

## 생명표를 이용한 회원유형별 인터넷 고객 수명 분석

박희창<sup>1</sup>

창원대학교 통계학과

접수 2008년 12월 15일, 수정 2009년 1월 7일, 게재확정 2009년 1월 13일

### 요약

점차 복잡화되고, 다양화되고 있는 고객들의 수요를 만족시키기 위해 많은 인터넷 회사에서는 고객관계관리를 통하여 마케팅 활동을 벌이고 있다. 단기간 판매실적을 극대화하기 보다는 고객서비스와 장기간의 고객만족도를 더 많이 강화하기 위해서는 마케팅 계량지표로서 고객생애가치를 활용하는 것이 바람직하다. 이러한 고객생애가치는 인터넷 비즈니스 분야에서 경쟁적인 힘을 증진시킬 수 있는 고객가치를 발견할 수 있는 핵심이며, 많은 인터넷 기업들이 고객생애가치에 관심을 가지고 있다. 본 논문에서는 회원 유형별로 인터넷 사이트의 회원 수명을 분석하기 위해 생명표 기법을 활용함으로써 여러 형태의 인터넷 회사에 다양한 생명표를 이용할 수 있는 기회를 마련하고자 한다.

주요용어: 고객관계관리, 고객생애가치, 마케팅 계량지표, 생명표, 인터넷 기업.

### 1. 서론

점차 다양화되고 복잡해지는 소비자의 욕구를 충족시키기 위해 기업들은 고객관계관리 (customer relationship management ; CRM)를 통한 마케팅 활동을 벌이고 있다. 기업의 마케팅 방식은 시대의 흐름과 기업 환경의 변화에 따라 지속적으로 변화하여 왔다. CRM은 사업전략, 영업정보, 영업 프로세스, 조직의 경영능력, 기술 인프라 등 고객 관리에 필수적인 요소들을 고객 중심으로 정리·통합하여 고객과의 상호작용을 개선함으로써, 고객과의 장기적인 관계를 구축하고 기업의 경영 성과를 개선하기 위한 새로운 경영방식이라고 볼 수 있다.

이러한 배경 하에 최근에 해결책으로 등장한 것이 고객생애가치 (customer lifetime value ; CLV)의 극대화라는 개념이다 (최정환, 2005). CLV에 대한 정의는 현재까지 많은 학자들에 의해 내려져 왔다. 대표적인 연구로는 Burger와 Nasr (1998)이 기업과 거래하는 기간 동안 고객에게서 얻는 수익 또는 손실로 정의하는 동시에 고객 평생에 걸쳐 얻어지는 전체 수익에 고객 확보, 판매 및 서비스 소요경비를 제한 다음 돈의 시간적 가치 개념을 추가함으로써 계산될 수 있다고 주장한 바 있다. 또한 Gupta 등 (2004)은 한 고객으로부터 발생하는 모든 미래 수익에 대한 현재가치라고 정의하였다. 국내 연구로는 Kang 등 (2003)이 고객이 일정한 기간 동안 그 기업에게 제공할 것으로 추정되는 잠재적인 수익의 합계로 정의를 내리는 동시에 데이터마이닝 기법을 이용한 CLV 측정 모형을 제안한 바 있다.

기업의 이익 측면에서 볼 때, 한 고객의 가치는 고객이 기업의 고객으로 유지되는 기간, 고객이 기업에게 제공하는 수입, 그리고 기업이 고객을 유지하고 유지하기 위해 지불하는 비용 등의 함수로 정의할 수 있으며, 이러한 관점에서 고객의 가치를 측정하려는 모형을 일반적으로 CLV 모형이라고 부른다 (정현수와 박성호, 2006). 특히 포화상태가 된 시장에서 기업은 새로운 이윤을 창출할 수 있는 새로운 돌파구가 절실히 필요하게 되었으며, 대다수의 기업들이 고객에 대해 더 많은 관심을 가지기 시작했다.

<sup>1</sup> (641-773) 경상남도 창원시 사림동, 창원대학교 통계학과, 교수. E-mail : hcpark@changwon.ac.kr

이러한 관점에서 볼 때, 충성도가 높은 고객의 수명을 가급적이면 연장하는 것이 더 많은 기업의 이익을 창출하며, 결국 경쟁에서 우위를 차지할 수 있는 방법이라고 할 수 있다. 인터넷 사이트의 회원 수명 분석에 적용 가능한 통계적 기법으로는 시간에 따른 회원자격 유지양상을 일반화하여 표현해주는 지수 함수 등과 같은 일정 함수를 이용하는 모수적 방법과 생명표 (life table) 기법, Chiang 기법, product-limit 기법, Littel 기법 등의 비모수적 방법이 있다 (안윤옥, 1990). 특히 Ali 등 (2005)은 기하분포를 따르는 모집단에 대한 비모수적 베이저안 다중 비교방안에 대해 연구한 바 있으며, Cho 등 (2006)은 이변량 지수분포를 따르는 모집단에 대한 베이저안 다중 비교방법을 제안한 바 있다. 본 논문에서는 인터넷 사이트의 회원유형에 따른 고객의 수명을 분석하기 위해 생명표 기법을 활용하는 방안에 대해 기술하고자 한다. 본 논문의 2절에서는 회원유형에 의한 수명분석을 위한 생명표 기법의 적용 방안을 구체적으로 제시한다. 3절에서는 가상의 자료를 이용하여 생명표를 작성하고 회원유형에 따른 고객수명분석 (customer lifetime value) 적용가능성을 타진한 후, 마지막으로 4절에서 결론을 내리고자 한다.

## 2. 회원유형에 따른 고객수명분석을 위한 생명표 기법

생명표 기법은 가장 보편적으로 사용되고 있는 비모수적 누적생존율 (cumulative survival rate)을 산출하는 방법으로 관찰기간을 일정 단위로 나누어 각 구간마다의 구간 생존 (또는 사망) 확률을 구하고 이들을 누적하여 일정기간까지의 누적 생존율 (또는 사망률)을 구하는 방법이다 (고응린, 1979).

먼저 생명표 기법을 인터넷 사이트의 고객 수명 분석에 활용하기 위해 치료 (또는 진단) 개시일을 회원 가입일로, 사망을 탈퇴로, 또한 소식 불명상태를 휴면상태로 변형하고자 한다. 그리고 회원 유형은 유료회원과 무료회원으로 구분한다. 이러한 사실로부터 회원 유형에 따른 고객수명분석을 위해 필요한 기호를 정의하면 다음과 같다.

$O_{(i)t}$  :  $t$ 구간의 최초시점에서 유형  $i$  (값이 1이면 무료회원, 2이면 유료회원)의 회원자격을 유지하고 있는 고객의 수

$O_{(i)t+1}$  :  $t$ 구간의 최종시점에서 유형  $i$ 의 회원자격을 유지하고 있는 고객의 수

$d_{(i)t}$  :  $t$ 구간에서 탈퇴하거나 회원유형을 변경한 고객의 수

$u_{(i)t}$  :  $t$ 구간에서 유형  $i$ 의 회원자격을 유지하고는 있으나 로그인을 한 번도 하지 않은 고객의 수

$w_{(i)t}$  :  $t$ 구간에서의 회원 유형별 관찰 중단 고객의 수

여기서  $d_{(1)t}$ 는 각 구간에서 탈퇴하거나 무료회원에서 유료회원으로 변경한 고객의 수를 의미하고,  $d_{(2)t}$ 는 각 구간에서 유료회원의 자격을 상실한 고객의 수를 의미한다. 또한  $w_{(i)t}$ 는 절단된 자료 (censored cases)로 추적 종결 시에도 유형  $i$ 의 회원자격을 유지하고 있는 고객의 수를 의미한다.

이러한 기호를 토대로 생명표 기법에서 필요한 항목을 나타내면 다음과 같다. 먼저 균등분포 (uniform distribution)를 따른다는 가정 하에 각 구간에서의 회원유형별 유효 관찰 고객의 수  $r_{(i)t}$ 는 다음과 같이 계산된다.

$$r_{(i)t} = O_{(i)t} - \frac{u_{(i)t} + w_{(i)t}}{2} \quad (2.1)$$

생명표 기법에서의 유형별 구간 유효 탈퇴자 수는 구간 내에서의 탈퇴자 수와 유형변경자수가 되고, 유효관찰 고객의 수는 탈퇴자와 유형변경자를 제외한 회원자격 유지 고객수 및 휴면 상태의 고객수의 총 관찰기간 (person-time)으로 하고 있다. 또한 특정 구간 내에서 탈퇴한 경우는 그 탈퇴시점이 해당 구간의 최종 시점에 가깝다고 할지라도 유효관찰수에는 반영되지 않는다. 각 구간에서의 회원유형별 탈퇴

을  $q_{(i)t}$ 는 다음과 같이 계산된다.

$$q_{(i)t} = \frac{d_{(i)t}}{O_{(i)t} - \frac{1}{2}(u_{(i)t} + w_{(i)t})} \quad (2.2)$$

각 구간에서의 유형별 회원자격 유지 비율  $p_{(i)t}$ 는 다음과 같이 계산된다.

$$p_{(i)t} = 1 - q_{(i)t} \quad (2.3)$$

각 구간에서의 유형별 회원자격 유지 누적 비율 (누적 생존율)  $S_{(i)t}$ 는 다음과 같이 계산된다.

$$S_{(i)t} = \prod_{i=1}^t p_{(i)t} \quad (2.4)$$

1,000 (또는 10,000)명의 고객들이 동시에 회원가입을 한다고 가정하였을 때  $t$ 기간 후에 각 유형별 회원자격을 유지하는 고객수  $l_{(i)t}$ 는 다음과 같이 계산된다.

$$l_{(i)t} = S_{(i)t} \times 1,000(\text{또는 } 10,000)\text{명} \quad (2.5)$$

### 3. 예제를 통한 고찰

본 절에서는 다음의 표 3.1과 같은 가상의 자료를 이용하여 고객수명분석을 위한 생명표를 작성해보기로 한다. 이 표는 2006년 1월 1일부터 회원으로 가입한 고객들을 추적 관찰하기 시작하여 2007년 12월 31일에 종료한 것으로 간주한다. 여기서 +는 중도절단된 경우 (censored case)를 의미한다.

표 3.1로부터 무료회원에 대해 추적한 관찰 결과를 나타내면 표 3.2와 같다. 여기서 ID가 1, 2, 4, 5, 7, 10, 12, 13, 17, 18, 20, 27, 28은 가입 당시부터 무료회원인 고객들이고, 6, 14, 16, 23은 가입 당시 무료회원에서 유료회원으로 변경한 고객들이며, 8, 11, 26은 가입 당시에는 유료회원이었다가 무료회원으로 변경한 고객들이다. 가입 당시 무료회원에서 유료회원으로 변경한 고객들의 경우에는 최종시점의 고객 상태는 탈퇴 (변경)가 되고, 유형변경일자가 탈퇴일자로 바뀌게 된다. 그리고 가입 당시 유료회원에 서 무료회원으로 변경한 고객들은 유형변경일자가 가입일자로 바뀌게 된다.

표 3.1로부터 유료회원에 대해 추적한 관찰 결과를 나타내면 표 3.3과 같다. 여기서 ID가 3, 9, 15, 19, 21, 22, 24, 25, 29, 그리고 30은 가입 당시부터 유료회원인 고객들이다. 가입 당시 무료회원에서 유료회원으로 변경한 고객들인 6, 14, 16, 23은 표 3.2와 중복하여 표 3.3에도 나타나지만 유형변경일자가 가입일자로 바뀌게 된다. 또한 가입 당시에는 유료회원이었다가 무료회원으로 변경한 고객들인 8, 11, 26도 표 3.2와 중복하여 표 3.3에도 나타나지만 최종시점의 고객 상태는 탈퇴 (변경)가 되고, 유형변경일자가 탈퇴일자로 바뀌게 된다.

생명표 기법에 의한 누적 생존율을 구하기 위해서는 관찰기간을 일정 단위의 구간으로 정해야 하는데 본 논문에서는 편의상 4개월을 그 구간단위로 하였다. 각 구간에서의 관찰결과를 다시 정리하면 <표 3.4>와 같다.

표 3.4에서 나타난 결과를 토대로 2절에서 기술한 각 비율의 공식을 이용하여 계산된 회원유지비율 산출 결과는 다음의 표 3.5와 같다.

이 표에서 각 구간에서의 누적 생존율인 회원유지비율  $S_t$ 는 그 이전 구간의 구간생존율을 누적하여 곱하면 된다. 이 표로부터 회원유형별로 각각에 대한 회원유지비율을 파악할 수 있는 동시에 각각의 관찰구간에 대해 무료회원과 유료회원간의 회원유지비율을 비교할 수도 있다. 이를 좀 더 구체적으로 기

표 3.1 가상의 회원가입 후의 추적 관찰 결과

ID	유형	가입 일자	변경상황	유형 변경일자	최종상태	탈퇴일자	관찰월수	유료회원 생존월수	무료회원 생존월수
1	무료	2006.04.03			탈퇴	2006.06.15	20		2
2	무료	2006.04.16			유지		20		20+
3	유료	2006.05.12			탈퇴	2006.10.18	19	5	
4	무료	2006.06.16			탈퇴	2006.08.04	18		1
5	무료	2006.06.18			휴면	(2006.09.25)	18		(3+)
6	무료	2006.07.26	유료	2007.01.25	유지		17	11+	6
7	무료	2006.08.16			탈퇴	2006.11.02	16		2
8	유료	2006.08.23	무료	2006.09.22	휴면	(2006.12.20)	16	1	(2+)
9	유료	2006.09.27			유지		15	15+	
10	무료	2006.10.16			유지		14		14+
11	유료	2006.10.26	무료	2007.09.29	탈퇴	2007.11.19	14	11	1
12	무료	2006.11.11			탈퇴	2007.12.20	13		13
13	무료	2007.01.14			유지		11		11+
14	무료	2007.01.20	유료	2007.05.20	유지		11	7+	4
15	유료	2007.02.05			휴면	(2007.06.15)	10	4+	
16	무료	2007.03.07	유료	2007.05.06	탈퇴	2007.12.03	9	6	2
17	무료	2007.04.14			유지		8		8+
18	무료	2007.04.17			휴면	(2007.08.05)	8		(3+)
19	유료	2007.04.26			탈퇴	2007.12.15	8	7	
20	무료	2007.05.09			탈퇴	2007.09.03	7		3
21	유료	2007.06.24			유지		6	6+	
22	유료	2007.06.27			탈퇴	2007.08.30	6	2	
23	무료	2007.07.23	유료	2007.10.22	유지		5	2+	3
24	유료	2007.08.05			유지		4	4+	
25	유료	2007.08.19			탈퇴	2007.10.25	4	2	
26	유료	2007.09.23	무료	2007.11.22	유지		3	2	1+
27	무료	2007.09.26			탈퇴	2007.11.05	3		1
28	무료	2007.09.26			유지		3		3+
29	유료	2007.10.04			유지		2	2+	
30	유료	2007.11.13			유지		1	1+	

표 3.2 무료회원의 추적 관찰 결과

ID	가입 일자	최종 고객 상태	탈퇴 일자	관찰월수	생존월수
1	2006.04.03	탈퇴	2006.06.15	20	2
2	2006.04.16	유지		20	20+
4	2006.06.16	탈퇴	2006.08.04	18	1
5	2006.06.18	휴면	(2006.09.25)	18	(3+)
6	2006.07.26	탈퇴 (변경)	2007.01.25	17	6
7	2006.08.16	탈퇴	2006.11.02	16	2
8	2006.09.22	휴면	(2006,12.20)	16	(2+)
10	2006.10.16	유지		14	14+
11	2006.10.26	탈퇴	2007.11.19	14	1
12	2006.11.11	탈퇴	2007.12.20	13	13
13	2007.01.14	유지		11	11+
14	2007.01.20	유지 (변경)	2007.05.20	11	4
16	2007.03.07	탈퇴 (변경)	2007.05.06	9	2
17	2007.04.14	유지		8	8+
18	2007.04.17	휴면	(2007.08.05)	8	(3+)
20	2007.05.09	탈퇴	2007.09.03	7	3
23	2007.07.23	유지 (변경)		5	3
26	2007.09.23	유지		3	1+
27	2007.09.26	탈퇴	2007.11.05	3	1
28	2007.09.26	유지		3	3+

술하면 먼저 무료회원 중에서 회원가입 후 4개월 미만의 고객들의 회원 유지비율은 54.3%, 회원 가입 후의 기간이 4.0-7.9개월인 고객과 8.0-11.9개월인 고객의 회원 유지비율은 38.8%, 12.0-15.9개월인 고

표 3.3 유료회원의 추적 관찰 결과

ID	가입 일자	최종 고객 상태	탈퇴 일자	관찰월수	생존월수
3	2006.05.12	탈퇴	2006.10.18	19	5
6	2007.01.25	유지		17	11+
8	2006.08.23	탈퇴	2006.09.22	16	1
9	2006.09.27	유지		15	15+
11	2006.10.26	탈퇴	2007.09.29	14	11
14	2007.05.20	유지		11	7+
15	2007.02.05	휴면	(2007.06.15)	10	4+
16	2007.05.06	탈퇴	2007.12.03	9	6
19	2007.04.26	탈퇴	2007.12.15	8	7
21	2007.06.24	유지		6	6+
22	2007.06.27	탈퇴	2007.08.30	6	2
23	2007.10.22	유지		5	2+
24	2007.08.05	유지		4	4+
25	2007.08.19	탈퇴	2007.10.25	4	2
26	2007.09.23	탈퇴 (변경)	2007.11.22	3	2
29	2007.10.04	유지		2	2+
30	2007.11.13	유지		1	1+

표 3.4 구간별 추적 관찰 결과

회원 유형	관찰구간 (월)	$O_t$	$O_{t+1}$	$d_t$	$u_t$	$w_t$
무료	0.0-3.9	20	7	8	3	2
	4.0-7.9	7	5	2	0	0
	8.0-11.9	5	3	0	0	2
	12.0-15.9	3	1	1	0	1
	16.0-19.9	1	0	0	0	1
유료	0.0-3.9	17	10	4	0	3
	4.0-7.9	10	3	3	1	3
	8.0-11.9	3	1	1	0	1
	12.0-15.9	1	0	0	0	1
	16.0-19.9	0	0	0	0	0

표 3.5 생명표 기법에 의한 회원유지비율 산출 결과

회원 유형	관찰구간 (월)	$d_t$	$r_t$	$q_t$	$p_t$	$S_t$	$l_t$
무료	0.0-3.9	8	17.5	0.457	0.543	0.543	543
	4.0-7.9	2	7	0.286	0.714	0.388	388
	8.0-11.9	0	4	0.000	1.000	0.388	388
	12.0-15.9	1	2.5	0.400	0.600	0.233	233
	16.0-19.9	0	0.5	0.000	1.000	0.233	233
유료	0.0-3.9	4	15.5	0.258	0.742	0.742	742
	4.0-7.9	3	8	0.375	0.625	0.464	464
	8.0-11.9	1	2.5	0.400	0.600	0.278	278
	12.0-15.9	0	0.5	0.000	1.000	0.278	278
	16.0-19.9	0	0	-	-	-	-

객과 16.0-19.9개월인 고객의 회원 유지비율은 23.3%인 것을 알 수 있다. 유료회원 중에서 회원가입 후 4개월 미만의 고객들의 회원 유지비율은 74.2%, 회원 가입 후의 기간이 4.0-7.9개월인 고객의 회원 유

지비율은 46.4%, 8.0-11.9개월인 고객과 12.0-15.9개월인 고객의 회원 유지비율은 27.8%인 것을 알 수 있다. 또한 관찰구간이 4개월 미만, 4.0-7.9개월, 12.0-15.9개월인 경우에는 회원유지비율이 유료회원인 경우가 더 높고, 관찰구간이 8.0-11.9개월인 경우에는 회원유지비율이 무료회원인 경우가 더 높은 것으로 나타났다.

#### 4. 결론

본 논문에서는 한 인구의 사망실태를 효과적으로 표현하는 종합표인 생명표를 이용하여 회원의 수명을 분석함으로써 회원유형별 등록고객 수명에 대한 생명표 기법 활용의 가능성을 알아보았다. 그 결과, 회원가입 후 특정 기간의 회원 유형별 유지비율, 회원가입 후 특정 기간 내에서 탈퇴한 회원의 수 등을 알 수 있었다. 이러한 결과로부터 인터넷 사이트의 정책 방향을 결정할 수 있을 것으로 생각된다. 더 나아가 고객세분화와 조직의 다양한 전략 수립에 도움을 줄 수 있을 것으로 생각된다.

향후 과제로는 관찰 중단이나 로그인을 하지 않은 고객에 대해 보다 적절한 확률분포를 적용하는 연구와 아울러 로그 분석을 통해 얻어지는 회원들의 여러 가지 행동 양상을 고려하여 종합적으로 회원 수명을 분석할 수 있는 연구가 수행되어야 할 것으로 생각된다.

#### 참고문헌

- 고응린 (1979). <계량의학통론>, 신광출판사, 서울.
- 안윤옥 (1990). <실용 의학 통계론>, 서울대학교출판부, 서울.
- 정현수, 박성호 (2006). 고객생애가치 측정모형의 B2B 비즈니스 적용연구. <한국경영과학회 2006년 추계학술대회 논문집>, 259-270.
- 최정환 (2005). <과학적 경영을 위한 CRM >, 다산출판사, 서울.
- Ali, M., Cho, J., and Begum, M. (2005). Nonparametric Bayesian multiple comparisons for Geometric populations. *Journal of the Korean Data & Information Science Society*, **16**, 1129-1140.
- Berger, P. D., and Nasr, N. I. (1998). Customer lifetime value : Marketing models and applications. *Journal of Interactive Marketing*, **12**, 17-30.
- Cho, J., Cho, K., and Choi, S. (2006). Bayesian multiple comparison of Bivariate Exponential populations based on fractional Bayes factor. *Journal of the Korean Data & Information Science Society*, **17**, 843-850.
- Gupta, S., Lehmann, D. R. and Stuart, J. A. (2004). Customer lifetime value : Marketing models and applications. *Journal of Marketing Research*, **41**, 7-18.
- Kang, H., Han, S., and Sin, H. (2003). On the development of customer lifetime value evaluation model by using data mining techniques. *Journal of the Korean Data Analysis Society*, **5**, 927-936.

## Internet customer life analysis by membership pattern using life table

Hee-Chang Park<sup>1</sup>

Department of Statistics, Changwon National University

Received 15 December 2008, revised 7 January 2009, accepted 13 January 2009

### Abstract

Many internet companies are holding marketing activity through customer relationship management to satisfy complicated and diversified consumer demands. Use of customer lifetime value as a marketing metric tends to place greater emphasis on customer service and long-term customer satisfaction, rather than on maximizing short-term sales. And so many internet companies have been interested in customer lifetime value, which is a primary key for discovery customer values to promote the competitive power in their business fields. In this paper, we apply a life table technique to lifetime analysis of internet site customers by membership pattern and provide the opportunity using revised life tables in several kinds of internet companies.

*Keywords:* Customer lifetime value, customer relationship management, internet companies, life table, marketing metric.

---

<sup>1</sup> Professor, Department of Statistics, Changwon National University, Changwon, Gyeongnam 641-773, Korea. E-mail: hcpark@changwon.ac.kr