

인터넷 쇼핑몰에서의 축차분석법 활용 방안

박희창¹

창원대학교 통계학과

접수 2009년 8월 1일, 수정 2009년 9월 29일, 게재확정 2009년 10월 7일

요약

인터넷은 우리나라뿐만 아니라 세계 도처에서 인간의 일상생활과 전통적인 상거래의 패러다임을 변화시켰으며, 새로운 비즈니스 모델을 구축할 수 있는 무한한 기회를 제공하였다. 이로 인하여 최근에는 상거래 수단으로서 인터넷 쇼핑몰이 등장하였다. 인터넷 쇼핑몰이 경쟁력을 갖기 위해서는 효과적인 고객만족서비스가 제공되어야 한다. 이를 위해 고객들에게 적절한 마케팅 프로모션을 실시하기 위해서는 시간의 변화에 따른 고객들의 구매행위 패턴을 예측할 수 있는 동적인 분석 방법이 필요하다. 본 논문에서는 통계적 추정 방법 중의 하나인 축차분석법을 이용하여 유사한 품목들 간의 매출액을 비교함으로써 고객들의 구매행위 패턴을 예측을 통해 매출 향상을 도모하는 방안에 대해 연구하고자 한다.

주요용어: 고객관계관리, 우도비, 전자상거래, 축차분석.

1. 서론

오늘날 정보통신기술의 발달과 정보통신 네트워크의 확산으로 인해 많은 기업과 조직이 전자상거래 (electronic commerce)를 적극적으로 도입하고 있으며, 시장규모도 급격히 증가하고 있다 (이성진, 2005). 특히 인터넷상의 가상공간을 매장으로 활용하고자 하는 온라인 쇼핑몰은 핵심적인 유통전략 및 정보기술의 전략적 활용대상의 하나로서 주요 이슈로 부상한 바 있다 (조남재와 송길영, 1998).

인터넷 비즈니스 산업의 비약적 성장 구조 속에서도 수많은 인터넷 쇼핑몰이 폐업하는 현실을 감안 할 때, 인터넷 쇼핑몰이 지속적인 성장을 추구하기 위해서는 성과모형의 설정이 요구된다 (이찬, 2007). 국내 인터넷 쇼핑몰 업계의 동향을 살펴보면 시장의 수요는 증가하지 않는 상황에서 고객이 쇼핑몰을 선택할 수 있는 폭은 넓어졌고 인터넷 쇼핑몰간의 경쟁은 더욱 치열해지고 있으며, 고객들의 요구는 더욱 다양화되었다. 따라서 공급자인 인터넷 쇼핑몰이 시장을 주도하던 시대에서 수요자인 고객이 시장을 주도하는 고객 중심의 환경으로 변모하고 있다. 이를 위해 고객관계관리 (Customer Relationship Management; CRM)기법이 전자상거래에서도 등장하게 되었다. CRM은 고객의 요구를 충족시킴으로써 고객과의 신뢰관계를 구축하고 궁극적으로 수익성이 높은 고객을 장기적으로 유지하기 위한 광범위한 비즈니스 활동을 의미한다. CRM에서 고객들에게 적절한 마케팅 프로모션을 실시하기 위해서는 시간의 변화에 따른 고객들의 구매행위 패턴을 예측할 수 있는 동적인 분석 방법이 필요하다 (하성호와 백경훈, 2004). 또 다른 방법으로는 데이터마이닝의 여러 가지 기법을 적용할 수도 있다 (Lee와 Park, 2008; Park, 2008).

본 논문에서는 통계적 추정 방법 중의 하나인 축차분석법을 이용하여 유사한 품목들 간의 매출액을 비교함으로써 고객들의 구매행위 패턴을 예측을 통해 매출 향상을 도모하고자 한다. 축차분석법 (method

¹ (641-773) 경남 창원시 사림동 9번지, 창원대학교 통계학과, 교수. E-mail: hcpark@changwon.ac.kr

of sequential analysis)은 표본의 크기가 미리 정해지지 않고, 표본을 순차적으로 하나씩 추출하되 각 단계에서 그 때까지의 정보를 분석하여 그 결과에 따라 판정하는 방법이다. 표본을 순차적으로 추출하고 각 단계에서 분석한 결과를 이용하여 추가적인 표본추출의 실시여부를 결정하는 축차적 표본추출에서는 추가추출의 실시여부가 사전에 준비된 정지규칙 (stopping rule)에 따르게 된다. 따라서 축차적 표본추출을 근거로 하는 축차분석에서는 표본의 크기가 상수가 아닌 과거의 관측값에 종속되는 확률변수가 된다. 표본의 크기가 확률변수이므로 확률실험의 횟수를 효과적으로 제한할 수 있으며, 이 방법은 다른 검정방법에 비해 평균적으로 적은 관찰횟수로 신뢰도가 높은 분석결과를 얻을 수 있다는 장점을 가지고 있다. 축차적 표본추출방법은 1930년대 초보적인 통계적 품질관리의 한 방법으로 적용되었으며, Wald (1947)에 의해 이론적 배경이 완성되었고, Armitage (1969)에 의해 일반화되었으며 축차확률비 검정이나 집단 축차검정법 등으로 발전하게 되었다. 논문의 2절에서는 쇼핑몰에서의 축차분석의 적용 절차에 대해 기술하고, 3절에서는 예제를 통하여 축차분석법의 적용가능성을 타진한 후 4절에서 결론을 맺는다.

2. 축차분석법의 적용

백화점이나 온라인쇼핑몰에서 매장 담당자들은 기존의 제품보다 매출액이 더 많은 신제품을 입고해야 할지에 대해 고민을 하게 된다. 예를 들어 기존의 매출비율이 π_0 인 경우 신제품의 매출비율이 이보다 높은 π_1 이라는 업체의 주장을 어떻게 받아들일 수 있을지에 대해 고민이 생긴다. 만약 업체의 주장이 거짓임에도 불구하고 기존의 제품 대신 신제품을 받아들이면 매출이 낮아지므로 업체 전체적으로 볼 때 위험에 처해질 수 있다. 이러한 경우에는 일정 기간 동안 두 제품을 모두 진열하여 비교 관찰한 후 기존제품 대신 신제품을 진열할 것인지에 대해 결정을 하기 위해 축차분석법을 활용할 수 있다.

축차분석법은 표본을 사전에 설정해놓고 관찰하는 것이 아니라 대상이 입수되는 대로 실험을 하여 그 결과를 추가해가면서 판정하게 된다. 축차분석법의 절차는 다음과 같다. 먼저 귀무가설과 대립가설을 설정한다.

$$H_0 : \Pi = \pi_0$$

$$H_1 : \Pi = \pi_1$$

여기서 Π 는 백화점에서는 매장 방문 고객 중에서 구매한 고객의 비율을 의미하며, 온라인 쇼핑몰에서는 제품을 클릭한 고객 중에서 구매한 고객의 비율을 의미한다. π_0 는 기존 제품의 매출 비율이고, π_1 은 업체가 기존의 π_0 보다 높다고 주장하는 매출비율을 의미한다.

이에 대한 판정기준으로는 다음과 같은 우도비 (likelihood ratio)를 이용할 수 있다.

$$\lambda = \frac{\binom{n}{r} \pi_0^r (1 - \pi_0)^{n-r}}{\binom{n}{r} \pi_1^r (1 - \pi_1)^{n-r}} \quad (2.1)$$

여기서 n 은 방문고객의 수이고, r 은 구매한 고객의 수이다. 의사결정기준의 값을 λ_0 와 λ_1 ($\lambda_1 < \lambda_0$)이라 할 때, 만약 $\lambda > \lambda_0$ 이면 귀무가설을 받아들이는 것으로 즉, 신상품을 철수하는 것으로 판정하고, $\lambda < \lambda_1$ 이면 대립가설을 받아들이는 것으로 즉, 기존상품을 철수하는 것으로 판정하며, $\lambda_1 < \lambda < \lambda_0$ 이면 판정을 유보한다. 여기서 귀무가설이 참인 모집단에서 뿔히게 될 가능성을 $1 - \alpha$ 라고 하고 대립가설이 참인 모집단에서 뿔히게 될 가능성은 β 이므로 $\lambda_0 = (1 - \alpha)/\beta$ 이고, $\lambda_1 = \alpha/(1 - \beta)$ 이다. 따라

서 $\lambda > (1 - \alpha)/\beta$ 이면 귀무가설을 받아들이게 되고, $\lambda < \alpha/(1 - \beta)$ 이면 대립가설을 받아들인다. 식 (2.1)의 양변에 자연대수를 취하면 다음의 식을 얻을 수 있다.

$$\ln \lambda = r \ln\left(\frac{\pi_0}{\pi_1}\right) + (n - r) \ln\left(\frac{1 - \pi_0}{1 - \pi_1}\right) \quad (2.2)$$

식 (2.2)에서 $\lambda = \lambda_0$ 로, $r = r_0$ 로 두고 r_0 에 대해 정리하면 다음의 식을 얻는다.

$$r_0 = a_0 + bn \quad (2.3)$$

여기서

$$a_0 = \frac{\ln[(1 - \alpha)/\beta]}{\ln(\pi_0/\pi_1) - \ln[(1 - \pi_0)/(1 - \pi_1)]}, \quad b = \frac{\ln[(1 - \pi_0)/(1 - \pi_1)]}{\ln(\pi_0/\pi_1) - \ln[(1 - \pi_0)/(1 - \pi_1)]}$$

이다.

또한 식 (2.2)에서 $\lambda = \lambda_1$ 로, $r = r_1$ 로 두고 r_1 에 대해 정리하면 다음의 식을 얻는다.

$$r_1 = a_1 + bn \quad (2.4)$$

여기서

$$a_1 = \frac{\ln[\alpha/(1 - \beta)]}{\ln(\pi_0/\pi_1) - \ln[(1 - \pi_0)/(1 - \pi_1)]},$$

이다.

식 (2.3)과 식 (2.4)로 표시되는 두 직선을 축차분석을 위한 도표 상에 그려 놓고 방문고객의 수인 n 에 해당하는 구매 고객의 수 r 의 값을 도표 상에 표시한 후 만일 이 점이 두 직선 r_0 와 r_1 밖으로 나가게 되면 관찰을 중단하나, r_0 와 r_1 사이에 있으면 관찰을 계속하게 된다. 그 때 그 점이 r_0 밖으로 벗어나면 귀무가설을 받아들이고, r_1 밖으로 벗어나면 대립가설을 받아들이게 된다.

예를 들어 기존제품의 매출 비율이 20%이고, 신제품업체에서는 자사제품이 50%의 매출비율을 기대할 수 있다고 주장한다면 이러한 사실을 확인하기 위해 축차분석법을 이용하여 다음과 판정기준을 정할 수 있다. $\alpha = 0.05$, $\beta = 0.1$ 이라고 가정하면 두 직선 r_0 와 r_1 은 다음과 같이 구해진다.

$$r_0 = -1.624 + 0.339n$$

$$r_1 = 2.085 + 0.339n$$

n 에 해당하는 r 의 값을 타점해 나가다가 이 두 직선의 바깥쪽에 나타나면 관찰을 중단하고 진열제품을 두 제품 중에 하나를 선택할 수 있다. 이러한 축차분석법은 백화점 측의 희망하는 매출비율을 귀무가설로 두고 제조업체가 주장하는 매출비율을 대립가설로 설정하는 경우에도 활용이 가능하다.

3. 적용 예제

이 절에서는 쇼핑물의 매출을 신장시키기 위해 예제를 통하여 축차분석법의 적용가능성을 타진해보고자 한다. 먼저 기존제품의 구매비율 π_0 를 0.05에서부터 0.25까지 0.05씩 증가시켜가면서 n 값의 변화에 따른 r_0 와 r_1 값이 변하는 양상을 살펴본 결과 표 3.1과 같다. 여기서 $\alpha = 0.05$, $\beta = 0.1$ 이고, 신제품의 구매비율 $\pi_1 = 0.30$ 으로 하였다. 이 표에서 보는 바와 같이 π_0 의 값이 π_1 의 값에 가까워질수록 r_0 와 r_1 값의 차이가 점점 커져서 판정을 유보하고 관찰을 계속할 가능성이 더 많아진다. 보다 구체

표 3.1 기존제품의 구매비율과 방문고객수의 변화에 따른 판정기준값

π_0	0.05		0.10		0.15		0.20		0.25	
n	r_0	r_1	r_0	r_1	r_0	r_1	r_0	r_1	r_0	r_1
50	6.21	8.66	7.64	11.45	8.40	14.20	8.21	17.75	4.77	25.23
60	7.66	10.12	9.50	13.31	10.59	16.39	10.69	20.23	7.51	27.97
70	9.12	11.57	11.36	15.17	12.78	18.57	13.17	22.70	10.26	30.72
80	10.58	13.03	13.23	17.03	14.97	20.76	15.64	25.18	13.00	33.46
90	12.03	14.48	15.09	18.90	17.16	22.95	18.12	27.66	15.75	36.21
100	13.49	15.94	16.95	20.76	19.34	25.14	20.60	30.14	18.49	38.95
110	14.94	17.40	18.81	22.62	21.53	27.33	23.07	32.61	21.24	41.70
120	16.40	18.85	20.67	24.48	23.72	29.52	25.55	35.09	23.99	44.44
130	17.86	20.31	22.53	26.34	25.91	31.70	28.03	37.57	26.73	47.19
140	19.31	21.76	24.40	28.20	28.10	33.89	30.51	40.05	29.48	49.93
150	20.77	23.22	26.26	30.07	30.29	36.08	32.98	42.52	32.22	52.68
160	22.23	24.68	28.12	31.93	32.47	38.27	35.46	45.00	34.97	55.43
170	23.68	26.13	29.98	33.79	34.66	40.46	37.94	47.48	37.71	58.17
180	25.14	27.59	31.84	35.65	36.85	42.64	40.42	49.96	40.46	60.92
190	26.59	29.05	33.70	37.51	39.04	44.83	42.89	52.43	43.20	63.66
200	28.05	30.50	35.57	39.37	41.23	47.02	45.37	54.91	45.95	66.41
3,000	435.78	438.23	556.84	560.65	653.91	659.71	739.05	748.58	814.63	835.09
4,000	581.40	583.85	743.01	746.82	872.73	878.52	986.79	996.33	1,089.15	1,109.61
5,000	727.02	729.47	929.18	932.99	1,091.54	1,097.34	1,234.53	1,244.07	1,363.68	1,384.14
6,000	872.63	875.09	1,115.35	1,119.15	1,310.36	1,316.15	1,482.27	1,491.81	1,638.21	1,658.67
7,000	1,018.25	1,020.70	1,301.51	1,305.32	1,529.17	1,534.97	1,730.01	1,739.55	1,912.74	1,933.20
8,000	1,163.87	1,166.32	1,487.68	1,491.49	1,747.99	1,753.78	1,977.75	1,987.29	2,187.27	2,207.73
9,000	1,309.49	1,311.94	1,673.85	1,677.66	1,966.81	1,972.60	2,225.49	2,235.03	2,461.79	2,482.25
10,000	1,455.11	1,457.56	1,860.02	1,863.83	2,185.62	2,191.42	2,473.23	2,482.77	2,736.32	2,756.78

적으로 $n = 3,000$ 인 경우에 대해 알아보자. 이 경우 π_0 의 값이 0.05인 경우에는 r_0 와 r_1 의 값이 각각 435.78과 438.23이고, π_0 의 값이 0.25인 경우에는 각각 814.63과 835.09이므로 π_0 가 0.25일 때 유보가능성이 더 커지는 것으로 나타났다.

또한 $\alpha = 0.05$, $\beta = 0.1$ 인 상태에서 기존제품의 구매비율 π_0 를 0.10으로 고정하고 신제품의 구매비율 π_1 의 값을 0.15에서부터 0.35까지 0.05씩 증가시켜가면서 n 값의 변화에 따른 r_0 와 r_1 값이 변하는 양상을 살펴본 결과는 다음의 표 3.2와 같다. 이 표에서 보는 바와 같이 π_1 의 값이 커질수록 즉, π_1 의 값이 π_0 의 값과 차이가 많이 날수록 판정에 대한 유보가능성이 점점 줄어드는 것을 알 수 있다. 이에 대해서도 좀 더 구체적으로 $n = 10,000$ 인 경우에 대해 알아보자. 이 경우 π_1 의 값이 0.15인 경우에는 r_0 와 r_1 의 값이 각각 1,230.66과 1,241.78이고, π_1 의 값이 0.35인 경우에는 각각 2,060.58과 2,063.84이므로 π_1 이 0.35일 때 유보가능성이 더 작아지는 것으로 나타났다.

4. 결론

경쟁이 더욱 치열해지고 고객들의 요구가 더욱 다양화되고 있는 쇼핑물에서 고객들의 구매행위 패턴을 예측을 통해 매출 향상을 도모하기 위해서는 시간의 변화에 따른 동적인 분석 방법이 필요하다. 이를 위해 본 논문에서는 통계적 추정 방법 중의 하나인 축차분석법을 이용하여 쇼핑물에서의 구매행위에 대한 동적 분석을 실시하는 방안에 대해 고찰하였다.

예제를 통하여 축차분석법의 적용가능성을 타진한 결과, 기존 제품의 구매비율이 신제품 구매 비율에 가까워질수록 판정 기준값들의 차이가 점점 커져서 판정을 유보하고 관찰을 계속할 가능성이 더 많아진다는 사실을 알 수 있었다.

표 3.2 신제품의 구매비율과 방문고객수의 변화에 따른 판정기준값

π_1	0.15		0.20		0.25		0.30		0.35	
n	r_0	r_1	r_0	r_1	r_0	r_1	r_0	r_1	r_0	r_1
50	1.31	12.43	4.49	10.83	6.25	10.93	7.64	11.45	8.88	12.14
60	2.55	13.66	5.94	12.28	7.91	12.59	9.50	13.31	10.95	14.20
70	3.78	14.90	7.39	13.73	9.57	14.25	11.36	15.17	13.01	16.27
80	5.02	16.13	8.84	15.18	11.23	15.91	13.23	17.03	15.07	18.33
90	6.25	17.37	10.30	16.64	12.89	17.57	15.09	18.90	17.13	20.39
100	7.49	18.60	11.75	18.09	14.55	19.23	16.95	20.76	19.19	22.45
110	8.72	19.84	13.20	19.54	16.21	20.89	18.81	22.62	21.26	24.51
120	9.96	21.07	14.65	20.99	17.87	22.55	20.67	24.48	23.32	26.58
130	11.20	22.31	16.11	22.45	19.53	24.21	22.53	26.34	25.38	28.64
140	12.43	23.55	17.56	23.90	21.18	25.86	24.40	28.20	27.44	30.70
150	13.67	24.78	19.01	25.35	22.84	27.52	26.26	30.07	29.50	32.76
160	14.90	26.02	20.46	26.80	24.50	29.18	28.12	31.93	31.57	34.82
170	16.14	27.25	21.92	28.26	26.16	30.84	29.98	33.79	33.63	36.89
180	17.37	28.49	23.37	29.71	27.82	32.50	31.84	35.65	35.69	38.95
190	18.61	29.72	24.82	31.16	29.48	34.16	33.70	37.51	37.75	41.01
200	19.84	30.96	26.27	32.61	31.14	35.82	35.57	39.37	39.81	43.07
3,000	365.79	376.91	432.96	439.30	495.82	500.50	556.84	560.65	617.17	620.43
4,000	489.34	500.46	578.20	584.54	661.78	666.46	743.01	746.82	823.37	826.63
5,000	612.90	624.01	723.45	729.79	827.73	832.41	929.18	932.99	1,029.58	1,032.83
6,000	736.45	747.56	868.69	875.03	993.69	998.37	1,115.35	1,119.15	1,235.78	1,239.03
7,000	860.00	871.12	1,013.93	1,020.27	1,159.64	1,164.32	1,301.51	1,305.32	1,441.98	1,445.23
8,000	983.56	994.67	1,159.18	1,165.52	1,325.60	1,330.28	1,487.68	1,491.49	1,648.18	1,651.43
9,000	1,107.11	1,118.22	1,304.42	1,310.76	1,491.56	1,496.24	1,673.85	1,677.66	1,854.38	1,857.63
10,000	1,230.66	1,241.78	1,449.67	1,456.01	1,657.51	1,662.19	1,860.02	1,863.83	2,060.58	2,063.84

향후 연구로는 본 논문에서 고려한 축차분석법과 기존의 데이터마이닝 기법을 연동하여 보다 구체적인 판정 기준값에 대한 방안이 모색되어야 할 것이다.

참고문헌

이성진 (2005). 인터넷 쇼핑몰의 구매모델에 관한 연구. <한국컴퓨터정보학회논문집>, **10**, 199-204.
 이찬 (2007). 인터넷 쇼핑몰의 성과모형에 관한 연구. <한국컴퓨터정보학회논문집>, **12**, 281-290.
 조남재, 송길영 (1998). 기업 배경에 따른 사이버쇼핑몰 경영전략: 비교사례연구. <한국전자거래학회지>, **3**, 117-136.
 하성호, 백경훈 (2004). 백화점 고객의 구매 분석 및 고객관계관리 전략 적용. <경영과학>, **21**, 66-69.
 Armitage, P. (1969). Sequential analysis in therapeutic trials. *Annual Review of Medicine*, **20**, 425-430.
 Lee, K. W. and Park, H. C. (2008). Application of k-means clustering for association rule using measure of association. *Journal of the Korean Data & Information Science Society*, **19**, 925-936.
 Park, H. C. (2008). The proposition of conditionally pure confidence in association rule mining. *Journal of the Korean Data & Information Science Society*, **19**, 1141-1151.
 Wald, A. (1947). *Sequential analysis*, Wiley, New York.

Application of sequential analysis in internet shopping malls

Hee Chang Park¹

Department of Statistics, Changwon National University

Received 1 August 2009, revised 29 September 2009, accepted 7 October 2009

Abstract

The Internet has changed the daily lives of human being in Korea and elsewhere in the world. It has changed the paradigms of traditional commercial activities and created immense opportunities for new business models. Recently, there has been much attention to the internet shopping mall as a means of commercial transaction. To make internet shopping mall competitive, effective customer satisfaction service should be provided and it is necessary to dynamic analysis method for customers' purchasing pattern. In this paper we apply the sequential analysis to comparison of two kinds of sales through the analysis of customers' purchasing pattern.

Keywords: CRM, electronic commerce, likelihood ratio, sequential analysis.

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