

■ 論 文 ■

이분산 로짓모형의 추정과 적용

Development and Application of the Heteroscedastic Logit Model

양인석
(국토연구원 연구원)

노정현
(한양대학교 도시대학원 교수)

김강수
(교통개발연구원 책임연구원)

목 차

- | | |
|------------------------------|-----------------------------|
| I. 서론 | 1. 표준 로짓모형 추정결과 |
| II. 통행시간길이에 따른 통행시간 차이의 이분산성 | 2. 통행시간의 이분산성을 고려한 이분산 로짓모형 |
| III. 이분산 로짓모형의 정립 | V. 결론 및 향후 연구과제 |
| IV. 모형의 추정 및 결과 | 참고문헌 |

Key Words : 수단선택, 통행시간, 이분산 로짓모형, 표준 로짓모형, 이분산성

요 약

로짓모형은 선택대안에 대한 확률 계산이 용이하고, 설명변수의 파라메타 추정이 용이하기 때문에 교통 수단 선택모형으로 널리 쓰여지고 있다. 그러나 이러한 로짓모형은 수단선택 효용함수의 오차항 분포가 선택 대안간에 독립적이고, 그 분산이 동일하다는(IID:Independent and Identically Distributed)가정을 내포한다.

본 연구는 수단선택 효용오차의 분산이 수단간에 동일하다는 가정을 완화시키는 이분산 로짓모형 추정에 관한 연구이다. 수단선택 효용오차항의 동분산성을 극복함으로써 보다 현실적인 통행자의 수단선택행태를 반영하는 로짓모형을 추정하는데 본 연구의 목적이 있다.

이를 위해 로짓모형 오차항의 분산과 직접적인 관련이 있는 규모인자(scale factor)를 도입하였다. 이는 대중교통과 승용차의 통행시간차이에 따른 이분산성을 고려하도록 정의되었으며, 이를 통행시간 파라메타 추정에 활용하였다.

본 연구에서 개발된 이분산 로짓모형의 추정 결과, 통행자의 통행시간이 증가하면서 대중교통수단과 승용차의 통행시간차이가 동일하더라도 통행자의 대중교통 수단선택확률이 차이를 보임으로 현실적인 통행자의 수단선택 행태를 반영하는 것으로 판명되었다.

I. 서론

모든 통행자는 통행을 위해 선택행위를 하게 된다. 통행 발생에서부터 목적지, 통행수단, 노선에 대한 선택을 하게 되며, 이러한 통행자의 선택행태를 모형화하는 것은 교통계획이나 교통정책 수립의 효과분석에 필수적이다.

통행자의 선택모형에는 프로빗모형과 로짓모형이 주로 사용되고 있다. 그 중 프로빗모형은 선택 대안이 증가함에 따라 선택대안의 선택확률 계산식이 복잡해져 실제 3~4개 이상의 선택 대안이 존재하는 경우에는 현실적인 적용에 어려움이 존재한다. 반면 로짓모형은 선택대안의 수가 증가할지라도 선택확률식이 단순하고, 또한 모형의 파라메타 추정을 위한 범용 소프트웨어가 존재해 최근까지 통행자의 선택모형으로 가장 널리 쓰이고 있다.

그러나 로짓모형은 효용함수의 오차항이 대안간 독립적이며 그 분산이 동일한 가정(IID:Independent and Identically Distributed) 때문에 비관련대안으로부터의 독립성(IIA:Independence from Irrelevant Alternative property)문제점을 지니고 있으며 이미 많은 문헌에서 제기된 바 있다.

사실 이러한 문제점 중 오차항의 분포가 각 대안에 대하여 독립적으로 분포한다는 가정은 네스티드 로짓모형(Nested Logit Model) 등에 의해 해결하려는 노력이 있어 왔으며, 그와 관련된 연구도 활발히 진행되어 오고 있다. 특히, 상업용 로짓모형추정 소프트웨어에 이러한 네스티드 구조의 로짓모형이 포함됨으로써 더 이상 현실적인 적용상의 문제점은 존재하지 않는게 사실이다.

그러나 로짓모형의 또 다른 가정인 오차항의 분산이 동일하다는 것이다. 이는 추정계수의 부정확성을 가져다 주어 실제 수단분담율 예측에 있어서 왜곡된 결과를 발생시키는 등(Yatchew et al, 1985) 많은 문제점이 있다. 그럼에도 불구하고 이에 대한 연구는 미흡한 실정이다.

본 연구는 서울시민의 출근시 수단선택 자료를 활용하여 실제 통행시간차이의 이분산성을 살펴보고 이러한 이분산성을 고려하는 로짓모형을 추정하는데 연구목적이 있다.

이를 위해 로짓모형의 규모인자(scale factor)를

도입하였다. 이를 통해 통행시간 차이에 따른 이분산성을 고려할 수 있도록 하였다. 그 결과를 표준 로짓모형과 통행시간대별 분할 로짓모형 추정결과와 비교·분석하여 이분산 로짓모형의 우수성을 입증하고자 한다. 단, 본 연구에서는 출근통행목적에 대한 대중교통과 승용차간의 수단분담을 대상으로 한다.

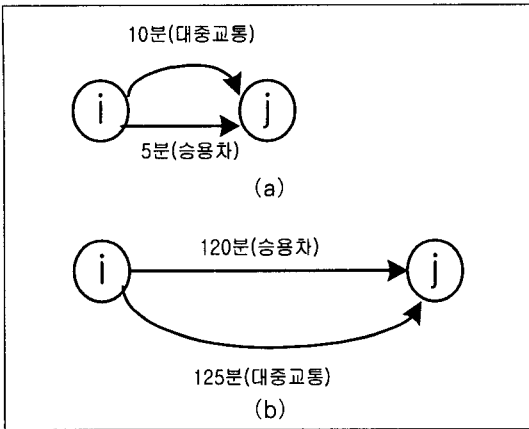
본 연구의 구성은 다음과 같다. 본 연구의 목적과 배경을 기술하는 서론부분에 이어 II단원에서는 통행시간증감에 따른 이분산성 문제를 제기한다. III단원에서는 오차항의 이분산성을 고려하는 규모함수(scale function)를 정의하고, 이 규모함수를 이용하여 이분산 로짓모형을 정립한다. IV단원에서는 기존 표준 로짓모형의 추정과 이분산 로짓모형의 추정 및 결과를 제시하고, 마지막으로 결론에서는 본 연구의 요약과 향후 연구내용을 제시한다.

II. 통행시간길이에 따른 통행시간 차이의 이분산성

통행시간길이에 따른 통행시간차이의 이분산성을 설명하기 위해 <그림 1>과 같은 가정상황을 살펴보자(Yosef Sheffi, 1985).

통행자들이 i 에서 j 로 통행하는데 두 가지 수단(승용차, 대중교통)이 존재하고, 승용차의 통행시간(분)과 대중교통의 통행시간(분)이 작은 (a)의 경우와 비교적 장거리인 (b)의 경우가 존재한다고 가정하자. (a)와 (b)의 경우 대중교통수단과 승용차의 통행시간은 단/장으로 구분할 수 있으나, 모두 두 수단의 통행시간 차이는 5분이다.

만약 두 교통수단에 대한 효용이 통행시간만으로 구성된다면 (a)의 경우 대부분의 통행자는 통행시간이 작은 승용차를 선택하게 된다. 그러나 비교적 장거리인 (b)의 경우 (a)와 동일하게 두 수단의 통행시간 차이가 5분일지라도 상당수(비록 모든 통행자의 반을 넘을 수가 없겠지만)의 통행자가 대중교통을 선택하게 된다. 즉, 단거리통행인 (a)와 장거리통행인 (b)의 경우에 있어 두 수단의 통행시간 차이가 비록 동일한 5분일지라도 통행자가 인식하는 5분 차이가 상이해 (b)의 경우 (a)와는 달리 훨씬 많은 수의 통행자가 대중교통수단을 선택하게 된다. 즉, 통행시간 길이에 따라 통행시간 차이의 이분산성이 존재한다.



〈그림 1〉 두 수단간 통행시간 차이의 이분산성

그러나 이러한 통행시간 차이의 이분산성을 앞에서 언급한 바와 같이 로짓모형은 설명하지 못한다. 왜냐하면 로짓모형은 기본적으로 오차항의 분산이 동일하다는 가정을 내포해 산출되어지는 두 수단의 수단선택확률이 단지 수단간의 통행시간 차이에 의해 결정되기 때문이다. 즉, 다시말해서 (a)와 (b)의 경우 동일한 대중교통과 승용차의 선택확률을 산출한다.

한편, 이러한 개인의 인식차이에 따른 이분산성은 독일의 심리학자 웨버(E.Weber)의 법칙으로도 설명할 수 있다(박영봉, 2001). 웨버는 두 자극 차이에서 차이를 인식하는 차이는 절대량이 아니라 최초의 자극에 대한 비교량이며 처음의 자극이 강하면 강할수록 다른 것으로 되는 두 번째 자극은 더 큰 부가적 강도를 요구한다고 주장하였다. 예를 들어 (a)의 승용차 통행시간 5분, (b)의 승용차 통행시간은 125분이기 때문에 (a)와 (b)의 최초자극이 상이하고, (b)의 경우 (a)와 같은 두 수단간의 통행시간 5분 차이에 따른 동일한 효과를 얻기 위해서는 대중교통과 승용차의 통행시간 차이가 훨씬 커야한다.

그러나 이러한 로짓모형의 동분산성 문제에 대한 연구는 미미하다. 외국의 경우에도 1990년대 중반 들어오면서 이러한 로짓모형의 오차항의 분산이 동일하다는 가정을 극복하려는 연구가 시작되었으나, 이에 대한 연구는 현재까지 미비한 실정이다. 본 연구와 유사한 연구를 살펴보면, Bhat(1995)가 Heteroscedastic

Extreme Value Model(HEV)을 개발하여 특정대안의 특성에 따른 이분산성 로짓모형을 개발하였고, Bhat(1996)은 이분산성을 수용하는 이분산 네스티드 로짓모형을 구축하였다.

Hensher(1998)은 RP데이터와 SP데이터를 이용하여 HEV모형과 네스티드 모형의 직접탄력성과 교차탄력성을 비교하였다.

하지만 본 연구에서는 위의 연구들과는 달리 개인의 효용인식차이에 따른 이분산성을 고려하는 이분산 로짓모형을 개발한다.

III. 이분산 로짓모형의 정립

로짓모형에서 가정한 동분산성 문제점을 극복하기 위해 정립되어지는 이분산 로짓모형은 다음과 같다.

어떤 개인 p 가 선택대안 i 에 대해 가지는 효용 U_{ip} 를 관측효용 V_{ip} 과 비관측효용 ϵ_{ip} 으로 구분한다면, 개인 p 의 선택대안 i 의 효용은 다음과 같이 표현할 수 있다.

$$U_{ip} = V_{ip} + \epsilon_{ip} \quad (1)$$

만약 ϵ_{ip} 의 분포를 다음과 같은 Weibull분포로 가정하면 ϵ_{ip} 의 확률밀도함수는

$$f(\epsilon_{ip}) = \frac{1}{\theta_{ip}} e^{-\epsilon_{ip}/\theta_{ip}} \cdot \exp(-e^{-\epsilon_{ip}/\theta_{ip}}) \quad (2)$$

이며, 분산(Var(ϵ_{ip}))은 다음과 같이 표현되어진다.

$$\text{Var}(\epsilon_{ip}) = \frac{1}{6} \frac{\pi^2}{\theta_{ip}^2} \quad (3)$$

여기서, θ_{ip} 는 규모인자(scale factor)를 나타낸다.

식(3)에서 보느냐와 같이 오차항의 분산(Var(ϵ_{ip}))은 개인 p 와 대안 i 에 따라 상이하게 정의되어진다.¹⁾

이제 어떤 개인 p 가 대안 i 를 선택할 확률(P_{ip})은 대안 i 와 j 의 관측 효용차이로 표현되는 로짓모형이다.

1) 오차항의 이분산성은 선택 대안간의 차이 또는 개인의 차이(설명변수의 인식차이)에 따라 발생할 수도 있으나 오차항의 이분산성에서는 이러한 특성이 혼재 되어 나타나 구별하기가 불가능하다. 본 논문에서의 이분산성은 각 개인에 따른 통행시간의 인식차이에서 비롯된다고 가정한다.

$$P_{ip} = \frac{1}{\sum_{j=1}^I \exp[-(\theta_{ip} V_{ip} - \theta_{jp} V_{jp})]} \\ = \frac{\exp \theta_{ip} V_{ip}}{\sum_{j=1}^I \exp \theta_{jp} V_{jp}} \quad (4)$$

식(4)에서 규모인자 θ_{ip} 는 관측효용 V_{ip} 에 포함되어 있는 파라메타와 분리하여 추정되지 못하므로 일반적으로 $\theta_{ip}=1$ 로 가정하는데(Ben-Akiva, M and Lerman, S. R, 1985), 이는 분산이 모든 개인 p 와 대안 i 에 대해 동일하다는 동분산성 가정이다. 왜냐하면 식(3)에서 보는바와 같이 $\theta_{ip}=1$ 인 경우 p 와 대안 i 에 대해서 $\text{Var}(\varepsilon_{ip}) = \frac{1}{6} \pi^2$ 이기 때문이다.

본 연구에서는 이처럼 오차항의 동분산성을 극복하기 위해 ($\theta_{ip}=1$) 오차항의 분산을 개인마다 각기 다른 값을 갖는 함수형태로 구성하여 통행자의 이분산성을 고려한다.

즉, 규모모수 θ_{ip} 를 S_p 로 표현되어진 규모함수를 이용하여 오차항의 동분산성을 극복한다.

$$\theta_{ip} = f(S_p) \quad (5)$$

따라서, 표준로짓모형에서 오차항의 분산은 $\frac{1}{6} \pi^2$ 이나 ($\theta_{ip}=1$ 일 때), 이분산 로짓모형에서 오차항의 분산은 $\frac{1}{6} \frac{\pi^2}{S_p^2}$ 으로 표현되어져 개인 p 의 특성을 반영할 수가 있다.

이제 오차항의 분산이 $\frac{1}{6} \frac{\pi^2}{S_p^2}$ 인 효용에서 개인 p 가 대안 i 를 선택할 확률은 다음과 같이 표현되어 진다.

$$P_{ip} = \frac{e^{V_{ip} S_p}}{\sum_{j=1}^I e^{V_{jp} S_p}} \quad (6)$$

한편, 본 연구에서 규모함수(S_p)²⁾를 다음과 같이 정의한다.

$$S_p = \exp(\alpha \cdot Z_p), \\ Z_p = \frac{T_{ip} - T_{jp}}{T_{ip} + T_{jp}} \quad (7)$$

여기서,

α : Z_p 의 파라메타

T_{ip} : 통행자 p 의 i 수단 통행시간

T_{jp} : 통행자 p 의 j 수단 통행시간

여기서, Z_p 는 통행자가 인식하는 두 대안간의 통행시간의 차이 변수로서 개인마다 다른 통행시간 차이를 나타내며, ($T_{jp} - T_{ip}$)는 통행자 p 의 두 수단의 절대적인 통행시간의 차이를 나타낸다. 따라서 규모함수는 통행자의 두 수단 통행시간차이에 따른 이분산성을 고려할 수 있다.

최종적으로 이분산 로짓모형에서 통행자 p 의 수단 선택확률 P_{ip} 은 식(8)과 같다.

$$P_{ip} = \frac{1}{\sum_{j=1}^I \exp[-(\exp(Z_p \cdot \alpha) V_{ip} - \exp(Z_p \cdot \alpha) V_{jp})]} \\ = \frac{\exp[\exp(Z_p \cdot \alpha) V_{ip}]}{\sum_{j=1}^I \exp[\exp(Z_p \cdot \alpha) V_{jp}]} \quad (8)$$

식(8)에서 보는 바와 같이 대안 i 에 대한 선택확률은 표준로짓모형과 유사하나 S_p 가 관측효용 V_{ip} 에 곱해짐으로써 비선형 효용함수식이 되어 파라메타 추정에 범용 소프트웨어를 사용하기가 불가능하다. 본 연구에서는 로짓모형의 파라메타 정산을 위해 GAUSS (Aptech Systems, 1994) 프로그램을 이용하여 모형을 추정하였다.

IV. 모형의 추정 및 결과

본 연구의 분석을 위해 1996년 서울시에서 실시한 일기식 패널조사자료 중 출근통행자료를 이용하였다. 분석의 용이함을 위해 통행수단은 승용차와 대중교통수단으로 대별하였다.

본 연구에서 개인의 통행효용함수에 사용되어진 설명변수는 본 연구가 통행시간 차이에 따라 이분산성을 고려한다는 측면에서 통행시간만을 사용하였으며, 수단고유상수로서 승용차에 대한 더미변수를 설정하였다. 통행시간만을 변수로 설정한 주된 이유는 변수의 증가시 이분산성의 출처가 명확하지 못하기 때문

2) 규모함수는 다양하게 정의될 수 있으나 본 연구의 목적이 일반적인 이분산 로짓모형의 개발과 적용에 있어 특정 규모함수 형태에 따른 효과는 향후 연구내용으로 제시한다. 규모함수의 조건과 형태는 김강수(1998)를 참고하기 바란다.

이다. 승용차의 더미변수는 승용차를 기준으로 할 때 승용차와 대중교통수단의 기본적인 효용의 차이를 보기 위해 설정하였다.

1. 표준 로짓모형 추정결과

식(9)는 표준 로짓모형을 추정하기 위한 승용차와 대중교통수단의 관측효용식이며, 추정을 위해 사용된 자료는 총 4,289개이다.

$$V_{car} = \beta_0 + \beta_1 T_{car},$$

$$V_{tran} = \beta_1 T_{tran} \tag{9}$$

여기서,

- V_{car} : 승용차의 효용함수
- V_{tran} : 대중교통수단의 효용함수
- T_{car} : 승용차 통행시간
- T_{tran} : 대중교통수단의 통행시간
- β_0, β_1 : 매개변수(Parameter)

〈표 1〉에서 제시된 모형 추정결과 모형전체의 적합도를 나타내는 ρ^2 는 0.1194로써 비교적 양호하지 못한 결과를 보였지만, 통행시간의 파라메타가 일반적으로 기대되는 것처럼 음(-)의 부호를 가지고 있고, χ^2 테스트 결과 자유도 2에서 유의수준 99%에 해당하는 χ^2 의 임계치인 10.597에 비해 훨씬 큰 것으로 나타나 전체적인 모형 추정결과는 타당한 것으로 나타났다. 또한 파라메타의 통계적 유의성을 나타내는 t값의 경우 충분히 큰 값을 갖고 있으므로 각 파라메타가 갖는 유의성은 매우 큰 것으로 나타났다.

〈표 1〉 표준 로짓모형 추정결과

통행시간 변수	승용차 상수	ρ^2	χ^2	빈도수
-0.0509 (-22.08)	-0.8504 (-19.80)	0.1194	709.99	4,289

1) 통행시간 증감에 따른 이분산성

대중교통과 승용차 두 수단의 통행시간 차이가 동일할지라도 통행시간 길이에 따라 두 수단의 선택 확률이 다를 것이라는 가정 하에 대중교통시간을 세분

〈표 2〉 대중교통 통행시간에 따라 분할된 모형의 추정결과

구분	통행시간 (분)	통행시간 변수	승용차 상수	ρ^2	χ^2	빈도수
전체 모형	전체	-0.0509 (-22.08)	-0.8504 (-19.80)	0.119	709.9	4,289
모형 1	20분 이내	-0.1083 (-8.022)	-0.9281 (-6.851)	0.133	105.7	571
모형 2	21-30	-0.0754 (-10.03)	-0.8063 (-8.615)	0.142	169.2	859
모형 3	31-40	-0.0589 (-8.834)	-0.8282 (-8.238)	0.132	131.1	715
모형 4	41-50	-0.0515 (-8.393)	-0.8928 (-7.314)	0.106	95.9	650
모형 5	51-60	-0.0520 (-10.35)	-1.0486 (-9.158)	0.145	163.1	810
모형 6	61-70	-0.0540 (-6.039)	-1.2036 (-4.926)	0.148	53.4	259
모형 7	71-80	-0.0498 (-5.000)	-1.1755 (-4.382)	0.129	33.8	188
모형 8	81-90	-0.0527 (-5.571)	-1.3434 (-4.791)	0.134	44.3	237

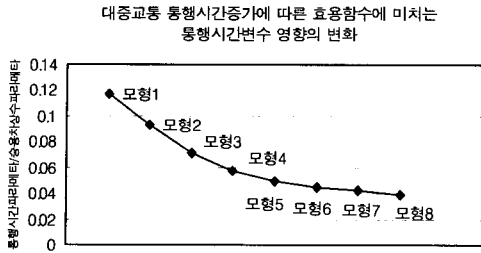
주 : ()은 t값

하여 모형을 각각 추정하였다. 만약 이분산성이 존재하지 않는다면 두 수단간의 통행시간차이가 동일할 때 대중교통의 통행시간 길이에 관계없이 동일한 수단분담율이 제시되어야 한다.

대중교통 통행시간은 20분이내에서 90분까지의 통행이 대부분을 차지하므로 20분에서 90분 사이의 통행시간을 10분 단위로 구분하여 모형을 추정하였다(모형1~모형8). 예를 들어 모형1은 대중교통 통행시간이 20분 이내 통행을 대상으로 구축한 모형이며, 모형2는 대중교통 통행시간이 21-30분 이내의 통행을 대상으로 구축한 모형이다. 통행시간에 따라 분할된 모형의 추정결과는 〈표 2〉와 같다.

〈표 2〉의 추정결과를 살펴보면 대중교통 통행시간이 증가함에 따라 대체로 통행시간 파라메타 크기가(절대값)이 작아진다. 이는 대중교통 통행시간이 증가함에 따라 통행시간이라는 변수가 효용함수에 미치는 영향이 작아져 결국 승용차와 대중교통수단간의 통행시간차이에 따른 수단선택확률에 대한 영향이 감소함을 나타낸다.

이를 좀더 자세하게 살펴보면 다음과 같다. 수단선택 효용은 통행시간과 승용차 상수로 구성되기 때문에



〈그림 2〉 대중교통 통행시간증가에 따른 효용함수에 미치는 통행시간변수 영향의 변화

통행시간만의 효과를 보기 위해 각 모형별로 통행시간 파라메타를 승용차 상수 파라메타로 나누었을 때 비교하였다. 〈그림 2〉는 모형별로 '통행시간 파라메타/승용차상수 파라메타'를 보여주고 있는데 모형1에서 모형8로 갈수록 그 비가 작아져 대중교통시간 증가에 따라 통행시간이라는 변수가 효용함수에 미치는 영향은 작아짐을 확인할 수 있다.

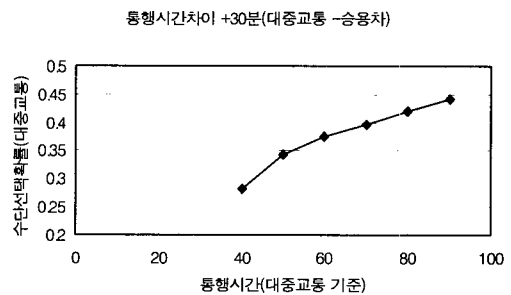
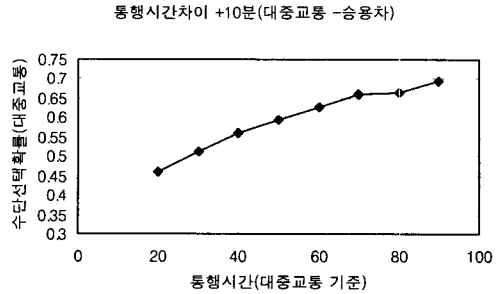
〈표 3〉은 〈표 2〉의 모형추정결과를 이용하여 두 수단의 통행시간 차이 (Δt =대중교통 통행시간 - 승용차 통행시간)에 따른 대중교통 수단선택확률을 나타내고 있다.

〈표 3〉에서 알 수 있듯이 대중교통 통행시간이 커짐에 따라 승용차와의 통행시간의 차는 동일해도 통행자가 인식하는 통행시간차이는 점점 작아져 대중교통수단의 수단분담율은 커지고 있다. 즉, 통행자의 통행시간에 이분산성이 존재한다.

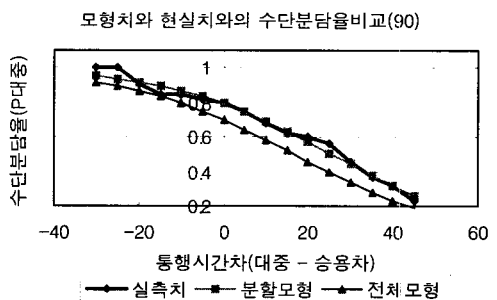
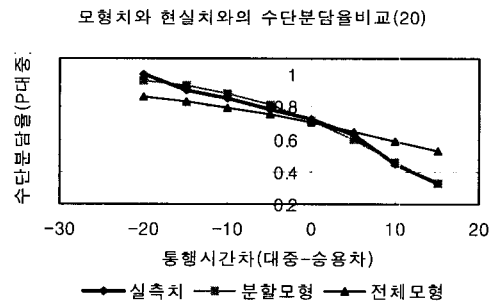
〈표 3〉 대중교통과 승용차의 통행시간차이가 10분과 30분일 때 대중교통 통행시간 변화에 따른 대중교통 수단선택확률변화

구분	대중 교통 통행 시간 (분)	대중교통 수단분담율 ($\Delta t = 10$ 분)	대중교통수단 분담율 ($\Delta t = 30$ 분)
전체모형	전체	0.5845	0.3370
모형1	20	0.4614	.
모형2	30	0.5129	.
모형3	40	0.5595	0.2813
모형4	50	0.5933	0.3424
모형5	60	0.6291	0.3749
모형6	70	0.6599	0.3971
모형7	80	0.6630	0.4207
모형8	90	0.6934	0.4409

한편, 수단분담율측면에서 통행시간에 따라 분할된 모형(모형1~모형8)의 수단선택확률과 관측 수단선택



〈그림 3〉 통행시간 변화에 따른 대중교통 수단선택확률의 변화



〈그림 4〉 통행시간별 모형치(모형1, 모형8, 전체모형)와 현실치의 수단분담율비교(대중교통시간 20분과 90분일 때)

확률을 비교하였다.

〈그림 4〉는 대중교통수단의 통행시간이 20분과 90분일 때 모형1과 모형8의 수단분담율과 실제 관측된 수단분담율을 비교한 결과이다.

〈그림 4〉에서 볼 수 있듯이 분할된 모형이 전체 모형에 비해 실측 수단분담율과 비슷한 곡선을 그리고 있는 것을 알 수 있다. 즉, 기존 표준 로짓모형을 사용함으로 이분산성에 대한 고려가 없을시 수단분담율 예측 측면에서 왜곡을 가져올 수 있음이 판명되었다. 이는 분할모형에서는 대중교통 통행시간대별로 분할하여 추정했기 때문에 분할된 그룹 내에서는 개인별로 인식하는 통행시간의 차이가 크지 않아 전체 로짓모형보다 현실을 보다 잘 반영하기 때문인 것으로 판단된다.

그러나 이러한 통행시간별 분할 모형이 전체모형보다 통행자의 선택행태 현실을 잘 반영한다하더라도 실제 교통계획에서 분할모형을 적용하는 것은 쉽지 않다. 예를 들어 독립변수가 통행시간이외의 변수들이 추가되었을 때는 본 연구에서처럼 통행시간대별로 분할할 기준이 모호해질 수 있으며, 분할기준도 자의적일 수밖에 없다. 또한 모형 추정에 있어서 효율적이지 못한 단점이 내재한다.

2. 통행시간의 이분산성을 고려한 이분산 로짓모형

통행자의 통행시간에 따른 분할모형 구축결과 통행시간의 이분산성이 존재함을 확인하였다. 즉, 대중교통 통행시간이 길어짐에 따라 승용차와 대중교통수단의 대안간의 통행시간차이가 수단선택에 미치는 영향이 감소하여 대중교통 수단분담율이 증가하였다.

본 단원에서는 이러한 이분산성을 고려한 로짓모형을 추정하였다. 단원Ⅲ을 기초로 모형식에 사용된 관측효용함수식은 다음과 같고 앞서 언급한 바와 같이 모형추정에는 GAUSS(Aptech Systems, 1994)를 사용하였다.

$$\begin{aligned}
 V'_{ip} &= (\beta_0 + \beta_1 T_{ip}) \cdot \exp(\alpha \cdot Z_p), \\
 V''_{ip} &= (\beta_1 T_{ip}) \cdot \exp(\alpha \cdot Z_p), \\
 Z_p &= \frac{T_{ip} - T_{jp}}{T_{ip} + T_{jp}} \quad (10)
 \end{aligned}$$

〈표 4〉 모형의 추정결과

구분	이분산 로짓모형			표준 로짓모형		
	파라메타	표준오차	t값	파라메타	표준오차	t값
통행시간	-0.052	0.250 E-02	-20.38	-0.051	0.230 E-02	-22.08
승용차상수	-0.820	0.464 E-01	-17.65	-0.850	0.429 E-01	-19.81
α	-0.142	0.2358	-3.148	·		
$LL(\hat{\beta})$	-2612.138			-2617.910		
$LL(0)$	-2972.908			-2972.908		
ρ^2	0.1213			0.1194		
$\overline{\rho^2}$	0.1203			0.1187		
χ^2	721.539			709.997		

여기서,

V_{ip} : 통행자 p 의 i 의 효용함수

V_{jp} : 통행자 p 의 j 의 효용함수

β_0, β_1 : 매개변수(Parameter)

이분산 로짓모형의 추정결과와 이를 표준 로짓모형과 비교한 결과는 〈표 4〉와 같다.

두 모형의 비교결과 통행시간 및 승용차 터미 값은 두 모형에서 비슷한 값을 나타냈으나, 통행시간 파라메타는 이분산 로짓모형이 약간 크며, 승용차상수는 표준 로짓모형의 값이 약간 더 크다. 두 모형의 파라메타가 크기가 차이나는 이유는 두 모형의 분산이 다르다는 것을 증명한다. 예를 들어 이분산 로짓모형의 경우 효용함수에서 오차항의 이분산을 고려했기 때문에 오차항의 분산이 표준 로짓모형과는 상이하다. t값의 경우는 이분산 로짓모형 값이 표준 로짓모형보다 전체적으로 작게 나타났으나 유의수준 99%수준에서 모두 유의성이 있는 것으로 나타났다.

전체적으로 다소 낮은 우도비를 보이고 있지만 이분산 로짓모형이 좀더 나은 수정 우도비를 보이고 있다.

두 모형 모두 각각 자유도 2와 3에서의 유의수준 99%에 해당되는 χ^2 의 임계치인 10.597과 12.838보다 훨씬 큰 값을 가짐으로 모형 자체적으로는 모두 통계적으로 매우 유의한 모형들인 것으로 판단된다.

이분산 로짓모형과 표준 로짓모형의 동일성을 검증하기 위해 LR검정³⁾을 실시하였다. LR검정값이 11.6으로 자유도 1에서 1% 유의수준 6.6보다 큼으로 두

3) 만약 계산되어진 검정통계량 LR값이 자유도 d 의 χ^2 분포의 임계치 보다 크면 두 모형의 각 파라메타가 동일하다는 가설이 기각되어져 두 모형이 서로 다른 모형이라는 결론을 내릴 수 있다.

모형이 서로 같다는 귀무가설을 기각하고, 이는 서로 다른 모형인 것으로 판정할 수 있다. 즉, 이분산을 고려하는 파라메타의 모형내 삽입은 정당함이 판명되었다.

한편, 통행자의 통행시간에 대한 이분산성을 나타내는 파라메타 α 인 경우 t 값이 유의하게 산출되어 통행자의 통행시간에 대해 이분산성이 존재하다는 것이 입증되었다.

1) 추정선택확률과 관측선택확률 비교

두 모형의 모형전체로서의 적합성을 판단하기 위하여 Δ , δ 지표를 식(11), 식(12)와 같이 설정하여 비교하였다.

$$\Delta = \sum_{p=1}^N \left(\frac{Y_p^a - Y_p^s}{Y_p^a} \right)^2 \tag{11}$$

여기서,

Y_p^s : 추정된 선택확률(simulated value)

Y_p^a : 관측된 선택확률(actual value)

N : 빈도수

$$\delta = \text{MAX} \left[\left(\frac{Y_p^a - Y_p^s}{Y_p^a} \right)^2 \right] \tag{12}$$

여기서,

δ : 실제값과 추정된 값의 최대오차

Δ 값은 추정선택확률이 관측선택확률과 얼마나 근사하게 추적하고 있는가를 판정하는 지표이다. 즉, Δ 값이 0에 가까울 수록 추정된 값이 실제 값을 잘 반영하고 있음을 나타낸다. 한편 δ 값은 실제 관측 선택 확률과 추정 선택확률의 최대오차를 나타낸다. δ 값이 작을수록 모형결과는 관측 선택확률을 정확하게 반영한다.

<표 5> 표준 로짓모형과 이분산 로짓모형의 Δ 값과 δ 의 결과이다.

<표 5> Δ 값과 δ 값의 비교

모형	표준 로짓모형	이분산 로짓모형
Δ	62.96	48.59
δ	14.30	11.18

<표 5>의 결과 실제값과 추정치의 오차들의 합을 나타내는 Δ 값이 이분산 로짓모형에서 더 작게 나타나고 있어 이분산 로짓모형이 실제관측치와의 오차가 더 작다는 것을 의미한다. 또 실제관측치와 추정치의 최대오차를 나타내는 δ 값도 이분산 로짓모형이 더 작게 나타나 이분산 로짓모형에 의해 추정된 값이 표준 로짓모형에 의해 추정된 값보다 현실의 수단 선택확률을 더 잘 반영하고 있는 것으로 분석되었다.

2) 통행시간의 차가 일정할 때 통행시간에 따른 수단 선택확률 비교

이분산 로짓모형이 표준로짓모형이 가정하는 동분산성문제를 극복하는지 살펴보기 위해 승용차와 대중교통의 통행시간이 길어지면서(짧아지면서) 대중교통의 수단선택확률이 변하는 결과를 산출하였다. 즉, 대중교통통행시간과 승용차의 통행시간 차이가 15분 ($\Delta t = +15$: Δt 는 대중교통통행시간-승용차통행시간)으로 일정할 때 대중교통 통행시간변화에 따른 두 수단의 선택확률을 분석하였다.

<표 6>은 두 수단간의 통행시간차이가 15분일 때 통행시간변화에 따른 수단선택확률을 비교한 결과이다.

<표 6>에서 보는 바와 같이 표준 로짓모형은 통행시간의 차이가 일정하면 대중교통 통행시간에 관계없이 동일한 선택확률 값을 나타낸다. 이것은 표준 로짓모형의 경우 오차항의 분산이 동일하게 분포한다고 가정해 수단선택확률이 두 수단의 통행시간 차이로만 결정되기 때문이다.

그러나 <표 6>는 대중교통 통행시간이 커짐에 따라 대중교통 수단선택확률이 이분산 로짓모형에서는

<표 6> $\Delta t = +15$ 일 때모형별 통행시간변화에 따른 수단 선택확률비교

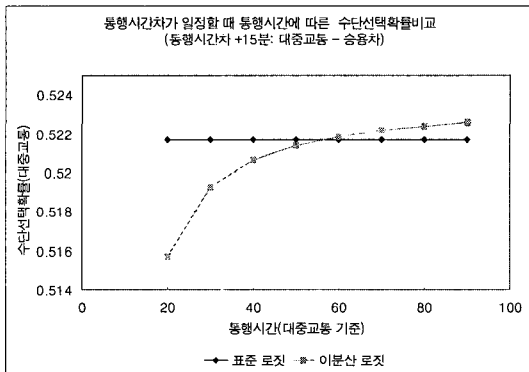
대중교통 시간	승용차 시간	통행시간 차이	표준로짓 모형	이분산 로짓모형
20	5	15	0.5216	0.5156
30	15	15	0.5216	0.5193
40	25	15	0.5216	0.5207
50	35	15	0.5216	0.5214
60	45	15	0.5216	0.5219
70	55	15	0.5216	0.5222
80	65	15	0.5216	0.5224
90	75	15	0.5216	0.5226

주 : $\Delta t = +15$: Δt 는 대중교통 통행시간-승용차통행시간

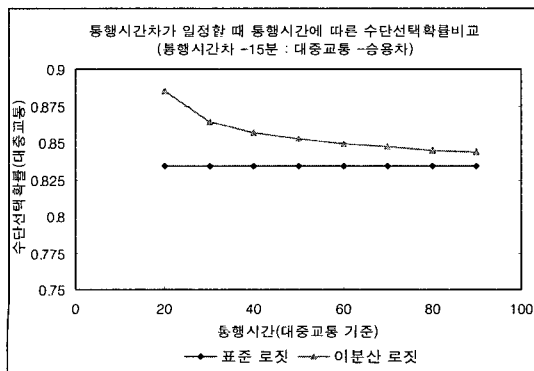
점점 커지고 있음을 보여주고 있다. 즉 대중교통 통행시간이 90분일 때 이분산 로짓모형에서의 대중교통 수단선택확률이 0.5226으로 20분일 때 대중교통 수단선택확률이 0.5156보다 크다. 이는 이분산 로짓모형이 대중교통수단과 승용차 통행시간 차이가 동일할 지라도 대중교통 통행시간의 크기에 따라 통행자가 인식하는 통행시간의 차이가 다른 이분산성을 반영하게 되어 상이한 대중교통 수단선택확률 결과가 산출되기 때문이다.

다시 말해서 대중교통 통행시간이 커짐에 따라 대중교통 통행자가 인식하는 승용차와의 통행시간의 차이가 작아져 $\Delta t = +15$ 가 효용함수에 미치는 영향이 상대적으로 작아지고 있음을 의미한다.

<그림 5>와 <그림 6>은 대중교통수단과 승용차의 통행시간차이가 +15분, -15분일 때 대중교통 수단선택확률 변화를 보여주고 있다.



<그림 5> 통행시간차가 일정할 때 통행시간 변화에 따른 수단선택확률비교(Δt = +15)



<그림 6> 통행시간차가 일정할 때 통행시간 변화에 따른 수단선택확률비교(Δt = -15)

V. 결론 및 향후 연구과제

본 연구에서는 하나의 모형 내에서 오차항의 이분산성을 고려할 수 있는 이분산 로짓모형을 추정하였다.

추정결과 적합도 측면에서 이분산 로짓모형이 기존 로짓모형보다 좀더 나은 수정 우도비를 보였다. 수단분담을 측면에서는 실제값과 추정치의 오차들의 합을 나타내는 Δ 값과 실제관측치와 추정치의 최대오차를 나타내는 δ 값이 이분산 로짓모형에서 더 작게 나타나 이분산 로짓모형에 의해 추정된 값이 표준 로짓모형에 의해 추정된 값보다 현실의 수단 선택확률을 더 잘 반영하고 있는 것으로 분석되었다.

또한, 두 대안간의 관측 효용차이만으로 수단선택확률이 계산되는 표준 로짓모형과는 달리 개발된 이분산 로짓모형은 통행시간이 길어짐에 따라 대안간 통행시간차이가 동일하더라도 수단선택확률이 변하는 모형으로써, 기존 표준로짓 모형보다 통행자의 수단 선택 행태를 보다 잘 반영할 수 있는 모형으로 결론지어 졌다.

향후 연구과제로써 오차항의 이분산성을 나타내는 개인간 특성이나 대안간 특성들 중에서 본 연구에서는 자료의 한계 때문에 통행자가 인식하는 통행시간의 차이 때문에 나타나는 이분산성만을 고려하였으나, 그 이외에도 다른 개인의 사회 경제적 지표를 포함시켜 분석할 필요가 있다. 그리고 이분산 로짓모형을 구축하기 위해 가정했던 규모함수에 대한 다양한 형태와 이에 따른 영향도 향후 필요한 연구과제가 될 것이다.

참고문헌

1. 강동진(1998), "로짓모형에 의한 통행시간가치의 산정 및 특성에 관한 연구", 한양대학교 석사학위논문.
2. 고준호(1999), "선택수단제약과 표본추출방법을 고려한 수단선택 모형에 관한 연구 -서울시 출근 목적 통행을 사례로", 서울대학교 환경대학원 석사학위논문.
3. 김강수(1998), "A Behavioural Approach to Freight Transport Modal Choice", University of Leeds.
4. 노정현(1999), "교통계획 : 통행수요이론과 모형",

남출판.

5. 박영봉(2001), "소비자 행동론, 박영사".
6. 서울시(1997), "서울시 교통센서스 및 데이터 베이스구축-가구통행실태조사".
7. 오재학·박지형(1997), "수도권 여객 통행행태의 조사: 개별행태모형의 정립을 중심으로", 교통개발연구원.
8. Aptech Systems(1994), "GAUSS User's Manuals", Aptech Systems Inc, Kent.
9. Ben-Akiva, M and Lerman, S.R.(1985), "Discrete Choice Analysis: Theory and Application to Travel Demand", M.I.T Press.
10. C.R. Bhat(1995), "A Heteroscedastic Extreme Value Model of Inter city Travel Mode Choice", Transportation Research, 28B, pp.471~483
11. C. R. Bhat(1996), "Covariance Heterogeneity in Nested Logit Model", Transportation Research, 31B, 1996, pp.11~21
12. Hensher et.al(1998), "A Comparison of Elasticities Derived from Multinomial Logit and Heteroscedastic Extreme Value SP-RP Discrete Choice Models", working paper.
13. Langche Zeng(1999), "A Hetero-scedastic GEV Discrete Choice Model".
14. Ortuzar J. de D. & L. G Willumsen (1994), "Modelling Transport", John wiley & Sons.
15. Yatchew, Adonis and Griliches, Zvi(1985), "Specification Error in Probit Models", Review of Economics and Statistics, Vol.67.
16. Yosef Sheffi(1985), "Urban Transportation Networks", Prentice-Hall Inc.

✉ 주 작 성 자 : 양인석

✉ 논문투고일 : 2003. 1. 14

논문심사일 : 2003. 2. 21 (1차)

2003. 6. 30 (2차)

심사판정일 : 2003. 6. 30

✉ 반론접수기한 : 2003. 12. 31

Accidents involving pedestrians at signalized intersections are classified by type of crosswalks by traffic volume, pedestrian volume at crosswalk, intersection geometry and phase in view of pedestrians' safety at 50 intersections in Gwangju. The Multiple regression models are applied to express the pedestrian accident rate. In addition, process deciding what type of crosswalk installs which includes accident rate involved pedestrian is changed into number of accident is represented to reduce number of accidents.

This paper presents what type of crosswalk installs in order to reduce pedestrian involved accidents at new or existing crosswalk.

Development and Application of the Heteroscedastic Logit Model

YANG, In Seok · RHO, Jeong Hyun · KIM, Kang-Soo

Because the Logit model easily calculates probabilities for choice alternatives and estimates parameters for explanatory variables, it is widely used as a traffic mode choice model. However, this model includes an assumption which is independently and identically distributed to the error component distribution of the mode choice utility function.

This paper is a study on the estimation of the Heteroscedastic Logit Model, which mitigates this assumption. The purpose of this paper is to estimate a Logit model that more accurately reflects the mode choice behavior of passengers by resolving the homoscedasticity of the model choice utility error component.

In order to do this, we introduced a scale factor that is directly related to the error component distribution of the model. This scale factor was defined so as to take into account the heteroscedasticity in the difference in travel time between using public transport and driving a car, and was used to estimate the travel time parameter.

The results of the Logit Model estimation developed

in this study show that Heteroscedastic Logit Models can realistically reflect the mode choice behavior of passengers, even if the difference in travel time between public and private transport remains the same as passenger travel time increases, by identifying the difference in mode choice probability of passengers for public transportation.

Warrants of Permissive Left-Turn Signal Systems Based on a Cross Road Volumes

KIM, Dong Nyong · CHOI, Jong Yun

The purpose of this study is to provide the criteria for implementing unprotected left turn at intersections with variation of traffic volume on a cross road approach. Using Transyt-7F model, the delays calculated from permissive and protected left turn signal system were compared by gradually increasing the left turn volume for a certain opposing through volume up to the volume limits to which permissive left turn is more effective. Average stopped delay of the intersection was used as the measure of effectiveness in this study.

The major conclusions are (1) the lighter the traffic gets in a cross road, the more the allowable left turn volume increases. The allowable left turn volume when the ratio of cross traffic to the concerned approach traffic is 0.6 appears about 50% more than the volume when the ratio is 1.0. (2) Comparing to the criteria of the manual of traffic safety facility, the results when the traffic ratio is 0.6 seem to be most similar the criteria of manual and the results when the traffic ratio are 0.8 and 1.0 appears to be lower than the criteria of manual. (3) The possible amount of making a left turn that is inversely proportional to the opposing through traffic, decreases as the number of opposing through lanes increases. The products of volume need to be used as the criteria of permissive left turn with considerable cautions because of its low consistency.