

고객 세분화에 기반한 생존분석을 활용한 고객수명 예측 모델

전희주^{1,a}

^a부산외국어대학교 데이터경영학과

요약

고객수명은 고객생애가치(CLV)와 함께 차별화된 마케팅 전략과 기업의 경쟁력 강화의 핵심수단으로 차별화된 고객관계관리(CRM) 마케팅의 매우 중요한 핵심요소이다. 그러나 보통 기업에서 사용하는 고객수명은 어떤 특정시점에서의 고객에 한정하여 고객 개별적인 특성을 반영하지 않고 단순 이탈율만을 가지고 고객수명을 예측하고 있다. 본 연구는 이러한 단점을 극복하고 현실적인 활용을 위해, 기업 고객들의 중도절단자료를 가지고 고객 세분화에 기반한 생존분석을 활용한 고객수명 예측 방법을 제시하고, 실제 활용을 위해 국내 A 이동통신사의 데이터를 활용하여 적용하였다. 본 연구에서 제시하고 있는 고객 세분화에 기반한 생존분석을 이용한 고객수명 예측방법은, 첫째, 가입시점이 다른 모든 고객을 반영하고 있고, 둘째, 고객의 개별적 특성을 반영하여 개별 고객수명에 대한 오차를 줄일 수 있으며, 셋째, 관측시점 이후의 수명을 예측함으로써 시간의 흐름에 따른 이탈율 또는 해지율의 변동추이를 반영하게 되어 더욱 현실성을 반영하고 있다.

주요용어: 고객생애가치, 고객수명, 중도절단자료, 생존분석.

1. 서론

오늘날 기업들은 다음의 3가지 커다란 도전에 직면해 있다. 첫 번째는 고객의 요구가 갈수록 증대되고 있다는 것이다. 많은 기업들은 고객의 선호 경향과 행동을 이해하는 쪽으로 변함으로써 다량마케팅(mass marketing)보다는 다량맞춤형(mass customization) 마케팅으로 변하고 있다. 두 번째는 이익 최적화(profit optimization)에 대한 필요성이다. 기업은 수익은 최대화를 하고 비용은 최소화하기 위해 ROI(return on investment)를 측정하고 예측하고자 한다. 세 번째로 경쟁에 대한 기업환경은 갈수록 더욱 더 치열해지고 있다는 것이다. 시장과 고객의 글로벌화는 가속화되고 있고 고객보호를 위한 시장규제(deregulation)는 더욱더 강화되고 있는 상황이다. 이러한 환경 속에서 기업의 입장에서 고객의 특성을 파악하고 고객의 가치를 측정하여 적절한 마케팅 전략을 수립하는 것은 고객가치를 상승시키고 지속적 경쟁우위를 유지하는데 있어서 매우 중요하다고 할 것이다.

고객가치에 대한 개념은 시장의 경쟁 환경 및 고객의 구매성향 변화와 정보기술의 발전 등으로 치열한 시장의 경쟁구도 속에서 고객과의 지속적인 관계를 유지함으로써 기업의 이익을 극대화하고자 하는 목적에서 출발하였다. 고객가치에 대한 올바른 측정과 고객의 가치분석은 고객가치에 따라 차별화된 대표적 마케팅 전략과 기업의 경쟁력 강화의 핵심수단으로 CRM(customer relationship management)의 중요한 요소가 되었다 (조용준과 허준, 2006). CRM은 정보기술의 발달로 축적된 고객의 다양한 데이터를 바탕으로 고객의 특성을 도출하여 고객의 가치에 따라 차별화된 마케팅을 제공하고자 하는 것이다. 따라서 기업들은 다양한 분석방법을 적용하여 고객에 대한 정보를 다차원적으로 파악하여 활용하는데 노력을 기울이고 있다.

이 논문은 2010년도 정부재원(교육과학기술부 인문사회연구역량강화사업비)으로 한국연구재단의 지원을 받아 연구되었음(NRF-2010-332-B00094).

¹ (608-060) 부산시 남구 석포로 15, 부산외국어대학교 데이터경영학과, 조교수. E-mail: hjchun@pufs.ac.kr

고객의 가치는 고객이 그 기업에게 가져다주는 이익과 그 고객이 그 기업의 상품을 사용하며 그 기업의 고객으로 유지되는 기간의 함수로 정의할 수 있다. 이러한 관점에서 고객의 생애가치를 측정하는 모형을 고객생애가치(Customer Lifetime Value; CLV) 모형이라고 한다. 고객생애가치는 CRM, 캠페인관리(campaign management), 리스크관리(risk management), 심지어 도용방지분석(fraud detection analysis)을 적용하는 데에도 적용할 수 있다 (Rosset 등, 2003). 그 중에서도 CRM 측면의 이탈분석(churn analysis), 고객유지 캠페인분석에서 고객생애가치는 중요한 추가 정보로 사용될 수 있을 것이다.

중요한 이슈는 기여도가 높은 즉 고객생애가치가 높은 고객은 유지하여, 고객유지 캠페인에 그들에게 자원과 마케팅을 집중하고, 그 고객들의 고객생애가치 변화에 관심을 집중하는 것이다 (Gupta 등, 2006). 따라서 기업의 고객관리 전략은 고객가치가 높은 고객은 유지하고, 고객가치가 낮은 고객은 고객가치를 높이며, 고객가치를 높일 수 없다고 생각되는 고객은 방치하거나 이탈을 유도하는 전략이다. 또한 기업은 현재 고객생애가치가 낮지만 고객생애가치를 올릴 수 있는 가능성이 높은 고객은 고객생애가치를 높이는 방향으로 마케팅 이동(migration)전략으로 나아가야 한다. 따라서 만일 고객생애가치를 안다면, 그 고객의 가치에 따라 고객을 분류하고, 그 기여도와 행동특성에 따라 차별적 마케팅 전략을 통해 기업의 수익성과 고객의 로열티를 더욱 높일 수 있을 것이다 (전희주, 2008). 예를 들어, 통신회사에서는 고객가치가 높은 고객에게는 더 많은 보조금을 지불할 수 있을 것이며, 은행이나 카드사 등의 금융업에서는 카드 연회비, 이체수수료 등 각종 수수료 혜택을 제공하여 그 기업에 유리한 고객을 유치하고 계속 유지시킬 수 있을 것이다. 그러므로 기업 현장에서는 정확한 고객가치를 측정하고 예측하는 것이 무엇보다 중요하다고 할 수 있다.

지금까지 고객가치에 대하여 많은 이론적 연구와 실증분석이 이루어져 왔으나 많은 연구가 주로 고객의 수명에 대한 예측은 고정된 값으로 가정하거나 기업의 전체적인 이탈율을 이용하여 진행되어 왔다 (Reichheld, 1996; 강현철 등, 2003). 이에 본 연구는 고객생애가치를 측정하는 요소들-고객의 주어진 기간 동안 가져다주는 가치, 고객의 서비스 기간과 할인율- 중 중도절단자료를 다루는 통계학의 생존분석을 활용하여 고객수명을 예측하는 방법을 제시하고자 한다. 이 생존분석 방법은 특정 관찰시점에 이탈고객과 유지고객들을 고려하여 관찰시점 이후에 해지율을 예측하여 전체 수명을 산출하는 방식이다.

기업에서 많이 다루는 개별 고객의 수명에 대한 데이터는 이미 이탈하여 해지 또는 종료된 고객의 기간과 현재 계속 남아 있는 고객으로 구성되어 있다. 이러한 형태의 자료는 중도절단자료(censored data)로 다른 데이터와 구별되는 특성을 가지고 있어 생존분석에서 다루어지고 있다 (이재원 등, 2007). 중도절단자료는 모든 연구 대상에 대한 결과가 모두 완성되어 있지 않고, 미완성으로 남아 있는 경우에 발생한다. 예를 들어 암환자의 생존시간을 관측하는 임상실험에서 몇몇의 환자들이 도중에 그만 두어 연락이 두절되는 경우이다. 이 경우에 우리는 그 환자가 최소한 마지막 시간까지 살아 있었음을 알 수 있을 뿐이다. 또한 실험의 종료로 인해 살아 있는 암환자는 현재까지 살아 있는 정보만 이용할 수 있을 것이다. 이러한 경우는 분석시점에 기업에서의 고객들에 대한 수명자료에도 동일하게 적용할 수 있을 것이다.

오늘날 많은 기업이 고객의 생애가치를 측정할 때, 고객 개별에 대한 생애가치를 정의하지 않고, 과거 몇 년 동안 그 회사의 과거 사용기간을 가지고 고객의 가치를 판단하고 있다. 또 고객의 수명예측은 고객 개별에 대한 고객수명을 예측하기보다는 집단의 월별 평균 해지율이나 월별 평균 이탈율을 가지고 집단의 평균적 수명을 예측하고 있는 실정이다. 따라서 가입시점이 다른 고객과 고객의 개별적 특성을 고려하여 오차를 줄일 수 있는 수명예측으로 좀 더 현실성 있는 방법이 요구되고 있다. 이에 본 연구는 기업의 고객 중도절단자료를 활용한 생존분석(survival Analysis)을 통한 고객수명 예측 방법을 제시하고, 실제 적용을 위해 국내 A 이동통신사의 데이터를 활용하여 적용하고자 한다.

2. 고객생애가치 정의

일반적으로 고객생애가치는 고객의 주어진 기간 동안 가져다주는 가치, 고객의 서비스 기간과 할인율의 3가지 구성요소로 구성되는 데, 고객으로부터 그 회사와 평생 동안 관계로부터 얻어지는 모든 미래가치의 현재가치로써 정의된다. 이 고객생애가치는 재무관론 혹은 경제성공학 분야에서 다루는 할인된 돈의 흐름과 비슷한 접근을 갖는다. 각 요소는 직접 얻은 값을 입력하거나 또는 직접 얻을 수 없는 경우에는 각각 분리하여 모형에 의한 추정을 통해 얻게 된다. Hoekstra와 Huizingh (1999)는 고객생애가치는 시스템에서 주어진 기간 동안 고객으로부터 얻어질 수 있는 전체 순이익으로 정의되고, 고객생애가치는 고객이탈, 수익과 캠페인관리를 분석하는 데 핵심요소이다.

고객생애가치에 대한 정의는 다음과 같이 표현된다 (Gupta 등, 2004; Reinartz와 Kumar, 2003).

$$CLV = \sum_{t=0}^T \frac{(p_t - c_t)r_t}{(1+i)^t} - AC, \tag{2.1}$$

여기서 p_t 는 t 시간에 고객에 의해 지불된 가격, c_t 는 t 시간에 고객에게 서비스하기 위해 지불된 비용, i 는 회사의 자본비용에 대한 할인율, r_t 는 t 시간에 고객으로 유지하고 또는 반복 구매할 확률, AC 는 고객 유치비용, T 는 고객 수명을 예측하는 시간의 길이이다.

Jain과 Singh (2002)는 다음의 고객생애가치 모형을 제시하였는데 식 (2.1)과 거의 유사한 모형이라 할 수 있다.

$$CLV = \sum_{i=0}^n \frac{(R_i - C_i)}{(1+d)^i}, \tag{2.2}$$

여기서 i 는 고객의 거래로부터 일어나는 현금흐름의 기간, d 는 현금흐름에 대한 할인율, R_i 는 i 기간에 고객으로부터 얻은 수익, C_i 는 i 기간에 수익 R_i 를 발생시키는 비용, n 는 고객 회사와 관계를 하는 기간을 더한 전체 수명이다.

모형 (2.1)과 (2.2)는 각 시간 기간의 끝 시점에서 모든 현금흐름을 가정한 것이다. Rosset 등 (2002)은 수학적으로 더 일반화한 다음의 모형을 제시하였다.

$$CLV = \int_0^{\infty} S(t)v(t)D(t)dt, \tag{2.3}$$

여기서 $v(t)$ 는 시간 t 에 대한 고객의 가치($t \geq 0$), $S(t)$ 는 고객이 서비스를 받는 시간의 길이, C_i 는 i 기간에 수익 R_i 를 발생시키는 비용, $D(t)$ 는 시간 t 에 있어서 현금 할인율이다.

식 (2.1), (2.2)와 (2.3)은 표현하는 방법은 차이가 있지만, 한 기업에 한 고객이 가져다주는 고객생애가치를 나타내고자 하는 면에서는 모두 일치한다고 볼 수 있다. 고객생애가치의 3가지 요소 중 주어진 기간 동안의 고객의 가치는 현재의 데이터를 가지고 사업적인 지식과 통계적인 분석을 통하여 구해진다. 다른 요소인 서비스 받는 시간의 길이 즉 수명은 주어진 기간 동안에 대한 이탈확률(churn probability)로써 나타낼 수 있으며, 생존분석에서 사용하는 t 시점에 고객이 여전히 살아 있는 확률인 생존함수(survival function) $S(t)$ 로 표현할 수 있을 것이다 (Helsen과 Schmittlein, 1993).

3. 고객수명 예측 방법론

A 이동통신사에서는 고객의 수명을 고객 개별적인 특성을 반영하지 않고 어떤 특정시점에서의 고객들로 한정된 고객의 해지율을 적용하여 고객의 수명을 계산하여 왔다. 이는 해당 고객 세그먼트

트(segment)의 평균 이탈을 또는 해지율을 사용하기에 실측 이탈을 또는 해지 이후의 이탈을 또한 일정하다고 가정하는 단점을 가지고 있다. 본 연구에서는 이러한 단점을 극복하고 개별고객에 대한 수명을 예측하고자 고객세분화를 기반으로 하는 임상실험이나 상품수명 등의 중도절단자료 예측에 많이 사용되는 통계학의 생존분석 방법인 준모수적(semi-parametric) 접근법인 Cox의 비례위험 모형(proportional hazards model, Cox, 1972)과 모수(parametric)적인 생존함수 추정을 결합한 모형적합 방법을 제시하고자 한다. 즉 각각 동질의 고객들로 이루어진 세그먼트의 과거 그 기업의 상품을 이용한 기간 동안의 정보를 기반으로 생존함수를 구해 각 개별고객의 미래 이용기간을 예측함으로써 각 개별 고객이 그 회사에 남아 있을 전체 수명을 예측하는 방법이다.

Cox의 비례위험 모형은 생존시간의 분포에 대한 가정을 필요로 하지 않고 시간에 따라 바뀌는 공변량(time-dependent variable)의 경우에도 분석할 수 있다는 장점이 있어 생존자료의 분석에서 매우 자주 사용되는 모형이다.

생존시간을 $T (\geq 0)$ 라고 할 때, T 가 확률밀도함수 $f(t)$ 와 분포함수 $F(t)$ 를 갖는다고 하면, T 의 생존함수(survival function) $S(t)$ 와 위험함수 $h(t)$ 는 각각 다음과 같이 정의된다. 이때 생존시간 T 에 영향을 주는 k 개의 공변량 x_1, \dots, x_k 가 있다면, Cox의 비례위험모형은

$$h(t) = h_0(t) \exp(\beta_1 x_1 + \dots + \beta_k x_k)$$

또는

$$\log h(t) = \log h_0(t) + \beta_1 x_1 + \dots + \eta_k x_k$$

와 같이 표현된다. 이는 공변량들을 고려했을 때의 위험함수(hazard function)이고, $h_0(t)$ 는 기저위험함수(baseline hazard function)로 공변량의 값들이 모두 0일 때의 위험함수의 값이다. 또한 시간 t 를 제외한 공변량들은 고정효과(fixed effect)를 나타낸다. 고객생애가치의 요소인 고객 수명 예측에 대한 모형은 중도절단된 고객을 다루는 생존분석의 특별한 경우로서 타당하다 할 것이다.

본 연구에서는 특정 관찰시점에 이탈고객과 유지고객들을 고려하여 관찰시점 이후에 해지율을 예측하여 전체 수명을 산출하는 방식인 생존분석 방법을 적용하고자 한다. 시간 즉 기업에서 고객의 이용기간은 예측하고자 하는 값이고 이탈 유무는 반응변수가 된다. 이 예측에 대한 모형을 구하는 절차는 다음과 같다.

- 1) 이탈 또는 유지기간(생존시간)에 영향을 주는 공변량 변수를 선택한다. 예를 들어 이동통신사의 경우 연령, 약정 유무, 가입연도, 기변회수, 해지 또는 측정시점 전의 3개월 사용량, 번호이동 유무 등.
- 2) Cox의 비례위험 모형을 이용하여 영향을 주는 공변량에 대한 누적이탈율 등의 생존함수를 추정하여 기저생존함수를 추정한다. 기저생존함수는 다음의 식에 의해 구해진다.

$$S_0(t) = \frac{\log S(t)}{\exp(\beta_1 x_1 + \dots + \beta_k x_k)}$$

- 3) 가장 적합한 모수적 생존시간의 분포와 누적 해지율을 찾는다. 그림 1은 누적 해지율 추정 방법을 그림으로 표시한 것이다.
- 4) 관측시점 이후는 적합한 생존시간 분포를 가지고 생성한다. 여기서 기저생존함수의 분포를 가정하면 관측시점 이후의 기저생존함수를 구할 수 있다.
- 5) 각 세그먼트(segment)를 나타내는 설명변수별에 대해 생존함수를 추정한다.

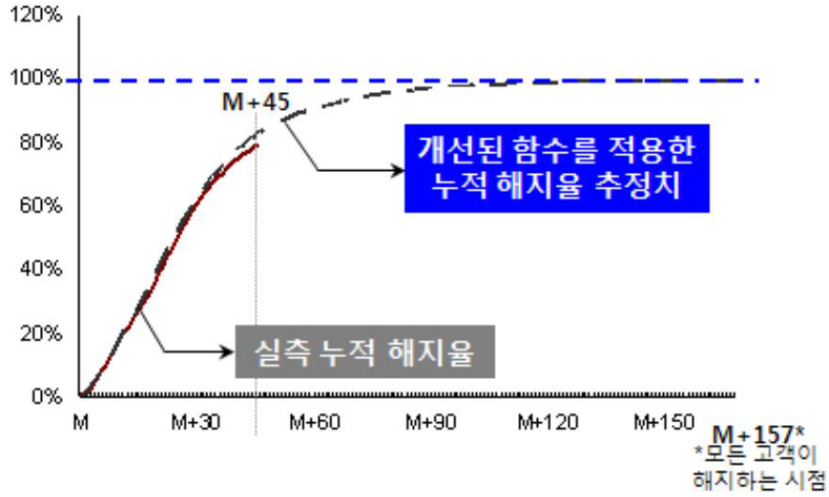


그림 1: 누적 해지율 추정

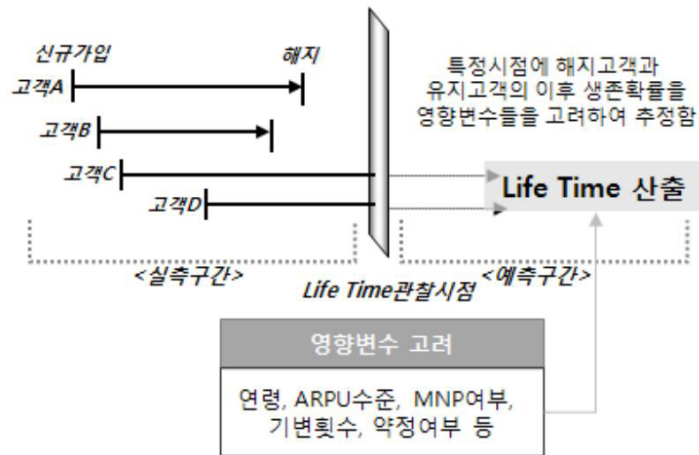


그림 2: 고객 세분화 기반의 생존분석 방법론

6) 생존함수를 반영한 개별고객에 대한 고객수명을 구한다.

그림 2는 본 논문에서 6단계로 구분해 설명한 고객 세분화 기반의 생존분석 방법론을 그림으로 도식화하여 요약한 것이다.

본 연구에서 제시하는 생존분석을 활용한 고객수명 예측은 다음과 같은 측면에서 장점을 들 수 있다. 첫째로 가입시점이 다른 모든 고객을 반영하고 있다는 것이다. 둘째, 고객들의 공변량을 모형의 영향변수로 포함하여 고객의 개별적 특성을 반영하여 개별 고객수명에 대한 오차를 줄일 수 있다. 셋째로, 관측시점 이후의 수명을 예측함으로써 시간의 흐름에 따른 이탈율 또는 해지율의 변동추이를 반영하게 되어 더욱 더 현실성을 반영한다고 할 수 있을 것이다. 그림 3은 이러한 관측시점 이후의 수명 예측 방법의 차이를 그림으로 자세히 표현한 결과이다.



그림 3: 가입 이후의 해지율 추정 비교

표 1: A 이동통신 해지 개인고객의 평균수명, ARPU와 고객생애가치

		평균 수명(월)	ARPU(원)	CLV(만원)
전체 개인		37.9	26,135	99.1
가입 구분	MNP 가입고객	45.1	32,305	145.8
	일반 가입고객	29.4	18,874	55.6
해지 년도	2005	37.5	19,695	73.8
	2006	40.2	22,162	89.2
	2007	40.0	30,917	123.5
	2008	34.6	29,847	103.3

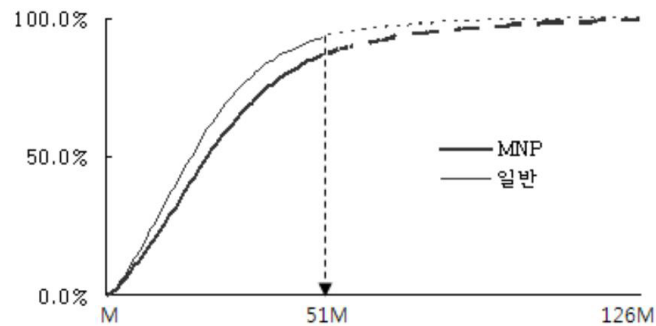


그림 4: 가입 구분(MNP, 일반)에 따른 누적 해지율 추정

4. 이동통신사의 적용 사례

본 연구에서 적용한 A 이동통신사의 자료는 이동통신사의 번호이동(MNP; mobile number portability) 시작 시점인 2004년 7월부터 2007년 12월까지 가입고객 전체에 대하여 5%의 표본인 개인 829,640회선을 가지고 적용하였다. 이 기간 동안 가입한 개인고객을 2009년 1월 시점까지의 수명과 해지유무를 나타낸 자료이기에 이런 이동통신사의 고객자료는 중도절단된 자료라 할 수 있을 것이다.

표 1은 개인고객의 실제 해지고객들을 대상으로 가입시점 이후의 수명과, ARPU(average rate per user, 월 평균 고객 1인당 이익)와 실제 고객생애가치를 보여준다. 개인 해지고객은 37.9개월 동안 가입 유지하였고 월 평균 이용금액은 26,135원이었고 고객생애가치는 99.1만원이었다. MNP고객이 일반고

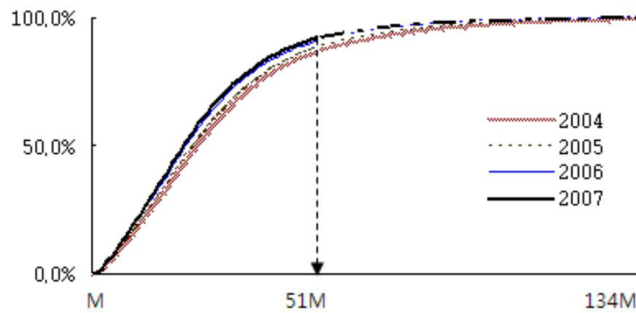


그림 5: 약정 구분(약정, 비약정)에 따른 누적 해지율 추정

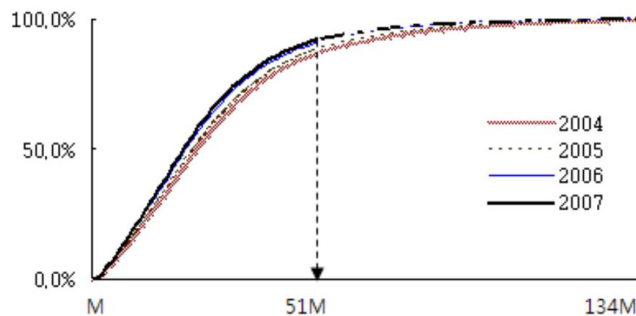


그림 6: 가입연도에 따른 누적 해지율 추정

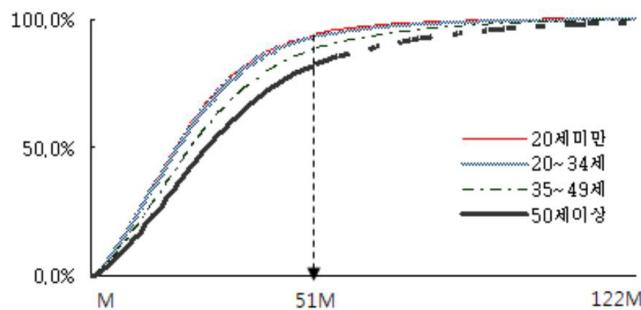


그림 7: 연령구분에 따른 누적 해지율 추정

고객보다 수명이 더 길며 ARPU 또한 2배 가까이 높은 경향을 보이고 있다. 해지연도별로는 상대적 차이는 크지 않게 나타났다.

그림 4~그림 7은 본 연구에서 제안한 각 개인고객의 가입구분, 약정여부, 가입연도, 연령그룹에 따른 세그먼트별에 따른 생존함수를 활용하여 해지율을 예측한 것이다. 일반 가입고객보다는 MNP 가입 고객의 해지율이 낮게 나타나며, 비약정 고객보다는 약정고객의 해지율이 낮으며, 연도별은 최근에 가까울수록 해지율이 높게 나타나며, 연령이 낮을수록 해지율은 낮은 현상을 보이고 있다.

표 2는 A통신사의 전체 개인고객에 대한 수명과 고객생애가치를 단순 해지율을 이용한 기존방법과 본 연구에서 제시한 고객 세분화 기반의 생존분석을 활용하여 해지율과 고객수명을 구한 개선된 방

표 2: 기존방법과 고객 생존분석에 의한 수명 비교

가입연도	평균수명(단위: 월)			CLV(단위: 만원)		
	기존 방법	개선된 방법	차이	기존 방법	개선된 방법	차이
2005	27.5	29.1	1.6	89.0	94.1	5.1
2006	22.9	26.7	3.8	77.1	89.8	12.7
2007	22.1	25.9	3.8	74.0	86.7	12.7

표 3: A 이동통신 개인고객의 가입년도에 따른 평균수명과 고객생애가치

		고객 수(명)	고객비중(%)	평균 수명(월)	ARPU(원)	CLV(만원)
전체		829,663	100.0	27.2	32,920	89.5
가입 구분	MNP	384,821	46.4	30.4	36,807	112.0
	일반	444,842	53.6	24.7	29,558	73.1
약정 여부	약정	25,591	3.1	30.0	53,658	160.8
	비약정	804,072	96.9	27.1	32,260	87.4
가입 연도	2004	74,159	8.9	30.6	29,490	90.1
	2005	192,962	23.3	29.1	32,349	94.1
	2006	215,027	25.9	26.7	33,666	89.8
	2007	347,515	41.9	25.9	33,508	86.7
연령 그룹	20세 미만	157,431	19.0	24.1	25,248	60.8
	20~34세	304,645	36.7	24.8	38,104	94.6
	35~49세	244,448	29.5	29.3	34,268	100.3
	50세 이상	123,139	14.8	34.8	27,229	94.8
ARPU 그룹	< 1.8만원	227,615	27.4	25.1	9,866	24.8
	< 2.5만원	137,923	16.6	32.3	21,381	69.1
	< 4.0만원	235,219	28.4	28.8	31,564	90.8
	≥ 4.0만원	228,906	27.6	25.1	64,190	161.4
기변 횟수	0회	592,349	71.4	23.8	30,574	72.7
	1회	119,098	14.4	37.1	36,661	135.9
	2회	54,215	6.5	33.4	37,372	124.8
	3회 이상	64,001	7.7	49.8	43,898	218.6

법을 비교하여 제시한 것이다. 표 2를 보면, 평균수명을 예측하는데 있어서 단순 해지율에 의한 기존 방법은 고객 세분화 기반의 생존분석 방법보다 고객의 수명을 저 평가하는 경향을 보이고 있다. 이는 고객 해지율의 트렌드를 반영하지 않고 단순 평균화하는 데서 수명을 낮게 평가하고 있다.

표 3은 본 연구에서 제안한 각 개인고객의 가입구분, 약정여부, 가입연도, 연령그룹, ARPU그룹, 기변횟수에 따른 세그먼트별에 따른 생존함수를 활용하여 고객수명과 고객생애가치를 구한 것이다. 일반 가입고객보다는 MNP 가입고객의 고객수명이 높게 나타나며, 비약정 고객보다는 약정고객의 고객수명이 높으며, 연도별은 최근에 가까울수록 고객수명이 높게 나타나는 데, 이는 새로운 신형단말기의 보급과 마케팅의 경쟁의 심화로 인해 매년 고객의 가입 수명은 짧아지는 경향을 보이고 있다. 당연히 기변횟수가 많은 고객의 고객수명은 그렇지 않은 고객보다 낮게 나타났다.

5. 맺음말

고객수명은 고객생애가치와 함께 차별화된 대표적 마케팅전략과 기업의 경쟁력 강화의 핵심수단으로 차별화된 CRM 마케팅의 매우 중요한 핵심요소이다. 그러나 보통 기업에서 사용하는 고객수명은 어떤 특정시점에서의 고객에 한정하여 고객 개별적인 특성을 반영하지 않고 있다. 그리고 해당 고객 세그먼트의 평균 이탈율 또는 해지율을 사용하기에 실측 이탈 또는 해지 이후의 이탈율 또한 일정

하다고 가정하는 단점을 가지고 있다. 그래서 본 연구에서는 이러한 단점을 극복하고 현실적인 활용을 위해, 기업 고객들의 중도절단자료를 가지고 고객 세분화에 기반한 생존분석(Survival Analysis)을 활용한 고객수명 예측 방법을 제시하고, 실제 적용을 위해 국내 A 이동통신사의 데이터를 활용하여 적용하였다. 본 연구에서 제시한 고객수명 예측방법은, 첫째, 가입시점이 다른 모든 고객을 반영하고 있고, 둘째, 고객들의 공변량을 모형의 영향변수로 포함하여 고객의 개별적 특성을 반영하여 개별 고객수명에 대한 오차를 줄일 수 있으며, 셋째, 관측시점 이후의 수명을 예측함으로써 시간의 흐름에 따른 이탈율 또는 해지율의 변동추이를 반영하게 되어 더욱 더 현실성을 반영한다고 할 수 있을 것이다.

근래 기업은 고객 가치가 높은 고객에 대해 차별성을 더욱 더 강조하고 이를 통해 여러 가지 자원의 배분을 효율적으로 배분하고자 한다. 본 연구에서 제안하는 생존분석을 이용한 고객수명 예측은 고객 유지기간 즉 고객수명에 대한 가치를 강조하는 은행, 카드사, 통신사, 증권회사, 백화점 등 모든 업종에서 멤버십 등의 고객 로열티 정책, 고객유지 및 이탈방지 마케팅 등 다양한 영역에서 활용가능하리라 여겨진다.

참고 문헌

- 강현철, 한상태, 신혜림 (2003). 데이터마이닝 기법을 이용한 고객생애가치 측정모형 개발, *Journal of the Korean Data Analysis Society*, **5**, 927-936.
- 이재원, 박미라, 유한나 (2007). <생명과학연구를 위한 통계적 방법>, 자유아카데미.
- 전희주 (2008). 고객 로열티 스코어 모델 개발, <응용통계연구>, **21**, 211-219.
- 조용준, 허준 (2006). 고객가치모형별 마케팅전략: 백화점 화장품 고객을 중심으로, *Journal of the Korean Data Analysis Society*, **8**, 335-348
- Cox, D. R. (1972). Regression models and life tables, *Journal of Royal Statistical Society, Series B*, **34**, 187-220.
- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Ravishanker, N. and Sriram, S. (2006). Modeling customer lifetime value, *Journal of Service Research*, **9**, 139-155.
- Gupta, S., Lehmann, D. R. and Stuart, A. S. (2004). Valuing customers, *Journal of Marketing Research*, **41**, 9-24.
- Helsen, K. and Schmittlein, D. C. (1993). Analyzing duration times in marketing: Evidence for the effectiveness of hazard rate models, *Marketing Science*, **11**, 395-414.
- Hoekstra, J. C. and Huizingh, E. K. R. E. (1999). The lifetime value concept in customer-based marketing, *Journal of Market Focused Management*, **3**, 257-274.
- Jain, D. and Singh, S. (2002). Customer lifetime value research in marketing: A review and future directions, *Journal of Interactive Marketing*, **16**, 34-46.
- Reichheld, P. F. (1996). *The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value*, Harvard Business School Press
- Reinartz, W. and Kumar, V. (2003). The impact of customer relationship characteristics on profitable lifetime duration, *Journal of Marketing*, **67**, 77-99.
- Rosset, S., Neumann, E., Eick, U. and Vatnik, N. (2003). Customer lifetime value models for decision support, *Data Mining and Knowledge Discovery*, **7**, 321-339.
- Rosset, S., Neumann, E., Eick, U., Vatnik, N. and Idan, Y. (2002). Customer lifetime value modeling and its use for customer retention planning, In *Proceedings of ACM International Conference on Knowledge Discovery and Data Mining*.

Customer Lifetime Value Model Using Segment-Based Survival Analysis

Heuiju Chun^{1,a}

^aDepartment of Data Management, Pusan University of Foreign Studies

Abstract

Customer Lifetime or Customer Lifetime Value is a essential metric of differentiated CRM marketing and differentiated marketing strategy as a company core competency. However, customer lifetime used in companies is easily obtained from a confined simple customer attrition rate at some specific time point regardless of customer characteristics. In this study, in order to overcome the constraints of previous simple methods and to make practical use of it in industries, we suggest a method that estimates a customer lifetime using a customer segment based survival analysis with the censored data of customers; in addition, we apply this method to A mobile telecom company data. A method using customer segment based survival analysis is suggested in this study 1) includes all customers having different subscription dates, 2) reduces individual error, 3) can reflect trends after the observed time point and is more realistic.

Keywords: Customer Lifetime Value(CLV), lifetime, censored data, survival analysis.

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government(NRF-2010-332-B00094).

¹ Assistant Professor, Department of Data Management, Pusan University of Foreign Studies, 15 Sukporo, Nam-Gu, Busan 608-738, Korea. E-mail: hjchun@pufs.ac.kr