

확산이론 관점에서 로지스틱 모형과 Bass 모형의 비교*

홍정식** · 구훈영***†

Comparison of the Bass Model and the Logistic Model from the Point of the Diffusion Theory

Jungsik Hong** · Hoonyoung Koo***

■ Abstract ■

The logistic model and the Bass model have diverse names and formulae in diffusion theory. This diversity makes users or readers confused while it also contributes to the flexibility of modeling. The method of handling the integration constant, which is generated in process of deriving the closed form solution of the differential equation for a diffusion model, results in two different 'actual' models. We rename the actual four models and propose the usage of the models with respect to the purpose of model applications. The application purpose would be the explanation of historical diffusion pattern or the forecasting of future demand. Empirical validation with 86 historical diffusion data shows that misuse of the models can draw improper conclusions for the explanation of historical diffusion pattern.

Keyword : Diffusion Theory, Bass Model, Logistic Model

1. 서 론

신기술, 신제품이나 신규서비스의 확산 모형에서

가장 많이 활용되는 모형은 로지스틱 모형과 Bass 모형이다. Bass 모형은 이를 토대로 한 논문이 750편에 이를 정도이고[13], 로지스틱 모형은 1845년

논문접수일 : 2011년 08월 25일 논문게재확정일 : 2012년 05월 23일

논문수정일(1차 : 2012년 04월 11일, 2차 : 2012년 05월 23일)

* 이 연구는 서울과학기술대학교 교내 학술연구비 지원으로 수행되었음.

** 서울과학기술대학교 IT정책전문대학원 산업정보시스템전공

*** 충남대학교 경상대학 경영학부

† 교신저자

에 Verhurst[18]에 의해 처음 소개되어 1920년대에 Reed[15]에 의해 본격적으로 거론된 이후, 확산뿐 아니라 다양한 분야에서의 S자형 성장모형에 가장 널리 활용되고 있다[5]. 그런데 이 두 모형은 잠재 자형요를 제외하면 확산 곡선을 설명하는 모수가 두 개로 절약적인 모형(parsimonious model)의 장점을 지니고 있지만 설명력과 적합도의 한계도 존재한다. 따라서 이들 모형은 추가 변수나 추가 모수의 도입을 통해, 다양한 방식으로 확장된 모형이 제시되고 있다. 가령 2000년 이후에도 이중시장 모형이나 조각확산 모형(piecewise diffusion model)으로 확장되거나[2, 11, 12, 17], 확산저항요인을 도입하거나[14], 다양한 로지스틱 모형을 중첩시키는 방식으로[10] 모형을 확장한 연구들이 제시되고 있다.

다른 한편으로 이들 모형을 비교한 연구도 꾸준히 제시되고 있다. 단순 비교가 이루어진 연구에서는 주로 특정 실제 사례의 확산과정을 설명하고 미래의 수요를 예측하는 데에 두 모형을 사용하고 결과를 비교한 연구들이 있다[1, 7, 20]. 또한 로지스틱과 Bass 모형뿐 아니라 다른 모형도 포괄적으로 포함한 상태에서 데이터에 대한 적합도나 예측의 정확성면에서 비교를 수행한 연구가 있다[8, 9]. 또 다른 측면의 비교는 특정 확산 현상에 영향을 미친 요인에 대한 가설채택을 두고 두 모형을 비교하는 것이다[4, 19]. 데이터에 대한 적합도나 가설채택 여부를 넘어 두 모형의 분포가 유도되는 과정과 특징을 비교한 논문도 있다[3]. 그런데, 이들 논문에 등장하는 로지스틱 모형과 Bass 모형의 형태는 논문마다 약간씩 다르며 따라서 결과도 상이하게 나오게 된다. 명칭 또한 다양하게 사용된다. 예를 들면, 로지스틱 모형은 내부영향 모형(internal influence model)이라고도 불리며, Bass 모형은 혼합영향 모형(mixed influence model)이라고 불리기도 하는데, 이 혼합영향 모형은 논문에 따라, 모수를 3개 포함한 것[6], 4개 포함한 것이 있고[4], 후자는 확장된 로지스틱 모형으로 불리기도 한다[8]. 형태와 명칭의 이러한 다양성은 정리될 필요가 있고, 이러한

과정에서 두 모형의 보다 근본적인 비교와 분석이 이루어질 수 있을 것이다.

본 논문에서는 확산 모형의 요인을 설명해주는 미분방정식에서 확산 모형의 닫힌 해(closed form solution)가 유도되는 과정과 이 과정에서 적분 상수가 처리되는 과정을 고찰함으로써 하나의 모형수식에서 두 개의 모형이 형성될 수 있음을 보인다. 여기서 두 개의 모형이 형성되는 것은 적분 상수를 다루는 두 가지 방식에 기인한다. 이 두 가지 방식은 (1) 초기치 설정을 통한 계산의 방식으로 구하는 것과, (2) 적분 상수를 모수로 취급하여 전체 데이터의 적합도를 최적화하는 추정의 방식으로 구하는 것으로 구분할 수 있다. 따라서 로지스틱 모형과 Bass 모형 역시 두 개의 모형이 도출되며, 이에 대해 현재의 모형 사용 관습을 고려하여, 적절한 명칭을 부여한다. 또한 확산 모형을 사용함에 있어, 과거 확산패턴을 설명하기 위한 목적과 미래 확산 추이를 예측하기 위한 목적으로 구분하여 로지스틱 모형과 Bass 모형의 모형틀 내에서 도출된 각각의 모형이 적절하게 사용되어야 하는 방안을 제안한다. 특히, 확산과정의 요인을 판별하는 논문에서 로지스틱 모형과 Bass 모형이 비교 대상일 경우, 각 모형틀 내에서 도출된 4개의 모형 중에서 어떤 모형이 서로 비교의 짝을 이루어야하는지를 밝힌다. 또한, 이 4개의 모형을 86개의 국내의 확산 데이터에 적용하고 그 결과를 토대로 각 모형의 특징과 장단점을 분석하고자 한다.

본 논문은 다음과 같이 구성된다. 제 2장에서는 연구배경으로 로지스틱 모형과 Bass 모형의 간략한 소개가 제시된다. 제 3장에서는 로지스틱 모형과 Bass 모형을 해석적으로 비교분석하고, 제 4장에서는 실제 데이터를 이용하여 두 모형의 다양한 형태들의 조합에 대해 적합도 및 예측정확도 비교를 수행하며, 또한 이를 바탕으로 기존 확산과정에서의 확산요인을 판별하는 작업의 타당성과 한계가 논의될 것이다. 마지막으로 제 5장에서는 결론이 기술된다.

2. 로지스틱 모형과 Bass 모형

$$t_2 = -\frac{a}{b} \tag{6}$$

2.1 로지스틱 모형

로지스틱 모형은 다양한 방식으로 표현되나[5], 일단 논의를 신제품 확산모형 연구분야로 한정하면, 다음과 같은 미분방정식으로 표현된다.

$$\frac{dN(t)}{dt} = (m - N(t))\left(\frac{b}{m}N(t)\right) \tag{1}$$

식 (1)에서 b 는 내부영향 계수로 불리며, m 은 잠재수요(market potential)로 불린다. 또한 $(m - N(t))$ 는 잠재수요 중 시점 t 에서 미채택자 수를 나타낸다. 따라서 식 (1)은 채택자 수의 증가가 기존 채택자의 영향을 받아서 생겨난다는 것을 의미한다. 이러한 모형을 내부 영향 모형(internal influence model)이라 한다. 여기서 $\frac{N(t)}{m} = F(t)$ 로 두고, 위험률(hazard rate)을 이용하여 정리하면 다음과 같다.

$$\frac{f(t)}{1 - F(t)} = b F(t) \tag{2}$$

제 3장에서 자세히 기술되지만, 로지스틱 모형은 $F(0) = 0$ 가 성립하지 않으므로 엄밀히 말하면 분포함수 $F(t)$ 표현을 쓸 수 없다. 그러나 본 논문에서는 편의상 $\frac{N(t)}{m} = F(t)$ 를 분포함수로 칭하기로 한다.

이러한 수식은 미분함수가 들어 있는 미분방정식이며, 이를 풀어서 표현하면 다음과 같다.

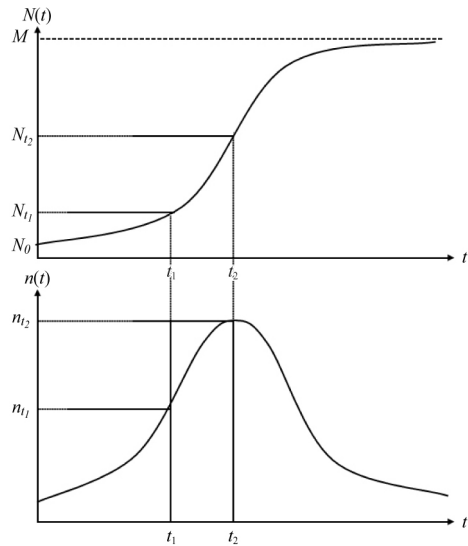
$$F(t) = \frac{1}{1 + ae^{-bt}} \tag{3}$$

$$N(t) = m \left[\frac{1}{1 + ae^{-bt}} \right] \tag{4}$$

로지스틱 모형의 당기 채택자 증가율이 최대가 되는 시점(t_1)과 당기 채택자 수가 최대가 되는 시점(t_2)은 각각 당기 채택자수와 누적 채택자수의 변곡점으로 다음과 같이 나타낼 수 있다.

$$t_1 = \frac{\ln 2 - a}{b} \tag{5}$$

로지스틱 모형에서 시간에 따라 누적 채택자수와 당기채택자수의 변화가 [그림 1]에 나와 있다.



[그림 1] 로지스틱 모형

2.2 Bass 모형

Bass 모형은 채택자 수의 증가가 기존 채택자수에 영향을 받는 것과 이와 무관하게 영향을 받는 것, 즉 두 가지 영향에 의해 생겨난다고 가정한다. 로지스틱 모형과 차이를 두기 위해 전자의 영향계수를 q 로 표기하고 후자의 영향계수를 p 로 표시한다. p 는 외부영향 계수라고 불린다. 이를 식 (1)과 같이 미분방정식으로 표현하면 다음과 같다.

$$\frac{dN(t)}{dt} = p(m - N(t)) + \frac{q}{m} N(t)(m - N(t)) \tag{7}$$

여기서, 로지스틱 모형과 마찬가지로 m 은 잠재수요이다. 따라서 식 (7)은 채택자 수의 증가가 마케팅이나 가격 등 외부의 영향과 함께 기존 채택자의 영향, 즉, 내부 영향을 동시에 받아서 생겨난다는 것을 의미한다. 이러한 모형을 혼합 영향 모형(mixed influence model)이라 불린다. q 가 0인 경우는 외부 영

항 모형(external influence model)이라 불린다. 로지스틱 모형과 마찬가지로 위험률을 사용하여 식 (7)을 다시 정리하면 다음과 같다.

$$\frac{f(t)}{1-F(t)} = p + qF(t) \tag{8}$$

식 (8)을 풀면 다음 수식이 유도된다.

$$F(t) = \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}} \tag{9}$$

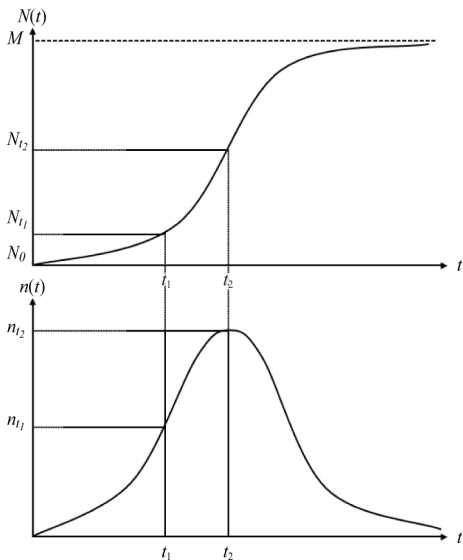
$$N(t) = m \left[\frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}} \right] \tag{10}$$

Bass 모형의 t_1 와 t_2 역시 로지스틱 모형과 같은 방식으로 유도되며 다음과 같다.

$$t_1 = \frac{1}{p+q} \ln \frac{q}{(2 + \sqrt{3})p} \tag{11}$$

$$t_2 = \frac{1}{p+q} \ln \frac{q}{p} \tag{12}$$

Bass 모형의 $N(t)$, $n(t)$ 를 그림으로 표시하면 [그림 2]와 같다.



[그림 2] Bass 모형

3. 로지스틱 모형과 Bass 모형의 분석 및 비교

로지스틱 모형과 Bass 모형에 있어 명칭 활용상의 혼란의 원인을 파악하기 위해서는 이들 모형의 미분 방정식(각각 식 (1)과 식 (7))으로부터 모형의 닫힌 해(각각 식 (3)과 식 (9))가 유도되는 과정을 고찰할 필요가 있다. 이러한 과정은 대부분의 논문이 상세히 다루고 있지 않다. 특히, 식 (1)이 변형된 식 (2)로부터 식 (3)이 도출되는 과정은 간단한 미분방정식의 풀이과정이므로 논문에서 다루어지지 않고 있고, 식 (7)의 변형인 식 (8)에서 식 (9)가 유도되는 과정 또한 좀 더 복잡한 미분 방정식의 풀이 과정으로 Bass 모형을 다루는 논문에서 좀처럼 다루어지지 않고 있다.

한편, 확산과정을 설명하는 미분방정식의 관점에서 보면, 로지스틱 모형이 Bass 모형의 특정 형태임에도 불구하고 적합도 성능비교에서 더 우월한 경우가 종종 발생한다. 이러한 이유 역시 미분 방정식으로써의 확산 모형과 이의 닫힌 해로써의 확산 모형의 차이에서 기인한다(닫힌 해의 형태로써의 확산 모형과 미분방정식 형태로써의 확산 모형을 구분하기 위해 후자를 때로 모형틀이라고 명명한다). 이제 각각의 모형의 미분 방정식에서 $N(t)$ 의 닫힌 해(closed-form solution)가 유도되는 과정을 살펴보자.

우선 식 (2)를 보자. 식 (2)의 우변항의 $F(t)$ 를 좌변으로 옮기면 다음과 같다.

$$\frac{f(t)}{[1-F(t)]F(t)} = b \tag{13}$$

식 (13)의 양변을 적분하면,

$$\int \frac{f(t)}{[1-F(t)]F(t)} dt = \int b dt \tag{14}$$

$$\ln \frac{F(t)}{1-F(t)} = bt + c \tag{15}$$

식 (15)에서 적분상수 c 가 나오게 되며, 이 적분 상수 c 는 미분방정식에서 초기해(initial condition)를 설정함으로써 특정 값을 취하게 된다. $F(t)$ 는 채택시간의 누적분포함수이므로, 초기해는 당연히 $F(0)=0$ 가 사용된다. 그런데 식 (15)를 보게 되면 $F(0)=0$ 라는 초기해 설정이 불가능하다. 이는 식 (1)로부터도 직관적으로 알 수 있다. $N(t)$ 의 변화량이 $N(t)$ 에만 의존하고 있기 때문에 $N(0)=0$ 가 되면 $N(t)$ 는 계속 0으로 남게 되는 것이다. Bass 모형의 경우는 외부영향 계수 p 가 있어 이러한 문제가 발생하지 않는다.

따라서, 대부분의 논문에서는 적분상수 c 를 그대로 두고 식 (15)를 풀어서 식 (3)이 나오게 되는 것이다.

$$F(t) = \frac{1}{1+ae^{-bt}}, \text{ 여기서 } a = e^c \text{이다.}$$

a 는 확산 모형의 미분방정식에 나오지 않아서 모수의 의미를 b 만큼 명시적으로 제시하기는 힘들지만 식 (3)을 다음과 같이 바꾸어보면 a 의 의미는 어느 정도 파악이 가능하다.

$$N(t) = m \left[\frac{1}{1+e^{-b(t-\tau)}} \right].$$

여기서 τ 는 위치모수라 불리며, $N(t)$ 의 변곡점을 뜻한다. 따라서 a 가 클수록 τ 가 커지며, 확산 현상의 당기 채택자수가 최대에 이르는 정점이 뒤로 밀리게 되므로 확산 현상이 지체된다는 것을 의미한다.

Hu[4]에서는 $N(0)=m_0$ 라는 초기치 설정을 하고, c 를 구하는 m_0 를 사용한 표현을 쓰기도 한다.

그러면, 식 (15)에서 $F(0)=\frac{m_0}{m}$ 를 대입하면,

$$\ln \left(\frac{\frac{m_0}{m}}{1-\frac{m_0}{m}} \right) = c, \quad a = e^c = \frac{\frac{m_0}{m}}{1-\frac{m_0}{m}} = \frac{m_0}{m-m_0}.$$

따라서, 식 (3)을 m_0 를 사용하여 표현하면,

$$N(t) = m \left[\frac{1}{1 + \frac{m_0}{m-m_0} e^{-bt}} \right]. \quad (16)$$

그런데, m_0 는 사실 대기 수요가 존재하는 신약이나 자동차의 신모델 등의 몇몇 신제품을 제외하면, 거의 0에 가까운 경우가 많다. 이 때 m_0 를 0에 가까운 값으로 설정하고 b 만을 추정하는 것은 로지스틱 모형의 적합도를 떨어뜨리는 결과를 초래한다. 그래서 [4]에서는 m_0 를 모수로 취급하여 데이터로부터 추정을 한다. 따라서 이는 실제 제품 출시에 대기하고 있는 채택자 수를 나타내는 것이 아니며, 전체 적합도를 향상시키기 위한 모수값인 것이다. [4]에서는 혼합영향 모형과 내부영향 모형, 외부영향 모형의 적합도를 다양한 확산 데이터를 이용하여 성능 비교한다고 주장하고 있는데, m_0 를 상수가 아닌 모수로 추정함으로써 이미 내부영향 모형이 아닌 모형을 내부영향 모형이라고 사용하게 된 것이다. [4]의 결과는 혼합영향 모형이 가장 우수한 것으로 나타났으나, 경우에 따라 잘못 사용된 내부영향 모형이 더 좋은 성능을 나타낼 수 있으며, 이러한 경우 확산과정이 내부영향에만 지배받는다라는 잘못된 결론을 내릴 수 있다. 이것에 대해서는 제4장에서 실제 데이터를 이용하여 다시 설명하도록 한다.

다음으로 식 (8)에서 식 (9)가 유도되는 과정을 보자. 식 (8)은 다음과 같은 형태로 표현된다.

$$f(t) = p + (q-p)F(t) - qF^2(t) \quad (17)$$

이는 리카티 수식(Ricatti Equation)의 단순형에 속하며[16], 이를 푸는 과정은 로지스틱 모형보다는 복잡하다. 자세한 내용은 [17]를 참조하라. 미분방정식 풀이를 통해 $F(t)$ 는 다음과 같이 얻어진다.

$$F(t) = \frac{c - \frac{q}{p+q} e^{-(p+q)t} + e^{-(p+q)t}}{c - \frac{q}{p+q} e^{-(p+q)t}} \quad (18)$$

로지스틱 모형과 달리 $F(0)=0$ 라는 초기해 설정을 통해 적분상수 c 가 다음 값을 갖게 된다.

$$c = \frac{-p}{p+q} \quad (19)$$

이 c 값을 식 (18)에 대입하면 식 (9)가 얻어진다. 식 (2)와 식 (8)을 비교하면, 로지스틱 모형이 Bass 모형의 특수형태라는 것을 알 수 있다. 즉 $q=b$, $p=0$ 으로 두면 Bass 모형에서 로지스틱 모형이 유도된다. 그런데 식 (9)에서 $q=b$, $p=0$ 의 설정 자체가 불가능하고, 따라서 식 (9)로부터 식 (3)이 유도되지 않는다. 한편, 식 (18)에서 $q=b$, $p=0$ 라고 두면 $F(t)$ 는 다음과 같다.

$$F(t) = \frac{c}{c - e^{-bt}} = \frac{1}{1 - \frac{1}{c}e^{-bt}} \quad (20)$$

식 (20)을 보면, 즉, $a = -\frac{1}{c}$ 로 두면 식 (20)이 바로 식 (3)이 되며, 로지스틱 모형이 Bass 모형의 특수형태라는 것을 알 수 있다. 이상의 논의를 요약하면, 기존 확산 논문에서 사용되는 대부분의 Bass 모형의 확산 모형식은 $F(0)=0$ 라는 초기치 설정을 둔 상태에서 얻어진 것이고, 이를 통해 적분상수가 제거되어 수식이 간단해졌다는 것이다. 그러나, 반대급부로는 로지스틱 모형을 특정 형태로 포함하지 못하고, 초기치를 제한함으로써 적합도 면에서의 손실을 가져온다는 것이다.

본 장의 모형 비교를 통해 다음의 사항을 확인하였다.

- 적분상수의 처리(적분상수 제거 또는 적분상수를 모수로 추정)에 따라 로지스틱 모형과 Bass 모형은 각각 2개의 모형이 존재할 수 있음
- 적분상수를 모수로 처리하면 혼합, 내부, 외부 등 영향요인에 대한 설명력을 훼손할 수밖에 없음
- $F(0)=0$, $F(1)=N(1)/m$ 등의 초기 조건을 이용해 적분상수를 제거하는 경우에 혼합영향 및 내부영향의 설명을 할 수 있음

따라서 연구 목적이 확산의 영향요인을 판별하여 과거 현상을 설명하고자 한다면 적분상수를 제거한 모형을 사용해야 한다. 또한 수요예측을 위해 데이터 적합도와 예측력을 극대화하기 위해서는 적분상수를 모수로 활용하여 적합도를 극대화하는 것이 바람직하다.

4. 실제 확산 데이터를 통한 모형 분석

실제 확산 데이터에 대해 로지스틱 모형과 Bass 모형의 적합도를 비교할 때, 대부분의 논문은 식 (4)와 식 (9)를 토대로 한다. 그런데 제 3장에서 논한 바와 같이 식 (9)는 $F(0)=0$ 라는 초기치 제한이 있고, 식 (4)는 이러한 제한이 없기 때문에 로지스틱 모형이 더 유리하게 나올 가능성도 존재한다. 따라서, 두 모형을 비교할 때, 비교 대상을 좀 더 명확하게 설정할 필요성이 있다.

우선, 제 3장에서 기술한 대로 적분상수 처리와 관련하여 로지스틱 모형은 다음 두 가지가 존재한다 (<표 1> 참조).

- (1) 초기치를 설정하여 적분상수를 제거하는 경우 : 이 경우 $N(0)=0$ 는 불가능하므로 $N(1)$ 을 이용한다. $N(1)$ 의 실측값을 이용하여 적분상수를 제거한다. 따라서 모수 2개(m, b)의 모형이 된다. 이를 내부영향 모형(internal influence model)이라 하고, L1이라 하자.
- (2) 적분상수를 모수로 처리하여 데이터 전체로부터 추정하는 경우 : 이 경우는 일반적인 로지스틱 모형으로 잠재수요 m 을 포함하여 3개의 모수로 구성된다. 이를 로지스틱 모형(logistic model)이라 하고, L2라 하자.

Bass 모형 역시 적분상수 처리와 관련 다음 두 가지가 존재한다.

- (1) $N(0)=0$ 을 설정하여 적분상수를 제거하는 경우 : 이 경우 모수 3개(m, p, q)의 모형이 된다. 이를 혼합영향 모형(mixed influence model) 또는

Bass 모형이라 하고, B1이라 하자.
 (2) 적분상수를 모수로 처리하여 데이터 전체로부터 추정하는 경우 : 적분상수 c 까지 모두 모수 4개 (m, p, q, c)의 모형이 된다. 이를 확장(extended) Bass 모형이라 하고, B2라 하자.
 이상을 정리하면 <표 1>과 같다.

<표 1>에 정리된 4가지 모형의 비교를 위해, 실제 국내의 86개 제품 및 기술에 대한 확산 데이터를 이용하여 분석해 보았다. 사용된 해외 데이터(49개)는 Bass.org에서 구하였고, 국내 데이터(37개)는 통계청, 관련 협회 등을 통해 입수하였다. 이는 우리가 아는 한, 국내 확산관련 논문 중 가장 많은 데이터를 다루는 것이다. 각 모형의 모수 추정은 Matlab 2010을 이용하였으며, 사용된 추정법은 NLS(non-linear least square) 방법이다. 모형 간 성능 비교를 위해서는 MSE(평균제곱오차, Mean Squared Error), MAD(평균절대오차, Mean Absolute Deviation)와 AIC(Akaike Information Criterion)을 사용하였다. MSE의 경우, 확산 데이터 전체를 사용하여 각 모형이 실제 데이터에 적합되는 정도의 척도로 추정 성능

을 확인하기 위해 사용하였다. MAD의 경우는 예측 정확도의 척도로 사용되었는데, 각 확산 데이터별로 “ \lfloor 데이터 개수 $\rfloor/2$ ”개의 데이터는 MSE로 추정하고, 추정된 모형으로 나머지 기간에 대해 예측하여 예측 오차를 MAD로 계산하였다. AIC의 경우는, 모수의 개수가 다른 경우 모형의 성능 비교를 가능케 해주는 척도의 하나로 MSE나 MAD가 모수의 개수를 반영하지 못하는 문제점을 해결하기 위해 사용되었다.

<표 2>는 본 연구에서 분석된 86개 확산 데이터에 대한 각 성능척도별 비율평균과 순위평가를 요약한 것이다. 비율평균이란 모형별 성능척도값을 합산한 값에 비하여 해당 모형의 성능척도값이 어느정도 차지하는가를 의미하는 것으로 각 확산 데이터별로 상이한 성능척도값의 크기를 평준화하기 위해 사용하였다. 순위 평균은 각 성능척도값의 순위를 성능이 가장 우수한 것을 1, 가장 열등한 것을 4로 순위를 부여한 값의 평균이다(상세한 순위값은 부록의 <표 6>, <표 7> 참조).

MSE에 대한 결과를 보면 B2가 가장 좋고, B1과 L2는 유사하고 L1이 가장 좋지 않은 성능을 보이고 있다. MAD의 경우도 유사한 결과를 보이는데,

<표 1> 적분상수 처리에 따른 4가지 확산모형

모형틀	구분	모수	모형	적용목적
로지스틱 모형틀	L1	a, b	내부영향 모형 (internal influence model)	과거 설명
	L2	m, a, b	로지스틱 모형 (logistic model)	미래 예측
Bass 모형틀	B1	m, p, q	혼합영향 모형 (mixed influence model) 또는 Bass 모형	과거 설명 (미래 예측)
	B2	m, p, q, c	확장(extended) Bass 모형	미래 예측

<표 2> 86개 확산 데이터에 대한 성능척도 결과

성능 척도	MSE				MAD				AIC			
	B1	B2	L1	L2	B1	B2	L1	L2	B1	B2	L1	L2
비율 평균	13%	9%	62%	16%	21%	19%	33%	27%	25%	30%	19%	26%
순위 평균	2.21	1.19	3.86	2.73	1.84	1.60	3.65	2.87	2.84	3.99	1.03	2.14

기본적으로 이는 모수의 수가 $B2 > B1 = L1 > L2$ 의 관계를 갖고 있기 때문에 적합도와 예측력의 차이를 가져오는 것으로 이해할 수 있다.

반면에, AIC에 대한 결과를 보면 L1이 가장 좋고 L2, B1, B2의 순임을 알 수 있다. 이는 AIC가 “ $2 \cdot (\text{모수의 수}) - 2 \cdot \ln \text{최대우도값}$ ”로 계산되는데, 모수의 수가 L1이 가장 작고, $L2 = B1 > B2$ 라는 관계에서 비롯된 것과 로지스틱 모형은 잠재수요를 Bass 모형에 비해 상대적으로 작게 추정하는 경향에서 찾을 수 있다. 모수의 수가 중요한 이유는 적합도 측면에서 보면 모수의 수를 증가시킬수록 적합도는 좋아지기 때문에, 적합도와 추정의 편의성, 해석의 용이성 등을 고려하면 모수의 수와 성능의 우수성 사이의 적절한 수준을 결정할 필요가 있다. 또한, AIC에서 사용되는 로그 우도함수가 잠재수요에 대한 음의 선형식으로 표현되기 때문에 작게 추정된 잠재수요는 로그 우도값을 크게 만드는 결과를 가져온다. AIC에 대한 결과로 볼 때, 로지스틱 모형이 Bass 모형보다 모수에 비해 추정성능이 떨어지는 모형은 아니라는 것을 알 수 있다.

〈표 3〉 모형 간 성능차이 빈도 및 비율

적도	비교	빈도*	비율
MSE	B1-L1	77	90%
	B2-L2	79	92%
	B1-L2	70	81%
	B2-L1	83	97%
	B1-B2	7	8%
	L1-L2	0	0%
MAD	B1-L1	78	91%
	B2-L2	73	85%
	B1-L2	79	92%
	B2-L1	75	87%
	B1-B2	28	33%
	L1-L2	9	10%
AIC	B1-L1	1	1%
	B2-L2	0	0%
	B1-L2	14	16%
	B2-L1	0	0%
	B1-B2	76	88%
	L1-L2	80	93%

주) * 빈도 : B1-L1의 경우, B1 모형이 L1 모형보다 해당 성능척도가 더 우수한 경우의 수.

세 가지 성능척도에 대해 확산 데이터별로 우위를 관별해 보면 <표 3>의 결과를 얻을 수 있다. 적합도 측면과 예측정확도 측면에서 대체적으로 Bass 모형이 우수한 것으로 볼 수 있다. 모수의 수까지 고려하는 AIC의 경우는 로지스틱 모형이 우수한 것으로 볼 수 있다.

〈표 4〉 확산 데이터 유형별 분산분석 결과

구 분	적도	유의확률
국내	MSE	2.50E-19
	MAD	8.69E-05
	AIC	5.21E-05
국외	MSE	1.68E-71
	MAD	3.27E-26
	AIC	7.15E-24
일반소비재	MSE	4.86E-15
	MAD	5.69E-05
	AIC	0.002258
산업재	MSE	4.64E-05
	MAD	0.233153
	AIC	0.038652

이상의 성능척도별 모형별 성능차이를 통계적으로 확인하기 위해 제품 유형별 분산분석과 모든 조합(성능척도별-모형별 조합)에 대해 대응표본 t검정을 수행하였다(<표 4>와 <표 5> 참조). 분산분석은 국내와 국외, 산업용과 일반소비자용으로 구분하여 각 척도에 따른 모형 간 차이를 살펴보았는데, 산업용 제품의 MAD에 대해서만 분산분석 결과가 귀무가설이 채택되었다. 이는 대부분의 제품유형에 대해 모형 간 차이가 있음을 의미하는 것이다. 보다 상세한 분석을 위해 수행된 대응표본 t검정의 결과로, 유의 수준 0.05로 MSE 사용 시 (B1, L2)만 유의미한 차이가 없으며, MAD 사용 시는 (B1, B2)만 유의미한 차이가 없으며, AIC의 경우는 모두 유의한 차이를 보이는 것으로 나타났다. 이는 MSE 관점에서는 B1이나 L2는 동일한 평균을 갖는다는 의미이므로, 적합도 측면에서 B2과 L2는 동일하다고 볼 수 있다. MAD 관점에서는 B1과 B2가 동일한 수준의 예측력을 갖는다고 볼 수 있다. AIC 관점에서는 <표 2>의 결과가 통계적으로 유의하다는 것을 의미한다.

<표 5> 성능척도별 모형 간 t검정 결과

척도	대응	평균	표준 편차	95% 신뢰구간		t 값	유의확률
				하한	상한		
MSE	B1-L1	-0.494	0.364	-0.572	-0.416	-12.591	0.000
	B2-L2	-0.076	0.146	-0.107	-0.045	-4.861	0.000
	B1-L2	-0.032	0.178	-0.070	0.007	-1.646	0.103
	B2-L1	-0.539	0.315	-0.606	-0.471	-15.860	0.000
	B1-B2	0.045	0.189	0.004	0.085	2.191	0.031
MAD	L1-L2	0.462	0.268	0.405	0.520	15.986	0.000
	B1-L1	-0.128	0.170	-0.164	-0.091	-6.978	0.000
	B2-L2	-0.072	0.128	-0.100	-0.045	-5.237	0.000
	B1-L2	-0.060	0.121	-0.086	-0.034	-4.619	0.000
	B2-L1	-0.140	0.167	-0.176	-0.104	-7.781	0.000
AIC	B1-B2	0.012	0.100	-0.009	0.034	1.151	0.253
	L1-L2	0.068	0.087	0.049	0.087	7.207	0.000
	B1-L1	0.051	0.087	0.033	0.070	5.470	0.000
	B2-L2	0.047	0.113	0.023	0.071	3.856	0.000
	B1-L2	-0.011	0.021	-0.016	-0.007	-5.176	0.000
	B2-L1	0.110	0.133	0.081	0.138	7.638	0.000
AIC	B1-B2	-0.059	0.105	-0.081	-0.036	-5.166	0.000
	L1-L2	-0.063	0.090	-0.082	-0.043	-6.423	0.000

<표 6> 추정된 모수간의 상관 분석(MSE 기준 추정치)

모형과 모수	B1		B2			L1		L2	
	p	q	p	q	c	a	b	a	b
B1	p	1	-	-	-	-	-	-	-
	q	-0.132	1	-	-	-	-	-	-
B2	p	0.348	-0.302	1	-	-	-	-	-
	q	-0.063	0.843	-0.713	1	-	-	-	-
	c	-0.748	0.205	-0.387	0.209	1	-	-	-
L1	a	-0.274	0.205	-0.183	0.208	0.182	1	-	-
	b	-0.086	-0.904	0.050	-0.664	0.006	-0.052	1	-
L2	a	-0.456	0.510	-0.593	0.618	0.347	0.655	-0.230	1
	b	-0.123	-0.955	0.221	-0.823	-0.015	-0.097	0.939	-0.346

B1의 혼합영향 중 내부영향 모수가 L1 및 L2의 모수 b와 어떠한 관계를 갖는지 확인하기 위해, 4개 모형의 각 모수에 대해 상관분석을 수행하였다 (<표 6> 참조). 상관계수의 절대값이 0.9 이상인 강한 상관관계를 갖는 경우를 진한 이탤릭체로 표현하였다. 이 세가지 경우를 살펴보면, B1의 q와 L2의 b, L1의 b와 L2의 b, B1의 q와 L1의 b의 조합이다. 이는 Bass 모형의 내부영향 요인이 로지스틱 모형에서도 유사한 패턴으로 추정됨을 의미한다. 여기서 흥미로운 것은 <표 1>의 설명한 바처럼 B1의 q와 L1의 b

의 상관성이 가장 높게 나타나지는 않았다는 것이다. 이는 상관성이 동일성을 의미하는 것은 아니나 <표 1>의 주장과 어긋나는 것처럼 보일 수도 있다. 그러나, L2의 경우나, L1의 경우 모두 q와 b는 강한 상관성을 갖는다는 것을 의미하며, L1의 경우 상관성이 떨어진 이유는 L1의 추정방식에서 기인하는 바가 크다. L1의 모수 추정을 위해서는 실측값인 $N(1)$ 값을 사용하게 되는데, 추정된 L1 모형의 $\hat{N}(1)$ 와 $N(1)$ 의 편차가 그대로 추정성능을 떨어뜨리는 결과로 나타나기 때문이다.

이상의 비교 분석 결과, [4]에서 사용된 내부영향 모형은 실제로 L2 모형이었으며, 분석 대상에 따라 혼합영향 모형보다 성능이 우수하게 나올 수 있었음을 <표 6>로부터 확인할 수 있다. 만약 분석 대상이 <표 3>에서 언급된 L2 모형이 더 우수한 15개의 확산 데이터였다면, 확산과정의 요인분석에서 내부적 영향이 지배적인 것으로 결론이 났을 것이다. 또한, [8]의 결과에서처럼, 내부영향 요인만을 설명한다고 알려진 로지스틱 모형이 우수한 적합도를 보인 이유는 L1(내부영향) 모형이 아닌 L2(로지스틱) 모형을 사용하였기 때문이다. 즉, 연구 목적에 따라 적절한 모형을 선택하지 못하거나, 구분을 명확하게 하지 않으면, 잘못된 결과를 얻을 수 있음을 알 수 있다.

5. 결 론

기존 확산이론에서 Bass 모형과 로지스틱 모형은 다양한 명칭과 수식으로 혼란을 야기해 왔다. 제 3절의 해석적 비교를 통해 혼란의 원인은 적분상수 처리를 혼재적으로 적용하였기 때문임을 확인하였다. 이는 확산요인에 대한 설명과 미래를 위한 예측을 동시에 취급하려 했기 때문이다. 본 논문에서는 설명력 관점과 예측력 관점을 보다 명확히 하여, 혼란을 야기하며 사용되었던 로지스틱 모형과 Bass 모형을 비교 분석함으로써 연구목적에 따른 적절한 모형 선택의 기준을 제시하였다.

본 논문에서 얻어진 주요 결과는 다음과 같다.

첫째, 이론적인 측면에서 Bass 모형과 로지스틱 모형의 닫힌 해가 도출되는 과정을 살펴보고, 각각의 모형들에서 적분 상수가 처리되는 과정에 따라 두 개의 모형이 얻어짐을 입증하였다.

둘째, 확산 과정에 영향을 주는 요인을 판별하는 문제와 향후 확산수요를 예측하는 문제를 구별하고, 각각의 문제에 적합한 모형을 제시하였다. 이에 따라, 통상적으로 사용되는 Bass 모형과 로지스틱 모형이 확산 요인을 판별하는 문제에 사용될 경우

잘못된 결론에 다다를 수 있음을 방대한 확산 데이터에 대한 다양한 통계분석을 통해 입증하였다.

셋째, Bass 모형의 경우 초기치 설정이 $F(0)=0$ 이므로, 초기에 확산 수요가 급격하게 오르는 경우를 제외하면, 적분 상수를 포함하지 않는 기존의 Bass 모형이 향후 수요예측에서도 충분히 활용가능함을 데이터 분석을 통해 보여 주었다.

이상의 연구결과를 활용하면, 확산이론 관점에서 로지스틱 모형들 내의 두 개의 모형과 Bass 모형들 내의 두 개의 모형을 연구목적에 맞게 선택적으로 적용하는데 도움이 될 것으로 판단된다. 다만, 로지스틱 모형의 경우 확산 요인을 판별하는 문제를 접근할 때, 적분상수를 계산하는 방식에 대한 주의 깊은 접근이 요구된다.

참 고 문 헌

- [1] 홍정식, 김태구, 임달오, “확산 모형에 의한 고가의 의료기기의 수요 확산의 특성분석 및 중장기 수요예측에 관한 연구”, 『보건행정학회지』, 제18권, 제4호(2008), pp.85-110.
- [2] 홍정식, 엄석준, “Bass 확산 모형의 이분 확장”, 『한국경영과학회지』, 제34권, 제4호(2009), pp. 15-26.
- [3] Bemmaor, A.C. and J. Lee, “The Impact of Heterogeneity and Ill-Conditioning on Diffusion Model Parameter Estimates,” *Marketing Science*, Vol.21, No.2(2002), pp.209-220.
- [4] Hu, Q., C. Saunders and M. Gebelt, “Research Report : Diffusion of Information Systems Outsourcing : A Reevaluation of Influence Sources,” *Information Systems Research*, Vol.8, No.3(1997), pp.288-301.
- [5] Kingsland, S., “The Refractory Model : The Logistic Curve and the History of Population Ecology,” *The Quarterly Review of Biology*, Vol.57, No.1(1982), pp.29-52.
- [6] Li, Y. and Z. Chen, “Diffusion of Innovations

- in a Small World Network,” *Wireless Communications, Networking and Mobile Computing*, (2007), pp.5617–5620.
- [7] Meade, N., “Forecasting Using Growth Curves—An Adaptive Approach,” *The Journal of the Operational Research Society*, Vol.36, No.12 (1985), pp.1103–1115.
- [8] Meade, N. and T. Islam, “Forecasting with growth curves : An empirical comparison,” *International Journal of Forecasting*, Vol.11, No.2(1995), pp.199–215.
- [9] Meade, N. and T. Islam, “Technological Forecasting—Model Selection, Model Stability, and Combining Models,” *Management Science*, Vol.44, No.8(1998), pp.1115–1130.
- [10] Meyer, P.S., J.W. Yung, and J.H. Ausubel, “A primer on logistic growth and substitution : The mathematics of the loglet lab software,” *Technological Forecasting and Social Change*, Vol.61(1999), pp.247–71.
- [11] Muller, E. and G. Yogev, “When does the majority become a majority? Empirical analysis of the time at which main market adopters purchase the bulk of our sales,” *Technological Forecasting and Social Change*, Vol.73(2006), pp.1107–1120.
- [12] Niu, S., “A Piecewise-Diff-usion Model of New-Product Demands,” *Operations Research*, Vol.54, No.4(2006), pp.678–695.
- [13] Ozkaya, E., Demand management in global supply chains, Ph. D. Thesis, Georgia Institute of Technology, 2008.
- [14] Philips, F., “On S-curves and tipping points,” *Technological Forecasting and Social Change*, Vol.74(2007), pp.715–730.
- [15] Reed, L.J. and R. Pearl, “On the summation of logistic curves,” *Journal of the Royal Statistical Society*, Vol.90, No.4(1927), pp. 729–746.
- [16] Satoh, D., “A discrete bass model and its parameter estimations,” *Journal of the Operations Research Society of Japan*, Vol.44, No.1 (2001), pp.1–18.
- [17] Van den Bulte, C. and Y.V. Joshi, “New Product Diffusion with Influentials and Imitators,” *Marketing Science*, Vol.26, No.3(2007), pp. 400–21.
- [18] Verhulst, P.F., “Recherches mathematiques sur la loi d’accroissement de la population,” *Mem. Acad. Roy. Belg.*, Vol.18(1845), pp.1–38.
- [19] Wang, H. and H. Doong, “Diffusion of mobile music service in Taiwan : an empirical investigation of influence sources,” *Management Decision*, Vol.48, No.9(2010), pp.1378–1387.
- [20] Wu, F. and W. Chu, “Diffusion models of mobile telephony,” *Journal of Business Research*, Vol.63, No.5(2010), pp.497–501.

〈부 록〉 86개 확산데이터 분석 결과

〈표 7〉 37개의 국내 확산 데이터 분석 결과(각 지표별 순위임)

#	제품/서비스/기술	MSE (Mean Squared Error)				MAD (Mean Absolute Deviation)				AIC(Akaike Information Criterion)			
		B1	B2	L1	L2	B1	B2	L1	L2	B1	B2	L1	L2
1	가스 렌지	2	1	4	3	2	1	4	3	3	4	1	2
2	건설용 크레인	2	1	4	3	2	1	4	3	3	4	1	2
3	경운기	2	1	4	3	2	1	4	3	3	4	1	2
4	광케이블	1	4	3	2	1	4	3	2	2	4	1	3
5	기계용 내연기관	2	1	4	3	2	1	4	3	3	4	1	2
6	김치 냉장고	4	1	3	2	4	1	3	2	3	4	1	2
7	녹음기	2	1	4	3	2	1	4	3	3	4	1	2
8	도메인 등록	2	4	3	1	2	4	3	1	3	4	1	2
9	무선호출	4	1	3	2	4	1	3	2	3	4	1	2
10	선반	2	1	4	3	2	1	4	3	3	4	1	2
11	선풍기	2	1	4	3	2	1	4	3	3	4	1	2
12	세탁기	1	2	4	3	1	2	4	3	3	4	1	2
13	액정모니터	2	1	4	3	2	1	4	3	3	4	1	2
14	엘리베이터	3	1	4	2	3	1	4	2	3	4	1	2
15	영상녹화 재생기	2	1	4	3	2	1	4	3	3	4	1	2
16	위성방송 서비스	4	1	3	1	4	1	3	1	2	4	1	3
17	위성방송 수신기	2	1	4	3	2	1	4	3	3	4	1	2
18	이동전화 등록	3	1	4	2	3	1	4	2	3	4	1	2
19	자판기	4	1	3	2	4	1	3	2	3	4	1	2
20	전기냉장고	2	1	4	3	2	1	4	3	3	4	1	2
21	전기밥솥	2	1	4	3	2	1	4	3	3	4	2	1
22	전자계산기	2	1	4	3	2	1	4	3	2	4	1	3
23	전자레인지	2	1	4	3	2	1	4	3	3	4	1	2
24	전자복사기	2	1	4	3	2	1	4	3	3	4	1	2
25	전축	2	1	4	3	2	1	4	3	3	4	1	2
26	전화기	4	1	3	2	4	1	3	2	3	4	1	2
27	캠코더	2	1	4	3	2	1	4	3	3	4	1	2
28	컬러티비	2	1	4	3	2	1	4	3	3	4	1	2
29	키폰전화기	2	1	4	3	2	1	4	3	3	4	1	2
30	타자기	2	1	4	3	2	1	4	3	3	4	1	2
31	팩시밀리	3	1	4	2	3	1	4	2	3	4	1	2
32	프레스기	2	1	4	3	2	1	4	3	3	4	1	2
33	현금지급기	2	1	4	3	2	1	4	3	4	3	1	2
34	흑백티비	3	1	4	2	3	1	4	2	3	4	1	2
35	CDP	2	1	4	3	2	1	4	3	3	4	1	2
36	MP3	4	1	3	2	4	1	3	2	3	4	1	2
37	X-DLS	1	3	4	2	1	3	4	2	2	4	1	3
	평균	2.35	1.24	3.78	2.59	1.84	1.78	3.49	2.81	2.92	3.97	1.03	2.08

<표 8> 49개의 국외 확산 데이터 분석 결과(각 지표별 순위임)

#	제품/서비스/기술	MSE				MAD				AIC			
		B1	B2	L1	L2	B1	B2	L1	L2	B1	B2	L1	L2
1	Aftermarket PC Monitors	2	1	4	3	2	1	4	3	3	4	1	2
2	Aftermarket Remote Controls	2	1	4	3	2	1	4	3	3	4	1	2
3	Analog Color TV	2	1	4	3	2	1	4	3	3	4	1	2
4	Analog Color TV with Stereo	2	1	4	3	2	1	4	3	3	4	1	2
5	Analog Handheld LCD Color TV	2	1	4	3	2	1	4	3	3	4	1	2
6	Analog Handheld LCD Monochrome TV	2	1	4	3	2	1	4	3	3	4	1	2
7	Analog Projection TV	2	1	4	3	2	1	4	3	3	4	1	2
8	Analog TV/VCR Combinations	2	1	4	3	2	1	4	3	3	4	1	2
9	Blank Audio Cassette	2	1	4	3	1	2	4	3	2	4	1	3
10	Blank Videocassettes	2	1	4	3	2	1	4	3	3	4	1	2
11	Boomboxes(non-CD)	2	1	4	3	1	2	4	3	2	4	1	3
12	Caller ID Devices	2	1	4	3	1	2	4	3	2	4	1	3
13	Camcorders	2	1	4	3	2	1	4	3	3	4	1	2
14	Cellular Phones	1	3	4	2	2	1	4	3	3	4	1	2
15	Compact Audio Systems	2	1	4	3	2	1	4	3	3	4	1	2
16	Corded Telephones	2	1	4	3	1	2	4	3	3	4	1	2
17	Cordless Telephones	2	1	4	3	2	1	4	3	3	4	1	2
18	DBS Satellite	4	1	3	2	4	3	2	1	3	4	1	2
19	Digital Cameras	2	1	4	3	2	1	4	3	3	4	1	2
20	Digital Direct-View Sets and Monitors	2	1	4	3	2	1	4	3	3	4	1	2
21	Digital Projection Sets and Monitors	2	1	4	3	1	3	4	2	2	4	1	3
22	Digital TV Sets and Monitors	2	1	4	3	2	1	4	3	3	4	1	2
23	DVD Players/Recorders	2	1	4	3	2	1	4	3	3	4	1	2
24	Family Radio Devices	2	1	4	3	2	1	4	3	3	4	1	2
25	Fax Machines	2	1	4	3	1	2	4	3	3	4	1	2
26	Home Radios	2	1	4	3	1	2	4	3	3	4	1	2
27	Home Theater-in-a-Box	3	1	4	2	3	4	1	2	3	4	1	2
28	Laserdisc Player	3	1	4	2	3	4	2	1	3	4	1	2
29	LCD TV(Digital and analog)	2	1	4	3	2	1	4	3	3	4	1	2
30	Modems / Fax Modems	1	3	4	2	1	3	4	2	3	4	1	2
31	Monochrome TV	2	1	4	3	2	1	4	3	1	4	3	2
32	PC Printers	2	1	4	3	2	1	4	3	3	4	1	2
33	Personal Computers	2	1	4	3	2	1	4	3	3	4	1	2
34	Personal Digital Assistants(PDA)	2	1	4	3	1	2	4	3	3	4	1	2
35	Personal Wordprocessors	2	1	4	3	2	1	4	3	2	4	1	3
36	Plasma DTV	1	4	3	2	3	1	2	4	3	4	1	2
37	Portable and Transportable Navigation	2	1	4	3	1	2	4	3	3	4	1	2
38	Portable CD Equipment	2	1	4	3	2	1	4	3	3	4	1	2
39	Portable Headset Audio	2	1	4	3	2	1	4	3	3	4	1	2
40	Portable MP3 Players	4	1	3	2	2	1	3	4	3	4	1	2
41	Portable Tape and Radio/Tape Players	2	1	4	3	2	1	4	3	2	4	1	3
42	Portable Tape Recorders	2	1	4	3	1	2	4	3	3	4	1	2
43	Rack Audio Systems	2	1	4	3	2	1	4	3	2	4	1	3
44	Telephone Answering Devices	4	1	3	2	1	2	3	4	3	4	1	2
45	Total CD Players	2	1	4	3	2	1	4	3	3	4	1	2
46	Turntables	2	1	4	3	1	2	4	3	2	4	1	3
47	VCR Decks	2	1	4	3	2	1	4	3	3	4	1	2
48	VCR Decks with Stereo	2	1	4	3	2	1	4	3	3	4	1	2
49	Videocassette Players	2	1	4	3	2	1	4	3	2	4	1	3
	평균	2.10	1.14	3.92	2.84	1.84	1.47	3.78	2.92	2.61	2.41	1.08	0.73