▣ 論 文 ▣

가산자료모형을 기초로 한 통행행태의 한계효과분석

Marginal Effect Analysis of Travel Behavior by Count Data Model

장 태 연
(전북대학교 건축도시공학부 교수)

목 차

Ⅰ. 서론
1. 연구의 배경 및 목적
2. 문헌고찰
Ⅱ. 가산자료모형
1. 음이항모형
2. Zero Altered Model(ZAM)
Ⅲ. 모형추정 및 최적모형선정
1. 변수설정
2. 과산포검정
3. 최적모형선정
Ⅳ. 한계효과분석
Ⅴ. 결론
참고문헌

Key Words : 비음정수, 가산자료모형, 비가정기반통행, 과산포, 한계효과분석

요 약

교통수요예측의 통행발생단계에서 일반적으로 선형회귀모형이 활용되고 있다. 이러한 선형회귀모형은 여러가지 방법론적 한계성과 실용적 지속성을 가지지 못하는 경향을 보인다. 첫째, 종속변수로 이용되는 통행발생의 경우 비음정수(non-negative integer : 0, 1, 2 등)의 이산분포특징을 보이나, 선형회귀모형에서는 종속변수가 연속확률분포 인 정규분포의 특징을 가진 것으로 가정한다. 둘째, 모형이 자료측정에 적용되었을 때 음(-)의 결과를 산정 할 수 있으며, 독립변수의 증감에 따라 결과 값을 너무 높게 혹은 낮게 예측하는 경우가 있다. 셋째, 예측된 값이 정상적인 범위 내에 있을 지라도 예측된 통행수만을 제시 할 뿐, 통행발생빈도에 대한 이산확률분포는 제공하지 않는다. 이같은 한계점을 극복하기 위해 주로 활용되어온 가산자료모형이 포와송모형이다. 그러나, 포와송모형의 경우 자료의 평균과 분산이 동일하다는 가정하에 활용되고 있어 자료상에 과산포가 존재할 경우 오차를 과소평가 할 경향이 높아 모형의 신뢰성에 문제가 발생됨으로 기타 다른 가산자료모형의 적용을 고려해야한다. 연구에서는 과산포검정을 통해 통행발생빈도상에 과산포 존재를 밝혀내고 포와송모형의 부적합함을 제시하였으며 Vuong 검정을 통해 최적의 모형을 선정하였다. 선정된 모형을 대상으로 우도비검정과 Theil 부등계수에 의해 모형의 신뢰도와 정확성을 조사하였다. 최종적으로 가구의 사회경제적 속성의 변화에 따른 통행발생의 변화를 측정하기 위한 민감도 분석을 실시하였다

논문은 2002년 한국학술진흥재단 연구비에 의해 연구됨(KRF-2002-003-B00367).

Ⅰ. 서론

1. 연구의 배경 및 목적

전통적인 4단계 교통수요예측의 통행발생단계에서는 통행발생을 예측하기 위해 일반적으로 다중선형회귀모형이 이용된다(Goulias & Kitamura 1989, Monzon et al. 1989, Goulias et al. 1990). 그러나, 이러한 선형회귀모형은 여러가지 방법론적 한계성과 실용적 지속성을 가지지 못하는 경향을 보인다(Lerman & Gonzales 1980, Ivan & O'Mara 1997).

즉, 기존의 통행발생의 선형회귀모형의 단점으로는 첫째, 종속변수로 이용되는 통행발생의 경우 비음정수(non-negative integer : 0, 1, 2 등)의 이산분포특징을 보이나, 선형회귀모형에서는 종속변수가 연속확률분포 인 정규분포의 특징을 가진 것으로 가정한다. 둘째, 모형이 자료측정에 적용되었을 때 음(-)의 결과를 산정 할 수 있으며, 독립변수의 증감에 따라 결과 값을 너무 높게 혹은 낮게 예측하는 경우가 있다. 셋째, 예측된 값이 정상적인 범위 내에 있을 지라도 예측된 통행수만을 제시 할 뿐, 통행발생빈도에 대한 이산확률분포는 제공하지 않는다.

통행발생과 같은 독립적이고 무작위적인 특징을 보이는 자료의 합리적 분석방법으로 포와송모형이 있다. 그러나, 포와송모형은 선형회귀모형의 단점을 극복할 수 있는 중요한 특성을 가지고 있음에도 불구하고 종속변수의 평균과 분산이 동일하다는 등분산(equidispersion) 가정을 전제로 하고 있다. 가산자료중에는 분산이 평균보다 클 경우가 있는 데 이를 과산포(overdispersion)라 부른다. 포와송모형을 활용하여 과산포가 있는 가산자료를 분석할 경우 계수에 대한 불합리한 유의성 검정이 도출되므로 주의가 요구된다(문진수 외 1999). 비선형특징을 보이는 통행행위에 가산자료모형을 적용하여 선형회귀분석의 단점을 극복하는 것이 연구배경이 된다.

연구의 목적은 가구내에서 발생되는 비가정기반통행에 가산자료모형을 적용하여 통행발생에 영향을 주는 결정요인을 실증적으로 분석하는 것이며, 다음의 경우가 세부적 목적이다. 첫째, 가구내 통행실태자료를 바탕으로 비가정기반통행으로부터 과산포를 측정하여 포와송모형을 적용하는데 있어 불합리성과 적절한 대안 모형의 적용가능성을 이끌어 낸다.

둘째, 가족구성원은 비가정기반통행을 발생시키지 않았을 때 영(0)으로 응답한다. 이는 설문조사 당일 비가정기반통행을 발생시키지 않았을 가능성이 있거나, 단순히 목적지와 집만을 통행하는 가족구성원으로서 비가정기반통행을 전혀 발생시키지 않는 경우로서 포와송모형은 이론적 영(zero counts)확률에 대해 과장함으로서 모형의 신뢰성을 떨어뜨린다. 이러한 상황을 고려할 수 있는 음이항모형을 비롯하여 수정가산모형 등 다양한 모형을 구축한다.

셋째, 다양한 모형중에서 자료의 속성을 가장 효과적으로 반영하는 모형을 Vuong 검정을 통해 선정하여 한계효과를 분석한다. 가구의 사회경제적속성이나 정책의 변화에 따라, 비가정기반통행발생은 변화될 수 있는데 이에 대한 잠재성을 검토한다.

2. 문헌고찰

가산자료에 대한 이론적 매력과 효과적 분석법으로서 포와송모형이 최근에 활용되고 있다. 이러한 가산자료모형은 교통분야에서도 특히 교통사고분석에서 매우 활용도가 높다(Poch & Mannering 1996, Shanker et al. 1998). 포와송모형은 과산포로 인한 모형결과의 신뢰성에 문제가 있어 수정가산모형이 적용되었는데, 이론적 측면에서 과산포현상과 사건 무발생에 대한 가중치 제공의 문제해결에 도움이 되는 것으로 나타났다(Jansakul & Hinde 2002, Xie & Goh 2001, Freund et. al 1999). 과산포현상을 보이는 가산자료에서 Gurmu(1998)은 수정가산모형의 중요성을 보여주고 있으며, McCarty(2001)는 음주운전 발생빈도를 분석하였는데, 다양한 수정가산모형 중에서 hurdle모형이 자료의 반영정도가 높음을 제시하고 있다.

Ma & Goulias(1999)가 포와송모형을 이용하여 개인의 활동발생에 대한 다양한 결과를 이끌어 내었으며 포와송모형은 개인이 하루동안 발생시키는 활동빈도가 매우 높을 경우 모형측정과 예측에서 좋은 결과를 주고 있지 못함을 제시하고 있다. 그러나, 연구에서는 가산자료내 존재하는 사건 무발생에 대한 상황을 설명하지 못하고 있다. Bhat, Carini & Misra (1999)는 가정에서 발생되는 활동중에서 쇼핑활동에 대한 가산자료분석을 실시하였는데, 고속도로 통행비용, 토지이용변화, 접근성을 기준으로 하여 활동빈도

의 관계를 이끌어 내었다. 일반적으로 가산자료인 종속변수는 이산사건발생을 의미하나 Misra(1999)는 비직장인들이 발생시키는 활동시간(duration)을 대상으로 시간의 일정길이를 이산화 한 가산자료모형을 적용하여 활동시간과 사회경제적 속성과의 인과관계를 분석하였다. 교통계획분야의 기존 연구에서는 수정가산모형의 적용이 미비하며, 통행발생에 있어서 과산포현상이나 활동 무발생의 정도가 높게 나올 수 있어 이러한 현상을 반영할 수 있는 모형구축이 요구된다.

Ⅱ. 가산자료모형

1. 음이항모형

가산자료분석이나 카테고리 자료를 분석하는 데 있어 주로 사용되는 방법이 포와송모형이며, 확률 밀도식과 모형은 다음과 같다.

$$P(Y=y_i)=\frac{e^{\lambda_i}\lambda_i^{y_i}}{y_i!} \quad (1)$$

$$\ln \lambda_i=\beta' X_i \quad (2)$$

여기서,
$P(Y=y_i)$: 사건 y_i 만큼 발생될 확률
λ_i : 시간당 발생될 평균 및 분산
β' : 추정계수벡터
X_i : 측정변수벡터
$y_i=0, 1, 2, 3, \cdots$

식(2)는 오차항을 포함하고 있지 않기 때문에 동질적(homogeneous)이고 평균과 분산이 동일한 자료에 단지 적용될 수 있음을 의미한다(Greene 1994). 평균 파라메타는 식(3)으로 산정 될 수 있으며 분산은 식(4)와 같다.

$$E[y_i|\beta, X_i]=\mu_i=e^{\beta X_i} \quad (3)$$

$$Var[y_i|\beta, X_i]=e^{\beta X_i} \quad (4)$$

Gradient와 Hessian은 최우추정법에 의해 측정될 수 있는데 우도함수는 식(5)와 같다.

$$L(\beta)=\prod_{i=1}\frac{\exp[-\exp(\beta X_i')][\exp(\beta X_i')]^{y_i}}{y_i!} \quad (5)$$

자료의 기대치와 분산이 동일하지 않거나 과산포현상이 나타날 경우, 계수에 대한 통계학적 중요성을 잘못 이끌어 내거나 과도하게 낙관적인 결과를 산출해 낸다(Karlaftis & Tarko 1998). 통행발생과 같은 가산자료에서는 과산포현상이 빈번히 발생하고 있어 자료의 이질성(heterogeneity)과 과산포현상을 반영하기 위해서는 식(2)에 오차항이 첨부된다.

$$\ln \lambda_i=\beta' X_i+\varepsilon_i \quad (6)$$

오차항의 확률분포에 따라 음이항모형이 결정되며, 공식(6)을 기초로 한 음이항 분포의 확률 밀도식은 식(7)과 같다. 과산포 파라메터 α가 0이 되면 자료는 등산포 특성을 보여 포와송모형과 같아진다.

$$P(Y=y_i|\varepsilon)=\frac{e^{\lambda_i}\lambda_i^{y_i}e^{\varepsilon}}{y_i!} \quad (7)$$

$\alpha \geq 0, \quad y=0, 1, 2, \cdots$

2. Zero Altered Model(ZAM)

단위시간당 사건이 발생되지 않았을 경우 일반적으로 2가지 경우를 고려해 볼 수 있는데 사건을 발생시켜오다가 조사 당일 사건을 발생시키지 않았을 경우와 처음부터 사건을 전혀 발생시켜오지 않은 경우이다. 즉, 연구의 목적에서 언급한 것처럼 비가정기반통행을 발생시키지 않았을 때(zero counts)의 경우는 조사 당일 발생시키지 않았을 가능성이 있거나, 가족구성원들이 단일 목적지와 집만을 통행하는 가구로서 비가정기반통행을 전혀 발생시키지 않을 경우이다. 이러한 경우 과도한 통행미발생은 자료상에 관측되지 않은 이질성을 내포함으로서 과산포현상을 보이게 되는데, 이를 고려할 수 있는 수정가산모형에는 zero inflated poission model(ZIP)과 zero inflated negative binomial model(ZINB) 등이 있다. 다음은 ZAM모형의 일반적인 형태를 나타내고 있다.

$$P(y_i=0|X_i)=q_i+(1-q_i)\left(\frac{\alpha^{-1}}{\alpha^{-1}+\lambda_i}\right)^{\alpha^{-1}} \quad (8)$$

$$P(y_i>0|X_i)=(1-q_i)\frac{\Gamma(\alpha^{-1}+y_i)}{\Gamma(\alpha^{-1})\Gamma(y_i+1)}\left(\frac{1}{1+\alpha\lambda_i}\right)^{\alpha^{-1}}\left(1-\frac{1}{1+\alpha\lambda_i}\right)^{y_i} \quad (9)$$

$y_i=1,\ 2,\ 3,\cdots$

q_i는 사건(i)을 발생시키지 않을 확률을 의미하며, 효용함수 $\tau\beta' X_i$를 가진 logistic 누적분포를 보이며 τ는 분리계수(estimated splitting parameter)이다. 일반적으로 q_i는 이항로짓모형이나 프로빗모형에 의해 측정된다.

$$q_i \sim Logistics(\tau\beta' X_i) \quad (10)$$

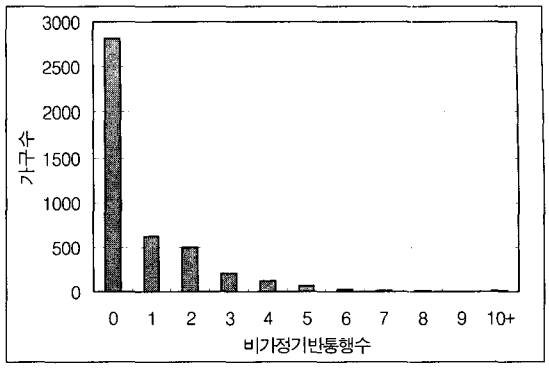
우도함수는 다음과 같으며 $I_{(.)}$는 사건발생의 지시함수로서 0 혹은 1의 값을 가진다.

$$L(\alpha, q, y)=\sum_{i=1}^{n} I_{(y_i=0)} \ln[P(y_i=0|X_i)] + \sum_{i=1}^{n} I_{(y_i>0)} \ln[P(y_i>0|X_i)] \quad (11)$$

Ⅲ. 모형추정 및 최적모형선정

1. 변수설정

가산자료모형을 구축하기 위해 활용된 자료는 2000년도에 4,416가구를 대상으로 조사된 전주시 통행실태자료이다. 일반적으로 통행발생분석은 교통존(traffic zone)단위로 행해지나 본 연구에서는 가구단위의 통행발생분석을 시도하였다. 가능한 설명변수를 포함하여 상관분석 후 통계적으로 유의성이 매우 낮은 변수는 제외하고 최종 변수를 선정하였다. 가장의 직업은 회사원, 공무원, 전문직 등을 white color로, 1차산업 종사자와 일용직, 생산, 자영업 종사자는 blue color로 설정하였다. 〈그림 1〉에서 대부분의 가구가 비가정기반통행을 발생시키지 않아 선형모형이나 포와송모형을 활용할 경우 과산포 발생과 통행 무발생이 과장 될 가능성이 있다.



〈그림 1〉 비가정기반 통행 발생에 따른 가구수

- 가 족 수 : 가구의 총가족수
- 취업자수 : 가구의 총 직장인수
- 소 득 : 1=100만원이하, 2=100~200만원이하, 3=200만원이상
- 주 택 : 1=단독주택, 0=기타
- 주택규모 : 가구의 주택 평수
- 차량대수 : 가구가 소유하고 있는 자동차수
- 나 이 : 가구내 가장의 연령
- 가장직업 : 1=white color, 2=blue color

2. 과산포검정

Cameron & Trivedi(1990)은 과산포검정법으로 회귀분석검정, conditional moment test, Lagrange multiplier test 등을 제시하고 있는데, 회귀분석검정을 활용하였다. μ_i는 포아송모형에서의 평균이며, $g(\mu_i)$는 μ_i의 함수이다. 대립가설양측을 μ_i로 나누면 식(12)의 과산포율을 나타낸다.

$$H_0 : Var[y_i]=\mu_i,$$
$$H_1 : Var[y_i]=\mu_i+\alpha g(\mu_i)$$
$$Var[y_i]/\mu_i=1+(\alpha g(\mu_i)/\mu_i) \quad (12)$$

만일, 자료에 과산포가 존재하면 과산포율이 1보다 크거나 혹은 작게되며 이때 α값을 가지게 된다. 만약 값이 0이면 포와송모형이, 양수 및 음수이면 음이항모형, 이항모형 및 기타 모형이 적합하다. LIMDEP에서는 μ_i를 기초로 하여 Z_i를 종속변수로 W_i를 독립변수로 한 회귀분석를 실시한다(Greene 1998).

$$Z_i=\frac{(y_i-\mu_i)^2-y_i}{\mu_i\sqrt{2}} \quad (13)$$

〈표 1〉 비가정기반통행 과산포검정

구분	α	t값
μ_i	1.34931	14.367
μ_i^2	1.50811	14.258

$$W_i = \frac{g(\mu_i)}{\mu_i\sqrt{2}} \tag{14}$$

회귀분석은 $g(\mu_i)=\mu_i$, $g(\mu_i)=\mu_i^2$의 2가지 가능성을 놓고 실시되며 각각에 대해 α값과 유의성 검증 위한 t값을 제시한다. 〈표 1〉의 검정결과에서 비가정기반통행에 대한 μ_i의 α는 1.34931, t값 14.367이고, μ_i^2의 α는 1.50811, t값 14.258로서 평균과 분산이 같다는 귀무가설은 기각된다. 결과적으로 α값이 모두 양의 값을 가지고 있어 포와송모형은 비가정기반통행 분석을 위해 적합하지 않다.

3. 최적모형선정

1) Vuong 검정

과산포 검정결과 포와송모형은 비가정기반통행 특성을 반영하는데 적합지 않음이 증명되었다. 〈표 2〉는 포와송모형을 제외한 모형 추정결과로서 최적모형의 선정을 위하여 Vuong 통계치가 활용되는데 만약, $F_1(y_i|x_i)$ 과 $F_2(y_i|x_i)$가 서로 대립관계에 있는 예측된 확률분포라면 Vuong 통계치는 다음과 같이 산정된다.

$$m_i = \log\left[\frac{F_1(y_i|x_i)}{F_2(y_i|x_i)}\right] \tag{15}$$

$$V = \frac{\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} m_i\right]}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_i-\overline{m})^2}} \tag{16}$$

여기서, V는 Vuong통계치, $\overline{m}$은 평균, n은 표본수를 의미한다. 95% 통계검증에서 V의 값이 1.96보다 크면 모형 $F_1(y_i|x_i)$이 자료를 잘 반영하고 있으며, -1.96보다 작게 되면 $F_2(y_i|x_i)$이 자료를 더 잘 반영하고 있다. V값이 -1.96과 1.96사이에 있게 되면 두 모형간의 차이가 없어 반영정도를 판단할 수 없다. Vuong 통계치를 산출한 결과, 자료반영정도가 ZINB 〉 ZIP 〉 음이항모형의 순서로 나타나 ZINB이 최적모형으로 선정되었다(〈표 3〉 참조).

〈표 3〉 최적모형선정을 위한 모형비교

구분	Vuong 통계치
ZINB vs 음이항모형	12.1531
음이항모형 vs ZIP	-3.5243
ZIP vs ZINB	-5.5845

〈표 2〉 모형의 추정

구분	음이항모형	ZIP	ZINB
상수	-2.6633(0.0000)	-0.07022(0.7077)	-1.24711(0.0000)
가족수	-0.0293(0.4609)	-0.08584(0.0002)	-0.01278(0.5907)
취업자수	0.2905(0.0000)	0.17766(0.0000)	0.18242(0.0000)
소득	0.2873(0.0000)	0.14610(0.0000)	0.19823(0.0000)
주택	0.1953(0.1197)	0.27491(0.0000)	0.12267(0.1084)
주택규모	0.0044(0.0446)	0.00218(0.0313)	0.00305(0.0153)
차량대수	0.4389(0.0000)	0.23984(0.0000)	0.26928(0.0000)
나이	0.0274(0.0000)	0.00430(0.1956)	0.01664(0.0000)
가장직업	-0.1416(0.0209)	-0.13728(0.0001)	-0.10360(0.0046)
n	4,416	4,416	4,416
$L(\beta)$	-5,058.346	-4,702.833	-4,452.127
L(0)	-5,685.678	-5,652.413	-5,614.283
ρ^2	0.110	0.168	0.207

주) ( ) : p값

2) 모형의 검정

ZINB의 추정된 설명변수 검정은 t값에 의해 이루어지나, 모형전체는 F검정, Lagrange multiplier검정, 우도비검정(*Likelihood ratio test*) 등으로 시행 되며, 본 연구에서는 우도비검정을 실시하였다(Ben-Akiva & Lerman, 1985).

H_0 : 상수를 포함한 모든 변수가 0임

H_1 : 변수들 중 적어도 하나이상은 0이 아님

LR 통계값은 $-2[L(0)-L(\beta)]$로 산정 되는데 $L(0)$는 모든 파라메터가 영(0)일 때의 로그우드함수를 $L(\beta)$는 제약되지 않은 상태에서의 로그우드함수의 최대값을 의미한다. LR통계량은 1개의 자유도를 가지는 χ^2 분포를 나타내며, 변수의 전체적인 유의수준을 검증하기 위해 χ^2 값이 이용된다. ZINB의 LR통계량과 자유도 9의 χ^2 통계량을 비교하여 검정을 실시한다. ZINB의 LR통계량은 2,324.31로 산정 되었으며, 자유도9의 유의수준 0.01에서 χ^2 값은 21.67, 유의수준 0.005에서 χ^2 값은 23.59이므로, 상수를 포함한 모든 변수가 0이라는 귀무가설은 기각되어 추정모형은 의미를 갖는다.

또한, 선정된 모형에 의해 계산된 추정값이 실제값을 얼마나 잘 설명하고 있는 가를 분석하였는데, Theil의 부등계수(U : Theil's inequality coefficient)가 이용된다.

$$U=\frac{\sqrt{\frac{1}{N}\sum_{n=1}^{N}(Y_n^s-Y_n^a)^2}}{\sqrt{\frac{1}{N}\sum_{n=1}^{N}(Y_n^s)^2}+\sqrt{\frac{1}{N}\sum_{n=1}^{N}(Y_n^a)^2}} \quad (17)$$

여기서, N은 표본수, Y_n^s는 예측값, Y_n^a는 실제값이며, U값은 회귀분석의 R^2와 비슷한 기능을 갖는다. 만약 모든 표본 n에 대해 $Y_n^s=Y_n^a$ 이면 U값이 0이 되는 데 예측값과 실제값이 완전히 일치하는 경우로서 완벽한 모형을 의미하며, 1이면 매우 나쁜 경우이다. 여기에서 분모부분인 평균제곱예측오차인 식(18)은 예측오차 원인에 대한 효과적인 분석을 위해 3가지방법, U^M, U^S, U^C로 분리될 수 있다(Pindyck & Rubinfeld, 1998).

$$\frac{1}{N}\sum(Y_n^s-Y_n^a)^2=(Y_{평균}^s-Y_{평균}^a)^2+(\sigma_s-\sigma_a)+2(1-\rho)\sigma_s\sigma_a \quad (18)$$

$$U^M=\frac{(Y_{평균}^s-Y_{평균}^a)^2}{\frac{1}{N}\sum(Y_n^s-Y_n^a)^2} \quad (19)$$

$$U^S=\frac{(\sigma_s-\sigma_a)^2}{\frac{1}{N}\sum(Y_n^s-Y_n^a)^2} \quad (20)$$

$$U^C=\frac{2(1-\rho)\sigma_s\sigma_a}{\frac{1}{N}\sum(Y_n^s-Y_n^a)^2} \quad (21)$$

$Y_{평균}^s$, $Y_{평균}^a$, σ_s, σ_a는 Y_n^s, Y_n^a의 평균과 표준편차를 의미하며, ρ는 상관계수이다. $U^M+U^S+U^C=1$이고 U^M, U^S, U^C는 U의 편차, 분산, 공분산 비율을 나타낸다. U^M과 U^S의 값이 상당히 클 경우 문제가 발생하여 예측모형의 재고가 필요한데, 가장 이상적인 결과는 $U^M=U^S=0$, $U^C=1$인 경우이다. 본 연구에서 U 값이 0.3403이며, U^M, U^S, U^C의 값은 1.9×10^{-6}, 0.26578, 0.73421로 모형을 재고할 필요성은 낮게 나타났다.

3) 추정결과의 분석

〈표 2〉에서 제시된 ZINB모형 결과를 기준으로 설명변수와 비가정기반통행발생의 인과관계를 분석하면, 첫째, 가족수는 통계적 유의성을 보이고 있지 않으며, 계수값 또한 매우 작아 비가정기반통행발생에 큰 영향을 주지 않는 것으로 나타났다. 둘째, 반면에 가구내 취업자수의 증가는 비가정기반통행의 증가를 가져오는 것으로 나타났으며 통계적으로도 유의성을 보이고 있다. 일반적으로 비가정기반통행은 직장 및 학교활동 등의 의무적활동을 제외한, 가구와 관련된 활동 및 자유재량활동을 위한 통행이 대부분일 것임으로 취업자수의 증가는 가족구성원 간에 이러한 자유재량활동에 대한 분담 감소를 의미하며 결과적으로 비가정기반통행발생빈도를 높이는 결과가 된다. 셋째, 소득은 통계학적으로 유의성을 보여주고 있다. 소득이 많은 가정일수록 사회적 활동이 왕성하고 이에 따른 비가정기반통행의 증가로 나타날 수 있음을 의미한다. 넷째, 단독주택의 경우가 비가정기반통행의 발생

확률이 높으나 통계적으로 유의하지 않다. 다섯째, 주택의 규모는 통계적으로 유의성을 보이고 있으나 계수값이 매우 적어 미미한 정도의 비가정기반통행발생에 양(+)을 작용하고 있다. 여섯째, 가정내 보유 차량수가 증가할수록 비가정기반통행의 증가를 보여주고 있는데, 차량은 가족구성원들로 하여금 보다더 활동을 행하게 하는 수단으로서 통행을 발생시킬 가능성을 높게 한다. 일곱째, 가장의 나이는 비가정기반통행발생빈도에 큰 영향을 주고 있지 않으나, 통계적으로 유의성은 높게 나타났다. 여덟째, 가장의 직업에 대한 계수가 음(-)의 값을 보여주고 있어 blue color 보다는 white color일수록 비가정기반통행발생빈도가 높음을 의미하며, 직업의 종류는 가구의 소득과 밀접한 관계가 있음을 가정할 때, 소득과 비슷한 경향을 보이고 있다.

Ⅳ. 한계효과분석

Vuong통계검정을 통해 선택된 모형을 대상으로 하여 종속변수에 대한 설명변수의 한계효과를 분석한다. 한계효과는 가구의 사회경제적속성 변화에 따라 비가정기반통행발생에 있어 변화를 의미한다. 설명변수의 변화에 따른 각각의 효과에 대한 비율변화를 제시함으로서 가구의 환경변화에 따른 비가정기반통행발생의 민감도를 분석한다. 모형에서 설명변수의 계수값을 기초로 한계효과를 분석할 수 있는데 기대값은 $\exp(\beta x_i)$이기 때문에 x_i에서 변화는 종속변수의 기댓값 변화를 의미한다. 즉, 설명변수 x_i에서 δ 만큼 변화한다면 $\exp(\beta\delta)$의 인자값이 산출되며 $100(\exp(\beta_i\delta)-1)$만큼 백분율(%)변화를 가져온다.

〈표 4〉는 한계효과분석에 대한 결과를 보여주고 있는데, 가구내 가족구성원의 증가는 비가정기반통행발생 가능성을 1.27%의 감소시킬 가능성이 있으며, 주택유형의 경우 13.05%의 증가효과를 가져올 것으로 나타났으나 통계적으로는 유의하지 않은 것으로 산정 되었다. 또한, 주택의 규모, 가장의 나이에 있어 변화는 통행발생에 영향력이 매우 낮음을 알 수 있다. 취업자수, 소득, 차량수 등은 비가정기반통행발생에 있어 각각 20.01%, 21.92%, 30.90%의 높은 증가효과를 나타내는 변수로 나타났으며, 가장의 직업은 -9.84%의 변화률을 반영하고 있데 이는 bluecolor일 경우 white color에 비해 9.84%의 비가정기반통행 발생확률이 낮음을 의미한다.

〈표 4〉 한계효과분석

구분	인자값 〔exp(βδ)〕	변화율(%) 〔100(exp(βiδ)−1)〕
가족수	0.9873	-1.27*
취업자수	1.2001	20.01
소득	1.2192	21.92
주택	1.1305	13.05*
주택규모	1.0030	0.30
차량대수	1.3090	30.90
나이	1.0167	1.67
가장직업	0.9016	-9.84

* : p=0.05 수준에서 유의하지 않음

Ⅴ. 결론

본 논문은 가구단위의 통행발생과 설명변수를 토대로 하여 가산자료의 특징을 보이는 통행발생빈도분석에 일반적으로 활용되는 선형회귀모형의 단점을 보완하기 위한 적절한 가산자료모형을 제시하였다. 가산자료특성을 보이는 통행발생을 예측하기 위해서는 실제 발생빈도가 영(0)에 집중되어 있는지 살펴보아야 되며, 발생빈도에 대한 예측치는 반드시 비음정수이어야 된다. 연구는 가산자료의 효율적인 분석을 위한 접근방법으로 통행행태 결과로 나타나는 가구내 비가정기반통행에 적용하였다. 등분산을 기본적인 가정으로 하여 설정되는 포와송모형의 단점 인 과산포 현상을 포괄적으로 분석할 수 있는 음이항모형 및 수정가산모형을 적용함으로서 향상된 변수측정과 함께 분산오차를 감소시키는 결과를 보여주었다. 가산자료 분석모형은 다양하게 존재하지만 자료에 따라 설정되는 모형도 다양할 것으로 예상되어 Vuong검정을 통해 최적모형을 선정하고, Theil의 부등계수를 활용하여 최적모형의 추정값에 대한 정확도를 측정하였다.

통행행태분야에 적합한 가산자료모형의 선정과정을 제시하여 진전된 통행수요예측의 가능성을 시도하였다. 민감도 분석을 통해 사회경제적 변수의 활용도에 따른 통행발생의 변화정도를 감지하였다. 이는 교통수요관리 등의 단기교통정책 도입시 통행행태의 변화와 정책의 성공여부를 판단하는데 도움이 될 것이다.

참고문헌

1. 문진수 · 김순관 · 임강원(1999), "가산자료모형을

이용한 버스이용횟수추정에 관한 연구", 대한교통학회지, 제17권 제5호, 대한교통학회, pp.123~135.

2. Bhat, C., and Carini, J. and Misra, R. (1998), "On Modeling the Generation and Organization of Household Activity Stops", Presented in the 1998 TRB, Washington, DC.
3. Freund, D. A., Kniesner, T. J. and LoSasso, A. T(1999), "Dealing with the Common Econometric Problems of Count Data with Excess Zeros, Endogenous Treatment Effects, and Attrition Bias", Economics Letters 62. pp.7~12.
4. Goulias, K. G. Pendyala, R. and Kitamura, R.(1990), "A Practical Method for the Estimation of Trip Generation and Trip Chaining", Transportation Research Record 1285. pp.47~56.
5. Goulias, K. G. and Kitamura, R.(1989), "Recursive Model System for Trip Generation and Trip Chaining", Transportation Research Record 1236. pp.59~66.
6. Greene, W.(1998) LIMDEP Version 7.0 User's Manual Revised Edition. Econometric Software, Inc.
7. Greene, W.(1994), "Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models", Working Paper Number EC-94-10. New York University.
8. Gurmu, S.(1998), "Generalized Hurdle Count Data Regression Models", Economics Letters 58. pp.263~268.
9. Ivan, J. N. and O' Mara, P. J.(1997), "Prediction of Traffic Accident Rates Using Poission Regression", Presented in the 1997 TRB, Washington, DC.
10. Jansakul, N. and Hinde, J. P.(2002), "Score Tests for Zero-inflated Poisson Models", Computational Statistics & Data Analysis. Elsevier Science. pp.1~22.
11. Lerman, S. R. and Gonzales, S. L. (1980), "Poisson Regression Analysis Under an Alternative Sampling Strategies", Transportation Science Vol. 14(4). pp.346~364.
12. Ma, J. and Goulias, K. G.(1999), "Application of Poisson Regression Models to Activity Frequency Analysis and Prediction", Transportation Research Record 1676. pp. 86~94.
13. McCarthy, P.(2001), "An Analysis of Drinking and Driving Behaviors in US", 9th WCTR.
14. Misra, R.(1999), "A Time Quantum Hurdle Model for Activity Type and Duration of Non-workers", Transportation Research Record 1676.
15. Monzon, J., Goulias, K. and Kitamura, R. (1989), "Trip Generation Models for Infrequent Trips", Transportation Research Record 1220 : 40-46.
16. Pindyck, R. and Rubinfeld, D.(1998), Econometrics Models and Economic Forecasts, 4th Edition, McGraw-Hill.
17. Poch, M. and Mannering, F.(1996), "Negative Binomial Analysis of Intersection Accident Frequencies", Journal of Transportation Engineering Vol.122, No.2. pp.105~113.
18. Shankar, V. Milton, J. and Mannering, F. (1998), "Modeling Accident Frequencies as Zero-altered Probability Process", Presented in the 1998 TRB, Washington DC.
19. Xie, M, He, B, and Goh, T. N.(2001), "Zero-inflated Poission Model in Statistical Process Control", Computational Statistics & Data Analysis 38. pp.191~201.

♧ 주 작 성 자 : 장태연
♧ 논문투고일 : 2003. 2. 28
논문심사일 : 2003. 4. 9 (1차)
2003. 5. 7 (2차)
심사판정일 : 2003. 5. 7
♧ 반론접수기한 : 2003. 10. 31

A Study on The Demand Management for Determination of Freeway Toll System

KWON, Yong Seok

Statistics shows that 80% of freeway users travel less than 50km, and only 8% of them do over 100km. It means that the freeway is used for commuting. As a result, the freeway is not used efficiently and social cost is high. The current toll system aims for the efficient usage of the freeway by restraining short-distance trips and inducing long-distance ones. Thus, policies such as minimum toll system and discount for long-distance trips have been carried out. However, these two policies take no account of demand management between the freeway and alternative roads. They merely consider the compensation for charge on users, so the assessment of discount rate and distinction of distance are not reasonable. Consequently, not only the effective demand management but also the evaluation of those policies has been difficult to be achieved.

The objective of this paper is to analyze those problematic policies and to establish a reasonable and improved toll system. In addition, this study presents the methodology to minimize social cost, which can be achieved by reducing short trips and encouraging long ones on the freeway. A new methodology of freeway toll system is applied and the results are presented.

Marginal Effect Analysis of Travel Behavior by Count Data Model

JANG, Tae Youn

In general, the linear regression model has been used to estimate trip generation in the travel demand forecasting procedure. However, the model suffers from several methodological limitations. First, trips as a dependent variable with non-negative integer show discrete distribution but the model assumes that the dependent variable is continuously distributed between $-\infty$ and $+\infty$. Second, the model may produce negative estimates. Third, even if estimated trips are within the valid range, the model offers only forecasted trips without discrete probability distribution of them. To overcome these limitations, a poisson model with a assumption of equidispersion has frequently been used to analyze count data such as trip frequencies. However, if the variance of data is greater than the mean, the poisson model tends to underestimate errors, resulting in unreliable estimates. Using overdispersion test, this study proved that the poisson model is not appropriate and by using Vuong test, zero inflated negative binomial model is optimal. Model reliability was checked by likelihood test and the accuracy of model by Theil inequality coefficient as well. Finally, marginal effect of the change of socio-demographic characteristics of households on trips was analyzed.

A Study on Airplane Maintenance Engineers and Pilots Relationships and Effectiveness of Flight Operation

KANG, Inwon · CHOI, Sejong

This paper investigates the effectiveness of flight operation between airplane maintenance engineers and pilots who work together. The developed model is to identify the relationship between the attitudes of airplane maintenance engineers and pilots and the effectiveness of flight operation.

The results indicate that the shared values and balanced power between airplane maintenance engineers and pilots are positively related with trust and job satisfaction, but conflict is negatively related. Further, trust and job satisfaction positively affect the effectiveness of flight operation.

These findings suggest that the improvements of the relationships between airplane maintenance engineers and pilots need more efforts to ensure the aviation safety and efficient flights.