

서비스 수요조사와 분류모형을 이용한 수요예측

김윤중¹⁾ 김용철²⁾

요약

수요예측은 모든 생산적 활동을 수립하기 위한 기반이 되기 때문에 수요가 어느 정도 발생할 것인가에 대한 방향성에 대하여 파악하려고 일반적으로 설문조사를 이용하지만 무응답 및 불성실한 응답으로 인하여 설문응답 자료만으로 수요를 예측하기에는 부족하다. 따라서 수요와 관련 있는 변수를 이용한 분류모형으로 설문조사의 수요예측을 보정하고자 하였다. 본 논문에서는 설문조사를 통하여 평가할 수 있는 직접적인 수요와 통계적 모형을 이용한 간접적 수요를 혼합하여 서비스 수요를 예측하는 혼합 모형을 제시하고자 한다.

주요용어: 로지스틱 회귀모형, 서비스 수요조사, 혼합모형.

1. 서론

어떤 기준에 의하여 수요를 예측한다는 것은 미래에 대한 불확실성을 줄이고 수행하고자 하는 정책에 대한 의사를 결정하는데 매우 중요한 역할을 한다. 일반적으로 정해진 어떤 상품에 대하여 구매자가 구입하고자 하는 수량을 수요라 정의하며 수요량은 상품의 단위가격이 하락하면 증가하고 단위가격이 상승하면 감소하는 특성이 있다. 수요는 구매자의 욕망상태, 소득, 가격 등에 의하여 결정된다. 수요예측은 모든 생산적 활동을 수립하기 위한 기반이 되기 때문에 대단히 중요하며 수요가 어느 정도 발생할 것인가에 대한 방향을 제시하여야 한다. 직접적으로 시장의 수요를 조사하기위해 일반적으로 설문조사를 이용하여 수요를 파악하지만 무응답 및 불성실한 응답으로 인하여 설문응답 자료만으로 수요를 예측하기에는 신뢰성이 부족하다. 또한 수요의 종류에 따라 설문조사 시 기술적 문제 때문에 외부 환경요인 및 수단적 요인의 현상이 고정되었다고 가정하고 수요를 예측하는 경우가 많다. 따라서 이들 요인에 어떤 변동이 나타나면 예측 값을 재검토해야 한다. 본 연구는 기관에서 수행하고 있는 서비스에 대한 수요예측에 초점을 두고 서비스 수요를 예측하고자 한다. 일반적으로 수요를 예측하는데 설문조사한 응답 자료만을 이용한다면 설문조사의 비 표본오차의 신뢰성에 의하여 결정된다. 따라서 수요에 영향력이 있는 변수를 이용한 로지스틱 회귀모형으로 보정함으로써 이러한 오차를 줄이려고 하였다.

1) (449-714) 교신저자. 경기도 용인시 처인구 삼가동 470, 용인대학교 컴퓨터정보학과, 교수.

E-mail: younkim@yongin.ac.kr

2) (449-714) 경기도 용인시 처인구 삼가동 470, 용인대학교 컴퓨터정보학과, 부교수.

E-mail: yckim@yongin.ac.kr

설문조사를 이용한 수요조사는 김범환 등 (2000)은 초고속장비 관련 시장 예측에 있어서 예측대상의 특성을 고려하여 문항을 설정하는 방법의 하나인 전문가에 의한 델파이 방법을 적용하여 응답의 신뢰성을 높이는 설문조사 기법을 소개하였다. 또한 남경희 등 (2001)은 정밀측정기술 수요조사에 대하여 분석하였고 설문응답에 대한 신뢰성을 높이기 위해 응답자에게 확인 전화조사를 병행하였다. 그러나 설문조사의 응답에 대한 신뢰성의 평가는 대부분은 주관적인 관점에서 해석하게 된다. 그러므로 직접적인 수요예측 방법인 설문조사 방법에 대하여 간접적으로 보정해 줄 수 있는 혼합모형이 필요하다.

본 논문에서는 설문조사를 통하여 평가 할 수 있는 직접적인 수요와 통계적 모형을 이용한 간접적 수요를 혼합하여 서비스 수요를 예측하는 혼합 모형을 제시하고자 한다. 다음 절에서는 서비스 수요예측에 필요한 설문문항과 수요 분류에 사용이 가능한 로지스틱 회귀모형에 대하여 논하고, 3절에서는 설문조사와 로지스틱 회귀모형을 적용한 서비스 수요 예측에 대한 예제를 논의하였다. 마지막 4절에서는 관련된 결론에 대해서 논의하였다.

2. 수요예측을 위한 설문항목 및 혼합모형

기관에서 수행하고 있는 서비스 수요를 예측하기 위한 설문조사에서 응답대상자의 정확한 서비스 수요를 알기 위하여 많은 항목을 설문하여야 하지만 특성상 항목의 수는 제한적이기 때문에 응답대상자들의 응답률 및 설문 자료의 효과성을 위하여 기본적인 특성만을 고려하여 본 연구에서는 서비스 수요 여부의 설문조사와 설문에 의한 서비스 수요의 신뢰성을 보정하기 위해 수요여부를 종속 변수로 하는 로지스틱 회귀모형을 이용하였다. 따라서 서비스 수요예측모형으로 설문조사와 모형의 가중치 합으로 혼합모형을 도출하였다.

2.1. 수요예측을 위한 설문 항목

서비스 수요 여부를 알기위한 직접적인 방법은 유효 모집단에 대하여 개인 대 개인, 우편 혹은 전화 인터뷰를 통해 수요 여부를 설문하는 방식을 사용하고 있다. 이런 조건 하에서 응답자들은 가상적 변화에 대해서 어느 정도 수요 의사가 있는지를 대답하게 된다. 하지만, 서비스 수요 여부는 응답자의 의사와 능력에 크게 의존하며 응답자들은 서비스 수요 여부에 대한 결정권자가 아니면 수요에 대한 불확실한 답변으로 서비스 수요 여부를 측정하는데 비 표본 오차가 발생할 수 있다. 서비스 수요에 대한 예측을 하기 위해 일반적으로 사용되어 지는 문항으로 다음과 같은 설문문항을 이용하여 자료를 얻는다.

문항) 귀하(귀 기관)는 업무에 관련된 서비스를 받을 의사가 있습니까?

- 1) 있다 2) 없다

직접 설문조사에서 발생할 수 있는 오차를 다음의 통계적 모형으로 보정하여 서비스 수요를 예측하고자 한다.

표 2.1: 설문 및 모형에 의한 수요분류

		실제수요여부		총합
		있다	없다	
설문에 의한 수요여부	있다	$n_{1C}^{(1)}$	$n_{1M}^{(1)} = n_1^{(1)} - n_{1C}^{(1)}$	$n_1^{(1)}$
	없다	$n_{2M}^{(1)} = n_2^{(1)} - n_{2C}^{(1)}$	$n_{2C}^{(1)}$	$n_2^{(1)}$
모형에 의한 수요여부	있다	$n_{1C}^{(2)}$	$n_{1M}^{(2)} = n_1^{(2)} - n_{1C}^{(2)}$	$n_1^{(2)}$
	없다	$n_{2M}^{(2)} = n_2^{(2)} - n_{2C}^{(2)}$	$n_{2C}^{(2)}$	$n_2^{(2)}$

2.2. 수요예측을 위한 혼합모형

설문조사와 로지스틱 회귀모형으로 산출된 예측수요와 실제수요를 비교하여 적중률이 높은 예측수요에 상대적으로 큰 가중치를 적용하여 다음과 같은 식으로 혼합 예측모형을 도출할 수 있다.

$$D_t = \omega_t D_t^{(1)} + (1 - \omega_t) D_t^{(2)}. \tag{2.1}$$

여기서 D_t 는 t 시점 기업의 수요확률, $D_t^{(1)}$ 은 t 시점 기업의 설문조사 수요, $D_t^{(2)}$ 은 t 시점 로지스틱 회귀모형의 수요확률 그리고 ω_t 은 t 시점에서의 가중치이다. 가중치의 결정은 다음과 같다.

위의 표 2.1에서 설문조사에서의 실질오류율(actual error rate)은

$$AER^{(1)} = \frac{n_{1M}^{(1)} + n_{2M}^{(1)}}{n_1^{(1)} + n_2^{(1)}} \tag{2.2}$$

이고 모형에서의 실질오류율(actual error rate)은

$$AER^{(2)} = \frac{n_{1M}^{(2)} + n_{2M}^{(2)}}{n_1^{(2)} + n_2^{(2)}} \tag{2.3}$$

이다. t 시점에서 설문조사의 실질오류율(actual error rate)과 모형의 실질오류율(actual error rate)을 이용하여 모형에 적용되는 가중치를 계산하면

$$\omega_t = \frac{1 - AER^{(1)}}{2 - AER^{(1)} - AER^{(2)}} \tag{2.4}$$

이다.

3. 수요예측 분석 사례

한국정보보호 산업협회에서는 2006년 8월 17일부터 2006년 10월 30일까지 정보보호 산업에 속한 사업체를 대상으로 실시하였고 그 조사에서 조사대상 모집단 200개 업체 중 유효 모집단 163개 업체와 유효 모집단에 포함된 2006년 8월 이후 실제로 발생한 업무수요 16개 업체를 대상으로 얻어진 결과를 업무수요에 대한 수요예측모형의 분석 사례로 제시하고자 한다.

표 3.1: 설문응답 업체와 실제로 수요가 발생한 업체의 분류

		실제로 업무수요가 있었던 업체		총합
		있다	없다	
설문응답에서 업무수요의 유무	있다	11 (6.75%)	53 (32.52%)	64 (39.26%)
	없다	5 (3.07%)	94 (57.67%)	99 (60.74%)
총합		16 (9.82%)	147 (90.18%)	163 (100.0%)

표 3.2: 로지스틱 회귀모형의 계수 추정 값 및 p -값

변수	계수 추정 값	p -값
기업부설 연구소인원	0.0579	0.0128
2005년 R&D 투자액	0.00224	0.0034
신지식 재산권 보유수	0.2105	0.0250

표 3.3: 모형으로 분류한 업체와 실제로 수요가 발생한 업체의 분류

		실제로 업무수요가 있었던 업체		총합
		있다	없다	
로지스틱 회귀모형에서 업무수요의 유무	있다	12 (7.36%)	24 (14.72%)	36 (22.09%)
	없다	4 (2.45%)	123 (75.46%)	127 (77.91%)
총합		16 (9.82%)	147 (90.18%)	163 (100.0%)

3.1. 설문조사에서의 업무수요 분류표

설문조사에서의 업무수요 분류표는 다음의 표 3.1과 같다.

설문응답에서의 업무수요가 있는 업체가 전체 163개 업체 중 64개 업체로 나타났으며 그 중 실제로 신청한 업체가 11개 업체로 17.19%가 실제 업무수요가 있었던 것으로 나타났다.

3.2. 로지스틱 회귀모형에서의 업무수요 분류표

사업체의 수요구조를 알 수 있는 부가적인 정보로써 기업부설연구소인원, 2005년 R&D 투자액, 신지식재산권보유수등을 로지스틱 회귀모형의 독립변수로 적용하여 계수를 추정 한 결과는 다음과 같다.

로지스틱 회귀모형으로 분류분석을 하는 경우에 종속변수인 업무수요가 있다는 확률이 일반적으로 사전지식이 없는 경우 0.5보다 크면 수요가 있다고 판정한다. 본 논문에서는 2006년 이후 실제로 평가인증을 신청한 개별 업체를 조사하여 수요 적중률이 최대가 되는 판정 기준 확률 값이 0.6으로 산출되었다. 유효 모집단 163개 업체를 대상으로 로지스틱 회귀모형을 이용한 업무수요에 대한 분류표는 표 3.3과 같다.

표 3.3에서 업무수요가 있는 업체가 전체 163개 업체 중 36개 업체로 나타났으며 그 중

표 3.4: 혼합 모형으로 분류한 업체와 실제로 수요가 발생한 업체의 분류

		실제로 업무수요가 있었던 업체		총합
		있다	없다	
혼합모형에서 업무수요의 유무	있다	10 (6.13%)	18 (11.04%)	28 (17.18%)
	없다	6 (3.68%)	129 (79.14%)	135 (82.82%)
총합		16 (9.82%)	147 (90.18%)	163 (100.0%)

실제로 신청한 업체가 12개 업체로 33.33%가 실제 업무수요가 있었던 것으로 나타났다.

3.3. 혼합 모형에서의 가중치 및 수요예측

가중치는 2006년 시점에서 설문조사에서의 실질오류율(actual error rate) 0.356과 로지스틱 회귀모형에서의 실질오류율(actual error rate) 0.172이며 식 (2.4)에 의하여 가중치는 0.4375이다. 2006년 수요예측은 다음의 식으로 표현이 가능하다.

$$D_{2006} = 0.4375D_{2006}^{(1)} + 0.5625D_{2006}^{(2)}$$

여기서 D_{2006} 는 2006년 시점의 기업의 수요확률, $D_{2006}^{(1)}$ 은 2006년 시점의 설문조사 수요, $D_{2006}^{(2)}$ 은 2006년 시점에서 로지스틱 회귀모형의 수요확률이다.

본 논문에서는 2006년 이후 실제로 평가인증을 신청한 개별 업체를 조사하여 수요 적중률이 최대가 되는 판정 기준 확률 값이 0.8로 산출되었다. 유효 모집단 163개 업체를 대상으로 혼합 모형을 이용한 업무수요에 대한 분류표는 다음의 표 3.4와 같다.

표 3.4에서 업무수요가 있는 업체가 전체 163개 업체 중 28개 업체로 나타났으며 그 중 실제로 신청한 업체가 10개 업체로 35.71%가 실제 업무수요가 있었던 것으로 나타났다. 2006년 시점에서 혼합 모형에서의 실질오류율(actual error rate)은 0.147로 2006년 시점에서 설문조사에서의 실질오류율(actual error rate) 0.356과 로지스틱 회귀모형에서의 실질오류율(actual error rate) 0.172보다 작게 나타났다.

4. 결론 및 향후 연구

본 논문은 서비스에 대한 수요를 조사하기 위하여 현장에서 사용되는 설문조사방법의 문제점을 해결하기 위한 방안으로 설문조사 전에 서비스에 대한 유효수요가 발생할 가능성을 측정하는 로지스틱 회귀모형을 제시하고, 설문조사를 통하여 예측된 수요와 로지스틱 회귀모형을 통하여 예측된 수요를 혼합하여 수요예측 모형을 제시하고자 하였다. 설문조사를 이용한 서비스 수요는 응답자의 의사와 능력에 크게 의존하며 응답자들은 서비스 수요 여부에 대한 결정권자가 아니면 수요에 대한 불확실한 답변으로 서비스 수요 여부를 측정하는데 비 표본 오차가 발생할 수 있으므로 설문조사에서 발생할 수 있는 오차를 로지스틱 회귀모형으로 보정하여 서비스 수요를 예측하였다. 본 연구의 결과는 수요자가 기업인 서비스에 대한 수요예측에서 설문조사로 예측된 수요의 실질오류율과 로지스틱 회귀모

형으로 예측된 수요의 실질오류율을 줄일 수 있는 혼합모형을 제시하였다. 제시한 수요예측 모형의 정도를 높게 하려면 설문에서의 오차와 모형에서의 변수선택이 중요한 변수가 될 것이며 중복수요에 대한 문제점도 해결하여야 할 것이다.

또한 본 논문에서 제시된 설문조사방법과 로지스틱회귀모형의 실질오류율과 혼합모형의 가중값은 한 가지 사례에 불과하며, 이를 동일한 현실문제에 그대로 적용하기에는 한계가 있다. 향후 다년간의 자료축적과 지속적인 연구를 통해 혼합모형의 가중값에 대한 검증을 실시하고 제시된 방법론에 대한 정확도를 측정하여 현장에서 유용하게 사용할 수 있는 모형으로 발전시켜야 할 것이다.

참고문헌

- 김범환, 박종봉, 김기영 (2000). 델파이 방법 적용 사례 연구(초고속망 장비관련 시장 예측과 관련하여), <기술혁신학회지>, 3, 102-112.
- 남경희, 안병덕, 안웅환 (2001). 정밀측정기술 수요조사연구, <기술혁신학회지>, 4, 110-126.
- 한국정보보호산업협회 (2006). 2006 국내 정보보호산업 통계조사, <한국정보보호진흥원>.

[2007년 8월 접수, 2007년 11월 채택]

Mixture Model with Survey and a Statistical Model

Youn-Jong Kim¹⁾ Yong-Chul Kim²⁾

ABSTRACT

In generally, we use a method of the survey to forecast the economic demands and non-economic demands for a market trend. But we have a difficult problem to estimate the demand for the marketplace objectively because the survey with the non-response and the inadequate understanding on questionnaires did not provided the strong and uniform forecast. Here, we proposed a method compounded of survey and a statistical model to estimate the demand for the marketplace and discussed the mixture model applied to the service demand on an agency.

Keywords: Logistic regression, survey, mixture model.

1) Corresponding author. Professor, Dept. of Computer and Information, Yongin University, 470 Samga-dong, Cheoin-gu, Yongin-si, Gyeonggi-do 449-714, Korea.

E-mail: younkim@yongin.ac.kr

2) Associate Professor, Dept. of Computer and Information, Yongin University, 470 Samga-dong, Cheoin-gu, Yongin-si, Gyeonggi-do 449-714, Korea.

E-mail: yckim@yongin.ac.kr