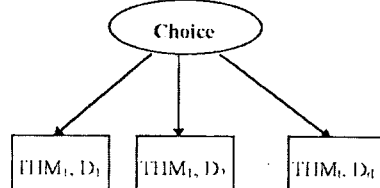
이와 같이, 上位代案의 선택에 영향을 미치는 변수(이하 上位代案관련 변수라고 칭함)와 下位代案의 선택에 영향을 미치는 변수(이하



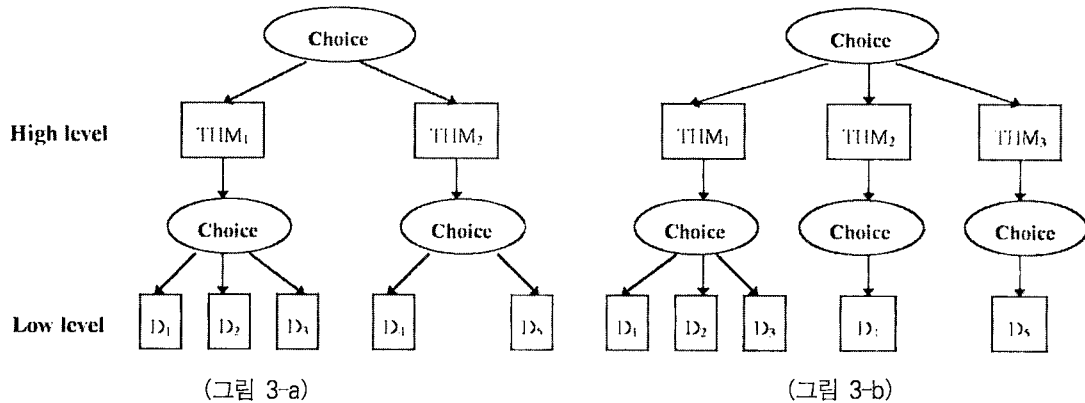
(그림 1) 二項 로짓모형



(그림 2) 多項 로짓모형



(그림 4) Joint 로짓모형



네스티드 로짓모델

THM: 유사대상지

D1: 에버랜드 D2: 롯데월드 D3: 서울랜드 D4: 드림랜드 D5: 어린이대공원

下位代案관련 변수라고 칭함)는 서로 논리적 연관성을 갖고 있으며 최종대상지의 선택시 다른 영향을 미치고 있지만, MNL은 독립변수를 단순히 上位代案관련 변수와 下位代案관련 변수의 조합으로 설정하기 때문에 多次元性이 무시된다. 또한, 대상지들간의 유사성으로 인해 上位代案과 下位代案이 순차적으로 선정되지만 MNL은 의사결정의 位階性을 무시하고 있다.

2. 모델의 형태 및 가정

NMNL은 MNL의 일반형이며, Generalized Extreme Value(GEV)모델의 특수형이다 (Ben-Akiva and Lerman, 1985; Hensher and Johnson, 1981). 이 모델은 (a)上位代案관련 변수(V_h)와 下位代案관련 변수(V_{hi}, V_i)를 따로 분리함으로써 독립변수의 多次元性을 고려하며, (b)독립변수들로 인한 오차(ϵ_h, ϵ_i)를 모델에 포함시킴으로써 대안들간의 類似性 정도를 고려하여, (c)개인 n 이 특정대상지에 대해 갖는 효용을 표현한다 (식 5). 그러나, (식 5)는 精算이 불가능하므로, 下位代案관련 독립변수 오차(ϵ_i)의 분산을 0으로 가정하면 (식 6)이 된다. 여기에 아래 3 가정을 추가하면, (a)효용 극대화 이론에 부합하면서, (b) I I A가정에 상관없이, (c)선택행동을 체계적으로 검증할 수 있

는 NMNL이 된다(식 7). 그러나, (식7)로는 μ^h/μ^i 의 비율만이 계산되므로 다른 로짓모델에서와 마찬가지로 low level의 scale parameter인 μ^i 를 1로 간주하면, NMNL에 의한 각 대상지의 선택확률이 최종 계산된다(식 8) (Ben-Akiva and Lerman, 1985; Maddala, 1983).

$$U_{hi} = V_h + V_i + V_{hi} + \epsilon_h + \epsilon_i + \epsilon_{hi} \text{ ----- (식 5)}$$

$$U_{hi} = V_h + V_i + V_{hi} + \epsilon_h + \epsilon_{hi} \text{ ----- (식 6)}$$

- U_{hi} : h^{th} 上位代案의 i^{th} 下位代案의 효용,
- V_h : 上位代案관련 변수,
- V_i : 下位代案관련 변수,
- V_{hi} : V_h 와 V_i 의 조합으로 구성된 下位代案관련 변수,
- ϵ_h : V_h 에 의한 오차,
- ϵ_i : V_i 와 V_{hi} 에 의한 오차,
- ϵ_{hi} : 나머지 오차

- ① 가정 1: ϵ_h 와 ϵ_{hi} 는 각각 독립적이다.
- ② 가정 2: ϵ_{hi} 는 독립적이며, 分散이 동일하며, scale parameter가 μ^i 인 Gumbel 분포한다.
- ③ 가정 3: 특정 上位代案에 포함된 下位代案의 최대효용은 scale parameter가 μ^h 인 Gumbel 분포한다.

$$P_n(hl) = P_n(h) \cdot P_n(l|h) = \left[\frac{\text{Exp}(V_h + I_n) \cdot \mu^h}{\sum_{h=1}^H \text{Exp}(V_h + I_n) \cdot \mu^h} \right] \cdot \left[\frac{\text{Exp}(V_{hl} + V_l) \cdot \mu^l}{\sum_{l=1}^L \text{Exp}(V_{hl} + V_l) \cdot \mu^l} \right] \quad \text{----- (식 7)}$$

$$I_n = (1/\mu^l) \cdot \ln \sum_{l=1}^L \text{Exp}(V_l + V_{hl}) \cdot \mu^l$$

$P_n(hl)$: 개인 n 이 h^{th} 上位代案의 l^{th} 下位代案을 선택할 확률

$P_n(h)$: 개인 n 이 h^{th} 上位代案을 선택할 限界確率 (marginal probability)

$P_n(l|h)$: 개인 n 이 h^{th} 上位代案을 선택하였을 경우, l^{th} 下位代案을 선택할 條件附確率(conditional probability)

H : 上位代案의 갯수, L : 下位代案의 갯수

I_n : h^{th} 上位代案의 포괄가격(inclusive value)

$$P_n(hl) = P_n(h) \cdot P_n(l|h) = \left[\frac{\text{Exp}(\alpha' V_h + (1-\theta)I_n)}{\sum_{h=1}^H \text{Exp}(\alpha' V_h + (1-\theta)I_n)} \right] \cdot \left[\frac{\text{Exp}[\beta'(V_l + V_{hl})]}{\sum_{l=1}^L \text{Exp}[\beta'(V_l + V_{hl})]} \right] \quad \text{----- (식 8)}$$

$$I_n = \ln \sum_{l=1}^L \text{Exp}[\beta'(V_l + V_{hl})] \quad \text{----- (식 9)}$$

α', β' : 정산할 독립변수의 계수

$1-\theta (= 1 - (\mu^h/\mu^l))$: 포괄가격 계수

3. Joint 로짓모델

Joint 로짓모델은 독립변수의 多次元性만을 고려하는 모델이다(그림 4). (식 6)에서 ϵ_h 가 생략되어 ϵ_{hl} 에 포함된 형태로(식 10), MNL의 독립변수(V_{in})를 V_h, V_l 및 V_{hl} 로 분리한 것과 동일하므로 선택확률은 (식 11)로 계산된다.

$$U_{hl} = V_h + V_l + V_{hl} + \epsilon_{hl} \quad \text{----- (식 10)}$$

$$P_n(h,l) = \frac{\text{Exp}(V_h + V_l + V_{hl})}{\sum_{h=1}^H \sum_{l=1}^L [\text{Exp}(V_h + V_l + V_{hl})]} \quad \text{----- (식 11)}$$

4. 포괄가격 (Inclusive Value)

포괄가격은 조건부확률의 분모에 log를 취한 형태로(식 9), 특정 上位代案에 포함된 下位代案들의 선택시 기대할 수 있는 최대효용을 나타낸다(Hensher, 1986). 또한 포괄가격은 한계 확률의 계산에도 직접 사용되기 때문에(식 8), NMNL과 MNL의 차이를 명확히 나타낸다.

NMNL은 유사대상지들을 上位代案으로 분류하므로, 동일 上位代案에 포함된 下位代案들간의 대체성은 동일해지는 반면, 다른 上位代案에 포함된 下位代案들과는 대체성이 다르게 된다. 따라서, 동일 上位代案에 포함된 下位代案들간에는 IIA가정이 성립하므로 NMNL의 조건부확률은 MNL과 동일한 형태로 계산할 수 있다. 그러나, 한계확률은 下位代案의 성격을 반영하고 있는 上位代案의 선택확률이어서 下位代案으로 부터의 영향력을 포함해야 하는데, 이런 역할을 담당하는 것이 포괄가격이다(Wissen and Rima, 1988; Börsch-Supan 1987).

포괄가격 계수(1- θ)중 θ 는 동일 上位代案에 포함된 대상지들간의 相關性 정도를 나타내는 것으로, (1- θ)값의 범위는 0과 1사이이다. 이 범위를 벗어나면 model specification error를 의미하는 것으로, 중요 독립변수들이 생략되어 있다는 것을 나타낸다(Lin et al., 1988; Siderelis et al., 1995; Hensher and Johnson, 1981).

- ① $0 < 1-\theta < 1$: NMNL이 성립
- ② $1-\theta = 1$: Joint 로짓모델
- ③ $1-\theta = 0$: MNL
- ④ $1-\theta < 0, 1-\theta > 1$: model specification error

5. 모델의 精算

NMNL은 최우추정법(maximum likelihood estimation)으로 정산하며, 정산시 low level과 high level을 분리하여 계산하는 순차적방법(sequential estimation)과 2 level을 동시에 고려해 계산하는 완전정보법(full information maximum likelihood estimation)으로 구분된다(Börsch-Supan, 1987, Hensher, 1986; Ben-Akiva and Lerman, 1985). 순차적방법은 (a)MNL로 조건부확률을 정산해 下位代案관련 변수들의 계수를 계산한 후, (b)정산된 下位代案의 계수로써 포괄가격을 계산해, 포괄가격을 마치 독립변수 처럼 간주한 후, (c)MNL로써 다시 한계확률을 정산해 上位代案관련 변수들의 계수를 정산한다. 이 방법은 low level을 먼저 계산하기 때문에 여러가지 문제를 일으킨다. 즉, 下位代案관련 변수의 계수로써 포괄가격을 계산하기 때문에 포괄가격은 "추정치들 사용한 추정치(estimate of estimates)"가 된다. 포괄가격 계수는 'Tree Structure'(그림 3-a, 그림 3-b)를 결정하는 요인이기 때문에 이 정산법의 약점은 명백하며, 이로 인하여 정산된 上位代案 변수의 계수들은 표준오차가 커져 유효성이 낮아지므로(inefficient) 따로 이를 補正해야 한다.

완전정보법은 순차적방법의 결점을 해결하기 위해 개발된 방법으로, 정산결과의 유효성이 높아(efficient) 표준오차의 補正이 필요없고, (식 8)의 순서대로 정산하므로 이론적이며, 순차적 방법보다 결과가 우수하다(Daly, 1987). 그러나, 이 방법은 선택 가능한 모든 대안을 응답자에게 제시해야 하며, MNL을 이용해 정산할 수 없어 특별한 software가 필요하다.

Ⅲ. 研究方法

1. 연구대상지

서울市內와 近郊에 있는 모든 주제공원을 연구대상지로 선정하였다. 구체적으로 롯데월드,

드림랜드, 능동 "어린이대공원(이하 대공원이 라고 칭함)", 서울랜드, 에버랜드를 연구대상지로 선정하였다.

2. 예비조사 및 분석

독립변수들을 선정하기 위하여 (a)주제공원 관련 연구시 수집한 원자료(raw data)를 본 연구의 목적에 맞추어 再分析하였으며(洪, 1995^a; 洪, 1995^b; 洪, 1994^a), (b)관련 문헌들을 조사하였다(Siderelis et al., 1995; 嚴 1994; 鄭, 1993; Backman and Crompton, 1991; Lin et al., 1988; 孫, 1985; 국민체육진흥공단, 1991). 上位代案 관련 변수들은 동기요인으로 구성하였으며, 下位代案관련 변수들은 사회·인구통계학적 변수, 상황변수, 예상 소요비용, 소요시간등으로 구성하였다(표 1).

선정된 上位代案관련 변수들의 사용가능성을 검토하기 위하여, 1995년 12월에 建國大學 校 農科大學 園藝科學科 학생 및 가족 150명을 대상으로 예비조사하였다. 5개의 주제공원 중 방문의도가 가장 높은 곳에 대해 설문을 평가케 한 결과, 대부분의 응답자들은 3 주제공원에 대해서만 평가하였다. 3 주제공원의 자료만을 要因分析(Factor Analysis)한 결과, 24개 모든 上位代案관련 변수들의 요인적재량/loading)이 0.5 이상이어서 본 설문에 포함시켰다.

3. 자료수집

(1) 표본크기의 결정

계산에 필요한 母集團 크기(N), 분산(σ^2), 정밀도(d), 신뢰도(z)값은 다음과 같이 결정하였다.

1) 휴일 이용자수는 평일의 2배라고 가정 한 후, 각 대상지의 1995년 5월 이용자수를 적용하여, 평일 하루와 휴일 이틀 동안의 이용자수를 계산하였다.

2) 예비분석에서 밝혀진 최대 분산인 4.224를 사용하였다.

〈표 1〉上位代案 및 下位代案관련 변수

上位代案 관련변수	SHOW1 SHOW2	다른곳 보다. 공연의 종류가 많아서 다른 곳에도 있지만, 이곳의 공연들이 더 재미있어서
	SHOW3 SPECIAL SEEING1	이곳에만 있는, 새로운 공연이 있어서 오늘, 특별행사 또는 야간공연이 있어서 다른곳 보다. 볼거리들의 종류가 많아서 (공연 제외)
	SEEING2	다른 곳에도 있지만, 이곳의 볼거리들이 더 재미 있어서(공연 제외)
	SEEING3	이곳에만 있는, 특이한 볼거리가 있어서 (공연 제외)
	PLAY1 PLAY2	다른 곳보다. 놀이기구의 종류가 많아서 다른곳에도 있지만, 이곳의 놀이 기구들이 더 재미있어서
	PLAY3	이곳에만 있는, 새로운 놀이기구가 있어서
	LAWN SCENE2 STRESS	탁 트인 잔디밭을 보고 싶어서 자연적 경치를 즐기기 위하여 일상생활에서 오는 스트레스를 풀기 위하여
	VARIETY	반복적인 일상생활로 부터, 변화를 추구하기 위하여
	RELAX EXER	휴식을 취하기 위하여 산책이나 가벼운 운동을 하기 위하여
	FRIEND FAMILY SAME	친구들과 함께 시간을 보내기 위해서 가족과 함께 시간을 보내기 위해서 나와 비슷한 부류의 사람들 또는 비슷한 나이 또래들과 같이 있고 싶어서
	EXOTIC	다른나라에 온 것 같은 분위기를 맛 보기 위해
	REMOTE	집에서 멀리 떨어져 있는 곳을 가고 싶어서
	SCENE1	사람이 만든, 인위적 경치를 감상하기 위하여
	NEW	본인 또는 같이 온 사람들에게, 무엇인가 배울 것이 있을 것 같아서
	ALONG	같이 온 사람들 모두가 즐길 수 있어서
下位代案 관련변수	AGE	나이
	HOLIDAY W/FRD	휴일 여부 친구와 동반 여부
	COST	예산 소요비용
	PSYCOST	예산 소요비용에 대한 주관적 평가
	TIME	출발지부터 대상지까지의 총 소요시간
	T-OUT	출발지부터 대상지까지의 소요 車外시간
	T-IN	출발지부터 대상지까지의 소요 車內시간
	FAR	출발지부터 대상지까지 거리에 대한 주관적 평가
	CAR	자가용 소유 여부
	LOYAL	忠誠度(Loyalty)
	IDLE CLIMATE	그저 시간을 보내기 위해서 날씨와 상관없이, 즐길 수 있어서

3) 연구자의 주관으로 정밀도는 0.13로 결정하였다.

4) 보편적 z값인 1.96을 사용하였다.

$$n = \frac{z^2 \delta^2 N}{(N-1)d^2 + z^2 \delta^2}$$

$$= \frac{(1.96)^2(4.224)(337,500)}{(337,500-1)(0.13)^2 + (1.96)^2(4.224)}$$

$$= \text{약 } 958\text{명}$$

(2) 標本抽出

표본은 choice-based sampling으로 추출하였다. 이 방법은 任意標出이나 層化標出에 비해 매우 경제적이다. 특히, 본 연구와 같이 대상지에서 標出해야만 응답자가 선택한 대상지를 알 수 있으며, 조사대상자들이 집단적으로 분포하는 경우, 또는 특정 대상지의 이용자수가 매우 적어 市場占有率에 맞추어 標出하는 것이 비경제적일 때 효과적이어서(Malhotra, 1984), 로짓 모델의 분석에 흔히 사용되고 있다(Lin et al., 1988). 이 방법은 層化한 후 응답자를 임의표출하는 점에서 층화표출과 동일하다. 그러나, 층화표출은 나이와 같은 외생(exogenous)변수로 층화하는 반면, 이 방법은 대상지를 기준으로 층화하므로 내생적(endogenous)이다.

Choice-based sampling은 대상지별 응답자들을 過大 또는 過小標出하기 때문에 정산된 계수는 偏倚(biased)되나, 각 대상지의 시장점유율을 고려해 정산하면 계수가 補正된다(Börsch-Supan, 1987; Cosslett, 1983).

(3) 자료수집

2인 1組로 구성된 조사원들이(총 18명) 1996년 5월 17일(金), 18일(土), 19일(日) 3일간 대상지 開場시간 부터 매 15분 간격으로 자료를 수집하였다. 조사지점은 각 대상지의 출입구나 주요 동선으로 설정하였다. 서울랜드는 실행상의 문제로 인하여 이 기간동안 필요 표본수의 1/3만을 수집하였으며, 5월 26(日)에 95개의 표본을 추가로 수집하여 표본의 휴일 비중이 매우 높아졌다(표 2).

<표 2> 대상지별 표본크기

	에버랜드	롯데월드	서울랜드	드림랜드	대공원	계
금요일	59	72	35	34	71	271
토요일	69	72	36	36	72	285
일요일	73	72	131	36	73	385
계	201	216	202	106	216	941

4. 자료분석

(1) 上位代案관련 변수

24개의 동기항목들을 要因分析(maximum likelihood, promax rotation)하였다. 모든 변수의 Kaiser's MSA가 0.5이상 이어서 (MSA평균: 0.85) 요인분석에 적절한 것으로 나타났으며, eigenvalue가 1 이상인 요인 5개를 추출하였다(표 3). 요인들을 각각 "공연요

<표 3> 上位代案관련 변수들의 요인분석

변수	Factor1	Factor2	Factor3	Factor4	Factor5
SHOW1	0.8230	0.0182	0.0456	-0.0563	-0.0153
SHOW2	0.7993	0.0860	0.0363	-0.0691	0.0134
SHOW3	0.7489	0.0308	0.0177	0.0182	-0.0000
SPECIAL	0.3970	-0.0341	0.0239	0.0998	0.0896
SEEING1	0.0497	0.7664	0.0789	-0.0151	-0.0050
SEEING2	0.0566	0.7940	0.0689	0.0055	-0.0365
SEEING3	0.1604	0.5597	0.0118	0.0787	0.0595
PLAY1	0.0329	0.0520	0.7208	-0.0186	-0.0386
PLAY2	0.0404	0.0026	0.3430	0.0005	-0.0188
PLAY3	0.1089	0.0796	0.6313	-0.0148	-0.0317
LAWN	-0.0374	0.1051	-0.1308	0.5902	-0.0335
SCENE2	-0.0448	0.1296	-0.1263	0.5617	-0.0580
STRESS	-0.0437	-0.0345	0.1845	0.5337	0.1279
VARIETY	-0.0062	-0.0080	0.1246	0.5329	0.0329
RELAX	-0.0311	-0.0380	0.0077	0.5228	0.0408
EXER	0.1443	-0.2209	-0.0623	0.4914	-0.1534
FRIEND	0.0110	-0.0231	0.0516	0.0377	0.7828
FAMILY	0.0143	-0.0129	0.1257	0.1293	-0.7852
SAME	0.1620	0.0201	-0.0240	0.1713	0.3794
EXOTIC	0.2117	0.0146	0.1240	0.3000	0.0640
REMOTE	0.1179	0.0222	0.0678	0.3080	0.1435
SCENE1	0.1559	0.1824	-0.0478	0.2495	-0.0052
NEW	0.1467	0.0990	-0.0604	0.2820	-0.1039
ALONG	-0.0841	0.1851	0.1130	0.2694	0.0127
Cum. Pct. of Var.	17.92	34.55	50.09	61.45	67.58

인”(Factor 1), “불거리요인”(Factor 2), “놀이기구요인”(Factor 3), “휴식/변화요인”(Factor 4), “친구요인”(Factor 5)으로 命名하였고, 응답자별 要因點數(factor score)를 上位代案관련 변수로 사용하였다(표 4).

(2) 下位代案관련 변수의 선정

사용 가능성이 있는 下位代案관련 변수들중에서(표 1), 名目水準변수(nominal)는 χ^2 검증으로, 序列水準(ordinal) 이상의 변수는 ANOVA로 유의성 있는 변수를 선발하였다(표 4). 上位代案관련 변수들은 요인점수이므로 다중공선성(multicollinearity)으로 인한 문제가 없지만, 下位代案관련 변수들은 이것이 문제이기 때문에 相關性분석으로 다중공선성이 심한 변수는 제외시켰다(李, 1989).

<표 4> 上位代案관련 변수 및 下位代案관련 변수의 계량법

	변수명	계량법
上位代案 관련변수	F1	공연요인의 요인점수
	F2	불거리요인의 요인점수
	F3	놀이기구요인의 요인점수
	F4	휴식/변화요인의 요인점수
	F5	친구요인의 요인점수
下位代案 관련변수	COST	예상 소비비용 (천원)
	TIME	출발지부터 대상지까지의 소요시간 (분)
	CLIMATE W/FRID	날씨와 상관없이, 즐길 수 있어서 친구와의 동반 興否에 관한 dummy변수 (1: 친구와 함께, 0: 기타)
	HOLIDAY	휴일 興否에 관한 dummy변수 (1:휴일, 0:평일)
	LOYAL	忠誠度 (1:방문중인 주제공원을 가장 좋아하며, 그 주제공원을 1년에 1회이상 방문 0: 기타)

(3) Tree Structure의 선정

본 연구는 5개의 대상지를 포함하기 때문에 가능한 Tree Structure(이하 모델構造라고 칭함)는 총 51개이다. 모든 모델構造에 대해서 Akaike Information Criterion 등으로

바람직한 structure를 일일이 확인할 수는 있지만(Börsch-Supan, 1987; Amemiya, 1980), 매우 많은 시간이 소요된다. 그래서, 본 연구는 군집분석(Cluster Analysis)과 판별분석(Discriminant Analysis)으로 적절한 모델構造를 결정하였다.

Pseudo F Statistic을 적용해 적절한 군집수를 2개로 결정하였으며(Calinski and Harabasz, 1974), F1, F2, F3 및 F4가 군집간 차이를 나타내는 주요 변수들로 나타났다(표 5). 분류 결과, 군집 1에는 에버랜드 이용자의 78.4%, 롯데월드의 68.1%, 서울랜드의 53.5%가 포함되었으며, 군집 2에는 드림랜드 이용자의 72.7%, 대공원의 78.3%가 포함되어, 두집단으로 구성된 모델構造인(그림 3-a)가 적합한 것으로 판단되었다.

밝혀진 모델構造의 신빙성을 검토하기 위해, 각 주제공원을 종속변수로 하여 판별분석하였다(표 6). 분석 결과, 3 대상지(에버랜드, 롯데월드, 서울랜드; 이하 THM 1 이라고 칭함)간에는 분류가 정확하지 않았다. 이는 예상하였던 것으로 下位代案관련 변수가 추가되어야 3 대상지간의 분류가 더 정확해 질 수 있다. 그러나, 드림랜드와 대공원의 경우는 예상과 달리 각각의 대상지가 비교적 정확히 분류되었다. 이는 드림랜드와 대공원이 따로 분류될 수 있는 가능성을 시사하는 것이어서, 세집단으로 구성된 모델構造인(그림 3-b)를 추가 정산하였다. 이 경우로 판별분석한 결과, F2, F3 및 F4가 중요 판별변수로 나타났다(표 7).

<표 5> 군집 평균

군집	F1	F2	F3	F4	F5
1	0.634	0.632	0.646	0.253	0.020
2	-0.701	-0.698	-0.714	-0.279	-0.022

* 군집 1: 에버랜드, 롯데월드, 서울랜드
 군집 2: 드림랜드, 대공원

<표 6> 판별분석 1

판별분석에 의한 분류 (obs./%)						
	에버랜드	롯데월드	서울랜드	드림랜드	대공원	계
에버랜드	76(38.2)	56(28.1)	29(14.6)	17(8.5)	21(10.6)	199(100)
롯데월드	49(22.7)	103(47.7)	25(11.8)	23(10.7)	16(7.4)	216(100)
서울랜드	41(20.3)	29(14.4)	45(22.3)	54(26.7)	33(16.3)	202(100)
드림랜드	10(9.4)	7(6.6)	15(14.2)	55(51.9)	19(17.9)	106(100)
대공원	14(6.5)	10(4.6)	22(10.2)	79(36.6)	91(42.1)	216(100)

* 에버랜드 자료중 2개는 missing

<표 7> 판별분석 2

Discriminant Functions Loading and Potency Index			
	FUNC 1	FUNC 2	potency index
F1	0.49283	0.42184	0.01933
F2	0.29013	0.92201	0.26414
F3	0.87760	0.21683	0.50582
F4	-0.24141	0.23819	0.18274
F5	0.06386	0.03320	0.00061

판별분석에 의한 분류 (obs./%)				
	THM 1	드림랜드	대공원	계
THM 1	393(63.7)	129(20.9)	95(15.4)	617(100)
드림랜드	18(17.0)	70(66.0)	18(17.0)	106(100)
대공원	24(11.1)	88(40.7)	104(48.2)	216(100)

(4) 모델 정산과 적합도

완전정보법으로 NMNL을 정산할 수 있는 LIMDEP version 7.0으로 MNL, NMNL 및 Joint 로짓모델을 각각 정산하였다(Greene, 1994). 완전정보법이 훨씬 우수한 정산법임에도 불구하고 순차적방법이 주로 사용되고 있는 이유는(Siderelis, 1995; Lin et al., 1988) 적절한 software의 부족이다(Daly, 1987). 이 정산법을 사용할 때는 MNL의 결과를 초기값으로 사용하는 것이 경제적이다(Ortuzar, 1983).

본 연구는 응답자들에게 방문중인 대상지만을 평가케 하였으므로 평가된 변수는 alternative specific variable 형태가 되었다. 대공원을

기준(base)으로 삼았으며, parameter의 偏倚를 補正하기 위하여 1995년 각 대상지의 시장점유율을 고려해 모델을 정산하였다.

정산된 모델의 적합도를 추정할 수 있는 최선의 방법은 아직 명확하지 않아 보통 1개 이상의 기준을 적용한다(Malhotra, 1984). 본 연구에서는 흔히 사용하고 있는 방법인 likelihood ratio index(ρ^2), 豫測力(accuracy of prediction)과 likelihood ratio test를 동시에 사용하였다. 일반적으로 0.2-0.4 사이의 ρ^2 값은 모델의 높은 적합도를 의미한다(Ben-Akiva and Lerman, 1985; Hensher and Johnson, 1981).

IV. 結果

1. 多項 로짓모델의 정산

(표 4)의 변수를 독립변수로 하여 MNL을 정산하였다. 본 연구는 choice-based sample을 사용하였으므로 대상지별 시장점유율을 고려하였고, 대공원을 base로 설정하였다. 유의성이 없는 변수를 차례로 제거한 결과, 예상소요비용(COST), 소요시간(TIME) 및 에버랜드의 휴일 여부(HOLIDAY)는 예상치 못한 부호로 정산되었다(표 8).

COST와 TIME의 정산 결과는 다음과 같이 해석할 수 있다. 본 연구는 특정 대상지에서, 응답자에게 그 주제공원에 대한 평가만을 의뢰하였다. 따라서, 응답자들은 그 대상지까지의 거리와 소요비용을 예상하고 그곳을 방문하였으며, 거리와 비용이 대공원보다 멀거나 높을지라도 문체시 많은 응답자들만이 분석되었기 때문이다. 이 2 변수의 계수값이 매우 낮은 것이 이같은 해석을 뒷받침하고 있다.

HOLIDAY의 경우, 에버랜드가 다른 대상지보다 멀리 떨어져 있는 곳임에도 불구하고 평일이 이용에 긍정적 영향을 미치는 것으로 판명된 것은, MNL이 주제공원의 선택 행동 연구에 적절한 분석법이 아님을 시사하고 있

다. Hausman-McFadden test로 IIA 가정을 검증한 결과(Hausman and McFadden, 1984) 'MNL이 적절하다'는 귀무가설이 기각되어($p < 0.000$), MNL의 일반형인 NMNL로 모델을 다시 정산하였다.

〈표 8〉 MNL의 정산결과

대상지	변수	계수	표준오차	유의성
에버랜드	HOLIDAY	-0.6602	0.1383	0.0000
	W/FRD	-0.3774	0.1228	0.0021
	F2	0.3528	0.1247	0.0047
	F3	2.2829	0.2016	0.0000
	F4	-0.3888	0.0882	0.0000
	F5	-0.3578	0.0731	0.0000
	COST	0.0822	0.0108	0.0000
	TIME	0.0262	0.0032	0.0000
롯데월드	CLIMATE	-0.3088	0.0456	0.0000
	HOLIDAY	-1.2137	0.2428	0.0000
	F1	0.3058	0.1535	0.0464
	F3	2.1578	0.2522	0.0000
	F4	-0.9679	0.1580	0.0000
서울랜드	COST	0.0774	0.0110	0.0000
	F1	0.3476	0.1319	0.0084
	F2	-0.4128	0.1799	0.0218
	F3	1.5090	0.2273	0.0000
	COST	0.0646	0.0114	0.0000
	TIME	0.0118	0.0037	0.0012
드림랜드	CLIMATE	-0.2319	0.0592	0.0001
	F2	-1.1163	0.2430	0.0000
	F3	1.6433	0.2969	0.0000

Log likelihood [-L(β)] -1005.184
 Restricted (slope=0) Log L [-L(0)] ... -1514.481
 $\rho^2 = 1 - [L(\beta)/L(0)]$ 0.336
 예측력: 에버랜드: 172/201= 85.57% 롯데월드: 69/216= 31.94%
 서울랜드: 46/202= 22.77% 드림랜드: 11/106= 10.38%
 대공원: 126/216= 58.33% 계: 424/941= 45.06%

2. 네스티드 로짓모델의 정산

MNL과 동일한 독립변수를 사용하였으나 (표 4) 上位代案관련 변수 및 下位代案관련 변수를 level별로 적용하였으며, MNL의 결과를 초기값으로 사용하여, 완전정보법으로 모델을 정산하였다. 가능성이 있는 2개의 모델構造 중, (그림 3-a)는 포괄가격 계수가 0과 1사이

에 포함되지 않아 제외하였다.

(그림 3-b)의 포괄가격 계수($1-\theta$)는 0.7788 이었다(표 9). 포괄가격 계수가 0과 1 사이에 포함되었다는 것은, 이 모델構造의 NMNL은 MNL보다 주제공원의 선택행동을 체계적으로 설명할 수 있다는 것을 나타낸다. 또한 포괄가격 계수의 값이 상당히 크다는 것은, 下位代案들간의 類似性($\theta=0.2212$)이 비교적 낮은 것을 의미한다(Siderelis et al., 1995, Lin et al., 1988). 즉, THM 1이 결정되면, 이용자들은 이들 3 주제 공원중 어느 곳을 선택할 것인지 쉽게 결정하는 것으로 나타났다. 정산된 모델의 적합도(ρ^2)와 예측력 모두 MNL보다 우수한 것으로 나타났다. 특히 ρ^2 값은 매우 우수하였다(0.402).

(1) 上位代案

예상한 대로, F2, F3과 F4가 上位代案의 선택행동시 주요 변수로 나타났다(그림 3-b)(표 7). THM 1의 선택은 놀이기구요인(F3)에 의해 크게 좌우되나, 휴식/변화요인(F4)은 부정적 영향을 주는 것으로 나타났다. 드림랜드 역시 놀이기구요인(F3)에 의해 긍정적 영향을 받고는 있으나 THM 1보다는 영향력이 적으며, 볼거리요인(F2)은 부정적 영향을 주었다. 단, 긍정적 또는 부정적 영향은 대공원에 비해 각 上位代案의 선택확률이 높아지거나 낮아진다는 것을 의미한다(Wissen and Rima, 1988).

(2) 下位代案

下位代案의 선택시 영향을 미치는 독립변수의 종류와 계수 크기는 MNL과 차이가 있었다. 이런 현상은 NMNL의 경우, 주제공원의 방문목적에 따라 上位代案을 먼저 결정한 후, 下位代案관련 변수들을 고려하여 1곳을 최종 선택하기 때문이다. 친구와의 동반여부(W/FRD), 휴일(HOLIDAY)여부, 忠誠度(LOYAL)및 날씨(CLIMATE)의 경우, 예상했던 계수의 부호와 크기가 정산되었다. 그러나, 예상소요비용(COST)과 소요시간(TIME)

은 계수의 부호가 '+'이며 크기는 매우 작았다. 이 결과는 MNL과 마찬가지로, 응답자들에게 방문중인 대상지만을 평가케 하였기 때문이며, 부호가 '+'인 곳들은 대공원보다 상대적으로 매력성(attraction)이 높은 곳으로 해석되었다.

1) 에버랜드의 경우, 친구와의 동반여부(W/FRD) 부호가 '-'로 나타났다. 이는, 에버랜드가 대공원보다 멀리 떨어져 있을 뿐만 아니라 교통수단도 불편하여 이곳의 선택은 친구보다는 다른 동반형태에 의해, 예를 들어 가족단위, 영향을 크게 받는 것으로 해석되었다.

2) MNL의 결과와는 달리, 휴일여부(HOLIDAY)가 에버랜드의 선택에 영향을 주지 않는 것으로 나타났다. 롯데월드는 COST의 계수가 매우 작을 뿐만 아니라 시내중심에 있어 평일이용이 용이하기 때문에, HOLIDAY의 부호가 '-'인 것으로 해석되었다. 서울랜드는 (a) 지하철로 접근성이 향상되었다는 점 이외에, (b)서울랜드는 휴일에 수집된 표본이 비중이 매우 높아(표 2) HOLIDAY의 계수가 매우 큰 값으로 정산되었다고 판단하였다.

3) MNL의 결과와는 달리, 최근 레크레이션 연구에서 중요성이 인정된 忠誠度(LOYAL)가 에버랜드나 롯데월드의 선택에 매우 긍정적 영향을 미치는 것으로 밝혀졌다. 忠誠度는 "오랜 기간동안, 1개 이상의 대안들 가운데 표현된 개인의 偏倚행동(biased behavior)"로 정의되며(Jacoby and Kyner, 1973), 충성스런 고객(loyal customer)들의 확보는 마케팅전략의 기본이 된다. 왜냐하면, 이들은 특정 대안에 대해 호의적 태도(favorable attitude)를 갖고 있을 뿐만 아니라 실제로도 그 대안을 지속적으로 이용하는 집단이기 때문에, 경쟁사들의 방문객 유치노력에 별로 영향을 받지 않아 안정적 시장의 확보에 필수적인 집단이기 때문이다(Schiffman and Kanuk, 1991; Backman and Crompton, 1991; Assael, 1984).

4) 롯데월드는 상당부분의 시설들이 실내에 있기 때문에 날씨(CLIMATE)가 선택에 영향을 주지 않는 것으로 나타났다. 그러나, 에버

랜드나 서울랜드는 대공원보다 예상소요비용과 소요시간이 매우 크기 때문에 부호는 '-'로 정산되었다고 판단하였다(표 10).

<표 9> NMNL의 정산결과

	대상지	변수	계수	표준오차	유의성
HL	THM 1	F3	1.8497	0.1970	0.0000
		F4	-0.4226	0.1195	0.0004
	THM 2	F2	-1.1939	0.1791	0.0000
		F3	1.5297	0.2580	0.0000
LL	에버랜드	W/FRD	-0.7437	0.1709	0.0000
		COST	0.0931	0.0286	0.0011
		TIME	0.0239	0.0033	0.0000
		LOYAL	1.8857	0.2552	0.0000
		CLIMATE	-0.3295	0.0569	0.0000
	롯데월드	HOLIDAY	-0.5264	0.1866	0.0048
		COST	0.0865	0.0275	0.0016
		LOYAL	0.5729	0.3136	0.0676
	서울랜드	HOLIDAY	1.1086	0.2472	0.0000
		COST	0.0703	0.0276	0.0109
		TIME	0.0089	0.0035	0.0098
	드림랜드	CLIMATE	-0.3068	0.0651	0.0000
		COST	0.0303	0.0149	0.0422
			TIME	-0.0159	0.0062
포괄가격 계수(1- θ)			0.7788	0.1430	0.0000

*HH: High Level, LL: Low Level
 THM1: 에버랜드, 롯데월드, 서울랜드 THM2: 드림랜드
 Log-likelihood -1024.778
 Restricted (slope=0) Log-L -1713.835
 ρ^2 0.402
 χ^2 1378.113 (p<0.00)
 예측력: 에버랜드: 176/201= 87.56% 롯데월드: 54/216= 25.00%
 서울랜드: 53/202= 31.19% 드림랜드: 28/106= 26.42%
 대공원: 115/216= 53.24% 계: 436/941= 46.33%

<표 10> 대상지별 예상소요비용과 소요시간의 평균 (천원, 분)

	에버랜드	롯데월드	서울랜드	드림랜드	대공원
소요비용	29.1	30.9	24.2	18.8	12.2
소요시간	83.6	60.9	59.8	40.5	45.6

3. Joint 로짓모델의 정산

NMNL 정산결과 포괄가격의 계수값은 1에 근접하였으므로(0.7788), Joint 로짓모델을 정산하였다(표 11). 예상하였던 대로 NMNL과 매우 유사한 결과가 나타났으나, ρ^2 와 예측력은 NMNL보다 떨어졌다.

<표 11> Joint 로짓모델의 정산결과

	대상지	변수	계수	표준오차	유의성
	THM 1	F3	1.9670	0.2062	0.0000
		F4	0.4260	0.1239	0.0006
	THM 2	F2	-1.2224	0.1822	0.0000
		F3	1.6140	0.2665	0.0000
	에버랜드	W/FRD	-0.7329	0.1650	0.0000
		COST	0.0719	0.0121	0.0000
		TIME	0.0223	0.0028	0.0000
		LOYAL	1.5942	0.1914	0.0000
		CLIMATE	-0.3313	0.0530	0.0000
	롯데월드	HOLIDAY	0.5587	0.1708	0.0011
		COST	0.0644	0.0105	0.0000
		HOLIDAY	0.9765	0.2171	0.0000
	서울랜드	COST	0.0507	0.0126	0.0001
		TIME	0.0076	0.0031	0.0137
		CLIMATE	-0.3075	0.0606	0.0000
	드림랜드	COST	0.0298	0.0146	0.0410
		TIME	-0.0144	0.0061	0.0184

Log-likelihood -1027.848
 Restricted (slope 0) Log L 1713.835
 ρ^2 0.400
 χ^2 1371.974 (p<0.00)
 예측력: 에버랜드: 179/201= 89.05% 롯데월드: 58/216= 26.85%
 서울랜드: 51/202= 25.25% 드림랜드: 28/106= 26.42%
 대공원: 116/216= 53.70% 계: 432/941= 45.91%

V. 結論 및 考察

본 연구는 서울市内와 近郊에 위치한 5개 主題公園의 利用者들을 대상으로, 이들의 선택 행동을 연구하였다. 기존에 사용하던 多項 로짓모델(MNL)은 부적절 代案으로 부터의 獨立性(IIA)을 가정하기 때문에, 잘 알려진 Red Bus/Blue Bus 문제가 발생한다. 이에

본 연구에서는 IIA가정이 문제되지 않으면서, 의사결정의 位階性과 多次元性を 동시에 고려할 수 있는 네스티드 로짓모델(NMNL)을 적용하였다. 자료는 choice-based sampling으로 수집하였고, LIMDEP version 7.0을 사용해 완전정보법으로 모델을 精算하였다. 완전정보법 사용시 계수가 偏倚(bias)되는 것을 방지하기 위하여 대상지별 市場占有率를 고려한 결과, 다음과 같은 결론을 도출하였다.

1. Hausman-McFadden test결과 IIA가정이 위배되어, MNL로써는 주제공원의 선택행동을 올바르게 이해할 수 없음이 확인되었다. 따라서, 본 연구에 MNL을 적용했다면, IIA가정이 위배되어 부적절한 결과가 유도되었을 것이다. 類似한 대상지들을 포함하는 선택행동의 연구는 NMNL을 사용해야 할 것이다.

2. 유의성 있는 변수들만을 고려해 NMNL을 정산한 결과, 포괄가격의 계수는 0.7788이었으며, 上位代案관련 변수는 불거리요인(F2), 놀이기구요인(F3)과 휴식/변화요인(F4)으로 밝혀졌고, 下位代案관련 변수는 예상 소요비용(COST), 출발지 부터 대상지까지의 소요시간(TIME), 친구와의 동반 여부(FRIEND), 휴일여부(HOLIDAY), 忠誠度(LOYAL) 및 날씨와 상관없이 즐길 수 있어서(CLIMATE)로 나타났다. 모델 적합성(ρ^2)은 0.402이었으며, 전체 응답자중 46.33% 응답자 선택을 바르게 예측하였다.

(1) 포괄가격 계수가 0과 1사이로 나타나(0.7788), NMNL은 NML보다 주제공원의 선택 행동을 체계적으로 설명할 수 있었다. 구체적으로 THM 1(에버랜드, 롯데월드, 서울랜드)은 드림랜드 또는 어린이대공원에 비해 상대적으로 均質하기 때문에 서로 집단화되어야 함이 밝혀졌다.

(2) 포괄가격의 계수값은 비교적 1에 가까웠다. 따라서, THM 1이 결정되었을 경우,

응답자들은 下位代案관련 변수를 고려해 3 주제공원중에서 1곳을 어렵지 않게 선택하는 것으로 나타났다. Joint 로짓모델의 분석결과가 이 결론을 뒷받침한다.

(3) 上位代案의 선택시, THM 1은 대공원에 비해 놀이기구요인(F3)에 의해서 긍정적인 영향을, 휴식/변화요인(F4)에 의해서는 부정적 영향을 받았다. 반면 드림랜드는 대공원에 비해 놀이기구요인(F3)에 의해서 긍정적인 영향을, 불거리요인(F2)에 의해서는 부정적 영향을 받았다.

(4) 下位代案의 선택은 대상지별로 종류가 다양하였다. 이는 5 주제공원별로 매력성, 위치, 소비비용등이 다르며, 상황변수들이 다르게 작용하였기 때문으로 해석되었다. 특히, 레크레이션 연구에서 최근 중요성이 인정된 忠誠度(LOYAL)가 에버랜드와 롯데월드의 선택에 매우 큰 영향을 미치고 있는 것으로 확인되었다.

NMNL이 주제공원의 선택행동 연구에 보다 적합한 기법이 되기 위해서는, 본 연구에서 적용치 못한 다음 사항들을 보완해야 한다.

1. 독립변수의 검토

유사 주제공원들로 나타난 에버랜드, 롯데월드와 서울랜드 중에서, 에버랜드의 예측력만이 높았다. 이는 롯데월드와 서울랜드의 특징을 반영하는 독립변수들이 모델에 반영되어 있지 못함을 의미할 수도 있다. 기존연구의 부족으로 본 연구에서는 포함되지 않았으나, 다음과 같은 변수들이 포함된다면 모델 설명력이 향상될 가능성이 매우 높다.

(1) 계절

본 연구는 5월중에 수집한 자료를 횡단분석(cross-sectional study)한 것이므로, 5월중의 선택행동만을 추정한 것이다. 선택행동을 일반화하기 위해서는 계절별로 자료를 수집해

분석해야 한다. 왜냐하면, choice-based sample에서 사용한 加重値가 변하면 정산결과도 변하기 때문이다. 특히, 롯데월드는 年평균 시장점유율과 5월중 시장점유율의 차이가 매우 크다.

(2) 의사결정자

전체 응답자의 35.1%만이 본인 스스로 방문지를 결정한 것으로 나타났다. 본인의 의사결정여부가 더미변수로 포함되는 것이 바람직하다.

(3) 忠誠度(Loyalty)

충성도는 행동특성과 태도특성으로 구성되기 때문에, 이들을 기준으로 4집단으로 분류하는 것이 보통이다(Backman and Crompton, 1991). 그러나, 본 연구는 이용빈도를 행동특성으로, 특정 주제공원의 선호도를 태도특성으로 단순화한 후, '충성스러운 고객(true loyal)' 1 집단만을 분류하였다. 보다 정교한 충성도의 측정척도를 사용해야 한다.

2. Tree Structure의 검토

적절한 Tree Structure(모델構造)를 설정할 수 있는 확실한 방법은 아직 없다. 전문가적인 판단이 도움이 될 수는 있다(Daly, 1987). 본 연구는 총 51개의 모델構造중 군집분석과 판별분석을 동시에 사용해 가능성이 높은 2개만을 검토하였다. 그 이유는 LIMDEP으로 다른 모델構造들을 정산하려면, 모델構造에 맞추어 자료를 다시 입력해야 하기 때문이다. 그러나, 군집분석은 (a)자료를 단순히 집단화하는 것이어서 통계학적 근거(foundation)가 약하며, (b)최적의 군집수를 설정하는 기준이 아직 명확하지 않기 때문에(Hair et al., 1995; Everitt, 1979), (c)전문가적 판단으로 군집분석과 판별분석의 결과를 해석한 후, (d)가능성 있는 모델構造들을 보다 광범위하게 검색해야 할 것이다.

(1) (그림 3-a)는 포괄가격 계수가 0과 1범위를 벗어나 분석에서 제외하였다. 그러나, 이

모델構造가 적절치 못하다고 단정할 수는 없다. 왜냐하면, model misspecification은 본 연구에서 사용한 독립변수를 고려했을 때의 결론이다. 본 연구에서 제외된 다른 변수들이 추가 도입된다면, 주제공원의 선택행동을 설명할 수 있는 다른 모델構造가 될 수 있다.

(2) "실내 주제공원인 롯데월드를 제외한 나머지 옥외 주제공원에서는 봄나들이 기간인 5월이 입장객수 최고월을 보이고 있다"(嚴, 1991). 따라서, 계절변수를 모델에 도입한다면, 본 연구에서 사용한 2-stage보다 3-stage가 바람직할 수 있다.

3. 자료수집

(1) 본 연구는 응답자에게 방문중인 대상지만을 평가케 하였다. 이로 인하여 COST와 TIME의 부호가 '+'로 나타났으며, MNL로는 계산이 불가능한 교차탄력성(Cross Elasticity)의 분석결과도 생략하였다. 이 변수들의 계수 크기가 매우 작기는 하였으나 부호가 '+'인 것은 비 합리적일 수 있다. 각 주제공원들까지의 소요시간 또는 예상소요비용을 각각 평가케 하는 것이 바람직하다.

(2) 서울랜드 자료의 64.9%가 휴일에 수집되었기 때문에, 휴일/평일 비율이 과대평가 되었다.

引用文獻

1. 국민체육진흥공단 (1991), 「한강조정/카누경기장 개발기본계획」, 국민체육진흥공단.
2. 김 성진, 안 진용 (1995), "소비자 공원선택행동의 모형화: 선택이론/모형의 검토와 적용", 『韓國造景學會誌』, 23(3):59-68.
3. 孫 大鉉 (1985), 『觀光마케팅論』, 日新社.
4. 嚴 瑞浩 (1994), "主題公園 서비스質의 測定 尺度 開發에 關한 研究", 『韓國造景學會誌』, 22(2):25-38.
5. 嚴 瑞浩 (1991), "주제공원의 워별 입장객수 추이분석과 마케팅 측면의 시사점", 『디벨로퍼』, 7/8:24-26.
6. 李 尙雨 (1993), 「大都市 市民의 遊樂 餘暇行動 選擇模倣

- 設定에 관한 연구: 서울을 중심으로, 서울시立大學校 碩士學位論文.
7. 李賢求 (1989), 「네스티드로짓모형을 이용한 쇼핑통행의 행태 분석에 관한 연구」, 서울大學校 環境大學院 環境計劃學科 碩士學位論文.
 8. 鄭必鎭 (1993), 「主題公園의 포지셔닝에 관한 연구」, 京畿大學校 觀光開發學科 碩士學位論文.
 9. 崔紀秀, 金漢培, 陳亮教, 陳相詰, 許美善 (1995), 「大都市 住民의 近隣屋外 餘暇施設 選擇模型을 기초로 한 施設地 配分에 관한 研究(I)」, 「韓國造景學會誌」, 23(1):123-139.
 10. 洪性權 (1995^a), 「主題公園의 競爭力 提高 方案에 關한 研究: Hybrid Conjoint Analysis의 適用」, 「韓國造景學會誌」, 23(2):1-16.
 11. 洪性權 (1995^b), 「重要度-成就度 分析을 利用한 都市公園의 管理方案에 關한 研究」, 「韓國造景學會誌」, 23(3):94-105.
 12. 洪性權 (1994^a), 「當日餘暇用 Recreation施設의 포지셔닝에 관한 研究: 서울市를 中心으로」, 「韓國造景學會誌」, 22(2): 13-24.
 13. 洪性權 (1994^b), 「서울市內와 近郊에 위치한 當日餘暇用 Recreation施設의 選擇行動 推定에 關한 研究: Generalized Logit Model의 適用」, 「韓國造景學會誌」, 22(3):1-12.
 14. Assael, Henry (1984), *Consumer Behavior and Marketing Action*, Boston, MA: Kent Publishing Company.
 15. Aldrich, J.H. and F.D. Nelson (1984), *Linear Probability, Logit, and Probit Models*, Newbury Park, CA: Sage publications.
 16. Amemiya, Takeshi (1980), "Selection of Regressors", *International Economic Review*, 21(2): 331-354.
 17. Backman, Sheila J. and John L. Crompton (1991), "Differentiating between High, Spurious, Latent, and Low Loyalty Participants in Two Leisure Activities", *Journal of Park and Recreation Administration*, 9(2): 1-17.
 18. Ben-Akiva, Moshe and Steven R. Lerman (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, Cambridge, MA: The MIT Press.
 19. Börsch-Supan, Axel (1987), *Econometric Analysis of Discrete Choice: with Applications on the Demand for Housing in the U.S. and West-Germany*, Berlin: Springer-Verlag.
 20. Calinski, T. and J. Harabasz (1974), "A Dendrite Method for Cluster Analysis", *Communications in Statistics*, 3(1): 1-27.
 21. Clark, Roger and Kent B. Downing (1985), "Why Here and Not There: The Conditional Nature of Recreation Choice", *Proceedings-Symposium on Recreation Choice Behavior*. USDA Forest Service, General Technical Report INT-184.
 22. Cosslett, Stephen R. (1983), "Efficient Estimation of Discrete-Choice Models", In C. F. Manski and Daniel McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications* pp.51-111, Cambridge, MA: The MIT Press.
 23. Daly, Andrew (1987), "Estimating 'Tree' Logit Models", *Transportation Research*, 21B(4): 251-267.
 24. Everitt, B. (1979), "Unsolved Problems in Cluster Analysis", *Biometrics*, 35: 169-181.
 25. Fesenmaier, Daniel R. (1990), "Theoretical and Methodological Issues in Behavioral Modeling: Introductory Comments", *Leisure Sciences*, 12(1): 1-7.
 26. Greene, W. H. (1994), *LIMDEP*, New York, NY: Econometric Software, Inc.
 27. Hair, Joseph F., Jr., Rolph E. Anderson, Ronald L. Tatham, and William C. Black (1995), *Multivariate Data Analysis*, Englewood Cliffs, NJ: Prentice Hall.
 28. Hausman, Jerry and Daniel McFadden (1984), "Specification Tests for the Multinomial Logit Model", *Econometrica*, 52(5): 1219-1240.
 29. Hensher, David A. (1986), "Sequential and Full Information Maximum Likelihood Estimation of a Nested Logit Model", *The Review of Economics and Statistics*, 68(Nov.): 657- 667.
 30. Hensher, D.A. and L. W. Johnson (1981), *Applied Discrete Choice Modelling*. London: Croom Helm.
 31. Jacoby, Jacob and David B. Kyner (1973), "Brand Loyalty vs. Repeat Purchasing Behavior", *Journal of Marketing Research*, 10(Feb.): 1-9.
 32. Lin, Yann-Jou, George L. Peterson, and Peter A. Rogerson (1988), "A Nested Urban Recreation Site Choice Model", *Leisure Sciences*, 10(1): 1- 15.
 33. Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.
 34. Malhotra, Naresh K. (1984), "The Use of Linear Logit Models in Marketing Research", *Journal of Marketing Research*, 21(Feb.): 20-31.
 35. McFadden, Daniel (1978), "Modelling of the Choice of Residential Location", In Anderson Karlqvist, Lars Lundqvist, Folke Snickars, and Jörgen W. Weibull (Eds.), *Spatial Interaction Theory and Planning Models*, Amsterdam: North-Holland Publishing Co.
 36. Ortuzar, Juan De Dios (1983), "Nested Logit Models for Mixed-Mode Travel in Urban Corridors", *Transportation Research*, 17A(4): 283-299.

37. Peterson, George L., John F Dwyer, and Alexander J. Darragh (1983). "A Behavioral Urban Recreation Site Choice Model", *Leisure Sciences*, 6(1): 61-81.
38. Schiffman, Leon G. and Leslie L. Kanuk (1991). *Consumer Behavior*, Englewood Cliffs, NJ: Prentice Hall.
39. Schreyer, Richard, Richard C. Knopf and Daniel R. William (1985). "Reconceptualizing the Motive/Environment Link in Recreation Choice Behavior", *Proceedings-Symposium on Recreation Choice Behavior*, USDA Forest Service, General Technical Report INT-184.
40. Siderelis, Christos, Gene Brothers and Phil Rea (1995). "A Boating Choice Model for the Valuation of Lake Access", *Journal of Leisure Research*, 27(3): 264-282.
41. Styne, Daniel J. and Geroge L. Peterson (1984). "A Review of Logit Models with Implications for Modeling Recreation Choices", *Journal of Leisure Research*, 16(4): 295-310.
42. Wissen, Leo Van and Annemarie Rima (1988). *Modelling Urban Housing Market Dynamics: Evolutionary Patterns of Households and Housing in Amsterdam*. Amsterdam: North-Holland.