

한국 성인 중 다빈도 외식소비자의 예측모형 개발: 데이터마이닝을 이용한 2001 국민건강·영양조사 자료 분석*

Developing a Model for Predicting Korean Adult
Consumers Who Frequently Eat Food-Away-From Home:
Data Mining of the 2001 National Health and Nutrition Survey*

성균관대학교 의과대학

연구조교수 정 상 진

이화여자대학교 통계학과

부 교수 강 승 호

석사과정 송 수 민

서울대학교 생활과학연구소

연수연구원 류 시 현

서울대학교 식품영양학과

조교수 & 생활과학연구소 겸임연구원 윤 지 현

School of Medicine, Sungkyunkwan University

Research Assistant Professor : Sang-Jin Chung

Department of Statistics, Ewha Womans University

Associate Professor : Seung-Ho Kang

Graduate Student : Su-min Song

Research Institute of Human Ecology, Seoul National University

Post-doc : Si Hyun Ryu

Department of Food & Nutrition & Research Institute of Human Ecology, Seoul National University

Assistant Professor : Jihyun Yoon

◀ 목 차 ▶

I. 서론

II. 연구방법

III. 결과 및 해석

IV. 논의 및 결론

참고문헌

Corresponding Author: Jihyun Yoon, Department of Food & Nutrition & Research Institute of Human Ecology, Seoul National University, San

56-1 Sillim-Dong, Kwanak-Gu, Seoul 151-742, Korea Tel: 82-2-880-8750, 019-422-6040 Fax: 82-2-884-0305

E-mail: hoonyoon@snu.ac.kr

* 이 논문은 2004년도 한국학술진흥재단의 지원에 의하여 연구되었음. (KRF-2004-042-C00017)

<Abstract>

The objective of this study was to develop a model for predicting Korean adult consumers who frequently eat food-away-from-home. A total of 7,032 adults aged 19 years and older from the 2001 National Health and Nutrition Survey in Korea were used as subjects. The data were analyzed using a data mining procedure including logistic regression and decile analysis. The model developed in the study was proven to be valid in predicting the consumers who frequently eat food-away-from-home(once a day or more often). This model showed that consumers eating food-away-from-home frequently tend to be younger men, living in a big city, working full time, receiving more stress and eating snacks and fried food more frequently. The model could be used to identify targets for nutrition and related education and consumer segments for the marketing of restaurant businesses.

주제어(Key Words): 외식(food-away-from-home), 소비자(consumers), 데이터 마이닝(data mining), 국민 건강·영양조사(national health and nutrition survey)

1. 서론

현대인의 식생활에서 외식이 차지하는 비중은 꾸준히 증가해 왔다. 우리나라의 경우, 특히 그 증가 속도가 1990년대 들어 가속화 되었는데, 우리나라 국민의 가구당 식료품비 중에서 외식비가 차지하는 비율은 1990년에 21.4%에서 2004년에 46.6%로 두 배 이상 증가하였다(통계청, 2005). 2001 국민건강·영양조사의 분석 결과에 의하면 전체 국민의 1/4 이상이 1일 1회 이상 외식을 하는 것으로 추정되었다(보건복지부, 2002).

이러한 외식비 및 외식인구의 증가는 외식산업의 성장으로 이어져 1990년에 10.6조원이었던 외식산업의 규모가 2003년 현재 44조원을 웃도는 것으로 보고되었다(통계청, 2004). 이에 외식업체나 외식산업을 대상으로 한 연구가 제 경영학 분야에서 활발히 이루어지고 있으며, 외식경영관련 전공 및 학과가 대학에 설치되기 시작하였다. 또한 1998년부터는 '외식경영연구'와 같은 학술지도 출판되어 외식에 대한 연구가 확대되고 있다. 그러나 이러한 외식관련 연구들은 경영학적인 관점에서 이루어져 외식산업의 효율성을 증대시키거나 외식산업을 활성화시키기 위한 목적을 가진 경우가 대부분이다. 약 20년 전부터 외식을 전통적인 가정기능의 사회화 현상으로 보는 관점에서 가정학자들의 연구가 시작된 것은 사실이나(모수미, 1986), 대한가정학회지(노정미, 유영상, 1989; 유정림, 박동연, 2001; 정은정, 김정숙, 2001; 주나미, 2004)나 한국가정관리학회지(류정순, 박

혜련, 1996)에 실린 논문 중 외식을 주제로 하는 논문이 소수에 불과하다는 사실에서 알 수 있듯이 그동안 가정학적 관점에서의 외식 연구가 활발히 진행되지는 못하였다. 외식이 식생활, 나아가 가정생활 전반에 미치는 영향이 증대되고 있음을 고려할 때, 이제 이러한 외식현상 및 외식소비자에 대한 이해와 연구가 가정학 분야에서 보다 활발히 이루어져야 할 필요가 있다.

1998년 이후 3년마다 한 번씩 우리나라에서는 국민건강 및 보건의식행태조사와 식품섭취 및 영양평가를 통합하여 국민건강·영양조사를 실시하고 있는데, 이 조사를 통하여 수집된 자료는 우리나라민의 대표적이고 일반적인 건강 및 식품영양섭취, 그에 관련된 요인 등에 대한 특성을 파악할 수 있는 자료로 사용되어왔다(보건복지부, 1999; 보건복지부, 2002). 그러나 대부분 이러한 자료는 변수와 내용이 방대하여 우리가 원하는 정보를 찾아내는 일을 보다 어렵게 만드는 경우가 많다. 이러한 상황에서 1990년대에 이르러 대용량의 자료로부터 이들 자료 내에 존재하는 유용한 지식을 찾아내는 데이터마이닝(data mining) 기법이 출현하게 되었고, 통계학, 전산학, 경영학 등의 분야에서 활발히 연구되고 있는 실정이다.

이에 본 연구에서는 2001년도 국민건강·영양조사의 외식관련 자료에 데이터마이닝 기법을 적용하여 외식비율이 상대적으로 높은 사람들과 관련된 변수들을 파악한 후 이 변수들을 이용하여 이들 대상을 파악할 수 있는 예측모형을 개발하고자 한다.

II. 연구방법

본 연구는 2001 국민건강·영양조사의 자료를 Rud(2001)가 제안한 데이터마이닝 기법을 이용하여 분석하였다

1. 분석에 사용한 자료와 변수

본 연구는 2001년 국민건강·영양조사 자료를 이용하였다. 국민건강·영양조사는 4가지의 조사종류로 나눌 수 있는데 이는 국민건강면접조사, 보건의식행태조사, 검진조사, 영양조사이다. 국민건강면접조사는 2000년 인구주택총조사에서 통계청 표본 조사구를 제외한 나머지 조사구에서 지역 층과 행정구역을 각각 층화변수로 한 2월 층화추출방법과 각 층의 조사구수에 비례하도록 표본 조사구를 배분하는 비례배분방법을 사용하여 총 600개 조사구를 추출하고 각 조사구에서 표본가구를 22가구 선정하여 총 13,200가구가 선정되었다. 표본인구는 표본 가구에 3개월 이상 거주하고 있는 혈연가구원들을 그리고 5인 이상의 비혈연 가구원 구성 가구에서는 가구주만을 포함하여 약 40,000명이었다. 보건의식행태조사, 검진조사, 영양조사는 총 조사구의 1/3인 200개 조사구를 계통 추출하여 실시되었다.

본 연구에서는 국민건강면접조사, 보건의식행태조사, 검진조사 그리고 영양조사에 포함되어 있는 자료 중 외식과 관련된 것으로 예측되는 관련 변수들을 추출하여 분석을 수행하였다. 200개 조사구에서 추출한 가구원들 중 19세 이상 성인 자료만을 고른 후 “외식횟수”와 “결혼여부” 변수가 결측치인 대상을 제외한 7,032명을 대상으로 데이터마이닝을 실시하였다. “외식횟수”는 본 연구의 가장 핵심적인 변수이고, “결혼여부”는 그 결측여부가 기타 변수의 결측과 함께 발생하는 경향을 보였으므로 두 변수 중 한 가지 변수에라도 답을 안 한 소비자는 분석대상에서 제외하였다. “외식횟수”의 경우 영양조사에서 “평상시 외식(매식)을 얼마나 자주합니까?”라는 설문으로 조사되었다.

대상을 선택하기 위해 데이터를 정리한 후 우선 로지스틱 회귀식(logistic regression)을 이용하여 종속변수(y)는 외식횟수 문항으로 하루 1회 이상을 1로, 하루 1회 미만을 0으로 변환하여 사용하였다. 데이터마이

닝은 광범위한 자료 중에서 외식을 예측하는 예측모형을 만드는 것이 목적이므로 외식여부에 영향을 주는 요인을 알아보기 위한 독립변수(x)로는 이용 가능한 변수를 가능한 많이 포함시켰다. 건강면접조사에서는 연령, 성별, 월 가구소득, 결혼여부, 학력, 직업활동분야, 종사상 지위를, 검진조사에서는 허리둘레를 포함시켰다. 보건의식행태조사에서는 주관적 건강상태, 이환횟수, 건강에 대한 염려도, 지난 1년 동안 느끼 슬픔·우울증, 건강관리실천, 흡연 여부, 평균 음주빈도, 체중조절여부, 땀에 젖고 숨이 가쁜 운동횟수, 일상생활 활동, 하루 평균 수면시간, 피로 인식도, 여가/휴가 보내는 방법, 스트레스 인식도를 포함시켰다. 영양조사에서는 결식이유, 간식횟수, 간식종류, 외식 시 느끼는 간의 정도, 튀긴 음식의 섭취횟수, 영양제 등의 복용동기, 임신 여부, 조사 1-2일 전의 아침, 점심, 저녁식사 여부, 비타민 또는 무기질의 복용여부, 건강보조식품 복용여부, 보신식품 복용여부, 보약 복용여부, 주거지역 등을 포함시켰다.

2. 분석 방법-변수의 탐색 및 변형, 선택

모든 통계분석은 SAS Version 9.1 (SAS Institute Inc., 2004)을 이용하였고, 국민건강·영양조사의 자료수집 가중치를 고려하여 분석하였다. 가중치는 각 사람이 뽑힐 확률의 역수로 정의되며, 국민건강·영양조사 자료에 각 사람의 가중치가 제시되어 있다.

1) 세분화(segmentation): 연속형 변수의 경우, 어떤 특정 값을 기준으로, 그 이상은 1, 그 이하는 0을 갖는 새로운 이진(dichotomous) 변수를 만들었다. 범주형 변수인 경우에는, 여러 범주를 모아서, 그 범주에 들어가면 1, 아니면 0을 갖는 새로운 이진 변수를 만들었다. 이러한 세분화 작업에서 기준이나 범주는 종속변수인 y를 가장 잘 예측할 수 있도록, y와의 연관성을 고려하여 선택하였다.

2) 결측값의 처리: 결측치가 있다는 것 그 자체만으로도 y와 상관성이 높을 수가 있으므로, 결측치가 많은 경우에는 일단 결측치가 있는지 없는지의 여부를 나타내주는 새로운 이진 변수를 만들었다. 그리고 원래 변수의 결측치는 남은 값들의 평균값으로 대체하였다.

3) 변환(transformation): 설명변수, 즉 x를 변환하

는 목적은 y 에 대한 예측력을 높이는 데 있다. 각각의 연속형 설명 변수에 대하여 총 15가지의 변환을 하여 15개의 새로운 변수($x^2, x^3, \sqrt{x}, \ln x, \frac{1}{x}$ 등)를 만들었다. 이렇게 각각의 연속형 변수에 대하여 총 16개의 변수(1개의 원래변수 + 15개의 변환변수)와 한 개의 결측치를 나타내는 변수 중 예측력이 높은 2개씩을 골라내었다.

4) 회귀식 추정: 변환을 하고, 이 중 각 설명변수에 대하여 유의한 변수 2개씩을 골라낸 후, 이들 변수들을 모두 이용하여 다중 로지스틱 회귀(multiple logistic regression) 분석을 실행하여 최종모형을 선택하였다. 본 논문에서는 Rud(2001)가 제안한 방법대로, 단계적 선택법(stepwise), 후진소거법(backward), 최적조합 선택법(score 방법)을 모두 사용하였다. 우선, 단계적 선택법과 후진소거법을 각각 실행했을 때 한 번이라도 선택된 모든 변수들을 고른 후에 최적조합선택법으로 최종 변수들을 선택하였다.

3. 모형식

위의 두 과정을 거쳐 얻고자 하는 최종 모형형태는 다음과 같다.

p 개의 설명변수(x)가 있다고 가정하고, $P(y=1|x_1 \dots x_p)$ 를 $x_1 \dots x_p$ 가 주어졌을 때의 종속변수 $y=1$ 일 확률이라 하면 로지스틱 회귀모형은 다음과 같다.

$$\ln \frac{P(y=1|x_1 \dots x_p)}{1 - P(y=1|x_1 \dots x_p)} = \alpha + \beta_1 x_1 + \dots + \beta_p x_p$$

4. 심분위 분석

얻어진 모형이 실제 자료의 소비자에측을 얼마나 잘 하는지를 평가하는 방법으로는 심분위분석법(decile analysis)(Rud, 2001)이 사용되었다. 심분위분석을 하기 위해서는, 먼저 각각의 사람들을 모형에 의한 예측값의 내림차순으로 정리하여 사람들의 수가 비슷한 크기가 되게 전체를 10개의 집단으로 나눈 후 각 집단에서 모형의 예측 값과 실제 y 가 1인 외식을 하루에 1회 이상 하는 사람들의 비율을 비교하고 리프트(lift) 측정법을 이용하여 모형을 평가하였다. 본 연구에서의 리프트 값은 각각의 심분위수 집단에서 실제 하루 1회 이상 외식하는 사람들의 비율을 실제 하루 1회 이상 외식하

는 사람들의 전체 평균비율로 나눈 후 100을 곱해 준 것이다. 예를 들어 어느 집단의 리프트 값이 2라면 이 집단에 속하는 사람들의 외식비율은 전체 사람들의 외식비율보다 2배 크다는 것을 의미한다.

일반적으로 로지스틱 회귀모형은 예측력은 높으나, 블랙박스와 같다. 즉 모형에 들어가는 변수들은 함께 모형에 들어갈 때만 그러한 예측력을 발휘하는 것이다. 한편 이러한 회귀모형에 의한 예측을 실제로 이용할 때는 정말 중요한 변수가 무엇인지 또한 얼마나 중요한지를 알고자한다. 그래서 많이 사용되는 방법이 중요한 변수들에 대하여 심분위분석을 하는 것이다. 본 연구에서도 이 분석을 통해 각 10개의 집단안의 대상자들의 특징을 파악해 보고자 하였다.

III. 결과 및 해석

1. 추정된 회귀식 모형

유의한 변수들이 채택되어 추정된 최종 회귀식은 다음과 같다.

$$\ln \frac{P(y=1|x_1 \dots x_{18})}{1 - P(y=1|x_1 \dots x_{18})} = -2.0897 + 8.11 \times 10^{-8} x_1^3 + \dots + 0.2687 \times x_{18}$$

여기서 x_1 은 나이이고, x_{18} 은 튀김음식섭취를 나타내는 변수이다. 얻어진 최종모형의 변수설명은 <표 1>에 나타내었다. 변수이름은 자료의 코딩설명서에 있는 국민 영양조사의 문항묘사를 사용하였고 범주형 변수 각 값의 설명은 여러 개를 묶어 1개의 값으로 만든 경우 그 특징을 묘사하였다.

우리나라 국민에서 하루 1회 이상의 외식하는 소비자의 특징을 추정된 회귀식으로 살펴보면 대도시에 거주할수록, 출근하는 직장을 가진 경우, 전일제 근로자일수록, 외식을 하루에 1회 이상하고 있었다. 연령이 낮을수록 외식을 자주 하는 경향이 있었고 남자들이 여자들보다, 결혼 상태는 미혼일수록 외식을 더 자주하는 경향을 보였다. 피로도 와 스트레스 인식수준이 높을수록 외식경향이 높은 것으로 나타났다. 이환겂수, 스트레스 인식 등의 변수에 대해 결측치를 가질수록 외식을 하는 경향을 보였고 여가 보내는 법, 결식이유, 외식을

<표 1> 우리나라 국민 중 외식을 하루 1회 이상 하는 소비자를 예측하기 위한 요인

변수 설명		회귀계수	표준오차	p-value
종속변수	외식횟수	하루1회이상=1, 하루1회미만=0		
	기울기		-2.0897	0.00450 <.0001
독립변수	연령	연령 ³	8.118E-8	9.641E-9 <.0001
	이환 개수 결측	결측=1, 결측 아닐때=0	0.2549	0.00173 <.0001
	거주지역	도시=1, 중소도시와 시골=0	0.5012	0.00155 <.0001
	성별	남자=1, 여자=0	0.7170	0.00157 <.0001
	결혼여부	미혼=1, 기혼=0	0.8129	0.00195 <.0001
	직업활동분야	고위 임직원/전문가/기술공/사무종사자/기능원/군인=1 서비스/판매/농업/기계/단순노무/학생/주부/무직=0	0.1631	0.00193 <.0001
	종사상 지위	전일근로자=1, 시간제근로자/무급가족종사자=0	0.4141	0.0019 <.0001
	피로인식도	언제나 느낌=1, 때때로, 거의/전혀 안느낌=0	0.2174	0.00201 <.0001
	여가법 결측	결측=1, 결측 아닐때=0	-0.7413	0.00209 <.0001
	스트레스인식 결측	결측=1, 결측 아닐때=0	0.8108	0.00310 <.0001
	스트레스인식도	많이 느낌=1, 조금 느낌/거의 느끼지 않음=0	0.2264	0.00169 <.0001
	결식이유 결측	결측=1, 결측 아닐때=0	-0.2323	0.00176 <.0001
	조사2일전 점심결식	섭취했으면=1, 섭취안했으면=0	0.5576	0.0041 <.0001
	결식이유	늦잠/소화 안돼서/간식을 먹어서=1 식욕없어/체중조절/습관=0	-0.4357	0.00277 <.0001
	간식횟수	하루3회이상=1, 하루3회미만=0	0.5103	0.00245 <.0001
	간식종류	빵/라면/튀김=1, 과자/떡/국수/과일/우유=0	0.4852	0.00169 <.0001
	외식음식간 결측	결측=1, 결측 아닐때=0	-17.5160	5.1894 0.0007
	튀김음식섭취	주1회 이상=1, 주 1회 미만=0	0.2687	0.00161 <.0001

식의 간에 대한 변수에 대해 결측치를 가지지 않은 경우가 그 이유는 알 수 없으나 외식을 자주하는 경향을 보였다. 점심결식을 하지 않을 때 외식을 좀 더 자주하고 있었고 이는 외식을 통해 점심을 함으로써 나타나는 결과로 보인다. 식욕이 없거나 습관적으로 결식을 한 사람이 외식을 더 자주 하는 경향을 보였다. 간식횟수가 많을수록 외식을 더 많이 하고 있음을 알 수 있었고 간식을 빵/라면/튀김 등의 식사대용식품으로 하는 사람이 외식을 더 자주하는 경향을 보였다. 튀김 음식섭취를 주 1회 이상 하는 사람이 외식을 하루 1회 이상을 하는 것으로 나타나 외식을 자주하는 여부와 튀김 섭취가 관련을 보이고 있었다.

2. 추정된 회귀식에 대한 심분위분석 결과

<표 2>의 결과는 이 모형에 의한 예측 값을 내림차순으로 정리하여 10개의 그룹으로 나눈 것이며 첫 번째

줄은 상위 10%에, 그 이후의 줄은 각각의 심분위 그룹에 속하는 사람들이다. 모형의 예측 값이란 모형에 따라 계산된 종속변수 y를 말하며 이는 외식을 하루 1회 이상 실시할 확률을 의미한다. 예를 들어 첫 번째 심분위에 들어가는 사람들은 모형에 의한 예측 값이 가장 높은 사람들의 그룹으로서 우리모형에서는 외식을 하루 1회 이상 실시할 확률이 0.749임을 보여준다.

두 번째 열의 수는 각 그룹에 속하는 사람들의 숫자이다. 이 수가 정확히 같지 않은 이유는 각 사람의 가중치가 조금씩 다르기 때문이다. 세 번째 열은 각 분위수 그룹에 속하는 사람들에게 대한 모형의 예측 값들의 평균 값이다. 첫 번째 심분위 그룹에서의 예측 값이 가장 높으며, 그 다음부터 차례로 낮아지는 것을 볼 수 있다. 네 번째 열은 각 그룹에서 실제로 외식 하루 1회 이상 한다고 답한 사람들의 비율을 가중치를 고려하여 구한 값이다. 이 또한 첫 번째 심분위에서 가장 높고, 그 이

〈표 2〉 외식소비자 예측 회귀모형의 적합성 검증

십분위 ¹⁾	대상수(명) ²⁾	모형의 예측값 ³⁾	실제 하루 1회 이상 외식하는 사람의 비율 ⁴⁾	실제 하루 1회 이상 외식하는 사람의 수(명) ⁵⁾	리프트(lift) ⁶⁾
1	586	0.749	0.721	423	241
2	652	0.604	0.578	377	193
3	679	0.508	0.546	371	182
4	745	0.419	0.442	329	148
5	730	0.319	0.366	267	122
6	732	0.219	0.204	149	68
7	737	0.140	0.114	84	38
8	752	0.037	0.023	18	8
9	669	0.000	0.000	0	0
10	750	0.000	0.000	0	0
평균		0.299	0.299	2017	

1) 분위수 집단은 내림차순임.

2) 대상수는 각 집단에 속하는 사람들의 숫자임. 이 수가 정확히 같지 않은 이유는 각 사람의 가중치가 조금씩 다르기 때문임.

3) 모형의 예측값이란 모형에 따라 계산된 종속변수 y . 즉 외식을 하루 1회 이상 실시할 확률임. 각 분위수 집단에 속하는 사람들에게 대한 모형의 예측값들의 평균임.

4) 실제로 외식을 하루 1회 이상 한다고 답한 사람들의 비율을 가중치를 고려하여 구한 값임.

5) 각 분위수 집단에서 실제 외식을 한 사람의 수임.

6) 리프트(lift) 측정법에 의한 리프트 값임($100 \times$ 해당 집단의 실제 하루 1회 이상 외식하는 사람들의 비율/전체의 하루 1회 이상 외식하는 사람들의 평균비율).

후로 차례로 감소하는 것을 볼 수 있는데, 이는 본 연구에서 개발한 모형의 우수성을 입증하고 있는 것이다. 다섯 번째 열은 각 그룹에서 실제 외식을 한 사람의 수이고, 여섯 번째 열은 리프트 측정법에 의한 리프트 값을 보여주는데 예를 들어 첫 번째 십분위 그룹의 241은 $(0.721/0.299) \times 100$ 에 의하여 얻어진다. 리프트는 모형을 만들지 않고 임의로 개인에게 접근했을 때(random approach) 또는 평균적인 모형의 적합성보다, 만들어진 모형이 얼마나 잘 적용되는지를 각 십분위(decile)에서 알아보는 것이다. 〈표 2〉의 첫 번째 십분위 집단의 경우, 평균보다 외식을 하는 사람이 2.41배 더 있음을 의미한다.

3. 주요 변수에 대한 십분위 분석 결과

〈표 3〉은 개발된 모형의 주요 독립변수들의 평균값을 각 십분위별로 보여주는데, 이러한 표를 통하여 각 십분위 집단에 속하는 사람들의 특성을 알 수 있다. 결과를 보면 연령이 낮을수록 첫 번째 십분위에 속하여, 외식을 자주하는 그룹에 속하는 것을 보여준다. 대도시

에 거주할수록, 전일제 근로자, 남자 그리고 미혼일수록, 점심결식을 하지 않을수록, 간식을 많이 할수록, 튀김음식을 자주 먹을수록 하루에 1회 이상 외식을 하는 그룹에 속했다.

IV. 논의 및 결론

본 연구에서는 2001년도 국민건강·영양조사의 외식관련 자료에 데이터마이닝 기법을 적용하여 우리나라 성인 중 하루 1회 이상 외식을 하는 소비자를 예측하기 위한 모형을 개발하고자 하였다. 본 연구의 주요 결과를 요약하고 논의하면 다음과 같다.

첫째, 하루 1회 이상 외식을 하는 소비자를 예측하기 위하여 총 18개의 설명변수를 포함하는 로지스틱 회귀식 모형이 개발되었다. 추정된 회귀식에 대한 십분위수 분석법을 실시한 결과, 본 연구의 모형이 대상자의 실제 외식비율을 잘 반영하고 있는 것으로 나타났다. 개발된 모형은 어떤 경향을 가진 사람들이 좀 더 자주 외

〈표 3〉 각 심분위에 속한 대상자의 주요 특징

심분위	나이 ¹⁾	성별	거주지역	결혼여부	직업활동 분야	피로인식	스트레스 인식	간식횟수	간식종류	튀김섭취 횟수
1	30.9	0.76	0.84	0.69	0.67	0.26	0.40	0.20	0.60	0.61
2	35.5	0.76	0.76	0.38	0.59	0.18	0.34	0.14	0.37	0.42
3	40.5	0.77	0.69	0.20	0.42	0.17	0.27	0.08	0.28	0.31
4	39.0	0.65	0.53	0.21	0.44	0.14	0.26	0.08	0.21	0.32
5	41.8	0.40	0.58	0.17	0.15	0.15	0.25	0.10	0.20	0.28
6	45.8	0.23	0.59	0.06	0.04	0.15	0.22	0.11	0.17	0.24
7	47.9	0.08	0.61	0.01	0.01	0.15	0.20	0.05	0.08	0.14
8	43.6	0.42	0.43	0.17	0.17	0.17	0.21	0.09	0.19	0.18
9	54.4	0.40	0.57	0.05	0.05	0.26	0.30	0.06	0.15	0.14
10	59.3	0.06	0.43	0.01	0.01	0.18	0.21	0.03	0.06	0.05
평균	43.9	0.45	0.60	0.19	0.25	0.18	0.27	0.09	0.23	0.27

1) 모형에는 연령의 삼승 즉 연령³의 형태로 연속형 독립변수로 포함되었으나 대상의 특징 파악을 위해 본 표에서는 원래 연령의 평균으로 보고 하였음.

식을 하는지에 대한 일반적인 정보를 주고 또한 모형에 포함된 몇 가지 항목을 개인에게 질문했을 때 그가 하루 1회 이상 외식을 할 확률을 예측할 수 있게 한다.

둘째, 개발된 모형은 연령이 낮을수록, 대도시에 거주할수록, 출근하는 직장을 가진 경우, 전일제 근로자일수록, 남자인 경우, 미혼인 경우에 하루에 1회 이상 외식을 하는 경향을 가지고 있음을 보여주었다. 또한 피로도도와 스트레스 인식 수준이 높을수록 외식경향이 높은 것을 보여주었다. 연령이 낮을수록 더 외식을 자주하는 경향은 젊은 연령이 외부활동이나 사회활동이 많기 때문에 나타난 결과로 판단된다. 대도시에 거주할수록, 출근하는 직장을 가진 경우, 전일제 근로자일수록, 남자, 미혼인 경우에 외식 경향이 높은 것은 대도시의 전일제 직장인들이 가정 밖에서 사회생활을 하면서 점심이나 저녁을 밖에서 해결하기 때문에 생기는 현상인 것으로 해석된다.

셋째, 개발된 모형에서 개인적인 식습관 패턴과 외식 빈도와와의 관계를 살펴보면 간식횟수가 많을수록 외식을 더 많이 하고 있음을 알 수 있었고, 간식을 빵/라면/튀김 등의 식사대용식품으로 하는 사람이 외식을 더 자주하는 경향을 보였는데, 이는 과자, 과일, 우유 등의 간편한 간식을 하는 경우에는 외식이 아닌 식품자체를 구매하여 섭취할 수 있어서라고 생각된다. 튀김 음식섭취를 주 1회 이상 하는 사람이 외식을 자주하는 경향을

보이고 있었는데 외식을 하다보면 튀김류를 많이 먹게 되는 것인지 튀김을 좋아하는 사람이 외식을 자주 하는 것인지의 인과관계는 알 수 없으나 튀김 섭취에 관한 질문을 통해 외식을 자주하는 사람을 예측할 수 있다는 것은 매우 흥미로운 사실로 보인다.

네째, 개발된 모형에 포함된 주요 설명변수들의 평균값을 각 심분위 집단별로 계산하여 각 집단에 포함된 사람들의 특성을 알아본 결과 나타난 외식을 자주하는 사람들의 특징은 개발된 모형 자체가 보여주는 결과와 유사했다. 즉 연령이 낮을수록, 대도시에 거주할수록, 전일제 근로자, 남자, 미혼일수록, 점심결식을 하지 않을수록, 간식을 많이 할수록, 튀김음식을 자주 먹을수록 하루에 1회이상 외식을 하는 집단에 속하는 것으로 나타났다.

본 연구의 결과는 정부의 보건부서가 외식과 관련된 식품영양정책을 수립하는데 고려해야할 중요한 정보를 제시하고 있다. 사회생활, 직장생활과 관련된 요인들이 외식을 더 자주하게 만든다면 맛벌이 가정이 늘고 많은 사람들이 사회생활을 하면서 더 많은 사람들이 외식을 할 것으로 예측된다. 이에 외식을 통한 지방섭취량이 가정식이나 급식을 통한 양보다 많다는 외국의 보고(O'Dwyer, Gibney, Burk, & McCarthy, 2005)와 본 연구결과에서의 간식과 튀김을 자주 섭취하는 사람들이 외식을 더 자주한다는 결과를 볼 때, 이들을 통해

섭취되는 식품 및 영양 그리고 만성질병에 대한 연구와 건강한 웰빙 외식음식개발에 대한 더욱 깊은 관심이 요구된다.

본 연구 결과는 또한 소비자 교육 및 영양교육을 위한 대상선정뿐 아니라 맞춤형 교육내용 개발에 이용될 수 있으리라 기대된다. 앞으로 보다 더 세분화된 자료에 대하여 유사한 분석이 이루어진다면 관련 교육 대상자 선정 및 교육내용 개발에 더욱 유용한 정보를 산출할 수 있으리라 생각된다. 예를 들면, 20대 미혼 여성들에 대해서만 분석을 하거나, 대도시에 사는 중년 남성들만을 대상으로 분석을 한다면 이는 각 집단별 특징의 파악을 가능하게 하여 세분화된 대상자들에게 각종 식생활 및 소비 교육을 효과적으로 수행할 수 있게 할 것이다.

본 연구를 통하여 개발된 모형과 모형에 포함된 주요변수의 심분위 집단별 평균값은 외식을 자주할 가능성을 가진 고객을 찾기 위해 외식업체가 어떤 사람을 대상으로 마케팅을 하는 것이 효과적인지를 제시하고 있다. 외식시장의 공급이 수요를 초과하고 환경에의 적응이 기업의 생존을 결정하는 중요한 변수로 자리 잡은 현대 시장에서 소비자를 그 성격에 따라 분류하여 접근하고자 하는 시장 세분화의 노력에도 본 연구에서 개발된 모형이 이용될 수 있으리라 생각된다.

본 연구는 이용한 이차 자료의 조사 대상자 선정이 우리나라 국민의 특징을 대표할 수 있도록 층화추출법을 이용하여 이루어졌으므로 그 결과를 일반화시킬 수 있다는 커다란 장점을 가진다. 그러나 종속변수인 외식 횟수에 대한 질문에 대상자들이 보통 외식과는 약간 성격이 다른 급식의 경우도 포함시켜 응답하였을 가능성이 있다는 연구의 제한점을 가지고 있지만 그를 구분하는 문항은 없었으므로 그에 대한 고찰은 가능하지 않았다. 또한 본 연구에서 사용한 문항들은 대부분 대상자 본인의 주관적인 응답이므로 이로 인한 오차가 존재하나 조사연구의 특성상 이를 고려하기는 어려웠다.

국민건강·영양조사 자료에서 외식 빈도 변수를 이용한 연구가 그 동안 발표된 적이 없었다는 사실이 본 연구의 의의를 더한다. 국민건강·영양조사에서는 외식 빈도뿐만 아니라 일일 섭취하는 각 끼니 및 간식에 대한 식사장소를 조사하고 있다. 그러나 이렇게 돈과

시간이 투자되어 자료는 모아지지만 이러한 변수에 대한 활용은 그동안 미진했었다. 최근 발표된 한 학위 논문(김지혜, 2005)에서 최초로 일일 섭취 음식의 식사장소 변수를 이용한 분석이 발표되었을 뿐이다. 외국의 경우, 국민을 대상으로 한 대규모 조사에 포함되어 있는 외식관련 변수를 이용한 연구가 보다 활발히 이루어져 각종 유용한 정보를 산출하고 있다는 점을 감안할 때(Binkley, Eales, & Jekanowski, 2000; Borrud, Enns, & Micke, 1996; Byrne, Capps, & Saha, 1998; Guthrie & Lin, 1997; Ham, Yoon, & Leong, 2003; Hu & McAleer, 1997; Min, Fang, & Li, 2004; Nayga, 1996; O'Dwyer *et al.*, 2005; Pol & Park, 1995; Stewart & Yen, 2004) 우리나라에서도 국민건강영양조사 자료에 대한 보다 다각적인 분석이 필요하다고 생각된다. 국내의 그 동안의 외식 소비관련 연구에 주로 쓰인 이차자료는 통계청의 도시가계 조사 자료이었다(류정순, 박혜련, 1996; 박혜련, 이경희, 류정순, 1997; 서정희, 이성림, 홍순명, 2003; Ham, Hwang & Kim, 2004; Lee, 2001). 그러나 국민건강·영양조사 자료의 외식변수를 이용한다면 건강과 관련된 기타 생활 습관 및 영양섭취상태와 관련하여 더욱 다양한 연구를 수행할 수 있으리라 생각된다.

■참고문헌

- 김지혜(2005). 가정식과 학교급식으로 점심식사를 한 초·중·고 재학생의 식사평가: 2001 국민건강영양조사 자료를 이용하여. 서울대학교 대학원 식품영양학과 생활과학석사학위논문.
- 노정미, 유영상(1989). 서울지역 대학생의 외식실태 및 기호도 조사연구(I). 대한가정학회지, 27(2), 65-74.
- 류정순, 박혜련(1996). 계층별 소비패턴 분석을 통한 외식비의 빈곤선 설정 -1993년 도시근로자 가계를 중심으로-. 한국가정관리학회지, 14(2), 277-287.
- 모수미(1986). 외식산업의 발달이 국민영양 및 식생활에 미치는 영향. 한국영양학회지, 19(2), 120-125.
- 박혜련, 이경희, 류정순(1997). 도시가계조사 자료를 이용한 소득계층별 식생활 패턴분석. 지역사회영

- 양학회지, 2(4), 633-646.
- 보건복지부(1999). 1998년도 국민건강 영양조사 - 총괄보고서-
- 보건복지부(2002). 2001년도 국민건강 영양조사 - 총괄보고서-
- 서정희, 이성림, 홍순명(2003). 도시가계 외식비 지출에 관한 시계열 분석: 1982년부터 2002년. 대한영양사협회 학술지, 9(2), 149-158.
- 유정림, 박동연(2001). 가족생활주기에 따른 외식행동의 차이. 대한가정학회: 학술대회지, 대한가정학회 제 54차 정기총회 및 추계학술대회, 112.
- 정은정, 김정숙(2001). 제주도 대학생의 외식행동에 관한 연구. 대한가정학회지, 39(7), 21-36.
- 주나미(2004). 자녀들의 패스트푸드점 선호에 대한 어머니의 인식 및 견해 연구. 대한가정학회지, 42(5), 1-10.
- 통계청(2004). 통계자료: 시도/산업별 도소매업 총괄. <http://www.nso.go.kr>. 2005.10.3 접속.
- 통계청(2005). 통계자료: 품목별 가구당 월평균 가계수지. <http://www.nso.go.kr>. 2005.10.3 접속.
- Binkley, J. K., Eales, J., & Jekanowski, M. (2000). The relation between dietary change and rising US obesity. *International Journal of Obesity*, 24(8), 1032-1039.
- Borrud, L., Enns, S. W., & Mickle, S. (1996). What we eat in America: USDA surveys food consumption changes. *Food Review*, 19(3), 14-19.
- Byrne, P. J., Capps, O. Jr., & Saha, A. (1998). Analysis of quick-serve, mid-scale, and up-scale food away from home expenditures. *International Food and Agribusiness Management Review*, 1(1), 51-72.
- Guthrie, J. F., & Lin, B. H. (1997). Trends in consumption of food away from home 1977-1995 and implications for achieving dietary guideline for fat intake. *Journal of the American Dietetic Association*, 97(9), A88.
- Ham, S., Hwang, J. H., & Kim, W. G. (2004). Household profits affecting food-away-from-home expenditures: a comparison of Korean and US households. *International Journal of Hospitality Management*, 23(4), 363-379.
- Ham, S., Yoon, B. J., & Leong, J. (2003). Demographic and socioeconomic characteristics affecting food away from home expenditure patterns of the US households. *Praxis-The Journal of Applied Hospitality Management*, 5(2), 22-34.
- Hu, B., & McAleer, M. (1997). A probit analysis of behaviour in rural China. *Mathematics and Computers in Simulation*, 43(3-6), 527-534.
- Lee, Y. G. (2001). Consumption patterns of elderly households: are they different between younger and older elderly in Korea? how do they differ between Korea and the US? *Consumer Interest Annual*, 47, 1-11.
- Min, I., Fang, C., & Li, Q. (2004). Investigation of patterns in food-away-from-home expenditure for China. *China Economic Review*, 15(4), 457-476.
- Nayga, R. M. Jr. (1996). Dietary fiber intake away-from-home and at-home in the United States. *Food Policy*, 21(3), 279-290.
- O'Dwyer, N. A., Gibney, M. J., Burke, S. J., & McCarthy, S. N. (2004). The influence of eating location on nutrient intake in Irish adults: implications for developing food-based dietary guidelines. *Public Health Nutrition*, 8(3), 258-265.
- O'Dwyer, N. A., McCarthy, S. N., Burke, S. J., & Gibney, M. J. (2004). The temporal pattern of the contribution of fat to energy and of food groups to fat at various eating

locations: implications for developing food-based dietary guidelines. *Public Health Nutrition*, 8(3), 249-257.

Pol, L. G., & Pak, S. (1995). Consumer unit types and expenditures on food away from home. *The Journal of Consumer Affairs*, 29(2), 403-428.

Rud, O. P. (2001). *Data Mining Cookbook: Modeling Data for Marketing, Risk, and Customer Relationship Management*. New

York: John Wiley & Sons.

SAS Version 9.1 [Computer Software]. (2004). Cary, North Carolina: SAS Institute Inc.

Stewart, H., & Yen, S. T. (2004). Changing household characteristics and the away-from-home food market: a censored equation system approach. *Food Policy*, 29(6), 643-658.

(2005년 10월 19일 접수, 2005년 11월 3일 채택)